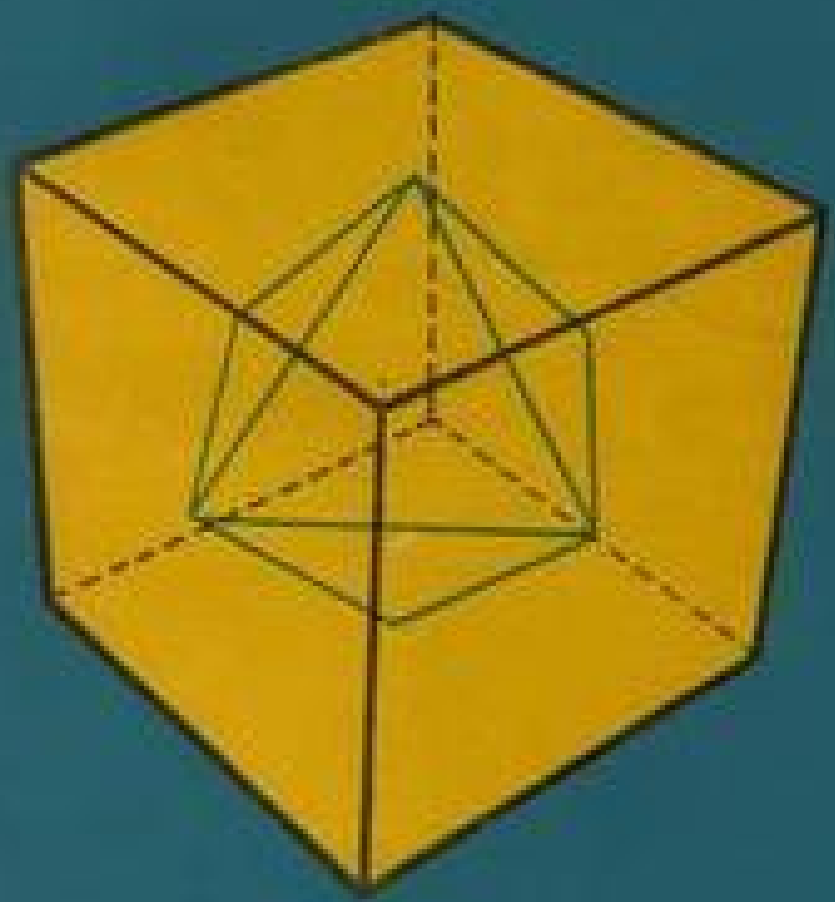
ALAMIRO ROBLEDO H.

LECCIONES DE ALGEBRA ELEMENTAL MODERNA

I I



UNIVERSIDAD DE CONCEPCION CHILE

LECCIONES DE ALGEBRA ELEMENTAL MODERNA

ALAMIRO ROBLEDO H.
Profesor Titular del Instituto de Matemática
Universidad de Concepción (Chile)

TOMO II

Santiago de Chile, 1973

Editorial Universitaria, S. A.
San Francisco 454
Santiago - Chile

INDICE GENERAL DEL TOMO II

Capítulo IX

DOMINIO ORDENADO. CAMPO ORDENADO

Capítulo X

TEORÍA ELEMENTAL DE ESPACIOS VECTORIALES Y TRANSFORMACIONES LINEALES

A) *Espacios vectoriales*

B) *Transformaciones lineales*

INTRODUCCION

Al escribir este TOMO II de los apuntes correspondientes a mi curso de "Lecciones de Algebra Elemental Moderna", he procurado lograr un constante paralelismo con el primero. Es decir, se ha procurado, por una parte, conseguir que el estudiante, al mismo tiempo que aprenda nuevas materias, las comprenda con facilidad, a pesar de la complejidad que algunas de ellas puedan tener, valiéndose para ello de las analogías que las relacionan con las contenidas en el TOMO I y renovando, así, el recuerdo de las contenidas en él.

Por otra parte, se ha procurado al máximo que la exposición de los nuevos temas se haga en forma tal que se siga el mismo orden y con razonamientos análogos a los que se dieron en el primer volumen.

En el primer tomo se atendió, además de las nociones de Lógica estrictamente necesarias para ciertos tipos de razonamientos habituales en Matemática, a una amplia información sobre los conceptos fundamentales de conjunto, de relaciones, de funciones, de operaciones y de homomorfismos, en general.

Nuestra primerísima idea fue de que el estudiante que recién ingresa a la Universidad se familiarizara bien pronto con el lenguaje de estos conceptos básicos y aprendiera a manejar con soltura estas operaciones conjuntistas. Por otra parte, se tuvo también la idea de no ser, en esta primera etapa de los estudios, demasiado riguroso, manteniéndonos, dentro de lo posible, en un plano intuitivo, pero sin renunciar por completo a las demostraciones.

En cambio, en este TOMO II el método es deductivo o axiomático; porque ya se cuenta con la suficiente madurez adquirida por el estudiante.

Los temas que en este tomo se tratan son los que todo profesional matemático está de acuerdo en considerar como indispensables para una primera formación matemática y que son: teorías elementales de grupos, de anillos, de cuerpos, de espacios vectoriales, de transformaciones lineales y de matrices. La elección de estos temas refleja de manera natural la evolución que ha venido experimentando la Matemática en el lapso de este último medio siglo.

También, en este segundo libro de estos apuntes, tal como se hizo en el primero, dado su carácter elemental, adecuado al estudiante que recién se inicia en el estudio de estas materias, se realizó todo el esfuerzo

posible para presentar los diversos temas en él contenidos mediant explicaciones lentas y cuidadosas. En realidad, en ambos tomos, nos hemos guiado por el sentimiento de que es mejor pecar por mucha explicación, que pecar por poca explicación.

Ahora bien, habiendo señalado perfectamente bien lo que está escrito en ambos tomos, se comprende fácilmente lo que no está contenido en ellos, de carácter fundamental, para que pudiera el TODO constituir un verdadero curso general de Algebra Elemental. Para conseguir este objetivo, faltan, pues, los sistemas numéricos usuales del álgebra ordinaria; es decir, los números naturales, los enteros, los racionales, los reales y los números complejos, y todos con sus propiedades aritméticas más fundamentales. También faltan, entre otros temas, la construcción del anillo de los polinomios con sus propiedades esenciales; las ecuaciones algebraicas, señalando sus propiedades fundamentales y su resolución numérica y, finalmente, la teoría de los determinantes y los sistemas de ecuaciones lineales.

Aquí, nuestra decisión se guió por la creencia de que las materias estudiadas en los dos primeros tomos constituyen sólo un lenguaje, mientras que las recientemente enunciadas constituyen a su vez el empleo o aplicación de dicho lenguaje. Cualquier intento que se hubiese hecho para introducir estas aplicaciones en el segundo volumen, habría dado como resultado una obra demasiado grande para presentarla honestamente al público estudiantil de nuestra Universidad.

Para remediar esta situación, se anuncia la pronta aparición de un tercer y último tomo que vendrá a completar la unidad de conocimientos necesarios que se persigue con estos Apuntes.

Ahora bien, como la lectura de una obra de matemática no ha de hacerse necesariamente en el orden de las páginas, sugerimos, por tanto, comenzar por el estudio del primer tomo de estos apuntes, seguir después con el tercero y consultando solamente las partes conceptuales de álgebra dadas en el segundo tomo.

Sería mi deseo haber acertado plenamente, y que la obra en su totalidad resultase verdaderamente útil a gran parte del estudiantado que se ha detenido atemorizado ante las dificultades matemáticas que se les presentaban.

Prof. ALAMIRO ROBLEDO HERRERA

Concepción, Ciudad Universitaria, mayo de 1972

Capítulo VII

TEORIA ELEMENTAL DE GRUPOS

7.0. En el presente capítulo vamos a recapitular los ejemplos dados al final del capítulo anterior, dando la definición siguiente:

Definición. Sea G un conjunto (finito o infinito) provisto de una "multiplicación" (operación binaria interna) que también podría recibir otro nombre,

$$\cdot : G \times G \to G$$
$$(x, y) \to x \cdot y$$

Vamos a establecer que G con esta operación es un GRUPO si satisface al siguiente sistema de axiomas (más débil, es decir, menos exigente o más modesto que el que se dio en el Capítulo VI, anterior):

a) $\forall\, a, b \in G \implies ab \in G$

b) $(ab)c = a(bc), \forall\, a, b, c \in G$

c) Existe por lo menos un elemento $e \in G$ que se llamará NEUTRO A LA IZQUIERDA tal que

$$ex = x, \quad \forall\ x \in G$$

d) Cualquiera sea $x \in G$, existe por lo menos un elemento $x^{-1} \in G$ que se llamará INVERSO A LA IZQUIERDA tal que

$$x^{-1}x = e$$

Admitiendo que se cumple este sistema "débil" de axiomas, damos a continuación las siguientes consecuencias:

Proposición 1. Se tiene:

$$x \cdot x^{-1} = e$$

Dem.

Tenemos, por los axiomas a), b), y c):

$$x^{-1}xx^{-1} = (x^{-1}x)\,x^{-1} = e\,x^{-1} = x^{-1}$$

Multiplicando por un "inverso a izquierda" de x^{-1}, es decir, por $(x^{-1})^{-1}$, se tiene:

$$(x^{-1})^{-1}(x^{-1}xx^{-1}) = (x^{-1})^{-1}x^{-1}$$
$$[(x^{-1})^{-1}x^{-1}]\,(xx^{-1}) = (x^{-1})^{-1}x^{-1}$$

o sea, $\quad e(xx^{-1}) = e$

de donde, $\quad xx^{-1} = e$

lo que demuestra la proposición.

Observación. La proposición que se acaba de demostrar nos enseña que un elemento que es inverso a izquierda, también lo es a derecha.

Corolario. Un inverso de x^{-1} es x.

En efecto, en la segunda parte de la demostración de la anterior proposición, vimos que:

$$(x^{-1})^{-1}\, x^{-1} = e$$

y habiendo probado que $xx^{-1} = e$, se concluye que:

$$(x^{-1})^{-1} = x$$

Proposición 2. Se tiene:

$$xe = x,\ \forall\, x \in G$$

Dem.

Tenemos:

$$xe = x(x^{-1}x) = (xx^{-1})\,x = ex = x \qquad \text{(c.q.d.)}$$

Observación. Esta segunda proposición nos enseña que, si un elemento es neutro a izquierda, también lo es a derecha.

Proposición 3. Dados a, b $\in$ G arbitrarios, entonces existen x, y EG tales que:

$$\begin{cases} a\,x = b \\ y\,a = b \end{cases}$$

es decir, se puede dividir *b* por *a* en dos sentidos (divisiones a la izquierda y a la derecha).

Dem.

Bastará indicar un "x" y un "y" que lo hace.

En efecto, basta formar:

$$\begin{cases} x = a^{-1}b \\ y = b\,a^{-1} \end{cases}$$

y tendremos,

$$\begin{cases} a\,x = a(a^{-1}b) = (aa^{-1})b = e\,b = b \\ y\,a = (b\,a^{-1})a = b(a^{-1}a) = b\,e = b \end{cases}$$

Probaremos, además, que estas divisiones son únicas, esto es:

$$\begin{cases} a\,x = a\,x' \implies x = x' \\ y\,a = y'\,a \implies y = y' \end{cases}$$

En efecto, supongamos que se tenga:

$$a\,x = a\,x'$$

luego, $a^{-1}(a\,x) = a^{-1}(a\,x')$

$$(a^{-1}a)x = (a^{-1}a)\,x'$$

$$e\,x = e\,x'$$

$$x = x'$$

Asimismo, al multiplicar los dos miembros de $y\,a = y'\,a$ por a^{-1}, se llega $y = y'$.

También, esto se expresa diciendo que en un grupo, cualquier elemento es simplificable, ya sea a izquierda o a derecha.

Corolario. Los elementos e y a^{-1} son únicos.

En efecto, *e* es la única solución de la ecuación:

$$x\,a = a \quad (a \text{ dado cualquiera})$$

a^{-1} es la única solución de la ecuación:

$$x\,a = e \quad (a \text{ dado cualquiera})$$

En resumen, las proposiciones 1, 2 y 3 nos han mostrado que el sistema "débil" de axiomas a), b), c) y d), caracteriza un grupo.

Es claro que, en este sistema "débil", en lugar de decir: existen por lo menos un elemento neutro a izquierda y un inverso a izquierda, pudo haberse dicho también, existen un elemento neutro a derecha y un inverso a derecha, y siempre este otro sistema "débil" de axiomas, caracteriza un grupo.

En la práctica, en casos concretos, para verificar que un conjunto presenta una estructura de grupo, basta aplicar uno cualquiera de los dos sistemas "débil" de axiomas.

7.1. *Otras caracterizaciones de un grupo*

Hay muchos otros sistemas de postulados o axiomas para caracterizar a los grupos. He aquí dos de ellos muy útiles:

Teorema 1. Si G es un conjunto no vacío, cerrado para una multiplicación asociativa, respecto a la cual todas las ecuaciones tales que:

$$\begin{cases} a\,x = b \\ \qquad\qquad ,\ \forall\ a, b \in G \\ y\,a = b \end{cases}$$

tienen soluciones *x* e *y* en G, entonces G es un grupo.

Dem. Sea c un elemento fijo arbitrario de G.

Sea *e* una solución de la ecuación:

$$x\,c = c$$

es decir, se cumple:

$$e\,c = c$$

Sea ahora *a* un elemento arbitrario de G; luego, por hipótesis, existe x tal que,

$$a = c\,x$$

por lo tanto,

$$e\,a = e\,(c\,x) = (e\,c)\,x = c\,x = a$$

por consiguiente, e es elemento neutro a la izquierda.

Supongamos en seguida que sea *a* dado arbitrariamente en G, y sea a^{-1} una solución de la ecuación:

$$y\ a = e$$

luego, $a^{-1}\ a = e$
es decir, a^{-1} es un elemento inverso de *a* a la izquierda.

Este resultado y el anterior muestran que G cumple los axiomas débiles que hemos mencionado anteriormente, y por tal motivo, G es un grupo.

Lo que demuestra el teorema.

Pasemos en seguida a enunciar el siguiente teorema, válido solamente para un grupo finito.

Teorema 2. Si G es un conjunto finito, cerrado para una multiplicación asociativa, respecto a la cual es válida la doble ley de cancelación, esto es, todos los elemento son regulares:

$$\begin{cases} a\,x = a\,x' \implies x = x' \\ y\,a = y'\,a \implies y = y' \end{cases}$$

entonces, G es un grupo.

Dem. Aprovechando el Teorema 1 anterior, bastará demostrar "la posibilidad de las divisiones en G".

Sea, pues, *a* un elemento fijo de G. Definamos la aplicación,

$$f : G \longrightarrow G \quad ,$$

así, $f(x) = a\,x,\ \forall\ x \in G$

Probemos que f es inyectiva. Tenemos:

$$f(x) = f(x')$$

o sea, $a\,x = a\,x'$
y como, por hipótesis, se verifica la ley de cancelación, resulta:

$$x = x'$$

luego, f es inyectiva.

Por otra parte, f(G) = conjunto de todas las imágenes de los elementos de la forma a x, es un subconjunto de G en correspondencia biunívoca con G.

Pero siendo, por hipótesis, G finito, f(G) no puede ser un subconjunto propio de G. (Un conjunto finito no es coordinable con ningún subconjunto propio). Luego,

$$f(G) = G$$

Así pues, f es una biyección de G sobre G.

En otras palabras, cada elemento de G es de la forma:

$$b = a\,x$$

para algún $x \in G$.

Así hemos demostrado que para *a*, *b* arbitrarios en G, existe x tal que:

$$a\,x = b$$

Asimismo, empleando en vez de f, la aplicación:

$$g : G \rightarrow G$$

definida por, $g(y) = y\,a$
se demuestra la solubilidad de la ecuación:

$$y\,a = b$$

Luego, en virtud del Teorema 1, G es un grupo.

Corolario. Sea G un grupo, y sea $G' \subset G$ tal que G' es finito y estable, entonces G' es un grupo (subgrupo).

Dem. Por ser G' estable, es cerrado para la operación de G, y si $a \in G'$ $\Rightarrow$ $a \in G$ $\Rightarrow$ *a* es regular, cualquiera que sea $a \in G$.

Ahora, como por hipótesis G' es finito y por ser todos sus elementos regulares, resulta por el teorema que G' es un grupo, subgrupo de G.

7.2. *Operación inversa*

Hemos visto que en un grupo multiplicativo G, las ecuaciones:

$$\begin{cases} a\,x = b \\ y\,a = b \end{cases}$$

cualesquiera sean *a* y *b* en G, tienen siempre solución y ésta es única. La unicidad es consecuencia de las leyes de cancelación.

En adelante, al hablar de "ecuación sobre un grupo G" (como también sobre cualquiera otra estructura), se entenderá cualquier expresión formal del tipo:

$$a\,x\,b = c;\quad a\,x = b;\quad x\,a = b;\ \ldots$$

donde a, b, c representan elementos de G y donde x representa un signo indeterminado (o simplemente una indeterminada). Cuando al reemplazar x por elementos de G en alguna de las ecuaciones anteriores se obtiene una expresión verdadera, diremos que x es *solución* de las mismas.

La resolución de la ecuación $a\,x = b$ equivale a decir que se pide, cono-

ciendo el resultado y el primer término de una operación, determinar el segundo; cuando, como en el caso de los grupos, la solución siempre existe y es única, a cada par de elementos a y b se le hace corresponder un elemento x del grupo y uno sólo; queda así definida una nueva operación en el grupo que se denomina la *Operación Inversa a la Derecha* de la operación dada.

Análogamente, partiendo de la ecuación y a = b, se define la *Operación Inversa a la Izquierda* de la operación dada.

En un grupo conmutativo o abeliano las dos operaciones inversas a la derecha y a la izquierda de una dada son la misma operación, que denominaremos entonces simplemente OPERACION INVERSA de la dada en el grupo. Así, por ejemplo, en un grupo multiplicativo y abeliano, la solución $x = b\,a^{-1}$ de la ecuación $x\,a = b$ que es idéntica a la ecuación $a\,x = b$, se escribe también en la forma,

$$x = \frac{b}{a}$$

y diremos que x es el COCIENTE o la RAZON de los dos elementos a y b; a es el numerador, b el denominador.

Definición. En un grupo multiplicativo y abeliano, llamaremos COCIENTE o DIVISION de dos elementos a y b, en este orden, un elemento que multiplicado por b da por resultado a y que es el elemento $a\,b^{-1}$, producto de a por el inverso de b. Se le representa $\frac{a}{b}$.

$$\frac{a}{b} = a\,b^{-1}$$

En particular,

$$\frac{1}{a} = 1 \cdot a^{-1} = a^{-1}$$

donde 1 representa el elemento unidad para la multiplicación.

Queda definida así una nueva operación: "La División", inversa de "La Multiplicación", y que es una aplicación de,

$$\div : G \times G \to G$$

Evidentemente es uniforme, es decir, si $a = c$ y $b = d$, entonces,

$$\frac{a}{b} = \frac{c}{d}$$

ya que $\frac{a}{b} = a\,b^{-1} = c\,d^{-1} = \frac{c}{d}$

7.3. Demostraremos ahora algunas reglas de cálculo que valen para los cocientes en un grupo multiplicativo o abeliano.

Teorema 1. Se tiene,

$$\frac{a}{b} = \frac{c}{d} \text{ si, y sólo si } a\,d = b\,c$$

Dem. Supongamos que $\frac{a}{b} = \frac{c}{d}$, es decir que $a\,b^{-1} = c\,d^{-1}$.

Entonces, podremos escribir:

$$ad = (a\,d) \cdot 1 = (ad) \cdot (b^{-1}b) = ((ad)\,b^{-1})\,b = ((da)b^{-1})\,b =$$
$$= (d\,(a\,b^{-1}))b = (d(c\,d^{-1}))\,b = (d(d^{-1}c))\,b =$$
$$((d\,d^{-1})c)b = (1 \cdot c)b = c\,b = b\,c$$

Recíprocamente, supongamos que $ad = bc$; entonces podremos escribir:

$$\frac{a}{b} = ab^{-1} = (ab^{-1}) \cdot 1 = (ab^{-1}) \cdot (dd^{-1}) = (b^{-1}a)(\,(dd^{-1}) =$$
$$= (\,(b^{-1}a)d\,)d^{-1} = (b^{-1}(ad))\,d^{-1}c = (b^{-1}(bc))\,d^{-1} =$$
$$= ((b^{-1}b)c)\,d^{-1} = (1 \cdot c)d^{1} = c\,d^{-1} = \frac{c}{d}$$

Este resultado y el anterior, demuestran el teorema.

Teorema 2. Se tiene, $\frac{a}{b} \cdot \frac{c}{d} = \frac{a\,c}{b\,d}$

Demo. Tenemos:

$$\frac{a}{b} \cdot \frac{c}{d} = (ab^{-1})\,(cd^{-1}) = ((ab^{-1})c)\;d^{-1} = (a\;(b^{-1}c))\,d^{-1} =$$
$$= (a(cb^{-1}))\;d^{-1} = ((ac)b^{-1})\;d^{-1} = (ac)\;(b^{-1}\,d^{-1})$$
$$= (ac)\,(db)^{-1} = (ac)\,(bd)^{-1}$$
$$= \frac{a\,c}{b\,d}$$

resultado que prueba el teorema y el cual nos da la regla de multiplicación de cocientes.

Teorema 3. Se tiene, $\frac{a}{b} = \frac{a\,c}{b\,c}$, $\forall\, c \in G$.

Dem. En virtud de las propiedades conmutativa y asociativa, tenemos:

$$a(b\,c) = b(a\,c)$$

y por el Teorema 1) esta relación es equivalente a la igualdad de los dos cocientes:

$$\frac{a}{b} = \frac{a\,c}{b\,c}$$

Asimismo, puede probarse que,

$$\frac{a}{b} = \frac{\frac{a}{c}}{\frac{b}{c}}$$

Observemos que aplicamos estas dos propiedades cuando amplificamos y simplificamos fracciones numéricas.

Observación. El teorema recién probado permite reducir dos cocientes, o dos razones al mismo denominador. Además, un elemento cualquiera puede ser considerado como un cociente.

$$a = a\,e^{-1} = \frac{a}{e} = \frac{a}{1} = \frac{a\,c}{c}$$

Teorema 4. Se tiene, $\left(\frac{a}{b}\right)^{-1} = \frac{b}{a}$

Dem. Habrá que probar que el inverso de $\frac{a}{b}$ es $\frac{b}{a}$, es decir que,

$$\frac{a}{b} \cdot \frac{b}{a} = 1$$

En efecto, por Teorema 2), se escribe:

$$\frac{a}{b} \cdot \frac{b}{a} = \frac{a\,b}{b\,a} = \frac{a\,b}{a\,b} = (ab)\,(ab)^{-1} = 1$$

Teorema 5. Se tiene,

$$\frac{\frac{a}{b}}{\frac{c}{d}} = \frac{a\,d}{b\,c}$$

Dem. Tenemos:

$$\frac{\frac{a}{b}}{\frac{c}{d}} = \frac{a}{b} \cdot \left(\frac{c}{d}\right)^{-1} = \frac{a}{b} \cdot \frac{d}{c} = \frac{ad}{bc}$$

En particular, se tiene:

$$\frac{\frac{a}{b}}{c} = a \cdot \left(\frac{b}{c}\right)^{-1} = a \cdot \frac{c}{b} = \frac{a}{1} \cdot \frac{c}{b} = \frac{a\,c}{b}$$

y, $$\frac{\frac{a}{b}}{c} = \frac{a}{b} \cdot c^{-1} = \frac{a}{b} \cdot \frac{1}{c} = \frac{a}{bc}$$

7.4. *Grupos abelianos*

Estudiaremos ahora el caso particular de los grupos abelianos, es decir, aquellos en que la ley de composición, cualquiera que ella sea, es conmutativa. En tales grupos se suele representar la operación en cuestión con el signo " + " de la adición ordinaria; el elemento neutro se representa con o (cero) y el inverso de cada elemento a se escribe $(-a)$ y se lo llama el simétrico o el opuesto de a.

Con esta notación especial (llamada aditiva, mientras que a la otra que hemos usado anteriormente, se la llama multiplicativa) los axiomas característicos de grupo abeliano, se escriben:

a) $\forall\, a, b \in G \implies a + b \in G$
b) $(a + b) + c = a + (b + c), \forall\, a, b, c \in G$
c) Existe un elemento $o \in G$ tal que, $x + o = x, \forall\, x \in G$
d) Para cada $x \in G$ existe un elemento $(-x) \in G$, llamado simétrico de x, tal que, $x + (-x) = o$
e) $a + b = b + a, \forall\, a, b \in G$

De acuerdo con lo demostrado anteriormente para grupos multiplicativos, en un grupo abeliano valen entonces las siguientes reglas:

1) Si $x + a = y + a$, entonces $x = y$ (Ley de cancelación)
2) $-(-a) = a$
3) $-(a + b) = (-a) + (-b)$
4) La ecuación $b + x = a$ tiene una única solución dada por $x = a + (-b)$

Definición. En un grupo ebeliano llamaremos DIFERENCIA de dos elementos a y b, en este orden, un elemento que sumado con b da por resultado a y que es el elemento $a + (-b)$, suma de a con el simétrico de b. Se lo representa por $(a - b)$

$$a - b = a + (-b)$$

Como el simétrico de b es único, dados dos elementos a y b, su diferencia es un elemento unívocamente determinado. Queda definida así una nueva operación binaria interna en los grupos abelianos: La SUSTRACCION, inversa de la ADICION, y que es una aplicación de,

$$- : G \times G \to G$$

Evidentemente es uniforme, es decir, si $a = c$ y $b = d$, entonces,

$$a - b = c - d$$

Demostraremos en seguida algunas reglas de cálculo que valen para las diferencias en un grupo aditivo, es decir, en un grupo abeliano.

Teorema 1. Se tiene,

$$a - (-b) = a + b$$

Dem. Tenemos:

$$a - (-b) = a + (-(-b)) = a + b$$

Teorema 2. Se tiene,

$$(a + b) - c = a + (b - c)$$

Dem. Tenemos:

$$(a + b) - c = (a + b) + (-c) = a + (b + (-c)) = a + (b - c)$$

Teorema 3. Se tiene,

$$(a - b) + c = a - (b - c)$$

Dem. Tenemos:

$$\begin{aligned}(a - b) + c &= (a + (-b)) + c = (a + (-b)) + (-(-c)) = \\ &= a + [(-b) + (-(-c))] = \\ &= a + [-(b + (-c))] = \\ &= a + [-(b - c)] = a - (b - c)\end{aligned}$$

Teorema 4. Se tiene,

$$(a - b) - c = a - (b + c)$$

Dem. Tenemos:

$$\begin{aligned}(a - b) - c &= (a + (-b)) + (-c) = a + [(-b) + (-c)] = \\ &= a + [-(b + c)] = a - (b + c)\end{aligned}$$

7.5. En resumen, al estudiar las propiedades fundamentales que derivan directamente de los axiomas de las estructuras de grupo multiplicativo conmutativo y de grupo abeliano (aditivo), hemos probado la mayoría de los teoremas de álgebra sólo para dos términos, pero pueden extenderse, por inducción, para cualquier número de términos.

Finalmente, toda la terminología utilizada en el estudio de monoides y semigrupos se aplica en particular a los grupos. Así, por ejemplo, un isomorfismo, o un homomorfismo de grupos $f : G \to G'$, significa un isomorfismo, o un homomorfismo de los semigrupos correspondientes. Conviene observar que tratándose de grupos, se tiene:

$$f(e) = e'$$

donde e es la unidad de G y e' la unidad de G'.

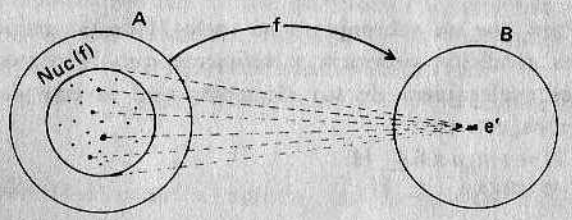
Veremos en seguida un resultado de suma utilidad, que nos es proporcionado por el siguiente:

Teorema. Sea G un grupo y sea A un semigrupo con unidad. Sea $f : G \to A$ un homomorfismo. Entonces, f es un monomorfismo si, y sólo si, Nuc $(f) = \{e\}$

Dem. Recordemos primeramente que el Nuc (f) de un homomorfismo entre dos monoides, o dos semigrupos A y B con unidad es el conjunto definido por:

$$\text{Nuc}(f) = \{x \in A : f(x) = e' \in B\}$$

es decir, el conjunto de todos los elementos de A que van a la unidad e' de B.



Con estos antecedentes probemos el teorema. Tenemos:

a) Supongamos que f sea un monomorfismo, es decir, la aplicación f es en este caso inyectiva, y probemos que en tal caso el núcleo se reduce al elemento unidad e del grupo G. En efecto, si $x \in$ Nuc (f) se tiene:

$$f(x) = e' = f(e)$$

y como f es inyectiva, resulta $x = e$.

Esto demuestra que Nuc $(f) = \{e\}$

b) Recíprocamente, supongamos que sea Nuc $(f) = \{e\}$ y probemos entonces que f es inyectiva.

Sean $x, x' \in G$ tales que,

$$f(x) = f(x')$$

y como G es un grupo, existe $x^{-1} \in G$; luego, se tiene:

$$f(x) \cdot f(x^{-1}) = f(x') \cdot f(x^{-1})$$

y como f es un homomorfismo, escribimos:

$$f(x \cdot x^{-1}) = f(e) = f(x' \cdot x^{-1})$$

o sea, $f(x' \cdot x^{-1}) = f(e) = e'$

de donde, $x' \cdot x^{-1} \in \text{Nuc}(f) = \{e\}$

y lo que implica, $x' \cdot x^{-1} = e \Longrightarrow (x' \cdot x^{-1})x = ex \Longrightarrow x'(x^{-1}x) =$

$= ex \implies x' = x$

Luego, si $f(x) = f(x')$ entonces $x = x'$, es decir f es inyectiva, y como f es, además, un homomorfismo, resulta que f es un monomorfismo.

Este resultado y el anterior, prueban el teorema.

7.6. *Subgrupos*

Dado un grupo G, puede ocurrir que los elementos de un subconjunto H de G formen, respecto a la operación de G, un grupo. En este caso diremos que H es un SUBGRUPO de G. Cada grupo G tiene dos subgrupos *impropios o triviales*; a saber, el mismo G y el grupo formado por la unidad *e* por sí sola; todos los subgrupos restantes se llaman *propios*.

Daremos a continuación algunos criterios para los subgrupos, habitualmente utilizados.

Creterio1. Para que un subconjunto no vacío H de un grupo G sea un subgrupo, es condición necesaria y suficiente que H contenga el producto de dos cualesquiera de sus elementos y el inverso de cualquiera de sus elementos; esto es:

a) $\forall\, a, b \in H$ se tenga $a\,b \in H$

b) $\forall\, a \in H$ se tenga $a^{-1} \in H$

Dem. Supongamos cumplidas estas propiedades. Entonces, vale la propiedad asociativa por ser la misma operación que en G; hay un elemento neutro, puesto que por a) y b) se tiene:

$$a \in H, a^{-1} \in H \implies a\,a^{-1} = e \in H;$$

por otra parte, cada elemento admite un inverso, y además el cierre de la operación en H se verifica.

Así hemos probado que la condición es suficiente.

Recíprocamente, sea H un subgrupo de G. Entonces, es evidente que contiene el producto de dos cualesquiera de sus elementos; sea e' el elemento neutro de H y *a* un elemento cualquiera de H, se tiene $a\,e' = a$; considerando *a* y *e'* como elemento de G se tiene también $a\,e = a$, siendo *e* el elemento neutro de G, luego, $a\,e' = a\,e$; en virtud de la ley de cancelación se tiene $e = e'$, es decir, G y H tienen el mismo elemento neutro. Luego los inversos de un elemento son los mismos en G y en H.

Así hemos probado que la condición es necesaria.

Este resultado y el anterior, demuestran el criterio.

Criterio 2. Para que un subconjunto no vacío H de un grupo G sea un subgrupo, es condición necesaria y suficiente que para todo par (a, b) de elementos contenidos en H, H contenga también ab^{-1}; esto es:

a) $\forall\, a, b \in H$ se tenga $a\,b^{-1} \in H$

Dem. La condición necesaria es evidente, ya que si H es un subgrupo, la condición a) se cumple automáticamente.

Probemos ahora que la condición es suficiente; esto es, se cumple a). Entonces, con el par (a, a), H contiene a $a^{-1} = e$; con el par (e, a), H contiene e $a^{-1} = a^{-1}$ y, en fin, con el par (a, b), H contiene a y b^{-1}; pues, contiene $a(b^{-1})^{-1} = ab$. Es decir, se cumplen todas las condiciones del Criterio 1); luego, H es un subgrupo de G.

Criterio 3. Un subconjunto no vacío H de un grupo *finito* G es un subgrupo si, y sólo si, $H^2 = H$, o sea, $a, b \in H \implies a \cdot b \in H$ es decir, si es cerrado respecto a la multiplicación.

Dem. La propiedad enunciada es consecuencia inmediata de un teorema ya estudiado que nos afirma que: todo conjunto finito, cerrado para una multiplicación asociativa, respecto de la cual todos los elementos son regulares (doble ley de cancelación), es un grupo.

Como se indica, este criterio es muy útil tratándose de grupos finitos.

Observación. Con notación aditiva, las condiciones expresadas en los criterios anteriores se escriben así:

Criterio 1'.

a) $\forall\, a, b \in H$ se tenga $a + b \in H$

b) $\forall\, a \in H$ se tenga $(-a) \in H$

Criterio 2'.

a) $\forall\, a, b \in H$ se tenga $a - b \in H$

Criterio 3'.

a) $\forall\, a, b \in H$ se tenga $a + b \in H$

7.7. *Grupo Cíclico*

Hemos hallado ya numerosos ejemplos de subgrupos: $\mathbb{Z}$ es un subgrupo aditivo de $\mathbb{Q}$ así como también de $\mathbb{R}$ el cual lo es de $\mathbb{C}$.

El grupo de translaciones del espacio euclideano es un subgrupo de los desplazamientos de este espacio. También las rotaciones alrededor de un punto fijo $\mathcal{O}$ forman un subgrupo del grupo de los desplazamientos. Pero, sobre el grupo de las rotaciones de centro $\mathcal{O}$ se pueden también encontrar subgrupos, por ejemplo, aquellos que corresponden a los ángulos $k\alpha$, donde $k \in \mathbb{Z}$ y α un ángulo dado.

Los múltiplos de cualquier entero positivo a,

$$\ldots, -3a, -2a, -a, o, +a, +2a, +3a, \ldots$$

forman subgrupos del grupo aditivo $\mathbb{Z}$ de los enteros relativos; las potencias de cualquier número racional forman un subgrupo del grupo de los racionales.

De modo más general, siendo G un grupo multiplicativo y a un elemento particular de G distinto de e, el conjunto de las potencias a^p con $p \in \mathbb{Z}$, constituyen un subgrupo de G, que denominaremos GRUPO CICLICO.

Este grupo es abeliano, ya que:

$$a^p \cdot a^q = a^{p+q} = a^{q+p} = a^q \cdot a^p$$

Se pueden producir dos eventualidades, a saber:

1°. Todos los a^p son distintos; el grupo está formado entonces por una infinidad de elementos:

$$\ldots, a^{-p}, a^{-(p-1)}, \ldots, a^{-1}, a^o = e, a^1, \ldots, a^{p-1}, a^p, \ldots$$

Designaremos, en este caso, al grupo cíclico por C_∞.

Este grupo es isomorfo a $\mathbb{Z}$, mediante la aplicación:

$$p \in \mathbb{Z} \rightarrow a^p \in C_\infty \quad \text{(biyectiva)}$$

2°. Existen $p, q \in \mathbb{Z}$ $p \neq q$, tales que:

$$a^p = a^q$$

luego, $a^{p-q} = e$

Es decir, existen enteros *no nulos* k tales que:

$$a^k = e$$

y también, $a^{-k} = (a^k)^{-1} = e^{-1} = e$

Uno por lo menos de k, $-k$, es positivo.

Por lo tanto, existen enteros positivos k tales que,

$$a^k = e.$$

Sea n el menor entero positivo tal que $a^n = e$

En este caso, el grupo consta de los n elementos distintos:

$$a^o = e, a, a^2, a^3, \ldots, a^{n-1}$$

es decir, es un grupo finito de orden n, que lo designaremos por C_n (grupo cíclico de orden n). Además, este grupo es isomorfo al grupo aditivo $\mathbb{Z}_n$, de las clases residuales módulo n.

En efecto, sea la serie de potencias enteras de a:

$$A = \{\ldots, a^{-n}, \ldots, a^{-1}, a^o = e, a^1, \ldots, a^n, \ldots\}$$

Sea $x \in A$ arbitrario, esto implica que x es de la forma:

$$x = a^m, m \in \mathbb{Z}$$

Probaremos que a^m es igual a uno de los n elementos:

$$e, a, a^2, \ldots, a^{n-1}$$

Si $o \leq m < n$, entonces a^m es uno de esos elementos y estaría demostrado.

Si $m \geq n$, entonces por el algoritmo de división se tiene:

$$m = q\,n + r, \quad o \leq r < n$$

luego,

$$a^m = a^{qn+r} = (a^n)^q \cdot a^r = e^q \cdot a^r = e \cdot a^r = a^r,$$

esto es,

$$a^m = a^r \text{ con } o \leq r < n$$

De aquí resulta que si $m \geq n$, entonces $x = a^m$ puede reemplazarse por su igual a^r, donde $o \leq r < n$.

Si $m < o$, entonces se puede escribir:

$$a^m = (a^{-m})^{-1} = (a^{-1})^{-m}$$

y como $a^n = e$, resulta que $a^{-1} = a^{n-1}$; sustituyendo este valor de a^{-1} en la igualdad anterior, escribimos:

$$a^m = (a^{n-1})^{-m} = a^{-m(n-1)}$$

y como $n - 1 \geq o$ y $-m > o$, resulta $-m(n-1) \geq o$; luego, poniendo $-m(n-1) = p$, $p \geq o$, se obtiene:

$$a^m = a^p \text{ con } p \geq o$$

esto es, toda potencia de exponente negativo es igual a una potencia de exponente positivo y de la misma base.

Por lo tanto, todas las potencias de exponente negativo se suprimen por ser iguales a potencias de exponente positivo, y las potencias de exponente positivo coinciden con uno de los elementos: $e, a, a^2, \ldots, a^{n-1}$.

Faltaría demostrar ahora que estos n elementos son todos distintos. En efecto, sean

$$o \leq i < n, \quad o \leq j < n, \quad i \neq j$$

y supongamos que fuese:

$a^i = a^j$

Para fijar las ideas, sea por ejemplo $i > j$; luego:

$$(1) \qquad a^{i-j} = a^0 = e$$

Pero, $i - j$ es un entero positivo menor que n; luego, (1) contradice la definición de n.

Así queda demostrado que los elementos:

$$e, a, a^2, \ldots, a^{n-1}$$

son todos distintos, y el grupo cíclico es:

$$C_n = \{e, a, a^2, \ldots, a^{n-1}\}$$

Finalmente nos queda por probar que C_n es isomorfo con el grupo aditivo $\mathbb{Z}_n$.

Antes de demostrar esto, probemos primeramente que:

$$a^p = a^q \iff p \equiv q \pmod{n}$$

En efecto, si $p \equiv q$ (mód. n), entonces $p - q = k\,n$, $k \in \mathbb{Z}$; luego,

$$a^p = a^{q+kn} = a^q \cdot (a^n)^k = a^q \cdot e^k = a^q \cdot e = a^q$$

por lo tanto, la condición es suficiente.

Para demostrar que la condición es necesaria, probaremos primero que:

$$a^p = e \Longrightarrow n \mid p$$

En efecto, por el algoritmo de división, se tiene:

$$p = q\,n + r, \qquad 0 \leq r < n$$

implica, $e = a^p = a^{qn+r} = (a^n)^q \cdot a^r = e \cdot a^r = a^r$

Pero, por la definición de n resulta $r = 0$; luego, $n \mid p$.

Bajo esta idea, supongamos ahora que:

$$a^p = a^q$$

luego,

$$a^{p-q} = e \Longrightarrow n \mid p - q \Longrightarrow p \equiv q \pmod{n}$$

Por consiguiente, todos los elementos de la clase de equivalencia del elemento p, es decir $\overline{p}$, tienen la propiedad de que:

$$a^p = a^q, \text{ donde } q \in \overline{p}$$

o sea, a^p depende sólo de la clase residual $\overline{p}$ de p módulo n.

La aplicación,

$$\overline{p} \longrightarrow a^p \qquad f$$

es una aplicación biyectiva de $\mathbb{Z}_n$ sobre C_n, en virtud de lo recién probado.

Sea, $f(\overline{p}) = a^p$; tenemos:

$$f(\overline{p} + \overline{q}) = f(\overline{p + q}) = a^{p+q} = a^p \cdot a^q = f(\overline{p}) \cdot f(\overline{q})$$

Luego, f es un isomorfismo.

Observaciones. 1. La adición de clases residuales módulo n tiene una interpretación geométrica interesante, aprovechada por Gauss en el estudio de les raíces n-ésimas de la unidad positiva 1 (las cuales forman un grupo). Si se consideran las rotaciones planas R_k de centros $\mathbb{O}$ y ángulo $k \cdot \frac{2\pi}{n}$, $k \in \mathbb{Z}$, hay sólo n distintas $R_0, R_1, R_2, \ldots, R_{n-1}$ (comprendida la rotación nula); corresponden a las clases residuales $\overline{0}, \overline{1}, \overline{2}, \ldots, \overline{n-1}$ y el estudiante verá sin dificultad que la *adición* de clases de congruencias módulo n corresponde al *producto* de las rotaciones de ángulos $k \cdot \frac{2\pi}{n}$; esto es:

$$R_k \longleftrightarrow \overline{k}, R_{k'} \longleftrightarrow \overline{k'}, R_k \cdot R_{k'} \longleftrightarrow (\overline{k} + k')$$

Tenemos así otro ejemplo interesante de isomorfismo. En particular, si las rotaciones del ángulo $\frac{2\pi}{n}$ se refieren a un polígono regular de n lados, un ánulo central del polígono valdrá:

$$\alpha = \frac{2\pi}{n}$$

y de lo dicho se deduce que si a cada rotación $k \cdot \alpha = k \cdot \frac{2\pi}{n}$ del polígono

le hacemos corresponder el elemento a^k del grupo cíclico C_n, establecemos una correspondencia isomorfa entre el grupo C_n y el grupo de las rotaciones de un polígono de n lados.

Las rotaciones de centro O y de ángulo $K\alpha$, siendo la medida de un ángulo en radianes, inconmensurable con π, forman un grupo cíclico infinito.

Las consideraciones precedentes nos mueven a dar las siguientes definiciones:

a) Llamaremos *grupos cíclicos finitos* los grupos que son isomorfos al grupo de las rotaciones de un polígono regular, o isomorfos al grupo aditivo de las clases residuales.

b) Llamaremos *grupos cíclicos infinitos* a los grupos que son isomorfos al grupo aditivo de los números enteros, o isomorfos al grupo de las rotaciones planas de centro O y de ángulo α medido en radianes e inconmensurable con π.

2. Si en vez de la notación multiplicativa se hubiera utilizado la aditiva, entonces los mencionados grupos cíclicos serían:

$$C_\infty = \{\ldots, -na, \ldots, -2a, -a, o, a, 2a, \ldots, na, \ldots\}$$

$$C_n = \{o, a, 2a, \ldots, (n-1)a\}$$

3. Es evidente que si dos grupos A y B son isomorfos al mismo grupo C, ellos son isomorfos entre sí. Por lo tanto, todos los grupos cíclicos infinitos son isomorfos entre sí. Por la misma razón todos los grupos cíclicos del mismo orden n son isomorfos.

4. Los resultados obtenidos nos permiten establecer la siguiente propiedad:

Todo elemento $a \neq e$ de un grupo G, genera un grupo cíclico, finito o infinito. Al orden de este grupo (subgrupo de G) lo llamaremos también *orden del elemento a;* es decir, el orden de un grupo cíclico es igual al orden del elemento generador.

Los grupos cíclicos, finitos o infinitos, pueden también definirse como sigue: diremos que un grupo es cíclico si él es generado por uno de sus elementos.

Decimos, por uno de sus elementos, porque, en general, el elemento generador no es único. A este respecto, afirmamos que si el elemento generador a es de orden n y si a^s (potencia de a) es tal que M.C.D. (n, s) = 1, entonces a^s es también un elemento generador del grupo cíclico:

$$C_n = \{a^0 = e, a, a^2, \ldots, a^{n-1}\}$$

En efecto, probaremos primeramente que el orden de cualquier potencia de a no puede superar al orden de a; esto es, es menor o igual que el orden de este elemento.

Sea n el orden de a y sea m el orden de a^s:

$$a^n = e, (a^s)^m = e;\ n, m \in \mathbb{Z},\ n > o, m > o$$

y demostremos que $m \leq n$.

Supongamos, por el contrario que sea $m > n$; esto es:

$$m = qn + r, \quad o < r < n$$

Tendremos:

$$e = (a^s)^m = (a^s)^{qn+r} = a^{sqn} \cdot a^{sr} = (a^n)^{qs} \cdot (a^s)^r =$$

$$= e \cdot (a^s)^r = (a^s)^r$$

lo cual indica que r ($r < m$) es el orden de a^s, absurdo, puesto que m es el orden de a^s. Luego, el orden de a^s es menor o igual al orden de a. La igualdad tiene lugar cuando n y s son primos entre sí. En efecto, si M.C.D. (n, s) = 1, entonces pueden hallarse dos enteros (uno positivo y el otro negativo) u y v tales que:

$$us + vn = 1$$

Luego, tenemos:

$$(a^s)^u = a^{su} = a^{1-vn} = a\ a^{-vn} = a\,(a^n)^{-v} =$$

$$= a\,e^{-v} = a\,e = a$$

resultado que nos enseña que a es potencia de a^s, y por la propiedad anterior concluímos que:

$$n \leq m$$
$$m \leq n$$

de donde, n = m, siendo n el orden de a y m el orden de a^s.

Así, hemos demostrado que los elementos a y a^s son del mismo orden, cuando n y s son primos entre sí. Por lo tanto, a^s es también un elemento generador del grupo cíclico,

$$C_n = \{e, a, a^2, \ldots, a^s, \ldots, a^{n-1}\}$$

Como consecuencia importante resulta, a modo de recíproco, que si un grupo finito de orden n contiene un elemento también de orden n, entonces el grupo es cíclico; porque si b es este elemento, entonces los elementos:

$$b^0 = e, b, b^2, \ldots, b^{n-1}$$

son todos distintos e iguales a los elementos del grupo.

7.8. *Subgrupo de un grupo cíclico.*

Sabemos C_∞ es un grupo cíclico e isomorfo a $\mathbb{Z}$. Todos los subgrupos de $\mathbb{Z}$ son los conjuntos $m\,\mathbb{Z}$ (múltiplos de un entero positivo m). Si $m \neq m'$, entonces $m\,\mathbb{Z} \neq m'\,\mathbb{Z}$.

Luego, para cada $m = 0, 1, 2, 3, \ldots$, tenemos un subgrupo de C_∞ que consta de todas las potencias de a^m; es decir, engendrado por a^m:

$$\{\ldots, a^{-pm}, \ldots, a^{-2m}, a^{-m}, a^0 = e, a, a^{2m}, \ldots, a^{pm}, \ldots\}$$

Estos subgrupos son distintos y son los únicos.

Por consiguiente, todos los subgrupos distintos de $\{e\}$ de un grupo cíclico infinito son grupos cíclicos infinitos.

Veamos ahora el caso de C_n, grupo cíclico finito de orden n.

A este respecto, demostraremos que todo subgrupo de C_n es también cíclico.

En efecto, sea H un subgrupo de C_n, generado por a. Al subgrupo H pertenece la unidad e, y si $a^p \in H$ también $\overrightarrow{a^p} \in H$.

Sea a^s el elemento de menor exponente contenido en H. Probaremos que todo elemento de H es de la forma:

$$a^m = (a^s)^q$$

o sea, que a^s es el elemento generador del subgrupo H. En efecto, sea $a^m \in H$ con $m > s$. Por el algoritmo de división tenemos:

$$m = qs + r, \quad 0 \leq r < s$$

entonces,

$$a^m = a^{qs+r} = a^{qs} \cdot a^r$$

de donde,

$$a^{m-qs} = a^r$$

Pero, $a^m \in H$ y $a^{-qs} = (a^s)^{-q} \in H$, luego $a' \in H$ y como $r < s$, y además s es el menor exponente de un elemento en H, resulta entonces que $r = o$.

Así pues, $s \mid m$ y $a^m = (a^s)^q$.

Es decir, que H consta de todas las potencias de a^s; luego, H es cíclico; esto es:

$$H = \{e, a^s, a^{2s}, \ldots, a^{(q-1)s}\}$$

Sea el grupo cíclico de n elementos C_n y sea s el menor entero positivo tal que a^s está en el grupo.

Afirmamos que $s \mid n$.

En efecto, $a^n = e \in H$.

Por otra parte sabemos que si $a^m \in H$, entonces $s \mid m \Rightarrow s \mid n$.

Recíprocamente, sea s un divisor positivo arbitrario de n, es decir:

$$n = sq, \quad q \in \mathbb{N}$$

El subgrupo engendrado por a^s es:

$$H = \{e, a^s, a^{2s}, \ldots, a^{(q-1)s}\}$$

luego, H es un subgrupo cíclico de orden q.

En resumen, para cada divisor q de n tenemos un subgrupo y uno solo de C_n de orden q. Este es un grupo cíclico engendrado por a^s, con $s = \frac{n}{q}$

Con conocer así todos los divisores de n, tendremos un subgrupo de C_n.

En consecuencia, si n es primo, entonces C_n no tiene subgrupos propios.

Ejemplo.

Hallar todos los subgrupos de C_{12}.

$$C_{12} = \{e, a, a^2, a^3, a^4, a^5, a^6, a^7, a^8, a^9, a^{10}, a^{11}\}$$

Los siguientes conjuntos son subgrupos:

Orden 1 : $\{e\}$

Orden 2 : $\{e, a^6\}$ es engendrado por a^6

Orden 3 : $\{e, a^4, a^8\}$ es engendrado por a^4

Orden 4 : $\{e, a^3, a^6, a^9\}$ es engendrado por a^3

Orden 6 : $\{e, a^2, a^4, a^6, a^8, a^{10}\}$ es engendrado por a^2

Orden 12 : $\{e, a, a^2, a^3, \ldots, a^{11}\}$ es engendrado por a

Proposición. Sea $(G;.)$ un grupo y sea $(H_i)_{i\in J}$ una familia de subgrupos del grupo G, entonces:

$$H = \bigcap_{i\in J} H_i$$

es un subgrupo de G.

Dem. Basta aplicar alguno de los criterios que tenemos para caracterizar un subgrupo, por ejemplo,

$$a, b \in H \Rightarrow a \cdot b^{-1} \in H$$

Sea pues, $a, b \in H = \bigcap_{i\in J} H_i \Rightarrow a, b \in H_i, \ \forall \ i \in J$

Como los H_i son subgrupos, tenemos:

$$a, b \in H_i \Rightarrow a \cdot b^{-} \in H_i, \ \forall \ i \in J$$

luego, $a \cdot b^{-1} \in H = \bigcap_{i\in J} H_i$, y por lo tanto, H es un subgrupo del grupo dado G.

H es, pues, el menor subgrupo (el más chico) que contiene algún o algunos elementos de G fijados de antemano. Diremos que H es el subgrupo de G generado por ese o esos elementos fijos en G. Si está generado por un solo elemento, el subgrupo H es cíclico, y si está generado por varios elementos fijados de antemano en G, entonces H está formado por esos elementos y por todos los productos de un número finito de factores del conjunto en cuestión con o sin exponentes naturales y asociados de todas las maneras posibles.

Ejemplos de subgrupos. 1. Anteriormente hemos dicho que los conjuntos de la forma $m\mathbb{Z}$ = múltiplos del entero positivo m, son subgrupos del grupo aditivo $\mathbb{Z}$. Ahora probaremos que ellos son los únicos subgrupos del grupo aditivo $\mathbb{Z}$, formado por todos los enteros.

En efecto, el que $m\mathbb{Z}$ sea un subgrupo de $\mathbb{Z}$ es una consecuencia inmediata de los criterios que hemos dado para subgrupos.

Recíprocamente, sea H un subgrupo de $\mathbb{Z}$ que contiene algún entero distinto de cero, es decir que H no sea un subgrupo trivial, entonces H contiene algún entero positivo.

En efecto, si $o \neq a \in H$, entonces por ser H un subgrupo resulta que $(-a) \in H$.

Ahora bien, como uno de los elementos a y $(-a)$ es positivo, concluimos que H contiene al menos un entero positivo; esto es:

$$H \cap \mathbb{N} \neq \varnothing$$

Ahora, siendo $H \cap \mathbb{N}$ un subconjunto no vacío de $\mathbb{N}$, entonces, por el Principio del mínimo entero positivo, existe $m \in H \cap \mathbb{N}$ tal que,

$$h \in H \cap \mathbb{N} \Rightarrow m \leq h$$

cualquiera sea h en el subconjunto $H \cap \mathbb{N}$.

Es claro que $m \in H$, y se tiene además:

$$m\mathbb{Z} \subseteq H \qquad (1)$$

Sea ahora $x \in H$ arbitrario; entonces, por el algoritmo de divisón, se puede escribir:

$$x = qm + r, \qquad o \leq r < m$$

De esta última relación resulta,

$$r = x - qm$$

Ahora bien, como x y qm son elementos de H, también lo será su diferencia, es decir r.

Por otra parte, sabemos que m es el menor elemento positivo de H, y por consiguiente, si $r < m$ está en H, debe ser necesariamente $r = o$, por la minimidad de m. De donde,

$$x = qm$$

Luego, x está en $m\mathbb{Z}$, o sea

$$H \subseteq m\mathbb{Z} \qquad (2)$$

De (1) y (2) se concluye que,

$$H = m\mathbb{Z}$$

y lo que demuestra nuestra aseveración.

Se suele escribir $m\mathbb{Z} = (m)$ y decimos que $m\mathbb{Z} = (m)$ es el subgrupo de $\mathbb{Z}$ generado por m.

2. Sea G un rgrupo y sea A un semigrupo con unidad e'.

Sea, además, $f : G \longrightarrow A$ un homomorfismo. Probaremos que se verifica:

a) Nuc (f) es un subgrupo de G.

b) Im(f) es un subgrupo de A.

Dem. a) Sean $x, y \in$ Nuc(f), entonces por definición de núcleo se tiene:

$$f(x) = e', \qquad f(y) = e'$$

Por otra parte, por ser f un homomorfismo, se escribe:

$$f(xy^{-1}) = f(x) \cdot f(y^{-1}) = (f(x) \cdot [f(y)]^{-1}$$

$$= e' \cdot e'^{-1} = e' \cdot e' = e'$$

luego, $xy^{-1} \in$ Nuc(f)

Por consiguiente, el Nuc(f) es un sybgrupo de G.

b) Que Im(f) es un grupo, es consecuencia inmediata del hecho de que algunas propiedades algebraicas se conservan en un homomorfismo ((revíselas), Tomo I)

Luego, podemos decir que la imagen homomorfa de un grupo es un grupo, en este caso un subgrupo del semigrupo ;A.

7.9. *Equivalencias Regulares en Grupos.*
Clases a izquierda y a derecha determinadas por un subgrupo en un grupo.

Sea G un grupo y $H \subset G$ un subgrupo de G.

Si $x, y \in G$ son tales que $x^{-1} y \in H$, escribiremos:

$$x \sim y \iff x^{-1}y \in H$$

o también, $x \equiv y$ (mód. H)

Es decir que, $x \equiv y$ (mód. H) si $x^{-1}y = h \in H$, o bien si $y = xh$ siendo $h \in H$.

Probemos que esta relación es verdaderamente una equivalencia, y para lo cual debemos hacer ver que tiene las tres propiedades que caracterizan a las equivalencias: reflexividad, simetría y transitividad.

En efecto,

a) $\forall x \sim x$, ya que $x^{-1}x = e \in H$

b) Sea $x \sim y$, es decir $x^{-1}y \in H \Rightarrow (x^1 y)^{-1} = y^{-1}x \in H \Rightarrow y \sim x$

c) Supongamos ahora que se tengan,

$$x \sim y \quad e \quad y \sim z$$

es decir, $x^{-1}y \in H \quad e \quad y^{-1}z \in H$

luego,

$$(x^{-1}y) \cdot (y^{-1}z) = x^{-1}(y\,y^{-1})z = x^{-1}z \in H \Rightarrow x \sim z$$

Por lo tanto, nuestra relación »$\sim$«, ó , $\equiv$ (mód. H) es una relación de equivalencia en el grupo considerado G.

Demostraremos, además, que también esta relación es regular a la izquierda, esto es:

$$x \sim y \Rightarrow a\,x \sim a\,y, \ \forall\, a \in G$$

En efecto, tenemos:

$$(ax)^{-1} \cdot ay = (x^{-1}a^{-1}) \cdot ay = x^{-1}(a^{-1}a)y = x^{-1}y \in H$$

y lo que implica que:

$$ax \sim ay$$

luego, es regular a la izquierda.

Recíprocamente, sea dada una equivalencia regular a la izquierda »$\sim$« en G.

Denotemos por H la clase de equivalencia del elemento *e*.

Afirmamos que H es un subgrupo de G.

En efecto, H no es vacío. Además supongamos $x, y \in H$.

Esto implica:

$$e \sim x, \qquad e \sim y$$

Multiplicando a la izquierda, la primera de estas relaciones, por x^{-1} obtenemos:

$$x^{-1}e = x^{-1}x, \text{ es decir, } x^{-1} \sim e \Rightarrow x^{-1} \in H$$

Finalmente, multiplicando, siempre la primera de estas relaciones, a la izquierda por y, tenemos:

$$ye \sim yx, \text{ es decir, } y \sim yx \Rightarrow yx \in H$$

Luego, H es un subgrupo.

Así hemos probado que existe una correspondencia biunívoca entre el conjunto de todas las relaciones de equivalencias regulares a la izquierda en G, y el conjunto de todos los subgrupos $H \subset G$, de modo que:

$$x \sim y \iff x^{-1}y \in H$$

(H clase de equivalencia de G).

Análogamente, con demostración fiel a la que se hizo para las equivalencias regulares a la izquierda, podemos decir también que existe una correspondencia biunívoca entre el conjunto de todas las equivalencias regulares a la derecha en G y el conjunto de todos los subgrupos $H \subset G$ de modo que:

$$x \sim y \iff xy^{-1} \in H$$

(H clase de equivalencia de G)

Sea »$\sim$« la equivalencia regular a la izquierda:

$$x \sim y \iff x^{-1}y \in H$$

Surge la pregunta: ¿qué es la clase de equivalencia de un elemento x?

Decir que $x \sim x'$ equivale a decir que,

$$x^{-1}x' = h \in H$$

o sea, $x' = xh, \quad h \in H$

Así pues, la clase de equivalencia de x es el conjunto de todos los elementos de la forma xh, donde h se pasea en H.

Es natural designar este conjunto por xH.

Estas clases de equivalencia xH las llamaremos COGRUPOS a la IZQUIERDA del grupo G relativo al subgrupo H.

Estos cogrupos forman una partición de G. Cada elemento de G pertenece a un cogrupo y solamente a uno.

Análogamente, se definen los conjuntos de la forma Hx, que denominaremos COGRUPOS a la DERECHA del grupo G relativo al subgrupo H.

Los cogrupos a la derecha forman también una partición de G, y cada elemento de G pertenece a un cogrupo a la derecha y solamente a uno.

Observaciones importantes. 1. Si H es finito, todos los cogrupos xH y Hx tienen el mismo número de elementos y todos son distintos.

En efecto, xH se deduce de H por la aplicación:

$$f : H \to xH, \qquad x \text{ fijo en } G$$

definida por, $f(h) = xh, \qquad \forall\, h \in H$

y como esta aplicación es biyectiva (pruébela), resulta que H y xH tienen el mismo número de elementos.

Todos los elementos de xH son diferentes, porque si $xh = xh'$, con $h \neq h'$ implica, por la ley de cancelación que se verifica siempre en un grupo que $h = h'$, lo que es un absurdo.

Esta contradicción, muestra que los elementos del cogrupo xH son todos distintos.

Análogamente se prueba que también los cogrupos a la derecha Hx tienen el mismo número de elementos, mediante la aplicación:

$$f : H \to Hx, \quad x \text{ fijo en } G$$

tal que, $f(h) = hx$

y además, todos ellos son distintos.

2. El número de cogrupos a la derecha es igual al número de cogrupos a la izquierda (siempre que uno de ellos sea finito).

En efecto, consideremos la simetría de G, es decir, la aplicación $x \to x^{-1}$.

Ella transforma xH en Hx^{-1}, lo cual significa que si $h \in H$, entonces:

$$xH \ni xh \Rightarrow (xh)^{-1} = h^{-1}x^{-1} \in Hx^{-1}$$

esto es, si h se pasea en H, su inverso h^{-1} también se pasea en H (porque H es grupo).

Luego, la transformación o aplicación:

$$x \to x^{-1}$$

transforma un cogrupo a la izquierda en un cogrupo a la derecha. La conclusión sigue del hecho de que la simetría de G es aplicación biyectiva.

3. Luego, el grupo dado G queda dividido, tanto en uno como en el otro caso, en clases disjuntas x H (ó H x) tales que cada una de ellas se compone de todos los elementos equivalentes o congruentes con uno de ellos; es decir, si $a \in x\ H$ y $a \sim b$ (ó $a \equiv b$ (mód. H)), entonces $b \in x\ H$. El subgrupo H mismo es una clase: la clase de todos los elementos equivalentes o congruentes con la unidad e, es decir tales que $e\ h = h \in H$.

4. Cuando, en el caso particular, el grupo G es finito, también el subgrupo $H = \{h_1, h_2, \ldots, h_p\}$ será finito, y lo será también el número de cogrupos a la izquierda y a la derecha. Este último número lo llamaremos INDICE de H en G, y lo designaremos por $G : H = i$.

Si el número de cogrupos es infinito, diremos que el índice es infinito.

5. Supongamos que G sea finito de orden n. Sea p el orden del subgrupo H = número de elementos en cada cogrupo.

Sea *i* el índice = número de cogrupos. Por lo tanto,

$$n = p\,i$$

De modo que se puede escribir:

$$\frac{\text{orden de } G}{\text{orden de } H} = G : H$$

Así pues, en este caso, el grupo G se reparte en *i* clases de equivalencia de *p* elementos cada una, resultando el siguiente cuadro:

$H : h_1, \ldots, h_p$
$H_1 : a_1 h_1, \ldots, a_1 h_p$
$H_2 : a_2 h_1, \ldots, a_2 h_p$
. . . . . .
. . . . . .
. . . . . .
$H_{i-1} : a_{i-1} h_1, \ldots, a_{i-1} h_p$

Este cuadro o esquema se denomina CUADRO DE LAGRANGE de las *clases laterales módulo H*. Como todo elemento de G está en un H_j, y dos H_j no tienen elementos comunes, resulta como lo vimos anteriormente:

$$n = p\,i$$

Obtenemos así el importante teorema:

Teorema de Lagrange. En un grupo finito, el orden de cualquier subgrupo divide al orden del grupo.

Corolario 1. Si G es un grupo finito de orden n, entonces el orden de cada elemento de G es un divisor de n.

Dem. Sea *a* un elemento de G de orden m; entonces G contiene los elementos:

$$a^0 = e, a, a^2, \ldots, a^{m-1}, (a^m = e)$$

los cuales forman un subgrupo cíclico de orden m. Por consiguiente, por el teorema de Lagrange, m es divisor de n.

Corolario 2. Un grupo de orden primo no tiene subgrupos propios y es necesariamente cíclico.

Dem. Si el orden del grupo es un número primo p, entonces el orden del subgrupo tiene que ser, ó 1, ó p, es decir, el subgrupo se compone o de {e} como único elemento o contiene todos los elementos p del grupo.

Si *a* es un elemento distinto de *e*, su orden, siendo mayor que 1, ha de ser necesariamente igual a *p*. Por consiguiente, los p elementos $a^0 = e$, $a, a^2, \ldots, a^{p-1}$, $(a^p = e)$ son todos los elementos del grupo considerado en otro orden. El grupo es, pues, cíclico.

Ejemplos

1. Consideremos el grupo cíclico:

$$C_{12} = \{e, a, a^2, a^3, a^4, a^5, a^6, a^7, a^8, a^9, a^{10}, a^{11}\}$$

Sea el subgrupo de orden tres, $H = \{e, a^4, a^8\}$
Los cuatro grupos o clases laterales de H son:

$H = \{e, a^4, a^8\}$
$aH = \{a, a^5, a^9\} = a^5 H = a^9 H$
$a^3 H = \{a^3, a^7, a^{11}\} = a^7 H = a^{11} H$
$a^6 H = \{a^6, a^{10}, a^2\} = a^2 H = a^{10} H$

Nótese que los elementos que multiplican al subgrupo H no están unívocamente determinados, porque si h es un elemento cualquiera de H, entonces:

$$xH = x\,(h\,H) = (x\,h)\,H$$

Así, pues, para nuestros fines, x puede sustituirse por xh.

2. Sea el grupo de orden seis dado por la tabla siguiente:

	e	a	b	c	d	f
e	e	a	b	c	d	f
a	a	b	e	f	c	d
b	b	e	a	d	f	c
c	c	d	f	e	a	b
d	d	f	c	b	e	a
f	f	c	d	a	b	e

El conjunto $H = \{e, c\}$ es un subconjunto de orden 2, ya que:

$$H^2 = H \cdot H = \{e, c\} \cdot \{e, c\} = \{e, c, c, c^2\}$$
$$= \{e, c, c, e\} = \{e, c\} = H$$

El índice de H en el grupo dado es $i = \frac{6}{2} = 3$; luego hay tres cogrupos, y los elementos x pueden tomarse como e, a, b, respectivamente, resultando los cogrupos:

$$H = \{e, c\}$$
$$aH = \{a, ac\} = \{a, f\} = fH$$
$$bH = \{b, bc\} = \{b, d\} = dH$$

que constituyen una partición del grupo considerado. Tendremos otra partición diferente si utilizamos los cogrupos a la derecha.

3. Consideremos el grupo aditivo $\mathbf{Z}$. Sabemos que todos los subgrupos son los conjuntos $m\mathbf{Z}$, $m = o, 1, 2, \ldots$ Consideremos uno de estos subgrupos, es decir m dado.

La equivalencia regular correspondiente es:

$$x \sim y \Longleftrightarrow x^{-1} y \in H$$

se traduce en este caso por:

$$x \sim y \Longleftrightarrow -x + y \in m\mathbf{Z} \Rightarrow y - x = m$$

que es la congruencia aritmética módulo m.

Los cogrupos son las clases residuales módulo m, y el número de ellas, es decir el índice, es:

$$\mathbf{Z} : m\mathbf{Z} = m \quad \text{(finito)}$$

4. Sea S_n el grupo simétrico de n elementos y sea A_n el grupo alternado. Entonces, A_n es subgrupo de S_n de índice dos, ya que $S_n : A_n = n! : \frac{1}{2} n! = 2$.

Luego, los dos cogrupos determinados por A_n en S_n son:

A_n = conjunto de todas las sustituciones pares.

$S_n - A_n$ = conjunto de todas las sustituciones impares.

7.10. *Subgrupo Normal*

Para que una equivalencia "$\sim$" regular a la izquierda (o regular a la derecha) asociada con un subgrupo H de un grupo G sea regular, es necesario y suficiente que los cogrupos a la derecha coincidan con los cogrupos a la izquierda, y viceversa; esto es:

$$xH = Hx, \quad \forall\, x \in G$$

Es decir, si para todo $x \in G$ y para todo $h \in H$ existe $h' \in H$ tal que,

$$xh = h'x$$

lo que equivale a que para todo $h \in H$ y todo $x \in G$ sea

$$xhx^{-1} = h' \in H$$

es decir, $xhx^{-1} \in H$, $\forall h \in H$ y $\forall x \in G$

Este hecho se escribe genéricamente en la forma:

$$xHx^{-1} = H, \quad \forall x \in G$$

Un subgrupo H que tiene esta propiedad lo llamaremos SUBGRUPO NORMAL o INVARIANTE, pues queda invariante al multiplicarlo por x a la izquierda y por x^{-1} a la derecha.

Si el grupo G es conmutativo, entonces todo subgrupo de él es invariante o normal, puesto que:

$$xhx^{-1} = xx^{-1}h = eh = h$$

o bien, teniendo en cuenta que $xh = hx$

En resumen, hay correspondencia biunívoca entre el conjunto de todas las equivalencias regulares (si es a la izquierda y a la derecha) en G y el conjunto de todos los subgrupos normales de G.

Observación importante. Más arriba definimos las relaciones de equivalencia o de congruencia:

$$1)\ x \sim y \Longleftrightarrow x^{-1}y \in H \Rightarrow x^{-1}y = h \Rightarrow y = x\,h$$
$$2)\ x \sim y \Longleftrightarrow x\,y^{-1} \in H \Rightarrow x\,y^{-1} = h \Rightarrow x = y\,h$$

donde h es un elemento del subgrupo H, y obtuvimos las clases laterales (cogrupos) a la izquierda x H y a la derecha H y. Estas clases dan lugar a dos particiones diferentes del grupo considerado G, es decir, a dos cuadros de Lagrange distintos.

Sean ahora H_1 y H_2 dos clases laterales arbitrarias de una misma partición de G, y sean:

$$x_1, y_1 \in H_1 \quad , \quad x_2, y_2 \in H_2$$

o sea, $x_1 \sim y_1$, $x_2 \sim y_2$

En general, no podemos afirmar que:

$$x_1\,x_2 \sim y_1\,y_2$$

es decir, no podemos afirmar que los elementos $x_1\,x_2$ e $y_1\,y_2$ pertenezcan a una misma clase de equivalencia H_3. Pero, si H es un subgrupo normal, sí podemos afirmarlo. En efecto, si $x_1, y_1 \in H_1$, podremos escribir:

$$x_1 = h_1\,y_1$$

análogamente, si $x_2, y_2 \in H_2$, escribimos:

$$x_2 = h_2\,y_2$$

luego, $x_1\,x_2 = (h_1\,y_1)\,(h_2\,y_2) = h_1\,(y_1\,h_2)\,y_2$

y como H es subgrupo normal, entonces $y_1\,h_2 = h_3\,y_1$,

luego, $x_1\,x_2 = h_1\,(h_3\,y_1)y_2 = (h_1\,h_3)y_1\,y_2 = h_4\,(y_1\,y_2)$,

por lo tanto:

$$x_1\,x_2 \sim y_1\,y_2$$

Es decir, el subgrupo normal H tiene la importante propiedad que la operación producto *no separa las clases.*

Esta propiedad nos permite definir el producto de dos clases, por la siguiente regla:

$$x\,H \cdot y\,H = x\,y\,H$$

y también, $H\,x \cdot H\,y = H\,x\,y$

Esta definición es unívoca (es decir, depende únicamente de las clases y no de elementos individuales escogidos en ellas) ya que la equivalencia "$\sim$" es regular.

Probaremos que con esta definición de producto de clases, el conjunto de todos los cogrupos forman un nuevo grupo, que designaremos por G/H.

En efecto, sea la aplicación de G sobre G/H definida por:

$$f(x) = x\,H$$

y entonces, la relación $x\,H \cdot y\,H = x\,y\,H$ equivale a:

$$f(x\,y) = f(x) \cdot f(y)$$

Luego, tendremos:

$$\begin{aligned}(x\,H \cdot y\,H) \cdot z\,H &= [\,f(x)\cdot f(y)\,]\,f(z) = f(x\,y)\cdot f(z) = \\ &= f\,[\,(x\,y)\,z\,] = [\,f\,x(y\,z)\,] = f(x)\cdot f(y\,z) = \\ &= f(x)\,[\,f(y)\cdot f(z)\,] = \\ &= x\,H \cdot (y\,H \cdot z\,H)\ \text{(asociatividad)}\end{aligned}$$

El elemento neutro de G/H es $f(e) = eH = H$.

El inverso de $f(x) = x\,H$ es:

$$f(x^{-1}) = x^{-1}\,H$$

Así hemos demostrado que el conjunto G/H es un grupo que llamaremos GRUPO COCIENTE de G por el subgrupo normal H.

Por otro lado, como la aplicación de G sobre G/H definida por, $f(x) = x\,H$ preserva el producto, concluimos que el grupo cociente G/H es una imagen homomórfica de G.

Ejemplos

1. Sea el grupo aditivo $\mathbb{Z}$ de los enteros racionales y sea $m\,\mathbb{Z}$ un subgrupo de $\mathbb{Z}$. Siendo $\mathbb{Z}$ un grupo conmutativo, entonces el subgrupo $m\,\mathbb{Z}$ es normal.

Como la equivalencia "$\sim$" asociada con $m\,\mathbb{Z}$ es:

$$x \sim y \Longleftrightarrow -x + y \in m\mathbb{Z} \Rightarrow y - x = m$$

es decir, la congruencia módulo m, entonces los cogrupos determinados por $m\mathbb{Z}$ en $\mathbb{Z}$ son las clases residuales:

$$\{0, 1, 2, \ldots, m-1\}$$

y las cuales constituyen el grupo-cociente, y que ahora llamaremos *grupo diferencia*, que denotamos por: $\mathbb{Z} - m\,\mathbb{Z}$. El elemento unidad es la clase 0 y el inverso de la clase i es $n - i$.

Este grupo diferencia es una imagen homomórfica de $\mathbb{Z}$.

2. La tabla siguiente define un grupo:

	A	B	C	D	E	F
A	A	B	C	D	E	F
B	B	A	D	C	F	E
C	C	F	A	E	D	B
D	D	E	B	F	C	A
E	E	D	F	B	A	C
F	F	C	E	A	B	D

Entonces:

a) $\{A, E\}$ y $\{A, D, F\}$ son subgrupos.
b) $\{A, E\}$ no es subgrupo normal.
c) $\{A, D, F\}$ es subgrupo normal.
d) Construir el grupo cociente $\{A, B, C, D, E, F\} / \{A, D, F\}$

En efecto, para probar que los conjuntos indicados en a) son subgrupos, basta aplicar el criterio que tenemos para subgrupos finitos; esto es, que ellos son cerrados para la multiplicación del grupo.

Tenemos:

$$\begin{aligned}\{A, E\}^2 &= \{A, E\} \cdot \{A, E\} = \{A^2, AE, EA, E^2\} \\ &= \{A, E, E, A\} = \{A, E\}\end{aligned}$$

luego, es grupo (subgrupo)

$$\begin{aligned}\{A, D, F\}^2 &= \{A, D, F\} \cdot \{A, D, F\} \\ &= \{A^2, AD, AF, DA, D^2, DF, FA, FD, F^2\} \\ &= \{A, D, F, D, F, A, F, A, D\} = \{A, D, F\}\end{aligned}$$

luego, es subgrupo.

Así hemos probado la afirmación a).

Demostraremos en seguida que el subgrupo $\{A, E\}$ no es normal. Para que lo fuese $\{A, E\} = H$ debe implicar:

$$x\,H\,x^{-1} = H \quad , \quad \forall\, x \in \{A, B, C, D, E, F\}$$

Tenemos:

$$\left.\begin{matrix} B \\ C \\ D \\ F \end{matrix}\right\} = \text{elementos que no están en } H = \{A, E\}$$

Probemos con B:

$$B\,H\,B^{-1} = \begin{cases} B\,A\,B^{-1} = B \cdot B = A \in H \\ B\,E\,B^{-1} = F \cdot B = C \notin H \end{cases}$$

Luego, por este solo hecho, ya $H = \{A, E\}$ no es normal.
Así hemos demostrado la afirmación b).
Ahora probaremos que el subgrupo $H = \{A, D, F\}$ es normal.
Los elementos que faltan en $\{A, D, F\} = H$ son B, C, E.
Probemos con B:

$$B\,H\,B^{-1} = \begin{cases} B\,A\,B^{-1} = B \cdot B = A \in H \\ B\,D\,B^{-1} = C \cdot B = F \in H \\ B\,F\,B^{-1} = E \cdot B = D \in H \end{cases}$$

Probemos con C:

$$C\,H\,C^{-1} = \begin{cases} C\,A\,C^{-1} = C \cdot C = A \in H \\ C\,D\,C^{-1} = E \cdot C = F \in H \\ C\,F\,C^{-1} = B \cdot C = D \in H \end{cases}$$

Probemos con E:

$$E\,H\,E^{-1} = \begin{cases} E\,A\,E^{-1} = E \cdot E = A \in H \\ E\,D\,E^{-1} = B \cdot E = F \in H \\ E\,F\,E^{-1} = C \cdot E = D \in H \end{cases}$$

Siendo $x\,H\,x^{-1} = H = \{A, D, F\}$ para todo $x \in \{A, B, C, D, E, F\}$, concluimos que este subgrupo $H = \{A, D, F\}$ es normal.

Así hemos demostrado la afirmación c).

Finalmente, determinemos los cogrupos que este subgrupo determina sobre el grupo considerado. Tenemos:

$$\begin{aligned} A\,H &= H = \{A, D, F\} \\ B\,H &= \{BA, BD, BF\} = \{B, C, E\} \\ C\,H &= \{CA, CD, CF\} = \{C, E, B\} = B\,H \\ D\,H &= H \\ E\,H &= \{EA, ED, EF\} = \{E, B, C\} = B\,H = C\,H \\ F\,H &= H \end{aligned}$$

Luego, hay solamente dos cogrupos distintos y que son:

$$H = \{A, D, F\} \quad y \quad \{B, C, E\}$$

porque el índice de H en G vale:

$$i = G : H = \frac{\text{orden de } G}{\text{orden de } H} = \frac{6}{3} = 2$$

Por consiguiente, el grupo cociente es:

$$\{A, B, C, D, E, F\} / \{A, D, F\} = \{\ \{A, D, F\}, \{B, C, E\}\ \}$$

3. Los dos subgrupos impropios $\{e\}$ y G de un grupo cualquiera G son ejemplos triviales de subgrupos normales.

Los grupos cocientes determinados por ellos son:

$$G/\{e\} = G, \quad G/G = \{e\}$$

4. Sea S_3 el grupo simétrico de tres elementos. Los elementos de este grupo son, como sabemos, las seis sustituciones:

$$s_1 = \begin{pmatrix} 1\,2\,3 \\ 1\,2\,3 \end{pmatrix}, \; s_2 = \begin{pmatrix} 1\,2\,3 \\ 1\,3\,2 \end{pmatrix}, \; s_3 = \begin{pmatrix} 1\,2\,3 \\ 2\,1\,3 \end{pmatrix}, \; s_4 = \begin{pmatrix} 1\,2\,3 \\ 2\,3\,1 \end{pmatrix},$$

$$s_5 = \begin{pmatrix} 1\,2\,3 \\ 3\,1\,2 \end{pmatrix}, \; s_6 = \begin{pmatrix} 1\,2\,3 \\ 3\,2\,1 \end{pmatrix}$$

Puede verificarse fácilmente que el subgrupo formado por las tres sustituciones s_1, s_4, s_6, es un subgrupo normal, no obstante que S_3 no es un grupo conmutativo.

Nótese que este subgrupo es el grupo alternado A_3.

En general, el grupo alternado A_n es un subgrupo normal o invariante del grupo simétrico S_n de todas las sustituciones de n elementos.

El índice de este subgrupo es siempre 2, ya que:

$$i = S_n : A_n = \frac{\text{orden de } S_n}{\text{orden de } A_n} = \frac{n!}{\frac{1}{2}\,n!} = 2$$

El teorema siguiente expresa una propiedad muy simple que es de uso frecuente:

Teorema. Un subgrupo de índice 2 es siempre subgrupo normal o invariante.

Dem. Sea H un subgrupo de un grupo cualquiera G y de índice 2. Entonces, existen solamente dos cogrupos de G respecto a H. De manera que si x es un elemento arbitrario de G que no pertenece a H, entonces ni x H ni Hx tienen elementos comunes con H; esto es:

$$\text{si } x \notin H, \text{ entonces } xH \cap H = \varnothing \text{ y } Hx \cap H = \varnothing$$

En efecto, si por el contrario, fuese $xh = h'$, $h, h' \in H$ se tendría $x = h' \cdot h^{-1} \in H$, absurdo, ya que $x \notin H$.

Por consiguiente, el grupo G se reparte así:

$$G = H + xH \quad y \quad G = H + Hx$$

y de aquí se deduce:

$$xH = Hx, \quad \forall\ x \in G$$

o bien, $x\,H\,x^{-1} = H, \quad \forall \quad x \in G$

resultado que muestra que H es un subgrupo invariante o normal.

De lo anterior, se concluye que A_n es un subgrupo invariante de S_n, y las dos clases del grupo cociente son el subgrupo A_n de las sustituciones pares y la clase $S_n - A_n$ de las sustituciones impares.

7.11. *Isomorfismo y Homomorfismo de grupos*

En la teoría de los grupos abstractos, en la que no se hace referencia alguna a la naturaleza de los elementos, considera un grupo sólo desde el punto de vista de la operación definida en él e interesa únicamente saber cuál es el elemento que resulta de la composición de dos elementos dados. Este enfoque está claramente expresado en las definiciones que siguen a continuación y que resulta de aplicar a los grupos el concepto

general de isomorfismo y homomorfismo entre dos estructuras algebraicas.

Definición 1. Una aplicación:

$f : G \to G'$

de un grupo (G; o) en un grupo (G'; *), es un *isomorfismo* si:

a) f es biyectiva

b) f preserva las operaciones de grupo; esto es:

$f(x \circ y) = f(x) * f(y), \quad x, y \in G$

Es claro que si f: G → G' es un isomorfismo, también la aplicación $f^{-1}: G' \to G$ es un isomorfismo.

Diremos que dos grupos G y G' son *isomorfos* si existe un isomorfismo f: G → G' (o al revés).

Los grupos isomorfos tienen la misma estructura, aunque pueden diferir en la notación y naturaleza de sus elementos. Por lo tanto, ellos son idénticos desde el punto de vista de la teoría abstracta de grupos. Esta teoría abstracta de grupos estudia solamente propiedades invariantes con respecto a los isomorfismos.

Un isomorfismo f : G → G de un grupo G en sí mismo, lo denominaremos *automorfismo.*

Teorema. El conjunto Aut (G) de todos los automorfismos de un grupo G es a su vez un grupo con respecto a la composición de aplicaciones.

Dem. Sea Aut (G) el conjunto de todos los automorfismos del grupo G. Si f ∈ Aut (G), entonces f es una biyección de G sobre G que preserva (respeta) el producto.

1. Cierre o clausura.

Probemos que $f_1, f_2 \in \text{Aut}(G) \iff f_1 \circ f_2 \in \text{Aut}(G)$.

Es claro que siendo f_1, f_2 biyectivas, entonces la compuesta o producto $f_1 \circ f_2$ es también biyectiva. Solamente nos falta probar que $f_1 \circ f_2$ preserva el producto.

En efecto, pongamos $f = f_1 \circ f_2$, tenemos:

$$f(a \cdot b) = (f_1 \circ f_2)(a \cdot b) = f_1(f_2(a\,b)) = f_1(f_2(a) \cdot f_2(b))$$
$$= f_1(f_2(a)) \cdot f_1(f_2(b)) =$$
$$= (f_1 \circ f_2)(a) \cdot (f_1 \circ f_2)(b) = f(a) \cdot f(b)$$

por lo tanto, $f = f_1 \circ f_2 \in \text{Aut}(G)$.

2. Propiedad asociativa.

Es válida en general para toda aplicación; luego,

$(f_1 \circ f_2) \circ f_3 = f_1 \circ (f_2 \circ f_3)$.

3. Elemento unidad.

Es la aplicación idéntica I_G, que es biyectiva y preserva el producto $a \cdot b = I_G(a \cdot b) = I_G(a) \circ I_G(b) = a \cdot b$; luego $I_G \in \text{Aut}(G)$.

Además, se tiene:

$f \circ I_G = I_G \circ f = f, \ \forall f \in \text{Aut}(G)$.

4. Elemento inverso.

Deberemos probar que $f \in \text{Aut}(G) \Rightarrow f^{-1} \in \text{Aut}(G)$.

En efecto, f^{-1} existe por ser f biyectiva, siendo también f^{-1} una biyección de G sobre G.

Demostremos que f^{-1} preserva el producto.

Sean a', b' ∈ G; luego existen a, b ∈ G tales que:

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

Es claro que $f^{-1}(a') = a$, $f^{-1}(b') = b$, entonces:

$$f^{-1}(a' \cdot b') = f^{-1}(f(a) \cdot f(b)) = f^{-1}(f(a \cdot b)) = a \cdot b =$$
$$= f^{-1}(a') \cdot f^{-1}(b')$$

Este resultado y los anteriores demuestran el teorema.

7.12. *Definición*

Sean a, x ∈ G, donde G es un grupo; entonces, diremos que el elemento $a\,x\,a^{-1}$ es el TRANSFORMADO de *x* por *a*. El elemento $a\,x\,a^{-1}$ (o bien, $a^{-1}\,x\,a$) recibirá también el nombre de CONJUGADO del elemento *x* por *a*.

La relación de conjugación es una relación de equivalencia; esto es:

$x_1 \sim x \iff x_1 = a\,x\,a^{-1}$

En efecto, tenemos:

a) $x \sim x \iff x = e\,x\,e^{-1}$

b) $x_1 \sim x_2 \iff x_1 = a\,x_2\,a^{-1} \Rightarrow a^{-1}\,x_1\,(a^{-1})^{-1} = x_2 \Rightarrow x_2 \sim x_1$

c) $x_1 \sim x_2$ y $x_2 \sim x_3 \Rightarrow x_1 \sim x_3$, ya que $x_1 = a\,x_2\,a^{-1}$ y $x_2 = b\,x_3\,b^{-1} \Rightarrow x_1 = a\,(b\,x_3\,b^{-1})a^{-1} = (a\,b)x_3\,(b^{-1}\,a^{-1}) =$
$= (ab)x_3\,(a\,b)^{-1} \Rightarrow x_1 \sim x_3$

Luego, la conjugación es una relación de equivalencia.

Por consiguiente, se puede descomponer el grupo G en clases disjuntas. Contrariamente, a lo que sucede con los cogrupos, estas nuevas clases de equivalencia no tienen necesariamente el mismo número de elementos. En particular, el elemento unidad e no es conjugado más que de sí mismo.

Proposición. Los transformados o conjugados de los elementos de un subgrupo $H = \{e, h_1, h_2, \ldots, h_p\}$ de un grupo G por un elemento fijo $a \in G$ forman un subgrupo llamado subgrupo conjugado de H.

Dem. Los elementos $a\ e\ a^{-1} = e$, $a\ h_1\ a^{-1} = h'_1$, $a\ h_2\ a^{-1} = h'_2, \ldots$ forman un conjunto denotado por $a\ H\ a^{-1}$.

Probaremos que $a\ H\ a^{-1}$ es un grupo (subgrupo de G). En efecto:

$$h'_1 \cdot h'_2 = (a\ h_1\ a^{-1})\ (a\ h_2\ a^{-1}) = a\ h_1\ (a^{-1}\ a)\ h_2\ a^{-1} = a\ h_1\ h_2\ a^{-1}$$

y como $h_1\ h_2 \in H$, entonces $h'_1\ h'_2 \in a\ H\ a^{-1}$

Por otro lado, $h'^{-1}_1 = (a\ h_1\ a^{-1})^{-1} = (a^{-1})^{-1}\ h_1^{-1}\ a^{-1} = a\ h_1^{-1}\ a^{-1}$, y como $h_1^{-1} \in H$, entonces $h'^{-1}_1 \in a\ H\ a^{-1}$.

Así hemos probado que el conjunto $a\ H\ a^{-1}$ es un subgrupo de G.

En esta terminología tenemos:

Un subgrupo H es normal si, y sólo si, para todo $h \in H$ todos los transformados de h están en H; es decir, H contiene con cada uno de sus elementos a todos sus transformados.

En otras palabras, un subgrupo H que coincide con todos sus subgrupos conjugados ($H = a\ H\ a^{-1}, \forall\ a \in G$) es normal.

Proposición. Si $a \in G$ es un elemento fijo del grupo, entonces la aplicación T_a de G en G definida como sigue:

$$Ta : x \to T_a(x) = a\ x\ a^{-1}, \forall\ x \in G$$

es un automorfismo de G.

Dem. Probemos primeramente que T_a es una biyección.

Sea *y* dado arbitrariamente en G, entonces la ecuación siguiente:

$$a\ x\ a^{-1} = y \text{ (siendo x la incógnita)}$$

tiene siempre una solución única:

$$x = a^{-1}\ y\ a$$

luego, no puede haber dos x distintas que tengan la misma imagen *y*.

Demostremos ahora que la aplicación T_a preserva los productos. En efecto:

$$T_a(x\ y) = a\ (x\ y)a^{-1} = (a\ x\ a^{-1})\ (a\ y\ a^{-1}) = T_a(x) \cdot T_a(y).$$

Este resultado y el anterior demuestran la proposición.

La aplicación T_a la llamaremos AUTOMORFISMO INTERIOR asociado con *a*.

Nota. Con esta nueva terminología tenemos:

Un subgrupo H es normal si, y sólo si, H es *invariante* por cualquier automorfismo interior de G.

Es decir, cualquier automorfismo interior de G transforma H en él mismo, ya que la imagen de H mediante T_a es:

$$T_a(H) = a\ H\ a^{-1} = H, \text{ porque es normal.}$$

De este hecho proviene la denominación de llamar *invariante* a un subgrupo normal.

Proposición. Los automorfismos interiores de un grupo G constituyen un subgrupo del grupo de *todos* los automorfismos de G.

Dem. Tenemos que probar los siguientes hechos:

a) $T_a \circ T_b = T_{ab}$.

b) T_e = automorfismo idéntico.

c) $T_{a^{-1}} = (T_a)^{-1}$.

En efecto, tenemos:

$$(T_a \circ T_b)(x) = T_a\ (T_b\ (x)) = T_a\ (b\ x\ b^{-1}) = a\ (b\ x\ b^{-1})a^{-1} = (ab)\ x\ (b^{-1}\ a^{-1}) = (a\ b)\ x\ (a\ b)^{-1} = T_{ab}(x), \quad \forall\ x \in G.$$

luego, $T_a \circ T_a = T_{ab}$.

Por otra parte, $I_G(x) = e\ x\ e^{-1} = x$; luego, T_e es unidad.

Finalmente, $T_a(x) = y \iff a\ x\ a^{-1} = y \iff x = a^{-1}\ y\ a = T_{a^{-1}}(y)$, luego, $T_{a^{-1}} = (Ta)^{-1}$.

Este resultado y los anteriores prueban la proposición.

Nota. Los automorfismos que no son interiores los llamaremos *automorfismos externos.*

Si el grupo G es conmutativo, todos los automorfismos interiores de G coinciden con la transformación idéntica:

$$T_a(x) = a\ x\ a^{-1} = a\ (x\ a^{-1}) = a\ (a^{-1}\ x) = (a\ a^{-1})x = e\ x = x.$$

El concepto de isomorfismo de grupos admite una generalización muy importante que resulta de considerar una aplicación epiyectiva ("sobre") de G dentro de G', pero no necesariamente inyectiva.

Esta consideración nos mueve a dar la segunda definición siguiente:

Definición 2. Una aplicación:

$$f : G \to G'$$

de un grupo (G; o) en un grupo (G'; *), es un *homomorfismo* si:

a) f es epiyectiva.

b) f preserva las operaciones de grupo; esto es:

$$f(x \circ y) = f(x) * f(y), \quad \forall\ x, y \in G.$$

Si existe un tal homomorfismo, entonces diremos que G'es una imagen *homomorfa* de G.

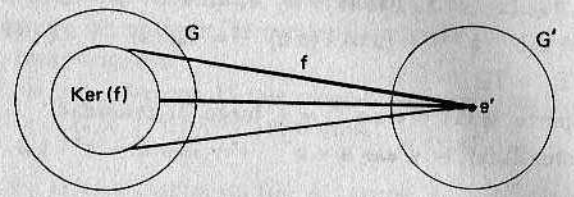
Un homomorfismo también puede ser de G "*en*" G' cuando es una aplicación de G "*en*" G' y preserva las operaciones de grupo.

Un homomorfismo $f : G \to G$ de un grupo G en si mismo, lo denominamos *endomorfismo.*

Por lo que vimos en el párrafo sobre el concepto general de homomorfismo entre dos estructuras algebraicas cualesquiera, podemos repetir ahora que, todo homomorfismo entre grupos transforma el elemento unidad e de G en el elemento unidad e' de G', y elementos inversos en elementos inversos.

Denifición. Llamaremos NUCLEO de un homomorfismo $f : g \to G'$ del grupo G sobre el grupo G' al subconjunto de N de G formado por los elementos que se transforman en el elemento unidad e' de G'.

Se le suele indicar al núcleo N con la notación Ker (f).



De modo que:

$$N = \text{Ker}(f) = \{x \in G : f(x) = e'\}.$$

Observación. Se cumple la equivalencia siguiente:

$$f(a) = f(b) \Longleftrightarrow a \cdot b^{-1} \in \text{Ker}(f).$$

En efecto, supongamos que $f(a) = f(b)$, entonces:

$$f(a \cdot b^{-1}) = f(a) \cdot f(b^{-1}) = f(a) \cdot [f(b)]^{-1} = f(b) \cdot [f(b)]^{-1} = e',$$

luego, $a \cdot b^{-1} \in \text{Ker}(f)$.

Recíprocamente, supongamos que $a \cdot b^{-1} \in \text{Ker}(f)$, entonces

$e' = f(a \cdot b^{-1}) = f(a) \cdot f(b^{-1}) = f(a) \cdot [f(b)]^{-1}$ pero, si $f(a) \cdot [f(b)]^{-1} = e'$, entonces:

$f(a) \cdot [f(b)]^{-1} f(b) = e' \cdot f(b)$

o sea, $f(a) = f(b)$.

Este resultado y el anterior prueban la equivalencia que afirmábamos. Esta equivalencia, equivale también a esta otra:

$$f(a) = f(b) \Longleftrightarrow f(a \cdot b^{-1}) = e'.$$

Teorema. El homorfismo $f : G \to G'$ es un isomorfismo si, y sólo si, Ker (f) $= \{e\}$.

Dem. a) Supongamos que f sea inyectiva y probemos que Ker(f) = e.

Sea $e \neq a \in \text{Ker}(f)$, entonces $f(a) = e' = f(e) \Longleftrightarrow a = e$, absurdo por la hipótesis hecha de que $a \neq e$.

Esta contradicción demuestra que si f es inyectiva, entonces el núcleo del homomorfismo f se reduce al elemento unidad e de G.

b) Recíprocamente, supongamos ahora que sea Ker(f) = e, y probemos que la aplicación f es inyectiva.

Sea f(a) = f(b), entonces por la equivalencia señalada anteriormente, tenemos:

$$a \cdot b^{-1} \in \text{Ker}(f) = \{e\}$$

es decir, $a \cdot b^{-1} = e \Rightarrow a = b$

luego, f es inyectiva.

El teorema está probado.

Teorema. El núcleo N = Ker(f) del homomorfismo $f : G \to G'$ es un subgrupo normal de G.

Dem. Probemos primero que N es un subgrupo de G.

Sean $n_1, n_2 \in N$, esto es $f(n_1) = e'$ y $f(n_2) = e'$, entonces:

$$f(n_1 \cdot n_2) = f(n_1) \cdot f(n_2) = e' \cdot e' = e'$$

luego, $n_1 \cdot n_2 \in N$.

Por otro lado, $f(n_1 \cdot n_2^{-1}) = f(n_1) \cdot f(n_2^{-1}) = f(n_1) \cdot [f(n_2)]^{-1} = e' \cdot e'^{-1} = e'$, luego, $n_1 \cdot n_2^{-1} \in N$.

Por lo tanto, N = Ker (f) es un subgrupo de G.

Demostremos ahora que N es normal. Tenemos,

$f(x\, n\, x^{-1}) = f(x) \cdot f(n)\, f(x^{-1}) = f(x) \cdot e' \cdot f(x^{-1}) =$

$= f(x) \cdot [f(x)]^{-1} = e'.$

luego, $x\, n\, x^{-1} \in N$, y N = Ker(f) es un subgrupo normal, como se quería demostrar.

Observación. En el caso de que el homomorfismo $f : G \to G'$ no es un homomorfismo epiyectivo, es decir, la aplicación f de G a G' es "en", entonces puede probarse que la imagen f(G) es un subgrupo de G'.

En efecto, sean f(x), f(y) dos elementos arbitrarios de f(G); entonces:

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

Luego, f (G) es un subgrupo de G'.

Ejemplo importantísimo. Sea G un grupo y H un subgrupo normal de G.

Sea g la aplicación de G sobre el grupo cociente G/H definida por:

$$g(x) = x\,H \text{ (ó } H\,x).$$

Por definición de la multiplicación en G/H tenemos:

$$g(x) \cdot g(y) = (x\,H) \cdot (y\,H) = x\,y\,H = g(x\,y)$$

luego, g es un homomorfismo.

Este homomorfismo g lo denominaremos el *homomorfismo natural o canónico* de G sobre G/H.

Cada homomorfismo de un grupo G se puede reducir a tal homomorfismo natural, merced al teorema siguiente:

Teorema. (Teorema de Homomorfía o Primer Teorema de Isomorfía).

Si f es un homomorfismo de G sobre G', entonces la imagen inversa $f^{-1}(e')$ de la unidad de G' es un subgrupo H normal de G, y se tiene que G/H es isomorfo a G'.

Dem. Por el teorema anterior sabemos que el núcleo del homomorfismo $f : G \to G'$ es un subgrupo normal de G; luego,

$$\text{Ker}(f) = H = f^{-1}(e') \text{ es un subgrupo normal.}$$

Sea $x\,H = H\,x$ una clase de equivalencia (cogrupo) determinada por el subgrupo normal $H = \text{Ker}(f) = f^{-1}(e')$, entonces,

$$x\,H = \{x\,h : h \in H,\ x \in G\}$$

y, $f(x\,H) = f(x) \cdot f(h) = f(x) \cdot e' = f(x)$, $\forall\ h \in H$ es decir, todos los elementos de $x\,H$ van a parar a un mismo elemento $f(x) = x'$ de G', o sea, $x\,H$ es la imagen inversa de x', y las demás clases $x\,H$ son imágenes inversas de los demás elementos x' de G'.

Definamos una aplicación g de G/H a G' como sigue:

$$g(x\,H) = f(x) = x'$$

Esta definición es unívoca ya que:

$$x\,H = y\,H \Longleftrightarrow f(x) = f(y).$$

Es evidente que esta aplicación es epiyectiva. Probemos que es también inyectiva.

En efecto, supongamos que:

$$g(x\,h) = g(y\,H)$$

es decir, $f(x) = f(y) \Rightarrow x\,H = y\,H$.

Además, esta aplicación preserva los productos, ya que:

$$g(x\,H \cdot y\,H) = g(x\,y\,H) = f(x\,y) = f(x) \cdot f(y) = g(x\,H) \cdot g(y\,H).$$

Así pues, g es un insomorfismo (unívocamente determinado por f) de G/H sobre G'.

En consecuencia, los núcleos $H = \text{Ker}(f) = f^{-1}(e')$ de los homomorfismos de un grupo G son sus subgrupos normales y sólo ellos. A cada homomorfismo $f : G \to G'$ corresponde un isomorfismo

$$g : G/\text{Ker}(f) \to G'.$$

Este isomorfismo g lo denominaremos *isomorfismo natural* asociado al homoformismo f.

Recíprocamente, si H es un subgrupo normal cualquiera de G, entonces la aplicación $g : G \to G/H$ que asigna a cada elemento $x \in G$ la clase (cogrupo) que lo contiene, es un homomorfismo.

Este homomorfismo ya le dimos anteriormente el nombre de *homomorfismo natural* del grupo G sobre el grupo cociente G/H.

En resumen, es equivalente conocer:

a) Todas las equivalencias regulares en G.

b) Todos los subgrupos normales de G.

c) Todos los homomorfismos de G.

Hay correspondencia biunívoca entre los elementos a), b) y c). Así, pues, una de estas tres proposiciones implica las otras dos.

En conclusión, los resultados anteriores sugieren que todas las imágenes homomórficas de un grupo G, pueden obtenerse a partir del mismo grupo G formando los grupos cocientes G/H por sus diferentes subgrupos normales H. O sea, todo subgrupo normal H es el núcleo de un homomorfismo, y los posibles grupos cocientes G/H dan, salvo isomorfismo, los posibles grupos homomorfos de G.

Así pues, tenemos la identidad de los dos conceptos o nociones: núcleo de homomorfismo y subgrupo normal.

Nota. Sea G un grupo cíclico y f un homomorfismo de G sobre G'.

Sea H el núcleo de f. Si a es un generador de G, entonces a H será un generador del grupo cociente G/H, o sea f(a) será un generador de G'.

En consecuencia, cada grupo cociente y cada imagen homomorfa de un grupo cíclico es un grupo cíclico.

Veamos un ejemplo. Sea:

$$C_8 = \{e, a, a^2, a^3, a^4, a^5, a^6, a^7\}$$

Tomemos:

$$H = \{e, a^4\}$$

que es un subgrupo normal de C_8 y, por lo tanto, núcleo de algún homomorfismo de C_8.

C_8/H consta de cuatro cogrupos que son:

$$H = \{e, a^4\},\ a\,H = \{a, a^5\},\ a^2\,H = \{a^2, a^6\},$$

$$a^3\,H = \{a^3, a^7\}$$

Luego, C_8/H es el grupo cíclico:

$C_8/H = \{H, a\,H, (a\,H)^2, (a\,H)^3\}$ de orden cuatro.

Algunos ejemplos del Teorema de Homomorfía. 1. Sea f un homomorfismo del grupo aditivo $\mathbb{Z}$ a un grupo G.

Sabemos que todos los subgrupos aditivos de $\mathbb{Z}$ son los conjuntos $m\mathbb{Z} = \{o, \pm m, \pm 2m, \pm 3m, \ldots\}$. Luego, cada imagen homomorfa del grupo aditivo $\mathbb{Z}$ es isomorfa a:

$$\mathbb{Z}/_{m\mathbb{Z}} = \begin{cases} \mathbb{Z}_m, \text{ si } m \neq o \\ \mathbb{Z}, \text{ si } m = o \end{cases} \quad \text{(grupos aditivos)}$$

2. Consideremos la aplicación de S_n al grupo multiplicativo

$\{1, -1\}$ definida así:

$$A \to \xi(A) = (-1)^s$$

donde $\xi(A)$ representa como ya sabemos la signatura de la sustitución A, y la cual asume el valor $(+1)$, si A es sustitución par, y el valor (-1), si A es impar.

Esta aplicación es un homomorfismo, ya que:

$$\xi(AB) = \xi(A) \cdot \xi(B), \quad \forall A, B \in S_n$$

El núcléo de este homomorfismo es A_n.

Así recobramos el hecho de que el grupo alternado A_n es un subgrupo normal de S_n de índice dos.

3. Sea G un grupo arbitrario y sea G' el grupo de todos los automorfismos interiores de G.

Consideremos la aplicación $f: G \to G'$ definida por $f(a) = T_a$.

Esta aplicación es un homomorfismo de G sobre G', puesto que:

$$T_{ab} = T_a \circ T_b$$

¿Qué es el núcleo de este homomorfismo?

El núcleo es por definición el conjunto de todos los $a \in G$ tales que $f(a) = e' \in G'$, es decir la aplicación idéntica de G, ya que $I_G(x) = e \cdot x \cdot e^{-1} = x$ es un automorfismo anterior.

Luego,

$$Ta(x) = a x a^{-1} = x, \quad \forall x \in G$$

o equivalentemente:

$$a x = x a, \quad \forall x \in G$$

Por lo tanto, el núcleo es el conjunto de los elementos de G que conmutan con todos los elementos de G; o lo que es lo mismo, el núcleo es el conjunto de todos los elementos autoconjugados de G, incluyendo a *e*.

Luego, este conjunto forma un subgrupo abeliano de G y que es, además, normal. A tal subgrupo lo denominaremos CENTRO de G y lo designaremos por Z_G.

Así pues, el grupo de los automorfismos interiores de G es isomorfo al grupo cociente G/Z_G.

En particular, si G es abeliano entonces $G = Z_G$, y el grupo de los automorfismos interiores se reduce a $\{e\}$.

Ahora cabe la pregunta: ¿En qué casos el grupo de los automorfismos interiores es isomorfo a G?

Para contestar esta pregunta, basta recordar que un homomorfismo es un isomorfismo si, y sólo si, el núcleo $\mathrm{Ker}(f) = \{e\}$. Luego, G es isoformo al conjunto de todos los automorfismos interiores de G si, y sólo si, el centro Z_G se reduce al elemento neutro de G, es decir, $Z_G = \{e\}$.

4. Sea W_n el grupo de las raíces n-ésimas de la unidad 1.

Esto es:

$$w_o = \cos o + i \operatorname{sen} o = 1$$

$$w_1 = \cos \frac{2\pi}{n} + i \operatorname{sen} \frac{2\pi}{n}$$

$$w_2 = \cos 2 \cdot \frac{2\pi}{n} + i \operatorname{sen} 2 \cdot \frac{2\pi}{n}$$

. . . . . .

$$w_j = \cos j \frac{2\pi}{n} + i \operatorname{sen} j \cdot \frac{2\pi}{n}$$

. . . . . .

$$w_{n-1} = \cos (n-1) \cdot \frac{2\pi}{n} + i \operatorname{sen} (n-1) \frac{2\pi}{n}$$

La aplicación $f: \mathbb{Z} \to W_n$ dada por:

$$f(m) = \cos m \cdot \frac{2\pi}{n} + i \operatorname{sen} m \frac{2\pi}{n}$$

es en virtud de la fórmula De Moivre:

$$(\cos\alpha + i \operatorname{sen}\alpha)(\cos\beta + i \operatorname{sen}\beta) = \cos(\alpha + \beta) + i \operatorname{sen}(\alpha + \beta)$$

un homomorfismo del grupo aditivo $\mathbb{Z}$ en el grupo multiplicativo W_n y es, evidentemente, un homomorfismo sobre.

Determinemos su núcleo Ker (f). Sea $k \in \text{Ker}(f)$, es decir

$$f(k) = \cos k\,\frac{2\pi}{n} + i \operatorname{sen} k\,\frac{2\pi}{n} = w_o = 1$$

lo que implica que:

$$\cos k \cdot \frac{2\pi}{n} = 1 \quad \text{y} \quad \operatorname{sen} k \cdot \frac{2\pi}{n} = o$$

lo cual se verifica si,

$$k \cdot \frac{2\pi}{n} = h \cdot 2\pi$$

y lo que exige que k sea divisible por n. Luego, el núcleo del homomorfismo f está formado por todos los múltiplos de n, es decir, es el conjunto $n\,\mathbb{Z}$.

$$\text{Ker}(f) = n\,\mathbb{Z}$$

Por lo tanto,

$$\mathbb{Z} / {}_n\,\mathbb{Z} \cong W_n$$

En base a este isomorfismo, resulta también que:

$$W_n \cong \mathbb{Z}_n = \{0, \overline{1}, \overline{2}, \ldots, \overline{n-1}\}$$

y también, W_n isomorfo al grupo de las rotaciones de centro 0 y de ángulo $\frac{2\pi}{n}$, $n \in \mathbb{N}$.

5. Sea $\mathbb{R}^* = \mathbb{R} - \{o\}$ el grupo multiplicativo de los números reales no nulos y sea $\mathbb{R}^+$ el grupo multiplicativo de los números reales positivos.

Sea la aplicación $f : \mathbb{R}^* \to \mathbb{R}^+$ definida por:

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

Esta aplicación es un homomorfismo, ya que:

$$f(x\,y) = (x\,y)^2 = x^2\,y^2 = f(x) \cdot f(y)$$

Afirmamos además que esta aplicación es "sobre": Más adelante pro-

baremos que todo número real positivo posee una raíz cuadrada en $\mathbb{R}$, o sea dado $r \in \mathbb{R}$, $r > o$, existe $y \in \mathbb{R}$ tal que,

$$y^2 = r$$

Determinemos el núcleo de este homomorfismo. Sea $x \in \text{Ker}(f)$, entonces,

$$x^2 = 1 \Longrightarrow x = 1, \quad o, \quad x = -1$$

Luego, $\text{Ker}(f) = \{1, -1\}$.

Por lo tanto, se tiene el isomorfismo:

$$\mathbb{R}^* / \{1, -1\} \cong \mathbb{R}^+$$

En este isomorfismo, $\mathbb{R}^* / \{1, -\}$ consiste en la totalidad de los pares $\{x, -x\}$ de números reales $x \neq o$, y el isomorfismo entre $\mathbb{R}^* / \{1, -1\}$ y $\mathbb{R}^+$ está dado por:

$$\{x, -x\} \to x^2$$

7.13. Finalizaremos el presente capítulo sobre grupos, viendo un modelo concreto de la estructura abstracta de grupo.

Teorema. Todo grupo abstracto G es isomorfo a un grupo de aplicaciones biunívocas de un conjunto sobre si mismo.

Dem. Sea G un grupo abstracto y pongamos X = G, es decir consideramos al grupo G apenas como conjunto; luego, $X \neq \varnothing$.

Consideremos el conjunto $X^X = \text{Aplic}(X)$ de todas las aplicaciones de X en sí mismo.

Por el teorema de representación de semigrupos, ya visto hace mucho, podemos afirmar que G como semigrupo con unidad es isomorfo a la familia $(f_a)\ a \in G$ de todas las aplicaciones de X en X definidas por:

$$f_a(x) = a\,x, \quad \forall\, x \in X \quad \text{y} \quad a \text{ fijo en } G,$$

siendo la aplicación de isomorfismo:

$$g : G \to (f_a)\ a \in G$$

definida por, $g(a) = f_a$, siendo $f_a(x) = a\,x$

Observemos que la familia $(f_a)_{a\in G}$ es la misma que consideramos para el caso de semigrupos con unidad, que ahora son biunívocas y sobre por el hecho que G ya no es semigrupo, sino grupo.

Ahora bien, se demostró que $(f_a)_{a\in G}$ era un semigrupo, luego para probar el teorema sólo restará demostrar que todo elemento de $(f_a)_{a\in G}$ tiene inverso en $(f_a)_{a\in G}$

Sea $f_a \in (f_a)_{a\in G}$; por ser f_a biunívoca y sobre (biyectiva), existe f_a^{-1} definida de la manera siguiente:

$$f_a^{-1}(y) = x \text{ si, y sólo si, } f_a(x) = y$$

Demostremos que $f_a^{-1} \in (f_a)_{a\in G}$. Tenemos:

$$f_a(x) = y = a\,x \Longrightarrow x = a^{-1}\,y$$

Pero, $a^{-1} \in G$, por ser G grupo; luego,

$$f_a^{-1}(y) = x = a^{-1}\,y, \text{ con } a^{-1}, y \in G$$

es decir, $f_a^{-1} = f_{a^{-1}} \in (f_a)_{a\in G}$ y por consiguiente, $(f_a)_{a\in G}$ es un grupo, puesto que,

$$(f_a \cdot f_{a^{-1}})(x) = f_a(f_{a^{-1}}(x)) = f_a(a^{-1}x) = a(a^{-1}x) = x = f_e = I_G$$

En consecuencia, hemos hallado un modelo concreto de cualquier grupo abstracto, y este modelo es un grupo de funciones y, por el isomorfismo, idenificable con el primero.
Luego,

$$(G;\cdot) = (f_a)_{a\in G}$$

donde $a = f_a$, en virtud del monomorfismo G.

Observación importante. Cuando un grupo abstracto es dado por su tabla de composición, puede descubrirse fácilmente su elemento unidad, el inverso de cada elemento, la propiedad de cierre o clausura y si el grupo es conmutativo; esto último, si el cuadro de la tabla de composición es simétrico con respecto a la diagonal principal. En cambio, sería muy laboriosa la comprobación directa de la ley asociativa, ya que tendríamos que examinar todas las igualdades de la forma,

$$(x\,y)\,z = x\,(y\,z)$$

Esta dificultad puede superarse con métodos indirectos, aprovechando, por ejemplo, el teorema de representación que acabamos de estudiar. En efecto, todas las relaciones que se deducen de la tabla de multiplicación tienen su contrapartida en relaciones entre funciones. Esto proporciona en realidad una comprobación indirecta del cumplimiento de la ley asociativa en el grupo abstracto que estamos considerando, ya que sabemos que esta ley se verifica para todas las composiciones de aplicaciones o funciones.

Otro modelo concreto de un grupo abstracto nos lo proporciona el siguiente teorema,

Teorema. Todo grupo abstracto es isomorfo a un grupo de sustituciones.

Dem. Sea $G = \{a_1, a_2, \ldots, a_n\}$ un grupo abstracto finito de orden n. Si a es uno cualquiera de sus elementos, entonces los productos:

$$a_1\,a, a_2\,a, \ldots, a_1\,a, \ldots, a_n\,a$$

son n elementos distintos de G, en general en otro orden, y por consiguiente representan una permutación de los n elementos $a_1, a_2, \ldots, a_n$ del grupo considerado G. Por lo tanto, a cada elemento $a \in G$ podemos asociar una sustitución, a saber:

$$A = \begin{pmatrix} a_1\ a_2 \ldots a_i \ldots a_n \\ a_1a\ a_2a \ldots a_1a \ldots a_na \end{pmatrix}$$

o más simplemente,

$$A = \begin{pmatrix} a_i \\ a_ia \end{pmatrix}, \quad i = 1, 2, \ldots, n$$

Los objetos con que opera esta sustitución son los mismos elementos del grupo G.

Es claro que, si a_i representa a todos y cada uno de los elementos del grupo G, lo mismo le ocurre al producto a_ix, en donde x es un elemento fijo en G. Luego, podrá escribirse:

$$A = \begin{pmatrix} a_i \\ a_ia \end{pmatrix} = \begin{pmatrix} a_i\,x \\ a_i\,x\,a \end{pmatrix}, i = 1, 2, \ldots, n$$

Sea ahora,

$$B = \begin{pmatrix} a_i \\ a_i b \end{pmatrix} \quad , \quad i = 1, 2, \ldots, n$$

la sustitución correspondiente al elemento $b \in G$.

Si a y b son dos elementos distintos de G, las sustituciones A y B son distintas evidentemente. Por consiguiente, hemos establecido una correspondencia biunívoca entre los elementos a, b,... de G y el conjunto de las sustituciones A, B,...

Determinemos en seguida el producto AB de las sustituciones A y B; obtendremos:

$$AB = \begin{pmatrix} a_i \\ a_i a \end{pmatrix} \cdot \begin{pmatrix} a_i \\ a_i b \end{pmatrix} = \begin{pmatrix} a_i \\ a_i a \end{pmatrix} \begin{pmatrix} a_i \\ a_i a\, b \end{pmatrix} = \begin{pmatrix} a_i \\ a_i a\, b \end{pmatrix}$$

El resultado es, pues, la sustitución que corresponde al elemento ab del grupo G.

En resumen:

$$a \longleftrightarrow A \text{ y } b \longleftrightarrow B \Rightarrow ab \longleftrightarrow AB$$

Así hemos demostrado que la correspondencia,

$$a \longleftrightarrow A$$

es efectivamente un isomorfismo.

Por consiguiente, el conjunto de sustituciones

$$G' = \{A, B, \ldots\}$$

forman un grupo (subgrupo de S_n) que posee la misma estructura que el grupo abstracto G considerado.

El elemento unidad de este grupo es, evidentemente, la sustitución idéntica I:

$$I = \begin{pmatrix} a_i \\ a_i e \end{pmatrix}, \; i = 1, 2, 3, \ldots n \text{ siendo } e \text{ la unidad de G.}$$

Capítulo VIII

TEORIA ELEMENTAL DE ANILLOS Y CUERPOS

8.0. Los grupos tienen una sola ley de composición. En cambio, los números enteros, los números racionales, los números reales y los números complejos, que forman el sostén de toda la matemática clásica, admiten dos leyes de composición fundamentales: la adición y la multiplicación.

De aquí que la idea de grupo, con todo su interés, no sea suficiente para caracterizar muchos de los conjuntos elementales con que opera la matemática. Hace falta, pues, suponer otra ley de composición con ciertas propiedades dadas como axiomas. Esto efectivamente es posible y da lugar, en particular, a la importante estructura algeibraica de *anillo*.

En general, un anillo es una terna ordenada (A; +, ·) cuyo primer elemento es un conjunto no vacío A y cuyos dos últimos elementos son operaciones en A.

De estas dos operaciones, a la primera la llamaremos habitualmente "adición" del anillo, pero esto no quiere decir en modo alguno, como lo dijimos en grupo, que tal adición coincida necesariamente con la adición ordinaria de números y sólo se trata de un nombre que se da por analogía y nada más.

De la misma manera, a la segunda de las mencionadas operaciones la llamaremos "multiplicación" del anillo, pero esto no quiere decir en modo alguno que ella deba coincidir necesariamente con la multiplicación ordinaria de números; se trata también aquí de un nombre que se da por analogía.

Las propiedades que debe cumplir esta terna (A; +, ·) están consignadas en la definición siguiente:

Definición. Llamaremos ANILLO a un conjunto no vacío A sobre el que están definidas dos operaciones binarias internas, una llamada "adición" (+) y otra llamada "multiplicación" (·), tal que A es un grupo abeliano con respecto a la adición y la multiplicación es asociativa y distributiva a izquierda y a derecha respecto de la adición. Es decir, se verifican las siguientes propiedades dadas como axiomas:

A1 : $\forall\, a, b \in A \Rightarrow a + b \in A$

A2 : $\forall\, a, b, c \in A$ se verifica $a + (b + c) = (a + b) + c$

A3 : $\forall\, a, b \in A$ se verifica $a + b = b + a$

A4 : $\exists\, o \in A$ tal que $a + 0 = a$, $\forall\, a \in A$

A5 : $\forall\, a \in A$ existe $(-a) \in A$ tal que $a + (-a) = 0$

M1 : $\forall\, a, b \in A \Rightarrow a \cdot b \in A$

M2 : $\forall\, a, b, c \in A$ se verifica $a(b\,c) = (ab)c$

M3 : $\forall\, a, b, c \in A$ se verifica $a(b + c) = ab + ac$ y $(b + c)a = b\,a + c\,a$

Observaciones. 1) Nótese que en un anillo cualquiera $(A; +, \cdot)$ la primera operación $(+)$ es siempre conmutativa, ya que por definición se exige que $(A; +)$ sea un grupo conmutativo o abeliano. Pero con respecto a la segunda operación $(\cdot)$, no se postula la conmutatividad de ella. Luego, puede ser que sea y puede ser que no sea conmutativa, según el anillo particular que se considere.

Si esta segunda operación también es conmutativa, diremos que el anillo es conmutativo o abeliano.

Luego, un anillo $(A; +, \cdot)$ es conmutativo, si la multiplicación del anillo es una operación conmutativa; es decir:

M4 : $\forall\, a, b \in A$ se verifica $a\,b = b\,a$

2) La definición de anillo no dice nada acerca de la existencia del elemento neutro o unidad de la segunda operación. En algunos casos puede existir tal elemento neutro, según el anillo particular que se considere. Cuando así ocurre, a este elemento neutro de la segunda operación lo llamaremos *unidad* del anillo, y diremos que el anillo en cuestión es un anillo con unidad o unitario; es decir:

M5 : $\exists\, 1 \in A$ tal que: $a \cdot 1 = 1 \cdot a = a$, $\forall\, a \in A$

Es claro que, si un anillo $(A; +, \cdot)$ tiene unidad, ésta es única; porque según lo visto en el párrafo sobre operaciones binarias, para cualquier operación, si hay elemento neutro, éste es único.

3) Finalmente, cabe formular consideraciones análogas a las que hicimos en el párrafo sobre grupos, a saber: En el anillo representado por la terna ordenada $(A; +, \cdot)$, suele llámarse anillo simplemente al conjunto subyacente A, quedando sobrentendidas las dos operaciones $(+)$ y $(\cdot)$. Quede bien claro que estas convenciones constituyen abusos de lenguaje, pero que son cómodas en los enunciados y escrituras.

Ejemplos.

1) Entre los conjuntos numéricos, forman anillos conmutativos con unidad, con respecto a la adición y a la multiplicación ordinarias, los siguientes: $\mathbb{Z}$, Q, $\mathbb{R}$ y C.

El conjunto $P = \{2n : n \in \mathbb{Z}\}$ forma un anillo conmutativo sin unidad.

2. El conjunto $\mathbb{Z}_n = \{\overline{0}, \overline{1}, \overline{2}, \ldots, \overline{n-1}\}$ algebrizado mediante las operaciones:

$$\overline{a} + \overline{b} = \overline{a + b}$$

$$\overline{a}\ \ \overline{b} = \overline{ab}$$

es un anillo conmutativo con unidad. ¡Pruébelo!

3. Consideremos el conjunto de las funciones numéricas:

$f : A \to \mathbb{R}$

siendo A un conjunto cualquiera no vacío y $\mathbb{R}$ el conjunto de los números reales.

Sabemos que la suma y el producto de estas funciones se definen por las reglas:

$$\begin{cases} (f + g)(x) = f(x) + g(x), & \forall\ x \in A \\ (f\,g)\ \ (x) = f(x) \cdot g(x), & \forall\ x \in A \end{cases}$$

Por las propiedades que vimos para la suma de funciones numéricas, muestran que el conjunto de tales funciones constituyen un grupo conmutativo con respecto a dicha operación.

Como además, la multiplicación de funciones numéricas es asociativa y distributiva con respecto a la suma, tenemos entonces que el conjunto de todas las funciones numéricas $f : A \to \mathbb{R}$ es un anillo con unidad.

El cero del anillo es la función $\overline{0} : A \to \mathbb{R}$ dada por $\overline{0}(x) = 0$ y la unidad es la función $\overline{1} : A \to \mathbb{R}$ dada por $\overline{1}(x) = 1$.

El opuesto o el simétrico del elemento $f : A \to \mathbb{R}$ es la función $(-f) : A \to \mathbb{R}$ dada por $(-f)(x) = -f(x)$, $\forall\, x \in A$.

En particular, si $A = \mathbb{R}$, entonces tenemos el anillo de todas las funciones reales de argumentos reales. En particular, el conjunto de los polinomios enteros de una variable constituye un anillo conmutativo con unidad para las dos leyes clásicas de adición y multiplicación.

Consideremos, en fin, el conjunto de las funciones continuas sobre un intervalo finito $[a, b]$; provisto de las leyes de composición, adición y multiplicación ordinarias, este conjunto tiene una estructura de anillo conmutativo unitario.

4. Sea A un anillo cualquiera sobre el cual están definidas dos leyes de composición internas denotadas aditivamente y multiplicativamente. Se supone que todo $x \in A$ es tal que:

$$x^2 = x$$

es decir, todo elemento de A es indempotente. Entonces, probemos que todo $x \in A$ es igual a su opuesto, y deducir de esto que A es un anillo conmutativo.

En efecto, tendremos por la idempotencia de los elementos del anillo A, por la propiedad distributiva de la multiplicación respecto a la adición y por la ley de cancelación del grupo aditivo del anillo, lo que sigue:

$$(a + a)^2 = a + a, \quad a \in A \text{ arbitrario}$$

o sea,

$$(a + a)(a + a) = a + a$$
$$a^2 + a^2 + a^2 + a^2 = a + a$$

o bien,

$$\not{a} + \not{a} + a + a = \not{a} + \not{a}$$
$$a + a = o$$

lo que prueba que cada elemento coincide con su opuesto; esto es,

$$a = -a$$

Sentado esto, tenemos ahora:

$$(a + b)^2 = a + b, \quad \forall \ a, b \in A$$

o sea,

$$(a + b)(a + b) = a + b$$
$$a^2 + ab + ba + b^2 = a + b$$
$$a + ab + ba + \not{b} = \not{a} + \not{b}$$
$$ab + ba = o$$
$$ab = -ba$$

y como, $-ba = ba$ resulta:

$$ab = ba$$

lo que prueba que nuestro anillo en cuestión, es conmutativo.

5) Sea E un conjunto no vacío sobre el que están definidas las operaciones binarias $(+)$ y $(\cdot)$, que por abuso de lenguaje llamaremos a $(+)$ *suma* y a $(\cdot)$ *producto*, verificando todos los axiomas de una estructura de anillo, menos la conmutatividad de la suma.

Probaremos que si el sistema algebraico $(E; +, \cdot)$ tiene unidad con respecto al producto, entonces, nuevamente por abuso de lenguaje, E es un anillo unitario (es decir, con unidad).

En efecto, sea 1 la unidad de E respecto al producto. Para que E sea anillo, basta solamente demostrar la conmutatividad de la suma, ya que los otros axiomas de la estructura de anillo se cumplen por hipótesis.

Tenemos, pues, por la propiedad distributiva generalizada:

(1) $(a + b)(1 + 1) = (a + b) \cdot 1 + (a + b) \cdot 1 = (a + b) + (a + b) = a + (b + a) + b$

Por otro lado, por la misma ley distributiva, se tiene:

(2) $(a + b)(1 + 1) = a(1 + 1) + b(1 + 1) = (a + a) + (b + b) = a + (a + b) + b$

luego, se obtiene de (1) y (2):

$$a + (b + a) + b = a + (a + b) + b$$

Sumando a la izquierda $(-a)$ y sumando a la derecha $(-b)$, resulta:

$$(-a) + a + (b + a) + b + (-b) = (-a) + a + (a + b) + b + (-b)$$
$$o + (b + a) + o = o + (a + b) + o$$
$$b + a = a + b$$

Por consiguiente, E es un anillo con unidad.

6. Probar que el conjunto de los números reales de la forma:

$m + n\sqrt{2}$, con $m, n \in \mathbb{Z}$

es un anillo (adición y producto habituales).

En efecto, probaremos que se verifican todos los axiomas de la estructura de anillo. Tenemos:

Adición:

a) El cierre se cumple, ya que:

$$(m + n\sqrt{2}) + (m' + n'\sqrt{2}) = (m + m') + (n + n')\sqrt{2}, \ \forall\ m, m', n, n' \in \mathbb{Z}$$

b) $[(m + n\sqrt{2}) + (m' + n'\sqrt{2})] + (m'' + n''\sqrt{2}) =$
$[(m + m') + (n + n')\sqrt{2}] + (m'' + n'')\sqrt{2} =$
$= [(m + m') + m''] + [(n + n') + n'']\sqrt{2}$
$= [m + (m' + m'')] + [n + (n' + n'')]\sqrt{2}$
$= (m + n\sqrt{2}) + [(m' + m'') + (n' + n'')\sqrt{2}]$
$= (m + n\sqrt{2}) + [(m' + n'\sqrt{2}) + (m'' + n'')\sqrt{2}]$ (asociativa)

c) La conmutatividad es evidente; luego:

$$(m + n\sqrt{2}) + (m' + n'\sqrt{2}) = (m' + n'\sqrt{2}) + (m + n\sqrt{2})$$

d) Existe $x + y\sqrt{2}$ tal que,

$$(m + n\sqrt{2}) + (x + y\sqrt{2}) = m + n\sqrt{2}$$

cualquiera sean m y n en $\mathbb{Z}$.

En efecto, tenemos:

$$x + y\sqrt{2} = o \implies x = o, \quad y = o$$

Así pues, el elemento neutro de la adición es:

$$o + o \cdot \sqrt{2}$$

e) Dado $m + n\sqrt{2}$, existe $x + y\sqrt{2}$ tal que:

$$(m + n\sqrt{2}) + (x + y\sqrt{2}) = o + o \cdot \sqrt{2}, \quad m, n \in \mathbb{Z}$$

Probaremos que, con estas dos leyes de composición, el conjunto A es un anillo conmutativo con unidad.

Demostremos primeramente que la diferencia simétrica hace de $A = P(X)$ un grupo conmutativo. Tenemos:

a) el cierre es evidente por la definición misma de Δ; luego,

$$x, y \in A \qquad x + y = x \Delta y = (x \cup y) - (x \cap y) \in A$$

b) $$(x + y) + z = (x \Delta y) \Delta z = [(x \cup y) - (x \cap y)] \Delta z$$
$$= (x \cup y) \cup z - (x \cap y) \cap z = x \cup (y \cup z) - x \cap (y \cap z)$$
$$= x \Delta [(y \cup z) - (y \cap z)] = x \Delta (y \Delta z)$$
$$= x + (y + z), \quad \forall\, x, y, z \in A$$

c) $$x + y = x \Delta y = (x \cup y) - (x \cap y) = (y \cup x) - (y \cap x)$$
$$= y \Delta x = y + x, \quad \forall\, x, y \in A$$

d) $$x + o = x \Delta \varnothing = (x \cup \varnothing) - (x \cap \varnothing) = x - \varnothing = x, \qquad x \in A$$

Luego, $\varnothing$ es el elemento neutro de la operación Δ.

e) $$x + (-x) = x \Delta x = (x \cup x) - (x \cap x) = x - x = \varnothing$$

Luego, el inverso de cada elemento es el mismo elemento; esto es:

$$x + x = o, \text{ o sea, } x \Delta x = \varnothing$$

Este resultado y los anteriores prueban que la operación Δ = diferencia simétrica hace del conjunto de partes de un conjunto cualquiera $X \neq \varnothing$, un grupo conmutativo o abeliano.

Por otra parte, el producto $x \cdot y = x \cap y$ es asociativo y también cerrado en $A = P(X)$. Para concluir que $A = P(X)$ es un anillo, solamente bastará demostrar la distributividad de la operación intersección con respecto a la operación diferencia simétrica, esto es:

$$x \cdot (y + z) = x \cap (y \Delta z) = (x \cap y) \Delta (x \cap z) = x \cdot y + x \cdot z$$

En efecto, tenemos:

$$xy + xz = (x \cap y) \Delta (x \cap z) = (x \cap y) \cup (x \cap z) - (x \cap y) \cap (x \cap z)$$
$$= [x \cap (y \cup z)] - [x \cap (y \cap z)] = x \cap [(y \cup z) - (y \cap z)]$$
$$= x \cap (y \Delta z) = x(y + z)$$

y como también $x \cap (y \Delta z) = (y \Delta z) \cap x$, concluímos que las operaciones Δ y $\cap$ hacen de $A = P(X)$ un anillo conmutativo. Además, $U(A) = \{X\}$ es el único elemento inversible de $A = P(X)$, ya que $X \cap X = X$; luego, la unidad del anillo $A = P(X)$ es X.

El anillo $(P(X); \Delta, \cap)$ se llama el *anillo de Boole de subconjuntos de X*.

Observación. En el anillo de Boole $(P(X); \Delta, \cap)$, cada elemento de $P(X)$ es idempotente con respecto a la segunda operación $\cap$; luego, en el ejemplo 4) de más arriba, el anillo A en el cual para cada $x \in A$ se tiene $x^2 = x$, es un anillo de Boole.

En consecuencia, llamaremos ANILLO DE BOOLE a un anillo en el cual

todos sus elementos son idempotentes con respecto a la segunda operación del anillo.

Lo anterior implica que cada elemento del anillo es igual a su inverso aditivo y, por consiguiente, el anillo es conmutativo.

8.1. *Consecuencias de la definición de anillo.*

En un anillo se verifican todas las propiedades de los grupos abelianos, en este caso de los grupos aditivos, más algunas otras que luego veremos.

Entre las propiedades de la suma de un anillo tenemos:

a) el cero de un anillo es único.

b) para cada elemento a de un anillo, el opuesto $(-a)$ es único,.

c) el opuesto de $(-a)$ es a, o sea

$$-(-a) = a,$$

d) el elemento opuesto de $a + b$ es,

$$-(a + b) = (-b) + (-a) = -b - a$$

o sea, por la conmutatividad de la adición:

$$-(a + b) = (-a) + (-b) = -a - b,$$

e) vale la ley de cancelación de la adición, o sea

$$a + b = a + c \Rightarrow b = c$$

es decir, se puede simplificar un sumando o término común a ambos miembros de una igualdad.

f) la ecuación $x + b = a$, cualesquiera sean a y b tiene siempre solución única en el anillo, y su solución es:

$$x = a + (-b) = a - b$$

Este hecho permite definir una nueva operación, la *Sustracción*, en el anillo.

Se tiene también, $-(a - b) = (-a) - (-b) = -a + b$.

Ahora pasaremos a ver otras propiedades que se verifican en un anillo y que son fundamentales.

Proposición 1. En todo anillo A la multiplicación es distributiva con respecto a la sustracción; esto es:

$$a(b - c) = ab - ac$$
$$(b - c)a = ba - ca$$

para elementos a, b, c cualesquiera de A.

Dem. Tenemos:

$$a(b - c) + ac = a[(b - c) + c] = a[b + (-c) + c] = a[b + (-c + c)] = a[b + o] = ab$$

y por tanto,

$$a(b - c) = ab - ac$$

Análogamente se demuestra la igualdad $(b - c)a = ba - ca$.

Proposición 2. En todo anillo A, el producto por cero es cero.

En términos más precisos, esto es:

$$a \cdot o = o \cdot a = o, \quad \forall a \in A$$

Dem. Si en las igualdades:

$$a(b - c) = ab - ac \text{ y } (b - c)a = ba - ca$$

consideramos el caso $b = c$, resulta:

$$a(b - b) = ab - ab \text{ y } (b - b)a = ba - ba$$

o sea, $a \cdot o = o$ y $o \cdot a = o$

Proposición 3. En todo anillo A se verifica la siguiente regla de signos:

$$(-a)b = -ab; \quad a(-b) = -ab; \quad (-a)(-b) = ab$$

Dem. Si en las igualdades,

$$a(c - b) = ac - ab \text{ y } (c - a)b = cb - ab$$

hacemos $c = o$, resultan:

$$a(-b) = -ab \text{ y } (-a)b = -ab$$

También se tiene:

$$(-a)(-b) - ab = (-a) \cdot (-b) + (-a)b = (-a)(-b + b) = o$$

y por tanto,

$$(-a)(-b) = ab.$$

Es decir, vale la regla de los signos de la multiplicación ordinaria.

Definiciones. En todo anillo se definen las potencias de exponente entero positivo del siguiente modo:

$$\begin{cases} a^1 = a \\ a^{n+1} = a^n \cdot a \end{cases} ; n \in \mathbb{Z}^+$$

y los múltiplos enteros de un elemento de la siguiente manera:

$$\begin{cases} 1a = a \\ (n + 1)a = na + a \end{cases} , n \in \mathbb{Z}$$

Proposición 4. En todo anillo A, se verifica:

a) $a^m \cdot a^n = a^{m+n}$; $(a^m)^n = a^{mn}$, $\forall m, n \in \mathbb{Z}^+$

b) $(ab)^n = a^n b^n$, si $ab = ba$, $\forall n \in \mathbb{Z}^+$

c) $n(a + b) = na + nb$, $\forall n \in \mathbb{Z}$

d) $(m + n)a = ma + na$, $\forall m, n \in \mathbb{Z}$

Dem. Todas estas igualdades se demuestran por inducción tal como lo hicimos al estudiar las potencias y múltiplos de un elemento en la teoría general de operaciones. ¡Hágala nuevamente como ejercicio!

En particular, en todo anillo se tiene la fórmula:

$$(a + b)^2 = a^2 + ab + ba + b^2$$

Si el anillo es conmutativo, entonces la fórmula:

$$(a + b)^2 = a^2 + 2ab + b^2$$

En general, en todo anillo conmutativo, vale la fórmula del binomio de Newton.

*****

Ya sabemos que la definición de múltiplo entero n a de un elemento a de un anillo significa:

$$na = a + a + \ldots + a \text{ (n sumandos)}$$

Si $n = o$, $o \cdot a = o$, mientras que si n es negativo:

$$na = (-a) + (-a) + \ldots + (-a) \text{ (n sumandos)}$$

Veamos ahora otras propiedades, consecuencia de la ley distributiva.

Proposición 5. En todo anillo A se verifica:

a) $(na)b = a(nb) = n(ab)$

b) $(ma)(nb) = mn(ab)$

Dem. De la ley general distributiva, resulta:

$$(na)b = (a + a + \ldots + a)b = ab + ab + \ldots + ab = a(b + b + \ldots + b) = n(ab).$$

Si n es positivo; pero si n es negativo:

$$(na)b = [(-a) + (-a) + \ldots + (-a)]b = (-ab) + (-ab) + \ldots + (-ab) = a[(-b) + (-b) + \ldots + (-b)] = n(ab)$$

lo que prueba a).

La regla $(a + a + m \ldots + a)(b + b + n \ldots + b) = ab + ab + \ldots + ab$, que es otra ley distributiva general, da:

$$(ma)(nb) = mn(ab)$$

que también vale para los enteros m y n positivos, negativos o nulos.

8.2. *Divisores de Cero*

Los distintos anillos presentan desigual comportamiento en aspectos que ahora pasamos a estudiar.

No siempre es cierta la propiedad recíproca de la Proposición 2); esto es: vimos que en todo anillo vale la siguiente propiedad,

$$a \cdot o = o \cdot a = o, \quad \forall a$$

En cambio, la recíproca no es en general cierta; es decir, existen anillos con elementos a, b no nulos, tales que su producto es el elemento nulo.

$$a \cdot b = o, \text{ y } a \neq o \text{ y } b \neq o$$

En tal caso diremos que a y b son DIVISORES de CERO y que el anillo posee divisores de cero.

También decimos que a es un *divisor de cero a izquierda* y b es un *divisor de cero a derecha.*

Definición. Un elemento a de un anillo A se llama DIVISOR de CERO, si siendo el mismo distinto de cero, existe $b \in A$, $b \neq o$, tal que:

$$(*)\ a \cdot b = o$$

Con más precisión: al elemento a lo llamaremos divisor de cero a izquierda, porque aparece a la izquierda en la fórmula (*), y al elemento b lo llamaremos divisor de cero a derecha, porque aparece a la derecha en la fórmula (*).

Existen anillos sin divisores de cero, es decir en los que se verifica la siguiente propiedad:

"Si $a \cdot b = o$, entonces es $a = o$, ó $b = o$".

o sea, $a \cdot b = o \Rightarrow a = o$, ó $b = o$.

Ejemplos. En el anillo $\mathbb{Z}_4 = \{\bar{o}, \bar{1}, \bar{2}, \bar{3}\}$, $\bar{2}$ es un divisor de cero tanto a izquierda como a derecha, ya que:

$$\bar{2} \cdot \bar{2} = \bar{o}$$

en donde $\bar{o}$ es el cero del anillo $\mathbb{Z}_4$.

En el anillo $\mathbb{Z}_6 = \{\bar{o}, \bar{1}, \bar{2}, \bar{3}, \bar{4}, \bar{5}\}$, los elementos $\bar{2}$ y $\bar{3}$, $\bar{3}$ y $\bar{4}$ son divisores de cero, pues:

$$\bar{2} \cdot \bar{3} = \bar{o} \text{ y } \bar{3} \cdot \bar{4} = \bar{o}$$

En cambio, el anillo $\mathbb{Z}$ de los enteros no tiene divisores de cero; tampoco lo tienen los anillos $\mathbb{Z}_3 = \{\bar{o}, \bar{1}, \bar{2}\}$ y $\mathbb{Z}_5 = \{\bar{o}, \bar{1}, \bar{2}, \bar{3}, \bar{4}\}$.

En general, el anillo $\mathbb{Z}_n = \{\bar{o}, \bar{1}, \bar{2}, \ldots, \overline{n-1}\}$ tiene divisores de cero si el módulo n no es un número primo, y es sin divisores de cero cuando n es primo.

En efecto, supongamos que sea $n = a \cdot b$, con a y b distintos de $(+1)$ y de (-1). Sabemos que:

$$n = a \cdot b \equiv o \text{ (mód. n)}, \quad a \neq o \text{ y } b \neq o$$

luego, $C_n = C_{a \cdot b} = C_a \cdot C_b = C_o$ siendo $C_a \neq C_o$ y $C_b \neq C_o$ en donde C_o es el elemento neutro del grupo aditivo de $\mathbb{Z}_n$, o sea el cero del anillo.

Luego, $\mathbb{Z}_n$ tiene divisores de cero si el módulo n no es primo.

Por otra parte, se prueba que, en cambio, si n es primo, el anillo $\mathbb{Z}_n$ de clases residuales módulo n no tiene divisores de cero.

En efecto, demostraremos que un producto en $\mathbb{Z}_n$, con n primo, no puede ser nulo más que cuando al menos uno de los factores es nulo.

Sea el producto,

$C_a \cdot C_b = Co$

es decir, $(a + \dot{n}) \cdot (b + \dot{n}) = \dot{n}$

$a \cdot b + \dot{n} = \dot{n}$

$a \cdot b = \dot{n} \Rightarrow n \mid ab \Rightarrow n \mid a$, ó, $n \mid b$.

Ninguna de estas dos alternativas es posible, a menos que sea $a = o$, ó, $b = o$, porque a y b son menores que n y primos con n.

Así pues, $C_a \cdot C_b = C_o$ solamente cuando $a = o$, ó $b = o$; es decir, si $C_a = C_o$, ó $C_b = C_o$. Por tanto,

$C_a \cdot C_b = C_o \Rightarrow C_a = o$, ó $C_b = o$

Luego el anillo $\mathbb{Z}_n$, con n primo, no tiene divisores de cero.

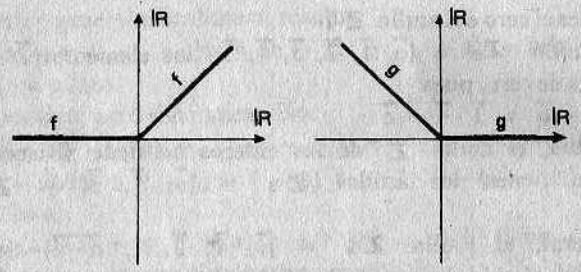
Consideremos ahora el anillo de todas las funciones reales de argumentos reales. En este anillo existen ejemplos de divisores de cero. En efecto, sean las funciones:

$f : \mathbb{R} \to \mathbb{R}$ y $g : \mathbb{R} \to \mathbb{R}$

definidas por:

$$f(x) = \max(o, x) = \begin{cases} o \text{ si } x \leq o \\ x \text{ si } x > o \end{cases} ; \quad g(x) = \min(o, x) = \begin{cases} x \text{ si } x < o \\ o \text{ si } x \geq o \end{cases}$$

Los gráficos de estas funciones son:



Por la composición punto a punto para el producto, tenemos para cada $x \in \mathbb{R}$:

$f(x) \cdot g(x) = o$

o sea, $f \cdot g : \mathbb{R} \to \mathbb{R}$

es la función constante nula, y la hemos ya denotado por $o(x) = o$.

Podemos escribir, entonces:

$f \cdot g = \bar{o}$

sin que f y g sean nulas; porque el cero del anillo en cuestión es la función constante nula y f, g no son funciones constante nula.

En los anillos conmutativos, como es el caso del que estamos considerando, todo divisor de cero a izquierda también lo es a derecha, y recíprocamente. Luego, para anillos conmutativos basta hablar simplemente de divisores de cero. Así, pues, en el anillo de las funciones reales de variable real, las funciones f y g definidas arriba, son divisores de cero.

Por último puede ocurrir también que existan elementos $a \neq o$ en un anillo tales que:

$a^n = o$

donde a^n es el producto de n veces a por sí mismo.

Estos elementos, cuando existan, los llamaremos NILPOTENTES.

Así, en el anillo $\mathbb{Z}_4 = \{\bar{o}, \bar{1}, \bar{2}, \bar{3}\}$, el elemento $\bar{2}$ es nilpotente, ya que

$\bar{2}^2 = \bar{2} \cdot \bar{2} = \bar{o}$

8.3. *Ley cancelativa de la multiplicación*

Hay una propiedad particularmente importante que un anillo puede verificar o no y que está relacionada con la no existencia de divisores de cero: la de la cancelación o simplificación de la multiplicación (a izquierda o a derecha).

Teorema. Un anillo A que verifica la ley de cancelación de la multiplicación no tiene divisores de cero y recíprocamente, si A no tiene divisores de cero verifica la ley de cancelación de la multiplicación.

Dem. a) Supongamos que A es un anillo en el que vale la ley de cancelación: $ab = ac$ (ó $ba = ca$) y si $a \neq o$, entonces $b = c$. Hay que probar que: si $a \cdot b = o$ entonces $a = o$, ó $b = o$

Si $a = o$ no hay nada que demostrar.

Si $a \neq o$, tenemos que mostrar que $b = o$.

La relación $a \cdot b = o$, puede ser escrita en la forma:

$a \cdot b = a \cdot o$.

Luego, aplicando la ley de cancelación a izquierda, resulta:

$b = o$

Asimismo se demuestra que si $b \neq o$, entonces $a = o$

Queda probada así la primera parte del teorema.

b) Supongamos ahora que A es un anillo sin divisores de cero: $a \cdot b = o \Rightarrow a = o$, ó $b = o$. Hay que probar que en A se verifica la ley de cancelación.

Supongamos que sea,

$ab = ac, a \neq o$

Sumando a ambos miembros el opuesto de ac, resulta por la propiedad uniforme:

$$ab - ac = o$$

o sea, $a(b - c) = o$

y puesto que no hay divisores de cero, debe ser:

$$b - c = o$$

o sea, $b = c$

Asimismo, $b\,a = c\,a$, con $a \neq o$, implica $b = c$.

Queda probada así la segunda parte del teorema.

Observaciones. a) El teorema recién probado muestra que la no existencia de divisores de cero y la ley de cancelación de la multiplicación son propiedades equivalentes.

b) La condición necesaria y suficiente para que un elemento no nulo *a*, de un anillo, verifique la ley de cancelación a izquierda,

$$(a \neq o, a\,b = a\,c) \Rightarrow b = c$$

es que dicho elemento *a* no sea divisor de cero a izquierda.

Análogamente, la condición necesaria y suficiente para que un elemento no nulo *a*, de un anillo, verifique la ley de cancelación a derecha,

$$(a \neq o, b\,a = c\,a) \Rightarrow b = c$$

es que dicho elemento *a* no sea divisor de cero a derecha.

c) Cuando el anillo es conmutativo, no interesa distinguir entre divisores de cero a izquierda o a derecha, y se habla simplemente de "divisores de cero". Por lo tanto, la condición necesaria y suficiente para que en un anillo valgan las dos leyes cancelativas (a izquierda y a derecha), es que en dicho anillo no haya divisores de cero. Este teorema, fundamental de la teoría de las ecuaciones, puede, pues, ser falso para un anillo cualquiera. Esto nos conduce a definir dos categorías de anillos. los sin divisores de cero y los con divisores de cero.

*****

Como casos particulares de anillos son interesantes los dominios de integridad y los cuerpos.

Definición. Llamaremos DOMINIO de INTEGRIDAD a todo anillo conmutativo con unidad y sin divisores de cero.

Ejemplo típico. El anillo de los enteros $\mathbb{Z}$ es un dominio de integridad.

También los números racionales, los reales y los complejos forman dominios de integridad.

Otros ejemplos de dominios de integridad son:

Los anillos $\mathbb{Z}_n$ con n primo; así, por ejemplo: $\mathbb{Z}_2$, $\mathbb{Z}_3$, $\mathbb{Z}_5$, $\mathbb{Z}_7$, $\mathbb{Z}_{11}$, etc.

El conjunto de todos los polinomios:

$$a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1} x + a_n$$

de coeficientes números racionales, algebrizado por las operaciones de adición y multiplicación de polinomios en el sentido del álgebra clásica.

8.4. *La noción de Cuerpo*

De todo lo anteriormente expuesto, concluimos que los elementos de un anillo A se pueden sumar, restar y multiplicar (también potenciar) respetando leyes formales análogas a las del cálculo numérico, pero en general en un anillo no existe la división. Por ejemplo, acabamos de decir que los números enteros, los racionales, los reales y complejos forman dominios de integridad. Pero entre el anillo de los enteros y los otros tres hay una diferencia fundamental. En estos últimos es posible la "división", es decir, dados dos números cualesquiera a y $b \neq o$, existe siempre un número x univocamente determinado, llamado *cociente* de a y b, tal que:

$$b \cdot x = a$$

No sucede lo mismo en el anillo de los enteros. En este último son posibles, sin excepción, las operaciones de adición, sustracción y multiplicación. Llamaremos por este motivo a tales operaciones con la denominación común de OPERACIONES ENTERAS.

Por otra parte, si la división no es una operación en $\mathbb{Z}$, o aplicación de $\mathbb{Z} \times \mathbb{Z}$ en $\mathbb{Z}$, por ejemplo, no asigna ningún correspondiente en $\mathbb{Z}$ al par $(3,5) \in \mathbb{Z} \times \mathbb{Z}$, ya que $\frac{3}{5} \notin \mathbb{Z}$, tampoco lo es en $\mathbb{Q}$, pues $\frac{3}{o}$ carece de sentido tanto en $\mathbb{Z}$ como en $\mathbb{Q}$, pero en cambio lo es en el conjunto $\mathbb{Q}^* = \mathbb{Q} - \{o\}$ de los números racionales no nulos. Asimismo, la división es una operación en cada uno de los conjuntos:

$$\mathbb{R}^* = \mathbb{R} - \{o\} \quad \text{y} \quad \mathbb{C}^* = \mathbb{C} - \{o\}$$

de los números reales no nulos y complejos no nulos.

De entre los conjuntos numéricos de las matemáticas clásicas, $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ $\mathbb{R}$ y $\mathbb{C}$, el conjunto $\mathbb{Q}$ de los números racionales es el de menor amplitud que goza de las dos propiedades siguientes:

a) En él son posibles las operaciones enteras: adición, sustracción y multiplicación;

b) En el conjunto $\mathbb{Q}^* = \mathbb{Q} - \{o\}$ que resulta de él suprimiendo el cero, es posible también la división.

Llamaremos por este motivo a tales cuatro operaciones: adición,

sustracción, multiplicación y división, con la denominación común de OPERACIONES RACIONALES.

Lo anterior nos mueve ahora a estudiar en particular aquellos anillos con unidad, no necesariamente conmutativos, en los que es siempre posible la división (excepto por cero); es decir, en los que las ecuaciones:

$b \cdot x = a$ y $x \cdot b = a, b \neq o$

tienen solución.

Surge así el concepto abstracto de cuerpo o campo de racionalidad.

Definición

Llamaremos CUERPO O CAMPO DE RACIONALIDAD a todo anillo con elemento unidad respecto del producto y tal que todo elemento distinto de cero tenga inverso.

En otras palabras: cuerpo es un anillo el cual es grupo no sólo respecto a la operación suma, sino también respecto de la operación producto, si se excluye el cero.

Los dos grupos mencionados en la definición anterior los llamaremos respectivamente *grupo aditivo y grupo multiplicativo del cuerpo*.

En resumen, un cuerpo K es un conjunto con más de un elemento sobre el que están definidas dos operaciones binarias internas, una llamada "adición" (+) y otra llamada "multiplicación" (·), de modo que se verifican los siguientes axiomas:

A_1 : $\forall\, a, b \in K \Rightarrow a + b \in K$
A_2 : $\forall\, a, b, c \in K$ se verifica $a + (b + c) = (a + b) + c$
A_3 : $\forall\, a, b \in K$ se verifica $a + b = b + a$
A_4 : $\exists\, o \in K$ tal que $a + o = a, \forall\, a \in K$
A_5 : $\forall\, a \in K$ existe $(-a) \in K$ tal que $a + (-a) = o$
M_1 : $\forall\, a, b \in K \Rightarrow a \cdot b \in K$
M_2 : $a, b, c \in K$ se verifica $a \cdot (b \cdot c) = (a \cdot b) \cdot c$
M_3 : $\exists\, 1 \in K$ tal que $a \cdot 1 = 1 \cdot a = a, \forall\, a \in K$
M_4 : $a \in K, a \neq o$, existe $a^{-1} \in K$ tal que $a \cdot a^{-1} = a^{-1} \cdot a = 1$
M_5 : $a, b, c \in K$ se verifica $a(b + c) = ab + ac$ y $(b + c)a = ba + ca$

Observaciones. 1) En la lista anterior de axiomas no se postula la conmutatividad de la multiplicación. Luego, ella puede ser que sea y puede ser que no sea conmutativa, según el cuerpo particular que se considere. Si ella es conmutativa, diremos que el cuerpo es conmutativo o abeliano, y se verificará el axioma:

M_6 : $\forall\, a, b \in K$ se cumple $a \cdot b = b \cdot a$

Algunos autores llaman solamente *campo* a los cuerpos conmutativos, y lo denotan por F.

2) En un cuerpo (o en un campo), es $1 \neq o$ y el cero no tiene inverso.

En efecto, si fuera $1 = 0$ resultaría para todo elemento a del cuerpo que,

$$a = a \cdot 1 = a \cdot 0 = 0$$

es decir, el cuerpo tendrá un solo elemento: 0, lo que contradice la definición dada más arriba. Por lo tanto es:

$$1 \neq 0$$

Ahora, si 0 tuviera un inverso x sería:

$$0 \cdot x = 1$$

o sea, $0 = 1$

luego, resultará $1 = 0$ y lo que contradice el hecho anterior de que $1 \neq 0$. Por lo tanto, el cero no tiene inverso multiplicativo.

3) Es oportuno recordar ahora que con la notación U (A) indicábamos el grupo de unidades del anillo A, y además hicimos notar que $0 \notin U(A)$. Estas consideraciones nos mueven a dar la siguiente definición:

Definición. Diremos que un anillo con unidad A es un *anillo de división* si:

$$U(A) = A - \{0\}$$

o sea, si todo elemento $a \in A$, $a \neq 0$, es inversible en A.

Por consiguiente, ahora podemos decir también que por *cuerpo* o por *campo* entendemos a todo anillo de división conmutativo.

4) Un cuerpo será representado por una cuaterna ordenada $(K; 1, +, \cdot)$, donde el primer término es un conjunto, el segundo un elemento particular del conjunto y el tercer y cuarto término son operaciones.

Tal como lo hicimos en grupos y anillos, indicaremos únicamente el cuerpo por el conjunto K (o F), quedando sobreentendidas las dos operaciones (+) y (·) y también el elemento unidad 1.

Quede bien claro que estas convenciones constituyen abusos de lenguaje, pero que son cómodas en los enunciados y escrituras.

8.5. *Consecuencias de la definición de cuerpo*

En un cuerpo se verifican todas las propiedades vistas en anillos más algunas otras que veremos a continuación.

Proposición 1. En un cuerpo no existen divisores de cero.

Dem. Tenemos que probar que en un cuerpo K si,

$$a \cdot b = 0, \text{ entonces es } a = 0, \text{ ó } b = 0$$

Si $a = 0$ no hay nada que demostrar.

Si $a \neq 0$, entonces existe un inverso a^{-1} de a.

Multiplicando ambos miembros de $a \cdot b = 0$ por a^{-1}, por la propiedad uniforme de la multiplicación, resulta:

$$a^{-1} \cdot (a \cdot b) = a^{-1} \cdot 0$$
$$(a^{-1} \cdot a) \cdot b = 0$$
$$1 \cdot b = 0$$
$$b = 0$$

Asimismo se demuestra que si $b \neq 0$, entonces $a = 0$. Lo que demuestra la proposición.

Como, por otra parte, $a = 0$, ó $b = 0$ implica $a \cdot b = 0$ resulta que: La condición necesaria y suficiente para que el producto de dos elementos de un cuerpo sea cero, es que sea cero por lo menos uno de esos elementos. Esta propiedad es fundamental en la teoría de las ecuaciones.

Corolario. Todo cuerpo conmutativo (campo) es un dominio de integridad.

En efecto, supongamos que $a \cdot b = 0, b \neq 0$

Si $b \neq 0$, existe b^{-1} y se tiene:

$$(ab)\, b^{-1} = 0 \cdot b^{-1}$$
$$a\,(b\, b^{-1}) = 0$$
$$a \cdot 1 = 0$$
$$a = 0$$

Se prueba en igual forma: $a \cdot b = 0$, $a \neq 0$, implica $b = 0$.

El recíproco de este corolario no es cierto en general.

Un recíproco parcial de él, es el siguiente:

Teorema. Todo dominio de integridad que tiene un número finito de elementos es un cuerpo.

Dem. Sea el dominio de integridad.

$$D = \{a_1, a_2, \ldots, a_1, \ldots, a_n\}$$

que consta de n elementos.

Para probar que D es un cuerpo, es necesario hacer ver que cada elemento $a \neq 0$ tiene inverso. Formemos todos los productos:

$$(*)\ a\,a_1, a\,a_2, \ldots, a\,a_i, \ldots, a\,a_n$$

que constituyen n elemento de D, todos distintos, porque la condición,

$$a\,a_i = a\,a_j \text{ exige } a_i = a_j$$

puesto que se trata de un dominio de integridad y $a \neq 0$.

El conjunto (*) comprende pues todos los elementos de D una vez y una solamente, en particular el elemento unidad 1, de donde $a\,a_j = 1$; a_i es el inverso de a.

Prácticamente, debe construirse la tabla de multiplicación de D para deducir inmediatamente el inverso de cada elemento.

Anteriormente, vimos que las clases residuales módulo n, $\mathbb{Z}_n$, forman dominios de integridad cuando n es un número primo. Luego, estas clases son igualmente cuerpos.

Así pues, son cuerpos los anillos (dominios de integridad):
$\mathbb{Z}_2, \mathbb{Z}_3, \mathbb{Z}_5, \mathbb{Z}_7, \ldots$

Tomemos uno de ellos, por ejemplo, $\mathbb{Z}_5 = \{\bar{0}, \bar{1}, \bar{2}, \bar{3}, \bar{4}\}$ o bien, $\mathbb{Z}_5 = \{\bar{0}, \bar{1}, \bar{2}, -\bar{2}, -\bar{1}\}$, ya que el opuesto de $\bar{3}$ es $(-\bar{2})$ y el opuesto de $\bar{4}$ es $(-\bar{1})$:

$$\bar{1} + \bar{4} = \bar{0} \Rightarrow \bar{4} = -\bar{1}$$
$$\bar{2} + \bar{3} = \bar{0} \Rightarrow \bar{3} = -\bar{2}$$

Los inversos multiplicativos son:

de $\bar{1}$ es $\bar{1}$; de $\bar{2}$ es $\bar{3}$, o bien, $-\bar{2}$; de $\bar{3}$ es $\bar{2}$; de $\bar{4}$ es $\bar{4}$, o bien, $-\bar{1}$.

Obsérvese la relación:

$$(-a)^{-1} = -a^{-1}$$

puesto que en un anillo se tiene,

$$(-a^{-1}) \cdot (-a) = a^{-1} \cdot a = 1$$

En resumen, en un cuerpo la existencia de inverso excluye la posibilidad de divisores de cero.

Proposición 2. En un cuerpo valen las leyes de cancelación de la multiplicación (a izquierda y a derecha).

Dem. Basta recordar que un cuerpo es en particular un dominio de integridad.

Proposición 3. En un cuerpo cada elemento diferente de cero tiene un único inverso.

Dem. Supongamos, por el contrario, que un elemento a, $a \neq 0$, tiene dos inversos a' y a'', luego,

$$a \cdot a' = 1 \text{ y } a \cdot a'' = 1$$

y por la unicidad del elemento unidad, es

$$a \cdot a' = a \cdot a''$$

y por la ley de cancelación (en este caso a izquierda) resulta:

$$a' = a'' \text{ (absurdo)}$$

Esta contradicción prueba la proposición.

Proposición 4. En un cuerpo, la ecuación de primer grado,

$b \cdot x = a$, donde $b \neq 0$

tiene solución y ésta es única.

Dem. Como $b \neq 0$, existe el inverso b^{-1} de b; entonces, multiplicando ambos miembros de la ecuación $b \cdot x = a$ por b^{-1} a la izquierda, en virtud de la propiedad uniforme del producto, resulta:

$b^{-1}(b\,x) = b^{-1}\,a$

o sea, $(b^{-1} \cdot b) \cdot x = b^{-1}\,a$

$1 \cdot x = b^{-1}\,a$

luego, $x = b^{-1}\,a$

Por consiguiente, como el inverso de *b* es único, si existe alguna solución x es única, porque necesariamente es igual a, $x = b^{-1}\,a$

También la unicidad de la solución se deduce de la ley de cancelación del producto.

Comprobemos que efectivamente $x = b^{-1}\,a$ es solución de la ecuación $b \cdot x = a$. Tenemos:

$b(b^{-1}\,a) = (b\,b^{-1})\,a = 1 \cdot a = a$

Así queda demostrado que la única solución de la ecuación $b \cdot x = a$ es $x = b^{-1}\,a$.

Proposición 5. En un cuerpo, la ecuación de primer grado

$x \cdot b = a$, con $b \neq 0$

tiene solución y ésta es única.

Dem. Análoga a la de la proposición anterior, y siendo $x = a \cdot b^{-1}$ la única solución de la ecuación $x \cdot b = a$.

Observaciones. 1. Las dos últimas proposiciones establecen que dados en un cuerpo dos elementos cualesquiera *a* y *b*, $b \neq o$, quedan unívocamente determinados un elemento x tal que $b \cdot x = a$ y otro elemento *y* tal que $y \cdot b = a$, siendo $x = b^{-1} \cdot a$ e $y = a \cdot b^{-1}$.

En general, si el cuerpo no es conmutativo, es $b^{-1} \cdot a \neq a \cdot b^{-1}$.

Por eso es necesario distinguir el orden de los factores y hablaremos de *cociente a derecha* ($x = b^{-1} \cdot a$) y *cociente a izquierda* ($y = a \cdot b^{-1}$) de *a* por *b*.

Naturalmente, esta distinción es superflua si el cuerpo es conmutativo o abeliano; es decir, si es un campo.

2. Si F es un cuerpo conmutativo (campo), entonces las ecuaciones $b \cdot x = a$ y $x \cdot b = a$, donde $b \neq o$, coinciden y su solución única es $x = b^{-1}\,a = a \cdot b^{-1}$.

Entonces, dado dos elementos *a* y *b*, $b \neq o$, queda unívocamente determinado un elemento que multiplicado por *b* da por resultado *a*, que lo

llamaremos COCIENTE de *a* por *b* y que es el elemento $a \cdot b^{-1}$, producto de *a* por el inverso de *b*. Lo representamos por $\frac{a}{b}$.

De modo que por definición, tenemos:

$$\frac{a}{b} = a \cdot b^{-1}$$

Este cociente lo llamaremos también RAZON o FRACCION de "numerador" *a* y de "denominador" *b*.

En particular, si $a \neq o$, tendremos:

$$\frac{1}{a} = 1 \cdot a^{-1} = a^{-1}$$

3. Queda definida así una nueva operación en el cuerpo, la DIVISION que es una función de $F \times (F - \{o\})$ en F.

Esta operación es evidentemente uniforme, es decir si $a = c$ y $b = d \neq o$, entonces:

$a \cdot b^{-1} = c \cdot d^{-1}$

o sea,

$$\frac{a}{b} = \frac{c}{d}$$

Esta operación conjuntamente con las de adición, sustracción y multiplicación constituyen las cuatro operaciones que hemos denominado Operaciones Racionales de un cuerpo o de un campo.

4. Como todo campo contiene, además del grupo aditivo, un grupo multiplicativo abeliano, resulta entonces que valen para un campo todas las reglas de cálculo que vimos para los cocientes en los grupos multiplicativos y abelianos. Estas reglas son:

a) $\frac{a}{b} = \frac{c}{d}$ si, y sólo si $ad = bc$ $(b \neq o, d \neq o)$

b) $\frac{a}{b} \pm \frac{c}{d} = \frac{ad \pm bc}{bd}$ $(b \neq o, d \neq o)$

c) $\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$ $(b \neq o, d \neq o)$

d) $\frac{a}{b} = \frac{ac}{bc}$ $(b \neq o, c \neq o)$

e) $\left(\frac{a}{b}\right) = \frac{b}{a}$ $(b \neq o, a \neq o)$

f) $\frac{\frac{a}{b}}{c} = \frac{ad}{bc}$ $(b \neq o, d \neq o, c \neq o)$

Además los cocientes en un campo verifican la siguiente regla de signos:

1. $\frac{-a}{b} = -\frac{a}{b}$ $(b \neq o)$
2. $\frac{a}{-b} = -\frac{a}{b}$ $(b \neq o)$
3. $\frac{-a}{-b} = \frac{a}{b}$ $(b \neq o)$

En efecto, bastará demostrar que $\frac{-a}{b}$ y que $\frac{a}{-b}$ son el opuesto de $\frac{a}{b}$. Tendremos, en virtud de propiedades ya vistas, que:

$$\frac{a}{b} + \frac{-a}{b} = \frac{ab + (-a)b}{b^2} = \frac{ab + [-(ab)]}{b^2} = \frac{o}{b^2} = o \cdot (b^2)^{-1} = o$$

Análogamente, se tiene:

$$\frac{a}{b} + \frac{a}{-b} = \frac{a(-b) + ab}{b(-b)} = \frac{-(ab) + ab}{-b^2} = \frac{o}{-b^2} = o \cdot (-b^2)^{-1} = o$$

Finalmente, aplicando sucesivamente las reglas 1 y 2 recién probadas y recordando que: $-(-x) = x$, resulta:

$$\frac{-a}{-b} = -\left(\frac{a}{-b}\right) = -\left(-\frac{a}{b}\right) = \frac{a}{b}$$

Otras reglas de cálculos son:

$$\frac{\frac{a}{b}}{c} = a \cdot \left(\frac{b}{c}\right)^{-1} = a \cdot \frac{c}{b} = \frac{a}{1} \cdot \frac{c}{b} = \frac{ac}{1 \cdot b} = \frac{ac}{b}$$

$$\frac{\frac{a}{b}}{c} = \frac{a}{b} \cdot c^{-1} = \frac{a}{b} \cdot \frac{1}{c} = \frac{a \cdot 1}{bc} = \frac{a}{bc}$$

En resumen, se ve que sobre un cuerpo o sobre un campo son válidas todas las reglas operatorias establecidas en la aritmética elemental de los enteros y fracciones (estos elementos pueden ser positivos o negativos) para las cuatro operaciones: adición, sustracción, multiplicación y división.

También pueden aplicarse los métodos clásicos de la aritmética para la resolución de ecuaciones y sistemas de ecuaciones lineales en cuerpos finitos.

Ejemplos:

1. Sobre el cuerpo $\mathbf{Z}_7$, resolver la ecuación:

$\overline{2} \cdot \overline{x} = \overline{5}$

Multiplicando ambos miembros por $\overline{4}$ (inverso de $\overline{2}$), se tiene:

$\overline{4} \cdot \overline{2} \cdot x = \overline{4} \cdot 5$
$\overline{8} \cdot \overline{x} = \overline{20}$
$\overline{1} \cdot \overline{x} = \overline{6}$
$\overline{x} = \overline{6}$

Comprobación: $\overline{2} \cdot \overline{6} = \overline{12} = \overline{5}$

2. Sobre el cuerpo $\mathbf{Z}_5$, resolver el sistema:

$$\begin{array}{|l} \overline{2} \cdot \overline{x} + \overline{3} \cdot \overline{y} = \overline{2} \\ \overline{1} \cdot \overline{x} + \overline{2} \cdot \overline{y} = \overline{4} \end{array}$$

Sumando miembro a miembro, se tiene:

$\overline{3} \cdot \overline{x} + \overline{5} \cdot \overline{y} = \overline{6}$
$\overline{3} \cdot \overline{x} + \overline{o} \cdot \overline{y} = \overline{1}$
$\overline{3} \cdot \overline{x} = \overline{1}$

Multiplicando por $\overline{2}$ (inverso de $\overline{3}$) resulta,

$\overline{6} \cdot \overline{x} = \overline{12}$
$\overline{1} \cdot \overline{x} = \overline{2}$
$\overline{x} = \overline{2}$

Multiplicando la segunda ecuación del sistema por $\overline{3}$ (inverso de $\overline{2}$) y sumando miembro a miembro, se obtiene

$\overline{5} \cdot \overline{x} + \overline{9} \cdot \overline{y} = \overline{14}$
$\overline{o} \cdot \overline{x} + \overline{4} \cdot \overline{y} = \overline{4}$
$\overline{4} \cdot \overline{y} = \overline{4}$

y multiplicando ambos miembros por $\overline{4}$ (inverso de $\overline{4}$), se tiene

$\overline{16} \cdot \overline{y} = \overline{16}$
$\overline{1} \cdot \overline{y} = \overline{1}$
$\overline{y} = \overline{1}$

3. Sobre el cuerpo $\mathbf{Z}_7$, resolver el sistema:

$$\begin{array}{|l|ll} -\overline{3}\,\overline{x} + \overline{2}\,\overline{y} = \overline{1} & \cdot\overline{3} & \cdot\overline{1} \\ \overline{1}\,\overline{x} + \overline{3}\,\overline{y} = -\overline{2} & \cdot\overline{2} & \cdot\overline{3} \end{array}$$

$$\begin{array}{|l} -\overline{9}\,\overline{x} + \overline{6}\,\overline{y} = \overline{3} \\ \overline{2}\,\overline{x} + \overline{6}\,\overline{y} = -\overline{4} \end{array}$$

$-\overline{11}\,\overline{x} = -\overline{7}$
$\overline{4} \cdot \overline{x} = \overline{o}$
$\overline{x} = \overline{o}$

$\overline{11}\,\overline{y} = -\overline{5}$
$\overline{4}\,\overline{y} = -\overline{5}$

$$\overline{8}\cdot\overline{y} = -\overline{10}$$
$$\overline{1}\cdot\overline{y} = -\overline{3}$$
$$\overline{y} = \overline{4}$$

Comprobación:

$$-\overline{3}\cdot\overline{0} + \overline{2}\cdot\overline{4} = \overline{0} + \overline{8} = \overline{8} = \overline{1}$$
$$\overline{1}\cdot\overline{0} + \overline{3}\cdot\overline{4} = \overline{0} + \overline{12} = \overline{12} = \overline{5} = -\overline{2}$$

8,6. *Subestructuras de un anillo, de un dominio de integridad y de un cuerpo*

En esta sección estudiaremos las subestructuras características de un anillo, de un dominio de integridad y de un cuerpo; esto es, subanillo, subdominio de integridad y subcuerpo. Sus definiciones son:

Definición 1. Diremos que una parte no vacía A' de un anillo A es un SUBANILLO de A, si A' es un anillo con respecto a las operaciones definidas en A.

Teorema. La condición necesaria y suficiente para que $A' \subseteq A$, $A' \neq \emptyset$, sea un subanillo de A, es que se verifiquen las siguientes propiedades:

S_1) $a, b \in A' \Rightarrow a - b \in A'$

S_2) $a, b \in A' \Rightarrow a \cdot b \in A'$

Dem. 1. Si A' es un subanillo de $(A; +, \cdot)$, entonces A' es un anillo $(A'; +, \cdot)$.

En particular, A' es un subgrupo del grupo aditivo de A; luego, $a, b \in A' \Rightarrow b \in A'$ b) $= a - b \in A'$, y se cumple la condición S_1).

Por otro lado, como A' es anillo, A' es cerrado respecto a la multiplicación; esto es, $a, b \in A' \Rightarrow a \cdot b \in A'$, y se cumple la condición S_2).

2. Recíprocamente, supongamos que A' verifica las condiciones S_1) y S_2), y probemos que es un anillo (subanillo de A).

En efecto, por S_1) podemos afirmar que A' es un subgrupo aditivo con respecto a la suma del grupo aditivo del anillo A.

Por otra parte, por S_2), A' es cerrado respecto al producto. Como en A' valen los axiomas M2 y M3, porque ellos valen en A, se sigue que A' es un subanillo del anillo A,

El teorema está demostrado.

Observaciones. 1. Puede suceder que un anillo A tenga unidad y que algún subanillo A' de A no la tenga.

Así por ejemplo, sea $A = \mathbb{Z}$ el anillo de los enteros, que usan anillo conmutativo y con unidad.

Sea $A' = \{0, \pm 2, \pm 4, \pm 6, \ldots \pm 2n, \ldots\}$ el conjunto de los enteros pares. Es claro que A' es un subanillo de $\mathbb{Z}$ y el cual carece de elemento unidad para la multiplicación. Luego, la unidad 1 de $\mathbb{Z}$ no pertenece a A'.

2) Del teorema recién probado (criterio para los subanillos), resulta como consecuencias.

a) $a \in A' \Rightarrow a, a \in A' \Rightarrow a - a = o \in A'$;

b) $o, a \in A' \Rightarrow o - a = -a \in A'$;

c) $a, b \in A' \Rightarrow a, (-b) \in A' \Rightarrow a - (-b) = a + b \in A'$.

En consecuencia, un subconjunto no vacío A' de un anillo A es un subanillo, si:

(1) $a, b \in A' \Rightarrow a + b, a \cdot b \in A'$

(2) $o \in A' \Rightarrow o \in A'$

(3) $a \in A' \Rightarrow (-a) \in A'$

De la misma manera que definimos los subanillos, se definen los subdominios de integridad y los subcuerpos.

Definición 2. Diremos que una parte no vacía D' de un dominio de integridad D es un SUBDOMINIO de D, si D' es asimismo un dominio de integridad respecto a las operaciones definidas en D.

Teorema. La condición necesaria y suficiente para que $D' \subseteq D$, $D' \neq \emptyset$, sea un subdominio del dominio D, es que:

D1) $a, b \in D' \Rightarrow a - b \in D'$ y $a \cdot b \in D'$

D2) $1 \in D \Rightarrow 1 \in D'$

En otros términos, si D' es un subanillo de D y contenga la unidad de D.

Dem. La primera condición D1) es evidente, y también lo es la suficiencia de la segunda, D2) (supuesto cumplida la primera). Veamos que es necesaria.

Si D' es un dominio de integridad debe contener un elemento neutro l' con respecto a la multiplicación y un elemento $a \neq o$. Tenemos entonces $a \cdot 1' = a$, pero como a es también elemento de D se tiene $a \cdot 1 = a$, luego por la ley de cancelación de D resulta $1 = 1'$; luego, $1 \in D'$.

Definición 3. Sea K un cuerpo (o bien un campo). Diremos que un subconjunto no vacío K' de K es un SUBCUERPO, si K' es asimismo un cuerpo respecto a las operaciones de adición y multiplicación en K.

Teorema. La condición necesaria y suficiente para que $K' \subseteq K$, $K' \neq \emptyset$, sea un subcuerpo del cuerpo K, es que:

K1) $a, b \in K' \Rightarrow a - b \in K'$ y $a \cdot b \in K'$
K2) $1 \in K \Rightarrow 1 \in K'$
K3) $a \in K', a \neq o$, implique $a^{-1} \in K'$

En otros términos, si K' es un subdominio de K y contenga con cada elemento distinto de cero, a su inverso.

Dem. ¡Hágalo como ejercicio!

Ejemplos

1) Sabemos que los conjuntos $m\,\mathbb{Z} = \{o, \pm m, \pm 2m, \ldots\}$ son los únicos subgrupos del grupo aditivo $\mathbb{Z}$ de los enteros relativos. Luego, ellos son también subanillos de $\mathbb{Z}$.

2) Cada uno de los conjuntos $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$ son subanillos de los siguientes. También son subdominios de integridad.

$$\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$$

Asimismo, cada uno de los conjuntos $\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$ son subcuerpos de los siguientes.

3) Anteriormente, en ejemplos dados sobre anillos, demostramos que los conjuntos:

$$A = \{ m + n\sqrt{2} : m, n \in \mathbb{Z} \}$$

y $B = \{ m + n\sqrt{-1} : m, n \in \mathbb{Z} \}$

son anillos, y lo probamos por medio de los axiomas de la estructura de anillo.

Ahora, lo demostraremos como subanillo de los números reales, el primero, y como subanillo de los complejos, el segundo. Tenemos: sea $a, b \in A$.

$$a = m + n\sqrt{2}, \quad b = m' + n'\sqrt{2}$$

entonces,

$$a - b = (m + n\sqrt{2}) - (m' + n'\sqrt{2}) = (m - m') + (n - n')\sqrt{2};$$

luego, $a - b \in A$.

Por otro lado, $a \cdot b = (m + n\sqrt{2})(m' + n'\sqrt{2})$

$= (m\,m' + 2n\,n') + (m\,n' + m'\,n)\sqrt{2}$; luego, $a \cdot b \in A$.

Así pues, el conjunto A es un subanillo de los reales.

Procediendo en la misma forma para el conjunto B, se tiene:

$$a = m + n\sqrt{-1}, b = m' = m' + n'\sqrt{-1}$$

$$a - b = (m - m') + (n - n')\sqrt{-1} \in B$$

$$a \cdot b = (m + n\sqrt{-1})(m' + n'\sqrt{-1}) = (m\,m' - n\,n') + (m\,n' + m'\,n)\sqrt{-1} \in B$$

Por lo tanto, el conjunto B es un subanillo del anillo de los números complejos.

8.7. La noción de Ideal.

Con referencia al ejemplo 1) recién señalado, dijimos que el conjunto $m\,\mathbb{Z} = \{o, \pm m, \pm 2m, \ldots\}$ de todos los múltiplos del entero positivo m es un subanillo del anillo de los enteros $\mathbb{Z}$ Por lo tanto, se cumple en él la implicación:

(*) $a \in (m\,\mathbb{Z})$ y $b \in (m\,\mathbb{Z}) \Rightarrow a \cdot b \in (m\,\mathbb{Z})$

es decir, el producto de dos múltiplos de m es múltiplo de m. Pero para que el producto de los enteros sea múltiplo de m basta con que lo sea uno de los factores, y entonces el subanillo $m\,\mathbb{Z}$ de $\mathbb{Z}$ cumple esta condición más fuerte que la (*):

(**) $a \in (m\,\mathbb{Z})$ y $b \in \mathbb{Z} \Rightarrow a \cdot b \in (m\,\mathbb{Z})$

Por esta razón diremos que $m\,\mathbb{Z}$ es un subanillo ideal, o simplemente *Ideal* del anillo $\mathbb{Z}$.

En general, en todo anillo A (supongamos conmutativo) se puede definir la noción de ideal.

Definición. Diremos que un subanillo (J; +, ·) de un anillo conmutativo (A; +, ·) es un IDEAL de este anillo, si el producto de un elemento cualquiera de I por un elemento cualquiera de A pertenece a I, es decir si se cumple la implicación.

$a \in I$ y $b \in A \Rightarrow a \cdot b \in I$.

Dada esta definición y del teorema que da el criterio para subanillos, resulta el siguiente teorema que da el criterio para ideales:

Teorema. Sea (A; +, ·) un anillo conmutativo y sea $I \subseteq A$. Para que (I; +, ·) sea un ideal de (A; +, ·), es condición necesaria y suficiente que se cumplan las implicaciones:

I1) $a \in I$ y $b \in I \Rightarrow a - b \in I$ (subgrupo aditivo de A)
I2) $a \in I$ y $c \in A \Rightarrow a \cdot c \in I$

Observaciones. 1) Si el anillo A no es conmutativo, entonces habrá que distinguir los dos casos siguientes:

i) $a \in I$ y $b \in I \Rightarrow a - b \in I$
$a \in I$ y $c \in A \Rightarrow c \cdot a \in I$
ii) $a \in I$ y $b \in I \quad a - b \in I$
$a \in I$ y $c \in A \quad a \cdot c \in I$

En el primer caso i) hablaremos de un *ideal a la izquierda de A*, y en el segundo caso ii) de un *ideal a la derecha de A*.

Por cierto que en el caso de un anillo conmutativo no es necesario hacer distinción entre estos dos conceptos y se hablará simplemente de ideal; es decir, utilizaremos la palabra ideal como sinónimo de ideal bilateral.

2) Para todo anillo A, los conjuntos $I = \{o\}$ e $I = A$ son ideales del anillo A. El primero $I = \{o\}$ lo llamaremos *ideal nulo*, y el segundo $I = A$, que contiene a todos los elementos de A, lo denominaremos *ideal unidad*. Estos dos ideales, por existir siempre, los llamaremos ideales triviales, y los otros, si existen, serán los ideales propios.

3) En un cuerpo sólo existen ideales triviales.

En efecto, sea K un cuerpo y sea I un ideal de K.

Si $I = \{o\}$, la afirmación queda probada.

Supongamos que $I \neq \{o\}$, luego existe $a \in I$ tal que $a \neq o$.

De $a \neq o$ resulta que existe $a^{-1} \in K$. De $a \in I$ y de $a^{-1} \in K$ resulta $a \cdot a^{-1} = 1 \in I$. Luego, $1 \in I$.

Probemos que $I = K$.

Es claro que $I \subseteq K$ (1)

Sea $a \in K$, entonces de $1 \in I$ y de $a \in K$ resulta:

$1 \cdot a = a \in I$

esto es, $a \in K \Rightarrow a \in I$, o sea $K \subseteq I$ (2)

De (1) y (2) concluimos que,

$I = K$

Por consiguiente, I es un ideal trivial.

4) De la proposición precedente resulta que basta que la unidad pertenezca a un ideal, para que el ideal coincida con el anillo entero. Por lo tanto, si A es un anillo con unidad y si I es un anillo ideal de A que contiene a la unidad 1, entonces $I = A$.

Demostración análoga a la anterior.

5) Si $(I_j)_{j \in J}$ es una familia de ideales de un anillo A, entonces su intersección $I = \bigcap_{j \in J} I_j$ es un ideal de A.

En efecto, si $I = \{o\}$ no hay nada que probar, y por esto supongamos sea $I \neq \{o\}$ y sean $a, b \in I_j \Rightarrow a, b \in \bigcap_{j \in J} I_j \Rightarrow a, b \in I_j$ para cualquier $j \in J$

$\Rightarrow a - b \in I_j \Rightarrow a - b \in \bigcap_{j \in J} I_j = I$

Sea ahora $a \in I$ y $x \in A$; $a \in I \Rightarrow a \in I_j, \forall j \Rightarrow xa, ax \in I_j, \forall J \Rightarrow xa, ax \in \bigcap_{j \in J} I_j = I$

Luego, $I = \bigcap_{j \in J} I_j$ es un ideal del anillo A.

Dado un subconjunto X de un anillo A, y sea $\bigcap_{j \in J} (I_j)$ la familia de todos los ideales que contienen al conjunto X.

Esta familia no es vacía, ya que $A \in (I_j)_j \in J$

Llamaremos *ideal engendrado por X* a la intersección de dicha familia de ideales, y escribiremos:

$$\overline{X} = \bigcap_{j \in J} I_j$$

$\overline{X} \subset I_j$

$\overline{X}$ es el ideal más pequeño que contiene al subconjunto X.

Para fijar ideas, supongamos que $X = \{a_1, a_2, \ldots, a_m\}$ y que el anillo A sea conmutativo y con unidad 1. Entonces, el ideal X contiene los elementos $a_1, a_2, \ldots, a_m$ y también debe contener todas las combinaciones lineales:

$$\sum_{i=1}^{m} x_i a_i$$

de estos elementos con coeficientes x_i en A.

Denotaremos este ideal $\overline{X}$ en la forma:

$$\overline{X} = (a_1, a_2, \ldots, a_m)$$

Luego,

(*) $(a_1, a_2, \ldots, a_m) = \{$Todos los elementos $\sum_{i=1}^{m} x_i a_i$ para $x_i \in A\}$

Este conjunto es un ideal, pues:

$$\sum_{i=1}^{m} x_i a_i - \sum_{i=1}^{m} y_i a_i = \sum_{i=1}^{m} (x_i - y_i) a_i$$

y $a(\sum_{i=1}^{m} x_i a_i) = \sum_{i=1}^{m} (a x_i) a_i$

que son las propiedades que caracterizan a un ideal.

Por otro lado, como el anillo A tiene un elemento unidad, cada a_i es uno de los elementos del conjunto $(a_1, \ldots, a_m)$, a saber, $a_i = o \cdot a_1 + o \cdot a_2 + \ldots + o \cdot a_{i-1} + 1 \cdot a_i + o \cdot a_{i+1} + \ldots + o \cdot a_m$.

Por consiguiente, el conjunto $(a_1, \ldots, a_m)$ definido por (*), es un ideal del anillo A que contiene a todas las a_i y está contenido en cualquier otro ideal que contenga a las a_i; de ahí es que él sea el menor de todos estos ideales que contienen al conjunto $X = \{a_1, \ldots, a_m\}$.

Llamaremos a este conjunto $X = \{a_1, \ldots, a_m\}$ la *base* del ideal $\overline{X} = (a_1, \ldots, a_m)$.

En particular, si $X = \{a\}$ entonces el ideal $\overline{X} = (a)$ generado por el solo elemento *a* será llamado *Ideal Principal*.

6) Dualmente a la interesección de dos ideales, consideremos la suma de dos ideales.

Sean I_1 e I_2 dos ideales de un anillo conmutativo A; consideremos el conjunto:

$I_1 + I_2 = \{$Todas las sumas $a_i + b_j$ tales que $a_i \in I_1$, $b_j \in I_2\}$

Probemos que este conjunto provisto de las operaciones de A es un ideal de este anillo. Tenemos:

sean $(a + b) \in I_1 + I_2$ y $(a' + b') \in I_1 + I_2$, entonces

$(a + b) - (a' + b') = (a - a') + (b - b') \in I_1 + I_2$

ya que $(a - a') \in I_1$ y $(b - b') \in I_2$

Por otra parte, para todo $x \in A$ se tiene:

$$(a + b)x = ax + bx \in I_1 + I_2$$

puesto que $ax \in I_1$ y $bx \in I_2$

Luego, $I_1 + I_2$ es un ideal del anillo A. Llamaremos a este ideal con la denominación de *ideal suma*.

No confundir $I_1 + I_2$ con $I_1 \cup I_2$, porque la unión no da un ideal, en general.

En general, si los ideales I_1 e I_2 en un anillo conmutativo están engendrados por las bases $\{a_1, a_2, \ldots, a_m\}$ y $\{b_1, \ldots, b_n\}$, respectivamente, es decir,

$$I_1 = (a_1, \ldots, a_m) \text{ e } I_2 = (b_1, \ldots, b_n),$$

entonces cualquier suma:

$$a + b = \sum_{i=1}^{m} x_i a_i + \sum_{j=1}^{n} y_j \cdot b_j \in I_1 + I_2$$

está engendrada por los a_i y b_j; de modo que

$$(a_1, \ldots, a_m) + (b_1, \ldots, b_n) = (a_1, \ldots, a_m, b_1, \ldots, b_n)$$

Análogamente se define la suma de varios ideales $I_1, I_2, I_3, \ldots$

7) El conjunto de los ideales de un anillo constituye una red cuando se ordena tal conjunto por la relación de inclusión.

Esto es:

Diremos que $I \leq J$ si, y sólo si $I \subseteq J$.

El ínfimo de dos ideales se define por:

$$I \wedge J = I \cap J$$

y el supremo por:

$$I \vee J = I + J$$

Como se verifica, para cualquier ideal I, la siguiente desigualdad:

$$\{o\} \subseteq I \subseteq A$$

o sea, $\{o\} \leq I \leq A$

concluimos que, los ideales de un anillo A bajo la relación de inclusión, constituyen una red, en que la intersección viene dada por $I \cap J$, y la reunión $I \cup J$ viene dada por la suma $I + J$.

8) La relación de inclusión entre ideales se asocia estrechamente con la relación de divisibilidad entre números.

En efecto, sabemos que en el anillo $\mathbf{Z}$ los conjuntos $m\mathbf{Z}$ que son múltiplos del entero m son los únicos subgrupos del grupo aditivo $\mathbf{Z}$, formado por todos los enteros.

Por otro lado, vimos también que estos conjuntos $m\mathbf{Z}$ son subanillos

e ideales de $\mathbf{Z}$. Por lo tanto, todo subanillo de $\mathbf{Z}$ es un ideal de $\mathbf{Z}$, y todo ideal de $\mathbf{Z}$ es principal.

Los múltiplos de m, es decir $m\mathbf{Z}$, constituyen pues el ideal principal (m). Con esta nueva nomenclatura escribimos,

$$m\mathbf{Z} = (m) = (-m)$$

puesto que:

$$x \in (m) \Rightarrow x = km = (-k)(-m),\ -k \in \mathbf{Z} \Rightarrow x \in (-m).$$

Así pues, el ideal $m\mathbf{Z}$ puede también ser generado por $-m$.

Probaremos ahora que la inclusión,

$$(m) \subset (n)$$

es equivalente a,

$$n \mid m \ (n \text{ divide a } m)$$

es decir, el elemento generador del ideal mayor es divisor del elemento generador del ideal menor.

En efecto, supongamos que sea $(m) \subset (n)$, entonces

$$m \in (n) \Rightarrow m = kn \Rightarrow n \mid m.$$

Recíprocamente, supongamos que $n \mid m$; luego, existe $k \in \mathbf{Z}$ tal que $m = kn$.

Sea $x \in (m) \Rightarrow x = hm = h(kn) = (hk)n \in (n)$, luego

$$(m) \subset (n)$$

Así hemos probado la equivalencia.

$$(m) \subset (n) \iff n \mid m$$

o sea, $m\mathbf{Z} \subseteq n\mathbf{Z} \iff n \mid m$.

De esta manera, por ejemplo, el ideal (6) de todos los múltiplos de 6 está contenido propiamente en el ideal (2) de todos los números pares.

En base a este importante hecho, suele ordenarse a la familia de los ideales de un anillo conmutativo cualquiera A por la siguiente relación:

$$I \leqq J \text{ sí, y sólo si } I \subseteq J$$

I "divide" a J.

9) El mínimo común múltiplo (M.C.M.) y el máximo común divisor (M.C.D.) tienen fácil interpretación en la teoría de los ideales.

Sean *a* y *b* dos enteros arbitrarios.

Consideremos la intersección de los ideales $a\mathbf{Z}$ y $b\mathbf{Z}$, engendrados respectivamente por *a* y *b*.

Como sabemos que la intersección de ideales es un ideal y que todos los ideales de $\mathbf{Z}$ son de la forma $m\mathbf{Z}$, $m \geq o$, existe entonces un entero único $m \leqslant o$ tal que:

$$a\mathbf{Z} \cap b\mathbf{Z} = m\mathbf{Z}$$

El entero positivo m lo llamaremos el M.C.M. de los enteros *a* y *b*, y lo designaremos por:

$m = \text{M.C.M.}(a, b) = a \vee b$.

Afirmamos que, de hecho, m es un múltiplo común de *a* y *b*, y que cada múltiplo común de *a* y *b* es un múltiplo de m.

En efecto, tenemos:

$$\begin{cases} m\,\mathbb{Z} \subseteq a\,\mathbb{Z} \iff a \mid m \\ m\,\mathbb{Z} \subseteq b\,\mathbb{Z} \iff b \mid m \end{cases}$$

con esto probamos ya que m es un múltiplo común de *a* y *b*.

Para demostrar que m es el mínimo múltiplo común de *a* y *b*, es decir, el M.C.M., bastará demostrar que cualquier otro múltiplo común m' de *a* y *b* es un múltiplo de m. Tenemos:

$$a \mid m' \iff m'\,\mathbb{Z} \subseteq a\,\mathbb{Z}$$
$$b \mid m' \iff m'\,\mathbb{Z} \subseteq b\,\mathbb{Z}$$

lo que implica que,

$$m'\,\mathbb{Z} \subseteq a\,\mathbb{Z} \cap b\,\mathbb{Z} = m\,\mathbb{Z} \iff m \mid m'$$

lo que prueba nuestra aseveración.

Análogamente, definimos el M.C.M. de varios números $a_1, a_2, a_3, \ldots, a_n$, como el ideal $m\,\mathbb{Z}$ tal que:

$$m\,\mathbb{Z} = a_1\,\mathbb{Z} \cap a_2\,\mathbb{Z} \cap \ldots \cap a_n\,\mathbb{Z}$$

La importancia del M.C.M. no reside en el hecho de que este número sea el mínimo entre todos los posibles múltiplos de los números $a_1, a_2, \ldots, a_n$, lo que es consecuencia de la buena ordenación de los enteros positivos, sino en el hecho que el M.C.M. es divisor de todos los múltiplos comunes de $a_1, a_2, \ldots, a_n$.

Sean, como antes, *a* y *b* dos enteros arbitrarios.

Consideremos el ideal engendrado por el conjunto {a, b}.

Este ideal de $\mathbb{Z}$, según ya hemos demostrado, es de la forma $d\,\mathbb{Z}$, $d \geq 0$, siendo *d* únicamente determinado por *a* y *b*.

El entero positivo *d* lo llamaremos el M.C.D. de los enteros *a* y *b*, y lo designaremos por:

$$d = \text{M.C.D}(a, b) = a \wedge b$$

Afirmamos que, de hecho, d es un divisor común de *a* y *b*, y que cualquier otro divisor común de *a* y *b* divide *d*.

En efecto, notando que $d\,\mathbb{Z}$ es también el ideal engendrado por la reunión a $\mathbb{Z} \cup b\,\mathbb{Z}$, tenemos:

$$\begin{cases} a\,\mathbb{Z} \subseteq d\,\mathbb{Z} \Rightarrow d \mid a \\ b\,\mathbb{Z} \subseteq d\,\mathbb{Z} \Rightarrow d \mid b \end{cases}$$

y así probamos que *d* es un divisor común de *a* y *b*.

Sea ahora d' otro divisor común de *a* y *b*, se tiene:

$$d' \mid a \Rightarrow a\,\mathbb{Z} \subseteq d'\,\mathbb{Z}$$
$$d' \mid b \Rightarrow b\,\mathbb{Z} \subseteq d'\,\mathbb{Z}$$

lo cual implica,

$$a\,\mathbb{Z} \cup b\,\mathbb{Z} \subseteq d'\,\mathbb{Z}$$

pero siendo $d\,\mathbb{Z}$ la intersección de todos los ideales que contienen a $\mathbb{Z} \cup b\,\mathbb{Z}$, él es el menor de todos, resulta:

$$d\,\mathbb{Z} \subseteq d'\,\mathbb{Z} \iff d' \mid d$$

lo que demuestra nuestra aseveración.

Cabe ahora la siguiente pregunta: ¿Cómo se presenta explícitamente el ideal de $\mathbb{Z}$ engendrado por el conjunto {a, b}?

Es claro que cualquier ideal que contiene *a* y *b*, contiene también todas las combinaciones lineales finitas de la forma:

$$s\,a + t\,b, \text{ con } s\,t \in \mathbb{Z}$$

Pero el conjunto de todas las expresiones de esta forma que acabamos de escribir, es un ideal de $\mathbb{Z}$, como se comprende inmediatamente.

Luego, el conjunto de todas las expresiones de la forma $s\,a + t\,b$, es el ideal engendrado por el conjunto {a, b}.

Es natural designar el conjunto de todas las expresiones de la forma $s\,a + t\,b$, por:

$\{s\,a + t\,b : s, t \in \mathbb{Z}\} = a\,\mathbb{Z} + b\,\mathbb{Z}$ (suma de ideales)

Así pues, si d es el máximo común divisor de *a* y *b*, es decir $d = a \wedge b$, tenemos:

$$d\,\mathbb{Z} = a\,\mathbb{Z} + b\,\mathbb{Z}$$

o sea, $(d) = (a) + (b)$

En consecuencia, $d = a \wedge b$ es el menor entero positivo de la forma:

$$d = s\,a + t\,b, \text{ con } s, t \in \mathbb{Z}.$$

Ejemplo:

Hallar el máximo común divisor de los números 336 y 270.

Tenemos, en virtud de la fórmula general ya vista:

$$(a_1, a_2, \ldots, a_m) + (b_1, b_2, \ldots, b_n) = (a_1, \ldots, a_m, b_1, \ldots, b_n)$$

que,

$$(336) + (270) = (336, 270)$$

Esta regla, combinada con la natural transformación de bases, da:

$$(336, 270) = (336\text{-}270, 270) = (66, 270) = (66, 270\text{-}4 \cdot 66) =$$
$$= (66, 6) = (66 - 10 \cdot 6, 6) = (6, 6) = 6$$

Luego, el M.C.D. $(336, 270) = 6$

Definición. Diremos que dos números enteros *a* y *b* son relativamente primos entre sí, si se tiene:

$$a \wedge b = 1$$

es decir, si el M.C.D. de ellos es 1.

Proposición. Para que a y b sean primos entre sí, es necesario y suficiente que existan ciertos enteros s y t tales que:

$$s\,a + t\,b = 1 \text{ (Relación de Bezout)}$$

Dem. 1) Supongamos a y b relativamente primos entre sí, entonces 1 es su máximo común divisor, y por consiguiente, 1 es un entero de la forma,

$$s\,a + t\,b = 1$$

2) Inversamente, supongamos que:

$$s\,a + t\,b = 1$$

y probemos que $a \wedge b = 1$

En efecto, si $s\,a + t\,b = 1$, entonces 1 es una combinación lineal de a y b, y es la mínima, pues no hay ningún entero positivo no nulo menor que 1; luego, $a \wedge b = 1$, lo que por definición implica que a y b son relativamente primos.

Observación. Otro razonamiento que conduce a este último resultado es el siguiente:

El ideal $a\,\mathbb{Z} + b\,\mathbb{Z}$ contiene a 1; luego, el ideal es:

$$a\,\mathbb{Z} + b\,\mathbb{Z} = \mathbb{Z} \qquad a \wedge b = 1$$

Teorema (de Euclides). Si $c \mid ab$ y si $c \wedge a = 1$, entonces $c \mid b$.

Dem. La hipótesis $c \wedge a = 1$ implica la existencia de enteros s, t tales que:

$$s\,a + t\,c = 1 \Rightarrow s\,ab + t\,bc = b$$

pero $c \mid ab$ (por hipótesis) y $c \mid bc$; por lo tanto, dividiendo c al primer miembro, también dividirá al segundo miembro; esto es,

$$c \mid b \qquad \text{(c.q.d.)}$$

Corolario. Si p es primo y $p \mid ab$, entonces $p \mid a$, ó, $p \mid b$

Dem. Por definición de número primo, los únicos divisores de p son 1 y p. Ahora, si la conclusión $p \mid a$ es falsa, el único divisor común de p y a es 1; así que 1 es el máximo común divisor de a y p, y por lo tanto.

$$s\,a + t\,p = 1 \Rightarrow s\,ab + t\,pb = b$$

Los dos términos del primer miembro son divisibles por p, luego b será divisible por p, que es la segunda alternativa del enunciado.

8.8. *Ideales principales en un anillo conmutativo.*

Anteriormente dijimos que los ideales engendrados por un solo elemento reciben el nombre de ideales principales.

Son muy importantes los anillos en los que los únicos ideales son los principales. Un ejemplo de tales anillos es, como sabemos, el anillo $\mathbb{Z}$ de los anteros.

Sea A un anillo conmutativo y sea a un elemento no nulo de A, que supondremos fijo.

Probaremos que el ideal engendrado por a, que denotaremos por I(a), coincide con el conjunto I de todos los elementos de la forma:

$$n\,a + r\,a$$

donde $r \in A$ y n es un número entero.

En efecto, teniendo presente que un ideal es en particular un subanillo, es decir cerrado con respecto a la suma y el producto, tendremos:

A. 1) si $n = 1$, $1 \cdot a = a \in I(a)$,

2) Si $n > 1$, $na = a + a + \overset{n}{\ldots} + a \in I(a)$, ya que $a \in I(a)$,

3) Si $n = 0$, $na = 0 \cdot a = 0 \in I(a)$, ya que 0 pertenece a cualquier ideal del anillo A,

4) si $n = -1$, $na = (-1)a = -a \in I(a)$, ya que: $0, a \in I(a) \Rightarrow 0 - a \in I(a) \Rightarrow (-a) \in I(a)$,

5) si $n < -1$, $na = (-a) + (-a) + \overset{n}{\ldots} + (-a) \in I(a)$, ya que $(-a) \in I(a)$.

Así hemos probado que el ideal I (a), engendrando por a, contiene al múltiplo natural na, cualquiera sea el entero $n \in \mathbb{Z}$.

B. Por otra parte, por una de las propiedades que caracterizan a un ideal, I (a) debe contener también todos los elementos de la forma.

$$r\,a, \text{ con } r \in A \text{ arbitrario}$$

También entonces, I(a) contiene todos los elementos de la forma:

$$n\,a + r\,a, \text{ con } n \in \mathbb{Z} \text{ y } r \in A.$$

De A) y B) resulta:

$$I = \{n\,a + r\,a : n \in \mathbb{Z}, r \in A\} \subseteq I(a) \qquad (1)$$

C. Demostraremos en seguida que el conjunto I es un ideal del anillo A. En efecto:

a) Sean $i_1, i_2 \in I$, entonces

$$i_1 - i_2 = (n_1\,a + r_1\,a) - (n_2\,a + r_2\,a)$$
$$= (n_1 - n_2)a + (r_1 - r_2)a$$

y siendo, $n_1 - n_2$ = número entero y $r_1 - r_2 \in A$, resulta entonces que, $i_1 - i_2 \in I$

b) Sean $i \in I$ y $s \in A$, entonces

$$s\,i = s(n\,a + r\,a) = (s\,n)\,a + (s\,r)\,a$$
$$= (s\,n + s\,r)a = 0 \cdot a + (s\,n + s\,r)a$$

en que 0 es un entero y $(s\,n + s\,r) \in A$; luego, $si \in I$

De a) y b) concluímos que I es un ideal de A que contiene también a a.

D. Como I(a) es el más pequeño ideal de A que contiene a a, resulta:

$I(a) \subseteq I = \{n\,a + r\,a : n \in \mathbb{Z}, r \in A\}$ (2)

De (1) y (2) se concluye que,

$I(a) = \{n\,a + r\,a : n \in \mathbb{Z}, r \in A\}$

En consecuencia, hemos demostrado que en un anillo conmutativo cualquiera A, el ideal engendrado por un elemento fijo $a \in A$, es el conjunto de todos los elementos de la forma: $n\,a + r\,a$, donde $n \in \mathbb{Z}$ y $r \in A$.

$I(a) = \{n\,a + r\,a : n \in \mathbb{Z} \text{ y } r \in A\}$

En particular, el conjunto de los "multiplos" en A de un elemento fijo $a \in A$:

$I = \{r\,a : r \in A\}$

Es un ideal de A (siendo A un anillo conmutativo).

En efecto, sean $i_1, i_2 \in I = \{r\,a : r \in A\}$, entonces:

$i_1 - i_2 = r_1\,a - r_2\,a = (r_1 - r_2)\,a \in I$

ya que $r_1 - r_2 \in A$

Por otro lado, se tiene también para $i \in I$ y $s \in A$:

$i\,s = (r\,a)\,s = r(a\,s) = r(s\,a) = (r\,s)\,a \in I$ con $r\,s \in A$.

Este resultado y el anterior demuestran que I es un ideal de A.

Pero, no es posible afirmar en general que el elemento *a* pertenezca al ideal $I = \{r\,s : r \in A\}$.

Proposición. Si A es un anillo conmutativo con unidad, entonces,

$I(a) = \{r\,a : r \in A\}$

Dem. Sabemos que $I(a) = \{n\,a + r\,a : n \in \mathbb{Z}, r \in A\}$.

Sea *e* la unidad de A, entonces $i = n\,a + r\,a \in I$ (a) puede escribirse en la siguiente forma:

$i = n(e\,a) + r\,a = (n\,e)\,a + r\,a = (n\,e + r)\,a$

donde $(ne + r)$ está en A, por lo tanto los elementos de I(a) son precisamente aquellos elementos de la forma:

$r\,a$, con $r \in A$,

luego, $I(a) \subseteq \{r\,a : r \in A\}$ (1)

Por otra parte, un elemento cualquiera del ideal $I = \{r\,a : r \in A\}$ puede escribirse también en la forma:

$r\,a = 0 \cdot a + r\,a \in I(a)$

luego, $\{r\,a : r \in A\} \subseteq I(a)$ (2)

De (1) y (2) resulta:

$I(a) = \{r\,a : r \in A\}$

lo que prueba la proposición.

Definición. Un ideal, en que cada elemento de él es un múltiplo de un elemento fijo *a* de un anillo conmutativo A con elemento unidad, se dirá

generado por *a*, lo llamaremos un IDEAL PRINCIPAL, y lo denotaremos con el símbolo (a).

En el caso de los enteros $\mathbb{Z}$, todo ideal es principal. Si m es un número fijo en $\mathbb{Z}$, el ideal $(m) = m\,\mathbb{Z}$ contiene justamente todos los enteros que son múltiplos de m.

$\mathbb{Z}$ mismo es un ideal generado por el elemento unidad 1, y cada elemento de $\mathbb{Z}$ es así un múltiplo de 1.

Llamaremos al anillo $\mathbb{Z}$ con el nombre de *ideal unidad*, y lo denotaremos por $\mathbb{Z} = (1)$.

Más generalmente, si el anillo A tiene unidad e, entonces es claro que,

$A = (e)$

ya que $I(e) = \{r\,e : r \in A\} = \{r : r \in A\} = A$

El ideal cero, (0), es obviamente un ideal principal.

Definición. Llamaremos ANILLO a IDEALES PRINCIPALES, a todo anillo conmutativo A tal que todo ideal en A es principal.

Puesto que suponemos que el anillo A tiene elemento unidad *e*, resulta que todo ideal de A se compone de los múltiplos x a de un elemento fijo $a \in A$ (x variable en A).

Dicho en otra forma: en este caso todo ideal es de la forma $(a) = \{x\,a : x \in A\}$

Estos anillos o ideales principales tienen la importante propiedad de que la descomposición de un elemento en factores primos es única, salvo el orden de los factores.

Definición. Llamaremos IDEAL PRIMO a un ideal P de un anillo conmutativo A, cuando cualquier producto *ab* de elementos arbitrarios de A perteneciente a P tenga al menos un factor *a*, ó *b*, o ambos pertenecientes a P.

Nótese que un número primo p engendra un ideal primo (p) en el anillo $\mathbb{Z}$ de los enteros relativos pues un producto *ab* de dos enteros es múltiplo de p si, y sólo si, uno de los factores *a* ó *b* es múltiplo de p. (recordar que si $p|ab \Rightarrow p|a$ ó $p|b$).

De esta manera, el ideal principal (7) es primo, desde que $a \cdot b \in (7)$ implica $7|a \cdot b \Rightarrow 7|a$, o $7|b$, y por lo tanto, $a \in (7)$, ó $b \in (7)$.

En cambio, el ideal (6) no es primo, desde que, por ejemplo, $12 \in (6)$, empero $12 = 3 \cdot 4 \Rightarrow 3 \notin (6)$ y $4 \notin (6)$, ya que 6 3 y 6 4.

Por consiguiente, en el anillo $\mathbb{Z}$ de los enteros, un ideal propio $P = m\mathbb{Z}$, $m \neq 0$, es un ideal primo si, y sólo si, m es un entero primo.

En cualquier anillo conmutativo con unidad, se puede también definir, además de la suma de dos o más ideales, el producto de dos o más ideales, de esta manera:

$A = (a_1, a_2, \ldots, a_m)$, $B = (b_1, b_2, \ldots, b_n)$

$A \cdot B = \left(\sum_{i=1}^{m} a_i x_i \right) \cdot \left(\sum_{j=1}^{n} b_j x_j \right) =$ {todas las sumas $a_1 b_1 + \ldots + a_m b_n$, para $a_i \in A$ y $b_j \in B$}

donde la base del ideal producto es:

$$(a_1, \ldots, a_m) \cdot (b_1, \ldots, b_n) = (a_1 b_1, a_2 b_2, \ldots a_m b_{n-1}, a_m b_n).$$

Este conjunto es, efectivamente, un ideal; está engendrado por todos los productos ab con un factor en A y otro en B, y es el menor ideal que contiene a tales productos.

En particular, el producto de dos ideales principales (a) y (b) es el ideal principal (ab).

Así, por ejemplo, en el anillo $\mathbb{Z}$ de los enteros racionales, formemos el producto de los dos ideales (2) y (3), entonces el ideal producto (2) (3) estará engendrado por todos los productos de un número finito de factores elegidos de todas las maneras posibles entre cada elemento del ideal (2) y cada elemento del ideal (3); es claro que se obtiene así todos los múltiplos de 6, es decir el ideal (6), Luego, $(2) \cdot (3) = (6)$

Ahora, decir que 6 es igual al producto de 2 por 3 equivale a decir que el ideal (6) es igual al producto del ideal (2) por el ideal (3), y descomponer un número entero racional en un producto de factores enteros o descomponer el ideal engendrado por este número en un producto de ideales principales, es el mismo problema.

Todo esto parece trivial cuando se permanece en el anillo de los números enteros racionales, pero no sucede lo mismo en otros anillos y se demuestra que, en todo anillo formado de números enteros algebraicos contenidos en un mismo cuerpo de los números algebraicos, si un número admite varias descomposiciones en factores primos, el ideal engendrado por este número no admite más que una sola descomposición en un producto de ideales primos. Por lo tanto, aquí la descomposición de un número y la descomposición del ideal que él engendra no son, como lo era en el anillo de los enteros racionales, problemas equivalentes.

Tampoco se puede edificar una teoría de la descomposición en factores primos en el seno de un cuerpo, porque, por definición misma de cuerpo, todo elemento a es divisible por todo elemento $b \neq 0$. La noción de elemento primo no existe pues en un cuerpo. Así, en el cuerpo de los números racionales, el número 7 es divisible por 3, porque:

$$3 \cdot \frac{7}{3} = 7$$

Por consiguiente, una teoría de la descomposición en factores primos no se puede edificar más que un anillo, y más especialmente en un anillo conmutativo con elemento unidad, porque en un anillo, la divisibilidad de un elemento por otro es una propiedad excepcional.

Finalizamos este párrafo sobre el concepto de ideal dando la definición de ideales máximos.

Definición. Diremos que un ideal M de un anillo A es un IDEAL MAXIMO, cuando los únicos ideales de A que contienen a M son M y el mismo anillo A; esto es, si se verifican las dos condiciones siguientes:

a) $M \neq A$ (es decir, M es un ideal propio)

b) Si I es un ideal de A tal que $M \subset I$, entonces es $I = A$, ó $I = M$

8.9. *Isomorfismo y Homomorfismo de anillos*

Aplicando a los anillos el concepto general de isomorfismo y homomorfismo, tenemos las siguientes definiciones:

Definición. 1) Sean $(A; +, \cdot)$ y $(A'; \oplus, \odot)$ dos anillos.

Diremos que una aplicación $f : A \longrightarrow A'$ es un isomorfismo (de anillo) si:

a) f es biyectiva,

b) f preserva las dos operaciones de anillo:

$f(x + y) = f(x) \oplus f(y), \forall\, x, y \in A$

$f(x \cdot y) = f(x) \odot f(y), \forall\, x, y \in A$

Un isomorfismo $f : A \longrightarrow A$ de un anillo A en sí mismo es un automorfismo de A. Entonces, la aplicación idéntica I_A es un automorfismo de (la estructura de anillo) A.

El concepto de isomorfismo de anillos admite una generalización importante que resulta de considerar una aplicación epiyectiva (es decir "sobre"), pero no necesariamente inyectiva de A sobre A'.

Definición. 2) Sean $(A; +, \cdot)$ y $(A'; \oplus, \odot)$ dos anillos.

Diremos que una aplicación $f : A \longrightarrow A'$ es un homomorfismo (de anillo) si,

a) f es epiyectiva,

b) f preserva las dos operaciones de anillo:

$f(x + y) = f(x) \oplus f(y)$

$f(x \cdot y) = f(x) \odot f(y)$

cualesquiera sean $x, y \in A$.

Un homomorfismo $f : A \to A$ de un anillo A en sí mismo es un endomorfismo.

Observaciones. 1) Por lo que vimos en los conceptos generales de isomorfismo y homomorfismo resulta que, todo isomorfismo y todo homomorfismo de anillos transforma el cero (0) de A en el cero (0') de A', la unidad (1) de A en la unidad (1') de A', elementos opuestos en elementos opuestos y elementos inversos en elementos inversos; esto es:

$$f(0) = 0', f(1) = 1'$$
$$f(-x) = -f(x), f(x^{-1}) = [f(x)]^{-1}$$

2) El núcleo o el Ker (f) de un homomorfismo de anillos $f : A \to A'$ es un ideal del anillo A. Más precisamente, el conjunto Nuc (f) de los elementos de A cuyo transformado por f es el cero (0') de A'.

$$\text{Nuc}(f) = \{n \in A : f(n) = o'\} = f^{-1}(o')$$

es un ideal (izquierdo, derecho o bilateral) del anillo A.

En efecto, probemos primero que Nuc(f) = Ker(f) es un subgrupo del grupo aditivo de A.

Sean $n_1, n_2 \in \text{Nuc}(f)$; esto es:

$$f(n_1) = f(n_2) = o'$$

y mostremos que $n_1 - n_2 \in \text{Nuc}(f)$; tenemos:

$$f(n_1 - n_2) = f[n_1 + (-n_2)] = f(n_1) + f(-n_2)$$
$$= f(n_1) + [-f(n_2)] = f(n_1) - f(n_2) = o' - o' = o';$$

luego, $n_1 - n_2 \in \text{Nuc}(f)$.

Por otra parte, sean $n \in \text{Nuc}(f)$ y $a \in A$ arbitrario, entonces,

$$f(a\,n) = f(a) \cdot f(n) = f(a) \cdot o' = o'$$

Asimismo, se tiene:

$$f(n\,a) = f(n) \cdot f(a) = o' \cdot f(a) = o'$$

Luego, $n \in \text{Nuc}(f)$ y cualquiera sea $a \in A$, resulta

$$a\,n \in \text{Nuc}(f) \text{ y } n\,a \in \text{Nuc}(f)$$

Este resultado y el anterior prueban que el Nuc(f) es un ideal del anillo A en el homomorfismo $f : A \to A'$.

Este resultado indica, además, que los ideales de un anillo son análogos a los subgrupos normales de un grupo.

3. Por un teorema visto en el homomorfismo de grupos, resulta que el homomorfismo $f : A \to A'$ es un isomorfismo sí, y sólo si

$$\text{Nuc}(f) = \{o\}$$

4. Sean A y A' dos anillos, siendo A un cuerpo. Entonces, si la aplicación $f : A \to A'$ es un homomorfismo, caben solamente dos únicas posibilidades:

a) f es el homomorfismo trivial $f(x) = o', \forall\ x \in A$, o

b) f es un monomorfismo.

En efecto, suponiendo que el caso a) no ocurre, entonces existe por lo menos un $x \in A$ tal que $f(x) \neq o'$.

Ahora, para probar que f es un monomorfismo es suficiente demostrar, según la observación 3) anterior, que

$$\text{Nuc}(f) = \{o\}$$

Sea pues $o \neq a \in A$, y siendo A un cuerpo existe $a^{-1} \in A$.

Luego, $o' \neq f(x) = f(x \cdot 1) = f(x \cdot a^{-1} \cdot a) = f(x \cdot a^{-1}) \cdot f(a)$

$$\Rightarrow \quad f(a) \neq o'$$

Así, hemos demostrado que si $a \neq o$, entonces $f(a) \neq o'$, o equivalente, $f(a) = o' = f(o) \Rightarrow a = o$, o sea $\text{Nuc}(f) = \{o\}$.

Luego, f es inyectiva; esto es, un monomorfismo.

5. Probaremos en seguida que el único endomorfismo de la estructura de anillo de los conjuntos $\mathbb{Z}$, $\mathbb{Q}$ y $\mathbb{R}$ es la aplicación idéntica $I_{\mathbb{Z}}$, $I_{\mathbb{Q}}$ e $I_{\mathbb{R}}$, respectivamente. Esto es, en otros términos, la aplicación idéntica (según el caso) es el único, automorfismo del anillo $\mathbb{Z}$, o del cuerpo de los números racionales, o del cuerpo de los números reales.

En efecto, sea $f : \mathbb{Z} \to \mathbb{Z}$ un endomorfismo de $\mathbb{Z}$.

Puesto que $f(1) = 1$, se sigue que, para $n \geq o$

$$f(n) = f(1 + 1 + \overset{n}{\ldots} + 1) = f(1) + \overset{n}{\ldots} + f(1)$$
$$= 1 + 1 + \overset{n}{\ldots} + 1 = n$$

y siendo, por otro lado, $f(-n) = -f(n) = -n$, concluimos que $f(x) = x$ cualquiera sea $x \in \mathbb{Z}$.

Luego, $f = I_{\mathbb{Z}}$, y como la aplicación idéntica I_A es un automorfismo de cualquier anillo A, se concluye que $I_{\mathbb{Z}}$ es el único automorfismo del anillo $\mathbb{Z}$ de los enteros racionales.

Sea ahora $f : \mathbb{Q} \to \mathbb{Q}$ un endomorfismo (la estructura del anillo) del cuerpo de los números racionales $\mathbb{Q}$.

Sea $\frac{n}{m} \in \mathbb{Q}$, $m \neq o$, $n, m \in \mathbb{Z}$, entonces:

$$f(1) = 1 \text{ y } f\left(\frac{n}{m}\right) = f(n \cdot m^{-1}) = f(n) \cdot f(m^{-1}) = f(n) \cdot [f(m)]^{-1}$$
$$= n \cdot m^{-1} = \frac{n}{m}$$

cualquiera sea $\frac{n}{m} \in \mathbb{Q}$. Luego, $f = I_{\mathbb{Q}}$

Así hemos probado también que la aplicación idéntica $I_{\mathbb{Q}}$ es el único automorfismo de $\mathbb{Q}$.

Finalmente, sea $\mathbb{R}$ el cuerpo de los números reales, y sea $f: \mathbb{R} \to \mathbb{R}$ un endomorfismo de (la estructura de anillo) $\mathbb{R}$. Entonces, puesto que $f(1) = 1$, se tiene también,

$$f(x) = x, \text{ si } x \in \mathbb{Z}$$

y también, $f\left(\frac{n}{m}\right) = \frac{n}{m}$, $m \neq o$, si $\frac{n}{m} \in \mathbb{Q}$

Sea ahora un número real r cualquiera. Queremos demostrar que la aplicación f de homomorfismo conserva el signo, es decir, que si $r > o$, también $f(r) > o$ y si $r < o$ también $f(r) < o$.

Basta para ello admitir dos propiedades muy finas de $\mathbb{R}$ (más tarde las probaremos), como son la existencia de raíces cuadradas de números positivos y la llamada "densidad" de $\mathbb{Q}$ en $\mathbb{R}$, que afirma que dados dos números reales cualesquiera a y b, $a < b$, existe siempre un número racional q que satisface a la relación:

$$a < q < b$$

Sentado esto, sea $r \in \mathbb{R}$, $r > o$, entonces existe $x \in \mathbb{R}$ tal que $x^2 = r$. Por lo tanto,

$$f(r) = f(x^2) = f(x \cdot x) = f(x) \cdot f(x) = f(x)^2$$

de manera que $f(r) \geq o$.

Pero, la situación $f(r) = o$ debe excluírse en virtud de la observación 4) de más arriba. Luego, hemos probado que,

$$r > o \Rightarrow f(r) > o$$

Sea ahora $r \in \mathbb{R}$ y $r < o$, entonces $-r > o$, y como $f(-r) = -f(r)$, resulta por lo anterior que

$$f(-r) = -f(r) > o \Rightarrow f(r) < o$$

Probado ya que el homomorfismo f conserva el signo, supongamos que para algún número real x fuera,

$$f(x) \neq x$$

entonces cabe las dos posibilidades siguientes:

a) $f(x) < x$

b) $x < f(x)$

En la situación a), eligiendo un número racional q tal que $f(x) < q < x$, sería:

$$f(x - q) = f(x) - f(q) = f(x) - q > o$$

ya que $x - q > o$ y $f(q) = q$ por ser q racional, por lo que se vio antes más arriba. Luego, resulta,

$$q < f(x)$$

en contra de la elección de q. Esta situación a) no es pues posible.

Veamos ahora la situación b). Eligiendo como antes un racional q tal que $x < q < f(x)$, tendríamos:

$$f(q - x) = f(q) - f(x) = q - f(x) > o$$

de donde, resulta:

$$q > f(x)$$

que también contradice la elección de q. Luego, tampoco esta segunda situación es posible. Por tanto debe tenerse en todos los casos,

$$f(x) = x, \quad \forall\ x \in R$$

En consecuencia, la aplicación idéntica I_R es el único endomor-

fismo de (la estructura de anillo) $\mathbb{R}$; o sea, el único automorfismo en el cuerpo de los números reales es la aplicación idéntica $I_{|R}$.

2. Sea A un anillo y sea A' un conjunto no vacío sobre el cual están definidas dos operaciones $\oplus$ y $\odot$ correspondientes a las del anillo A.

Sea f un homomorfismo de A sobre A', es decir

$$f(a + b) = f(a) \oplus f(b)$$
$$f(a \cdot b) = f(a) \odot f(b)$$

cualquiera sea $a, b \in A$.

Entonces, diremos que A' es una imagen homomórfica de A y que f es el homomorfismo (de anillo) correspondiente.

Teorema. Toda imagen homomórfica de un anillo es un anillo.

Dem. Sea A un anillo y se A' una imagen homomórfica de A, y sea f el homomorfismo correspondiente.

Denotando con el mismo signo las operaciones correspondientes u homólogas en A y A', probaremos que A' es un anillo.

Sean a', b', c', ... elementos arbitrarios de A'. Como f es una aplicación "sobre" A', existen a, b, c, ... en A tales que $f(a) = a'$, $f(b) = b'$, $f(c) = c'$, ...

Bastará demostrar que sobre A' se cumplen todos los axiomas que caracterizan la estructura de anillo. Tenemos:

A_1) $a', b' \in A' \Rightarrow a' + b' \in A'$

Sea $c' = a' + b' = f(a) + f(b) = f(a + b)$

y como $a + b \in A$, resulta $c' = f(a + b) \in A'$

A_2) $a' + (b' + c') = (a' + b') + c'$. En efecto,

$$a' + (b' + c') = f(a) + [f(b) + f(c)] = f(a) + f(b + c) =$$
$$= f[a + (b + c)] = f[(a + b) + c] = f(a + b) + f(c) =$$
$$= [f(a) + f(b)] + f(c)$$
$$= (a' + b') + c'$$

A_3) Existe $o' \in A'$ tal que $a' + o' = a'$, $\forall\ a' \in A'$

Sea $o' = f(o) \in A'$, entonces

$$a' + o' = f(a) + f(o) = f(a + o) = f(a) = a'$$

A_4) Dado $a' \in A'$, existe $x' \in A'$ tal que $a' + x' = o'$

Sea $x' = f(-a)$, entonces

$$a' + x' = f(a) + f(-a) = f[a + (-a)] = f(o) = o'$$

Escribiremos $x' = -a'$, esto es

$$f(-a) = -f(a)$$

A_5) $a' + b' = b' + a'$

$$a' + b' = f(a) + f(b) = f(a + b) = f(b + a) = f(b) + f(a)$$

luego, $a' + b' = b' + a'$

Por consiguiente, A' es un grupo conmutativo (aditivo en este caso).

Veamos en seguida los axiomas de la multiplicación.

$M_1)$ $a', b' \in A' \Rightarrow a' \cdot b' \in A'$

Sea $c' = a' \cdot b' = f(a) \cdot f(b) = f(a \cdot b)$

y como $a \cdot b \in A$, resulta $c' = f(a \cdot b) \in A'$

$M_2)$ $a' \cdot (b' \cdot c') = (a' \cdot b') \cdot c'$

$a' \cdot (b' \cdot c') = f(a) \cdot [f(b) \cdot f(c)] = f(a) \cdot f(b \cdot c) =$
$= f[a \cdot (b \cdot c)] = f[(a \cdot b) \cdot c] = f(a \cdot b) \cdot f(c) =$
$= [f(a) \cdot f(b)] \cdot f(c)$
$= (a' \cdot b') \cdot c'$

$M_3)$ $a' \cdot (b' + c') = a' \cdot b' + a' \cdot c'$

$a' \cdot (b' + c') = f(a) \cdot [f(b) + f(c)] = f(a) \cdot f(b + c) =$
$= f[a \cdot (b + c)] = f(a \cdot b + a \cdot c) =$
$= f(a \cdot b) + f(a \cdot c) = f(a) \cdot f(b) + f(a) \cdot f(c) =$ $=$
$= a' \cdot b' + a' \cdot c'$

Análogamente se prueba la otra ley distributiva.

Luego, A' es un anillo.

Nótese que si A es un anillo con unidad, entonces A' es también un anillo con unidad.

En efecto, basta tomar $1' = f(1)$, y se tiene

$a' \cdot 1' = f(a) \cdot f(1) = f(a \cdot 1) = f(a) = a'$

Si A es un anillo conmutativo, A' es también un anillo conmutativo.

En efecto, tenemos,

$a' + b' = f(a) + f(b) = f(a + b) = f(b + a) = f(b) + f(a) =$
$= b' + a'$

Puede suceder que A no tenga divisores de cero y A' si los tenga.

Supongamos que el anillo A' tenga divisores de cero, es decir,

$f(a) \cdot f(b) = o'$, $f(a) \neq o'$ y $f(b) \neq o'$

o sea, $f(a \cdot b) = f(o) \Rightarrow a \cdot b \in \mathrm{Nuc}\,(f)$

Ahora bien, como f es un homomorfismo, la relación anterior no implica siempre que,

$a \cdot b = o$

En cambio, sí la implica si f es un isomorfismo.

8.10. *Equivalencias regulares en anillos*

Aquí vamos ahora a analizar la compatibilidad de una relación de equivalencia "$\sim$" definida sobre un anillo A, con respecto a la estructura de anillo.

Definición. Diremos que una equivalencia "$\sim$" en un anillo A es REGULAR si es "compatible con la adición y multiplicación del anillo"; es decir:

$$\left.\begin{matrix} x \sim x' \\ y \sim y' \end{matrix}\right\} \Rightarrow \begin{cases} 1)\ x + y \sim x' + y' & (\sim \text{ regular resp. a la adición}) \\ 2)\ x \cdot y \sim x' \cdot y' & (\sim \text{ reg. resp. a la multiplic.}) \end{cases}$$

Diremos que "$\sim$" es REGULAR a la IZQUIERDA si es "compatible a la izquierda con la estructura de anillo" de A:

$$\forall\, a \in A \quad y \quad x \sim y \Rightarrow \begin{cases} a + x \sim a + y \\ a \cdot x \sim a \cdot y \end{cases}$$

Diremos que "$\sim$" es REGULAR a la DERECHA si es "compatible a la derecha con la estructura de anillo" de A:

$$\forall\, a \in A \quad y \quad x \sim y \Rightarrow \begin{cases} x + a \sim y + a \\ x \cdot a \sim y \cdot a \end{cases}$$

Se ve inmediatamente que:

Para que una equivalencia "$\sim$" sea regular con respecto a la multiplicación, es necesario y suficiente que sea regular a la izquierda y a la derecha.

Las características de las relaciones de equivalencia regulares a la izquierda están dadas por el siguiente teorema:

Teorema. Sea A un anillo. Entonces:

a) Si "$\sim$" es una equivalencia regular a la izquierda, entonces el subconjunto $I \subset A$ definida por

$I = \{x : x \sim o\}$

es un ideal a la izquierda de A que satisface la relación:

$x \sim y \Longleftrightarrow y - x \in I$

b) Recíprocamente, si I es un ideal a la izquierda de A, entonces la relación,

$x \sim y \Longleftrightarrow y - x \in I$

es una relación de equivalencia sobre A regular a la izquierda.

Dem. a) Restringida la equivalencia "$\sim$" a la estructura aditiva de A, sabemos de la teoría de grupos que hay correspondencia biunívoca entre todas las equivalencias "$\sim$" en un anillo A *regulares* con respecto a la *adición* y todos los subgrupos aditivos $I \subseteq (A; +)$, de modo que:

$x \sim y \Longleftrightarrow y - x \in I$

Vamos a probar que el subgrupo aditivo I es un ideal a la izquierda de A.

En efecto, supongamos "$\sim$" regular a la izquierda con respecto a la multiplicación. Tenemos, en particular:

$$x \sim o \Rightarrow a \cdot x \sim a \cdot o = o, \quad \forall\, a \in A,$$

es decir, $x \in I \Rightarrow a \cdot x \in I, \quad \forall\, a \in A$

Luego, I es un ideal a la izquierda del anillo A.

b) Recíprocamente, supongamos que I es un ideal a la izquierda de A. Entonces sea,

$$x \sim y \Longleftrightarrow y - x \in I, \text{ y } a \in A.$$

Tenemos:

$$a\,y - a\,x = a(y - x) \in I$$

porque I es un ideal a la izquierda. Luego,

$$a\,x \sim a\,y$$

Por lo tanto, "$\sim$" es regular a la izquierda con respecto a la multiplicación.

Asimismo se demuestra que:

Para que la equivalencia "$\sim$" asociada con un subgrupo I del grupo aditivo (A; +) del anillo A sea regular a la derecha con respecto a la multiplicación, es necesario y suficiente que I sea un ideal a la derecha.

La combinación de esta proposición, con la anterior enunciada en el teorema recién probado, nos permite enunciar el siguiente:

Teorema. Para que la equivalencia "$\sim$" asociada con el subgrupo I sea regular, es necesario y suficiente que I sea un ideal bilateral del anillo A.

Enunciamos otra vez;

Teorema. Hay correspondencia biunívoca entre todas las equivalencias "$\sim$" regulares en A y todos los ideales bilaterales I de A, de modo que:

$$x \sim y \Longleftrightarrow x - y \in I$$

y que se suele escribir: $x \equiv y$ (mód. I).

El resultado obtenido es pues similar al que obtuvimos al estudiar las relaciones de equivalencia regulares en grupos que caracterizaron a los subgrupos normales.

Por consiguiente, es posible ahora, de modo similar a lo efectuado en grupos, plantearnos el problema de definir una estructura de anillo en el conjunto cociente $A/\sim$ de un anillo A por una relación de equivalencia "$\sim$" regular o compatible con la estructura del anillo A, que convierta a la aplicación canónica:

$$f : A \to A/\sim \; = A/I$$

en un homomorfismo de anillo.

Tenemos así el siguiente teorema.

Teorema. Sea A un anillo y sea "$\sim$" una equivalencia regular en A, o equivalentemente el ideal bilateral I correspondiente.

Sea $A/\sim \; = A/I$ el conjunto cociente de A por "$\sim$", o por I, y sea $f : A \to A/I$ la aplicación canónica definida por:

$$x \to \bar{x} = x + I = \text{Clase de equivalencia de } x;$$

donde $\bar{x} = x + I = \{$ conjunto de todos los elementos de A de la forma, $\underline{|x + i,|}$ con i variable en I$\}$

Entonces:

a) En A/I existe una única estructura de anillo que convierte a f en un homomorfismo de anillos.

b) Nuc(f) = I, es decir el ideal bilateral asociado a la equivalencia regular "$\sim$" definida sobre A.

Dem. a) En A/I definimos una adición y una multiplicación así:

$$\begin{cases} \bar{a} + \bar{b} = \overline{a + b}, \text{ o sea, } (a + I) + (b + I) = a + b + I \\ \bar{a} \cdot \bar{b} = \overline{a \cdot b}, \text{ o sea, } (a + I) \cdot (b + I) = ab + I \end{cases}$$

para todo a, b $\in$ A.

Como la equivalencia, $x \sim y \Leftrightarrow x - y \in I$, o bien

$$x \equiv y \text{ (mód. I)}$$

es regular en A, es fácil verificar que el resultado de las dos operaciones (+) y (·) definidas anteriormente, no depende del elemento particular o individual que se elija en cada clase de equivalencia $\bar{a}$ y $\bar{b}$.

También es fácil probar que estas dos operaciones dotan al conjunto A/I de una estructura de anillo.

Este anillo A/I lo llamaremos el ANILLO COCIENTE del anillo A por la equivalencia regular "$\sim$", o por el ideal bilateral I asociado a "$\sim$", y él es una imagen homomorfa de A por la aplicación canónica:

$$f : A \to A/\sim \; = A/I$$

que denominamos homomorfismo canónico del anillo A.

b) Por lo visto en el teorema de homomorfismo de grupos, tenemos:

$$\text{Nuc}(f) = \{x : x \sim o\}$$

y siendo "$\sim$" una relación de equivalencia compatible con la estructura de anillo de A, resulta que Nuc(f) es un ideal bilateral I asociado a "$\sim$".

Este resultado y el anterior demuestran el teorema.

Casos particulares triviales.

Si $I = \{o\}$, entonces A/I es isomorfo a A.

Si $I = A$, es decir si I es el ideal unidad, entonces A/I se reduce a cero (anillo de un solo elemento).

El ideal $I = A$ lo hemos llamado "unidad" por la siguiente razón: Si el anillo A tiene elemento unidad 1, entonces A entero es el ideal engendrado por 1.

Como complemento del teorema anterior tenemos el siguiente:

Teorema. Sea A un anillo y sea I un ideal bilateral de A.

Entonces, son equivalentes entre sí las proposiciones:

a) I es ideal bilateral de A.

b) Existe un anillo Q y un homomorfismo $f : A \to Q$ tal que $\text{Nuc}(f) = I$.

Dem. Tenemos que probar las dos implicaciones lógicas:

$a) \Rightarrow b)$ y $b) \Rightarrow a)$

La primera $a) \Rightarrow b)$, es consecuencia inmediata del teorema anterior, ya que basta tomar $Q = A/I$ y elegir f como el homomorfismo canónico,

$$f : A \to A/I = Q$$

Probemos ahora que $b) \Rightarrow a)$.

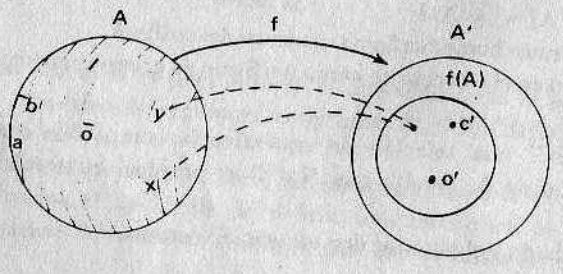
En efecto, por ser $f: A \to Q$ un homomorfismo de anillos y cuyo núcleo $I = \text{Nuc}(f)$ es un subgrupo de $(A; +)$, tenemos para cada $x \in A$ y cada $a \in I$:

$$f(x\,a) = f(x) \cdot f(a) = f(x) \cdot o = o$$
$$f(a\,x) = f(a) \cdot f(x) = o \cdot f(x) = o$$

de donde resulta que $xa \in I$ y $ax \in I$.

Luego, $I = \text{Nuc}(f)$ es un ideal bilateral de A.

Observación importante. Como lo vimos en grupos, veremos también que cada homomorfismo de un anillo A se puede reducir al homomorfismo canónico $f : A \to A/I$ merced al teorema siguiente:



Teorema. (Teorema de homomorfía para anillos).

Sea f un homomorfismo de un anillo A a un anillo A'.

Entonces, $f^{-1}(\{o'\})$ es un ideal bilateral I de A y $f(A)$ es un anillo isomorfo a A/I.

Dem. Definamos en A una relación "$\sim$" así:

$$x \sim y \iff f(x) = f(y)$$

(es decir, x e y tienen la misma imagen en $f(A)$).

Se comprueba inmediatamente que "$\sim$" es una relación de equivalencia regular en A.

En efecto,

$$a \sim a' \iff f(a) = f(a')$$
$$b \sim b' \iff f(b) = f(b')$$

y como f es un homomorfismo, se tiene:

$$f(a) \cdot f(b) = f(a') \cdot f(b')$$
$$f(ab) = f(a'\,b')$$

de donde $a\,b \sim a'\,b'$

La clase de equivalencia del $o \in A$ es $f^{-1}(\{o'\})$. Luego, $f^{-1}(\{o'\})$ es un ideal bilateral I del anillo A.

Las otras clases de equivalencia son los conjuntos:

$$a + I, \quad b + I, \ldots$$

elementos del conjunto cociente A/I.

Observemos que si $x \in a + I$, entonces $f(x) = f(a) = a'$.

En efecto, si $x \in a + I \Rightarrow x = a + i$, donde $i \in I$. Luego,

$$f(x) = f(a + i) = f(a) + f(i) = f(a) + o' = a' + o' = a'$$

Por lo tanto,

$$a + I = f^{-1}(a') \quad \text{(imagen completa inversa)}.$$

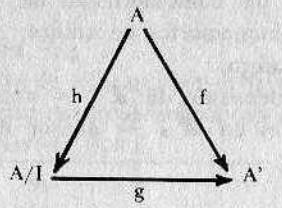
Definamos una aplicación $g : A/I \to f(A)$ así:

$$g(a + I) = f(a) = a'$$

Es inmediato que la definición de g es unívoca.

Se comprueba también, tal como lo hicimos en grupos, que g es un isomorfismo de A/I sobre $f(A)$.

Por consiguiente, es conmutativo al diagrama:



o sea, $g \circ h = f$.

Esto nos sugiere que todas las imagenes homomórficas de un anill se pueden obtener dentro del anillo. Por lo tanto, podemos afirmar, e consecuencia, que todas las imágenes homomórficas de un anillo, salv un isomorfismo, pueden obtenerse en esta forma, es decir, haciendo e cociente por un ideal bilateral de elementos representados en cero.

En resumen: sea A un anillo. Hay correspondencia biunívoca entre:

a) Equivalencias regulares en A.
b) Ideales bilaterales en A.
c) Homomorfismos de A.

De manera más precisa: un anillo A' es homomorfo de un anillo conmutativo A si, y sólo si, A' es isomorfo de A/I, anillo cociente de A por el ideal bilateral I de elementos de A representados en el cero del anillo A'.

En este isomorfismo, el cero (o') de A' es homólogo de la clase residual I ϵ A/I, y los demás elementos de A', de una manera general, son homólogos de las clases residuales módulo el ideal I.

En particular, si I = {o}, la congruencia a $\equiv$ b (mód. I) es una *igualdad* y A y A' son isomorfos.

Si I = A = (1), la congruencia módulo I es la *equivalencia absoluta*, es decir, sólo hay una clase de equivalencia en A, el mismo A, y en tal caso A' se reduce al solo elemento o'.

Ejemplos:

1. Siendo los únicos ideales de $\mathbb{Z}$ los ideales principales m $\mathbb{Z}$ = (m), m $\geq$ o, entonces las únicas imagenes homomorfas de $\mathbb{Z}$ son: $\mathbb{Z}$ mismo y los $\mathbb{Z}_m = \{\bar{0}, \bar{1}, \ldots, \overline{m-1}\}$, donde, como ya sabemos, $\mathbb{Z}_m$ es un anillo conmutativo con unidad, y este anillo es cuerpo si, y sólo si, m es primo.

2. Sean m y n dos enteros positivos.

Cabe la pregunta: ¿Existe un homomorfismo del anillo $\mathbb{Z}_n = \{\bar{0}, \bar{1}, \ldots, \overline{n-1}\} = \{o_n, 1_n, 2_n, \ldots\}$ sobre el anillo $\mathbb{Z}_m = \{\bar{0}, \bar{1}, \ldots, \overline{m-1}\} = \{o_m, 1_m, 2_m, \ldots\}$?

Afirmamos que la respuesta es positiva si y sólo si m| n.

En efecto, si existe un homomorfismo de $\mathbb{Z}_n$ sobre $\mathbb{Z}_m$, entonces por el teorema de homomorfía de anillos, $\mathbb{Z}_m$ será isomorfo a un anillo cociente de $\mathbb{Z}_n$. Luego,

m = número de elementos de $\mathbb{Z}_m$ = número de cogrupos o clases laterales correspondientes en $\mathbb{Z}_n$ = divisor del número de elementos de $\mathbb{Z}_n$, o sea

m| n

La condición es pues necesaria.

Recíprocamente, supongamos que m| n.

Vamos a mostrar que existe un homomorfismo f *único* de $\mathbb{Z}_n$ sobre $\mathbb{Z}_m$.

Designaremos por α_n la clase residual del entero α módulo n, y por α_m la clase residual del entero α módulo m.

Ahora, si f es un homomorfismo de $\mathbb{Z}_n$ a $\mathbb{Z}_m$, entonces f debe aplicar el elemento unidad 1_n sobre el elemento 1_m, o sea

$$f(1_n) = 1_m$$

Luego, si α es un entero positivo arbitrario, se tiene:

$$f(\alpha_n) = f(1_n + 1_n + \overset{\alpha}{\ldots} + 1_n) = f(1_n) + \overset{\alpha}{\ldots} + f(1_n) = 1_m + 1_m + \overset{\alpha}{\ldots} + 1_m = \alpha_m.$$

Asimismo, si α es un entero negativo o cero. Luego, f, si existe, es único, a saber:

$$(*)\ f(\alpha_n) = \alpha_m$$

Definamos una aplicación f de $\mathbb{Z}_n$ a $\mathbb{Z}_m$ por la propia fórmula (*).

La definición de f es unívoca. En efecto, si

$$\alpha_n = \beta_n \Rightarrow n|\alpha - \beta \Rightarrow m|\alpha - \beta$$

ya que, por hipótesis, m| n.

Luego, $m|\alpha - \beta \Rightarrow \alpha_m = \beta_m$

Por otro lado, f es evidentemente una aplicación epiyectiva (sobre).

Además, tenemos:

$$f(\alpha_n + \beta_n) = f[(\alpha + \beta)_n] = (\alpha + \beta)_m = \alpha_m + \beta_m = f(\alpha_n) + f(\beta_n)$$

y asimismo,

$$f(\alpha_n \cdot \beta_n) = f[(\alpha \cdot \beta_n)] = (\alpha \cdot \beta)_m = \alpha_m \cdot \beta_m = f(\alpha_n) \cdot f(\beta_n)$$

Luego, f es un homomorfismo de $\mathbb{Z}_n$ sobre $\mathbb{Z}_m$.

El núcleo Ker(f) es el conjunto de todas las clases α_n tales que $m|\alpha$, es decir las que van al o_m de $\mathbb{Z}_m$ mediante f.

Ejemplo: Homomorfismo de $\mathbb{Z}_{12}$ sobre $\mathbb{Z}_3$

$$\mathbb{Z}_{12} = \{o_{12}, 1_{12}, 2_{12}, \ldots, 10_{12}, 11_{12}\}$$
$$\mathbb{Z}_3 = \{o_3, 1_3, 2_3\}$$

Núcleo de f es,

$$\text{Ker}(f) = \{o_{12}, 3_{12}, 6_{12}, 9_{12}\}$$

es decir, todos los elementos de $\mathbb{Z}'_{12}$ que son divisibles por 3.

Formamos así el siguiente cuadro de cogrupos y sus imágenes por f:

COGRUPOS	IMAGENES POR f
{o, 3, 6, 9}	0_3
{1, 4, 7, 10}	1_3
{2, 5, 8, 11}	2_3

Como ejercicios, fórmese sumas y productos entre estos cogrupos y averíguese qué imágenes les corresponde.

*****

Veamos ahora qué condiciones debe cumplir el ideal I de un anillo A, conmutativo con unidad, para que el anillo cociente A/I sea un dominio de integridad o un campo.

Las proposiciones siguientes dan las respectivas respuestas.

Proposición 1. Si A es un anillo conmutativo con elemento unidad, entonces el anillo cociente A/I será un dominio de integridad si, y sólo si, I es un ideal primo en A.

Dem. a) Supongamos que el ideal I sea primo, y sea

$$f : A \longrightarrow A/I$$

el homomorfismo canónico correspondiente.

Sea $\bar{a}, \bar{b} \in A/I$ y supongamos que:

(1) $\bar{a} \cdot \bar{b} = \overline{a \cdot b} = \bar{o}$

siendo $\bar{a}$ y $\bar{b}$ las clases de equivalencia según el módulo I de los elementos *a* y *b*, respectivamente, en A.

Ahora bien, una clase de equivalencia $\bar{x}$ según I es cero, sí, y sólo si, *x* se encuentra en el ideal I. Así pues, la condición (1) se traduce por:

(2) $a \cdot b \in I$

Pero, siendo, por hipótesis, I un ideal primo, entonces (2) implica $a \in I$ ó $b \in I$, es decir $\bar{a} = \bar{o}$ ó $\bar{b} = \bar{o}$.

Así hemos probado que cualquier producto en A/I es cero, si al menos un factor es cero; es decir, si A/I es un dominio de integridad.

Luego, I = ideal primo $\Rightarrow$ A/I = dominio de integridad.

b) Recíprocamente, supongamos que A/I es un dominio de integridad, es decir:

$$\bar{a} \cdot \bar{b} = \bar{o} \Rightarrow \bar{a} = \bar{o} \text{ ó } \bar{b} = \bar{o}$$

o sea, $a \cdot b \in I \Rightarrow a \in I$ ó $b \in I$

Resulta, pues, que I cumple exáctamente la definición de ideal primo.

Luego, A/I = dominio de integridad $\Rightarrow$ I = ideal primo.

Este resultado y el anterior demuestran la proposición.

Proposición 2. Si A es un anillo conmutativo con unidad, entonces el anillo cociente A/I será un campo sí, y sólo si, I es un ideal máximo en A.

Dem. a) Supongamos que I es máximo en A, y sea

$$f : A \longrightarrow A/I$$

el homomorfismo canónico correspondiente.

Sea $\bar{a} = a + I$ un elemento *no nulo* de A/I, es decir

$$a \notin I$$

Para probar que A/I es un cuerpo, bastará demostrar que se pueden resolver en A/I las ecuaciones:

(1) $\bar{a} \cdot \bar{x} = \bar{b}$

(2) $\bar{y} \cdot \bar{a} = \bar{b}$

siendo $\bar{b} = b + I$ un elemento arbitrario dado de A/I.

Nos limitaremos a mostrar que se puede resolver (1), ya que para la (2) el razonamiento es similar.

Consideremos el ideal a la derecha H engendrado por la reunión $I \cup \{o\}$.

Es claro que H debe contener todos los elementos de la forma:

$$H = \{ax + i x : x \in A, i \in I\} = \{ax + j : x \in A, j \in I\}$$

H es efectivamente un ideal de A, ya que si $i_1, i_2, \in H$

se tiene:

$$i_1 = a x + j_1 \ , \ i_2 = a x + j_2$$

entonces,

$$i_1 - i_2 = j_1 - j_2 \in I \subset H$$

$$\forall \ z \in A, i \cdot z = (ax + j) z = a(xz) + jz \in H$$

lo que prueba que H es un ideal a la derecha de A.

Por otra parte, siendo I un ideal máximo de A y siendo I contenido arbitrariamente en H, ($I \subset H$), tenemos H = A.

Luego, si *b* es un elemento arbitrario de A, existe $x \in A$ y $j \in I$ tales que,

$$a x + j = b$$

Tomando las clases de equivalencia módulo I, resulta:

$$\bar{a} \cdot \bar{x} + \bar{o} = \bar{b}$$

o sea, $\bar{a} \cdot \bar{x} = \bar{b}$ (que resuelve (1)).

b) Recíprocamente, supongamos que A/I es un cuerpo.

Sea, por ejemplo, H un ideal a la derecha de A que contiene I estrictamente, esto es $I \subset H$. Hay que mostrar que H = A.

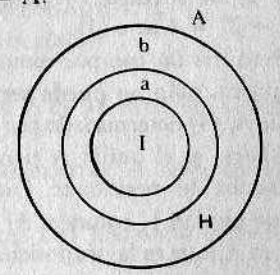
Sea $a \in H$, $a \notin I$,

Sea *b* un elemento arbitrario de A.

Por hipótesis se puede resolver la ecuación,

$$\bar{a} \cdot \bar{x} = \bar{b}$$

es decir, existe $x \in A$ tal que $\bar{a} \cdot \bar{x} = \bar{b}$.



Equivalentemente, $\bar{a} \cdot \bar{x} - \bar{b} = \bar{o} \Rightarrow \overline{a \cdot x - b} = \bar{o} \Rightarrow$

(*) $a x - b \in I \subset H$

Pero, $a \in H, x \in A \Rightarrow ax \in H$, por ser H un ideal a la derecha.

Resulta ahora de (*) que, $b \in H$.

Como *b* es un elemento arbitrario de A, resulta

$$A \subseteq H$$

y como, $H \subseteq A$

concluimos que $H = A$, e I es un ideal máximo.

Este resultado y el anterior prueban la proposición.

Observaciones. 1) Como cualquier campo (cuerpo conmutativo) es en particular un dominio de integridad, resulta entonces, por las dos últimas proposiciones, que cualquier ideal máximo es primo.

Por lo tanto, cualquier ideal máximo de un anillo conmutativo con unidad 1, es primo.

Así, en $\mathbb{Z}$ los ideales $m\mathbb{Z} = (m)$ son primos y maximales a la vez, aquellos donde $m > o$ es un número primo.

2. En la Proposición 2), inmediatamente anterior, se probó que si A/I es un cuerpo, entonces I es un ideal máximo.

Este resultado es válido en cualquier anillo (con o sin unidad 1).

El recíproco no es válido, en general; esto es, puede ser I maximal y el cociente A/I no ser cuerpo.

Consideremos, en efecto, un contra-ejemplo. Sea A el conjunto de los enteros pares,

$$A = \{o, \pm 2, \pm 4, \pm 6, \ldots\}$$

que es, como sabemos, un anillo cuando se lo algebriza mediante las operaciones ordinarias de adición y multiplicación.

En este anillo, el ideal (4), de los múltiplos de 4 es maximal, ya que si n es un entero par perteneciente a A, y no divisible por 4, entonces el M.C.D. $(4, n) = 2$, por lo tanto el ideal $(4, n) = (2) = A$.

Así pues, $(4) \subset (2) \Rightarrow (2) = A$; luego, (4) es maximal.

Sin embargo, el anillo cociente A/(4) no es un cuerpo, desde que $4 = 2^2 \equiv o$ (mód. (4)), mientras que $2 \not\equiv o$ (mód. (4)). Luego, $2 \not\equiv o$ $(4) \Rightarrow 2^2 = 4 \equiv o\,(4)$.

Pero, en un cuerpo, como lo veremos más tarde, el cuadrado de un elemento no nulo no puede ser igual a cero. Por lo tanto, concluimos que el anillo A/(4) determinado por el ideal maximal (4), no es un cuerpo.

Ahora, si el anillo A tiene unidad 1, entonces sí vale el recíproco que acabamos de mencionar. Esto es, si A es un anillo conmutativo con elemento unidad 1, entonces A/I es un campo sí, y sólo si, I es un ideal maximal en A (que es la proposición 2) anteriormente demostrada).

*****

Consideremos en el mismo anillo de los enteros pares $A = \{o, \pm 2, \pm 4, \pm 6, \pm 8, \ldots\}$ el ideal (6) de los múltiplos de 6. Es fácil constatar como antes que tal ideal es también maximal en A. Pero, en este caso no ocurre que exista un número entero par n no divisible por 6 tal que $n^2 \equiv o$ $(6), n \equiv o\,(6)$.

Además, aquí sí que A/(6) es un campo, por el hecho de que A/(6) es isomorfo con $\mathbb{Z}/(3)$.

Este ejemplo conjuntamente con el inmediatamente anterior nos mueven a enunciar el siguiente resultado general.

Teorema. Si I es un ideal en un anillo conmutativo A, entonces A/I es un campo si, y sólo si, se satisfacen las dos condiciones siguientes:

a) I es un ideal maximal en A,

b) Si $x \notin I$, entonces $x^2 \notin I$, o equivalentemente, si $x^2 \equiv o(I)$ entonces $x \equiv o(I)$.

Dem. Sea $f : A \to A/I$ el homomorfismo canónico correspondiente.

Probaremos primeramente que f(M) es un ideal de A/I, si M es ideal de A. Sean $i_1', i_2' \in f(M)$, entonces existen $i_1, i_2 \in M$ tales que:

$$f(i_1) = i_1', \qquad f(i_2) = i_2'$$

Como M es un ideal, resulta:

$i_1 - i_2 \in M \Rightarrow f(i_1 - i_2) \in f(M) \Rightarrow f(i_1) - f(i_2) = i_1' - i_2' \in f(M)$

Sea ahora, $i' \in f(M)$ y $a' \in A/I$; entonces existen $i \in M$ y $a \in A$ tales que:

$$f(i) = i' \quad y \quad f(a) = a'$$

Como M es un ideal, $i \in M$ y $a \in A \Rightarrow ia \in M$; luego,

$f(ia) \in f(M)$

o sea, $f(i) \cdot f(a) = i' \cdot a' \in f(M)$

Asimismo se prueba que $i' \in f(M)$ y $a' \in A/I \Rightarrow a' \cdot i' \in f(M)$

Por tanto, f(M) es un ideal bilateral del anillo cociente A/I.

Así hemos probado que en el homomorfismo canónico

$$f : A \to A/I$$

la imagen directa f(M) de cada ideal del anillo A es un ideal del anillo cociente A/I.

Con este antecedente previo, pasemos a demostrar el teorema enunciado. Tenemos:

a) Supongamos que A/I es un campo, y sea f el homomorfismo natural o canónico,

$$f : A \to A/I$$

I es el núcleo de f, es decir $f(I) = o'$

Supongamos que existe un ideal M tal que,

$$I \subset M$$

entonces f(M) es un ideal de A/I.

Como por hipótesis A/I es un campo, los únicos ideales de A/I son los triviales. Esto es,

$$f(M) = \{o'\} \quad \text{ó} \quad f(M) = A/I.$$

Luego, ó $I = M$, ó $M = A$.

Por consiguiente, I es un ideal máximo en A.

Así hemos probado a).

Sea ahora $x \notin I$, entonces $f(x) = x' \neq o'$, y mostremos que $x^2 \notin I$.

Supongamos, por el absurdo, que $x^2 \in I$, esto es $f(x^2) = o'$

Pero $f(x^2) = f(x \cdot x) = f(x) \cdot f(x) = x' \cdot x' = o'$, con $x' \neq o$, absurdo, ya que en un cuerpo no existen divisores de cero. Por tanto, $x \notin I \Rightarrow x^2 \notin I$.

Así hemos probado b).

Este resultado y el anterior prueban que la condición es necesaria.

b) Recíprocamente, sea I un ideal máximo en A que verifica la condición: $x \notin I \Rightarrow x^2 \notin I$. Probaremos entonces que A/I es un campo.

En efecto, sea $x' \in A/I$ y $x' \neq o'$. Entonces, existe $x \in A$ tal que $f(x) = x'$.

Como $x' \neq o'$, entonces $x \notin I$. Luego, por la hipótesis $x^2 \notin I$.

Ahora bien, los múltiplos de x, que son elementos de la forma $a\,x$, con $a \in A$, forman un ideal M que contiene en particular a, $x^2 = x \cdot x$.

Luego, M no está contenido en I. Pero,

$$I \subset I + M, \qquad M \subseteq I + M$$

y como $x^2 \in M$, $x^2 \notin I$, resulta $x^2 \in I. + M$; luego, I está contenido estrictamente en $I + M$; o sea

$$I \subset I + M$$

Pero I es por hipótesis un ideal máximo en A, luego

$$I + M = A$$

Ahora, para concluir que A/I es un cuerpo, bastará probar que la ecuación de primer grado:

$$x' \cdot z' = y', \quad \text{con } x' \neq o'$$

tiene solución en A/I.

Consideremos el conjunto G' de todos los elementos no nulos de A/I. Probaremos que la ecuación,

$$(1)\ x' \cdot z' = y', \text{ con } x', y' \in G'$$

tiene solución $z' \in A/I$.

En efecto, como $y' \in G'$, entonces por la definición de G' es $y' \neq o'$. Luego, $z' \neq o'$. Por lo tanto, la ecuación (1) tiene solución en G'. Por consiguiente, por lo que vimos en grupos, G' es un grupo respecto a la multiplicación. Entonces, podemos afirmar que G' tiene una unidad e'. Esto es,

$$x' \cdot e' = x', \forall\, x' \in G'$$

Además, $o' \cdot e' = o'$

Luego, e' es también una unidad de A/I.

Finalmente, nos queda por mostrar que todo $x' \neq o'$ tiene un inverso multiplicativo, y lo cual es cierto, ya que la ecuación,

$$x' \cdot z' = e'$$

como lo vimos antes, tiene solución en G'.

Así hemos probado que la condición es suficinte.

Corolario. Si A es un anillo conmutativo con unidad, entonces la condición necesaria y suficiente para que A/I sea un campo, es que I sea un ideal máximo en A.

Este corolario no es otro que la Proposición 2) anteriormente enunciada y demostrada.

8.11. *Homomorfismos de Cuerpos.*

Sabemos que un cuerpo K no tiene más ideales que los triviales: el ideal nulo (o) y el ideal unidad K.

Resulta de esto, que solamente pueden existir dos anillos cocientes de K y que son:

$$K/(o) = K \text{ y } K/K = (o)$$

Por lo tanto, solamente pueden existir dos homomorfismos de un cuerpo K sobre un conjunto $K' = f(K)$, a saber:

a) Aquel que corresponde a $I = (o)$, es un isomorfismo de K y del conjunto K', que por lo tanto es un cuerpo;

b) Aquel que corresponde a $I = K$, K' es un anillo reducido a un solo elemento o'.

Luego, todo homomorfismo de un cuerpo K sobre K' es un isomorfismo.

Aplicación importante del teorema de homomorfia

8.12. *Característica de un Anillo*

Sea A un anillo con elemento unidad *e*.

Si x es un elemento arbitrario de A y $n \in \mathbb{Z}$, definimos:

$$(*)\ n\,x = \begin{cases} x + x + \overset{n}{\ldots} + x, & \text{si } n > o \\ o, & \text{si } n = o \\ -(x + x + \overset{-n}{\ldots} + x), & \text{si } n < o \end{cases}$$

En particular, si $x = e$ se tiene:

$$(**)\ n\,e = \begin{cases} e + e + \overset{n}{\ldots} + e, & \text{si } n > o \\ \text{ó}, & \text{si } n = o \\ -(e + e^{-n} + \ldots + e), & \text{si } n < o \end{cases}$$

Nótese también, en el caso $n < o$, la validez de,

$$n\,e = (-e) + (-e) + \ldots + (-e) = (-n)\cdot(-e)$$

Nótese además que $f(1) = e$

Definamos ahora una aplicación $f : \mathbb{Z} \to A$ así:

$$f(n) = n\,e$$

Probemos que esta aplicación es un homomorfismo.

Sean $m, n \in \mathbb{Z}$. Entonces, tendremos que considerar los siguientes casos posibles:

(1) $m > o,\ n > o$
(2) $m < o,\ n > o$
(3) $m > o,\ n < o$
(4) $m < o,\ n < o$

Tenemos:
En (1), es inmediato que $f(m + n) = f(m) + f(n)$
En (2) se tiene:

$$\begin{aligned} m\,e + n\,e &= (-m)\cdot(-e) + n\cdot e \\ &= (-m)\cdot(-e) + (-n)(-e) \\ &= [(-m) - n]\cdot(-e) \\ &= [n - (-m)]\cdot e \\ &= [n + m]\,e \end{aligned}$$

Por lo tanto, $f(m + n) = f(m) + f(n)$
También en (3) se tiene:

$$f(m + n) = f(m) + f(n)$$

por ser análogo a (2).
Finalmente, en (4) se tiene:

$$\begin{aligned} m\,e + n\,e &= (-m)\cdot(-e) + (-n)\cdot(-e) \\ &= [(-m) + (-n)]\cdot(-e) \\ &= [-(m + n)]\cdot(-e) = (m + n)\,e \end{aligned}$$

o sea, $f(m + n) = f(m) + f(n)$

Los casos en que $m = o$, ó $n = o$ son inmediatos.

Así hemos probado que en todos los casos nuestra aplicación f es un homomorfismo de las estructuras aditivas de los anillos $\mathbb{Z}$ y A.

De la ley distributiva generalizada,

$$m\,e \cdot n\,e = (e + \overset{m}{\ldots} + e)\cdot(e + \overset{n}{\ldots} + e) = e + \overset{mn}{\ldots} + e = (mn)e$$

se concluye que también $f(mn) = f(m)\cdot f(n)$

Luego, f es también un homomorfismo de las estructuras multiplicativas de los anillos $\mathbb{Z}$ y A.

En consecuencia, f es un homomorfismo de $\mathbb{Z}$ a A.

Probaremos en seguida que la aplicación f es el único homomorfismo de $\mathbb{Z}$ dentro de A.

En efecto, si $g: \mathbb{Z} \to A$ es otro homomorfismo, entonces él debe satisfacer $g(1) = e$, $g(2) = g(1) + g(1) = e + e = 2\,e$, $g(-2) = -g(2) = -2\,e = (-2)e$, y así en general, $g(n) = ne$ cualquiera sea $n \in \mathbb{Z}$.

Luego, de la definición de igualdad de funciones, resulta:

$$g = f$$

lo que muestra la unicidad del homomorfismo de $\mathbb{Z}$ en A.

Por el teorema de homomorfía, el conjunto de los elementos de A de la forma *ne*, $n \in \mathbb{Z}$, es un subanillo de A isomorfo a un anillo cociente de $\mathbb{Z}$:

$$\mathbb{Z}/_{(o)} = \mathbb{Z},\quad \mathbb{Z}/_{(n)} = \mathbb{Z}/_{n} = \{\overline{o}, \overline{1}, \ldots, \overline{n-1}\}$$

es decir, bien a $\mathbb{Z}$ mismo, o bien a un $\mathbb{Z}_n$, siendo n un entero positivo oportuno.

En el primer caso, diremos que la CARACTERISTICA del anillo A es *cero*. En el segundo caso, el entero positivo n lo llamaremos la CARACTERISTICA de A.

En resumen, la unicidad del homomorfismo $f: \mathbb{Z} \to A$, demostrado arriba, nos permiten asociar unívocamente,

$$A \to n$$

al anillo A el entero *no negativo* n ($n \geq o$), que denominamos la *característica del anillo A*.

Veamos a continuación la propiedad fundamental del número no negativo n, que indica la característica de A.

Sea $m \in \mathbb{Z}$ tal que $m \cdot e = o'$ en el anillo A.

Examinaremos este hecho en cada una de las dos eventualidades:

a) $\mathbb{Z}/(o) = \mathbb{Z} \cong \{ne : n \in \mathbb{Z}, e \in A\} \subset A$

b) $\mathbb{Z}/(n) = \mathbb{Z}_n \cong \{ne : n \in \mathbb{Z}, e \in A\} \subset A$

en que se presenta de una manera única el homomorfismo $f : \mathbb{Z} \to A$.

En a) tenemos, $Nuc(f) = (n) = (o) \Rightarrow n = o$, por tratarse de una aplicación inyectiva f de $\mathbb{Z}$ en A. Por lo tanto,

$$m \neq o \Rightarrow f(M) = me \neq o', o' \in A$$

Luego, $m \cdot e = o' \Rightarrow m = o$

También, cualquiera sea $a \in A$, se tiene:

$$ma = m(ea) = (me)a = o' \cdot a = o'$$

Por consiguiente, en el caso de un anillo A de característica cero, el producto *ne* es siempre distinto de cero, para cualquier entero positivo n.

El anillo $\mathbb{Z}$ de los enteros racionales tiene característica cero.

En la situación b), tenemos:

$$Nuc(f) = (n), n \neq o$$

y entonces, $f(m) = me = o' \Leftrightarrow m \in Nuc(f) = (n) \Leftrightarrow n \mid m$

Análogamente, como en el caso a, cualquiera sea $a \in A$, se tiene:

$$ma = m(ea) = (me)a = o' \cdot a = o'$$

Por consiguiente, en este segundo caso, diremos que un anillo A tiene característica finita $n > o$, si $na = o'$ cualquiera que sea $a \in A$, y ningún entero positivo menor que n tiene esta propiedad.

Si algún otro entero positivo m tiene la propiedad $ma = o'$, entonces $n \mid m$, y recíprocamente; esto es, son equivalentes las dos proposiciones siguientes:

1) $ma = o'$

2) $n \mid m$ (m es divisible por n).

En efecto, probemos que 1) $\Rightarrow$ 2), esto es:

Si $ma = o$, entonces $m = kn, k \in \mathbb{Z}$

Como se supone $m > n$, entonces por el algoritmo de división se escribe,

$$m = kn + r, \quad o < r < n$$

Entonces,

$$o' = ma = (kn + r)a = k(na) + ra = o' + ra = ra,$$

esto es, $ra = o'$, lo que es absurdo ya que $r < n$ y n es el menor entero positivo que verifica la condición,

$$na = o'$$

Luego, $m = kn$, es decir $n \mid m$.

Mostremos ahora que 2) $\Rightarrow$ 1), esto es:

Si $n \mid m$, entonces $ma = o'$.

En efecto, por hipótesis tenemos, $m = kn, k \in \mathbb{Z}$.

Entonces, $ma = (kn)a = k(na) = k \cdot o' = o'$

Este resultado y el anterior prueban la equivalencia:

$$1) \Leftrightarrow 2)$$

o sea, $ma = o' \Leftrightarrow n \mid m$

donde n es la característica finita (es decir no nula) del anillo A.

Observaciones Importantes. Resulta difícil la clasificación de todos los anillos, pero en los dominios de integridad y en los cuerpos se llega a resultados más simples, como pasamos a verlos.

Sea A un dominio de integridad (es decir, un anillo conmutativo, con unidad y sin divisores de cero).

Sea $a \in A$ un elemento de orden n, tendremos:

$$na = o, a \neq o$$

Probaremos que si $b \in A$, $b \neq o$, es otro elemento cualquiera, también se tiene:

$$nb = o$$

En efecto, si $na = o$, entonces

$$(na) \cdot b = o$$

o sea, $n(a\,b) = n(b\,a) = (n\,b)\cdot a = o = o\cdot a$
y como vale la ley de cancelación de la multiplicación, resulta de,

$$(n\,b)\,a = o\cdot a \Rightarrow n\,b = o$$

En consecuencia, en el grupo aditivo de un dominio de integridad, todos los elementos no nulos tienen el mismo orden.

Este orden común de los elementos no nulos en el grupo aditivo de un dominio de integridad lo llamaremos la característica del dominio.

Como el elemento unidad *e* no es nulo, podemos decir también que la característica del dominio es el menor entero positivo n tal que,

$$n\,e = o$$

Si *n e* es siempre distinto de cero, para cualquier entero positivo n, entonces diremos que la característica del dominio es cero, o sin característica.

Teorema. La característica de un dominio de integridad es bien cero, o bien un número entero primo p.

Dem. Si la característica es cero, entonces

$$n\,a = a + a + \overset{n}{\ldots} + a = o \Rightarrow a = o, \forall\, a$$

Si la característica no es cero, entonces la relación

$$n\,e = o \Rightarrow n = \text{número primo.}$$

Para demostrarlo, supondremos que, por el contrario, existe algún dominio con característica compuesta, como $n = r\,s$.

Entonces, por la ley distributiva general:

$$(m\,a)\cdot(n\,b) = (mn)\,(ab)$$

escribimos, $n\,e = (r\,s)\,e = (r\,e)\cdot(s\,e) = o$

y como no hay divisores de cero, resulta:

$$o \text{ bien } r\,e = o, \quad o \text{ bien } s\,e = o$$

Luego, la característica de nuestro dominio en cuestión deberá ser un divisor de r o de s, pero no *n*, como habíamos supuesto.

Esta contradicción prueba el teorema.

En consecuencia, todo dominio de integridad tiene bien característica cero, en cuyo caso $n\,a = o \Rightarrow a = o$, o bien característica prima p, en cuyo caso vale,

$$p\,a = a + a + \overset{p}{\ldots} + a = o, \text{ con } a \neq o$$

El dominio fundamental de característica cero es $\mathbb{Z}$, y el de característica p es $\mathbb{Z}_p = \{\overline{o}, \overline{1}, \overline{2}, \ldots, \overline{p-1}\}$

El sentido de estas dos afirmaciones es el siguiente:

Teorema. 1) En cualquier dominio D de característica cero, el subgrupo aditivo cíclico engendrado por el elemento unidad *e* es un subdominio isomorfo con el dominio de los enteros racionales, y

2) En cualquier dominio D de característica p, el subgrupo aditivo cíclico engendrado por el elemento unidad *e* es un subdominio isomorfo con el dominio $\mathbb{Z}_p$ de los enteros módulo p.

Demostraciones. 1) El subgrupo cíclico infinito en cuestion es:

$$S = \{\ldots, -m\,e, \ldots, -2\,e, -e, o, e, 2\,e, \ldots, m\,e, \ldots\}$$

y que está formado por todos los múltiplos naturales del elemento unidad *e*.

En virtud de que,

$$(m\,e)\cdot(n\,e) = (mn)\cdot e$$

concluímos de que S es también multiplicativamente cerrado.

Luego, S es un subdominio de D.

Este subdominio puede ser puesto en correspondencia biunívoca con el dominio $\mathbb{Z}$ de los enteros por la regla:

$$me \longleftrightarrow m$$

que preserva sumas y productos; esto es:

$$f(m\,e + n\,e) = f(me) + f(ne) = m + n$$

$$f[(me)\cdot(ne)] = f(mn\,e) = mn$$

Luego, la aplicación $f : S \to \mathbb{Z}$ es un isomorfismo.

2) Aquí el subgrupo cíclico finito S en cuestión es:

$$S = \{e, 2e, 3e, \ldots, (p-1)e, pe = o\}$$

que está formado por todos los múltiplos naturales distintos del elemento unidad e.

Este subgrupo S es también multiplicativamente cerrado, ya que $(me) \cdot (ne) = (mn)e$, siendo este resultado $(mn)e$ uno de los elementos $o, e, 2e, \ldots, (p-1)e$, como lo vimos en grupos cíclicos finitos (C_n).

Por lo tanto, S es un subdominio de D, que consta de p elementos distintos y puede ser puesto en correspondencia biunívoca con el dominio $\mathbb{Z}_p$ de los enteros módulo p por la regla:

$$me \longleftrightarrow \overline{m}$$

Probemos que la aplicación $f : S \to \mathbb{Z}_p$ definida por la regla anterior es una biyección. Sea,

$$f(me) = f(ne)$$

o sea, $\overline{m} = \overline{n} \Rightarrow m \equiv n \text{ (mód. p)} \Rightarrow m - n \equiv o \text{ (mód. p)}$ y esto, a su vez, equivale a

$$f[(m-n)e] = o$$

$$f[me - ne] = o$$

y lo que equivale a

$$me - ne = o \Rightarrow me = ne$$

Luego, f es inyectiva, y como, bajo la regla de correspondencia $me \longleftrightarrow \overline{m}$, cada elemento de $\mathbb{Z}_p$ es el correspondiente de un elemento de S, e inversamente, concluímos que la aplicación f es también una biyección.

Por otra parte, es fácil constatar que esta aplicación f preserva sumas y productos, esto es,

$$f(me + ne) = f[(m+n)e] = \overline{m+n} = \overline{m} + \overline{n} = f(me) + f(ne)$$

$$f[(me) \cdot (ne)] = f[(mn)e] = \overline{mn} = \overline{m} \cdot \overline{n} = f(me) \cdot f(ne)$$

Luego, f es justamente un isomorfismo de S sobre $\mathbb{Z}_p$. El teorema queda demostrado.

Observaciones. 1) Una regla de cálculo en dominios de característica prima p es la siguiente:

$$(a+b)^p = a^p + b^p$$

En efecto, por el teorema del binomio, tenemos:

$$(a+b)^p = a^p + b^p + \sum_{k=1}^{n-1} \binom{p}{k} a^{p-k} b^k$$

donde, $\binom{p}{k} = \frac{p(p-1)\ldots(p-k+1)}{k!}$, $k = 1, 2, \ldots, p-1$

Pero, siendo p primo, p es relativamente primo con $1, 2, 3, \ldots, p-1$; luego, ningún factor de k! es divisor de p, y como p es mayor que k, tampoco p puede dividir a ningún factor de k!; es decir, $(p-1)(p-2)\ldots(p-k+1)$ debe ser divisible por k!, ya que $\binom{p}{k}$ es un entero.

Luego, en el desarrollo de $(a+b)^p$ todo coeficiente, excepto el primero y el último, es múltiplo de p. O sea,

$$p \mid \binom{p}{k} \Rightarrow \binom{p}{k} a^{p-k} b^k = o$$

Por consiguiente:

$$(a+b)^p = a^p + b^p$$

Consecuencias

$$(a_1 + a_2 + \ldots + a_r)^p = a^p{}_1 + a^p{}_2 + \ldots + a^p{}_r$$

También,

$$(a+b)^{p^2} = ((a+b)^p)^p = (a^p + b^p)^p = a^{p\,p} + b^{p\,p} =$$

$$= a^{p^2} + b^{p^2}$$

En general,

$$(a+b)^{p^q} = a^{p^q} + b^{p^q}, \ \forall\, q \in \mathbb{N}$$

También se demuestra que en un dominio de característica prima p es:

$$(a-b)^p = a^p - b^p$$

En efecto, si p es impar, se tiene,

$$(a - b)^p = [a + (-e)\,b]^p = a^p + (-e)^p\, b^p$$

y como $(-e)^p = -e$, resulta,

$$(a - b)^p = a^p + (-e)\,b^p = a^p - b^p$$

Si p es un primo par, debe ser necesariamente $p = 2$, y en tal caso $(-e)^p = (-e)^2 = +e$, pero también $+e = -e$, puesto que $2e = e + e = o$. Por lo tanto, en cualquier caso se tiene,

$$(a - b)^p = a^p - b^p$$

2. El subdominio $S = \{me : e \in D, m \in \mathbb{Z}\}$ del dominio D indicdicado en el teorema anterior lo llamaremos DOMINIO de INTEGRIDAD PRIMO de D.

Puesto que todo subdominio de D con la unidad e debe necesariamente contener sus múltiplos me, esto es, debe contener a S, resulta que S es el *mínimo dominio de integridad contenido en D*, o intersección de todos los subdominios de D.

En virtud de esta consideración y del teorema recién probado, llegamos a la siguiente e importante conclusión final: Todo dominio de integridad D puede considerarse como una extensión (ampliación) de su dominio primo $S = \{me : e \in D, m \in \mathbb{Z}\}$, que es el mínimo dominio de integridad contenido en D, o intersección de todos los subdominios de D. El dominio S sólo puede ser de dos tipos: o bien isomorfo al dominio $\mathbb{Z}$ de los enteros, o bien isomorfo al dominio $\mathbb{Z}_p$ de los restos módulo un número primo p; lo cual lo expresamos diciendo que la característica de D es, respectivamente, cero o p.

En el primer caso, todo elemento no nulo de D es de orden infinito, en el segundo de orden p, en el grupo aditivo de D.

El dominio primo S es el mismo para todo subdominio y toda extensión de D.

3. *Característica de un campo* (cuerpo). Puesto que un campo K (o un cuerpo) es en particular un dominio de integridad en el cual ahora es posible la división (excepto por cero), las anteriores consideraciones sobre la característica se aplican sin más a los campos o cuerpos.

En este caso, el subconjunto $S = \{me : e \in K, m \in \mathbb{Z}\}$ es un subcampo (subcuerpo) de K y que lo llamaremos el CUERPO PRIMO de K. Este cuerpo o campo primo sólo puede ser de dos tipos: si K tiene característica prima

p, entonces, por el teorema anterior, el subgrupo aditivo de K engendrado por su elemento unidad e es un subcampo, y es isomorfo con el campo finito de los restos módulo p. Es decir, todo cuerpo de característica prima p tiene un subcuerpo isomorfo al cuerpo de los enteros módulo p, que se representa por $\mathbb{Z}_p = \{\overline{0}, \overline{1}, \overline{2}, \ldots, \overline{p-1}\}$; si K tiene característica cero, el subcampo engendrado por el elemento unidad e es isomorfo con el campo $\mathbb{Q}$ de los números racionales, siendo la correspondencia de isomorfismo la siguiente,

$$\frac{me}{ne} \longleftrightarrow \frac{m}{n}, \text{con } n \neq o$$

Este isomorfismo conserva las cuatro operaciones racionales de K.

Veremos más tarde, al estudiar el Problema de la Extensión de una estructura, que el subcampo engendrado por e isomorfo a $\mathbb{Q}$ es el campo de cocientes del subdominio de todos los múltiplos me.

Ahora, si convenimos en identificar cada elemento de la forma $\frac{me}{ne}$, $n \neq o$, con el correspondiente número racional $\frac{m}{n}$, entonces todo campo de característica cero viene a contener todos los números racionales $\frac{m}{n}$, con $n \neq o$.

Por un convenio similar, cualquier campo de característica prima p viene a contener al campo $\mathbb{Z}_p$. En este sentido, cualquier campo K es una extensión de uno de los dos campos mínimos $\mathbb{Q}$ y $\mathbb{Z}_p$, intersecciones de todos los subcampos de K, y son los mismos para todo subcampo y toda extensión de K.

Capítulo IX
DOMINIO ORDENADO. CAMPO ORDENADO

9.0. Vamos ahora a considerar el caso de dominios de integridad y de cuerpos en los que está definida una relación de orden.

Definición I. Sea D un dominio de integridad.

Diremos que D es un DOMINIO ORDENADO si en él está definida una relación de orden, de modo que se verifican las siguientes propiedades:

01): Entre dos elementos cualesquiera a, b de D se verifica una y sólo una de las tres relaciones:

$a = b, a < b, b < a$ (Ley de tricotomía)

02): $a < b$ y $b < c \Rightarrow a < c$ (Ley transitiva)

03): Si $a < b$, entonces $a + c < b + c$ (Ley de monotonía de la adición)

04): Si $a < b$ y $c > o$, entonces $ac < bc$ (Ley de monotonía de la multiplicación).

La propiedad 01) establece que D, como conjunto, es un conjunto totalmente ordenado, y las propiedades 03) y 04) vinculan la relación de orden con las operaciones de adición y multiplicación definidas en D.

Luego, cuando el dominio de integridad D cumple el sistema 01), 02), 03) y 04), se dirá que el par $(D; \leq)$ es un *Dominio de Integridad Ordenado*. No confundir esto cuando en D, como conjunto, esté definida una relación de orden. Lo esencial en un dominio de integridad ordenado es la vinculación entre el orden "$\leq$" y las operaciones "$+$" y "$\cdot$" definidas en el dominio, dadas por 03) y 04), respectivamente.

Reduciendo la comparación con respecto al elemento neutro o de la adición, podemos definir los elementos positivos y los elementos negativos de D, como sigue:

Definición. En un dominio de integridad ordenado D, un elemento $a \in D$ es un ELEMENTO POSITIVO si a es mayor que cero y se escribe $a > o$, y un elemento $b \in D$ es un ELEMENTO NEGATIVO si b es menor que cero, y se escribe $b < o$.

De esta definición y de la propiedad 01) de la Definición 1), resulta que dado un elemento cualquiera a de un dominio de integridad ordenado, entonces se verifica una y sólo una de las siguientes relaciones:

$$a = o, o < a, a < o$$

Es decir, a es necesariamente nulo, positivo o negativo. Por lo tanto, D

está dividido en tres subconjuntos disjuntos dos a dos: el de los elementos positivos, el de los elementos negativos y el que tiene por único elemento a o.

Proposición 1. La suma de dos elementos positivos es un elemento positivo.

Dem. Sean $o < a$ y $o < b$ dos elementos positivos.

Por 03) resulta, $o + b < a + b$

o sea, $b < a + b$

De $o < b$ y $b < a + b$, por 02) resulta,

$o < a + b$

lo que demuestra la proposición.

Proposición 2. El producto de dos elementos positivos es un elemento positivo.

Dem. Sean $o < a$ y $o < b$ dos elementos positivos.

Aplicando 04) se tiene,

$o \cdot b < a \cdot b$

o sea, $o < a \cdot b$

lo que prueba la proposición.

Otro modo de definir un orden en un dominio de integridad es la siguiente, y también de mucho uso:

Definición II. Sea D un dominio de integridad.

Diremos que D es un DOMINIO ORDENADO, si hemos distinguido un subconjunto $D^+ \subset D$ (los elementos de D^+ se llaman *positivos),* verificándose las tres leyes siguientes:

D01): $\exists\ D^+ \subset D$

D02): Para $a \in D$ arbitrario se verifica siempre una y una sola de las tres posibilidades:

$a = o, a \in D^+, (-a) \in D^+$

D03): $a, b \in D^+ \Rightarrow a + b \in D^+$ y $ab \in D^+$

Los elemento *a* tales que $(-a) \in D^+$, se llaman *negativos* y constituyen el subconjunto $D^- \subset D$.

Nótese que las Proposiciones 1) y 2) deducidas de la definición anterior I), en la nueva definición II) se dan como axiomas. (Ver D03).

Probaremos ahora que los sistemas de axiomas dados en las definiciones I) y II) son equivalentes. Por lo tanto, escoger uno u otro es sólo una cuestión de convenio.

Demostraremos que I $\Leftrightarrow$ II. Tenemos:

En efecto, admitamos que 01), 02), 03) y 04) se verifica; y definamos D^+ como el conjunto de aquellos elementos $a \in D$ tales que,

$$\boxed{(*)\quad o < a}$$

Luego, se cumple D01).

Por otra parte, por 01), para cada $a \in D$ se verifica:

$a = o$, ó $o < a$, ó $a < o$

Pero, por 03), $a < o \Rightarrow a + (-a) < o + (-a) \Rightarrow o < -a$

Así pues, se cumple D02).

Finalmente, sean $a, b \in D^+$, esto es,

$o < a, o < b$

Por 03), se tiene:

$o < a \Rightarrow o + b < a + b \Rightarrow b < a + b$

y como $o < b$, entonces, por 02), es $o < a + b \Rightarrow a + b \in D^+$

De $o < a$, $o < b$, resulta por 04), $o \cdot b < a\,b \Rightarrow o < ab \Rightarrow ab \in D^+$; luego, se cumple D03).

Así hemos probado que I $\Rightarrow$ II

Recíprocamente, admitamos ahora que el sistema D01), D02) y D03) se verifica; y definamos $a < b$ como equivalente de $b - a \in D^+$.

$$(**)\boxed{a < b \text{ significa } b - a \in D^+}$$

y probemos que se cumple el sistema 01), 02), 03) y 04).

Tendremos:

Dados $a, b \in D$, por D02), se verifica uno y sólo uno de los casos:

$a - b = o, a - b \in D^+, -(a - b) = b - a \in D^+$

o sea, $a = b, b < a, a < b$

Luego, se cumple 01).

Si $a < b$ y $b < c$, entonces por la definición (**), resulta que,

$b - a \in D^+$ y $c - b \in D^+$

y por D03), resulta:

$(b - a) + (c - b) = c - a \in D^+ \Rightarrow a < c$

Así, pues, se cumple 02).

De $a < b$, por la definición (**), se tiene,

$b - a = (b + c) - (a + c) \in D^+ \Rightarrow a + c < b + c$

y se cumple 03).

De $a < b \Leftrightarrow b - a \in D^+$ y $c \in D^+$, por D04), resulta,

$(b - a)\,c \in D^+$

o sea, $b\,c - a\,c \in D^+ \Rightarrow a\,c < b\,c$

y se cumple 04).

Así hemos probado que II $\Rightarrow$ I.

Este resultado y el anterior, demuestra la equivalencia I $\Longleftrightarrow$ II de los dos sistemas de axiomas dados en las definiciones I) y II).

En lo sucesivo emplearemos de preferencia el sistema dado en la Definición I), por estar más habituado al uso "$\leq$" de su notación de orden.

Evidentemente, D02) implica que si $a \in D^+$, entonces $(-a) \in D^+$ y recíprocamente.

Luego, $a \in D^+$ significa $o < a$, ó $(-a) < o$
y $a \in D^-$ significa $a < o$, u $o < -a$.

Finalmente, como de costumbre, se definen las notaciones:

$a \leq b$ como $a \not> b$
$a < b$ como $b > a$
$a \geq b$ como $a \not< b$
$a > b$ como $b < a$

Por otra parte, por 01),

$a \leq b$ significa que debe ser $a < b$, ó $a = b$
$a \geq b$ significa que debe ser $a > b$, ó $a = b$

Sabemos que un campo (cuerpo conmutativo) es también un dominio de integridad. Por tanto, podemos definir un campo ordenado.

9.1 *Definición.*

Un campo F se llama *ordenado,* si F es un dominio de integridad ordenado.

En otras palabras, diremos que un campo F está ordenado, si contiene un subconjunto F^+ de elementos "positivos" con las propiedades aditiva, multiplicativa y tricotomía expuestas en la Definición II), o las de su equivalente dadas en la Definición I).

Pasaremos en seguida a demostrar algunas propiedades de los campos ordenados (que valen también en algunos casos para dominios de integridad ordenados) que se deducen de la definición.

Estas propiedades son, en el sistema 01), 02), 03) y 04):

Teorema 1. En un campo ordenado, un elemento *a* es menor que otro *b* sí, y sólo si, la diferencia $b - a$ es positiva; esto es:

$$a < b \iff b - a > o$$

Dem. a) Si $a < b$, entonces sumando a ambos miembros $(-a)$, por 03), resulta,

$$a + (-a) < b + (-a)$$

o sea, $o < b - a$

b) Si $b - a > o$, entonces sumando a ambos miembros a, resulta, por 03)

$$o + a < (b - a) + a$$

o sea, $a < b$

Este resultado y el anterior prueban la equivalencia,

$$a < b \iff b - a > o$$

Teorema 2. En un campo ordenado, un elemento es positivo sí, y sólo si, su simétrico es negativo; esto es:

$$a > o \iff (-a) < o$$

Dem. a) Si $a > o$, o lo que es lo mismo $o < a$, entonces sumando a ambos miembros $(-a)$, resulta por 03):

$$o + (-a) < a + (-a)$$

o sea, $(-a) < o$

b) Si $(-a) < o$, entonces sumando *a* a ambos miembros, por 03, se obtiene:

$$(-a) + a < o + a$$

o sea, $o < a$

Este resultado y el anterior prueban la equivalencia:

$$a > o \iff (-a) < o$$

Teorema 3. En un campo ordenado, un elemento es negativo sí, y sólo si, su simétrico (u opuesto) es positivo.

Dem. Demostración análoga a la del Teorema 2). O bien, por el Teorema 2) sabemos que,

$$a > o \iff (-a) < o$$

y substituyendo *a* por $(-a)$, resulta:

$$(-a) > o \iff -(-a) < o$$

o sea, $(-a) > o \iff a < o$

y como la equivalencia es simétrica, obtenemos

$$a < o \iff (-a) > o$$

resultado que prueba el teorema.

Teorema 4. En un campo ordenado, el cuadrado de cualquier elemento diferente de cero es positivo.

Dem. Sea $a \neq o$; luego es $o < a$, ó $a < o$

Si $o < a$, entonces multiplicando ambos miembros por *a*, resulta por 04):

$$o \cdot a < a \cdot a$$

o sea, $o < a^2$

Ahora, si $a < o$, entonces por el Teorema 3) es $(-a) > o$.

Por tanto, según lo recién demostrado, resulta:

$$o < (-a)^2$$

Pero, $(-a)^2 = (-a) \cdot (-a) = a^2$.

Luego, $o < a^2$

Así, pues, en todos los casos, el cuadrado de cualquier elemento no nulo es positivo.

Corolario. Se tiene, $o < 1$; es decir, el elemento unidad del campo es positivo.

Dem. Sabemos que $1 \neq o$; luego, por el Teorema 4) es $1^2 > o$, o sea, $1 > o$.

Teorema 5. En un campo ordenado, si se suman miembro a miembro dos

desigualdades del mismo sentido o una igualdad y una desigualdad, se obtiene otra desigualdad del mismo sentido. Esto es:

1) $a < b$ y $c < d \Rightarrow a + c < b + d$
2) $a \leq b$ y $c < d \Rightarrow a + c < b + d$
3) $a < b$ y $c \leq d \Rightarrow a + c < b + d$
4) $a \leq b$ y $c \leq d \Rightarrow a + c \leq b + d$

Dem. Haga (1), (3) y (4) como ejercicio. Solamente demostraremos la (2).

Supongamos que $a \leq b$, entonces $o \leq b - a$

Si, por otra parte, se tiene $c < d$, entonces

$o < d - c$

sumando a ambos miembros de esta última desigualdad $b - a$, se tiene por 03):

$o + (b - a) < (d - c) + (b - a)$

o sea, $b - a < (b + d) - (a + c)$

y como $o \leq b - a$

entonces, por 02), se obtiene,

$o < (b + d) - (a + c)$

y lo que es equivalente a:

$a + c < b + d$

Teorema 6. En un campo ordenado, ambos miembros de una desigualdad pueden ser multiplicados por un mismo elemento, dando lugar a otra desigualdad de igual sentido si el multiplicador es positivo, y de distinto sentido si el multiplicador es negativo; esto es

(1) $a < b$ y $o < c \Rightarrow a \cdot c < b c$
(2) $a < b$ y $c < o \Rightarrow a \cdot c > b c$

Dem. La (1) es evidente por el axioma 04). Por esto, sólo demostraremos (2).

Si $a < o$, entonces es $o < b - a$.

Si $c < o$, entonces por Teorema 3), es $(-c) > o$

Luego, aplicando 04) se tiene:

$o \cdot (-c) < (b - a) \cdot (-c)$

o sea, $o < ac - b c$

lo que equivale a:

$b c < a c$,

y el teorema queda demostrado.

Este teorema tiene en particular como consecuencia inmediata, la siguiente propiedad:

En un campo ordenado, el producto de dos elementos negativos es un elemento positivo y el producto de un elemento positivo por un elemento negativo es un elemento negativo.

En efecto, si $a < o$ y $b < o$, entonces por el Teorema 6) es $o \cdot b < a \cdot b$, o sea, $o < ab$.

Ahora, si $o < a$ y $b < o$, entonces por el mismo Teorema 6), resulta $a \cdot b < o \cdot b$, es decir, $ab < o$.

Teorema 7. En un campo ordenado, ambos miembros de una desigualdad pueden ser divididos por un mismo elemento, dando lugar a otra desigualdad de igual sentido si el divisor es positivo, y de distinto sentido si el divisor es negativo; esto es:

(1) Si $ac < bc$ y $c > o$, entonces es $a < b$
(2) Si $a c < b c$ y $c < o$, entonces es $b < a$

Dem. a) Sea $a c < b c$, entonces es $o < b c - a c$, es decir,

$o < (b - a) c$

Si $c > o$, entonces $(b - a)$ no puede ser nulo ni negativo, ya que $(b - a) \cdot c$ es positivo.

Luego, debe ser necesariamente $b - a > o$, lo que equivale a, $a < b$.

b) Supongamos como antes $a c < b c$, esto es $o < (b - a) \cdot c$

Si $c < o$, entonces $(b - a)$ no puede ser nulo ni positivo, ya que $(b - a) \cdot c$ es positivo.

Luego, $b - a$ es necesariamente negativo, es decir,

$b - a < o$

y lo que equivale a, $b < a$

Este resultado y el anterior, demuestran el teorema.

Teorema 8. En un campo ordenado, un elemento es positivo sí, y sólo si, su inverso es positivo; esto es

$a > o \iff a^{-1} > o$

Dem. Si $a > o$, como $a \cdot a^{-1} = 1$ y $1 > o$, debe ser necesariamente $a^{-1} > o$, pues si fuera negativo o nulo, entonces el producto $a \cdot a^{-1}$ también sería negativo o nulo, lo cual es imposible. Luego, $a^{-1} > o$.

Recíprocamente, si $a^{-1} > o$ entonces por un razonamiento análogo al anterior, resulta ser $a > o$.

Este resultado y el anterior prueban la equivalencia:

$a > o \iff a^{-1} > o$

Teorema 9. En un campo ordenado, un elemento es negativo sí, y sólo si, su inverso es negativo.

Dem. La demostración es análoga a la del teorema anterior.
¡Hágala como ejercicio!

Teorema 10. En un campo ordenado, se verifica:

$$a - 1 < a < a + 1$$

cualquiera sea *a* en el campo.

Dem. Sabemos que, $o < 1$ (Por corolario del Teorema 4)). Sumando a ambos miembros *a* por 03), resulta:

$$a + o < a + 1$$

o sea, $a < a + 1$

cualquiera que sea *a* en el campo; luego, en particular para $(a - 1)$, se tiene:

$$a - 1 < (a - 1) + 1$$

o sea, $a - 1 < a$

De este resultado y el anterior concluimos por 02) que,

$$a - 1 < a < a + 1$$

y lo que demuestra el teorema.

9.2. Propiedad arquimediana

Un dominio de integridad se denominará ARQUIMEDIANO (del nombre del ilustre geómetra griego Arquímides, del siglo III A. C., a quien se atribuye el haber hecho notar por primera vez la importancia de la propiedad análoga en Geometría), si dados dos elementos *a* y *b* positivos cualesquiera existe siempre un número natural n tal que el n − ésimo múltiplo natural de uno de ellos supera o iguala al otro; esto es, por ejemplo, $n\,a \geq b$, donde $n\,a = a + a + \ldots + a$ (n sumandos).

Así por ejemplo, el dominio ordenado $\mathbb{Z}$ de los enteros racionales es arquimediano.

En efecto, sean *a* y *b* dos enteros positivos tales que $a < b$.

Tendremos que probar que existe un entero natural n tal que: $na \geq b$.

Si $a = +1$, bastará tomar $n = b + 1$, ya que:

$$b + 1 > b$$

luego, $1 \cdot (b + 1) > 1 \cdot b = b$

o sea, $a \cdot n > b$

Pero si $a \neq +1$, entonces basta tomar $n = b$, ya que:

$$a > 1 \Rightarrow a \cdot b > 1 \cdot b = b$$

luego, nuevamente, tenemos $a \cdot n > b$.

En cambio, el dominio de integridad de los polinomios en una indeterminada x, no es arquimediano.

En efecto, sean los dos polinomios positivos x^2 y x. Es claro que

el polinomio: $x^2 - n\,x$, o bien el trinomio $x^2 - n\,x + 0$ es positivo para todo *x* mayor que la raíz $x = n$. Luego,

$$x^2 - n\,x > 0$$

de donde, $n\,x < x^2$ y la propiedad arquimediana no se cumple.

El campo ordenado Q de los números racionales: $\frac{a}{b}$, $b \geq 0$, con respecto al orden natural "<" es arquimediano; esto es, dados dos racionales positivos $\frac{a}{b}$ y $\frac{c}{d}$, existe un número natural n tal que:

$$n \cdot \frac{a}{b} > \frac{c}{d}$$

En efecto, sean $\frac{a}{b} > 0$ y $\frac{c}{d} > 0$, de modo que los cuatro enteros, a, b, c y d son positivos. Por consiguiente, tendremos:

$$b\,c < b\,c + b\,c = 2\,b\,c$$

Por otra parte, se tiene:

$$ad \geq 1$$

de donde, $\frac{1}{ad} \leq 1$

Tendremos entonces:

$$\frac{c}{d} = \frac{a}{b} \cdot \frac{bc}{ad} \leq \frac{a}{b} \cdot b\,c < \frac{a}{b} \cdot 2\,bc$$

es decir, que la propiedad arquimediana se verifica con el número natural $n = 2\,bc$.

9.3. Cuerpo Ordenado Completo

Diremos que un cuerpo ordenado F es un CUERPO ORDENADO COMPLETO, si todo subconjunto de F, no vacío y acotado superiormente tiene extremo superior o supremo en F.

Es decir, si F es un cuerpo ordenado completo y S es un subconjunto ordenado cualquiera de F, no vacío y acotado superiormente, entonces el conjunto C de todas las cotas superiores de S tiene primer elemento.

Teorema. En un cuerpo ordenado completo F, todo subconjunto no vacío y acotado inferiormente tiene extremo inferior o ínfimo en F.

Dem. Sea S un subconjunto no vacío y acotado inferiormente de F.

Consideremos el conjunto C de todas las cotas inferiores de S. Como S es acotado inferiormente, entonces $C \neq \varnothing$ y además este conjunto es acotado superiormente, ya que cada uno de los elementos de S es cota superior de C.

Luego, como F es cuerpo ordenado completo, entonces el subcon-

junto no vacío C tiene extremo superior o supremo en F. Sea z este supremo.

Probaremos que z es el extremo inferior o el ínfimo de S.

En efecto, tendremos:

$$z \leq x, \forall\, x \in S$$

por ser z el primer elemento del conjunto de todas las cotas superiores de C. Además, si z' es un elemento de F tal que:

$$z' \leq x, \forall\, x \in S$$

entonces $z' \in C$ y, por lo tanto, $z' \leq z$; es decir, z es la mayor de las cotas inferiores de S; luego, S tiene extremo inferior o ínfimo en F. Queda así probado que S tiene extremo inferior.

Análogamente se demostraría que si en un cuerpo ordenado F todo subconjunto no vacío y acotado inferiormente tiene extremo inferior o ínfimo en F, entonces todo subconjunto no vacío y acotado superiormente, tiene extremo superior o supremo en F.

Luego, ambas propiedades son equivalentes.

Por consiguiente, en un cuerpo ordenado completo F cualquier subconjunto no vacío y acotado tiene ínfimo y supremo en F.

Además, se demuestra (nosotros no lo haremos) que dos cuerpos ordenados completos son necesariamente isomorfos; es decir, salvo isomorfismos, *existe un único cuerpo ordenado completo*. Por lo tanto, se tiene así una caracterización abstracta de un sistema muy importante en Matemática: *el Sistema de los Números Reales*.

Axioma de los números reales: El sistema de los números reales es un cuerpo ordenado completo.

Es decir, postulamos que el sistema de los números reales es un conjunto abstracto en el cual están definidas dos operaciones binarias internas, llamadas "adición" y "multiplicación" y una relación de orden en forma tal que se verifican las propiedades de cuerpo ordenado completo.

Por consiguiente, a la pregunta: ¿qué es un número real?, podemos contestar: "Un número real es un elemento de un cuerpo ordenado completo".

Al sistema de los números reales se lo designa habitualmente con la notación $\mathbb{R}$

Todo lo que hemos estudiado y visto sobre cuerpos, cuerpos ordenados y cuerpos ordenados completos, se aplica también al sistema $\mathbb{R}$ de los números reales.

9.4. *Aplicaciones al Campo Ordenado de los Números Reales.* Consideremos, en particular, el campo $\mathbb{R}$ de los números reales ordenado por la relación "$\leq$".

Desde luego, en este campo ordenado valen todos los teoremas que hemos venido estudiando en el presente capítulo. Además, veremos algunos otros que nos serán útiles.

Teorema. Se pueden restar ordenadamente dos desigualdades que se verifican en sentido contrario, resultando otra desigualdad que se verifica en el mismo sentido de la que sirvió de minuendo.

Dem. Sean las dos desigualdades:

$$a < b$$

$$y\ c > d$$

Luego, $b - a > 0$ y $c - d > 0$.

Como la suma de positivos es positivo, se tiene,

$$(b - a) + (c - d) > 0$$

o bien, $(b - d) - (a - c) > 0$

lo que es equivalente a:

$$a - c < b - d$$

y el teorema está probado.

Teorema. Se pueden multiplicar ordenadamente varias desigualdades que se verifican en un mismo sentido, si los dos miembros de cada una son positivos, permaneciendo la nueva desigualdad en el mismo sentido.

Dem. Sean las desigualdades, cuyos miembros son positivos,

$$a > b, c > d, e > f, \ldots$$

Entonces, el producto a c e ... de los primeros miembros será positivo, lo mismo que el producto b d f ... de los segundos miembros; además, como los factores del primer producto son respectivamente mayores que los del segundo, se tendrá evidentemente,

$$ace\ldots > bdf\ldots$$

Corolario. Se puede elevar una desigualdad a una potencia de exponente entero y positivo, siempre que los dos miembros sean positivos.

En efecto, por el teorema recién probado, de la desigualdad $a < b$ se deduce que, siendo n un número entero y positivo y positivas las cantidades *a* y *b*, entonces es $a^n < b^n$.

Teorema. Toda desigualdad se puede elevar a una potencia de grado impar.

Dem. Sea la desigualdad $a < b$.

Si *a* y *b* son positivas, ya hemos demostrado que,

$a^{2n+1} < b^{2n+1}$

Si $a < o$ y $b > o$, entonces $a^{2n+1} < o$ y $b^{2n+1} > o$

Luego se tendrá evidentemente, $a^{2n+1} < b^{2n+1}$.

Si $a < o$ y $b < o$, entonces las potencias a^{2n+1} y b^{2n+1} serán también negativas; pero siendo $a < b$, el valor numérico de *a* será menor que el de *b*; luego, el valor numérico de a^{2n+1} será también menor que el de b^{2n+1}, y por consiguiente será,

$$a^{2n+1} < b^{2n+1}$$

con lo cual queda demostrado el teorema.

Teorema. Se puede extraer una raíz de índice o grado impar de los dos miembros de una desigualdad.

Dem. Sea la desigualdad $a < b$.

Si *a* y *b* son positivos, sus raíces de grado impar también lo serán, y además la primera será menor que la segunda.

Si $a < o$ y $b > o$, sus raíces de grado impar serán, la primera negativa y la segunda positiva, y por consiguiente, la primera menor que la segunda.

Por último, si $a < o$ y $b > o$ negativas, serán también sus raíces de grado impar, menor en valor numérico la primera que la segunda.

Luego, si se tiene $a < b$, se tendrá siempre cualesquiera que sean los signos de *a* y *b*, la desigualdad:

$$\sqrt[2n+1]{a} < \sqrt[2n+1]{b}$$

Teorema. Se puede extraer una raíz de grado par de los dos miembros de una desigualdad, siempre que se tomen para estas raíces cantidades positivas.

Dem. Sea la desigualdad $a^{2n} < b^{2n}$.

Sabemos que la raíz de grado 2n de una catidad a^{2n} es, $\pm$ a. Por consiguiente, las raíces de grado 2n de los dos miembros de la desigualdad propuesta, serán $\pm$ a y $\pm$ b.

Ahora bien, como a^{2n} y b^{2n} son cantidades positivas, ya sean positivas o negativas las cantidades *a* y *b*, el valor numérico de *a* tiene que ser menor que el de *b*, puesto que se tiene $a^{2n} < b^{2n}$; luego, si tomamos el signo de estas raíces de modo que las cantidades $\pm$ a y $\pm$ b sean positivas, se verificará el enunciado.

Así, pues, si *a* y *b* son positivas, se deducirá de la desigualdad $a^{2n} < b^{2n}$, que $a < b$.

Si *a* y *b* son negativas, se deducirá entonces que $-a < -b$.

Si $a > o$ y $b < o$, se tendrá $a < -b$.

Por último, si $a < o$ y $b > o$, será $-a < b$

Luego, en general, de $a^{2n} < b^{2n}$ se deduce $|a| < |b|$.

(El significado del símbolo $|a|$ se verá más adelante).

Finalizaremos este párrafo viendo el teorema que sigue y sumamente útil en el estudio de desigualdades:

Teorema. Si tenemos una serie de fracciones $\frac{a_1}{b_1}, \frac{a_2}{b_2}, \frac{a_3}{b_3}, \ldots, \frac{a_n}{b_n}$, siendo cualesquiera sus numeradores, pero sus denominadores positivos, y formamos la fracción:

$$\frac{a_1 + a_2 + \ldots + a_n}{b_1 + b_2 + \ldots + b_n}$$

cuyo numerador sea la suma de los numeradores de las propuestas y el denominador la suma de sus denominadores, entonces esta nueva fracción estará comprendida entre la menor y la mayor de dichas fracciones.

Dem. Supongamos las fracciones dadas dispuestas por orden, de manera que, $\frac{a_1}{b_1}$ sea la menor y $\frac{a_n}{b_n}$ la mayor. Evidentemente tendremos:

$$a_1 = b_1 \cdot \frac{a_1}{b_1}, a_2 > b_2 \cdot \frac{a_1}{b_1}, a_3 > b_3 \cdot \frac{a_1}{b_1}, \ldots, a_n > b_n \cdot \frac{a_1}{b_1}$$

Ahora, si sumamos todas estas desigualdades, miembro a miembro, resultará:

$$a_1 + a_2 + \ldots + a_n > (b_1 + b_2 + \ldots + b_n) \cdot \frac{a_1}{b_1}$$

y de donde,

$$\frac{a_1 + a_2 + \ldots + a_n}{b_1 + b_2 + \ldots + b_n} > \frac{a_1}{b_1}$$

Partiendo, por el contrario, de la igualdad $a_n = b_n \cdot \frac{a_n}{b_n}$ hallaríamos, $a_{n-1} < b_{n-1} \cdot \frac{a_n}{b_n}$, etc., y así llegaríamos a deducir que,

$$\frac{a_1 + a_2 + \ldots + a_n}{b_1 + b_2 + \ldots + b_n} < \frac{a_n}{b_n}$$

Este resultado y el anterior demuestran el teorema; esto es:

$$\frac{a_1}{b_1} < \frac{a_1 + a_2 + \ldots + a_n}{b_1 + b_2 + \ldots + b_n} < \frac{a_n}{b_n}$$

9.5. *Inecuaciones*

Llamaremos INECUACION a una desigualdad que contiene una o varias cantidades indeterminadas o incógnitas, y que hay que determinarlas con la condición de que dicha desigualdad se verifique.

Esto se consigue mediante el uso de los teoremas vistos en campos ordenados y también los vistos en el campo ordenado de los números reales. Estos teoremas permiten hacer transformaciones análogas a las que se han hecho con las ecuaciones, y que dan por resultado, no el valor de la incógnita o cantidad indeterminada, como en las ecuaciones se obtiene, sino un límite, ya sea inferior, ya sea superior, de los valores que la incógnita puede recibir, y a veces dos límites que comprenden un número reducido de valores, precisando cuanto es posible la cuestión.

Luego, en general, con las inecuaciones no se pueden hacer todas las transformaciones que se hacen con las ecuaciones; ya que las primeras están sujetas a ciertas restricciones dadas por los teoremas recientemente estudiados y que son necesarios tener muy presentes para no incurrir en graves errores.

Las inecuaciones, lo mismo que las ecuaciones, pueden ser de primero, de segundo, de tercero, etc. grado.

Nosotros sólo nos preocuparemos preferentemente de las inecuaciones de primer y de segundo grado con una sola incógnita.

9.6. *Inecuaciones de Primer Grado con una incógnita*

A toda inecuación de primer grado con una incógnita, siempre podremos quitarle los denominadores, si los tiene, fundándonos en el Teorema (6); transponer un término de un miembro a otro con signo contrario al que tenía, fundándonos en el axioma 03), y reducirla de este modo a la forma:

$$a\,x > b \text{ ó } a\,x < b$$

en la cual *a* es una cantidad entera y positiva, y *b* una cantidad entera, positiva o negativa.

Para ello, como lo hemos hecho presente, no hay más que quitar los denominadores si los hay, multiplicando los términos por el mínimo común múltiplo de todos los denominadores, dejando la inecuación en el mismo sentido o sentido contrario, según que el mínimo común múltiplo sea positivo o negativo.

Después se transponen todos los términos que contiene la incógnita al primer miembro y los que no la contienen, al segundo miembro. En seguida se efectúa la reducción de los términos semejantes si la inecuación es numérica, o se saca factor común a la incógnita si es algebraica o literal. De esta manera la inecuación dada queda reducida a una de las dos formas indicadas al principio; esto es:

$$a\,x > b, \text{ ó } a\,x < b$$

De donde resulta:

$$x > \frac{a}{b}, \text{ ó } x < \frac{b}{a}$$

En el primer caso, la cantidad $\frac{b}{a}$ es un *Límite Inferior* de *x*, es decir que, para verificar la desigualdad propuesta a $x > b$, podrán darse a *x* todos los valores imaginables, mayores que $\frac{b}{a}$.

En el segundo caso, la cantidad $\frac{b}{a}$ es un *Límite Superior* de *x*, de suerte que, pueden darse a esta incógnita todos los valores imaginables menores que este límite.

De modo pues, en las inecuaciones de primer grado con una incógnita, puede recibir ésta, según hemos visto, una infinidad de valores, ya que sólo determinamos un límite inferior o superior de los que dicha incógnita puede recibir. En algunos casos no se excluye la igualdad, y entonces *x* no sólo puede recibir valores mayores o menores, según el caso, que $\frac{b}{a}$, sino que también puede recibir este valor.

Ejemplos:

a) Resolver la inecuación:

$$3x - \frac{1}{4} > 20 - \frac{2x}{3}$$

Multipliquemos esta inecuación por 12 = M.C.M. (3,4)

$$36x - 3 > 240 - 8x$$

Transponiendo términos semejantes, se obtiene:

$$36x + 8x > 240 + 3$$

$$44x > 243$$

$$x > \frac{243}{44} = 5\frac{23}{44} \text{ (L. I.)}$$

Si solamente se exigen valores enteros, se tendrá:

$x = 6, 7, 8, 9, \ldots$

b) Resolver la inecuación:

$$\frac{3x + 8}{4} > 2x - 5$$

Tendremos:

$$3x + 8 > 8x - 20$$

$$3x - 8x > -20 - 8$$

$$-5x > -28$$

$$x < \frac{-28}{-5} = \frac{28}{5} = 5\frac{3}{5} \text{ (L. S.)}$$

Si sólo interesan valores enteros, se tendrá

$x = 5, 4, 3, 2, 1, 0, -1, -2, \ldots$

*****

Cuando la incógnita x, esté sujeta a satisfacer a dos inecuaciones diferentes, podrán dársele una infinidad de valores, si dichas inecuaciones determinan a la vez dos límites inferiores, o dos límites superiores de esta incógnita; sólo que deberemos partir del mayor límite inferior o del menor límite superior. Pero, si deducimos de las dos inecuaciones un límite inferior y un límite superior de x, sólo podremos dar a esta incógnita valores comprendidos entre dichos límites, de suerte que, si el límite inferior (L. I.) hallado no es menor que el límite superior (L. S.) encontrado, será imposible satisfacer simultáneamente a las inecuaciones propuestas.

Ejemplos

a) Encontrar los valores de x, positivos o negativos, pero enteros, que verifican simultáneamente las inecuaciones siguientes:

$$6x + \frac{5}{7} > 4x + 7, \quad \frac{8x + 3}{2} < 2x + 25$$

Resolviendo separadamente cada inecuación, se obtiene:

$$42x + 5 > 28x + 49, \quad 8x + 3 < 4x + 50$$

$$42x - 28x > 49 - 5 \qquad 4x < 47$$

$$14x > 44 \qquad x < \frac{47}{4} = 11\frac{3}{4} \quad \text{(L.S.)}$$

$$x > \frac{44}{14} = 3\frac{1}{7} \text{ (L. I.)}$$

Como el límite inferior $3\frac{1}{7}$ es menor que el límite superior $11\frac{3}{4}$, entonces es posible satisfacer a la vez a las dos inecuaciones dadas. Ahora, como sólo se piden valores enteros para x, éstos serán todos los comprendidos en la limitación:

$$3\frac{1}{7} < x < 11\frac{3}{4}$$

o sea, los números: 4, 5, 6, 7, 8, 9, 10, 11.

b) Hallar los valores de x que satisfacen simultáneamente a las siguientes inecuaciones:

$$8x - 5 > \frac{15x - 8}{2},$$

$$2(2x - 3) > 5x - \frac{3}{4}$$

Tenemos:

$$16x - 10 > 15x - 8, \quad 4x - 6 > 5x - \frac{3}{4}$$

$$x > 2 \text{ (L. I.)} \qquad 16x - 24 > 20x - 3$$

$$-4x > 21$$

$$x < \frac{21}{-4} = -5\frac{1}{4} \text{ (L.S.)}$$

Como aquí el límite inferior 2 no es menor que el superior $-5\frac{1}{4}$, será imposible satisfacer a la vez a las dos inecuaciones propuestas; es decir, el problema no tiene solución.

Si se tuviesen varias inecuaciones con una incógnita

$$ax > b, a'x > b', a''x > b'', a'''x > b''', \ldots$$

se hallará para límites respectivos

$$x \gtrless \frac{b}{a}, \quad x \gtrless \frac{b'}{a'}, \quad x \gtrless \frac{b''}{a''}, \quad x \gtrless \frac{b'''}{a'''}$$

según sean los signos de los coeficientes de la x, los cuales estarían todos en un mismo sentido, o unos en un sentido y otros en sentido contrario.

En el primer caso, dando a la x valores mayores que el mayor de los límites, o menores que el menor de los límites, según que sean inferiores o superiores, se tendrán todos los valores de x que satisfacen a las inecuaciones propuestas.

En el segundo caso, se darán a la x valores comprendidos entre los límites inferiores y superiores más próximos; y las inecuaciones serán incompatibles si dichos dos límites fueran contradictorios.

9.7. *Inecuaciones de primer grado con varias incóginitas*

En primer lugar, consideraremos un caso muy especial, cual, una ecuación y una inecuación con dos incógnitas. Sea el sistema:

$$\left.\begin{aligned} x - 2y &= 4 \\ 2x + 5y &> 45 \end{aligned}\right|$$

De la ecuación, se tiene $x = 4 + 2y$

Este valor lo sustituimos en la inecuación, dando sucesivamente:

$$2(4 + 2y) + 5y > 45$$

$$8 + 4y + 5y > 45$$

$$9y > 37$$

$$y > \frac{37}{9} = 4\frac{1}{9}$$

Luego, la incógnita y podrá tomar todos los valores después de $4\frac{1}{9}$ hasta el infinito. A cada uno de estos valores de y, la relación $x = 4 + 2y$ dará el valor correspondiente de la incógnita x. Este valor de x con el indicado para y, constituirá una solución del sistema propuesto. Luego, este sistema tiene infinitas soluciones.

Pasaremos en seguida a considerar el caso más general, cual, un sistema de inecuaciones con dos incógnitas. Sea el sistema:

$$\left.\begin{array}{r} ax + by > c \\ a'x + b'y > c' \end{array}\right|$$

De ellas se deduce:

$$x \gtrless \frac{c - by}{a}, \text{ según que } a \gtrless o$$

$$x \lessgtr \frac{c' - b'y}{a'}, \text{ según que } a' \gtrless o$$

Si $a > o$ y $a' > o$, entonces resultan dos límites inferiores de x, y podremos por lo tanto, dar a la incógnita y un valor arbitrario β, y dar al mismo tiempo a la x todos los valores mayores que la mayor de las cantidades:

$$\frac{c - b\beta}{a} \quad y \quad \frac{c' - b'\beta}{a'}$$

Si $a < o$ y $a' < o$, entonces también podrá darse a la y un valor arbitrario β, recibiendo la x al mismo tiempo todos los valores inferiores a la menor de las cantidades:

$$\frac{c - b\beta}{a} \text{ y } \frac{c' - b'\beta}{a'}$$

Si a y a' son de signos contrarios, y si por ejemplo, $a > o$ y $a' < o$, tendremos:

$$x > \frac{c - by}{a}, x < \frac{c' - b'y}{a'}$$

necesitándose para esto que sea:

$$\frac{c - by}{a} < \frac{c' - b'y}{a'}$$

lo cual determina un límite de la incógnita y.

Entonces, daremos a y todos los valores mayores que este límite, caso de que este límite sea inferior; o menores que él, si fuese un límite superior.

Ahora, relacionando con cada valor de y, todos los valores de x, comprendidos entre los dos límites correspondientes de esta incógnita, tendremos soluciones del sistema propuesto.

Veamos un ejemplo numérico. Sea el sistema:

$$\left.\begin{array}{r} 2x + 5y > 15 \\ 2x - y > -4 \end{array}\right|$$

Resolviendo cada una de estas inecuaciones con respecto a y, se tiene:

$$y > \frac{15 - 3x}{5}, \quad y < 2x + 4$$

De donde se concluye la inecuación:

$$\frac{15 - 3x}{5} < 2x + 4$$

$$15 - 3x < 10x + 20$$

$$-13x < 5$$

$$x > -\frac{5}{13}$$

Se puede atribuir a x todos los valores más grandes que $-\frac{5}{13}$, y a cada valor de x corresponder dos límites de valores de y.

Sea, por ejemplo, $x = o$, entonces $y > \frac{15}{5} = 3$ e $y < 4$

Para $x = 1$, entonces $y > \frac{12}{5} = 2\frac{2}{5}$ e $y < 6$; etc.

El sistema propuesto es, pues, verificado por una infinidad de valores de x y de y.

Observaciones. 1) Nótese que las desigualdades del sistema dado son del mismo sentido; por lo tanto pueden ser sumadas miembro a miembro. Luego, multiplicando por 5 la segunda y sumándolas, se obtiene:

$$13x > -5$$

$$x > -\frac{5}{13}$$

quedando eliminada la y.

Pero este procedimiento no permite eliminar la x, porque después de haber multiplicado la primera inecuación por 2 y la segunda por 3, no se puede restar miembro a miembro dos desigualdades del mismo sentido.

2) Las mismas consideraciones conducirían a la determinación de los límites de las dos incógnitas x e y, en un sistema de cualquier número de inecuaciones con dos incógnitas. Sea, por ejemplo, el sistema:

$$\left.\begin{array}{l} 2y - x > o \\ 1 - 2x - 3y > o \\ 7 + 4x + y > o \end{array}\right|$$

que queremos satisfacer por medio de valores enteros de x y de y.

Resolviéndolas con la relación a x, deduciremos:

$$x < 2y, x < \frac{1 - 3y}{2}, x > -\frac{7 + y}{4}$$

es necesario, pues, que tengamos simultáneamente:

$$\left.\begin{array}{l} -\frac{7 + y}{4} < 2y \\ -\frac{7 + y}{4} < \frac{1 - 3y}{2} \end{array}\right|$$

De aquí deducimos:

$$y > -\frac{7}{9} \text{ e } y < \frac{9}{5}$$

o sea, $-\frac{7}{9} < y < \frac{9}{5} = 1\frac{4}{5}$

Luego, ya que y debe ser una cantidad entera, no podremos asignarle más valores que o y 1.

Ahora, para $y = o$, tendremos: $x < o$, $x < \frac{1}{2}$, $x > -\frac{7}{4} = -1\frac{3}{4}$.

Luego, sólo podemos señalar $x = -1$, como valor de x, pues éste es el unico valor entero, comprendido entre 0 y $(-1\frac{3}{4})$.

Para $y = 1$, tendremos: $x < 2$, $x < -1$, $x > -2$.

Pero no habiendo números enteros, comprendidos entre (-1) y (-2), entonces a $y = 1$ no corresponde ningún valor de x; de suerte que, el sistema propuesto sólo puede verificarse, en valores enteros, por el único par:

$$y = o, x = -1$$

9.8. *Análisis Indeterminado de primer grado.*

Sabemos que, cuando el número de las incógnitas es mayor que el de las ecuaciones, que se proponen para ser resueltas, estas ecuaciones o sistemas son indeterminadas o indeterminados, es decir, que existe una infinidad de sistemas valorables de tales incógnitas, y que pueden verificarlas.

Pero, si se pide que los valores de las incógnitas sean *enteros*, el número de soluciones podrá no ser ya infinito, y aún a veces la cuestión será imposible, sobre todo, si además se exige que, estos valores sean, no sólo enteros, sino también *positivos*.

La especificación de esta clase de valores forman el objeto del *Análisis Indeterminado*.

De lo anterior se desprende que *no* se puede hablar de la *resolución* de esta clase de ecuaciones (o sistemas de ecuaciones), sino de hallar métodos para determinar las soluciones enteras (o enteras y positivas) de esta clase de ecuaciones, o sistemas de ecuaciones.

El primero que se ocupó de esta clase de ecuaciones fue Diofanto de Alejandría, quien a principios del siglo IV de nuestra era las trató en su célebre obra sobre aritmética, y por esto estas ecuaciones se llaman diofánticas. En nuestros tiempos la discusión de estas ecuaciones o sistemas se llaman "Análisis Indeterminado del primer grado".

La forma más sencilla de una ecuación indeterminada del primer grado, es la siguiente que contiene sólo dos incógnitas:

(*) $a x + b y = c$

en la cual a, b, c representan tres números enteros.

Caben ahora las siguientes preguntas:

1) ¿Bajo qué condición será posible satisfacer la ecuación (*) por valores enteros de las incógnitas x e y?

2) ¿Cuántas soluciones enteras tiene en este caso la ecuación (*)?

3) ¿En qué caso es posible satisfacer la ecuación (*) por valores enteros y positivos de las incógnitas y cuántas de estas soluciones hay?

En primer lugar, se puede siempre suponer los coeficientes enteros a, b, c primos entre sí, porque si no lo fueran, se los dividiría por su máximo común divisor; por ejemplo, la ecuación:

$6x - 2y = 18$ será supuesta escrita: $3x - y = 9$

En contestación a la primera pregunta, si los coeficientes a y b de las incógnitas x e y no son primos entre sí, el problema es imposible, es decir, no existen valores enteros de x e y que satisfagan la ecuación (*).

En efecto, si a y b no son primos entre sí, admiten un divisor común distinto de la unidad; y si x e y son enteros, este divisor común a a y b divide también los múltiplos ax y by de a y b, y por consiguiente divide su suma $ax + by$, o sea a c, lo que es contrario a la hipótesis hecha de que a, b, c son supuestos primos entre sí.

Luego, a y b deben ser primos entre sí para que la ecuación $ax + by = c$ admita soluciones enteras

Así hemos dado respuesta a la primera pregunta formulada anteriormente.

La segunda pregunta queda contestada por el siguiente, *Teorema*.

Cuando la ecuación $ax + by = c$ admite una solución entera, entonces admite una infinidad de ellas.

Dem. Sea $x = m$, $y = n$ una solución entera de la ecuación,

$$ax + by = c$$

esto es, $am + bn = c$

De donde, $ax + by = am + bn$

o bien, $a(x - m) = -b(y - n)$

o aún, $x - m = -\frac{b}{a}(y - n) = \frac{-b(y - n)}{a}$

Como $(x - m)$ e $(y - n)$ son números enteros y a y b primos entre sí, entonces necesariamente, por el Teorema de Euclides, a es un divisor de $(y - n)$; luego,

$$y - n = at$$

y por tanto, $x - m = -bt$

siendo t un entero arbitrario.

De donde resulta que:

$$(1)\begin{cases} x = m - bt \\ y = n + at \end{cases}$$

La ecuación (*) admite, pues, una infinidad de soluciones enteras dadas por las fórmulas (1), con lo cual queda contestada la segunda pregunta.

Ahora, si sólo se admiten soluciones enteras y positivas de la ecuación (*), es preciso averiguar si tales soluciones son posibles, lo que sucede siempre, si *a* y *b* son de distinto signo y en este caso hay un número infinito de soluciones enteras y positivas. Pero si *a* y *b* son positivos, y subsiste la condición de que $a + b > c$, es evidente que ni la solución más pequeña $x = y = 1$ puede satisfacer la ecuación; luego, en este caso, debe ser $a + b < c$ para que pueda haber soluciones enteras y positivas. Además se ve por las igualdades (1) que el número t debe elegirse de modo que:

$$m - bt > o, \text{ luego } t < \frac{m}{b}$$

$$n + at > o, \text{ luego } t > -\frac{n}{a}$$

Si estas dos últimas condiciones no se pueden satisfacer simultáneamente por un valor entero de t, no habrá soluciones enteras y positivas; en caso contrario, habrá tantas soluciones enteras y positivas como números enteros hay en la limitación:

$$-\frac{n}{a} < t < \frac{m}{b}$$

con lo cual queda contestada la tercera pregunta.

En resumen, el análisis de las ecuaciones indeterminadas de primer grado se limita al caso más frecuente de obtener soluciones enteras y positivas.

Recordemos primeramente que para que la ecuación,

$$ax + by = c$$

pueda resolverse en números enteros, es absolutamente indispensable que el máximo común divisor de *a* y *b* divida a *c*.

Recordemos igualmente que si *a* y *b* son primos entre sí, la ecuación tiene una infinidad de soluciones enteras, que se obtienen por las fórmulas:

$$\begin{cases} x = m - bt \\ y = n + at \end{cases}$$

donde *m* y *n* es ya una solución entera, y el parámetro t toma todos los valores enteros de $(-\infty)$ a $(+\infty)$.

En realidad, las fórmulas generales son:

$$\begin{cases} x = m \mp bt \\ y = n \pm at \end{cases}$$

pues los valores de *x* e *y* que se obtienen con los valores positivos de t son: $x = m - bt, y = n + at$, y los valores de *x* e *y* que se obtienen con los valores negativos de t son:

$$x = m + bt, y = n - at.$$

Finalmente, el problema consiste en hallar una solución entera cualquiera $x = m$ e $y = n$, y aplicar en seguida las fórmulas citadas.

Para encontrar una primera solución pueden emplearse varios procedimientos. Nosotros indicaremos algunos.

En particular, si el coeficiente de una de las incógnitas es la unidad, por ejemplo:

$$x + by = c$$

se tiene de inmediato una solución, que es:

$$x = c, \quad y = o$$

y las fórmulas $x = m - bt$, $y = n + at$, dan todas las otras soluciones.

Basándonos en este hecho especial, nosotros podremos ahora, dada la ecuación general,

$$ax + by = c$$

encontrar otra ecuación en la cual el coeficiente de una incógnita sea igual a la unidad. Nosotros aplicaremos el procedimiento por medio de un ejemplo.

Sea la ecuación,

$$14x + 23y = 139$$

Despejemos la incógnita de menor coeficiente, en este caso la x, y obtenemos:

$$x = \frac{139 - 23y}{14}$$

y efectuamos esta división tomando los cocientes por defecto o por exceso de manera de obtener siempre los restos más pequeños; esto es, menores que la mitad del divisor, tendremos:

$$x = \frac{139 - 23y}{14} = 10 - \frac{1}{14} - (2y - \frac{5y}{14}) = 10 - 2y + \frac{-1 + 5y}{14}$$

En la práctica se efectúa, desde luego, la división como indica el último miembro, escribiendo primero los cocientes y en seguida los restos se van escribiendo sobre la raya de la fracción complementaria.

Si ahora hacemos $\frac{-1 + 5y}{14} = z$, será,

De donde resulta que:

$$(1)\begin{cases} x = m - bt \\ y = n + at \end{cases}$$

La ecuación (*) admite, pues, una infinidad de soluciones enteras dadas por las fórmulas (1), con lo cual queda contestada la segunda pregunta.

Ahora, si sólo se admiten soluciones enteras y positivas de la ecuación (*), es preciso averiguar si tales soluciones son posibles, lo que sucede siempre, si *a* y *b* son de distinto signo y en este caso hay un número infinito de soluciones enteras y positivas. Pero si *a* y *b* son positivos, y subsiste la condición de que $a + b > c$, es evidente que ni la solución más pequeña $x = y = 1$ puede satisfacer la ecuación; luego, en este caso, debe ser $a + b < c$ para que pueda haber soluciones enteras y positivas. Además se ve por las igualdades (1) que el número t debe elegirse de modo que:

$$m - bt > o, \text{ luego } t < \frac{m}{b}$$

$$n + at > o, \text{ luego } t > -\frac{n}{a}$$

Si estas dos últimas condiciones no se pueden satisfacer simultáneamente por un valor entero de t, no habrá soluciones enteras y positivas; en caso contrario, habrá tantas soluciones enteras y positivas como números enteros hay en la limitación:

$$-\frac{n}{a} < t < \frac{m}{b}$$

con lo cual queda contestada la tercera pregunta.

En resumen, el análisis de las ecuaciones indeterminadas de primer grado se limita al caso más frecuente de obtener soluciones enteras y positivas.

Recordemos primeramente que para que la ecuación,

$$ax + by = c$$

pueda resolverse en números enteros, es absolutamente indispensable que el máximo común divisor de *a* y *b* divida a *c*.

Recordemos igualmente que si *a* y *b* son primos entre sí, la ecuación tiene una infinidad de soluciones enteras, que se obtienen por las fórmulas:

$$\begin{cases} x = m - bt \\ y = n + at \end{cases}$$

donde *m* y *n* es ya una solución entera, y el parámetro t toma todos los valores enteros de $(-\infty)$ a $(+\infty)$.

En realidad, las fórmulas generales son:

$$\begin{cases} x = m \mp bt \\ y = n \pm at \end{cases}$$

pues los valores de *x* e *y* que se obtienen con los valores positivos de t son: $x = m - bt$, $y = n + at$, y los valores de *x* e *y* que se obtienen con los valores negativos de t son:

$$x = m + bt, y = n - at.$$

Finalmente, el problema consiste en hallar una solución entera cualquiera $x = m$ e $y = n$, y aplicar en seguida las fórmulas citadas.

Para encontrar una primera solución pueden emplearse varios procedimientos. Nosotros indicaremos algunos.

En particular, si el coeficiente de una de las incógnitas es la unidad, por ejemplo:

$$x + by = c$$

se tiene de inmediato una solución, que es:

$$x = c, \quad y = o$$

y las fórmulas $x = m - bt$, $y = n + at$, dan todas las otras soluciones.

Basándonos en este hecho especial, nosotros podremos ahora, dada la ecuación general,

$$ax + by = c$$

encontrar otra ecuación en la cual el coeficiente de una incógnita sea igual a la unidad. Nosotros aplicaremos el procedimiento por medio de un ejemplo.

Sea la ecuación,

$$14x + 23y = 139$$

Despejemos la incógnita de menor coeficiente, en este caso la x, y obtenemos:

$$x = \frac{139 - 23y}{14}$$

y efectuamos esta división tomando los cocientes por defecto o por exceso de manera de obtener siempre los restos más pequeños; esto es, menores que la mitad del divisor, tendremos:

$$x = \frac{139 - 23y}{14} = 10 - \frac{1}{14} - (2y - \frac{5y}{14}) = 10 - 2y + \frac{-1 + 5y}{14}$$

En la práctica se efectúa, desde luego, la división como indica el último miembro, escribiendo primero los cocientes y en seguida los restos se van escribiendo sobre la raya de la fracción complementaria.

Si ahora hacemos $\frac{-1 + 5y}{14} = z$, será,

$$5y - 14z = 1$$

ecuación mucho más sencilla que la propuesta.

En esta ecuación despejamos ahora y, por tener el menor coeficiente, y se tiene:

$$y = \frac{1 + 14z}{5} = 3z + \frac{1 - z}{5}$$

Haciendo ahora $\frac{1 - z}{5} = u$, se tendrá

$$5u + z = 1$$

ecuación que por tener un coeficiente de una de las incógnitas igual a la unidad, tiene ya por solución entera la:

$$z = 1, \quad u = o$$

Luego, las fórmulas generales serán:

$$\begin{cases} z = 1 - 5t = 1 - 5t \\ u = o + t = t \end{cases}$$

Volviendo a los valores de x e y, obtendremos:

$$y = 3z + u = 3(1 - 5t) + t = 3 - 14t$$
$$x = 10 - 2y + z = 10 - 2(3 - 14t) + (1 - 5t)$$
$$x = 5 + 23t$$

En resumen, las soluciones generales son:

$$\begin{cases} x = 5 + 23t \\ y = 3 - 14t \end{cases}$$

es conveniente comprobar las soluciones dando a t algunos valores; por ejemplo:

Para $t = o$, será $x = 5$, $y = 3$; la ecuación es:

$$14 \cdot 5 + 23 \cdot 3 = 70 + 69 = 139$$

Para $t = 1$, será $x = 28$, $y = -11$; la ecuación es:

$$14 \cdot 28 - 23 \cdot 11 = 392 - 253 = 139$$

Las soluciones enteras y positivas se obtienen si se imponen las condiciones:

$$5 + 23t > o \text{ y } 3 - 14t > o$$

de donde, $t > -\frac{5}{23}$ y $t < \frac{3}{14}$

o bien, $-\frac{5}{23} < t < \frac{3}{14}$

y así vemos que, t no puede recibir más que el valor o, al cual corresponde la única solución entera positiva:

$$x = 5, \quad y = 3$$

*****

Otro procedimiento consiste también en despejar una incógnita, la de menor coeficiente y siempre que este coeficiente sea pequeño, como ocurre, por ejemplo, con la ecuación:

$$3x + 11y = 26$$

$$x = \frac{26 - 11y}{3}$$

Bastará dar valores a y desde cero en adelante y pronto hemos de encontrar un valor entero para x, pues para $y = o, 1, 2, \ldots$ hay un valor entero para x, y nada más que uno. Así, para $y = o$ es $x = \frac{26}{3}$, que no es entero; para $y = 1$ da $x = \frac{15}{3} = 5$

Luego, las fórmulas generales serán:

$$\begin{cases} x = 5 - 11t \\ y = 1 + 3t \end{cases}$$

Se comprende que si los coeficientes de las incógnitas son grandes, este procedimiento sería penoso.

Indicaremos, por ahora, un último procedimiento, cuando se advierte casi de inmediato una primera solución.

Los veremos por medio de algunos ejemplos.

Si la ecuación es $6x + 5y = 42$ en la que 42 es múltiplo de 6, entonces será evidentemente $x = 7$, $y = o$ una primera solución y, por tanto, las soluciones generales son:

$$\begin{cases} x = 7 - 5t \\ y = 0 + 6t \end{cases}$$

Si la ecuación es $ax + by = o$, evidentemente $x = o$, $y = o$, y las otras soluciones serán:

$$\begin{cases} x = -bt \\ y = at \end{cases}$$

Si la ecuación es $5x + 2y = 11$, se buscan combinaciones de múltiplos de 5 y de 2, que por suma o resta den 11.

Por ejemplo: $15 - 4 = 11$, $5 \cdot 3 + 2(-2) = 11$; luego es

$$x = 3, \quad y = -2$$

Las otras soluciones serán:

$$\begin{cases} x = 3 - 2t \\ y = -2 + 5t \end{cases}$$

sea la ecuación,

$$153x + 13y = 270$$

Observaremos que 153 y 270 son divisibles por 9.

Dividiendo la ecuación por 9, tendremos,

$$17x + \frac{13y}{9} = 30$$

Haciendo $y = 9y'$, se tendrá:

$$17x + 13y' = 30$$

ecuación más sencilla que la propuesta. Tenemos:

$$y' = \frac{30 - 17x}{13} = 2 - x + \frac{4 - 4x}{13} = 2 - x + 4 \cdot \frac{1 - x}{13}$$

Poniendo $\frac{1 - x}{13} = z$, resulta $x + 13z = 1$; luego, una primera solución de esta última ecuación es,

$$x = 1, \quad z = o$$

por lo tanto, las soluciones generales serán:

$$\begin{cases} x = 1 - 13t \\ z = o + t \end{cases}$$

Luego, $y' = 2 - x + 4z = 2 - (1 - 13t) + 4t = 1 + 17t$, por tanto, $y = 9y' = 9 + 153t$

Por consiguiente, las soluciones generales serán:

$$\begin{cases} x = 1 - 13t \\ y = 9 + 153t \end{cases}$$

Siempre que algún coeficiente de una incógnita tenga algún factor común con el término conocido, tomaremos una incógnita auxiliar como acaba de explicarse.

9.9. *Sistemas simplemente indeterminados*

Llamaremos sistemas simplemente indeterminados a aquellos en que el número de incógnitas exceda al de las ecuaciones en una unidad.

El caso más sencillo es el de una ecuación con dos incógnitas,

$$ax + by = c$$

que ya sabemos resolver en números enteros.

Consideraremos ahora un sistema simplemente indeterminado de dos ecuaciones con tres incógnitas.

$$(**) \quad \begin{cases} a\,x + b\,y + c\,z = d \\ a'x + b'y + c'z = d' \end{cases}$$

Como ya lo hemos dicho, para que una ecuación entera de primer grado con dos (o más) incógnitas, pueda verificarse para valores enteros de estas incógnitas, es necesario que el máximo común divisor de todos los coeficientes de ellas divida a la cantidad constante o término conocido.

Supongamos cumplida esta condición en el sistema dado (**) esto es, que ya los coeficientes a, b, c son primos entre sí, así como a', b', c'.

En caso de que dos coeficientes tuviesen un divisor común con el segundo miembro, se dividirá por él tomando una incógnita auxiliar, como ya es sabido (ver la observación anterior).

Si eliminamos la incógnita z, por ejemplo, el sistema propuesto será equivalente al formado por una de ellas y por la ecuación resultante.

$$\begin{cases} ax + by + cz = d \\ Ax + By = C \end{cases}$$

y resolviendo esta segunda, se tendrá:

$$(1) \quad \begin{cases} x = m - Bt \\ y = n + At \end{cases}$$

valores que substituidos en la primera, nos darán una ecuación de la forma,

$$Dt + cz = E$$

Resolviendo esta última, tendremos,

$$\begin{cases} t = p - cs \\ z = q + Ds \end{cases}$$

Y substituyendo estos valores en las (1), tendremos finalmente para soluciones enteras:

$$(2) \quad \begin{cases} x = m - B(p - cs) = (m - Bp) + Bcs \\ y = n + A(p - cs) = (n + Ap) - Acs \\ z = q + Ds \end{cases}$$

Para determinar las soluciones enteras y positivas, tendremos:

$$(m - Bp) + Bcs > o,\ (n + Ap) - Acs > o,\ q + Ds > o$$

o sea,

$$s > \frac{Bp - n}{Bc}, s < \frac{n + Ap}{Ac},\ s > -\frac{q}{D}$$

Los dos límites, primero y último, son L. I. y el segundo es L. S.

De verificarse o no la desigualdad,

$$\frac{Bp - m}{Bc} < s > \frac{n + Ap}{Ac}$$

dependerá que haya o no soluciones positivas. Si se verifica, los valores que admite s, substituídos en los de x, y, z, nos darán las soluciones enteras y positivas del sistema propuesto (**).

Ahora, ocurre preguntar, ¿qué incógnita debe eliminarse?

A este respecto, sólo afirmaremos que la incógnita que conviene eliminar es aquella cuyos coeficientes sean primos entre sí.

Ejemplo:

$$\left.\begin{array}{r} 3x + 5y + 6z = 104 \\ 9x + 3y + 8z = 164 \end{array}\right|$$

Como los coeficientes de la y son primos entre sí, eliminaremos esta incógnita, y por reducción tendremos:

$$\left.\begin{array}{r} 9x + 15y + 18z = 312 \\ 45x + 15y + 40z = 820 \end{array}\right|$$

$$36x + 22z = 508$$

Dividiendo por 4 se tiene,

$$9x + \frac{11}{2}z = 127$$

y poniendo $z = 2x'$, resulta,

$$9x + 11z' = 127$$

$$x = \frac{127 - 11z'}{9} = 14 - z' + \frac{1 - 2z'}{9}$$

Haciendo $\frac{1 - 2z'}{9} = u$, resulta $9u + 2z' = 1$.

Se advierte que una primera solución de esta última ecuación, es $u = 1, z' = -4$. Luego, las fórmulas generales son:

$$\begin{cases} u = 1 - 2t \\ z' = -4 + 9t \end{cases}$$

Por lo tanto,

$$x = 14 - (-4 + 9t) + (1 - 2t) = 19 - 11t$$

$$z = 2z' = 2(-4 + 9t) = -8 + 18t$$

Substituyendo estos dos valores de x y z en una de las ecuaciones del sistema propuesto, por ejemplo en la primera, obtendremos:

$$3(19 - 11t) + 5y + 6(-8 + 18t) = 104$$

$$5y + 75t = 95$$

$$y + 15t = 19$$

$$y = 19 - 15t$$

Luego, las fórmulas generales que dan todas las soluciones enteras del sistema son:

$$\begin{cases} x = 19 - 11t \\ y = 19 - 15t \\ z = -8 + 18t \end{cases}$$

Las soluciones enteras y positivas serán:

$$19 - 11t > o, 19 - 15t > o, -8 + 18t > o$$

o sea,

$$t < \frac{19}{11}, t < \frac{19}{15}, t > \frac{8}{18}$$

De donde

$$\frac{8}{18} < t < \frac{19}{15}$$

y el único valor entero para t es 1. Luego,

$$x = 8, y = 3, z = 10$$

es la única solución entera y positiva del sistema propuesto.

Observación. Cuando no haya ninguna incógnita con coeficientes primos entre sí, se elimina aquella cuya razón de coeficientes sea la más sencilla.

Por ejemplo, en el sistema:

$$\left.\begin{array}{r} 6x + 9y + 14z = 77 \\ 4x + 15y + 7z = 51 \end{array}\right|$$

conviene eliminar la z, por ser $\frac{14}{7} = 2$ la razón más sencilla.

El método acabado de explicar se extiende fácilmente al caso de n ecuaciones con n + 1 incógnitas.

En efecto, eliminaremos una misma incógnita entre una de las ecuaciones propuestas y cada una de las demás, lo cual dará (n − 1) ecuaciones entre las otras n incógnitas.

De la misma manera eliminaremos una de las n incógnitas restantes, entre una de éstas (n − 1) ecuaciones y cada una de las demás, lo cual dará (n − 2) ecuaciones con (n − 1) incógnitas.

Y continuando así, llegaremos por último a una ecuación con dos incógnitas.

De esta manera podremos reemplazar el sistema de las ecu iones propuestas por el formado con una ecuación con (n + 1) incógnitas, otra con n incógnitas, otra con (n − 1) incógnitas, etc., y otra con dos incógnitas. Resuelta esta última con una indeterminada t, no hay más que substituir de abajo hacia arriba, y se van obteniendo los valores de todas las incógnitas en función de la única indeterminada t.

Ejemplo:

Sea el sistema,

$$\left|\begin{array}{l} 2x+3y-3z+5u-4v=-1 \\ 3x+2y-3z+4u-5v=-11 \\ 4x-5y+2z-3u-2v=-22 \\ 2x+5y-4z+2u+3v=23 \end{array}\right.$$

Eliminemos la incógnita x entre la primera ecuación y cada una de las demas. Así tendremos:

$$\left|\begin{array}{l} 2x+3y-3z+5u-4v=-1 \quad \cdot 3 \\ 3x+2y-3z+4u-5v=-11 \quad \cdot 2 \end{array}\right.$$

$$\left|\begin{array}{ll} (1) & 5y-3z+7u-2v=19 \\ & 2x+3y-3z+5u-4v=-1 \quad \cdot 2 \\ & 4x-5y+2z-3u-2v=-22 \end{array}\right.$$

$$\left|\begin{array}{ll} (2) & 11y-8z+13u-6v=20 \\ & 2x+3y-3z+5u-4v=-1 \\ & 2x+5y-4z+2u+3v=23 \end{array}\right.$$

(3) $\quad 2y-z-3u+7v=24$

Las ecuaciones (1), (2) y (3) constituyen el sistema:

$$\left|\begin{array}{l} 5y-3z+7u-2v=19 \\ 11y-8z+13u-6v=20 \\ 2y-z-3u+7v=24 \end{array}\right.$$

Procediendo como antes, eliminemos la incógnita z.

$$\left|\begin{array}{l} 5y-3z+7u-2v=19 \quad \cdot 8 \\ 11y-8z+13u-6v=20 \quad \cdot 3 \end{array}\right.$$

$$\left|\begin{array}{ll} (4) & 7y+17u+2v=92 \\ & 5y-3z+7u-2v=19 \\ & 2y-z-3u+7v=24 \quad \cdot 3 \end{array}\right.$$

(5) $\quad y-16u+23v=53$

Las ecuaciones (4) y (5) forman el sistema,

$$\left|\begin{array}{l} 7y+17u+2v=92 \\ y-16u+23v=53 \quad \cdot 7 \end{array}\right.$$

Eliminando la incógnita y, resulta:

$$-129u+159v=279$$

o bien, $-43u+53v=93 \qquad (6)$

Resolvamos la ecuación (6).

$$u=\frac{53v-93}{43}=v-2+\frac{10v-7}{43}=v-2+t$$

donde, $t=\frac{10v-7}{43}$, o sea, $10v-43t=7$.

Resolviendo esta última ecuación, se obtiene:

$$v=\frac{43t+7}{10}=4t+\frac{3t+7}{10}=4t+t'$$

donde, $t'=\frac{3t+7}{10}$, o sea, $10t'-3t=7$.

Para esta ecuación se advierte una primera solución que es: $t'=1$, $t=1$; luego,

$$\begin{cases} t' = 1+3t'' \\ t = 1+10t'' \end{cases}$$

Retrocediendo, hallaremos primeramente los valores de v y de u en función de la indeterminada o parámetro t''. Se tiene:

$$v=4t+t'=4(1+10t'')+1+3t''=5+43t''$$

$$u=v-2+t=5+43t''-2+1+10t''=4+53t''$$

Para hallar los valores de las demás incógnitas, procedemos de la siguiente manera:

Hallemos y en la ecuación (5):

$$y=53+16u-23v=53+16(4+53t'')-23(5+43t'')$$

$$y=2-141t''$$

Hallemos z en la ecuación (3):

$$z=2y-3u+7v-24=2(2-141t'')-3(4+53t'')+7(5+43t'')-24$$

$$z=3-140t''$$

Finalmente, hallaremos x en la primera de las ecuaciones del sistema propuesto:

$$x=\frac{-1-3y+3z-5u+4v}{2}=\frac{2-90t''}{2}=1-45t''$$

Así hemos obtenido las fórmulas generales siguientes:

$$\begin{cases} x = 1-45t'' \\ y = 2-141t'' \\ z = 3-140t'' \\ u = 4+53t'' \\ v = 5+43t'' \end{cases}$$

que dan las infinitas soluciones enteras del sistema dado, cuando el parámetro o indeterminada t'' varía desde ($-\infty$) hasta ($+\infty$).

Es interesante que el lector determine algunas soluciones y compruebe que ellas satisfacen al sistema propuesto.

Las soluciones enteras y positivas se determinan mediante las condiciones:

$1 - 45\,t'' > 0,\ 2 - 141\,t'' > 0,\ 3 - 140\,t'' > 0,\ 4 + 53\,t'' > 0,$
$5 + 43\,t'' > 0$, o sea,

$$t'' < \frac{1}{45},\ t'' < \frac{2}{141},\ t'' < \frac{3}{140},\ t'' > -\frac{4}{53},\ t'' > -\frac{5}{43}$$

Se observa que $t'' = 0$, es el único valor que satisface a estos límites, resultando:

$$x = 1,\ y = 2,\ z = 3,\ u = 4,\ v = 5$$

como la única solución entera y positiva del sistema propuesto.

9.10. *Sistemas más que indeterminados*

Llamaremos sistemas más que indeterminados a aquellos en que el número de incógnitas excede al de ecuaciones en más de una unidad. El caso más sencillo es el de una ecuación con tres incógnitas,

(***) $a\,x + b\,y + c\,z = d$

que supondremos simplificada en lo posible, de suerte que los tres coeficientes a, b, c sean primos entre sí, sin lo cual nuestra ecuación no admitiría ninguna solución en números enteros.

En seguida examinaremos si dos de los tres coeficientes mencionados son primos entre sí, y haremos pasar al segundo miembro al que tenga un factor común con uno de los otros dos. Si, pués, *a* y *b* son primos entre sí, resultará:

$$a\,x + b\,y = d - c\,z = d'$$

haciendo, para abreviar, $d - c\,z = d'$.

Luego, resolveremos la ecuación,

$$a\,x + b\,y = d'$$

y obtendremos:

$$x = m - b\,t,\quad y = n + a\,t$$

siendo m y n funciones de d'. Reemplazando estos valores a *d'* por el suyo, $d - c\,z$, hallaremos expresiones de la forma:

$$x = A + A'Z - b\,t,\quad y = B + B'z + a\,t$$

y, a todos los valores enteros de *z* y de *t* corresponderán otros, también enteros, para *x* y para *y*.

Ahora, en el caso donde los tres coeficientes a, b, c no son primos entre

sí dos a dos, también una de las incógnitas será transportada al segundo miembro, lo cual producirá una ecuación de la forma:

$$p\,a'\,x + p\,b'\,y = d - c\,z$$

donde p representa el factor común que tienen los coeficientes *a* y *b* de los términos que quedan en el primer miembro, y por *a'* y *b'* los coeficientes primos entre sí, que resultan de dividir *a* y *b* por el factor común p. Tendremos, que la ecuación propuesta (***), se podrá poner bajo la forma:

$$a'\,x + b'\,y = \frac{d - c\,z}{p}$$

Y como el primer miembro ha de ser un número entero, el segundo también tendrá que serlo; de modo que tendremos, llamando t a este número entero,

$$\frac{d - c\,z}{p} = t$$

de lo cual resultará,

$$a'\,x + b'\,y = t$$

y, resolviendo esta última ecuación, hallaremos:

$$\begin{cases} x = m - b'\,t' \\ y = n + a'\,t' \end{cases}$$

Aquí m y n son funciones de la indeterminada t, cuyos valores enteros deben, juntamente con otros también enteros de z, verificar la ecuación:

$$d - c\,z = p\,t$$

o bien, $c\,z + p\,t = d$

en la cual *c* y *p* son primos entre sí.

Resolveremos, pues, esta ecuación y deduciremos:

$$\begin{cases} z = m' - p\,t'' \\ t = n' + c\,t'' \end{cases}$$

Sustituyendo este valor de $t = n' + c\,t''$ en la ecuación,

$$a'\,x + b'\,y = t = n' + c\,t''$$

resulta esta nueva ecuación que reemplazará a la propuesta, de la cual deduciremos los valores:

$$\begin{cases} x = m'' - b'\,t''' \\ y = n'' + a'\,t''' \end{cases}$$

siendo m'' y n'' números que dependen de la indeterminada t''; por consiguiente, estas dos incógnitas *x* e *y* dependerán de las dos indeterminadas t'' y t''', mientras que la tercera incógnita *z*, sólo depende de la indeterminada t''.

En resumen, podemos decir que los valores de *x*, *y*, *z*, serán de la forma:

$$\begin{cases} x = A + A't'' - b't''' \\ y = B + B't'' + a't''' \\ z = m' - pt'' \end{cases}$$

De donde concluímos que, cuando los coeficientes de las incógnitas que quedan en el primer miembro tienen un factor común, hay que aplicar el método del análisis indeterminado a una ecuación más que en el caso anterior.

Ejemplo:

Sea la ecuación,

$$6x + 22y + 15z = 77$$

Aquí los coeficientes de y y de z son primos entre sí; luego, transponemos el término 6 x al segundo miembro, y resultará:

$$22y + 15z = 77 - 6x = d$$

y en seguida vemos que tiene una solución este enunciado, haciendo:

$$y = -2d, z = 3d, \text{ ya que,}$$

$$22 \cdot (-2d) + 15 \cdot 3d \doteq -44d \dot{+} 45d = d$$

De suerte que,

$$\begin{cases} y = -2d - 15t \\ z = 3d + 22t \end{cases}$$

o reemplazando d por su valor $d = 77 - 6x$,

$$\begin{cases} y = -2(77 - 6x) - 15t = -154 + 12x - 15t \\ z = 3(77 - 6x) + 22t = 231 - 18x + 22t \end{cases}$$

y, a todos los valores enteros de x y de t corresponderán otros, también enteros, para y y para z.

Ahora, si queremos resolver la ecuación dada en números enteros y positivos, empezaremos por los de y y z, haciendo:

$$-154 + 12x - 15t > 0, 231 - 18x + 22t > 0$$

después, deduciremos de estas desigualdades:

$$t < \frac{-154 + 12x}{15}, t > \frac{-231 + 18x}{22}$$

lo cual determina un límite superior y otro inferior de t.

Para que estos límites no sean contradictorios, necesitaremos que se verifique,

$$\frac{-231 + 18x}{22} < \frac{-154 + 12x}{15}$$

o sea, $-3465 + 270x < -3388 + 264x$

$$6x < 77$$

$$x < \frac{77}{6} = 12\frac{5}{6}$$

Luego, x podrá recibir los valores enteros y positivos desde 1 hasta 12, los cuales iremos sustituyendo en los límites de t, y hallaremos:

$$\text{Para } x = 1, \begin{cases} t < \dfrac{-154 + 12}{15} = -\dfrac{142}{15} = -9\dfrac{7}{15} \\ t > \dfrac{-231 + 18}{22} = -\dfrac{213}{22} = -9\dfrac{15}{22} \end{cases} \text{ Contradictorios}$$

$$\text{Para } x = 2, \begin{cases} t < \dfrac{-154 + 24}{15} = -\dfrac{130}{15} = -8\dfrac{10}{15} \\ t > \dfrac{-231 + 36}{22} = -\dfrac{195}{22} = -8\dfrac{19}{22} \end{cases} \text{ Contradictorios}$$

$$\text{Para } x = 3, \begin{cases} t < \dfrac{-154 + 36}{15} = -\dfrac{118}{15} = -7\dfrac{13}{15} \\ t > \dfrac{-231 + 54}{22} = -\dfrac{177}{22} = -8\dfrac{1}{22} \end{cases} \quad t = -8$$

Llevado los valores $x = 3$ y $t = -8$ a las expresiones de y y de z, se encuentra:

$$y = -154 + 36 + 120 = 2$$

$$z = 231 - 54 - 176 = 1$$

Luego, los valores $x = 3, y = 2, z = 1$ constituyen una solución entera y positiva de la ecuación propuesta.

Puede comprobarse que para los demás valores de x hasta $x = 12$, dan para t límites contradictorios.

Observación. Hemos resuelto la ecuación $6x + 22y + 15z = 77$ considerando la incógnita x como una indeterminada. Sin embargo, podemos prescindir de este hecho, procediendo del modo siguiente:

$$6x + 22y + 15z = 77$$

Dejamos como antes en el primer miembro las dos incógnitas y y z cuyos coeficientes son primos entre sí

$$22y + 15z = 77 - 6x$$

Pongamos $77 - 6x = t$; luego, tendremos las dos ecuaciones:

$$22y + 15z = t, 77 - 6x = t$$

Consideremos la identidad siguiente:

$$44 - 45 = -1$$

o bien, $22 \cdot 2 + 15 \cdot (-3) = -1$

o también, $22 \cdot (-2) + 15 \cdot 3 = 1$

o aún, $22 \cdot (-2t) + 15 \cdot (3t) = t$

Luego, la primera ecuación $22y + 15z = t$ admite por solución:

$y = -2t, z = 3t$

Entonces las fórmulas generales serán:

$$\begin{cases} y = -2t - 15t' \\ z = 3t + 22t' \end{cases}$$

Análogamente, consideremos la identidad:

$72 + 5 = 77$

o sea, $6 \cdot 12 + 1 \cdot 5 = 77$

Luego, la segunda ecuación $77 - 6x = t$, o sea, $6x + t = 77$ admite por solución $x = 12, t = 5$

Por tanto, las fórmulas generales serán:

$$\begin{cases} x = 12 + t'' \\ t = 5 - 6t'' \end{cases}$$

Sustituyendo este valor de $t = 5 - 6t''$ en las expresiones de los valores de y y de z, tendremos:

$y = -2(5 - 6t'') - 15t' = -10 + 12t'' - 15t'$

$z = 3(5 - 6t'') + 22t' = 15 - 18t'' + 22t'$

En resumen, se tendrá:

$$\begin{cases} x = 12 + t'' \\ y = -10 + 12t'' - 15t' \\ z = 15 - 18t'' + 22t' \\ t = 5 - 6t'' \end{cases}$$

Con este procedimiento logramos expresar las tres incógnitas de la ecuación propuesta en función de dos indeterminadas, t' y t''.

Para obtener las soluciones enteras y positivas, si existe, habrá que considerar las tres desigualdades siguientes:

$x = 12 + t'' > 0, y = 10 + 12t'' - 15t' > 0, z = 15 - 18t'' + 22t' > 0$

y de las cuales deduciremos:

$$t'' > -12$$

$$t'' > \frac{15t' + 10}{12}$$

$$t'' < \frac{22t' + 15}{18}$$

Para que haya compatibilidad han de verificarse las desigualdades siguientes:

$$\begin{cases} -12 < \dfrac{22t' + 15}{18} \\ \dfrac{15t' + 10}{12} < \dfrac{22t' + 15}{18} \end{cases}$$

Resolviendo este sistema de inecuaciones, se encontrará:

$$\begin{array}{l|l} 22t' + 15 > -216 & 3(15t' + 10) < 2(22t' + 15) \\ 22t' > -231 & 45t' + 30 < 44t' + 30 \\ t' > -\dfrac{231}{22} = -10\dfrac{11}{12} & 45t' < 44t' \\ & t' < 0 \end{array}$$

De donde, $-10\frac{11}{12} < t' < 0$; o sea, t' podrá recibir los valores:

$-10, -9, -8, -7, \ldots, -2, -1$

Para $t' = -10$, satisfacemos $t'' > -12$, y se tendrá:

$$\left.\begin{array}{l} t'' > \dfrac{15 \cdot (-10) + 10}{12} = -\dfrac{140}{12} = -11\dfrac{8}{12} \\ t'' < \dfrac{22 \cdot (-10) + 15}{18} = -\dfrac{205}{18} = -11\dfrac{7}{18} \end{array}\right\} \text{Incompatibles}$$

Para $t' = 9$, satisfacemos $t'' > -12$, y se tendrá:

$$\left.\begin{array}{l} t'' > \dfrac{15 \cdot (-9) + 10}{12} = -\dfrac{125}{12} = -10\dfrac{5}{12} \\ t'' < \dfrac{22 \cdot (-9) + 15}{18} = -\dfrac{183}{18} = -7\dfrac{17}{18} \end{array}\right\} \text{Contradictorios}$$

Para $t' = 8$, será $t'' > -12$, y tendremos:

$$\left.\begin{array}{l} t'' > \dfrac{15 \cdot (-8) + 10}{12} = -\dfrac{110}{12} = -9\dfrac{2}{12} \\ t'' < \dfrac{22 \cdot (-8) + 15}{18} = -\dfrac{161}{18} = -8\dfrac{17}{18} \end{array}\right\} \quad t'' = -9$$

Para este valor de $t'' = -9$, las incógnitas x, y, z tienen los valores siguientes:

$$\begin{cases} x = 12 + (-9) = 3 \\ y = -10 + 12(-9) - 15 \cdot (-8) = -10 - 108 + 120 = 2 \\ z = 15 - 18(-9) + 22 \cdot (-8) = 15 + 162 - 176 = 1 \end{cases}$$

Que es la solución obtenida anteriormente para la ecuación propuesta.

Puede verificarse que para los demás valores de t' hasta (-1), resultan para la indeterminada t'' límites contradictorios.

*****

Si se trata, ahora, de hallar las soluciones enteras y positivas de una ecuación con un número cualquiera de incógnitas (mayor que tres), se pasan todas menos dos al segundo miembro, teniendo cuidado de dejar en el primer miembro aquellas dos cuyos coeficientes sean primos entre sí; se resuelve la ecuación que resulte con relación a estas dos incógnitas, cuyas fórmulas generales vendrán en función entera de las demás incógnitas y de una indeterminada t, como ha sucedido ante-

riormente; y dando después a la t y a estas incógnitas restantes valores enteros, hallaremos también valores enteros para las otras dos.

Si cualesquiera que sean las incógnitas que se pasen al segundo miembro, las dos que quedan en el primero tienen en sus coeficientes un factor común, se aplica el método que ya se ha aplicado anteriormente para este caso.

Ejemplo.

Sea la ecuación,

$$3x + 2y - z + 5u - v = 3$$

Tendremos:

$$z - v = 3x + 2y + 5u - 3$$

Poniendo $3x + 2y + 5u - 3 = k$, tendremos la ecuación:

$$z + v = k$$

en la que una primera solución podrá ser: $z = 2k$, $v = -k$; luego, las fórmulas generales serán:

$$\begin{cases} z = 2k - t \\ v = -k + t \end{cases}$$

o bien, $\begin{cases} z = 2(3x + 2y + 5u - 3) - t = 6x + 4y + 10u - t - 6 \\ v = -(3x + 2y + 5u - 3) + t = -3x - 2y - 5u + t + 3 \end{cases}$

donde figuran cuatro indeterminadas, y dando a estas cuatro indeterminadas: x, y, u, t, valores enteros, hallaremos también valores enteros para z y para v.

Para soluciones enteras y positivas, tendremos:

$$z = 6x + 4y + 10u - t - 6 > 0$$

$$v = -3x - 2y - 5u + t + 3 > 0$$

y de las cuales deduciremos:

$$\begin{cases} t < 6x + 4y + 10u - 6 \\ t > 3x + 2y + 5u - 3 \end{cases}$$

Para que halla compatibilidad ha de verificarse la desigualdad:

$$6x + 4y + 10u - 6 > 3x + 2y + 5u - 3$$

o sea, $3x + 2y + 5u > 3$

Luego, habrá que dar a la expresión $(3x + 2y + 5u)$ valores enteros mayores que 3.

Así, pues, para $3x + 2y + 5u = 4$, tendremos:

$$3x + 2y = 4 - 5u$$

Una primera solución será: $x = 4 - 5u$, $y = -(4 - 5u)$.

Luego, las fórmulas generales serán:

$$\begin{cases} x = (4 - 5u) - 2t' = 4 - 5u - 2t' > 0 \\ y = -(4 - 5u) + 3t' = -4 + 5u + 3t' > 0 \end{cases}$$

De donde,

$$u < \frac{4 - 2t'}{5}, \quad u > \frac{4 - 3t'}{5}$$

Para que halla compatibilidad debe verificarse,

$$\frac{4 - 2t'}{5} > \frac{4 - 3t'}{5}$$

$$4 - 2t' > 4 - 3t'$$

$$t' > 0$$

Sea $t' = 1$, entonces será,

$$\begin{cases} x = 4 - 5u - 2 = 2 - 5u > 0 \\ y = -4 + 5u + 3 = -1 + 5u > 0 \end{cases} \Rightarrow \begin{cases} u < \frac{2}{5} \\ u > \frac{1}{5} \end{cases}$$

límites que son contradictorios.

Para $3x + 2y + 5u = 5$ se tendrá:

$$3x + 2y = 5 - 5u$$

con una primera solución: $x = 5 - 5u$, $y = -(5 - 5u)$

Luego,

$$\begin{cases} x = (5 - 5u) - 2t' = 5 - 5u - 2t' > 0. \\ y = -(5 - 5u) + 3t' = -5 + 5u + 3t' > 0 \end{cases}$$

de donde,

$$u < \frac{5 - 2t'}{5}, \quad u > \frac{5 - 3t'}{5}$$

$$\frac{5 - 2t'}{5} > \frac{5 - 3t'}{5}$$

$$5 - 2t' > 5 - 3t'$$

$$t' > 0$$

Sea $t' = 1$, entonces será:

$$\begin{cases} x = 5 - 5u - 2 = 3 - 5u > 0 \\ y = -5 + 5u + 3 = -2 + 5u > 0 \end{cases} \Rightarrow \begin{cases} u > \frac{3}{5} \\ u < \frac{2}{5} \end{cases}$$

límites que también son contradictorios.

Hagamos $3x + 2y + 5u = 7$; tendremos:

$$3x + 2y = 7 - 5u$$

Luego,

$$\begin{cases} x = (7 - 5u) - 2t' = 7 - 5u - 2t' > 0 \\ y = -(7 - 5u) + 3t' = -7 + 5u + 3t' > 0 \end{cases}$$

$$u < \frac{7 - 2t'}{5}, \quad u > \frac{7 - 3t'}{5}$$

Debe ser para la compatibilidad,

$$\frac{7-2t'}{5} > \frac{7-3t'}{5}$$

$$7-2t' > 7-3t'$$

$$t' > 0$$

Para $t' = 1$, será,

$$\left.\begin{array}{l} x = 7-5u-2 = 5-5u > 0 \\ y = -7+5u+3 = -4+5u > 0 \end{array}\right\} \Rightarrow \left\{\begin{array}{l} u < 1 \\ u > \frac{4}{5} \end{array}\right.$$

límites que son incompatibles por el hecho de que u debe ser un número entero.

Para $t' = 2$, será,

$$\left.\begin{array}{l} x = 7-5u-4 = 3-5u > 0 \\ y = -7+5u+6 = -1+5u > 0 \end{array}\right\} \Rightarrow \left\{\begin{array}{l} u < \frac{3}{5} \\ u > \frac{1}{5} \end{array}\right.$$

límites que son también incompatibles para el entero u.

Para $t' = 3$, se tendrá,

$$\left.\begin{array}{l} x = 7-5u-6 = 1-5u > 0 \\ y = -7+5u+9 = 2+5u > 0 \end{array}\right\} \Rightarrow \left\{\begin{array}{l} u < \frac{1}{5} \\ u > -\frac{2}{5} \end{array}\right\} \Rightarrow u = 0$$

Este único valor entero para $u = 0$ con $t' = 3$, dan

$$\begin{cases} x = 7-0-6 = 1 \\ y = -7+0+9 = 2 \end{cases}$$

Sustituyendo estos valores en las expresiones de z y v, encontraremos:

$$\begin{cases} z = 6x+4y+10u-t-6 = 6+8+0-t-6 = 8-t > 0 \\ v = -3x-2y-5u+t+3 = -3-4-0+t+3 = \\ \qquad = -4+t > 0 \end{cases}$$

De donde resulta,

$$t < 8, t > 4$$

es decir, $4 < t < 8$

por tanto, $t = 5, 6, 7$

Para $t = 5$: $x = 1, y = 2, z = 3, u = 0, v = 1$

Para $t = 6$: $x = 1, y = 2, z = 2, u = 0, v = 2$

Para $t = 7$: $x = 1, y = 2, z = 1, u = 0, v = 3$

que son soluciones enteras y positivas de la ecuación propuesta.

Procediendo de esta manera, podemos dar a la expresión $(3x + 2y + 5u)$ nuevos valores enteros mayores que 7, para obtener nuevas soluciones enteras y positivas.

9.11. Finalmente, si se da un sistema cualquiera de ecuaciones, con un número cualquiera también de incógnitas, pero que le exceda en más de una unidad, pasaremos a los segundos miembros el número de incógnitas suficientes para que en los primeros miembros no quede más que un número de incógnitas que sólo excede al de ecuaciones en una unidad.

En seguida aplicaremos al sistema resultante el método que ya hemos explicado anteriormente para tales sistemas de ecuaciones.

En resumen, si se tienen m ecuaciones con $m + n$ incógnitas, dejaremos en los primeros miembros las mismas $(m + 1)$ incógnitas, quedando las $m + n - (m + 1) = n - 1$ incógnitas restantes, como indeterminadas en los segundos miembros; y considerando estos segundos miembros como conocidos, estamos en presencia de un sistema simplemente indeterminado. Luego, los valores de esas $(m + 1)$ incógnitas vendrán en función de una indeterminada t y de las $(n - 1)$ incógnitas que se pasaron a los segundos miembros. Estas $(n - 1) + 1$ indeterminadas podrán recibir toda clase de valores enteros, para hallar las soluciones enteras del sistema propuesto.

Para las soluciones enteras y positivas, habrá que establecer y resolver, las desigualdades de condición correspondiente.

Ejemplo.

Sea el sistema,

$$\begin{array}{l} 5x+7y-4z+3u = 12 \\ 8x+2y+3z+4u = 36 \end{array}$$

Observamos que, haciendo $z = 2z'$, podremos dividir todos los términos de la segunda ecuación por 2, de modo que nuestro sistema se transformará en:

$$\begin{array}{l} 5x+7y-8z'+3u = 12 \\ 4x+y+3z'+2u = 18 \end{array}$$

Eliminando y, resulta por reducción:

$$23x + 29z' + 11u = 114,$$

y el último sistema se transformará en su equivalente:

$$\begin{array}{l} 4x+y+3z'+2u = 18 \\ 23x+29z'+11u = 114 \end{array}$$

La segunda de estas ecuaciones puede escribirse,

$$23x + 11u = 114 - 29z' = k$$

o sea, $23x + 11u = k, k + 29z' = 114$

La primera de estas últimas queda evidentemente satisfecha por $x = k$, $u = -2k$; luego, las fórmulas generales serán:

$$\begin{cases} x = k - 11t \\ u = -2k + 23t \end{cases}$$

y la segunda de estas mismas ecuaciones, admite una primera solución que es: $k = -2, z' = 4$; luego, las fórmulas generales serán:

$$\begin{cases} k = -2 - 29t' \\ z' = 4 + t' \end{cases}$$

Sustituyendo estos últimos valores en las fórmulas generales anteriores, resulta,

$$x = -2 - 29t' - 11t$$
$$u = 4 + 58t' + 23t'$$

y como, $z = 2z' = 8 + 2t'$

Sustituyendo, a su vez, estos valores de *x*, *u*, *z* en una de las ecuaciones del sistema transformado del propuesto, hallaremos el valor de *y*, y tendremos:

$$y = 18 - 4x - 3z' - 2u = 6 - 3t' - 2t$$

Luego, las relaciones,

$$\begin{cases} x = -2 - 29t' - 11t \\ y = 6 - 3t' - 2t \\ z = 8 + 2t' \\ u = 4 + 58t' + 23t \end{cases}$$

dan para toda clase de valores enteros de las indeterminadas t y t', las soluciones enteras del sistema propuesto.

Para las soluciones enteras y positivas, habrá que establecer y resolver, las desigualdades:

$$\begin{cases} x = -2 - 29t' - 11t > 0 \\ y = 6 - 3t' - 2t > 0 \\ z = 8 + 2t' > 0 \\ u = 4 + 58t' + 23t > 0 \end{cases}$$

De, $8 + 2t' > 0$, resulta $t' > -4$, o sea, $t' = 3, -2, -1, 0, 1, \ldots$

Para $t' = -3$, se obtendrá:

$$z = 8 - 6 = 2$$

y entonces,

$$x = -2 + 87 - 11t > o \Rightarrow t < 7\frac{8}{11}$$

$$y = 6 + 9 - 2t > o \Rightarrow t < 7\frac{1}{2}$$

$$u = 4 - 174 + 23t > o \Rightarrow t > 7\frac{9}{23}$$

límites contradictorios, a causa de que t debe ser entero.

Para $t' = -2$, será $z = 8 - 4 = 4$, y entonces:

$$x = -2 + 58 - 11t > o \Rightarrow t < 5\frac{1}{11}$$

$$y = 6 + 6 - 2t > o \Rightarrow t < 6$$

$$u = 4 - 16 + 23t > o \Rightarrow t > 4\frac{20}{23}$$

de donde, resulta para t la limitación,

$$4\frac{20}{23} < t < 5\frac{1}{11}$$

y el único valor para t será 5.

Por lo tanto, el sistema propuesto admite la solución entera y positiva siguiente:

$$x = 1,\ y = 2,\ z = 4,\ u = 3$$

Nótese que hemos resuelto el sistema propuesto de manera que ninguna de las incógnitas se considere como una indeterminada; esto es, todas las incógnitas se expresan en función entera de dos indeterminadas, t y t', que no son ninguna de las incógnitas.

Resolveremos nuevamente nuestro sistema de modo que una de las incógnitas desempeñe el papel de indeterminada.

Volvamos nuevamente al sistema:

$$\left.\begin{array}{l} 5x + 7y - 8z' + 3u = 12 \\ 4x + y + 3z' + 2u = 18 \end{array}\right|$$

transformado del propuesto mediante la substitución $z = 2z'$.

Eliminando, como antes, la incógnita *y*, tendremos:

$$23x + 29z + 11u = 114$$

y resolvámosla dejando una incógnita indeterminada.

$$23x + 11u = 114 - 29z' = k$$

ecuación que queda evidentemente satisfecha por:

$$x = k,\ u = -2k$$

Luego, las fórmulas generales serán:

$$\begin{cases} x = k - 11t = 114 - 29z' - 11t \\ u = -2k + 23t = -228 + 58z' + 23t \end{cases}$$

Sustituyendo estos valores en la segunda de las ecuaciones del sistema transformado, quedará para *y* el valor:

$$y = 18 - 4(114 - 29z' - 11t) - 3z' - 2(-228 + 58z' + 23t)$$

$$y = 18 - 3z' - 2t$$

Luego, las relaciones siguientes:

$$\begin{cases} x = 114 - 29z' - 11t \\ y = 18 - 3z' - 2t \\ u = -228 + 58z' + 23t \end{cases}$$

darán valores enteros de estas tres incógnitas x, y, u, cuando demos valores enteros a la indeterminada t y a la incógnita z' que hace la función de indeterminada. Cada conjunto de valores x, y, u, z' que así resulta, será una solución entera del sistema propuesto. Recuérdese que $z = 2z'$.

Para las soluciones enteras y positivas, habrá que establecer y resolver las desigualdades siguientes:

$$x = 114 - 29z' - 11t > o \Rightarrow t < \frac{114 - 29z'}{11}$$

$$y = 18 - 3z' - 2t > o \Rightarrow t < \frac{18 - 3z'}{2}$$

$$u = -228 + 58z' + 23t > o \Rightarrow t > \frac{228 - 58z'}{23}$$

Para que haya compatibilidad, tendrá que verificarse simultáneamente,

$$\frac{114 - 29z'}{11} > \frac{228 - 58z'}{23}$$

$$\frac{18 - 3z'}{2} > \frac{228 - 58z'}{23}$$

$$\text{osea, } \left.\begin{matrix} 29z' < 114 \\ 47z' > 42 \end{matrix}\right\} \Rightarrow \begin{cases} z' < 3\frac{27}{29} \\ z' > \frac{42}{47} \end{cases}$$

cuyos valores enteros positivos posibles son: z' = 1, 2, 3.

Para z' = 1, resulta:

$$x = 114 - 29 - 11t > o \Rightarrow t < 7\frac{8}{11}$$

$$y = 18 - 3 - 2t > o \Rightarrow t < 7\frac{1}{2}$$

$$u = -228 + 58 + 23t > o \Rightarrow t > 7\frac{9}{23}$$

límites contradictorios, porque t debe ser entero.

Para z' = 2, se tendrá:

$$x = 114 - 58 - 11t > o \Rightarrow t < 5\frac{1}{11}$$

$$y = 18 - 6 - 2t > o \Rightarrow t < 6$$

$$u = -228 + 116 + 23t > o \Rightarrow t > 4\frac{20}{23}$$

Luego, resulta que,

$$4\frac{20}{23} < t < 5\frac{1}{11} \Rightarrow t = 5$$

Entonces, tendremos:

$$x = 1,\ y = 2,\ u = 3,\ z = 2,\ z' = 4$$

la solución entera y positiva, que antes habíamos obtenido por el primer procedimiento.

Creemos que con todo lo ya visto sobre "Análisis indeterminado del primer grado" es suficiente para abordar el estudio de esta parte tan interesante y útil del Algebra.

9.12. *Inecuaciones de Segundo Grado con una incógnita*

Llamaremos inecuación de segundo grado con una incógnita, a toda relación de la forma:

$$ax^2 + bx + c > o, \text{ ó también } ax^2 + bx + c < o.$$

Resolver una desigualdad de este tipo, es encontrar entre qué límites puede variar la incógnita x para que la desigualdad se verifique.

Un método bastante sencillo reposa sobre el estudio de la variación del signo del trinomio de segundo grado, que vimos en el Tomo I), Capítulo IV) de estos apuntes.

Luego, a este respecto conviene tener presente que:

Cuando un trinomio de segundo grado tiene sus raíces reales y distintas, entonces él toma un valor numérico del signo de su primer coeficiente para todos los valores de x exteriores al intervalo de las raíces, y un valor numérico de signo contrario al de su primer coeficiente para todos los valores de x interiores al intervalo de las raíces.

Pero, si las raíces son reales e iguales o imaginarias (complejas conjugadas), entonces él toma siempre el signo de su primer coeficiente, cualquiera sea el valor que se da a x.

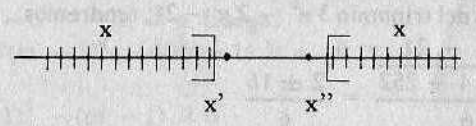
Por consiguiente, al resolver la inecuación, por ejemplo,

$$ax^2 + bx + c > o$$

habrá que considerar dos casos, según que las raíces del trinomio $ax^2 + bx + c$, sean reales y distintas, y reales e iguales o imaginarias.

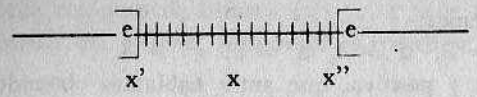
Así, pues, si las raíces x', x'' son reales diferentes y si $a > o$, entonces la inecuación $ax^2 + bx + c > o$ se satisfará para todos los valores de x exteriores al intervalo x', x'', esto es, para todo x tal que:

$$x < x',\ \text{ó}\ x > x''$$



Pero, si $a < o$, entonces la inecuación se satisfará para todo x tal que:

$$x' < x < x''$$



En cambio, si las raíces x' y x'', son reales iguales o imaginarias, y si $a > o$, entonces la inecuación $ax^2 + bx + c > o$ se satisfará para cualquier valor dado ax; y si $a < o$, la inecuación $ax^2 + bx + c > o$ no se verificará jamás.

De una manera análoga se resuelve la inecuación,

$$ax^2 + bx + c < o$$

con un razonamiento totalmente al revés de lo expresado anteriormente.

Veamos algunos ejemplos.

1. Sea la inecuación $x^2 - 6x + 5 > o$.

Las raíces del trinomio $x^2 - 6x + 5$ son: $x' = 1$, $x'' = 5$.

Como el primer coeficiente es $a = 1 > o$, entonces la inecuación propuesta se verifica para todo $x < 1$ y para todo $x > 5$.

Observación. Otro método bastante simple consiste en factorizar el trinomio dado en función de sus factores radicales, que en nuestro ejemplo es: $(x - 1)(x - 5)$.

Ahora bien, como debe ser

$$(x - 1)(x - 5) > o$$

entonces el producto de dos factores es positivo si, y sólo si, ambos son positivos o ambos negativos, tendremos:

$$\begin{cases} x - 1 > o \\ x - 5 > o \end{cases} \Rightarrow \begin{cases} x > 1 \\ x > 5 \end{cases} \Rightarrow x > 5$$

y

$$\begin{cases} x - 1 < o \\ x - 5 < o \end{cases} \Rightarrow \begin{cases} x < 1 \\ x < 5 \end{cases} \Rightarrow x < 1$$

Luego, la solución de la inecuación dada es:

$$x < 1, \text{ ó } x > 5$$

que es la misma que obtuvimos con el primer método.

2. Sea la inecuación $3x^2 - 2x - 21 < o$.

Para las raíces del trinomio $3x^2 - 2x - 21$, tendremos:

$$3x^2 - 2x - 21 = o$$

$$x = \frac{2 \pm \sqrt{4 + 252}}{6} = \frac{2 \pm 16}{6}$$

$$x' = -\frac{7}{3} = -2\frac{1}{3}, \quad x'' = 3$$

Luego, la inecuación propuesta se verifica para todos los valores de x tales que,

$$-2\frac{1}{3} < x < 3$$

ya que el trinomio es negativo solamente para valores comprendidos en el intervalo de sus raíces.

Por el método de la descomposición factorial, la inecuación,

$$3x^2 - 2x - 21 < o$$

se resuelve así:

$$3(x + \frac{7}{3})(x - 3) < o$$

y como el producto de dos factores es negativo sí, y sólo si, ambos son de distinto signo, resultará:

$$\begin{cases} x + \frac{7}{3} > o \\ x - 3 < o \end{cases} \Rightarrow \begin{cases} x > -\frac{7}{3} = -2\frac{1}{3} \\ x < 3 \end{cases} \Rightarrow -2\frac{1}{3} < x < 3$$

y

$$\begin{cases} x + \frac{7}{3} < o \\ x - 3 > o \end{cases} \Rightarrow \begin{cases} x < -\frac{7}{3} = -2\frac{1}{3} \\ x > 3 \end{cases} \quad \text{(Incompatible)}$$

Luego, el único caso que puede darse corresponde a la limitación,

$$-2\frac{1}{3} < x < 3$$

ya encontrada por el primer procedimiento.

3. Las inecuaciones tales como:

$$ax^2 + bx + c > k, \text{ ó } ax^2 + bx + c < k$$

se resuelven como las anteriores, pasando previamente al primer miembro la cantidad constante k.

4. Dada la ecuación,

$$x^2 - 2(m - 1)x + (m - 1) = o$$

encontrar los límites en que debe variar el parámetro m para que las raíces x', x'' de la ecuación sean reales.

Sol. Para que las raíces sean siempre reales es necesario que la discriminante $\Delta = b^2 - 4ac \geq o$. Luego,

$$\Delta = 4(m - 1)^2 - 4(m - 1) \geq o$$

o bien,

$$(m - 1)^2 - (m - 1) \geq o$$

$(m - 1)(m - 1 - 1) \geq o$

$(m - 1)(m - 2) \geq o$

De donde se concluye que:

$m \leq 1$, ó $m \geq 2$

5. Resolver la desigualdad,

$$\frac{3x - 5}{x + 2} > o$$

Multiplicando ambos miembros por la cantidad positiva $(x + 2)^2$, el problema se reduce a resolver la desigualdad equivalente,

$(3x - 5)(x + 2) > o$

o bien,

$3(x - \frac{5}{3})(x + 2) > o$

Esta inecuación se satisfará para,

$x < -2$, o $x > \frac{5}{3}$

Si la desigualdad es de la forma,

$$\frac{3x - 5}{x + 2} < o$$

se obtendrá, al multiplicarla por $(x + 2)^2$

$(3x - 5)(x + 2) < o$

$3(x - \frac{5}{3})(x + 2) < o$

inecuación que ahora se satisface para,

$$-2 < x < \frac{5}{3}$$

6. Resolver la inecuación,

$$\frac{x^2 - 3x + 2}{x^2 + 2x - 15} > o$$

Multiplicando ambos miembros por la cantidad positiva $(x^2 + 2x - 15)^2$, el problema se reduce a resolver la inecuación equivalente,

$(x^2 - 3x + 2)(x^2 + 2x - 15) > o$

Estudiaremos en seguida el signo de cada trinomio factor separadamente.

1er Trinomio: $x^2 - 3x + 2 = (x - 1)(x - 2)$

2° Trinomio: $x^2 + 2x - 15 = (x - 3)(x + 5)$

Las raíces: 1, 2, 3, −5, ordenadas en orden de magnitud creciente, se escriben,

$-5, 1, 2, 3$

Formemos el siguiente cuadro:

x	Signo 1er Trinomio	Signo 2° Trinomio	Conclusiones
$-\infty$			
,			
,	+	+	valores que convienen al problema
,			
−5			
,			
,	+	−	
,			
1			
,			
,	−	−	valores que convienen al problema
,			
2			
,			
,	+	−	
,			
3			
,			
,	+	+	valores que convienen al problema
,			
$+\infty$			

Luego, la inecuación propuesta se satisface para:

$x < -5$, $x > 3$, $1 < x < 2$

Este mismo método puede ser empleado para la resolución de inecuaciones simultáneas. Veamos algunos ejemplos.

7. Resolver el sistema simultáneo,

$$\left.\begin{array}{l} 6x + 11 > o \\ x^2 - 7x + 6 > o \end{array}\right|$$

La primera inecuación tiene por soluciones los números,

$$x > -\frac{11}{6} = -1\frac{5}{6}$$

La segunda inecuación de segundo grado tiene por raíces 1 y 6. Escribiendo estos números por orden de magnitud creciente, formaremos la siguiente tabla:

x	Signo 1ª Inecuación	Signo 2ª Inecuación	Conclusiones
$-\infty$, ,			
, ,	−	+	
$-1\frac{5}{6}$, ,			
, ,	+	+	Conviene al problema
1 , ,			
, ,	+	−	
6 , ,			
, ,	+	+	Conviene al problema
$+\infty$			

Luego, los valores de x que verifican al sistema propuesto son:

$$-1\frac{5}{6} < x < 1, \quad x > 6$$

8. Resolver el sistema,

$$\begin{aligned} 3x^2 + x - 14 &> 0 \\ x^2 - 9x + 18 &< 0 \\ x^2 - 6x + 5 &< 0 \end{aligned}$$

Las raíces de estos trinomios son, respectivamente:

$$2 \text{ y } -\frac{7}{3}, 6 \text{ y } 3; \quad 5 \text{ y } 1$$

Ordenando en magnitud creciente estas seis raíces, formamos la siguiente tabla que nos dará los signos de estos tres trinomios, cuando la variable x crece en el intervalo $]-\infty, +\infty[$

x	Signo 1er Trinomio	Signo 2° Trinomio	Signo 3er Trinomio	Conclusiones
$-\infty$, ,				
, ,	+	+	+	
$-\frac{7}{3}$, ,				
, ,	−	+	+	
1 , ,				
, ,	−	+	−	
2 , ,				
, ,	+	+	−	
3 , ,				
, ,	+	−	−	Valores que convienen al sistema
5 , ,				
, ,	+	−	+	
6 , ,				
, ,	+	+	+	
$+\infty$				

Luego, el sistema propuesto se satisface para todos los valores de x tales que: $3 < x < 5$.

9.13. *Aplicación de las Inecuaciones a la discusión de las ecuaciones de segundo grado*

1) Dada la ecuación,

$$P(x) = 4x^2 + (m-2)x + (m-5) = o$$

donde m designa un indeterminado (parámetro).

¿Qué valores puede tomar m para que las raíces de esta ecuación sean reales diferentes y menor que 2?

Sol. Para que las raíces sean reales y distintas deberá tenerse:

$$\Delta = b^2 - 4ac = (m-2)^2 - 16(m-5) > o$$
$$m^2 - 4m + 4 - 16m + 80 > o$$
$$m^2 - 20m + 84 > o$$
$$(m-6)(m-14) > o$$
$$m < 6, \quad \text{ó} \quad m > 14$$

Por otra parte, como el primer coeficiente $a = 4 > o$, el trinomio primer miembro de la ecuación propuesta tendrá fuera del intervalo de sus raíces el signo positivo, y por esto deberá cumplirse también la condición adicional siguiente:

$$P(2) = 16 - 2(m-2) + (m-5) > o$$

ya que ambas raíces deben ser menores que 2.

Resolviendo esta última inecuación, se tendrá:

$$m > -\frac{7}{3}$$

que unida con las condiciones ya encontradas $x < 6$ y $m > 14$, dan la respuesta a la pregunta. De modo que, el parámetro debe cumplir las condiciones:

$$-\frac{7}{3} < m < 6, \quad \text{ó} \quad m > 14$$

2) Determinar los valores del parámetro m en la ecuación,

$$(m-2)x^2 + 2(m-1)x + m - 3 = o$$

para que sus raíces sean positivas.

Sol. Como las raíces deben ser reales, una primera condición que debe cumplir m es:

$$\Delta = b^2 - 4ac = 4(m-1)^2 - 4(m-2)(m-3) \geq o$$
$$\text{o sea, } m^2 - 2m + 1 - m^2 + 5m - 6 \geq o$$
$$3m - 5 \geq o$$
$$m \geq \frac{5}{3} = 1\frac{2}{3}$$

Por otra parte, si las raíces han de ser positivas, entonces tanto su producto P como su suma S deben ser también positivas. Luego, las otras condiciones que debe cumplir también m son:

$$P = \frac{m-3}{m-2} > o \ , \ S = -\frac{2(m-1)}{m-2} > o$$

o sea,

$$(m-2)(m-3) > o, \quad (m-1)(m-2) < o$$

y de donde resulta:

$$m < 2 \text{ y } m > 3, \ 1 < m < 2$$

Estas condiciones unidas con la condición ya encontrada $m \geq \frac{5}{3}$ dan la solución del problema; esto es,

$$\frac{5}{3} \leq m < 2$$

A este mismo resultado se habría llegado, utilizando el siguiente cuadro:

m	Δ	P	S	Conclusión sobre las raíces
$-\infty$				
,				
,	−	+		Raíces imaginarias
,				
1				
,				
,	−	+		Raíces imaginarias
,				
5/3				
,				
,	+	+	+	Raíces positivas
,				
2				
,				
,	+	−	−	Raíces de signos contrarios
,				
3				
,				
,	+	+	−	Raíces negativas
$+\infty$				

Antes de continuar con algunas otras aplicaciones de las inecuaciones a la discusión de las raíces de una ecuación de segundo grado con un parámetro, veremos primeramente el siguiente problema fundamental.

Problema. Dada la ecuación de segundo grado,

$$f(x) = a x^2 + b x + c = o$$

encontrar las condiciones para que un número dado k sea:

1. Superior a las dos raíces de la ecuación,
2. Inferior a estas dos raíces,
3. Comprendido entre estas raíces.

Sol. Para que el número dado k sea mayor que las raíces de la ecuación, es necesario que estas raíces existan, es decir que se tenga $\Delta = b^2 - 4ac > o$.

Cumplida esta primera condición, para que el número k sea superior a las raíces $x' < x''$, se exige ahora que f(k) tenga el mismo signo que el primer coeficiente *a*, esto es que se tenga: $a f(k) > o$.

Cumplidas estas dos condiciones, tendremos:

$$x' < k$$
$$x'' < k$$

de donde, $x' + x'' < 2k$, o sea, $k > \frac{x' + x''}{2} = -\frac{b}{2a}$

Con lo cual hemos dado respuesta a la pregunta 1).

Ahora, si el número dado k es inferior (menor) que las dos raíces $x' < x''$, entonces deberá ser f (k) también del mismo signo que el coeficiente *a*, es decir: $af(k) > 0$

Luego tendremos:

$$k < x'$$
$$k < x''$$

de donde, $2k < x' + x'$, o bien, $k < \frac{x' + x''}{2} = -\frac{b}{2a}$

Con lo que damos respuesta a la pregunta 2).

Finalmente, para que el número dado k esté comprendido entre las raíces $x' < x''$, es necesario ahora que f (k) sea de signo contrario al coeficiente *a*, es decir que se tenga: $f(k) < 0$.

Esta sola condición es suficiente, porque a $f(k) < 0$ expresa que las raíces son reales.

Por último veremos el siguiente teorema, que sólo enunciamos, y que nos será sumamente útil.

Teorema. Si en una ecuación con coeficientes reales $f(x) = 0$, se sustituyen en lugar de x, dos números reales α y β ($\alpha < \beta$), entre ambos habrá un número impar de raíces reales distintas si los dos resultados $f(\alpha)$ y $f(\beta)$ son de distinto signo, y un número par de raíces (que puede ser cero) si son del mismo signo.

En otros términos:

Si $f(\alpha) \cdot f(\beta) < 0$, entonces existe entre α y β un número impar de raíces reales de la ecuación $f(x) = 0$.

Si $f(\alpha)\ f(\beta) > 0$, entonces existe entre α y β un número par de raíces reales de la ecuación $f(x) = 0$, o no existe ninguna raíz.

En particular, para la ecuación de segundo grado,

$$f(x) = a x^2 + b x + c = 0$$

tenemos:

1°. Si los resultados $f(\alpha)$ y $f(\beta)$ son de signos contrarios, entonces la ecuación tiene raíces reales distintas, o una de ellas está comprendida entre los números α y β.

2°. Si los resultados $f(\alpha)$ y $f(\beta)$ son del mismo signo, entonces los números α y β comprenden un número par de raíces reales distintas de la ecuación; es decir, cero o dos.

Ejemplos.

1) Determinar los valores de m, para los cuales las raíces de la ecuación,

$$2x^2 + (m - 3)x + 3 - m = 0$$

estén comprendidas entre (-2) y $(+3)$.

Sol. Sea $x' < x''$ las raíces y tales que:

$$-2 < x' < x'' < 3$$

Tendrán que cumplirse las siguientes condiciones:

a) $\Delta = b^2 - 4ac = (m - 3)^2 - 8(3 - m) > 0$

b) $\frac{S}{2} = \frac{x' + x''}{2} = -\frac{m - 3}{4}$ tal que,

$$-2 < -\frac{m - 3}{4} < 3$$

c) $f(-2) \cdot f(-3) = [8 - 2(m - 3) + 3 - m]\,[18 + 3(m - 3) + 3 - m] > 0$

Resolviendo estas tres desigualdades, tendremos:

a) $(m - 3)^2 - 8(3 - m) > 0$

$(m - 3)^2 + 8(m - 3) > 0$

$(m - 3)(m - 3 + 8) > 0$

$(m - 3)(m + 5) > 0$

luego, $m < -5$, $m > 3$, para que las raíces sean reales y distintas.

b) $-2 < -\frac{m-3}{4} < 3$

$-8 < -m + 3 < 12$

$-11 < -m < 9$

o bien, $-9 < m < 11$

para que la semisuma de las raíces esté comprendida entre los dos números dados (-2) y $(+3)$.

c) $f(-2) \cdot f(3) = (-3m + 17)(2m + 12) > 0$

o sea, $-6\,(m - \frac{17}{3})(m + 6) > 0$ luego, $-6 < m < \frac{17}{3} = 5\frac{2}{3}$

m	Δ	f(−2) · f(3)	Conclusiones
$-\infty$			
	+	−	
-9			
	+	−	
-6			
	+	+	Convienen al problema
-5			
	−	+	
3			
	+	+	Convienen al problema
$5\frac{2}{3}$			
	+	−	
11			
	+	−	
$+\infty$			

Ordenando estos límites de m en orden de magnitud creciente, tendremos la serie:

$$-9,\ -6,\ -5,\ 3,\ 5\frac{2}{3},\ 11$$

Luego, podremos observar el cuadro de signos anterior.

Luego, se satisfacen las condiciones pedidas si:

$$-6 < m < -5,\ 3 < m < 5\frac{2}{3}$$

2) Dada la ecuación,

$f(x) = (12m + 7)x^2 - 3(14 - 3m)x + 11 - 3m = 0$

hallar los valores de m para que:

a) Las raíces sean reales,

b) La unidad esté comprendida entre las raíces.

Sol. a) $\Delta = b^2 - 4ac = 9(14 - 3m)^2 - 4(12m + 7)(11 - 3m) \geq 0$

Reduciendo, obtendremos:

$225m^2 - 1200m + 1456 \geq 0$

$225\,(m - \frac{28}{15})(m - \frac{52}{15}) \geq o$

de donde, $m < \frac{28}{15}, m > \frac{52}{15}$

b) Para que la unidad 1 esté comprendida entre las raíces, es necesario que a $f(1) = (12m + 7)(18m - 24) < o$

o sea, $12 \cdot 18\,(m + \frac{7}{12})(m - \frac{24}{18}) < o$

De donde, $-\frac{7}{12} < m < \frac{24}{18} = \frac{4}{3}$

Luego, tendremos el cuadro de signos:

m	Δ	a f(1)	Conclusiones
$-\infty$	+	+	
$-\frac{7}{12}$	+	−	Valores que convienen

m	Δ	a f(1)	Conclusiones
,			
,	+	−	Valores que convienen
,			
$\frac{4}{3}$			
,			
,	+	+	
,			
$\frac{28}{15}$			
,			
,			
,	−		Raíces imaginarias
$\frac{52}{15}$			
,			
,	+	+	
,			
$+\infty$			

Por lo tanto, los valores de m que satisfacen las condiciones del problema son: $-\frac{7}{12} < m < \frac{4}{3}$

3) Dada la ecuación:

$x^2 - 2mx - (1 - m^2) = 0$

determinar los valores del parámetro m para que las raíces estén comprendidas entre (-2) y $(+4)$.

Sol. Este problema es análogo al 1), pero será ahora resuelto por otro razonamiento que es muy útil conocer.

Sean $x' < x''$ las dos raíces de la ecuación, supuestas reales y distintas; esto es, que se verifique:

$\Delta = b^2 - 4ac = 4m^2 + 4(1 - m^2) = 4 > 0$

Luego, las raíces son siempre reales en este ejemplo.

Ahora, aplicando el Problema Fundamental visto al comienzo de este párrafo, expresaremos las condiciones para que una sola raíz esté comprendida entre (-2) y $(+4)$.

Para esto formaremos las expresiones:

$f(-2) = m^2 + 4m + 3$, $f(4) = m^2 - 8m + 15$

y en seguida las expresiones: a $f(-2)$ y a $f(4)$, y como a $= 1$, resultan ser las mismas $f(-2)$ y $f(4)$, respectivamente.

Luego, para que (-2) sea inferior a la raíz x', debe ser:

$a\,f(4) = f(-2) > 0$

y para que $(+4)$ sea superior a la raíz x', debe ser:

$a\,f(4) = f(4) > 0$

Por lo tanto, tendremos las desigualdades:

$f(-2) = m^2 + 4m + 3 = (m + 3)(m + 1) > 0$

$f(+4) = m^2 - 8m + 15 = (m - 3)(m - 5) > 0$

De donde resultan las condiciones:

$$\begin{cases} m < -3 \\ m > -1 \end{cases} \quad ; \quad \begin{cases} m < 3 \\ m > 5 \end{cases}$$

m	Δ	a f(−2)	a f(4)	*Conclusiones*
$-\infty$				
	+	+	+	
− 3				
	+	−	+	Raíces separadas por (-2)
− 1				
	+	+	+	
3				
	+	+	−	Raíces separadas por $(+4)$
5				
	+	+	+	
$+\infty$				

Por otra parte, como ambas raíces deben ser menores que $(+4)$, su semisuma $\frac{S}{2} = \frac{x' + x''}{2} = m$ debe ser también menor que $(+4)$, es decir $m < 4$; del mismo modo también mayor que (-2), o sea $m < -2$. Por lo lo tanto, el intervalo útil es solamente:

$$-1 < m < 3$$

Los otros dos intervalos: $]-3, -1[$ y $]3, 5[$ no comprenden más que una sola raíz.

9.14. *Valor absoluto.* En un campo ordenado, llamaremos VALOR ABSO-

LUTO de un elemento cualquiera x, y se lo representa por $|x|$, al elemento definido de la siguiente manera:

$$|x| = \begin{cases} x, & \text{si } x \geq o \\ -x, & \text{si } x < o \end{cases}$$

Es consecuencia inmediata de la definición de $|x|$, que:

$-|x| \leq x \leq |x|$, cualquiera sea x.

Las principales propiedades del valor absoluto, son enunciadas en los teoremas que siguen:

Teorema 1) Para cualquier elemento x del campo ordenado, se tiene:

1. $|x| \geq o$; $|x| = o$ sí y sólo si, $x = o$
2. $|x|^2 = x^2$
3. $\sqrt{x^2} = |x|$

Dem. La 1) se deduce de la definición de valor absoluto y del hecho de que si x es un elemento del campo ordenado, entonces una y sólo una de las siguientes relaciones es cierta:

$$x > o, \quad x = o, \quad x < o$$

Luego, si $x > o$ es $|x| = x > o$, si $x = o$ es $|x| = o$ y si $x < o$ es $|x| = -x > o$.

En todos los casos es $|x| \geq o$; es decir, el valor absoluto de x nunca es negativo.

La 2) se prueba de la manera siguiente:

$$|x|^2 = |x| \cdot |x|$$

Si $x \geq o$, entonces $|x|^2 = x \cdot x = x^2$; si $x < o$, entonces $\frac{3}{4}$

$$|x|^2 = |x| \cdot |x| = (-x) \cdot (-x) = x^2$$

Luego, en todos los casos es $|x|^2 = x^2$

Finalmente, la 3) es un caso particular del campo ordenado de los números reales.

En los próximos capítulos, al estudiar los números reales, veremos que si *a* es un número real no negativo, entonces existe el número real *no negativo b* con la propiedad de que $b^2 = a$. Llamaremos a este número *b* con el nombre de *raíz cuadrada principal* de *a*, y lo denotaremos por $\sqrt{a}$.

Por consiguiente, tendremos: $\sqrt{x^2} = x$ si $x \geq o$, y $\sqrt{x^2} = -x$ si $x < o$.

Vemos, pues, que $\sqrt{x^2}$ se define por:

$$\sqrt{x^2} = \begin{cases} x, \text{ si } x \geq o \\ -x, \text{ si } x < o \end{cases}$$

Es decir, $\sqrt{x^2}$ satisface la definición de $|x|$.

Por lo tanto hemos demostrado que $\sqrt{x^2} = |x|$

Este resultado y los anteriores prueban el teorema.

Teorema 2) Si $a \geq o$, entonces:

1. $|x| \leq a$ si, y sólo si, $-a \leq x \leq a$
2. $|x| \geq a$, sí, y sólo si, $x \leq -a$, ó, $x \geq a$

Dem. 1. De la definición de valor absoluto, se tiene,

$$-|x| \leq x \leq |x|$$

Supongamos que sea $|x| \leq a$, de donde $-|x| \geq -a$.

Luego, $-a \leq -|x| \leq x \leq |x| \leq a$

esto es, $-a \leq x \leq a$

Así hemos demostrado el "sólo si".

Recíprocamente, supongamos que $-a \leq x \leq a$

Entonces, si $x > o$, tenemos $|x| = x \leq a$; mientras que si $x < o$, tenemos $|x| = -x \leq a$, ya que $-a \leq x \Rightarrow a \geq -x = |x|$.

En ambos casos tenemos $|x| \leq a$, lo que prueba el »sí«.

Este resultado y el anterior, prueban la equivalencia:

$$|x| \leq a \iff -a \leq x \leq a$$

2. Supongamos que $|x| \geq a$. Entonces:

Si $x > o$, tenemos $x = |x| \geq a$, mientras que si $x < o$, tenemos $-x = |x| \geq a$, o sea, $x \leq -a$.

Luego, $|x| \geq a$ implica $x \leq -a$, ó $x \geq a$, y así hemos demostrado el »sólo si«.

Inversamente, supongamos ahora $x \leq -a$, ó $x \geq a$.

Entonces, si $x > o$, tenemos $|x| = x \geq a$, mientras que si $x < o$, tenemos $|x| = -x \geq a$, ya que si $x \leq -a$ entonces $-x \geq a$. En ambos casos tenemos $|x| \geq a$, lo que demuestra el »sí«.

Este resultado y el anterior, prueban la equivalencia:

$$|x| \geq a \Longleftrightarrow x \leq -a, \text{ ó } x \geq a,$$

y el teorema queda demostrado.

Teorema 3) El valor absoluto tiene las propiedades:

1. $|a + b| \leq |a| + |b|\,; |a - b| \leq |a| + |b|$
2. $||a| - |b|| \leq |a - b|$
3. $|a \cdot b| = |a| \cdot |b|$
4. $\left|\frac{a}{b}\right| = \frac{|a|}{|b|}$, si $b \neq o$

Dem. 1. Por definición sabemos que:

$$-|a| \leq a \leq |a| \quad y \quad -|b| \leq b \leq |b|$$

Sumando miembro a miembro estas desigualdades que se verifican en el mismo sentido, se obtiene:

$$-|a| - |b| \leq a + b \leq |a| + |b|$$

o bien,

$$-(|a| + |b|) \leq a + b \leq (|a| + |b|)$$

Luego, en virtud del teorema 2), concluímos que,

$$|a + b| \leq |a| + |b|$$

La igualdad, $|a + b| = |a| + |b|$, tiene lugar únicamente cuando los elementos *a* y *b* son del mismo signo. Así, por ejemplo, en el campo ordenado de los números reales, tenemos:

$$|3 + 5| = 8 = |3| + |5|$$
$$|-3 - 5| = 8 = |-3| + |-5|$$
$$|3 - 5| = 2 \text{ y } |3| + |-5| = 8 \Rightarrow |3 - 5| < |3| + |5|$$

Por otra parte, si en la relación $|a + b| \leq |a| + |b|$ sustituímos b por $(-b)$, resulta:

$$|a - b| \leq |a| + |-b|$$

o bien, $|a - b| \leq |a| + |b|$, ya que $|-b| = |b|$

En otros términos, hemos probado que el valor absoluto de la suma o diferencia entre dos elementos es menor o igual a la suma de los valores absolutos de estos elementos.

Es evidente que la parte 1) del teorema se generaliza para una suma algebraica cualquiera, esto es:

$$|a + b - c + d - e - f| \leq |a| + |b| + |c| + |d| + |e| + |f|$$

2) Podemos escribir,

$$a = b + (a - b)$$

y aplicando la propiedad 1) recién demostrada, se tiene:

$$|a| \leq |b| + |a - b|$$

De aquí resulta:

$$|a| - |b| \leq |a - b| \qquad (1)$$

Análogamente obtendremos:

$$b = a + (b - a)$$
$$|b| \leq |a| + |b - a|$$
$$|b| - |a| \leq |b - a|$$

o bien, $|a| - |b| \geq -|b - a|$

y como, $|b - a| = |a - b|$, resulta

$$|a| - |b| \geq -|a - b| \qquad (2)$$

De (1) y (2), resulta:

$$-|a - b| \leq |a| - |b| \leq |a - b|$$

y, por Teorema 2) concluímos que,

$$||a| - |b|| \leq |a - b|$$

3) Por el Teorema 1), podemos escribir:

$$|a \cdot b| = \sqrt{(ab)^2} = \sqrt{a^2 \cdot b^2} = \sqrt{a^2} \cdot \sqrt{b^2} = |a|\ |b|$$

En general,

$$|a \cdot b \cdot c \cdot d \ldots| = |a| \cdot |b| \cdot |c| \cdot |d| \ldots$$

En otros términos, el valor absoluto de un producto es igual al producto de los valores absolutos de los factores.

4) Podemos escribir,

$$a = \frac{a}{b} \cdot b, \qquad b \neq o$$

Luego, por la propiedad 3) recién probada, se tiene,

$$|a| = \left|\frac{a}{b}\right| \cdot |b|$$

de donde,

$$\left|\frac{a}{b}\right| = \frac{|a|}{|b|}$$

En otras palabras, el valor absoluto de un cociente con denominador no nulo es igual al cociente de los valores absolutos de sus términos.

Este resultado y los anteriores, demuestran el teorema.

9.15. *Aplicaciones del valor absoluto*

Resolver las ecuaciones e inecuaciones siguientes:

1) $|x - 3| = 5$

Tenemos: $x - 3 = \pm 5 \Rightarrow x - 3 = 5$ y/o $x - 3 = -5$

$x = 8, \quad x = -2$

2) $|3x + 2| = 5 - x$

Por definición de valor absoluto, debe ser,

$5 - x \geq 0 \Rightarrow x \leq 5$

y también, $3x + 2 = 5 - x$ y/o $3x + 2 = -5 + x$

$$x = \frac{3}{4} \quad | \quad x = -\frac{7}{2}$$

soluciones que satisfacen la condición $x \leq 5$.

3) $|x + 4| = |x + 2|$

Tenemos: $x + 4 = x + 2$, ó $x + 4 = -x - 2$

$x = -3$ (única solución)

Otro método es el siguientes:

$\sqrt{(x+4)^2} = \sqrt{(x+2)^2}$

Elevando al cuadrado, resulta:

$(x + 4)^2 = (x + 2)^2$

$x^2 + 8x + 16 = x^2 + 4x + 4$

$4x = -12$

$x = -3$

4) $|x - 2| < 3$

Se tiene, $-3 < x - 2 < 3$

$-1 < x < 5$

5) $|5x| < 2$

Sol. $-2 < 5x < 2$

$$-\frac{2}{5} < x < \frac{2}{5}$$

6) $|2x - 3| < 5$

Sol. $-5 < 2x - 3 < 5$

$-2 < 2x < 8$

$-1 < x < 4$

7) $|3x - 5| > 4$

Sol. $3x - 5 < -4$ y/o $3x - 5 > 4$

$3x < 1$, $3x > 9$

$x < \frac{1}{3}$, $x > 3$

8) $|x - \frac{1}{4}| \geq \frac{3}{4}$

Sol. $x - \frac{1}{4} \leq -\frac{3}{4}$ y/o $x - \frac{1}{4} \geq \frac{3}{4}$

$x \leq -\frac{1}{2}$ $x \geq 1$

9) $\left|\frac{x+2}{3-x}\right| < 1$

Sol. $-1 < \frac{x+2}{3-x} < 1$

La cual se descompone en las dos condiciones simultáneas:

$\frac{x+2}{3-x} > -1$ y $\frac{x+2}{3-x} < 1$

Resolviendo estas inecuaciones, se encontrará:

$\frac{x+2}{3-x} + 1 > 0$, $\frac{x+2}{3-x} - 1 < 0$

$\frac{5}{3-x} > 0$, $\frac{2x-1}{3-x} < 0$

$3 - x > 0$, $(2x - 1)(3 - x) < 0$

$x < 3$ $2(x - \frac{1}{2})(x - 3) > 0$

$x < \frac{1}{2}$, $x > 3$

Luego, debe ser $x < \frac{1}{2}$ como solución

10) Resolver la inecuación,

$$\left|\frac{x^2 - 2x + 3}{x^2 - 5x + 6}\right| > \frac{1}{5}$$

Tendremos:

$$-5 < \frac{x^2 - 5x + 6}{x^2 - 2x + 3} < 5$$

y lo que da orígen a las dos inecuaciones simultáneas:

$\frac{x^2 - 5x + 6}{x^2 - 2x + 3} + 5 > 0$ y $\frac{x^2 - 5x + 6}{x^2 - 2x + 3} - 5 < 0$

o sea,

$$\frac{6x^2 - 15x + 21}{x^2 - 2x + 3} > 0 \quad \text{y} \quad \frac{-4x^2 + 5x - 9}{x^2 - 2x + 3} < 0$$

Puede observarse que los cuatro trinomios indicados en ambas inecuaciones tienen sus raíces imaginarias. Por lo tanto, por el sentido que tienen las inecuaciones, ellas se satisfacen por cualquier valor real de x; es decir, la solución de la inecuación propuesta es todo el campo de los números reales.

*****

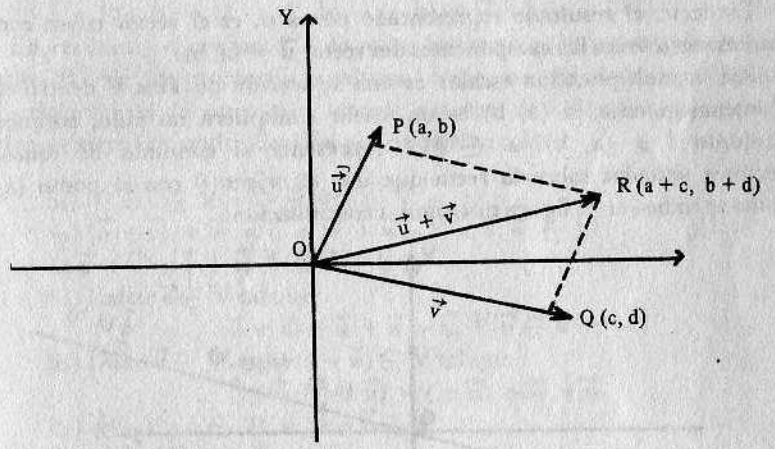
Capítulo X

TEORIA ELEMENTAL DE ESPACIOS VECTORIALES Y TRANSFORMACIONES LINEALES

A) *Espacios vectoriales*

10.1. Completaremos el estudio de las estructuras fundamentales del álgebra, viendo ahora el concepto de espacio vectorial. Este concepto es uno de los más importantes de la Matemática Moderna.

El plano cartesiano y el espacio cartesiano de la Geometría Analítica, representados respectivamente por $\mathbb{R}^2$ y $\mathbb{R}^3$, son ejemplos típicos, de lo que se llama *Espacio Vectorial Real*. Así, por ejemplo, en $\mathbb{R}^2$ cada punto o vector, es un par ordenado (a, b) de números reales cuyas coordenadas *a* y *b* reciben el nombre de *componentes* de dicho vector. Geométricamente el vector $\vec{u} = (a, b)$ puede ser representado por medio de una flecha que va desde el origen O (o, o) del sistema de coordenadas al punto P (a, b), tal como se ilustra en la figura adjunta



Si designamos por V_2 el conjunto de todos los vectores del plano cartesiano $\mathbb{R}^2$, entonces una operación de adición puede ser introducida de la manera siguiente:

(1) $\vec{u} + \vec{v} = (a, b) + (c, d) = (a + c, b + d)$

esto es el paralelógramo de adición usado en las ciencias físicas para sumar fuerzas y velocidades. Así pues, si (a, b) y (c, d) son considerados como dos lados de un paralelógramo, entonces la suma (a, b) + (c, d) es la diagonal de este paralelogramo, como puede comprobarse fácilmente.

De la ecuación (1) se concluye de inmediato que la adición vectorial recién definida es asociativa y conmutativa, es decir, se verifican las relaciones:

(2) $(u + v) + w = u + (v + w)$

(3) $u + v = v + u$

El vector (o, o), de longitud nula, lo llamamos el *vector cero*, desde que

4) $(a, b) + (o, o) = (a + o, b + o) = (a, b)$

Por otra parte, el vector $(-a, -b)$ obtenido por simetría de (a, b) con respecto al origen, es el *aditivo inverso* u *opuesto* de (a, b); desde que

(5) $(a, b) + (-a, -b) = (a - a, b - b) = (0, 0)$

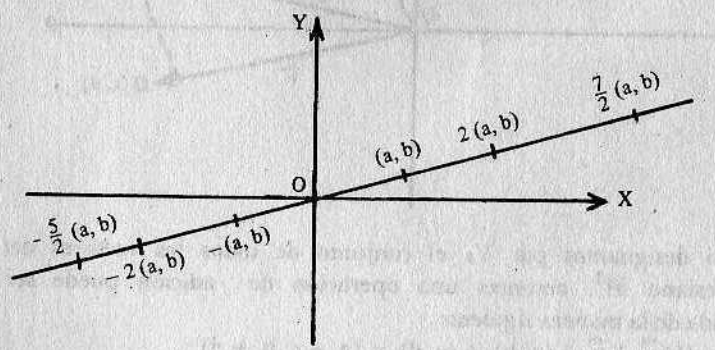
De (1), (2), (3), (4) y (5) concluimos que V es un grupo abeliano respecto a la adición.

Por otra parte, así como es posible sumar vectores en $V = \mathbb{R}^2$, también la operación de multiplicación escalar (operación, binaria externa), esto es, el *producto* de un número real α por un vector $\vec{u} = (a, b)$ puede definirse como sigue:

(6) $\alpha \vec{u} = \alpha (a, b) = (\alpha a, \alpha b)$

Es decir, el resultado representado por α u, es el vector cuyas componentes son α veces las componentes del vector $\vec{u} = (a, b)$.

Así la multiplicación escalar es una aplicación de $\mathbb{R} \times V$ dentro de V. Por consiguiente, si (a, b) es un vector cualquiera no nulo, entonces el conjunto $\{ \alpha (a, b) : \alpha \in \mathbb{R} \}$ representa el conjunto de todos los vectores situados sobre la recta que une el origen 0 con el punto (a, b), como se indica en la figura que sigue a continuación:



Sean ahora $\vec{u}, \vec{v} \in V$ y $\alpha, \beta, \in \mathbb{R}$; entonces se tiene:

$$\begin{aligned}\alpha (\vec{u} + \vec{v}) &= \alpha [(a, b) + (c, d)] \\ &= \alpha [a + c, b + d] \\ &= [\alpha (a + c), \alpha (b + d)] \\ &= [\alpha a + \alpha c, \alpha b + \alpha d]\end{aligned}$$

$$\begin{aligned}&= [\alpha a, \alpha b] + [\alpha c, \alpha d] \\ &= \alpha (a, b) + \alpha (c, d)\end{aligned}$$

(7) $\alpha (\vec{u} + \vec{v}) = \alpha \vec{u} + \alpha \vec{v}$

Por otro lado tenemos.

$$\begin{aligned}(\alpha + \beta) \vec{u} &= (\alpha + \beta) (a, b) \\ &= [(\alpha + \beta) a, (\alpha + \beta) b] \\ &= [\alpha a, \alpha b] + [\beta a, \beta b] \\ &[\alpha a, \alpha b] + [\beta a, \beta b] \\ &= \alpha (a, b) + \beta (a, b)\end{aligned}$$

(8) $(\alpha + \beta) \vec{u} = \alpha \vec{u} + \beta \vec{u}$

Por otra parte resulta que

$$\begin{aligned}\alpha (\beta \vec{u}) &= \alpha [\beta (a, b)] \\ &= \alpha (\beta a, \beta b) \\ &= [\alpha (\beta a), \alpha (\beta b)] \\ &= [(\alpha \beta) a, (\alpha \beta) b] \\ &= (\alpha \beta) (a, b)\end{aligned}$$

(9) $\alpha (\beta \vec{u}) = (\alpha \beta) \vec{u}$

Finalmente, se tiene

$1 \cdot \vec{u} = 1 (a, b) = (1 \cdot a, 1 \cdot b) = (a, b) = \vec{u}$

En resumen: el conjunto V de todos los vectores libres del plano Cartesiano $\mathbb{R}^2$ de la geometría analítica, tiene las siguientes propiedades, en relación con la operación binaria interna de adición y la operación binaria externa de multiplicación por números reales:

S_1) $\forall \vec{u}, \vec{v} \in V \Rightarrow \vec{u} + \vec{v} \in V$

S_2) $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w}), \forall \vec{u}, \vec{v}, \vec{w}, \in V$

S_3 $\vec{u} + \vec{v} = \vec{v} + \vec{u}, \forall \vec{u}, \vec{v} \in V$

S_4) Existe $\vec{o} \in V$ tal que

$$\vec{o} + \vec{u} = \vec{u} + \vec{o} = \vec{u}, \forall \vec{u} \in V$$

S_5) Dado $\vec{u} \in V$, existe $(-\vec{u}) \in V$ tal que

$$\vec{u} + (-\vec{u}) = (-\vec{u}) + \vec{u} = \vec{o}$$

P_1) $\vec{u} \in V$ y $\alpha \in \mathbb{R} \Rightarrow \alpha \vec{u} \in V$

P_2) $\alpha (\vec{u} + \vec{v}) = \alpha \vec{u} + \alpha \vec{v}$

P_3) $(\alpha + \beta) \vec{u} = + \beta \vec{u} +$ [illegible]

P_4) $\alpha (\beta \vec{u}) = (\alpha \beta) \vec{u}$

P_5) $1 \cdot \vec{u} = \vec{u}$

Todo lo que hicimos anteriormente en el plano euclidiano $\mathbb{R}^2$ puede hacerse también en el espacio euclidiano $\mathbb{R}^3$ de la geometría analítica, verificándose las propiedades S_1) hasta S_5) y las P_1) hasta P_5) señaladas anteriormente.

El nombre que se ha dado de ESPACIO VECTORIAL deriva de que el concepto en cuestión ha sido desglosado del estudio de los vectores libres

del plano y del espacio ordinario de la geometría analítica, que constituyen por esto, un primer ejemplo de espacio vectorial, siendo la ley binaria interna la adición geométrica y la ley binaria externa la multiplicación por un número. Pero, muy pronto se reconoció que podían ser considerados también como espacios vectoriales otros muchos conjuntos de entes matemáticos, como ser: el conjunto de los polinomios sobre un cuerpo, el conjunto de las funciones continuas en un intervalo, las soluciones de ecuaciones diferenciales lineales sin segundo miembro, etc.

El concepto de espacio vectorial ha sido generalizado por Peano.

10.2. *Espacio Vectorial Abstracto*

Sea (K; + , ·) un cuerpo de elementos α, β, γ,..., que supondremos aquí conmutativo, y sea (V; +) un grupo aditivo de elementos $\vec{u}, \vec{v}, \vec{w}, \ldots$

Diremos que V constituye un ESPACIO VECTORIAL (y también "espacio lineal") sobre K, denotado por V (K), si en V además de la ley aditiva interna que verifica las condiciones:

S_1) $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$

S_2) $\vec{u} + \vec{v} = \vec{v} + \vec{u}$

S_3) $\vec{o} + \vec{u} = \vec{u} + \vec{o} = \vec{u}$

S_4) $\vec{u} + (-\vec{u}) = (-\vec{u}) + \vec{u} = o$

existe una ley externa con operadores en K que verifican las condiciones:

P_1) $\alpha\ (\beta\ \vec{u}) = (\alpha\ \beta)\ \vec{u}$ (asociatividad de la multiplicación externa)

P_2) $\alpha\ (\vec{u} + \vec{v}) = \alpha\ \vec{u} + \alpha\ \vec{v}$ (distributividad respecto a los elementos de V)

P_3) $(\alpha + \beta)\ \vec{u} = \alpha\ \vec{u} + \beta\ \vec{u}$ (distributividad respecto a los elementos de K)

P_4) $1\ \vec{u} = \vec{u}$ (neutralidad para el elemento unidad de K)

Los elementos $\vec{u}, \vec{v}, \ldots$ de V se llamarán *vectores* y los elementos α, $\beta, \ldots$ de K se llamarán *escalares*.

Las reglas S_1, S_2, S_3, S_4, P_1, P_2, P_3 y P_4 son válidas para todos los vectorres u, v,..., y para todos los escalares.

Nota. Si el cuerpo K es el de los números reales $\mathbb{R}$ entonces el espacio V ($\mathbb{R}$) será un espacio vectorial real. Luego, un *Espacio Vectorial Real* es una colección de objetos llamados *vectores*, junto con operaciones de adición y de multiplicación por números reales, que satisfacen a la lista de axiomas S_1, S_2, S_3, S_4, P_1, P_2, P_3 y P_4 dada anteriormente.

10.3. *Algunas propiedades algebraicas*

A) Por el hecho de ser V un grupo aditivo, podemos afirmar que:

1°. Existe un único elemento neutro respecto de la operación " + ". Este elemento será designado con el símbolo $\vec{o}$ (vector nulo).

2°. Para cada $\vec{u} \in V$, hay en V un único elemento simétrico, que designamos con $(-\vec{u})$ tal que $\vec{u} + (-\vec{u}) = \vec{0}$.

3°. $-(-\vec{u}) = \vec{u}$

4°. $-(\vec{u} + \vec{v}) = (-\vec{u}) + (-\vec{v})$

5°. La ecuación $\vec{a} + \vec{u} = \vec{b}$ tiene por única solución $\vec{u} = \vec{b} + (-\vec{a})$.

Convenimos en escribir $\vec{b} + (-\vec{a}) = \vec{b} - \vec{a}$

6°. Siendo la operación " + " asociativa, eliminaremos los paréntesis en sumas reiteradas, y no nos cuidaremos del orden de los sumandos dado que la operación " + " es conmutativa.

B) En un espacio vectorial V (K) valen además las siguientes propiedades:

1) $0 \cdot \vec{u} = \vec{o}$; $\forall\ \vec{u} \in V$, donde $o \in K$ (el escalar nulo).

En efecto, desde que

$$\vec{u} + o \cdot \vec{u} = (1 + 0) \cdot \vec{u} = 1 \cdot \vec{u} = \vec{u}$$

y por otro lado, $\vec{u} + \vec{o} = \vec{u}$

se sigue que, $\vec{o} \cdot \vec{u}$ y $\vec{o}$ son soluciones de la ecuación $\vec{a} + \vec{u} = \vec{b}$, y como esta solución es única, resulta

$$o \cdot \vec{u} = \vec{o}$$

2) $\alpha \cdot \vec{o} = \vec{o}$, donde $\vec{o} \in V$ (el vector nulo)

En efecto tenemos

$$\alpha \cdot \vec{o} = \alpha\,(o \cdot \vec{o}) = (\alpha \cdot o) \cdot \vec{o} = o \cdot \vec{o} = \vec{o}$$

(1) (1)

3) La ecuación $\alpha \cdot \vec{u} = \vec{v}$, $\alpha \neq o$, tiene una y una sola solución, y que es $\vec{u} = \alpha^{-1}\vec{v}$.

En efecto, se tiene

$$\alpha^{-1}(\alpha\,\vec{u}) = \alpha^{-1}\vec{v}$$
$$(\alpha^{-1}\alpha)\vec{u} = \alpha^{-1}\vec{v}$$
$$1 \cdot \vec{u} = \alpha^{-1}\vec{v}$$
$$\vec{u} = \alpha^{-1}\vec{v}$$

Comprobación.

$$\alpha(\alpha^{-1}\vec{v}) = (\alpha\,\alpha^{-1})\vec{v} = 1 \cdot \vec{v} = \vec{v}$$

4) Si $\alpha\ \vec{u} = \vec{o}$, entonces es, $\alpha = o$, ó $\vec{u} = \vec{o}$

En efecto, si $\alpha = o$, entonces por 1) la igualdad se satisface, ya que $o \cdot \vec{u} = \vec{o}$

Ahora, si $\alpha \neq o$, entonces por 3) resulta

$$\vec{u} = \alpha^{-1} \cdot \vec{o} = \vec{o}$$

1)

5) $(-1)\vec{u} = -\vec{u}$

En efecto, habrá que probar que $(-1)\vec{u}$ es el simétrico de $\vec{u}$.

Pues bien, podremos escribir

$$\vec{u} + (-1)\vec{u} = 1 \cdot \vec{u} + (-1)\vec{u} = [1 + (-1)]\vec{u} = o \cdot \vec{u} = \vec{o}$$

Luego, resulta que

$$(-1)\vec{u} = -\vec{u}$$

6) $(-\alpha)\vec{u} = -(\alpha \vec{u})$

En efecto, podemos escribir

$$(-\alpha)\vec{u} = [(-1)\alpha]\vec{u} = (-1)(\alpha \vec{u}) = -(\alpha \vec{u})$$

5)

7) $\alpha(-\vec{u}) = -(\alpha \vec{u})$

En efecto, se puede escribir

$$\alpha(-\vec{u}) = \alpha[(-1)\vec{u}] = [\alpha(-1)]\vec{u} = [(-1)\alpha]\vec{u} = (-1)(\alpha \vec{u}) = -(\alpha \vec{u})$$

5) 5)

8) $(-\alpha)(-\vec{u}) = \alpha \vec{u}$

En efecto, tenemos

$$(-\alpha)(-\vec{u}) = (-\alpha)[(-1)\vec{u}] = [(-\alpha)(-1)]\vec{u} = \alpha \vec{u}$$

5)

9) Si $\alpha \vec{u} = \alpha \vec{v}$, y $\alpha \neq o$, entonces es $\vec{u} = \vec{v}$

En efecto, se tiene

$$\alpha^{-1}(\alpha \vec{u}) = \alpha^{-1}(\alpha \vec{v})$$
$$(\alpha^{-1}\alpha)\vec{u} = (\alpha^{-1}\alpha)\vec{v}$$
$$1 \cdot \vec{u} = 1 \cdot \vec{v}$$
$$\vec{u} = \vec{v}$$

Observación Importante. Puede observarse de inmediato de que hemos usado el mismo signo »+« en dos situaciones diferentes en la definición que hemos dado de espacio vectorial. Así pues, en el axioma P_2) : $\alpha(\vec{u} + \vec{v}) = \alpha \vec{u} + \alpha \vec{v}$, el signo »+« indica la suma de elementos en el grupo V, mientras que en el axioma P_3): $(\alpha + \beta)\vec{u} = \alpha \vec{u} + \beta \vec{u}$, el signo »+«, del primer miembro indica la suma de elementos en el grupo aditivo del cuerpo K, en tanto que el signo »+« del segundo miembro es la suma de elementos del grupo V.

Solamente advertimos estos hechos para evitar confusiones, ya que nos son suficientemente claras la adición de elementos de V o de elementos de K.

Otra observación de interés está en el axioma P_1): $\alpha(\beta \vec{u}) = (\alpha \beta)\vec{u}$ en donde el producto $\alpha \beta$ del segundo miembro es el producto de dos elementos del cuerpo base K, mientras que el primer miembro $\alpha(\beta \vec{u})$ es la multiplicación escalar del vector $\beta \vec{u}$ por el escalar α de K.

10.4. *Isomorfismos de espacios vectoriales.*

Sean V(K) y V'(K) dos espacios vectoriales sobre el mismo cuerpo K, y sea

$$f : V \rightarrow V'$$

una aplicación biunívoca de V sobre V'.

Si f verifica las condiciones siguientes:

$$(*) \begin{cases} f(\vec{u} + \vec{v}) = f(\vec{u}) + f(\vec{v}) = \vec{u}' + \vec{v}' \\ f(\alpha \vec{u}) = \alpha f(\vec{u}) = \alpha \vec{u}' \end{cases}$$

para todo $\vec{u}, \vec{v} \in V$ y para todo $\alpha \in K$, siendo $\vec{u}', \vec{v}'$, las imágenes de $\vec{u}$, $\vec{v}$ por f, entonces diremos que la aplicación f es un isomorfismo de los espacios V (K) y V' (K).

Las condiciones (*) pueden también escribirse bajo la sola forma,

$$(**) f(\alpha \vec{u} + \beta \vec{v}) = \alpha f(\vec{u}) + \beta f(\vec{v}) = \alpha \vec{u}' + \beta \vec{v}'.$$

Pruebe a modo de ejercicio que las condiciones (*) y (**) son equivalentes.

10.5. Indicamos a continuación algunos ejemplos de espacios vectoriales.

1) En la sección 9.1. vimos que el conjunto de todos los vectores libres del plano ordinario $\mathbb{R}^2$ y del espacio ordinario $\mathbb{R}^3$ de la geometría analítica, satisfacen a todas las condiciones de S_1) a S_4) y de P_1) a P_4). Luego, los conjuntos de todos los vectores de $\mathbb{R}^2$ y de $\mathbb{R}^3$ forman o constituyen un espacio vectorial sobre el cuerpo $\mathbb{R}$ de los números reales.

El alumno, seguramente, habrá notado que lo esencial en el tratamiento algebraico que hicimos de los vectores en el plano, como también lo afirmamos para los vectores del espacio ordinario, consiste en que se identifican estos vectores geométricos, $\vec{OP} = \vec{u}$, con los pares, o ternas ordenadas de números reales; esto es

$$\vec{u} = (a, b), \vec{u} = (a, b, c)$$

y para sus pares o ternas se pueden definir directamente la igualdad y las operaciones, y finalmente estudiar sus propiedades. Es evidente que no necesitaríamos ninguna figura ni representación geométrica para tratar esos vectores y demostrar todos los teoremas.

Este hecho nos permitirá ahora prescindir por completo de toda intuición geométrica, y generalizar la noción de VECTOR REAL.

Definición. Llamaremos *vector real,* o *vector n-dimensional,* a una n – upla de números reales

$$\vec{u} = (a_1, a_2, \ldots, a_n)$$

y las a_i $(i = 1, 2, \ldots, n)$ las llamaremos las *componentes* del vector. También, el vector $\vec{u} = (a_1, a_2, \ldots a_n)$ puede llamarse *punto.*

Denotemos por R^n al conjunto de todos estos vectores reales n-dimensionales.

Análogamente, llamaremos *vector complejo n-dimensional* a una n-upla de números complejos.

$$\vec{z} = (z_1, z_2, \ldots, z_n)$$

y denotemos con C^n al conjunto de todos estos vectores complejos n-dimensionales.

Como en el caso real, los elementos de C^n serán llamados también *puntos,* y los elementos de C se llamarán escalares.

Anteriormente hemos considerado al plano ordinario como el $\mathbb{R}^2$ y al espacio ordinario como el $\mathbb{R}^3$. Ahora, es claro que para $n > 3$, ya se pierde toda intuición geométrica, y nuestros razonamientos se harán por vía puramente algebraica; pero, esto no obsta, lo cual es muy útil, conservar en algunas cuestiones el lenguaje geométrico usual, aún cuando esté desprovisto de toda significación concreta.

En realidad, obsérvese que no hay nada de misterioso en el hecho de que n puede valer más de tres. Es cierto, por supuesto, que los gráficos no pueden dibujarse en la forma usual, pero esto no es un defecto serio. De hecho, la misma intuición geométrica tridimensional, es aún razonablemente exacta para $\mathbb{R}^n$ cuando $n > 3$, a pesar de la falta de una representación visual para estos espacios.

Desde el punto de vista de máxima abstracción en que se coloca el Algebra Abstracta, la cual prescinde de todas las cualidades físicas, químicas y geométricas, etc. de los objetos o elementos, todos los seres reales y abstractos son equivalentes, consideraremos ahora un cuerpo conmutativo cualquiera K. Entonces, una generalización natural de $\mathbb{R}^n$, o del C^n, es obtenida como sigue:

Definición: Llamaremos un *n-vector* sobre el cuerpo K, a una sucesión $(x_1, x_2, \ldots x_n)$ de n elementos de K, en la cual las componentes $x_1, x_2, \ldots x_n$ serán denominadas las *coordenadas* del n-vector $x = (x_1, x_2, \ldots, x_n)$.

Desde luego, como muy bien sabemos, en el concepto de sucesión está

implícito el hecho de que algunas o todas las componentes $x_1, x_2, \ldots, x_n$ del algún n-vector puedan coincidir.

Denotaremos con K^n al conjunto de todas estas sucesiones de n elementos cada una, siendo n un número fijo.

Definición. Diremos que dos n-vectores de K^n son iguales solamente si tienen las mismas coordenadas, esto es:

$$(a_1, a_2, \ldots a_n) = (b_1, b_2, \ldots b_n)$$

si, y sólo si

$$a_1 = b_1, a_2 = b_2, \ldots, a_n = b_n$$

Ahora, una operación de »adición« y una de »multiplicación escalar« pueden definirse en K^n como sigue:

$$(a_1, a_2, \ldots, a_n) + (b_1, b_2, \ldots, b_n) = (a_1 + b_1, a_2 + b_2, \ldots a_n + b_n)$$

$$k(a_1, a_2, \ldots a_n) = (ka_1, ka_2, \ldots, ka_n)$$

Afirmamos, en seguida, que K^n con estas dos operaciones así definidas, es un espacio vectorial.

La demostración de que K^n es un espacio vectorial es enteramente idéntica de la que hicimos en la sección 9.1. para el espacio R^2 Por consiguiente, podemos efectivamente afirmar que K^n con las operaciones definidas recientemente, es un espacio vectorial sobre el cuerpo K. Sin embargo, a modo de ejercicio, aconsejamos al estudiante comprobar nuevamente para el caso K^n los axiomas S_1, S_2, S_3, S_4, P_1, P_2, P_3 y P_4 que caracterizan a la estructura de espacio vectorial.

En K^n, el n-vector nulo es la sucesión $(0,0,\ldots,0)$ (n veces), y el n-vector opuesto del $(a_1, a_2, \ldots, a_n)$ es $(-a_1, -a_2, \ldots, -a_n)$. En K^n, los vectores que tienen (n-1) coordenadas iguales a cero y la otra coordenada igual a uno, tal como, por ejemplo,

$$(0,0,\ldots,0,\underset{n}{1},0,\ldots,0)$$

son de importancia especial, como veremos más tarde. En consecuencia, si δ_{ij} denota la delta de Kronecker, esto es:

$$\delta_{ij} = 0 \text{ si } i \neq j, \quad \delta_{ii} = 1 \quad \text{si } i = j$$

entonces el n-vector de K^n

$$\vec{u} = (\delta_{i1}, \delta_{i2}, \ldots, \delta_{in})$$

es el vector con 1 (unidad del cuerpo K) en la i-ésima coordenada y ceros en cualquier otra parte.

Finalmente, si cambiamos ahora K por $\mathbb{R}$, entonces $\mathbb{R}^n$ se convierte, en un espacio vectorial sobre el cuerpo IR de los números reales.

Llamaremos a $\mathbb{R}^n$ el ESPACIO VECTORIAL n-DIMENSIONAL REAL. Asimismo, si cambiamos K por C, entonces C^n es también un espacio vectorial que llamaremos el ESPACIO VECTORIAL N-DIMENSIONAL COMPLEJO sobre el cuerpo C de los números complejos.

Otros ejemplos de espacios vectoriales son los que a continuación se indican.

2) a) K (K), es decir todo cuerpo es un espacio vectorial sobre el mismo cuerpo.

b) $\mathbb{R}$ ($\mathbb{R}$); Q (Q); $\mathbb{R}$ (Q); C ($\mathbb{R}$)

Compruebe que en los ejemplos indicados en los puntos a) y b) se satisfacen todos los axiomas S_1, S_2, S_3, S_4, P_1, P_2, P_3 y P_4 que caracterizan a la estructura de espacio vectorial.

3) El conjunto de todos los polinomios de grado $\leq n$ en una indeterminada o variable x, con coeficientes reales o complejos, constituyen un espacio vectorial si se definen la adición y la multiplicación por un número real o complejo en la forma habitual de estas operaciones aplicadas a polinomios.

Nótese que se exige que los polinomios, que ahora son llamados vectores, sean de grado $\leq n$, porque puede suceder, en ciertos casos, que la suma de dos polinomios de grado n sea menor que n. Por lo tanto, si imponemos solamente la condicional de que todos los polinomios sean de igual grado n, entonces la propiedad de cierre de la adición de ellos no se verifica en general.

Por consiguiente, el conjunto de todos los polinomios de la forma:

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_2 x^2 + a_1 x + a_0$$

con coeficientes reales o complejos y cuyos grados son los números: 0, 1, 2, 3, ..., n, forman un espacio vectorial sobre el cuerpo de los números reales $\mathbb{R}$, o el de los complejos C.

El polinomio nulo

$$0 \cdot x^n + 0 \cdot x^{n-1} + \ldots + 0 \cdot x^2 + 0 \cdot x + 0$$

representa al vector nulo $\vec{o}$, y el vector opuesto del vector

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$$

es el vector

$$-p(x) = (-a_n)x^n + (-a_{n-1})x^{n-1} + \ldots + (-a_1)x + (-a_0)$$

4) El conjunto de todas las funciones numéricas del tipo $f: X \to \mathbb{R}$, donde el dominio X es un conjunto no vacío cualquiera no necesariamente numérico forman un espacio vectorial sobre $\mathbb{R}$, si la adición de dos funciones f(x) y g(x) es la suma punto a punto:

$$(f + g) : X \to \mathbb{R}$$

definida por

$$(f + g)(x) = f(x) + g(x), \quad \forall \ x \in X$$

vista en la sección 4.17. del tomo I, y la multiplicación escalar es la función

$$(kf) : X \to \mathbb{R}$$

definida por

$$(kf)(x) = kf(x), \quad \forall \ x \in X \cdot k \in \mathbb{R}$$

vista también en la misma sección 4,17. del primer tomo de estos apuntes.

Aquí, el vector cero es la función constante de valor cero, es decir, aquella que aplica cada $x \in X$ en el $o \in R$; y el vector opuesto de f (x) es la función

$$-f : X \to \mathbb{R}$$

definida por $(-f)(x) = -f(x) = -f(x)$ para cada $x \in X$.

5) Si en el ejemplo anterior, el cuerpo $\mathbb{R}$ de los reales se sustituye por otro cualquiera K, entonces el conjunto de todas las funciones

$$f : X \to K$$

definidas en X con valores en K, algebrizado por las mismas operaciones de adición y de multiplicación escalar vistas en el ejemplo 4), y sigue siendo un espacio vectorial sobre el cuerpo K.

6(Sea ahora K el cuerpo de los enteros módulo 3, esto es, $K = \{0, 1, 2\}$. Entonces, el espacio vectorial K^2 está formado de los vectores siguientes:

$\vec{o} = (o,o)$ $\vec{u_1} = (o,1)$, $\vec{u_2} = (o,2)$, $\vec{u_3} = (1,o)$, $\vec{u_4} = (1,1)$, $\vec{u_5} = (1,2)$, $\vec{u_6} = (2,o)$, $\vec{u_7} = (2,1)$, $u_8 = (2,2)$

Construya, a modo de ejercicio, la tabla de adición y de multiplicación escalar. Nosotros sólo determinaremos algunas sumas y algunos productos de un vector de K^2 por un escalar perteneciente al cuerpo de base K.

$\vec{u_3} + \vec{u_5} = (1,o) + (1,2) = (1 + 1, o + 2) = (2,2) = u_8$

$\vec{u_4} + \vec{u_5} = (1,1) + (1,2) = (1 + 1, 1 + 2) = (2,o) = u_6$

$2\vec{u_5} = 2(1,2) = (2,4) = (2,1) = u_7$; etc.

10.6. *Subespacio.*

Sea V(K) un espacio vectorial y U una parte no vacía de V. Entonces, diremos que U (K) es un SUBESPACIO VECTORIAL DE V(K), si U(K) es un espacio vectorial con respecto a las operaciones de adición y de multiplicación escalar que están definidas en V (K).

De donde, el siguiente criterio para averiguar si un subconjunto U de V, $U \subset V$, es un subespacio.

Criterio. Si V(K) es un espacio vectorial, y si U es un subconjunto de V, entonces U(K) es el subespacio de V(K) si y sólo si las siguientes condiciones son verificadas:

S_1) Si $\vec{u}, \vec{v} \in U$, entonces $\vec{u} + \vec{v} \in U$

S_2) Si $\alpha \in K$ y $\vec{u} \in U$, entonces $\alpha \vec{u} \in U$

Demostración. Las condiciones S_1) y S_2) muestran que la adición y la multiplicación por escalares, son operaciones bien definidas en U.

La asociatividad, y la conmutatividad de la operación "+" valen en U, porque ellas valen en V, puesto que $U \subset V$.

El elemento nulo (es decir, el vector nulo) de V pertenece también a U, ya que $\vec{o} = 0 \cdot \vec{u}$ con $\vec{u} \in U$.

Por otra parte, $\vec{u} \in U \Rightarrow (-\vec{u}) \in U$, ya que $(-\vec{u}) = (-1)\vec{u} \in U$.

Así hemos probado que U es un grupo aditivo.

Por otro lado, las propiedades de la multiplicación escalar indicadas en los axiomas P_1, P_2, P_3 y P_4 que caracterizan la estructura de espacio vectorial, valen también en U, porque ellas se verifican en V.

Por consiguiente, U(K) es un espacio vectorial (subespacio de V(K)).

Recíprocamente, si U(K) es un espacio vectorial, entonces las condiciones dadas en S_1) y S_2) se verifican.

Este resultado y el anterior demuestran el criterio.

Observaciones

a) Las condiciones S_1) y S_2) del criterio enunciado son equivalentes a la única condición.

S) Si $\alpha, \beta \in K$ y $\vec{u}, \vec{v} \in U$, entonces $\alpha \vec{u} + \beta \vec{v} \in U$.

¡Pruebe esta equivalencia!

b) Además, conviene observar también que si U (K) es un subespacio y si $\vec{u}, \vec{v} \in U$, entonces $\vec{u} - \vec{v} \in U$.

En efecto, si $\vec{v} \in U$ entonces $(-\vec{v}) = (-1)\vec{v} \in U$ (K), y por consiguiente, $\vec{u} + (-1)\vec{v} = \vec{u} - \vec{v} \in U$ (K).

c) Un subespacio U(K) del espacio V(K) será denominado también con el nombre *variedad lineal* de V(K).

Ejemplos.

1. Todo espacio vectorial V(K) incluye siempre dos subespacios, a saber: el espacio total V(K) y el subespacio constituido únicamente por el vector cero.

Llamaremos a estos dos subespacios con el nombre común de *subespacios triviales*, y todo otro distinto de ellos será llamado un *subespacio propio*.

2. Es claro que el conjunto U (K) de todos los polinomios

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0,$$

en una indeterminada x, de grados $\leq n$, n fijo, en un subespacio del espacio V (K) de todos los polinomios en la indeterminada x en el cuerpo K.

Así, por ejemplo, son subespacios del espacio considerado V(K) los siguientes: a) el conjunto de todos los polinomios de grado ≤ 2; b) el conjunto de todos los polinomios de grado ≤ 3; etc. sobre el cuerpo K de los reales IR y de los complejos C.

3) Existen muchos subconjuntos U de K^n, cuyos vectores pueden ser sumados, restados y multiplicados por escalares para obtener de nuevo un vector de U. Estos subespacios, muchos de ellos, pueden fácilmente ser obtenidos al considerar todos aquellos vectores de K^n que tienen nula una o más coordenadas de un lugar fijo. Así, por ejemplo, supongamos que U está formado por todos los vectores de K^4 cuya tercera coordenada es cero. Los vectores de U serán, pues, de la forma $(a_1, a_2, 0, a_4)$.

Entonces, aplicando el criterio que tenemos para los subespacios, escribimos:

$$(a_1, a_2, 0, a_4) + (b_1, b_2, 0, b_4) = (a_1 + b_1, a_2 + b_2, 0, a_4 + b_4)$$

$$k(a_1, a_2, 0, a_4) = (ka_1, ka_2, 0, ka_4)$$

Es decir, la suma de dos vectores de U y los múltiplos escalares de los vectores de U son también vectores de U. Luego, el espacio U(K) de todos estos vectores es un subespacio del espacio K^4.

Particularizando el cuerpo K al cuerpo $\mathbb{R}$ de los números reales, hay algunos subespacios de $\mathbb{R}^3$ que son interesantes. Así pues, por ejemplo, consideremos el conjunto U de todos los vectores de $\mathbb{R}^3$ que son de la forma

$$U = \{(a, b, o): a, b \in \mathbb{R}\}$$

Este subespacio de $\mathbb{R}^3$ es evidentemente el plano (XY) que contiene a todos los vectores cuya tercera componente es cero. Luego, cada plano coordenado del sistema (XYZ) del espacio ordinario, es un subespacio de $\mathbb{R}^3$.

Asimismo, el conjunto U de todos los vectores de $\mathbb{R}^3$ que son de la forma

$$U = \{(a, o, o): a \in \mathbb{R}\}$$

es el subespacio constituido por el eje coordenado OX, que contiene a todos los vectores cuya segunda y tercera componente es cero.

Luego, cada eje de coordenadas del sistema (XYZ) del espacio ordinario $\mathbb{R}^3$, es un subespacio de $\mathbb{R}^3$.

Otro caso de interés es ahora el siguiente. Sea $\vec{u} = (a, b, c)$ un vector fijo de $\mathbb{R}^3$, y sea U el conjunto de todos los múltiplos $k(a, b, c) = k\vec{u}$ de este vector fijo $\vec{u}$. Entonces, dicho conjunto es un subespacio de $\mathbb{R}^3$, ya que se verifican las condiciones que caracterizan a un subespacio:

S_1) $k(a, b, c) + k'(a, b, c) = (ka, kb, kc) + (k'a, k'b, k'c)$

$= (ka + k'a, kb + k'b, kc + k'c)$

$= ((k+k')a, (k+k')b, (k+k')c)$

$= (k + k')\,(a, b, c)$

$S_2)\ h\,[k\,(a, b, c)] = (hk)\,(a, b, c)$

En consecuencia, la familia de todos los múltiplos de un vector forman subespacios. A un subespacio de estas condiciones le daremos el nombre de DIRECCION LINEAL. Hay infinitos de estos subespacios de dirección lineal en $\mathbb{R}^3$.

4) Consideremos ahora el espacio de todas las funciones numéricas $f: \mathbb{R} \to \mathbb{R}$. Entonces, el subconjunto U de $\mathbb{R}^{\mathbb{R}}$ formado por

$$U = \{f: f(2) = f(5)\}$$

es decir, el conjunto de todas las funciones reales de variable real que asignan el mismo valor a los argumentos 2 y 5, forman un subespacio de $\mathbb{R}^{\mathbb{R}}$ puesto que se verifica la condición:

$$(\alpha f + \beta g)(2) = \alpha f(2) + \beta g(2) = \alpha f(5) + \beta g(5) = (\alpha f + \beta g)(5)$$

que caracteriza de una manera general a los subespacios. Ahora, si el conjunto U está formado por todas las funciones pares o impares, esto es, las funciones f tales que $f(-x) = f(x)$, o $f(-x) = -f(x)$, entonces cada uno de estos conjuntos es también un subespacio del espacio $\mathbb{R}^{\mathbb{R}}$, ya que se verifica:

$$(\alpha f + \beta g)(-x) = \alpha f(-x) + \beta g(-x) = \alpha f(x) + \beta g(x) = (\alpha f + \beta g)(x)$$

para las funciones pares, y

$$(\alpha f + \beta g)(-x) = \alpha f(-x) + \beta g(-x) = -\alpha f(x) - \beta g(x) = -[\alpha f(x) + \beta g(x)] = -(\alpha f + \beta g)(x)$$

para las funciones impares.

5) Ahora, ya que sabemos lo que es un subespacio, podemos indagar cómo se pueden hallar todos los subespacios de un espacio vectorial dado. En general, esto es un problema difícil, como lo muestran algunos de los ejemplos anteriores en los cuales existen una infinidad de subespacios de un espacio dado. Pero, para ciertos espacios, la respuesta a la pregunta anterior puede ser dada fácilmente. Así, por ejemplo, vimos que los únicos subespacios propios de $\mathbb{R}^3$ son líneas y planos que pasan por el origen, por ser el origen el vector nulo, y el cual pertenece tanto al espacio total como a todos sus subespacios. Hecha esta observación, se ve con claridad de que la intersección de dos subespacios de $\mathbb{R}^3$ es nuevamente un subespacio de $\mathbb{R}^3$. Este hecho no es casual, sino general, como lo indica el teorema que sigue.

Teorema. Si $(H_i)_{i \in I}$ es una familia de subespacios de un espacio vectoral V (K), entonces su intersección es también un subespacio de V (K).

Dem. Sea

$$H = \bigcap_{i \in I} H_i$$

y probaremos que H es un subespacio de V (K).

En efecto, sean $\vec{u}, \vec{v} \in H$, entonces $\vec{u}, \vec{v} \in H_i$, $\forall\, i \in I$.

Como los H_i son subespacios, entonces

$$\alpha \vec{u} + \beta \vec{v} \in H_i,\ \forall\, i \in I$$

luego, $\alpha \vec{u} + \beta \vec{v} \in H = \bigcap_{i \in I} H_i$

El teorema está demostrado.

6) Sean U (K) y W (K) dos subespacios arbitrarios de un espacio vectorial V (K).

Por el ejemplo 5) sabemos que la intersección $U \cap W$ es también un subespacio de V (K). Pero la unión $U \cup W$ no es en general un subespacio de V (K). En cambio, el conjunto de todas las sumas posibles de un vector de U con un vector de W es un subespacio de V (K), que llamaremos SUMA DE LOS SUBESPACIOS U y W, y que expresaremos por U + W. Luego,

$$U + W = \{\vec{u} + \vec{w} : \vec{u} \in U, \vec{w} \in W\}$$

En efecto, aplicando el conocido criterio de los subespacios tendremos:

Sean $\vec{u} + \vec{w} \in U + W$ y $\vec{u}' + \vec{w}' \in U + W$ tales que

$\vec{u}, \vec{u}' \in U$ y $\vec{w}, \vec{w}' \in W$

Entonces,

$$(\vec{u} + \vec{w}) + (\vec{u}' + \vec{w}') = (\vec{u} + \vec{u}') + (\vec{w} + \vec{w}') \in U + W$$

y $\alpha(\vec{u} + \vec{w}) = \alpha \vec{u} + \alpha \vec{w} \in U + W$, $\forall\, \alpha \in K$

lo que demuestra nuestra aseveración.

Nótese que U + W incluye a la intersección $U \cap W$, ya que el vector nulo está tanto en U como en W. Además, cualquier subespacio T(K) de V(K) que contenga a $U \cup W$ contendrá también a la suma U + W.

Veamos algún ejemplo.

Sea $\mathbb{R}^3$ el espacio ordinario de la geometría analítica, y sean $U = \{(a, b, o) : a, b \in \mathbb{R}\}$ y $W = \{(o, b', c') : b', c' \in \mathbb{R}\}$ dos subespacios, como sabemos, de $\mathbb{R}^3$. Estos subespacios son, respectivamente, el plano (XY) y el plano (YZ). Es claro que $U + W = \mathbb{R}^3$, puesto que cualquier vector en $\mathbb{R}^3$ es la suma de un vector de U y de un vector de W. Así, por ejemplo:

$$(2, 7, 5) = (2, 3, 0) + (0, 4, 5)$$

Obsérvese que estas sumas no son únicas, ya que, por ejemplo, se tiene

$$(2, 7, 5) = (2, -8, 0) + (0, 15, 5)$$

Pero, si en la suma U + W todo vector $\vec{x}$ de U + W puede escribirse en una y sólo en una manera como:

$$\vec{x} = \vec{u} + \vec{w}, \text{ con } \vec{u} \in U \text{ y } \vec{w} \in W$$

entonces llamaremos a la suma de los subespacios U y W con el nombre de SUMA DIRECTA, denotándola por

$$U \oplus W$$

Por ejemplo, en $\mathbb{R}^3$, sea $U = \{(a, b, o) : a, b \in \mathbb{R}\}$ el plano (XY) y sea $W = \{(o, o, c) : c \in \mathbb{R}\}$ el eje 0Z, que como sabemos, ambos son subespacios de $\mathbb{R}^3$. Pero entonces, cualquier vector de $\mathbb{R}^3$ puede escribirse como la suma de un vector de U y de un vector de W de una y sólo de una manera, puesto que,

$$(a, b, c) = (a, b, o) + (o, o, c)$$

Luego, en este caso R^3 es la suma directa de los subespacios dados U y W; esto es,

$$\mathbb{R}^3 = U + W$$

Para reconocer si la suma directa de dos subespacios U y W existe, tenemos el teorema que sigue a continuación:

Teorema. El espacio S (K) es la suma directa de los subespacios U (K) y W (K), si, y sólo si se verifican las condiciones siguientes:

a) $S = U + W$

b) $U \cap W = \{\vec{o}\}$

Demostración. Supongamos que las condiciones a) y b) se satisfagan. Probaremos entonces que la suma es directa, esto es,

$$S = U \oplus W$$

En efecto, sea $\vec{x} \in S$, entonces por a) existen $\vec{u} \in U$ y $\vec{w} \in W$ tales que $\vec{x} = \vec{u} + \vec{w}$.

Necesitamos ahora demostrar que esta suma es única. Por el contrario, supongamos que existan $\vec{u}' \in U$ y $\vec{w}' \in W$ tales que también se tenga $\vec{x} = \vec{u}' + \vec{w}'$. Entonces, es

$$\vec{u} + \vec{w} = \vec{u}' + \vec{w}'$$

y por tanto, $\vec{u} - \vec{u}' = \vec{w}' - \vec{w}$

Pero, $\vec{u} - \vec{u}' \in U$ y $\vec{w}' - \vec{w} \in W$; y como por b) es $U \cap W = \{\vec{o}\}$, resulta entonces que debe ser necesariamente

$$\vec{u} - \vec{u}' = \vec{o} \text{ y } \vec{w}' - \vec{w} = \vec{o}$$

o sea, $\vec{u} = \vec{u}'$ y $\vec{w} = \vec{w}'$ (absurdo).

Esta contradicción prueba que la suma $\vec{x} = \vec{u} + \vec{w}$ es única, y por esto resulta que $S = U \oplus W$.

Recíprocamente, supongamos ahora que la suma sea directa; es decir, que se tiene $S = U \oplus W$, y probemos entonces que en tal caso se satisfacen las condiciones a) y b) indicadas en el enunciado del teorema.

En efecto, siendo por hipótesis $S = U \oplus W$, entonces cualquier $\vec{x} \in S$ puede escribirse en forma única como

$$\vec{x} = \vec{u} + \vec{w}, \text{ con } \vec{u} \in U \text{ y } \vec{w} \in W$$

Luego, en particular se tiene

$$S = U + W$$

resultado que prueba a).

Supongamos ahora que sea x un elemento cualquiera de la intersección $U \cap W$, es decir, $\vec{x} \in U \cap W$.

Entonces, se podría escribir

$$\vec{x} = \vec{x} + \vec{o}, \text{ con } \vec{x} \in U \text{ y } \vec{o} \in W \qquad (1)$$

y también, $\vec{x} = \vec{o} + \vec{x}$, con $\vec{o} \in U$ y $\vec{x} \in W$ (2)

Ahora bien, como por hipótesis la suma $\vec{x}$ es única, resulta entonces que el vector $\vec{x} \in U \cap W$ que satisface a la vez a (1) y (2) debe ser necesariamente nulo, y como es además arbitrario en $U \cap W$, resulta entonces que

$$U \cap W = \{\vec{o}\}$$

resultado que prueba b), y el teorema queda demostrado.

Veamos un ejemplo.

Sea $\mathbb{R}^3$ el espacio ordinario de la geometría analítica, y sean los subespacios $U = \{(a, a, a) : a \in \mathbb{R}\}$ y $W = \{(o, b, c) : b, c \in \mathbb{R}\}$

Probar que $\mathbb{R}^3 = U \oplus W$.

En efecto, mostremos primeramente que $U \cap W = \{\vec{o}\}$.

Tenemos: Sea $\vec{x} \in U \cap W \Rightarrow \vec{x} \in U$ y $\vec{x} \in W$, o sea, por una parte es $\vec{x} = (a, a, a)$, y por otra, $\vec{x} = (o, b, c)$; lo cual es posible solamente cuando $a = b = c$ y $a = 0$, es decir, si el vector $\vec{x}$ es el $(o, o, o) = \vec{0}$. Luego,

$$U \cap W = \{o\}$$

En segundo término mostraremos ahora que se verifica:

$$\mathbb{R}^3 = U + W$$

En efecto, sea $(a, b, c) \in \mathbb{R}^3$ arbitrario. Entonces, se tiene

$$(a, b, c) = (a, a, a) + (o, b - a, c - a)$$

resultado que prueba nuestra aseveración.

Así hemos demostrado que, por verificarse las condiciones $U \cap W = \{\vec{o}\}$ y $\mathbb{R}^3 = U + W$, entonces la suma de los subespacios U y W es directa, y además que $\mathbb{R}^3 = U \oplus W$.

7) Sea $X = \{(2, -3, 1, 0), (-5, 3, 8, 1)\}$ un subconjunto del espacio $\mathbb{R}^4$.

Consideremos en seguida el subconjunto S de $\mathbb{R}^4$ formado por todos los vectores de la forma

$$\vec{x} = \alpha (2, -3, 1, 0) + \beta (-5, 3, 8, 1)$$

donde α y β son dos números reales cualesquiera.

Probaremos que el subconjunto S de $\mathbb{R}^4$ es un subespacio de $\mathbb{R}^4$. En efecto, según el criterio de los subespacios, se tiene

$$[\alpha_1 (2, -3, 1, 0) + \beta_1 (-5, 3, 8, 1)] + [\alpha_2 (2, -3, 1, 0) + \beta_2 (-5, 3, 8, 1)] = (\alpha_1 + \alpha_2)(2, -3, 1, 0) + (\beta_1 + \beta_2)(-5, 3, 8, 1) \in S$$

y por otra parte,

$\gamma\ [\alpha\ (2, -3, 1, 0) + \beta\ (-5, 3, 8, 1)] = (\alpha\ \gamma)\ (2, -3, 1, 0) + (\alpha\ \beta)$
$(-5, 3, 8, 1) \in S$

Estos dos resultados demuestran que el subconjunto de $\mathbb{R}^4$,

$$S = \{\alpha\ (2, -3, 1, 0) + \beta\ (-5, 3, 8, 1) : \alpha, \beta \in \mathbb{R}\}$$

es un subespacio del espacio vectorial $\mathbb{R}^4$ que contiene a X.

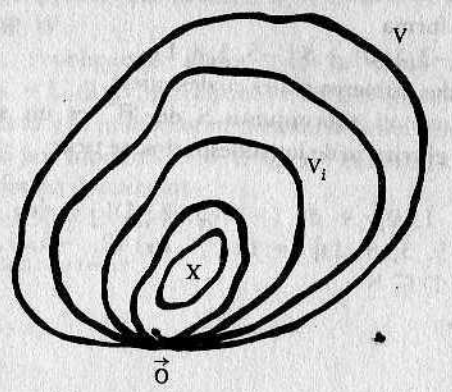
10.7. *Espacio vectorial engendrado por una familia de vectores*

En los ejemplos 5) y 6) de la sección anterior 9.6. nos propusimos hallar todos los subespacios de un espacio vectorial dado V (K). Vimos que el problema no es fácil. Por este motivo, en vez de atacar directamente el problema, es aconsejable seguir mejor el siguiente procedimiento: Sea X un subconjunto no vacío cualquiera de vectores del espacio V. Entonces, tal como se vio en el ejercicio 7) de la sección 9.6., hay cuando menos un subespacio de V (K) que contiene a X, a saber el espacio total V (K). Siendo esto así, vamos a intentar encontrar el "más pequeño" subespacio que contiene al subconjunto X. Con esto queremos expresar a aquel subespacio de V (K) que contiene al conjunto X y que está a su vez contenido en todos los subespacios de V (K) que contienen al subconjunto X de vectores de V.

Ahora bien, para mostrar que realmente existe tal subespacio menor, procederemos de la siguiente manera:

Sea $(V_i)_{i \in I}$ la familia de todos los subespacios V_i de V que contienen al conjunto X. Esta familia no es vacía, puesto que existe por lo menos un subespacio que contiene a X, a saber V.

Sea $U = \bigcap_{i \in I} V_i$ la intersección de todos estos subespacios. Por el teorema indicado en el ejemplo 5) de la sección 9.6. sabemos que esta intersección $U = \bigcap_{i \in I} V_i$ es también un subespacio del espacio considerado V (K). Es claro que esta intersección U contiene también al subconjunto X, en virtud de propiedades muy conocidas vistas en la teoría de conjuntos (Ver Capítulo II, Tomo I).



Por otro lado, como la intersección de dos o más conjuntos es un subconjunto de cada uno de ellos, resulta que efectivamente U es el menor subespacio que contiene el conjunto X de vectores de V.

Luego, $U = \bigcap_{i \in I} V_i$ es el subespacio deseado.

Diremos que U es el subespacio *engendrado (o generado)* por el conjunto de vectores X. Lo escribiremos en la forma $\overline{X}$; es decir, $\bigcap_{i \in I} V_i = \overline{X}$.

La definición del subespacio $\overline{X}$, como lo acabamos de ver, hace intervenir la familia de todos los subespacios que contienen al conjunto X; por lo tanto, es interesante que se pueda encontrar un método fácil de determinar $\overline{X}$ en términos de los vectores pertenecientes al conjunto X. Si esto se logra conseguir, habremos progresado un poco más en la determinación de los subespacios del espacio dado V (K).

Ahora, para poder expresar los vectores del subespacio X en función de los vectores del conjunto considerado X, será necesario que introduzcamos la definición siguiente.

Definición: Dados los vectores $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ de un espacio V(K), todo vector $\vec{u}$ de la forma

$$\vec{u} = \alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \ldots + \alpha_n \vec{a}_n = \sum_{i=1}^{n} \alpha_i \vec{a}_i$$

se dirá una *combinación lineal* de los vectores $\vec{a}_1, \vec{a}_2, \ldots \vec{a}_n$, y donde los escalares $\alpha_1, \alpha_2, \ldots$ y α_n, llamados *coeficientes* de la combinación lineal considerada, son elementos cualesquiera del cuerpo K

Las combinaciones lineales de un vector $\vec{a}$ son entonces todos los vectores múltiplos de $\vec{a}$, es decir, los de la forma $\alpha \cdot \vec{a}$. En particular, $\vec{a} = 1.\vec{a}$ es combinación lineal de $\vec{a}$.

Por otra parte, es inmediato que todo vector $\vec{a}_i$ es combinación lineal de toda familia $\{a_1, a_2, \ldots a_i, \ldots a_n\}$ que lo contiene, puesto que:

$$\vec{a}_i = 1 \cdot \vec{a}_i = 0 \cdot \vec{a}_1 + 0 \cdot \vec{a}_2 + \ldots + 1 \cdot \vec{a}_i + \ldots + 0 \cdot \vec{a}_n$$

También es inmediato que el vector nulo $\vec{0}$ es combinación lineal de cualquier familia arbitraria de vectores, puesto que

$$\vec{0} = 0 \cdot \vec{a}_1 + 0 \cdot \vec{a}_2 + \ldots + 0 \cdot \vec{a}_n$$

Si $\vec{u}$ es una combinación lineal de ciertos vectores $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$, y si $\vec{b}_1, \vec{b}_2, \ldots \vec{b}_m$ son vectores arbitrarios, entonces $\vec{u}$ es combinación lineal del sistema o familia

$\{a_1, \vec{a}_2, \ldots, \vec{a}_n, \vec{b}_1, \vec{b}_2, \ldots, \vec{b}_m\}$, ya que de

$$\vec{u} = \alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \ldots + \alpha_n \vec{a}_n$$

resulta, $\vec{u} = \alpha_1 \vec{a}_1, + \alpha_2 \vec{a}_2 + \ldots + \alpha_n \vec{a}_n + 0 \cdot \vec{b}_1 +$

$$+ 0 \cdot \vec{b}_1 + 0 \cdot \vec{b}_2 + \ldots + 0 \cdot \vec{b}_m$$

Finalmente, si $\vec{u}$ es una combinación lineal de $\{\vec{a}_1, \vec{a}_2, \ldots \vec{a}_n\}$ y si cada $\vec{a}_i$ (i = 1, 2,... n) es combinación lineal del sistema $\{\vec{b}_1 \vec{b}_2, \ldots \vec{b}_p\}$, entonces $\vec{u}$ es combinación lineal de los vectores de esta última familia.

En efecto, tenemos

$$\vec{u} = \sum_{i=1}^{n} \alpha_i \vec{a}_i$$

$$\vec{a}_i = \sum_{j=1}^{p} \beta_j \vec{b}_j$$

resulta, $\vec{u} = \sum_{i=1}^{n} \alpha_i \left(\sum_{j=1}^{p} \beta_j \vec{b}_j \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{p} \alpha_i \beta_j b_j \right)$

o bien:

$$\vec{u} = \sum_{ij}^{n\,p} (\alpha_i \beta_j) \vec{b}_j$$

lo que demuestra nuestra aseveración.

Ahora, en posesión de la definición que acabamos de dar y de los antecedentes expuestos, estamos en condiciones de describir el subespacio $\overline{X}$ generado por la familia X de vectores del espacio V (K), que anteriormente sólo definíamos. A este respecto, tenemos el teorema que sigue.

Teorema. Dado un conjunto, o familia no vacía $X = \{\vec{a}_1, \vec{a}_2, \ldots \vec{a}_n\}$ de vectores de un espacio vectorial V (K). Entonces, el subespacio $\overline{X}$ engendrado por la familia X está formado por todas las combinaciones lineales finitas de los vectores de X.

En otras palabras, dados n vectores. $n \geq 1$, $\vec{a}_1, \vec{a}_2, \ldots \vec{a}_n \in V$, entonces se verifican las propiedades siguientes:

1) $\overline{X}$ es un subespacio de V (K)

2) $X \subseteq \overline{X}$

3) Entre todos los subespacios que contienen al conjunto X, $\overline{X}$ es el más pequeño.

Demostración.

1. Probemos que $\overline{X}$ es un subespacio, si lo es, tendrá por el criterio conocido, que ser cerrado para las operaciones de adición y de multiplicación por escalares. En efecto, si $\vec{u}, \vec{v} \in \overline{X}$ y $\alpha \in K$ tenemos:

$$\vec{u} = \sum_{i=1}^{n} \alpha_i \vec{a}_i, \vec{v} = \sum_{i=1}^{n} \alpha'_i a_i$$

$$\vec{u} + \vec{v} = \sum_{i=1}^{n} (\alpha_i + \alpha'_i) \vec{a}_i \in \overline{X}$$

$$\alpha \vec{u} = \alpha \sum_{i=1}^{n} \alpha_i \vec{a}_i = \sum_{i=1}^{n} (\alpha \alpha_i) \vec{a}_i \in \overline{X}$$

Así hemos probado que $\overline{X}$ es un subespacio de V (K). Llamaremos también a $\overline{X}$ con el nombre de *variedad lineal engendrada por el sistema* $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$.

Luego, hemos demostrado la propiedad 1) enunciada arriba.

2. Como cada $\vec{a}_i \in X$ es combinación lineal de sí mismo, como también de cualquier conjunto de vectores que lo contiene, resulta de inmediato que $X \subseteq \overline{X}$.

Así hemos probado la propiedad 2) del enunciado del teorema.

3. Para probar este tercer y último punto del enunciado, deberemos hacer ver que si U (K) es un subespacio de V (K) tal que si $\overline{X} = \{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\} \subseteq \cup$, entonces es $\overline{X} \subseteq U$ (K). Para esto debemos mostrar entonces que U (K) contiene a cualquiera combinación lineal de los vectores $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$. Esto es cierto, pues $\vec{a}_1, \vec{a}_2, \ldots \vec{a}_n$, por hipótesis, son elementos de U (K), y por lo que ya se ha visto, toda combinación lineal de elementos del subespacio $\overline{X}$, es también un elemento de U (K).

Este resultado y los anteriores demuestran el teorema.

En particular, si $\overline{X} = V$ (K), entonces decimos que los vectores $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ son generadores del espacio vectorial V (K).

Finalmente, no olvidar que, cuando decimos que el espacio $\overline{X}$ que acabamos de definir es el "más pequeño", o "el menor" espacio vectorial que contiene a la familia $X = \{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$ de vectores del espacio V (K), queremos expresar con esto de que cualquier otro subespacio U (K) de V (K) que contenga a los vectores $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ de la familia o sistema X, debe por por las propiedades del cierre de la adición vectorial y de la multiplicación por un escalar, contener a todos los vectores.

$$\alpha_1 \vec{a}_1 + \alpha_2 a_2 + \ldots + \alpha_n \vec{a}_n$$

Esto es, U (K) debe contener a $\overline{X}$.

Ejercicios. 1. Expresar el vector $(-8,27,27)$ de $\mathbb{R}^3$ como una combinación lineal de los vectores $\vec{a}_1 = (-1, 2, 3)$, $\vec{a}_2 = (5, -3, 1)$ y $\vec{a}_3 = (1, 3, 4)$.

Sol. Denotando por x, y, z los escalares correspondientes de la combinación lineal, escribimos:

$(-8, 27, 27) = x(-1, 2, 3) + y(5, -3, 1) + z(1, 3, 4)$

$(-8, 27, 27) = (-x + 5y + z, 2x - 3y + 3z, 3x + y + 4z)$

y por la definición de igualdad de vectores en los espacios de n-uplas, resulta el sistema lineal de ecuaciones siguientes:

$$\begin{array}{rcl|l|l} -x + 5y + z = -8 & & \cdot 3 & \cdot 4 \\ 2x - 3y + 3z = 27 & & \cdot 1 & \\ 3x + y + 4z = 27 & & & \cdot 1 \end{array}$$

$$\begin{array}{rcl|l} 5x - 18y = 51 & & \cdot 19 \\ 7x - 19y = 59 & & \cdot 18 \end{array}$$

$$\begin{array}{rcl} 95x - 18 \cdot 19y = 969 \\ 126x - 18 \cdot 19y = 1062 \end{array}$$

$$31x = 93$$

$$x = 3$$

$$y = \frac{5x - 51}{18} = \frac{15 - 51}{18} = -\frac{36}{18} = -2$$

$$z = -8 + x - 5y = -8 + 3 + 10 = 5$$

que son los escalares de la combinación lineal.

Observación. Como el sistema lineal de ecuaciones es compatible, entonces resulta que el problema propuesto tiene solución, es decir, es posible expresar el vector $(-8, 27, 27)$ en función de los vectores $(-1, 2, 3)$, $(5, -3, 1)$ y $(1, 3, 4)$. Esto es, la combinación lineal

$(-8, 27, 27) = 3(-1, 2, 3) - 2(5, -3, 1) + 5(1, 3, 4)$

2. Escribir el polinomio $p(x) = 3x^2 + 8x - 5$ como una combinación lineal de los polinomios $q(x) = 2x^2 + 3x - 4$, $r(x) = x^2 - 2x - 3$ sobre el cuerpo $\mathbb{R}$ de los números reales.

Sol.. Denotando por h y k los escalares de la combinación lineal correspondiente, se escribe:

$$p(x) = hq(x) + kr(x)$$

o sea,

$$3x^2 + 8x - 5 = h(2x^2 + 3x - 4) + k(x^2 - 2x - 3)$$

$$3x^2 + 8x - 5 = (2h + k)x^2 + (3h - 2k)x - (4h + 3k)$$

Como esta igualdad tiene lugar para cualquier valor de x, es necesario entonces que los coeficientes de las mismas potencias de x de ambos miembros sean iguales; luego, el sistema

$$\begin{array}{rcl} 2h + k = 3 \\ 3h - 2k = 8 \\ 4h + 3k = -5 \end{array}$$

El sistema formado por las dos primeras ecuaciones tiene la solución $h = 2$ y $k = -1$. Como esta solución satisface también la tercera ecuación del sistema, entonces dicho sistema es compatible y por esto el problema propuesto tiene solución, y la cual es

$$p(x) = 2q(x) - 1 \cdot r(x)$$

es decir,

$$3x^2 + 8x - 5 = 2(2x^2 + 3x - 4) - (x^2 - 2x - 3).$$

3. Sea $\mathbb{R}^3$ el espacio ordinario de la geometría analítica. Entonces vamos a considerar los siguientes subespacios generados,

a) Sea $X = \{\vec{a}\}$.

Entonces el subespacio $\overline{X}$ generado por el sistema $X = \{\vec{a}\}$ es $\overline{X} = \{\alpha \vec{a} : \alpha \in \mathbb{R}\}$, es decir, es la recta que pasa por el origen y determinada por el vector dado $\vec{a}$.

Si el vector $\vec{a}$ es enteramente arbitrario en $\mathbb{R}^3$, entonces $\overline{X} = \{\alpha \vec{a} : \vec{a} \in \mathbb{R}^3 \text{ y } \alpha \in \mathbb{R}\}$ representa las infinitas rectas que pasan por el origen (haz de rectas).

b) Sea $X = \{\vec{a}, \vec{b}\}$.

Entonces, si $\vec{a}$ y $\vec{b}$ son dos vectores no colineales, el subespacio $\overline{X}$ engendrado por el sistema $X = \{\vec{a}, \vec{b}\}$ es el plano que pasa por el origen determinado por las direcciones de $\vec{a}$ y $\vec{b}$, puesto que toda combinación lineal de $\vec{a}$ y $\vec{b}$, es decir

$$\vec{u} = \alpha \vec{a} + \beta \vec{b}$$

está ciertamente contenida en este plano. Recíprocamente, podemos expresar cualquier vector de este plano mediante una combinación lineal de los vectores del sistema dado $X = \{\vec{a}, \vec{b}\}$.

Como en el caso anterior, si los vectores $\vec{a}$ y $\vec{b}$ son enteramente arbitrarios en $\mathbb{R}^3$, entonces el subespacio $\overline{X}$ generado por $X = \{\vec{a}, \vec{b}\}$ representa el haz de planos que pasan por el origen.

En resumen: los subespacios de un espacio vectorial de tres dimensiones son: el espacio nulo $\{\vec{o}\}$, las rectas que pasan por el origen, los planos que pasan por el origen y el espacio todo.

Nótese que todos estos subespacios pasan por el origen que representa, como sabemos, el vector nulo $\vec{o}$ y que es común a todos los subespacios del espacio entero.

Veamos algunos ejemplos, que se refieren al punto b).

(1) ¿Generan los vectores siguientes:

$\vec{a}_1 = (-1, 2, 0), \vec{a}_2 = (3, 1, -1), \vec{a}_3 = (1, 0, -1)$

al espacio $\mathbb{R}^3$?

Sol. Necesitamos demostrar si un vector cualquiera (a, b, c) $\in \mathbb{R}^3$ es o no una combinación lineal de $\vec{a}_1, \vec{a}_2$ y $\vec{a}_3$.

Sea (a, b, c) = x(−1, 2, 0) + y(3, 1, −1) + z(1, 0, −1) de donde x, y, z indican los escalares correspondientes de la combinación lineal buscada.

(a, b, c) = (−x + 3y + z, 2x + y, −y − z)

Entonces, formamos el sistema de ecuaciones simultáneas que sigue.

$$\begin{array}{r} -x + 3y + z = a \\ 2x + y \quad\;\; = b \\ -y - z = c \end{array}$$

Este sistema es equivalente al siguiente

$$\begin{array}{r} -x + 3y + z = a \\ 2x + y \quad\;\; = b \\ -x + 2y = a + c \end{array}$$

y éste al que sigue a continuación:

$$\begin{array}{r} -x + 3y + z = a \\ 2x + y \quad\;\; = b \\ 5x \quad\;\; = 2b - a - c \end{array}$$

y como este último tiene solución, concluimos que los vectores dados $\vec{a}_1, \vec{a}_2$, y $\vec{a}_3$ generan el espacio $\mathbb{R}^3$.

(Nótese que no es necesario encontrar los valores de los escalares x, y, z; solamente se necesita conocer si existe una solución).

Análogamente, averigüaremos también si los vectores $\vec{b}_1 = (2, -2, 3)$, $\vec{b}_2 = (3, -1, -5)$ y $\vec{b}_3 = (-1, -5, 27)$ generan o no a $\mathbb{R}^3$.

Procediendo como antes, se encuentra:

(a, b, c) = x(3, −1, −5) + y(−1, −5, 27) + z(2, −2, 3)

(a, b, c) = (3x − y + 2z, −x −5y − 2z, −5x + 27y + 3z)

de donde el sistema simultáneo

$$\begin{array}{r} 3x - y + 2z = a \\ -x - 5y - 2z = b \\ -5x + 27y + 3z = c \end{array}$$

o sea,

$$\begin{array}{r} 3x - y + 2z = a \\ -16y - 4z = a + 3b \\ 76y + 19z = 5a + 3c \end{array}$$

o sea,

$$\begin{array}{r} 3x - y + 2z = a \\ -16y - 4z = a + 3b \\ 0 = 39a + 57b + 12c \end{array}$$

Este resultado nos enseña que el vector (a, b, c) $\in \mathbb{R}^3$ pertenece al espacio generado por los vectores $\vec{b}_1$, $\vec{b}_2$, y $\vec{b}_3$ si y sólo si el sistema anterior es compatible, y lo es si y sólo si, 39a + 57b + 12c = 0. Por tal motivo, concluimos que los vectores $\vec{b}_1$, $\vec{b}_2$ y $\vec{b}_3$ no generan todos el espacio $\mathbb{R}^3$.

(2) Probar que el plano (YZ), U = {(o, b, c) : b, c $\in \mathbb{R}$} en el espacio ordinario $\mathbb{R}^3$ es generado por los vectores $\vec{a}_1 = (0, 1, 1)$ y $\vec{a}_2 = (0, 2, -1)$.

En efecto, basta mostrar que cualquier vector (o, b, c) del plano (YZ) es combinación lineal de $\vec{a}_1$ y $\vec{a}_2$; tenemos

(o, b, c) = x(0, 1, 1) + y(0, 2, −1)

(o, b, c) = (0, x + 2y, x − y)

De donde el sistema

$$\begin{array}{r} o = o \\ x + 2y = b \\ x - y = c \end{array}$$

o bien,

$$\begin{array}{r} x + 2y = b \\ 3x = b + 2c \end{array}$$

Ahora, como este sistema es compatible y tiene, por tanto, solución, resulta entonces que el (YZ) es generado por los vectores $\vec{a}_1 = (0,1,1)$ y $\vec{a}_2 = (0,2,-1)$.

(3) Demostrar que en $\mathbb{Q}^3$, el subespacio engendrado por $\vec{a}_1 = (1, 2, 1)$ y $\vec{a}_2 = (1, 3, 2)$ es el mismo que el subespacio generado por $\vec{b}_1 = (1, 1, 0)$ y $\vec{b}_2 = (3, 8, 5)$.

En efecto, sabemos que un vector cualquiera del primer subespacio es de la forma

$$\vec{u} = \alpha \vec{a}_1 + \beta \vec{a}_2$$

Ahora bien, para que este subespacio sea el mismo que el generado por los vectores $\vec{b}_1$ y $\vec{b}_2$, deberá tenerse

$$\alpha \vec{a}_1 + \beta \vec{a}_2 = \vec{b}_1 \quad \text{y} \quad \alpha \vec{a}_1 + \beta \vec{a}_2 = \vec{b}_2$$

para valores convenientes de los escalares α y β.

Por consiguiente tendremos:

α (1, 2, 1) + β (1, 3, 2) = (1, 1, 0); α (1, 2, 1) + β (1, 3, 2) = (3, 8, 5)

o sea,

$$(\alpha + \beta,\ 2\alpha + 3\beta,\ \alpha + 2\beta) = (1, 1, 0);\ (\alpha + \beta,\ 2\alpha + 3\beta, \alpha + 2\beta) = (3, 8, 5)$$

De donde resultan los dos sistemas simultáneos

$$\left.\begin{array}{r} \alpha + \beta = 1 \\ 2\alpha + 3\beta = 1 \\ \alpha + 2\beta = 0 \end{array}\right| \quad ; \quad \left.\begin{array}{r} \alpha + \beta = 3 \\ 2\alpha + 3\beta = 8 \\ \alpha + 2\beta = 5 \end{array}\right|$$

Restando la primera ecuación de la segunda previamente multiplicada por dos, y en seguida restando la primera a la tercera, se encuentran dos sistemas equivalentes a los anteriores, respectivamente

$$\left.\begin{array}{r} \alpha + \beta = 1 \\ \beta = -1 \\ \beta = -1 \end{array}\right| \quad ; \quad \left.\begin{array}{r} \alpha + \beta = 3 \\ \beta = 2 \\ \beta = 2 \end{array}\right|$$

o bien,

$$\left.\begin{array}{r} \alpha + \beta = 1 \\ \beta = -1 \end{array}\right| \quad ; \quad \left.\begin{array}{r} \alpha + \beta = 3 \\ \beta = 2 \end{array}\right|$$

y cuyas soluciones son, respectivamente las siguientes:

$$\begin{cases} \alpha = 2 \\ \beta = -1 \end{cases} \quad ; \quad \begin{cases} \alpha = 1 \\ \beta = 2 \end{cases}$$

y los vectores $\vec{b}_1$ y b_2 son: $(1, 1, 0) = 2(1, 2, 1,) - (1, 3, 2)$ y $(3, 8, 5) = (1, 2, 1) + 2(1, 3, 2)$

Luego, cada uno de los vectores del segundo subespacio es un vector, que pertenece al primer subespacio, y de donde concluimos que ambos subespacios coinciden.

Observación. Si denotamos ambos espacios por las notaciones

$$[\vec{a}_1, \vec{a}_2] \quad \text{y} \quad [\vec{b}_1, \vec{b}_2]$$

podremos escribir las igualdades siguientes:

$$[\vec{a}_1, \vec{a}_2] = [\vec{b}_1, \vec{b}_2]$$
$$[\vec{b}_1, \vec{a}_1, \vec{a}_2] = [\vec{a}_1, \vec{a}_2]$$
$$[\vec{b}_1, \vec{b}_2, \vec{a}_1, \vec{a}_2] = [\vec{a}_1, \vec{a}_2]$$

Este ejemplo nos mueve a enunciar el teorema siguiente.

Teorema. Si $[\vec{b}_1, \vec{b}_2, \ldots \vec{b}_m, \vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n] = [\vec{a}_1, \vec{a}_2, \ldots \vec{a}_n]$, entonces cada uno de los vectores $\vec{b}_j$ $(j = 1,2,\ldots,m)$ es combinación lineal de los $\vec{a}_i$ $(1 = 1, 2, \ldots, n)$ y recíprocamente.

En otras palabras, no se cambia el subespacio engendrado por un sistema dado de vectores agregando nuevos generadores que son combinaciones lineales de los primeros.

Demostración. En primer término es evidente que si se tiene

$$[\vec{b}_1, \ldots, \vec{b}_m, \vec{a}_1, \ldots \vec{a}_n] = [\vec{a}_1, \ldots, \vec{a}_n],$$

entonces cada b_j $(j = 1, 2, \ldots, m)$, por pertenecer a este último subespacio $\vec{a}_1, \ldots \vec{a}_n$, debe ser combinación lineal de los vectores

$\vec{a}_i$ $(i = 1, 2, \ldots, n)$.

Recíprocamente, si suponemos ahora que cada $\vec{b}_j$ $(j = 1,2,\ldots,m)$ es combinación lineal de los $\vec{a}_i$ $(i = 1, 2, \ldots, n)$, vamos entonces a probar que los subespacios indicados en el enunciado del teorema coinciden.

En efecto, es claro que al aumentar los vectores generadores, el subespacio que ellos engendran no puede disminuir, pero podría aumentar, entonces para demostrar tal coincidencia nos bastará demostrar que toda combinación lineal de los $\vec{a}_i$, $\vec{b}_j$ es una combinación lineal solamente de los vectores $\vec{a}_i$ $(i = 1, 2, \ldots, n)$. Efectivamente, por hipótesis tenemos

$$\vec{b}_1 = \beta_{11}\vec{a}_1 + \beta_{12}\vec{a}_2 + \ldots + \beta_{1n}\vec{a}_n$$
$$b_2 = \beta_{21}\vec{a}_1 + \beta_{22}a_2 + \ldots + \beta_{2n}\vec{a}_n$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\vec{b}_m = \beta_{m1}\vec{a}_1 + \beta_{m2}\vec{a}_2 + \ldots + \beta_{mn}\vec{a}_n$$

y entonces una combinación lineal de los $\vec{a}_i$, $\vec{b}_j$ tal como

$$\alpha_1\vec{a}_1 + \alpha_2\vec{a}_2 + \ldots + \alpha_n\vec{a}_n + \gamma_1\vec{b}_1 + \gamma_2\vec{b}_2 + \ldots + \gamma_m\vec{b}_m$$

se expresa así:

$$\alpha_1\vec{a}_1 + \alpha_2\vec{a}_2 + \ldots + \alpha_n\vec{a}_n + \gamma_1 (\beta_{11}\vec{a}_1) + \ldots + \beta_{1m}\vec{a}_n) + \gamma_2 (\beta_{21}\vec{a}_1 + \ldots + \beta_{2n}\vec{a}_n)) +$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$+ \gamma_m (\beta_{m1}\vec{a}_1 + \ldots + \beta_{mn}\vec{a}_n)$$

es decir, en definitiva como combinación lineal de los $\vec{a}_i$, y el teorema está demostrado.

Observación importante. Sea $X = [\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n]$ un sistema dado de vectores de un espacio vectorial V (K). Entonces, cabe la pregunta: ¿es posible engendrar el subespacio $X = [a_1, \vec{a}_2, \ldots, \vec{a}_n]$ con un número menor de vectores?

La respuesta a esta pregunta será dada cuando nos aboquemos un poco más adelante al estudio del concepto de la independencia lineal de vectores.

10.8. *Operaciones Elementales*

Daremos en primer lugar la definición siguiente.

Definición. Diremos que dos sistemas $X = \{\vec{a}_1, \ldots, \vec{a}_n\}$ e $Y = \{\vec{b}_1, \ldots, \vec{b}_m\}$ de vectores de un espacio vectorial V(K) son *equivalentes* si, y sólo si, todo vector de cada sistema es combinación lineal de los vectores del otro.

Se trata pues, de una relación de equivalencia en el sentido preciso del término. Esto es:

a) $X \sim X$

b) $X \sim Y \Rightarrow Y \sim X$

c) $X \sim Y$ e $Y \sim Z \Rightarrow X \sim Z$.

Sea ahora $X = \{a_1, a_2, \ldots a_n\}$ un sistema de generadores del subespacio $\overline{X}$.

Indicaremos en seguida ciertas transformaciones bastante simples del sistema dado X que forman sistemas equivalentes a él. Ellas son las que se indican en el teorema que sigue a continuación.

Teorema. No se modifica el subespacio engendrado por el sistema $x = \{\vec{a}_1, \ldots, \vec{a}_n\}$, cuando los vectores de este sistema experimentan una o varias de las siguientes transformaciones:

1) El cambio del orden de los generadores,
2) La multiplicación de un generador por un escalar diferente de cero,
3) La adición de uno de los generadores de un múltiplo de otro generador.

Demostración. Es preciso probar que todo vector $\vec{u} \in \overline{X}$, combinación lineal del sistema original $X = \{\vec{a}_1, \ldots, \vec{a}_n\}$ es también combinación lineal del nuevo sistema $Y = \{\vec{b}_1, \vec{b}_2, \ldots, b_n\}$.

En efecto, 1), resulta de la conmutatividad de la adición, puesto que los $\vec{b}_i$ son idénticos a los $\vec{a}_i$, en otro orden solamente. Para probar 2), basta reemplazar $\vec{a}_i$ por $\vec{b}_i = \lambda \vec{a}_i$, permaneciendo sin alteración los demás generadores.

Entonces se tiene

$$\vec{u} = \sum_{i=1}^{n} \alpha_i \vec{a}_i = \alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \ldots + \frac{\alpha_i}{\lambda} \vec{b}_i + \ldots + \alpha_n \vec{a}_n$$

Finalmente, para demostrar 3), basta sustituir $\vec{a}_i$ por $\vec{b}_i = a_i = \vec{a}_i + \lambda \vec{a}_j$, permaneciendo como en el caso 2) sin alteración el resto de los generadores. Entonces, se tiene

$$\vec{u} = \sum_{i=1}^{n} \alpha_i \vec{a}_i = \alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \ldots + \alpha_i \vec{b}_i +$$

$$+ \ldots + (\alpha_j - \lambda \alpha_i) \vec{a}_j + \ldots + \alpha_n a_n.$$

En consecuencia, hemos visto que toda combinación lineal del sistema dado X es también combinación lineal del sistema nuevo Y.

Por otra parte, es fácil ver que las operaciones 1), 2) y 3) son inversibles, esto es, las operaciones inversas son del mismo tipo. Por lo tanto, los sistemas de generadores X e Y son equivalentes, y el teorema está demostrado.

NOTA. Las transformaciones 1), 2) y 3) del enunciado del teorema las llamaremos *Operaciones Elementales.*

10.9. *Dependencia e Independencia Lineal*

En un espacio vectorial puede ocurrir que se tenga:

$$\alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \ldots + \alpha_n \vec{a}_n = \beta_1 \vec{a}_1 +$$

$\beta_2 a_2 + \ldots + \beta_n a_n.$

sin que sea simultáneamente.

$$\alpha_1 = \beta_1, \alpha_2 = \beta_2 \ldots \alpha_n = \beta_n$$

Así, pues, podemos decir que, dos combinaciones lineales de $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$, pueden ser iguales sin que los coeficientes correspondientes lo sean. En el plano ordinario, por ejemplo, los vectores $\vec{a}_1 = (0,1), \vec{a}_2 = (1,0), \vec{a}_3 = (1,1)$, verifican la relación:

$$1 \cdot \vec{a}_1 + 1 \cdot \vec{a}_2 + 0 \cdot \vec{a}_3 = 0 \cdot \vec{a}_1 + 0 \cdot \vec{a}_2 + 1 \cdot \vec{a}_3$$

sin que los coeficientes sean iguales.

Definición. Diremos que un sistema $X = \{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$ de un espacio vectorial V (K) es LINEALMENTE INDEPENDIENTE (abreviadamente, L, I.) si la igualdad

$$\alpha_i \vec{a}_1 + \alpha_2 \vec{a}_2 + \ldots + \alpha_n \vec{a}_n = \beta_1 \vec{a}_1 + \beta_2 \vec{a}_2 + \ldots + \beta_n \vec{a}_n$$

implica:

$$\alpha_1 = \beta_1, \alpha_2 = \beta_2, \ldots, \alpha_n = \beta_n$$

En otras palabras, un sistema $X = \{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$ es L. I., si cada vector $\vec{a}_i \in X$ no es una combinación lineal de los vectores restantes de X; esto es, que la igualdad:

$$\alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \ldots + \alpha_i \vec{a}_i + \ldots + \alpha_n \vec{a}_n = \vec{o}$$

no puede ser verificada más que para los coeficientes α, todos nulos. Por lo tanto, la sola combinación lineal de los vectores del sistema $X = \{a_1, a_2, \ldots, a_n\}$ que sea nula corresponde a los coeficientes nulos.

$\alpha_1 = \alpha_2 = \ldots = \alpha_i = \ldots = \alpha_n = 0$

Los vectores $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$, de un sistema $X = \{\vec{a}_1, \vec{a}_2, \ldots, a_n\}$ L. I., los llamaremos *vectores independientes.*

Ejemplo.

Sean en $\mathbb{R}^4$ los vectores $\vec{a}_1 = (1, 0, 0, 0)$, $\vec{a}_2 = (0, 1, 0, 0)$ y $\vec{a}_3 = (0, 0, 1, 0)$.

Si se tiene

$$\alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \alpha_3 \vec{a}_3 = (\alpha_1, \alpha_2, \alpha_3, 0) = o = (0, 0, 0, 0)$$

entonces, por la definición de igualdad, resulta

$$\alpha_1 = 0, \alpha_2 = 0, \alpha_3 = 0$$

Por consiguiente, los vectores $\vec{a}_1, \vec{a}_2, \vec{a}_3$ son linealmente independientes.

La negación lógica de la noción de sistema L. I. conduce a la noción de sistema **LINEALMENTE DEPENDIENTE** (abreviadamente, L. D).

En otros términos, decimos que el sistema $X = \{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$ es L. D. si, y sólo si, existe al menos un conjunto de escalares $\alpha_1, \alpha_2, \ldots \alpha_n$ *no todos nulos*, tales que:

$$\alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \ldots + \alpha_n \vec{a}_n = \vec{o}$$

Los vectores $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ de un sistema $X \{\vec{a}_1, \vec{a}_2, \ldots, a_n\}$ L. D., los llamaremos *vectores dependientes*

Ejemplo.

Probar que en $\mathbb{R}^4$, los vectores $\vec{a}_1 = (1, 1, 1, 0)$, $\vec{a}_2 = (0, 1, 1, 0)$ $\vec{a}_3 = (0, 1, 0, 0)$ y $a_4 = (0, 0, 1, 0)$ son linealmente dependientes.

En efecto, si $\alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \alpha_3 \vec{a}_3 + \alpha_4 \vec{a}_4 = \vec{o}$

o sea, $(\alpha_1, \alpha_1, + \alpha_2 + \alpha_3, \alpha_1 + \alpha_2 + \alpha_4, 0) = \vec{o} = (0, 0, 0, 0)$ entonces, por la definición de igualdad en $\mathbb{R}^n$, resulta:

$$\left.\begin{array}{rl} \alpha_1 & = o \\ \alpha_1 + \alpha_2 + \alpha_3 & = o \\ \alpha_1 + \alpha_2 + \alpha_4 & = o \\ o & = o \end{array}\right|$$

de donde,

$$\alpha_1 = 0$$
$$\alpha_2 = -\alpha_3$$
$$\alpha_3 = \alpha_4$$

Luego, por ejemplo, para $\alpha_1 = o$, $\alpha_2 = 2$, $\alpha_3 = -2$, $\alpha_4 = -2$, tendremos una solución del sistema y, por tanto, la combinación lineal nula siguiente:

$$o \cdot (1, 1, 1, 0) + 2(0, 1, 1, 0) - 2(0, 1, 0, 0) - 2(0, 0, 1, 0) = (0, 0, 0, 0)$$

en la cual los coeficientes no son todos nulos.

Luego, los vectores dados $\vec{a}_1, \vec{a}_2, \vec{a}_3$ y $\vec{a}_4$ son linealmente dependientes.

Teorema. Para que los vectores $\vec{a}_1, \vec{a}_2, \ldots \vec{a}_n$ sean linealmente dependientes, es necesario y suficiente que uno de ellos (al menos) sea combinación lineal de los restantes.

Demostración. Condición necesaria (sólo si). Supongamos que los vectores $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ sean linealmente dependientes. Entonces existen coeficientes $\alpha_1, \alpha_2 \ldots \alpha_n$ no todos nulos tales que:

$$\alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \ldots + \alpha_n \vec{a}_n = o$$

Sea α_k uno de los coeficientes que no son nulos, entonces podemos escribir

$$\alpha_k \vec{a}_k = - \vec{a}_1 - \alpha_2 \vec{a}_2 - \ldots - \vec{\alpha}_{k-1} a_{k-1}$$
$$\vec{\alpha}_{k-1} a_{k-1} - \ldots - \alpha_n \vec{a}_n$$

de donde.

$$\vec{a}_k = -\frac{\alpha_1}{\alpha_k}\vec{a}_1 - \frac{\alpha_2}{\alpha_k}\vec{a}_2 - \ldots - \frac{\alpha_{k-1}}{\alpha_k}\vec{a}_k - \frac{\alpha_n}{\alpha_k}\vec{a}_{k+1} - \ldots - \frac{\alpha_n}{\alpha_k}\vec{a}_n$$

Por tanto, el vector a_k es combinación lineal de los restantes vectores $\vec{a}_1, \vec{a}_2 \ldots, \vec{a}_{k-1}, \vec{a}_{k+1}, \ldots \vec{a}_n$

Condición suficiente (si). Supongamos ahora que uno de los vectores, por ejemplo, $\vec{a}_k$ es combinación lineal de los restantes.

Sea

$$\vec{a}_k = \alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \ldots + \alpha_{k-1} \vec{a}_{k-1} + a_{k-1} \vec{a}_{k-1} + \ldots + \alpha_n a_n$$

entonces, tenemos

$$\alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \ldots + \alpha_{k-1} \vec{a}_{k-1} + (-1)\vec{a}_k + a_{k-1} + \ldots + \alpha_n \vec{a}_n = \vec{o}$$

y como el coeficiente de a_k, que es (-1), es distinto de cero, entonces los vectores $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k, \ldots, \vec{a}_n$ son linealmente dependientes.

Este resultado y el anterior demuestran el teorema. Ahora, pasaremos a enunciar un teorema más fuerte que el anterior y que tiene muchas consecuencias importantes.

Teorema. Los vectores no nulos $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ de un espacio vectorial V(K) son linealmente dependientes si, y sólo si, algún $\vec{a}_k$ con $2 \leq k \leq n$, es combinación lineal de los precedentes; esto es, de $\vec{a}_1, \vec{a}_2, \ldots \vec{a}_{k-1}$.

Demostración. Condición suficiente. Supongamos que el vector $\vec{a}_k$ es combinación lineal de los precedentes $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_{k-1}$;

entonces,

$$\vec{a}_k = \alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \ldots + \alpha_{k-1} \vec{a}_{k-1}$$

o bien,

$$\vec{a_k} = \alpha_1 \vec{a_1} + \alpha_2 \vec{a_2} + \ldots + a_{k-1} + o \cdot \vec{a_{k+1}} + \ldots + o \cdot \vec{a_n}$$

y por lo tanto $\vec{a_k}$ es combinación lineal de los restantes; luego, por el teorema anterior, los vectores $\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}$ son linealmente dependientes.

Condición necesaria. Supongamos ahora que los vectores no nulos a_1, $\vec{a_2}, \ldots, \ldots, \vec{a_n}$ son linealmente dependientes, esto es, existen coeficientes no todos nulos tales que:

$$\alpha_1 \vec{a_1} + \alpha_1 + \alpha_2 \vec{a_2} + \ldots + \alpha_n \vec{a_n} = \vec{o}$$

Sea k el índice máximo tal que $\alpha_k \neq o$, y probemos que $k \neq 1$.

Si fuera $k = 1, \alpha_1, \neq 0, \alpha_2 = \alpha_3 = \ldots = \alpha_n = 0$, resultaría

$$\alpha_1 \vec{a_1} = \vec{o}$$

y como $\alpha_1 \neq 0$, tendríamos $\vec{a_1} = \vec{o}$, lo que es imposible puesto que por hipótesis los vectores no son nulos, luego debe ser necesariamente $k > 1$.

Ahora bien, como hemos dicho que $\alpha_k \neq 0$ es el último coeficiente no nulo, tendremos entonces

$$\alpha_1 \vec{a_1} + \alpha_2 \vec{a_2} + \ldots + \alpha_{k-1} + \alpha_k \vec{a_k} = \vec{o}$$

y de donde, por $\alpha_k \neq 0$; $\alpha_k = -\frac{\alpha_1}{\alpha_k} \vec{a_1} - \frac{\alpha_2}{\alpha_k} \vec{a_2} - \ldots - \frac{\alpha_{k-1}}{\alpha_k} \vec{a_{k-1}}$

luego, el vector $\vec{a_k}$ es combinación lineal de los precedentes, y las condiciones del teorema se cumplen.

Teorema. Si los vectores a_1, $\vec{a_2}, \ldots, \vec{a_n}$ son independintes, entonces ninguno de ellos es combinación lineal de los precedentes.

Demostración. En primer término deberemos observar que el vector nulo es dependiente, ya que la relación $\alpha_o = \vec{o}$ puede verificarse para $\alpha \neq 0$.

Por lo tanto, si por hipótesis los vectores $\vec{a_1}, \vec{a_2}, \ldots \vec{a_n}$ son independientes, entonces ninguno de ellos puede ser nulo.

Esto supuesto, por reducción al absurdo, si alguno de los vectores a_1, $\vec{a_2}, \ldots, \vec{a_n}$ fuera combinación lineal de los precedentes, entonces, por el teorema recién probado, los vectores $\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}$ serían linealmente dependientes, resultando así un absurdo contrario a la hipótesis de que los vectores $\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}$ eran linealmente independientes.

Este resultado contrario a la hipótesis, demuestra el teorema.

10.10. *Consecuencias de las definiciones de dependencia e independencia lineal.*

Comenzaremos indicando algunas consecuencias útiles de la definición de dependencia lineal.

1°. El vector cero $(\vec{0})$ es dependiente, puesto que la igualdad, $\alpha o = \vec{0}$ puede verificarse para $\alpha \neq o$

2°. Todo sistema de vectores que contiene al vector cero, es de pendiente,

En efecto, una manera posible de verificar la relación

$$\alpha_1 \vec{a_1} + \alpha_2 \vec{a_2} + \ldots + \alpha_i \vec{a_i} + \ldots + \alpha_n \vec{a_n} = \vec{0}$$

consiste en elegir nulos todos los coeficientes $\alpha_1, \alpha_2, \ldots, \alpha_n$ salvo el del vector nulo $\vec{a_i} = \vec{0}$

3°. Si un sistema $X = \{\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}\}$ es L. D. adjuntando otros vectores $\vec{b_1}, \vec{b_2}, \ldots, \vec{b_m}$, el sistema obtenido $Y = \{\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}, b_1, b_2, \ldots, b_m\}$ sigue siendo L.D.

En efecto, si el sistema X es L. D., entonces existe una combinación lineal nula con coeficientes $\alpha_1, \alpha_2, \ldots, \alpha_n$ no todos iguales a cero:

$$\alpha_1 \vec{a_1} + \alpha_2 \vec{a_2} + \ldots + \alpha_n \vec{a_n} = \vec{0}$$

Luego, también se verifica

$$\alpha_1 \vec{a_1}, \alpha_2 \vec{a_2} + \ldots + \alpha_n \vec{a_n} + 0 \cdot \vec{b_1} + 0 \cdot \vec{b_2} + \ldots + 0 \cdot \vec{b_m} = \vec{0}$$

con coeficientes no todos nulos.

Por consiguiente, el nuevo sistema $Y = \{\vec{a_1}, \ldots, \vec{a_n}, \vec{b_1}, \ldots, \vec{b_m}\}$ es también L. D.

4°. Sea $X = \{\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}\}$ un sistema L. D. y sea $\overline{X}$ el subespacio engendrado por X. Entonces probaremos que todo vector $u \in \overline{X}$ admite infinitas representaciones de la forma:

$$\vec{u} = \alpha_1 \vec{a_1} + \alpha_2 \vec{a_2} + \ldots + \alpha_n \vec{a_n}$$

En efecto, sabiendo, por otra parte, que los $\vec{a_i}$ (i = 1, 2, ... n) son linealmente dependientes, se tendrá

$$\beta_1 \vec{a_1} + \beta_2 \vec{a_2} + \ldots + \beta_n \vec{a_n} = \vec{o}$$

con algún coeficiente $\beta_i \neq 0$.

También, cualquiera sea γ, será

$$\gamma \beta_1 \vec{a_1} + \gamma \beta_2 \vec{a_2} + \ldots + \gamma \beta_i \vec{a_i} + \ldots + \gamma \beta_n \vec{a_n} = \vec{o}$$

De donde, resulta que

$$\vec{u} = \vec{u} + \vec{o} = (\alpha_1 + \gamma \beta_1) \vec{a_1} + \ldots + (\alpha_i + \gamma \beta_i) \vec{a_i} + \ldots + (\alpha_n + \gamma \beta_n) \vec{a_n}$$

Pero, para el β_i, $\neq 0$, el coeficiente $(\alpha_i + \gamma \beta_i)$ toma infinitos valores al variar γ ($\gamma \in K$, cuerpo infinito); por tanto, el vector $\vec{u} \in \overline{X}$, arbitrario, admite, como anunciamos, infinitas representaciones de la forma:

$$\vec{u} = \alpha_i \vec{a_i} + \alpha_2 \vec{a_2} + \ldots + \alpha_n \vec{a_n}$$

Ahora, pasamos a indicar también algunas consecuencias, que nos serán muy útiles, de la definición de independencia lineal.

1°. Un vector, no nulo es independiente.

En efecto, la igualdad $\alpha \vec{a} = \vec{o}$ implica, $\alpha = 0$, ó, $\vec{a} = \vec{o}$, y esta última alternativa está excluida en el enunciado.

2°. Todos los vectores de un sistema L. I. son distintos de cero, y distintos entre si dos a dos.

Esto es inmediato.

3°. Si $X = \{\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}\}$ es un sistema L. I., entonces todo vector $\vec{u}$ del subespacio $\bar{X}$ admite una única representación de la forma:

$$\vec{u} = \alpha_1 \vec{a_1} + \alpha_2 \vec{a_2} + \ldots + \alpha_n \vec{a_n}$$

En efecto, si suponemos que también admite, por el contrario, una segunda representación distinta:

$$\vec{u} = \beta_1 \vec{a_1} + \beta_2 \vec{a_2} + \ldots + \beta_n \vec{a_n}$$

entonces será:

$$(\alpha_1 - \beta_1)\vec{a_1} + (\alpha_2 - \beta_2)\vec{a_2} + \ldots + (\alpha_n - \beta_n)\vec{a_n} = \vec{o}$$

y por ser, por hipótesis, $\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}$ independientes, tendremos

$$\alpha_1 - \beta_1 = 0, \alpha_2 - \beta_2 = 0, \ldots, \alpha_n - \beta_n = 0$$

o sea, $\alpha_1 = \beta_1, \alpha_2 = \beta_2, \ldots, \alpha_n = \beta_n$ (absurdo)

Esta contradicción prueba que la representación es única; esto es, los α_i $(i = 1, 2, \ldots n)$ son unívocamente determinados.

Como consecuencia de lo anterior resulta que, si $\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}$ son linealmente independientes, entonces para que sea

$$\alpha_1 \vec{a_1} + \alpha_2 \vec{a_2} + \ldots + \alpha_n \vec{a_n} = \vec{o}$$

es necesario que sea

$$\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$$

En otros términos, de modo equivalente, los vectores $\vec{a_1}, \vec{a_2}, \ldots \vec{a_n}$, serán linealmente independientes si la única combinación lineal $\alpha_1 \vec{a_1} + \alpha_2 \vec{a_2} + \ldots + \alpha_n \vec{a_n}$ que es nula es aquella para la cual $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$.

4°. Todo subconjunto de un sistema L. I. es también un sistema L. I. En efecto, si el subconjunto considerado fuera, por el contrario, un sistema L. D., entonces, adjuntándoles los otros vectores restantes del sistema original se tendría otro sistema L. D., y lo cual es un absurdo contrario al hecho de que los vectores del sistema dado primitivamente eran L. I.

5°. En virtud de la definición de sistema $X = \{\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}\}$ linealmente independiente de vectores y del último teorema estudiado al final de la sección 9.9., podemos enunciar el teorema que sigue:

Teorema. Sea $X = \{\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}\}$, $n \geq 1$, un sistema de vectores de un espacio vectorial. Entonces, las siguientes condiciones son equivalentes:

1) $X = \{\vec{a_1}, \vec{a_2} \ldots, \vec{a_n}\}$ es L. I.

2) $\vec{a_1} \neq o$ y cada $\vec{a_i}$ $(2 \leq i \leq n)$ no es combinación lineal de los precedentes $\vec{a_1}, \vec{a_2}, \ldots \vec{a_{i-1}}$

3) Si $\alpha_1 \vec{a_1} + \alpha_2 \vec{a_2} + \ldots + \alpha_n \vec{a_n} = \vec{o}$, entonces $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$

4) Si $\alpha_1 \vec{a_1} + \alpha_2 \vec{a_2} + \ldots + \alpha_n \vec{a_n} = \beta_1 \vec{a_1} + \beta_2 \vec{a_2} + \ldots + \beta_n \vec{a_n}$, entonces es $\alpha_1 = \beta_1, \alpha_2 = \beta_2, \ldots, \alpha_n = \beta_n$

Demostraciones. En primer término, si $n = 1$, entonces la condición 1), $X = \{\vec{a_1}\}$, linealmente independiente implica la condición 2) que se reduce $\vec{a_1} \neq \vec{o}$, y ésta a su vez implica, la condición 3), ya que $\alpha_1 \vec{a_1} = \vec{o}$ trae como consecuencia $\alpha_1 = 0$, por ser $\vec{a_1} = \vec{o}$.

Asimismo, la condición 3) implica la 4), puesto que

$$\alpha_1 \vec{a_1} = \beta_1 \vec{a_1} = (\alpha_1 - \beta_1)\vec{a_1} = \alpha_1 - \beta_1 = 0, = \alpha_1 = \beta_1$$

Finalmente, la condición 4) implica la 1), puesto que si $\alpha_1 \vec{a_1} a_1 = \beta_1 \vec{a_1}$, que por 4) resulta $\alpha_1 = \beta_1$, entonces la relación $\alpha_1 \vec{a_1} = \beta_1 \vec{a_1}$ se reduce a $(\alpha_1 - \beta_1)\vec{a_1} = \vec{0}$, o sea $0 \cdot \vec{a_1} = \vec{0}$ y lo que implica $\vec{a_1} \neq \vec{o}$, es decir, el sistema $X = \{\vec{a_1}\}$, es linealmente independiente.

Se ve pues, el teorema es trivial para el caso $n = 1$.

Por este motivo, él será demostrado para $n > 1$. Deberemos entonces probar las siguientes implicaciones lógicas:

1) $=> 2)$, 2) $=> 3)$, 3) $=> 4)$ y 4) $=> 1)$

Tenemos:

1) $=> 2)$. Sea $X = \{\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}\}$ linealmente independiente.

Entonces, el vector nulo no pertenece al subconjunto X, $\vec{o} \notin X$; luego, en particular, resulta $\vec{a_1} \neq \vec{o}$, que es la primera afirmación de la condición 2).

Por otra parte, probaremos que un vector $\vec{a_i}$, $2 \leq i \leq n$, no puede ser combinación lineal de los precedentes $\vec{a_1}, \vec{a_2}, \ldots, \vec{a_{i-1}}$, porque si lo fuera, $\vec{a_i}$ sería también combinación lineal de los restantes vectores de X y entonces X no sería L. I., contra lo que se ha supuesto.

Esto prueba la segunda afirmación de la condición 2).

2) $\Longrightarrow$ 3). Supongamos que el sistema $X = \{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$ verifique ahora la condición 2); probaremos entonces que él también verifica la 3); esto es:

$$\alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \ldots + \alpha_n \vec{a}_n = \vec{o} \Longrightarrow \alpha_1 = \alpha_2 = \ldots \alpha_n = 0$$

Ahora, si así no fuese, es decir, si algún α_i no es nulo, consideraremos entre los coeficientes que no se anulan el de mayor índice. Sea éste α_i, $1 \leq i \leq n$

Si $i = 1$, entonces la implicación anterior se reduce a

$$\alpha_1 \vec{a}_1 = \vec{o}, \text{ con } \alpha_1 \neq 0$$

de donde resulta que, $\vec{a}_1 = \vec{o}$, y lo que no puede ser, por la primera afirmación de 2), que por hipótesis se cumple. Luego, $i \neq 1$.

Sea, pues, $i > 1$, entonces la implicación anterior se reduce a,

$$\alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \ldots + \alpha_i \vec{a}_i = \vec{o}$$

Despejando $\vec{a}_i$, lo que es posible ya que $\alpha_i \neq 0$, tenemos

$$a_i = -\frac{\alpha_1}{\alpha_i} a_i - \frac{\alpha_2}{\alpha_i} a_2 - \ldots - \frac{\alpha_{i-1}}{\alpha_i} a_{i-1}$$

Pero esto contradice la segunda afirmación de la condición 2). Este resultado y el anterior muestran que es absurdo suponer que hay algún coeficiente α_i que no se anula.

En consecuencia, hemos probado que 2) $=>$ 3)

3) $\Longrightarrow$ 4). Supongamos ahora que el sistema de vectores

$X = \{\vec{a}_1 \vec{a}_2, \ldots, \vec{a}_n\}$ verifica la condición 3) y probaremos que también verifica la condición 4).

En efecto, supongamos que se tiene

$$\alpha_1 a_1 + \alpha_2 \vec{a}_2 + \ldots + \alpha_n \vec{a}_n = \beta_1 \vec{a}_1 + \ldots + \beta_n \vec{a}_n$$

Esta igualdad se transforma en la siguiente:

$$(\alpha_1 - \beta_1) \vec{a}_1 + (\vec{\alpha_2} - \beta_2) \vec{a}_2 + \ldots + (\alpha_n - \beta_n) \vec{a}_n = \vec{o}$$

y como, por hipótesis, se cumple 3), resulta

$$\alpha_1 - \beta_1 = 0, \alpha_2 - \beta_2 = 0, \ldots, \alpha_n - \beta_n = 0$$

luego, $\alpha_1 = \beta_1, \alpha_2 = \beta_2, \ldots, \alpha_n = \beta_n$

Así hemos demostrado que 3) $=>$ 4).

Finalmente, deberemos probar que 4) $=>$ 1).

Es decir, supondremos ahora que el sistema $X = \{\vec{a}_1, \vec{a}_2, \ldots \vec{a}_n\}$ verifica la condición 4) y, entonces, deberemos probar que este sistema $X = \{\vec{a}_1, \vec{a}_2, \ldots \vec{a}_n\}$ de vectores del espacio vectorial considerado en el enunciado del teorema, es un sistema linealmente independiente.

Razonando por el absurdo, suponemos que tal sistema no es linealmente independiente, es decir, existe al menos un $\vec{a}_i \in X$ que es combinación lineal de los restantes vectores de X ; esto es

$$\vec{a}_i = \alpha_1 a_1 + \alpha_2 \vec{a}_2 + \ldots + \alpha_{i-1} \vec{a}_{i-1} + \alpha_{i+1} \vec{a}_{i+1} + \ldots + \alpha_n a_n + \beta_n \vec{a}_n$$

Esta igualdad puede escribirse en la forma:

$$0 \cdot \vec{a}_1 + \ldots + 0 \cdot \vec{a}_{i-1} + 1 \cdot \vec{a}_i + 0 \cdot \vec{a}_{i+1} + \ldots + 0 \cdot \vec{a}_n = \alpha_1 \vec{a}_1 + \ldots + \alpha_{i-1} \vec{a}_{i-1} + 0 \cdot a_i + \alpha_{i+1} \vec{a}_{i+1} + \alpha_n \vec{a}_n$$

Ahora bien, como estamos suponiendo por hipótesis que se verifica la condición 4), entonces deberá tenerse la igualdad de los coeficientes. Pero esto conduce al absurdo de que $1 = 0$, y lo que proviene de suponer que el sistema $X = \{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$ no es linealmente independiente.

Así hemos demostrado que 4) $=>$ 1).

Por consiguiente, las cuatro condiciones señaladas en el enunciado del teorema que hemos estudiado, son lógicamente equivalentes.

1) $\Longleftrightarrow$ 2) $\Longleftrightarrow$ 3) $\Longleftrightarrow$ 4)

Así por ejemplo, si queremos hacer ver la equivalencia de 2) con 3), tenemos:

Hemos visto que: 1) $\Longrightarrow$ 2), 2) $\Longrightarrow$ 3), 3) $\Longrightarrow$ 4), 4) $\Longrightarrow$ 1).

Luego, ya tenemos por una parte, que

(1) 2) $\Longrightarrow$ 3)

y por otra, por transitividad de la implicación lógica, se tiene: 3) $\Longrightarrow$ 4) y 4) $\Longrightarrow$ 1) trae como consecuencia que 3) $\Longrightarrow$ 1), y como, 1) $\Longrightarrow$ 2), resulta entonces que

(2) 3) $\Longrightarrow$ 2)

De (1) y (2), por la propiedad antisimétrica de la implicación lógica, obtenemos el resultado de que:

2) $\Longleftrightarrow$ 3)

Análogamente, se demuestra cualquiera de las otras equivalencias enunciadas.

Aplicación. El teorema recién probado nos hace ver que el siguiente proceso conduce a vectores $\vec{a}, \vec{b}, \vec{c}, \ldots$, linealmente independientes.

Este proceso se compone de los pasos que a continuación se indican:

1°. Se toma un vector $\vec{a} \neq \vec{o}$, arbitrario.

2°. Se toma otro vector de los dados que no pertenezca al subespacio [a] generado por el vector $\vec{a}$; esto es $\vec{b} \notin [\vec{a}]$, arbitrario.

3°. Se toma un tercer vector $\vec{c}$, arbitrario, tal que $\vec{c} \notin [a, b]$.

4°. Se toma, asimismo, un cuarto vector, si ello es posible, arbitrario también, que no pertenezca al subespacio $[\vec{a}, \vec{b}, \vec{c}]$; es decir

$$\vec{d} \in [\vec{a}, \vec{b}, \vec{c}].$$

5°. Proseguir, mientras sea posible, este proceso de la manera indicada.

10.11. *Ejercicios sobre dependencia e independencia lineal de vectores*

En el estudio teórico de las secciones anteriores 9.9. y 9.10. frecuentemente supusimos que podía determinarse la dependencia e independencia lineal de un sistema de vectores $X = \{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$ perteneciente a un espacio vectorial dado V(K).

En la mayoría de los casos, debido a la generalidad de las demostraciones, bastará suponer que se conocen escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$, no todos nulos, tales que

$$\alpha_1\vec{a}_1 + \alpha_2\vec{a}_2 + \ldots + \alpha_n\vec{a}_n = \vec{o}$$

si los vectores en cuestión son linealmente dependientes, o todos los escalares iguales a cero, si los vectores mencionados son linealmente independientes. Pero, cuando apliquemos esta discusión a ejemplos específicos de R^n, debemos dar un procedimiento, muy utilizado en la práctica, para hallar el valor de las constantes o escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$.

Veremos este procedimiento mediante ejemplos, o ejercicios.

1. Determinar si los siguientes vectores de $\mathbb{R}^4$: $\vec{a}_1 = (1, 2, -1, 4)$, $\vec{a}_2 = (2, 1, 3, 5)$, $\vec{a}_3 = (1, -1, 3, -1)$ y $\vec{a}_4 = (3, 0, 4, 0)$ son o no linealmente independientes.

Solución. Por definición sabemos que la dependencia e independencia lineal de estos vectores, es equivalente a la existencia o no existencia de números $\alpha_1, \alpha_2, \alpha_3, \alpha_4$, no todos cero, tales que se tenga

$$\alpha_1(1, 2, -1, 4) + \alpha_2(2, 1, 3, 5) + \alpha_3(1, -1, 3, -1) + \alpha_4(3, 0, 4, 0) = (0, 0, 0, 0)$$

Entonces,

$$(\alpha_1 + 2\alpha_2 + \alpha_3 + 3\alpha_4,\ 2\alpha_1 + \alpha_2 - \alpha_3,\ -\alpha_1 + 3\alpha_3 + 3\alpha_3 + 4\alpha_4,\ 4\alpha_1 + 5\alpha_2 - \alpha_3) = (0, 0, 0, 0)$$

Igualando las componentes o coordenadas correspondientes, obtenemos el sistema lineal homogéneo equivalente.

$$(*)\qquad \begin{array}{rcl} \alpha_1 + 2\alpha_2 + \alpha_3 + 3\alpha_4 & = & 0 \\ 2\alpha_1 + \alpha_2 - \alpha_3 & = & 0 \\ -\alpha_1 + 3\alpha_2 + 3\alpha_3 + 4\alpha_4 & = & 0 \\ 4\alpha_1 + 5\alpha_2 - \alpha_3 & = & 0 \end{array}$$

Para resolver este sistema indicaremos un procedimiento que nos va a permitir obtener todas las soluciones de (*), que llamaremos *método de reducción*, o de la *forma escalonada.*

Dejando la primera ecuación, por ejemplo, y eliminando entre ella y cada una de las demás las incógnitas α_1, obtenemos el siguiente nuevo sistema equivalente al (*).

$$(**)\qquad \begin{array}{rcl} \alpha_1 + 2\alpha_2 + \alpha_3 + 3\alpha_4 & = & 0 \\ -3\alpha_2 - 3\alpha_3 - 6\alpha_4 & = & 0 \\ 5\alpha_2 + 4\alpha_3 + 7\alpha_4 & = & 0 \\ -5\alpha_2 - 5\alpha_3 - 12\alpha_4 & = & 0 \end{array}$$

Dejando en este último sistema las dos primeras ecuaciones y eliminando entre la segunda de ellas y la tercera y cuarta restantes la incógnita α_2, se halla un nuevo sistema que es equivalente al anterior (**) y, por consiguiente también al (*).

$$(***)\qquad \begin{array}{rcl} \alpha_1 + 2\alpha_2 + \alpha_3 + 3\alpha_4 & = & 0 \\ -3\alpha_2 - 3\alpha_3 - 6\alpha_4 & = & 0 \\ -\alpha_3 - 3\alpha_4 & = & 0 \\ -2\alpha_3 - 6\alpha_4 & = & 0 \end{array}$$

y para lo cual fue necesario multiplicar la segunda ecuación del sistema (**) por $\frac{5}{3}$ y sumarla en seguida a la tercera ecuación, y restar finalmente esta misma segunda ecuación a la cuarta.

Siguiendo el mismo procedimiento, dejamos las tres primeras ecuaciones del sistema (***) y eliminamos ahora la incógnita α_3 entre la tercera y la cuarta ecuaciones, y lo cual conduce al nuevo sistema equivalente a los anteriores.

$$(****)\qquad \begin{array}{rcl} \alpha_1 + 2\alpha_2 + \alpha_3 + 3\alpha_4 & = & 0 \\ -3\alpha_2 - 3\alpha_3 - 6\alpha_4 & = & 0 \\ -\alpha_3 - 3\alpha_4 & = & 0 \\ 0\cdot\alpha_4 & = & 0 \end{array}$$

De este sistema "diagonal" o "escalonado" pueden obtenerse todas las soluciones. La última ecuación deja a la incógnita α_4 indeterminada, la tercera nos da la incógnita α_3 en función de α_4, $\alpha_3 = -3\,\alpha_4$, la segunda nos da α_2 en función de α_3 y α_4, lo que es lo mismo en función de α_4 solamente, $\alpha_2 = \alpha_4$, y finalmente, la primera ecuación nos da la incógnita $\alpha_1 = -2\,\alpha_4$. Estos valores

$$\alpha_1 = -2\,\alpha_4, \alpha_2 = \alpha_4, \alpha_3 = -3\,\alpha_4$$

satisfacen al sistema original (*), y por consiguiente, a la relación combinación lineal nula:

$$-2\,\alpha_4\,(1,2,-1,4) + \alpha_4\,(2,1,3,5) - 3\,\alpha_4\,(1,-1,3,-1) + \alpha_4\,(3,0,4,0) = (0,0,0,0)$$

Ya que α_4 está indeterminada, entonces puede tomar un valor arbitrario. Poniendo por ejemplo, $\alpha_4 = 1$, resultará:

$$-2(1,2,-1,4) + (2,1,3,5) - 3(1,-1,3,-1) + (3,0,4,0) = (0,0,0,0)$$

Por consiguiente, se concluye, del estudio hecho, que los cuatro vectores considerados son linealmente dependientes,. Por lo tanto, uno cualquiera de ellos puede ser expresado como combinación lineal de los otros tres, y el subespacio generado por ellos cuatro coincide con el subespacio generado por sólo tres de ellos.

2) Queremos determinar si los siguientes vectores de $\mathbb{R}^3$ son dependientes o no:

$\vec{a_1} = (1,2,-3)$, $\vec{a_2} = (1,-3,2)$ y $\vec{a_3} = (2,-1,5)$.

Solución: Procediendo como en el ejemplo anterior, tenemos

$$\alpha_1\,(1,2,-3) + \alpha_2\,(1,-3,2) + \alpha_3\,(2,-1,5) = (0,0,0)$$

luego el sistema

$$(*)\quad \begin{cases} \alpha_1 + \alpha_2 + 2\,\alpha_3 = 0 \\ 2\,\alpha_1 - 3\,\alpha_2 - \alpha_3 = 0 \\ -3\,\alpha_1 + 2\,\alpha_2 + 5\,\alpha_3 = 0 \end{cases}$$

y de éste al sistema equivalente que sigue

$$(**)\quad \begin{cases} \alpha_1 + \alpha_2 + 2\,\alpha_3 = 0 \\ -5\,\alpha_2 - 5\,\alpha_3 = 0 \\ 5\,\alpha_2 + 11\,\alpha_3 = 0 \end{cases}$$

y el cual a su vez se transforma en el sistema equivalente que se indica a continuación:

$$\begin{cases} \alpha_1 + \alpha_2 + 2\,\alpha_3 = 0 \\ -5\,\alpha_2 - 5\,\alpha_3 = 0 \\ 6\,\alpha_3 = 0 \end{cases}$$

De esta expresión diagonal o escalonada la única solución es: $\alpha_1 = 0, \alpha_2 = 0, \alpha_3 = 0$

Luego, los tres vectores dados son linealmente independientes.

Observación. Utilizando las operaciones elementales, que no modifican, como sabemos, el subespacio engendrado por un sistema de vectores, visto como teorema final en la sección 9.8., se puede indicar para el espacio $\mathbb{R}^n$ un criterio de independencia de uso frecuente en la práctica.

Sea $X = \{\vec{a_1}, \vec{a_2}, \ldots \vec{a_n}\}$ el sistema de generadores dado, e ignorando si él es dependiente o independiente.

Se considera estos vectores en un orden determinado; pero el subespacio $\bar{X}$ generado por ellos es, evidentemente, independiente de este orden.

Se utiliza las operaciones elementales para formar un nuevo sistema de generadores constituido generalmente por vectores independientes. Las relaciones de dependencia entre los vectores dados conducen a generadores nulos para el nuevo sistema. Naturalmente, estos generadores son descartados, ya que la presencia del vector nulo no modifica un subespacio.

Para aclarar las ideas, consideraremos nuevamente los dos ejemplos anteriores.

1) Sea en $\mathbb{R}^4$ el sistema formado por los vectores $\vec{a_1} = (1,2,-1,4)$, $\vec{a_2} = (2,1,3,5)$, $\vec{a_3} = (1,-1,3,-1)$ y $\vec{a_4} = (3,0,4,0)$

Entonces, fijado previamente un orden de ellos, escribimos sus respectivas componentes en la siguiente disposición:

$$\begin{pmatrix} 1 & 2 & -1 & 4 \\ 2 & 1 & 3 & 5 \\ 1 & -1 & 3 & -1 \\ 3 & 0 & 4 & 0 \end{pmatrix}$$

que llamaremos la *matriz* de las componentes, o más simplemente, la matriz de los vectores, que son sus respectivas filas o renglones.

Por medio de las operaciones elementales entre vectores, transformamos esta matriz en la siguiente:

$$\begin{pmatrix} 1 & 2 & -1 & 4 \\ 0 & -3 & 5 & -3 \\ 0 & -3 & 4 & -5 \\ 0 & -6 & 7 & -12 \end{pmatrix}$$

que se obtiene de la dada, dejando o conservando el primer renglón que corresponde al primer vector en el orden prefijado, y reemplazando los otros renglones de manera que sea cero su primera componente.

De esta segunda matriz pasamos ahora a la siguiente:

$$\begin{pmatrix} 1 & 2 & -1 & 4 \\ 0 & -3 & 5 & -3 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & -3 & -6 \end{pmatrix}$$

que se obtiene de la segunda, conservando el primer y segundo renglón y sustituyendo los renglones restantes de manera que sean ahora también cero su segunda componente.

Ahora, de esta tercera matriz se pasa a la cuarta que sigue a continuación:

$$\begin{pmatrix} 1 & 2 & -1 & 4 \\ 0 & -3 & 5 & -3 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

que se obtiene de la tercera dejando sin variar las componentes del primer, segundo y tercer renglón y sustituyendo los restantes de manera que sean ahora cero su tercera componente, además naturalmente de la primera y segunda componente nulas.

Como aquí el proceso de operaciones elementales ha terminado, decimos que la matriz dada primitivamente ha sido transformada en otra equivalente que tiene *forma triangular* o *forma escalonada*.

En el ejemplo presente, hemos mostrado que los renglones distintos de cero de la matriz en forma escalonada, es decir, de los vectores correspondientes del nuevo sistema de generadores, son independientes.

Este resultado se cumple en general, y por esto, lo estableceremos formalmente como un teorema en vista que se utilizará con frecuencia.
Teorema. Los renglones distintos de cero de una matriz en forma triangular o escalonada son linealmente independientes.

Volviendo a nuestro ejemplo, observamos además que el cuarto renglón del nuevo sistema de generadores es el vector cero. Luego, el sistema obtenido, equivalente al dado, es linealmente dependiente; resultado que está de acuerdo con el que antes habíamos obtenido.

2) Sea en $\mathbb{R}^3$ el sistema formado por $\vec{a_1} = (1, 2, -3)$, $\vec{a_2} = (1, -3, 2)$ y $\vec{a_3} = (2, -1, 5)$.

Entonces, tenemos

$$\begin{pmatrix} 1 & 2 & -3 \\ 1 & -3 & 2 \\ 2 & -1 & 5 \end{pmatrix}$$

o bien,

$$\begin{pmatrix} 1 & 2 & -3 \\ 0 & -5 & 5 \\ 0 & -5 & 11 \end{pmatrix}$$

por último

$$\begin{pmatrix} 1 & 2 & -3 \\ 0 & -5 & 5 \\ 0 & 0 & 6 \end{pmatrix}$$

que es una matriz escalonada en la cual ningún renglón es cero. Luego, los tres vectores dados son linealmente independientes, tal como antes se vio.

Observación. En el método de la matriz triangular o escalonada que acabamos de utilizar en los ejemplos anteriores, no pone de manifiesto los coeficientes de la combinación lineal nula en el caso de que el sistema dado de vectores es linealmente dependiente. Pero si las operaciones elementales que se van empleando las vamos anotando en todos los pasos que hay que dar para obtener la matriz en forma escalonada, entonces los coeficientes de la combinación lineal correspondiente pueden determinarse fácilmente.

En efecto, consideremos nuevamente el ejemplo 1) de esta misma sección. En este ejemplo se vio que en $\mathbb{R}^4$, los vectores $\vec{a_1} = (1, 2, -1, 4)$, $\vec{a_2} = (2, 1, 3, 5)$, $\vec{a_3} = (1, -1, 3, -1)$ y $\vec{a_4} = (3, 0, 4, 0)$ son linealmente dependientes, encontrándose la siguiente combinación lineal nula:

$$-2\,(1, 2, -1, 4) + (2, 1, 3, 5) - 3\,(1, -1, 3, -1) + (3, 0, 4, 0) = (0, 0, 0, 0)$$

Nuevamente, nos proponemos encontrar los coeficientes $\alpha_1 = -2$, $\alpha_2 = 1$, $\alpha_3 = -3$, $\alpha_4 = 1$ de esta combinación lineal aplicando el método de la matriz en forma triangular o escalonada. Tenemos:

$$\begin{array}{ll} a_1: & \\ a_2: & \\ a_3: & \\ a_4: & \end{array} \begin{pmatrix} 1 & 2 & -1 & 4 \\ 2 & 1 & 3 & 5 \\ 1 & -1 & 3 & -1 \\ 3 & 0 & 4 & 0 \end{pmatrix}$$

$$\begin{array}{ll} a_1 & : \\ a_2' = a_2 - 2\,a_1 & : \\ a_3' = a_3 - a_1 & : \\ a_4' = a_4 - 3\,a_1 & : \end{array} \begin{pmatrix} 1 & 2 & -1 & 4 \\ 0 & -3 & 5 & -3 \\ 0 & -3 & 4 & -5 \\ 0 & -6 & 7 & -12 \end{pmatrix}$$

$$
\begin{array}{lll}
a_1 & & \\
a'_2 & & \\
a''_3 = a'_3 - a'_2 & : & \\
a''_4 = a'_4 - 2\,a'_2 & : &
\end{array}
\begin{pmatrix} 1 & 2 & -1 & 4 \\ 0 & -3 & 5 & -3 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & -3 & -6 \end{pmatrix}
$$

$$
\begin{array}{lll}
a_1 & & \\
a'_2 & & \\
a''_3 & & \\
a'''_4 = a''_4 - 3a''_3 & :
\end{array}
\begin{pmatrix} 1 & 2 & -1 & 4 \\ 0 & -3 & 5 & -3 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix}
$$

Luego, concluimos que el vector $\vec{a}'''_4 = \vec{a}_4 - 3\vec{a}''_3 = 0$. Pero, según las notaciones que hemos hecho al aplicar operaciones elementales, se tiene

$$
\begin{aligned}
a''_4 &= a'_4 - 2\,a'_2 \\
&= (a_4 - 3a_1) - 2\,(a_2 - 2\,a_1) \\
&= a_1 - 2\,a_2 + a_4 \\
a''_3 &= a'_3 - a'_2 = (a_3 - a_1) - (a - 2\,a_1) \\
&= a_1 - a_2 + a_3
\end{aligned}
$$

Luego, sustituyendo estos valores en la relación

$$a'''_4 = a''_4 - 3\,a''_3 = \vec{o}$$

resulta,

$$(a_1 - 2\,a_2 + a_4) - 3\,(a_1 - a_2 + a_3) = \vec{o}$$

o sea,

$$a_1 - 2\,a_2 + a_4 - 3\,a_1 + 3\,a_2 - 3\,a_3 = \vec{o}$$

o bien,

$$-2\,\vec{a}_1 + \vec{a}_2 - 3\,\vec{a}_3 + \vec{a}_4 = \vec{o}$$

que es exactamente la misma que anteriormente habíamos encontrado.

10.12. *Número mínimo de vectores que genera un subespacio engendrado por una familia dada de vectores*

En la observación hecha al final de la sección 9.7. nos formulamos la pregunta de que si era posible, dado un sistema $X = \{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$, engendrar el mismo subespacio $\overline{X}$ con un número menor de vectores. Ahora bien, en posesión del concepto de dependencia e independencia lineal estamos en condiciones de dar respuesta a esa interrogante.

La idea básica es obvia; justamente eliminar tantos vectores linealmente dependientes en el sistema $X = \{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$ como sea posible. Así, por ejemplo, si $\vec{a}_n$ es linealmente dependiente respecto a los vectores $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_{n-1}$, esto es

$$\vec{a}_n = \alpha_1\vec{a}_1 + \alpha_2\vec{a}_2 + \ldots + \alpha_{n-1}\vec{a}_{n-1}$$

entonces, un vector cualquiera del subespacio $\overline{X}$, tal como

$$\vec{u} = \beta_1\vec{a}_1 + \beta_2\vec{a}_2 + \ldots + \beta_{n-1}\vec{a}_{n-1} + \beta_n\vec{a}_n$$

se podrá escribir también en la forma:

$$\vec{u} = \beta_1\vec{a}_1 + \beta_2\vec{a}_2 + \ldots + \beta_{n-1}\vec{a}_{n-1} + \beta_n\,(\alpha_1\vec{a}_1 + \alpha_2\vec{a}_2 + \ldots + \alpha_{n-1}\vec{a}_{n-1})$$

o sea,

$$\vec{u} = (\beta_1 + \alpha_1\beta_n)\,\vec{a}_1 + (\beta_2 + \alpha_2\beta_n)\,\vec{a}_2 + \ldots + (\beta_{n-1} + \alpha_{n-1}\beta_n)\,\vec{a}_{n-1}$$

resultado que nos muestra que el vector $\vec{u} \in \overline{X}$ es combinación lineal de los $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_{n-1}$, y por consiguiente,

$$[\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_{n-1}, \vec{a}_n] = [\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_{n-1}]$$

Cuando esto suceda, eliminaremos el vector $\vec{a}_n$ del sistema dado $X = \{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$. Por el contrario, si a_n no es linealmente dependiente respecto a los vectores $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_{n-1}$, entonces se lo conserva.

Ahora bien, si se repite este procedimiento con cada uno de los $\vec{a}_i$, eliminándolo cuando es linealmente dependiente y conservándolo cuando no lo es, es claro que se obtendrá finalmente un subconjunto del sistema $X = \{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$ que será linealmente independiente que engendra el mismo subespacio $\overline{X}$.

Esto nos mueve, pues, a enunciar el teorema siguiente.

Teorema Fundamental de la Independencia. Dado el sistema $X = \{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$ de vectores de un espacio vectorial V(K), podemos extraer de él un subconjunto de vectores linealmente independientes que genere el mismo subespacio $\overline{X} = [\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n]$.

Dicho subconjunto, que puede eventualmente coincidir con el conjunto dado X, se obtiene eliminando sucesivamente aquellos vectores $\vec{a}_i$ que sean combinaciones lineales de los anteriores; esto es, de los precedentes.

Demostración. Utilizaremos en esta demostración el proceso que se dio en la aplicación correspondiente al último teorema de la sección 10.10., para obtener sistemas de vectores linealmente independientes. Así tendremos:

Supongamos que sea $\vec{a}_1 \neq \vec{o}$, arbitrario (téngase presente que podemos renumerar de nuevo, en otro orden, los vectores del sistema dado $X = \{\vec{a}_1\,\vec{a}_2, \ldots, \vec{a}_n\}$).

Supongamos que, si $\vec{a}_2$ es combinación lineal de $\vec{a}_1$

$$\vec{a}_2 = \alpha_1 \vec{a}_1$$

entonces los subespacios $[\vec{a}_1]$ y $[\vec{a}_1, \vec{a}_2]$ coinciden

$$[\vec{a}_1, \vec{a}_2] = [\vec{a}_1]$$

y no tiene, por tanto, objeto conservar el vector $\vec{a}_2$ para generar el subespacio $[\vec{a}_1]$. Por el contrario, si $\vec{a}_2$ no es combinación lineal de $\vec{a}_1$, entonces el sistema $\{\vec{a}_1, \vec{a}_2\}$ es linealmente independiente y el vector $\vec{a}_2$ habrá que conservarlo, aumentando el subespacio generado que ahora es $[\vec{a}_1, \vec{a}_2]$.

Prosiguiendo así, supongamos que, renumerando de nuevo los vectores si fuese necesario, hayamos llegado a obtener los vectores $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_{k-1}$, tales que ninguno de ellos sea combinación lineal de los restantes. Ahora, si el siguiente $\vec{a}_k$ es combinación lineal de los $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_{k-1}$, se tendrá

$[\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_{k-1}, \vec{a}_k] = [\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_{k-1}]$, y dejaremos de lado el vector $\vec{a}_k$; en caso contrario, al agregarlo aumenta el subespacio generado.

Procederemos de este modo hasta que lleguemos a obtener todo el subespacio $\overline{X} = [\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n]$ generado por el sistema $X = \{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$; y suponiendo que esto se obtiene en la etapa m, concluiremos que:

$$[\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m] = [\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n]$$

en donde $m \leq n$.

Finalmente probaremos por inducción sobre m que el sistema $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\}$ así obtenido es linealmente independiente.

En efecto, primeramente, $\vec{a}_1$ es independiente por haberlo tomado diferente del vector $\vec{o}$. Luego, el sistema $\{\vec{a}_1\}$ es linealmente independiente.

Supongamos en seguida haber demostrado que el sistema parcial $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_{m-1}\}$ es L. I., y probaremos que también lo es el sistema $\{\vec{a}_1, a_2, \ldots, \vec{a}_{m-1}, \vec{a}_m\}$.

Por el contrario, si este último sistema fuera dependiente, resultaría entonces que $\vec{a}_m$ es combinación lineal de los vectores $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_{m-1}$, lo cual es absurdo por la construcción que hemos venido haciendo de estos vectores. Por consiguiente, por inducción resulta así probado que el sistema $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\}$ es L. I. y el teorema queda demostrado.

En resumen: en el subespacio $\overline{X}$ engendrado por un sistema de n generadores $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$, todo subsistema linealmente independiente $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\}$ (renumerando los vectores $\vec{a}_i$ si es preciso) que genere el mismo subespacio $\overline{X}$, contiene un número m de vectores que verifica la relación $m \leq n$.

Ejemplo:

En el espacio $\mathbb{R}^5$ se da el sistema $\vec{a}_1 = (1,2,-4,3,1)$, $\vec{a}_2 = (2,5,-3,4,8)$, $\vec{a}_3 = (6,17,-7,10,22)$ y $\vec{a}_4 = (1,3,-3,2,0)$.

Si el sistema formado por estos cuatro vectores no es L.I., entonces hallar un subsistema de él formado por vectores linealmente independientes que engendre el mismo subespacio formado por los cuatro vectores dados.

Solución: Utilizaremos el método de la matriz escalonada. Tendremos entonces:

$$\begin{array}{ll} \vec{a}_1 & \\ \vec{a}_2 & \\ \vec{a}_3 & \\ \vec{a}_4 & \end{array} \begin{pmatrix} 1 & 2 & -4 & 3 & 1 \\ 2 & 5 & -3 & 4 & 8 \\ 6 & 17 & -7 & 10 & 22 \\ 1 & 3 & -3 & 2 & 0 \end{pmatrix}$$

$$\begin{array}{ll} \vec{a}_1 & \\ \vec{a}'_2 = \vec{a}_2 - 2\vec{a}_1 & \\ \vec{a}'_3 = \vec{a}_3 - 6\vec{a}_1 & \\ \vec{a}'_4 = \vec{a}_4 - \vec{a}_1 & \end{array} \begin{pmatrix} 1 & 2 & -4 & 3 & 1 \\ 0 & 1 & 5 & -2 & 6 \\ 0 & 5 & 17 & -8 & 16 \\ 0 & 1 & 1 & -1 & -1 \end{pmatrix}$$

$$\begin{array}{ll} \vec{a}_1 & \\ \vec{a}'_2 & \\ \vec{a}''_3 = \vec{a}'_3 - 5\vec{a}'_2 & \\ \vec{a}''_4 = \vec{a}'_4 - \vec{a}'_2 & \end{array} \begin{pmatrix} 1 & 2 & -4 & 3 & 1 \\ 0 & 1 & 5 & -2 & 6 \\ 0 & 0 & -8 & 2 & -14 \\ 0 & 0 & -4 & 1 & -7 \end{pmatrix}$$

$$\begin{array}{ll} a_1 & \\ a'_2 & \\ a''_3 & \\ a'''_4 = a''_4 - \frac{1}{2}a''_3 & \end{array} \begin{pmatrix} 1 & 2 & -4 & 3 & 1 \\ 0 & 1 & 5 & -2 & 6 \\ 0 & 0 & -8 & 2 & -14 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Este resultado nos muestra que solamente tres de los cuatro vectores dados son linealmente independientes.

Si queremos establecer la relación de dependencia entre los cuatro vectores dados, tendremos:

$$\vec{a}'''_4 = \vec{a}''_4 - \tfrac{1}{2}\vec{a}''_3 = \vec{o}$$

pero, $\vec{a}''_4 = \vec{a}'_4 - \vec{a}'_2 = \vec{a}_4 - \vec{a}_1 - (\vec{a}_2 - 2\vec{a}_1) = \vec{a}_4 - \vec{a}_2 + \vec{a}_1$

$$\vec{a}''_3 = \vec{a}'_3 - 5\vec{a}'_2 = \vec{a}_3 - 6\vec{a}_1 - 5(\vec{a}_2 - 2\vec{a}_1) = \vec{a}_3 + 4\vec{a}_1 - 5\vec{a}_2$$

Luego, resulta

$$\vec{a}_4 - \vec{a}_2 + \vec{a}_1 - \frac{1}{2}(\vec{a}_3 + 4\vec{a}_1 - 5\vec{a}_2) = \vec{o}$$

o sea,

$$-2\vec{a}_1 + 3\vec{a}_2 - \vec{a}_3 + 2\vec{a}_4 = \vec{o}$$

Tres cualesquiera de estos cuatro vectores engendra el mismo subespacio; pero el sistema formado por los vectores, $\vec{a}_1$, $\vec{a}'_2$ y $\vec{a}''_3$, esto es por los vectores:

$$\vec{a}_1 = (1, 2, -4, 3, 1)$$
$$\vec{a}'_2 = \vec{a}_2 - 2\vec{a}_1 = (0, 1, 5, -2, 6)$$
$$\vec{a}''_3 = \vec{a}_3 + 4\vec{a}_1 - 5\vec{a}_2 = (0, 0, -8, 2, -14)$$

que verifican la condición de independencia es más cómodo en casi todas las aplicaciones.

10.13. *Axioma de la dimensión*

Los conjuntos linealmente independientes de vectores, gozan de una posición privilegiada en el estudio de los espacios vectoriales, y entre tales conjuntos, aquellos que generan el espacio entero son particularmente importantes. Sobre tales conjuntos o sistemas de vectores nos ocuparemos en esta sección.

Sea, pues, V(K) un espacio vectorial arbitrario; y consideremos los conjuntos o sistemas de vectores linealmente independientes que podemos detraer o deducir sobre V(K). Sólo dos hipótesis son posibles. Ellas son:

1°. El número de vectores independientes está acotado; esto es, podemos encontrar al menos un sistema de n vectores $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ linealmente independientes, tal que si agregamos un nuevo vector cualquiera del espacio V(K), el nuevo sistema $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n, \vec{u}\}$ es linealmente dependiente. Entonces, diremos que el sistema $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ es de *orden máximo*, o *maximal*, o también que es una **BASE** del espacio V(K).

Si un espacio vectorial V(K) tiene n elementos en una base, entonces al número n lo llamaremos la **DIMENSION** del espacio V(K), y la indicaremos escribiendo.

$$\dim V(K) = n$$

Un espacio vectorial V(K) que verifique esta primera hipótesis o postulado (axioma), se dirá de *dimensión finita.*

Nota. Por convención, al espacio vectorial que consiste sólo del vector cero, le asignaremos la dimensión 0.

$$\dim(\vec{o}) = 0$$

2°. El número de vectores independientes detraídos del espacio V(K) no está acotado; esto es, cualquiera que sea el número natural n, existe al menos un sistema de n vectores independientes.

Un espacio vectorial V(K) que verifique esta segunda hipótesis o postulado (axioma), se dirá de *dimensión infinita.*

Definición. Llamaremos **BASE** de un espacio vectorial a cualquier subconjunto suyo linealmente independiente que engendra al espacio entero. Es decir, si $B \subset V$ es una base del espacio V(K), entonces todo vector $\vec{u} \in V(K)$ puede escribirse de una manera y de una sola, como combinación lineal de un número finito de elementos de B.

En resumen: Las dos hipótesis 1° y 2° consideradas anteriormente nos mueven a postular (aunque ello se demuestre) que todo espacio vectorial posee al menos una base finita o infinita. A tal postulado lo llamaremos el **AXIOMA DE LA DIMENSION.**

Nosotros nos ocuparemos exclusivamente de espacios vectoriales V(K) de dimensión finita, es decir, de los casos donde $\dim V(K) = n \geq 0$, n número natural.

Por brevedad diremos que V(K) tiene dimensión n ($n \geq 1$) si las condiciones siguientes se verifican:

a) Si hay n vectores linealmente independientes, y

b) Si no hay n +1 vectores linealmente independientes.

Nota. La dimensión de un subespacio $\overline{X}$ de V(K) será por definición la dimensión de $\overline{X}$ considerado como espacio vectorial; es claro que

$$\dim \overline{X} \leq \dim V(K)$$

Ejemplos:

1. En el plano $\mathbb{R}^2$, los vectores $\vec{e}_1 = (1, 0)$ y $\vec{e}_2$ $(0, 1)$ constituyen una base de este espacio, ya que

$$\vec{u} = (a, b) = a\vec{e}_1 + b\vec{e}_2$$

es la única forma de expresar un vector arbitrario de $\mathbb{R}^2$ como combinación lineal de $\vec{e}_1$ y $\vec{e}_2$.

Esta base particular, la llamaremos *base canónica* para $\mathbb{R}^2$.

Asimismo, en el espacio ordinario $\mathbb{R}^3$ de la geometría analítica, una base está constituida por los vectores $\vec{e}_1 = (1, 0, 0)$, $\vec{e}_2 = (0, 1, 0)$ y $\vec{e}_3 = (0, 0, 1)$, y que es su base canónica. Para $\vec{u} = (a, b, c) \in \mathbb{R}^3$, arbitrario, se tiene:

$$\vec{u} = (a, b, c) = a\vec{e}_1 + b\vec{e}_2 + c\vec{e}_3$$

como única forma de expresar dicho vector como una combinación lineal

de $\vec{e}_1, \vec{e}_2$ y $\vec{e}_3$.

En general, en el espacio vectorial numérico $\mathbb{R}^n$ de todas las n-uplas de números reales, los vectores $\vec{e}_1 = (1, 0, \ldots, 0)$, $\vec{e}_2 = (0, 1, 0, \ldots, 0), \ldots, \vec{e}_n = (0, 0, \ldots, 0, 1)$, es decir, los vectores que tienen una componente igual a 1 y las demás igual a cero, constituyen una base, la base canónica para $\mathbb{R}^n$.

En efecto, es fácil constatar que ellos son linealmente independientes, y además, cualquier otro vector de R^n se expresa de manera única como combinación lineal de ellos; esto es:

$$(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n) = a_1 \vec{e}_1 + a_2 \vec{e}_2 + \ldots + a_n \vec{e}_n$$

2. No hay que creer, sin embargo, que en cada espacio vectorial existe una sola base. Así por ejemplo, en el espacio ordinario $\mathbb{R}^3$ de la geometría analítica, también los vectores $\vec{a}_1 = (2, -1, 0)$, $\vec{a}_2 = (1, -1, 1)$ y $\vec{a}_3 = (0, 2, 3)$ constituyen una base de $\mathbb{R}^3$.

En efecto, bastará probar que ellos son L.I. Tendremos:

$$\alpha_1 (2, -1, 0) + \alpha_2 (1, -1, 1) + \alpha_3 (0, 2, 3) = (0, 0, 0)$$

o sea,

$$(2\alpha_1 + \alpha_2, -\alpha_1 - \alpha_2 + 2\alpha_3, \alpha_2 + 3\alpha_3) = (0, 0, 0)$$

de donde el sistema lineal homogéneo siguiente:

$$\left.\begin{array}{rcl} -\alpha_1 - \alpha_2 + 2\alpha_3 &=& 0 \\ 2\alpha_1 + \alpha_2 &=& 0 \\ \alpha_2 + 3\alpha_3 &=& 0 \end{array}\right|$$

que es equivalente a

$$\left.\begin{array}{rcl} -\alpha_1 - \alpha_2 + 2\alpha_3 &=& 0 \\ -\alpha_2 + 4\alpha_3 &=& 0 \\ \alpha_2 + 3\alpha_3 &=& 0 \end{array}\right|$$

y el cual se transforma a su vez en el siguiente que le es equivalente.

$$\left.\begin{array}{rcl} -\alpha_1 - \alpha_2 + 2\alpha_3 &=& 0 \\ -\alpha_2 + 4\alpha_3 &=& 0 \\ 7\alpha_3 &=& 0 \end{array}\right|$$

y que evidentemente admite como única solución $\alpha_1 = \alpha_2 = \alpha_3 = 0$.

Resultado que muestra que los vectores dados $\vec{a}_1 = (2, -1, 0)$, $\vec{a}_2 = (1, -1, 1)$ y $\vec{a}_3 = (0, 2, 3)$ son independientes y forman, por esto, una base para $\mathbb{R}^3$.

También en $\mathbb{R}^n$, los vectores

$\vec{f}_1 = (1, 0, \ldots, 0)$
$\vec{f}_2 = (1, 1, 0, \ldots, 0)$
$f_3 = (1, 1, 1, 0, \ldots, 0)$
. . . . . .
. . . . . .
$f_n = (1, 1, 1, \ldots, 1)$

donde f_i es una n-upla ordenada de números reales que contiene a 1 en los i primeros lugares y ceros en los lugares restantes:

$f_i = (1, 1, \ldots, 1, 0, \ldots, 0)$

constituyen una base para el espacio vectorial numérico $\mathbb{R}^n$. ¡Pruébelo como ejercicio!

3) Si denotamos con P_n el conjunto de todos los polinomios de la forma

(1) $\vec{a}_{n-1} x^{n-1} + a_{n-2} x^{n-2} + \ldots + a_1 x + a_0$

donde $a_0, a_1, a_2, \ldots, a_{n-1}$ son números reales arbitrarios, y n es un entero positivo, entonces, como ya sabemos, P_n es un espacio vectorial.

Los polinomios $1, x, x^2, \ldots, x^{n-1}$ forman una base para este espacio P_n, ya que cada polinomio con coeficientes reales y de grado menor que n, junto con el polinomio nulo, puede escribirse de una y sólo de una manera de la forma (1).

En cambio, el conjunto P de todos los polinomios con coeficientes reales de todos los grados posibles, forman también un espacio vectorial, pero de dimensión infinita, y tiene como base el conjunto

$\{1, x, x^2, \ldots, x^n, \ldots\}$

4) En el espacio $\mathbb{R}^4$ determinar una base del subespacio engendrado por los vectores $\vec{a}_1 = (1, 2, 2, 1)$, $\vec{a}_2 = (5, 6, 6, 5)$, $\vec{a}_3 = (-1, -3, 4, 0)$ y $\vec{a}_4 = (0, 4, -3, -1)$

Solución: Disponemos los cálculos de la siguiente manera:

$$\begin{array}{l} \vec{a}_1: \\ \vec{a}_2: \\ \vec{a}_3: \\ \vec{a}_4: \end{array} \begin{pmatrix} 1 & 2 & 2 & 1 \\ 5 & 6 & 6 & 5 \\ -1 & -3 & 4 & 0 \\ 0 & 4 & -3 & -1 \end{pmatrix}$$

$$
\begin{array}{ll}
\vec{a}_1: & \begin{pmatrix} 1 & 2 & 2 & 1 \\ 0 & -4 & -4 & 0 \\ 0 & -1 & 6 & 1 \\ 0 & 4 & -3 & -1 \end{pmatrix} \\
\vec{a}'_2 := \vec{a}_2 - 5\vec{a}_1: & \\
\vec{a}'_3 = \vec{a}_3 + \vec{a}_1: & \\
\vec{a}_4: &
\end{array}
$$

$$
\begin{array}{ll}
\vec{a}_1: & \begin{pmatrix} 1 & 2 & 2 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 6 & 1 \\ 0 & 4 & -3 & -1 \end{pmatrix} \\
\vec{a}'_2 = -\frac{1}{4}\vec{a}'_2: & \\
\vec{a}'_3: & \\
\vec{a}_4: &
\end{array}
$$

$$
\begin{array}{ll}
\vec{a}_1: & \begin{pmatrix} 1 & 2 & 2 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 7 & 1 \\ 0 & 0 & -7 & -1 \end{pmatrix} \\
\vec{a}''_2: & \\
\vec{a}''_3 = \vec{a}'_3 + \vec{a}''_2: & \\
\vec{a}''_4 = \vec{a}_4 - 4\vec{a}''_2: &
\end{array}
$$

$$
\begin{array}{ll}
\vec{a}_1: & \begin{pmatrix} 1 & 2 & 2 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 7 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\
\vec{a}''_2: & \\
\vec{a}''_3: & \\
\vec{a}''_4 = \vec{a}''_4 + \vec{a}''_3: &
\end{array}
$$

El subespacio es de dimensión tres, y una base está formada por los vectores:

$\vec{a}_1 = (1, 2, 2, 1)$, $\vec{a}''_2 = (0, 1, 1, 0)$ y $\vec{a}''_3 = (0, 0, 7, 1)$.

Existe, por lo tanto, una relación de dependencia entre los cuatro vectores dados $\vec{a}_1, \vec{a}_2, \vec{a}_3$ y $\vec{a}_4$.

Para hallar esta relación se tiene:

$$\vec{a}''_4 = \vec{a}''_4 + \vec{a}''_3 = \vec{o}$$

pero, $\vec{a}''_4 = \vec{a}_4 - 4\vec{a}''_2 = \vec{a}_4 - 4.\left(-\frac{1}{4}\vec{a}'_2\right) = \vec{a}_4 + \vec{a}'_2 = \vec{a}_4 + \vec{a}_2 - 5\vec{a}_1$

y $\vec{a}''_3 = \vec{a}'_3 + \vec{a}''_2 = \vec{a}_3 + \vec{a}_1 - \frac{1}{4}\vec{a}'_2 = \vec{a}_3 + \vec{a}_1 - \frac{1}{4}(\vec{a}_2 - 5\vec{a}_1) = \vec{a}_3 + \frac{9}{4}\vec{a}_1 - \frac{1}{4}\vec{a}_2$

Luego, sustituyendo, se encuentra

$$\vec{a}_4 + \vec{a}_2 - 5\vec{a}_1 + \vec{a}_3 + \frac{9}{4}\vec{a}_1 - \vec{a}_2 = \vec{o}$$

o sea, $4\vec{a}_4 + 4\vec{a}_2 - 20\vec{a}_1 + 4\vec{a}_3 + 9\vec{a}_1 - \vec{a}_2 = \vec{o}$

de donde, la relación de dependencia lineal

$$-11\vec{a}_1 + 3\vec{a}_2 + 4\vec{a}_3 + 4\vec{a}_4 = \vec{o}$$

Tres cualesquiera de los vectores dados constituyen una base para el subespacio generado por todos ellos; pero la base (1, 2, 2, 1), (0,1,1,0) y (0, 0, 7, 1) indicada anteriormente que verifica el criterio de independencia es bastante cómoda en casi todas las aplicaciones.

4) Sea V ($\mathbb{R}$) el espacio vectorial de los polinomios de grado menor o igual a tres sobre el cuerpo $\mathbb{R}$ de los números reales.

Determinar si los vectores:

$\vec{a}_1 = \vec{x}^3 - 2x^2 + 4x + 1$, $\vec{a}_2 = 2x^3 - 3x^3 - 3x^2 + 9x - 1$, $\vec{a}_3 = x^3 + 6\vec{x} - 5$ y $\vec{a}_4 = 2x^3 - 5\vec{x}^2 + 7x + 5$, son independientes o dependientes.

Hallar además, una base y la dimensión del subespacio engendrado por el sistema $X = \{\vec{a}_1, \vec{a}_2, \vec{a}_3, \vec{a}_4\}$.

Solución. Tomando una combinación lineal de los polinomios $\vec{a}_1$, $\vec{a}_2$, $\vec{a}_3$ y $\vec{a}_4$, e igualándola al polinomio nulo (todos los coeficientes son ceros), tendremos:

$$\alpha(x^3 - 2x^2 + 4x + 1) + \beta(2x^3 - 3x^2 + 9x - 1) + \gamma(x^3 + 6x - 5) + \delta(2x^3 - 5x^2 + 7x + 5) = \vec{0}$$

o sea,

$$(\alpha + 2\beta + \gamma + 2\delta)x^3 + (-2\alpha - 3\beta - 5\delta)x^2 + (4\alpha + 9\beta + 6\gamma + 7\delta)x + (\alpha - \beta - 5\gamma + 5\delta) = \vec{0}$$

Ahora bien, como un polinomio es nulo cuando todos sus coeficientes son ceros, se tendrá entonces el sistema lineal homogéneo para los escalares $\alpha, \beta, \gamma, \delta$ de la combinación lineal:

$$\begin{array}{r} \alpha + 2\beta + \gamma + 2\delta = 0 \\ -2\alpha - 3\beta \quad\quad - 5\delta = 0 \\ 4\alpha + 9\beta + 6\gamma + 7\delta = 0 \\ \alpha - \beta - 5\gamma + 5\delta = 0 \end{array}$$

Aplicando el método de reducción, o de la forma escalonada, se encuentra

$$\begin{array}{r} \alpha + 2\beta + \gamma + 2\delta = 0 \\ \beta + 2\gamma - \delta = 0 \\ \beta + 2\gamma - \delta = 0 \\ -3\beta - 6\gamma + 3\delta = 0 \end{array}$$

y de donde,

$$\begin{array}{r} \alpha + 2\beta + \gamma + 2\delta = 0 \\ \beta + 2\gamma - \delta = 0 \\ 0 = 0 \\ 0 = 0 \end{array}$$

o simplemente,

$$\begin{array}{r} \alpha + 2\beta + \gamma + 2\delta = 0 \\ \beta + 2\gamma - \delta = 0 \end{array}$$

Como este sistema en forma escalonada tiene dos incógnitas libres (es decir, que disponemos a voluntad), tal sistema tiene entonces una solución distinta de cero; esto es, que la combinación lineal nula

$$\alpha \vec{a}_1 + \beta \vec{a}_2 + \gamma \vec{a}_3 + \delta \vec{a}_4 = 0$$

no implica la solución nula $\alpha = \beta = \gamma = \delta = o$. Luego, los cuatro polinomios dados son linealmente dependientes.

Las infinitas soluciones del sistema escalonado se encuentran como sigue:

$$\begin{array}{l} 2\beta + \gamma = -\alpha - 2\delta \\ \beta + 2\gamma = \delta \end{array}$$

$$\begin{array}{l} \beta + \gamma = -\dfrac{\alpha}{3} - \dfrac{\delta}{3} \\ \beta - \gamma = -\alpha - 3\delta \end{array}$$

$$\begin{array}{l} \beta = -\dfrac{2}{3}\alpha - \dfrac{5}{3}\delta \\ \gamma = \dfrac{1}{3}\alpha + \dfrac{4}{3}\delta \end{array}$$

Una solución será:

$$\alpha = 1, \beta = 1, \gamma = -1, \delta = -1$$

Otra manera de probar la dependencia de estos cuatro polinomios, es como se indica a continuación.

Representemos cada polinomio dado como un vector, en este caso particular, de $\mathbb{R}^4$ cuyas componentes son sus coeficientes; esto es,

$\vec{a}_1 = x^3 - 2x^2 + 4x + 1 = (1, -2, 4, 1)$
$\vec{a}_2 = 2x^3 - 3x^2 + 9x - 1 = (2, -3, 9, -1)$
$\vec{a}_3 = x^3 + 6x - 5 = (1, 0, 6, -5)$
$\vec{a}_4 = 2x^3 - 5x^2 + 7x + 5 = (2, -5, 7, 5)$

Formando en seguida la matriz cuyos renglones son estos vectores, y llevándola después a la forma triangular o escalonada se encuentra:

$$\begin{pmatrix} 1 & -2 & 4 & 1 \\ 2 & -3 & 9 & -1 \\ 1 & 0 & 6 & -5 \\ 2 & -5 & 7 & 5 \end{pmatrix}$$

$$\begin{pmatrix} 1 & -2 & 4 & 1 \\ 0 & 1 & 1 & -3 \\ 0 & 2 & 2 & -6 \\ 0 & -1 & -1 & 3 \end{pmatrix}$$

$$\begin{pmatrix} 1 & -2 & 4 & 1 \\ 0 & 1 & 1 & -3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Como esta matriz escalonada tiene filas nulas (dos en este caso), los vectores, es decir los polinomios correspondientes son dependientes. Luego los cuatro polinomios dados generan un subespacio de dimensión 2, en donde los polinomios $p_1(x) = x^3 - 2x^2 + 4x + 1$ y $p_2(x) = x^2 + x - 3$ forman una base.

5) Los ejemplos 3) y 4) estudiados anteriormente nos mueven a enunciar el teorema siguiente:

Teorema. Sea $\overline{X}$ el subespacio engendrado por el sistema $X = \{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$.

Entonces, existe al menos una base de $\overline{X}$ detraída de X.

Dem. Si el sistema X es linealmente independiente, entonces él constituye una base de $\overline{X}$, y el teorema demostrado. Ahora, si esto no ocurre, entonces existe al menos un vector $\vec{a}_i$ que es combinación lineal de los vectores restantes. Suprimiendo este vector, se forma un nuevo sistema $X' = \{a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n\}$ que engendra también a $\overline{X}$.

Podremos aplicar el mismo razonamiento anterior al sistema X'; si él es linealmente independiente, constituye una base; si esto no sucede, se suprimirá otro vector a_j, obteniéndose un tercer sistema $X'' = \{a_1, \ldots a_{i-1}, a_{i+1}, \ldots, a_{j-1}, a_{j+1}, \ldots, a_n\}$.

Es claro que, repitiendo un número finito de veces esta operación se llegará a obtener una base de $\overline{X}$ conteniendo un número finito m de vectores. Este número m satisface, pues, a la relación $m \leq n$.

En resumen: Si el conjunto $X = \{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$ es un sistema de generadores de X y si $X_1 = \{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_m\}$ es un sistema maximal detraído de X, entonces se verifica

$$\overline{X} = \overline{X}_1$$

En efecto, desde que cada b_j es algún a_i, entonces todo elemento de $\overline{X}_1$ es también un elemento de $\overline{X}$; o sea

$$\overline{X}_1 \subseteq \overline{X} \qquad (1)$$

Por otra parte, el sistema $\{a, b_1, b_2, \ldots, b_m\}$ es L. D., y por lo tanto,

$$\alpha \vec{a}_i + \beta_1 \vec{b}_1 + \ldots + \beta_m \vec{b}_m = \vec{0}$$

para lo menos algún coeficiente no nulo.

Es claro que debe ser necesariamente $\alpha \neq 0$, puesto que el sistema X_1 es L.I. Por lo tanto,

$$\vec{a}_1 = -\alpha^{-1}(\beta_1 \vec{b}_1 + \ldots + \beta_m b_m)$$

y esto prueba que cada $\vec{a}_1$ está también en $\overline{X}_1$; esto es

$$\overline{X} \subseteq \overline{X}_2 \qquad (2)$$

De (1) y (2) resulta $\overline{X} = \overline{X}_1$, y $X_1 = \{\vec{b}_1, \ldots, \vec{b}_m\}$ es una base para el subespacio $\overline{X}$

Esta base X_1 para $\overline{X}$ no es única; existen una infinidad de bases distintas de $\overline{X}$; pero estas bases tienen el mismo número de elementos como lo afirma el teorema que sigue.

Teorema. Sea V(K) un espacio vectorial de dimensión finita. Todas las bases de V(K) son sistemas finitos que comprenden un mismo número de vectores.

Demostración: Consideremos dos bases cualesquiera de B_1 y B_2 del espacio V(K), detraídas o no del sistema de generadores $X = \{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_p\}$. El teorema estudiado en la sección 9.12., llamado de la Independencia Lineal, conjuntamente con el inmediatamente anterior de esta sección, nos enseñan que cada una de estas dos bases contienen un número finito de vectores; digamos, n_1 para B_1 y n_2 para B_2, verificándose según los teoremas mencionados las relaciones:

$$n_1 \leq p, n_2 \leq p$$

Por otra parte, cada base constituye también un sistema de generadores del espacio V(K).

Si consideramos como sistema de generadores a la base B_1, entonces en V(K) se tiene n_2 vectores independientes que generan a V(K), y por el teorema fundamental de la independencia, resulta que $n_2 \leq n_1$.

Asimismo, si se considera ahora a la base B_2 como un sistema de generadores de V(K), tendremos también, por el teorema fundamental de la independencia, n_1 vectores independientes que generan V(K), y se tiene $n_1 \leq n_2$.

Como $n_2 \leq n_1$ y $n_1 \leq n_2$ implican $n_1 = n_2$, el teorema queda demostrado.

Así hemos probado que si el espacio vectorial V(K) posee varias bases, estas bases tienen el mismo número de elementos y este número, es como ya lo hemos dicho en otra ocasión, la dimensión de V(K).

Por consiguiente, *la dimensión de V(K)* no depende de la base particular elegida.

Si el espacio V(K) es de dimensión n, escribimos

$$\dim V(K) = n$$

y todo sistema linealmente de n vectores constituye una base para V(K).

Representaremos también por $V^n(K)$ este espacio *n-dimensional.*

10.14. Teorema de la base incompleta

En la sección 9.6. definimos la suma directa de dos subespacios $V_1(K)$ y $V_2(K)$ de un espacio vectorial V(K), que denotamos por $V_1(K) \oplus V_2$, si todo

vector $\vec{u} \in V(K)$ se puede escribir de una y de una sola manera en la forma:

$$\vec{u} = \vec{u}_1 + \vec{u}_2 \text{ con } \vec{u}_1 \in V_1(K), \vec{u}_2 \in V_2(K)$$

Es claro que la suma directa se confunde con el espacio entero V(K); esto es

$$V(K) = V_1(K) \oplus V_2(K)$$

En este caso se puede denominar a los subespacios $V_1(K)$ y $V_2(K)$ con el nombre de *subespacios suplementarios.*

En base a la definición de suma directa de dos subespacios, podemos enunciar el teorema que sigue.

Teorema de la base incompleta

Sea V(K) un espacio vectorial de dimensión n.

Si $m < n$ elementos son linealmente independientes en V(K), entonces siempre se les puede añadir otros n-m elementos de V(K), de manera que el conjunto de estos n elementos constituya una base de V(K).

Demostración. Los m vectores independientes engendran un subespacio $V_1(K)$ de dimensión m, y en el cual ellos mismos constituyen una base.

Ahora bien, es suficiente elegir un subespacio suplementario $V_2(K)$ de $V_1(K)$ y una base de $V_2(K)$ de n-m elementos. Es claro que, en virtud de la definición de suma directa, el conjunto de los $m + (n - m) = n$ vectores obtenidos forman una base de V(K).

En consecuencia, hemos demostrado que dado un espacio vectorial V(K) de dimensión n, y si el sistema $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_m\}$ o es una base de un subespacio de dimensión m, entonces existen n-m vectores $\vec{e}_{m+1}$, $\vec{e}_{m+2}, \ldots, \vec{e}_n$ en V(K) tales que los vectores $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_m$, $e_{m+1}, \ldots, e_n$ forman una base para V(K)

Ejemplos:

1) En el espacio $\mathbb{R}^4$, hallar una base tal que una parte de ella está constituida por una base del subespacio $V_1(\mathbb{R})$ teniendo por generadores los vectores $\vec{a}_1 = (3, -4, 1, 2)$ y $\vec{a}_2 = (1, -2, 5, -7)$

Solución. Estos vectores son independientes, puesto que la matriz reducida a la forma escalonada verifica el criterio de independencia lineal

$$\begin{pmatrix} 1 & -2 & 5 & -7 \\ 3 & -4 & 1 & 2 \end{pmatrix}$$

de donde

$$\begin{pmatrix} 1 & -2 & 5 & 7 \\ 0 & 2 & -14 & 23 \end{pmatrix}$$

Entonces, una base del subespacio $V_1(\mathbb{R})$ podría estar formada de los vectores

$\vec{e_1} = (1, 0, 0, 0)$

$\vec{b} = (0, 2, -14, 23)$

Luego, para completar una base para el espacio $\mathbb{R}^4$ en las condiciones del problema, bastará tomar los vectores $\vec{e_3} = (0, 0, 1, 0)$ y $\vec{e_4} = (0, 0, 0, 1)$, los cuales conjuntamente con los vectores $\vec{e_1}$ y $\vec{b}$ verifican el criterio de independencia lineal, puesto que cualquier combinación lineal nula de ellos

$$\alpha \vec{e_1} + \beta \vec{b} + \gamma \vec{e_3} + \delta \vec{e_4} = \vec{o}$$

implica $\alpha = \beta = \gamma = \delta = 0$.

2) En el espacio $\mathbb{R}^5$ se consideran los cuatro vectores $\vec{a_1} = (1, 2, -4, 3, 1)$, $\vec{a_2} = (2, 5, -3, 4, 8)$, $\vec{a_3} = (6, 17, -7, 10, 22)$ y $\vec{a_4} = (1, 3, -3, 2, 0)$

Hallar los vectores necesarios para completar una base para $\mathbb{R}^5$.

Solución. En primer lugar deberemos encontrar una base para el subespacio generado por los cuatro vectores dados.

Tenemos:

$$\begin{pmatrix} 1 & 2 & -4 & 3 & 1 \\ 2 & 5 & -3 & 4 & 8 \\ 6 & 17 & -7 & 10 & 22 \\ 1 & 3 & -3 & 2 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & -4 & 3 & 1 \\ 0 & 1 & 5 & -2 & 6 \\ 0 & 5 & 17 & -8 & 16 \\ 0 & 1 & 1 & -1 & -1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & -4 & 3 & 1 \\ 0 & 1 & 5 & -2 & 6 \\ 0 & 0 & -8 & 2 & -14 \\ 0 & 0 & -4 & 1 & -7 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & -4 & 3 & 1 \\ 0 & 1 & 5 & -2 & 6 \\ 0 & 0 & -4 & 1 & -7 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Luego, como sólo hay tres vectores linealmente independientes, entonces el subespacio engendrado por el sistema de los cuatro vectores dados es de dimensión 3. Una base para este subespacio puede ser la

formada por los vectores $(1, 2, -4, 3, 1)$, $(0, 1, 5, -2, 6)$ y $(0, 0, -4, 1, -7)$ que verifica el criterio de independencia lineal. Por consiguiente, para formar una base para $\mathbb{R}^5$ bastará elegir dos vectores que tengan lo más posible componentes nulas y que verifiquen al mismo tiempo con los tres vectores anteriormente considerados, el criterio de independencia. Estos dos vectores que faltan pueden ser entonces, por ejemplo, $\vec{e_4} = (0, 0, 0, 1, 0)$ y $\vec{e_5} = (0, 0, 0, 0, 1)$

10.15. *Relación entre las dimensiones de la suma e intersección de dos subespacios.*

Sea ahora V(K) un espacio vectorial de dimensión n. Para dos subespacios S y T de V(K); existe una relación interesante entre las dimensiones de los subespacios S, T, $S \cap T$ y $S + T$, relación que es independiente de la dimensión n del espacio entero V(K).

Teorema. Si S y T son subespacios de un espacio vectorial V(K), entonces existe la relación

$$\dim S + \dim T = \dim (S \cap T) + \dim (S + T)$$

Demostración. Sea $\{\vec{x_1}, \vec{x_2}, \ldots, \vec{x_i}\}$ una base de la intersección $S \cap T$. Entonces, por el teorema de la base incompleta, podemos extender este conjunto de vectores de manera para formar una base para cada uno de los subespacios S y T.

Sea pues, $\{\vec{x_1}, \ldots, \vec{x_i}, \vec{y_1}, \vec{y_2}, \ldots, \vec{y_{m-i}}\}$ una base para S y $\{\vec{x_1}, \ldots, \vec{x_i}, \vec{z_1}, \vec{z_2}, \ldots, \vec{z_{p-i}}\}$ una base para T.

Probaremos ahora que el sistema de vectores:

$$X = \{\vec{y_1}, \ldots, \vec{y_{m-i}}, \vec{x_1}, \ldots, \vec{x_i}, \vec{z_1}, \ldots, \vec{z_{p-i}}\}$$

forman una base para el subespacio suma $S + T$. En efecto, en primer término, todos los vectores del sistema X son, con certeza, elementos de la suma $S + T$ y, en segundo término, cualquier vector de $S + T$ es suma de un vector de S y de un vector de T y, además, puede expresarse como combinación lineal de los vectores del sistema X. Por lo tanto, el sistema X constituye un sistema de generadores para el subespacio $S + T$.

Supongamos ahora que entre los vectores del sistema X existiera una relación de dependencia, tal como

$$\sum_{\alpha=1}^{m-i} a_\alpha \vec{y}_\alpha + \sum_{\beta=1}^{i} b_\beta \vec{x}_\beta + \sum_{\gamma=1}^{p-i} c_\gamma \vec{z}_\gamma = \vec{o}$$

en donde los a_α, b_β y c_γ son elementos del cuerpo K. De donde podemos escribir la relación

$$\sum_{\alpha=1}^{m-i} a_\alpha \vec{y}_\alpha = -\left(\sum_{\beta=1}^{i} b_\beta \vec{x}_\beta + \sum_{\gamma=1}^{p-i} c_\gamma \vec{z}_\gamma \right)$$

que muestra que el vector $\sum_{\alpha=1}^{m-i} a_\alpha y_\alpha \in T$. Pero este vector está

también en el subespacio S, y evidentemente en la intersección $S \cap T$. Por lo tanto, él es una combinación lineal de los $\vec{x}_\beta$

$$\sum_{\alpha=1}^{m-i} a\alpha\, y\alpha = d_1\vec{x}_1 + d_2\vec{x}_2 + \ldots + d_i\vec{x}_i$$

Por otra parte, los elementos $\vec{x}_1, \ldots, \vec{x}_i, \vec{y}_1, \ldots, \vec{y}_{m-1}$, que son una base de S, son linealmente independientes, y por lo tanto todas las a_α ($\alpha = 1, 2, \ldots, m-i$) deben ser iguales a cero.

Con un razonamiento similar se muestra también que todas las c_γ ($\gamma = 1, 2, \ldots, p-i$) deben ser iguales a cero, y por ende también todas las b_β ($\beta = 1, 2, \ldots, i$) son nulas.

Así hemos probado que los vectores $X = \{\vec{y}_1, \ldots, y_{m-i}, \vec{x}_1, \ldots, \vec{x}_i, \vec{z}_1, \ldots, \vec{z}_{p-i}\}$ que engendran al subespacio $S + T$ son linealmente independientes y constituyen por lo tanto una base para la suma $S + T$.

Este resultado demuestra que

$$\dim (S + T) = (m - i) + i + (p - i) = m + p - i$$

en donde $m = \dim S$, $p = \dim T$, $i = \dim (S \cap T)$

Luego, $\dim (S + T) = \dim S + \dim T - \dim (S \cap T)$

de donde,

$$\dim S + \dim T = \dim (S + T) + \dim (S \cap T)$$

y el teorema está demostrado.

En particular, cuando la suma $S + T$ es directa, esto es $S \oplus T$, entonces la intersección se reduce al vector cero, y por esto, $\dim S \cap T = 0$, y la relación anterior se reduce a la siguiente:

$$\dim S + \dim T = \dim (S \oplus T)$$

Por consiguiente, la reunión de una base de S y de una base de T forma una base de $S \oplus T$.

Ejercicios. 1. Determinar una base de la suma y de la intersección de los subespacios engendrados, respectivamente, por los sistemas siguientes:

$X = \{(1, 2, -1, -2), (3, 1, 1, 1), (-1, 0, 1, -1)\}$; $S = \overline{X}$

$Y = \{(2, 5, -6, -5), (-1, 2, -7, -3)\}$; $T = \overline{Y}$

Solución: Determinar una base del subespacio suma $S + T$ es relativamente fácil, ya que basta formar la matriz cuyas filas son los cinco vectores dados, y reducirla en seguida a la forma escalonada; luego se tiene

$$\begin{pmatrix} 1 & 2 & -1 & -2 \\ 3 & 1 & 1 & 1 \\ -1 & 0 & 1 & -1 \\ 2 & 5 & -6 & -5 \\ -1 & 2 & -7 & -3 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & -1 & -2 \\ 0 & -5 & 4 & 7 \\ 0 & 2 & 0 & -3 \\ 0 & 1 & -4 & -1 \\ 0 & 4 & -8 & -5 \end{pmatrix}, \text{ o bien, } \begin{pmatrix} 1 & 2 & -1 & -2 \\ 0 & 1 & -4 & -1 \\ 0 & -5 & 4 & 7 \\ 0 & 2 & 0 & -3 \\ 0 & 4 & 8 & -5 \end{pmatrix}$$

reduciendo esta última por las operaciones elementales, se obtiene,

$$\begin{pmatrix} 1 & 2 & -1 & -2 \\ 0 & 1 & -4 & -1 \\ 0 & 0 & -16 & 2 \\ 0 & 0 & 8 & -1 \\ 0 & 0 & 8 & -1 \end{pmatrix}$$

y de donde, por operaciones elementales

$$\begin{pmatrix} 1 & 2 & -1 & -2 \\ 0 & 1 & -4 & -1 \\ 0 & 0 & -16 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Como esta matriz escalonada tiene tres filas distintas de cero, el subespacio suma $S + T$ tiene una base formada por los vectores $(1, 2, -1, -2)$, $(0, 1, -4, -1)$ y $(0, 0, -8, 1)$, los cuales, como sabemos verifican el criterio de independencia lineal. Por lo tanto,

$$\dim (S + T) = 3$$

Ahora, para determinar una base para el subespacio $S \cap T$, deberemos tener presente que un vector cualquiera $\vec{x} \in S \cap T$ deberá tener la forma

$$\vec{x} = a(1,2, -1,2) + b(3,1,1,1) + c(-1,0,1, -1)$$

por pertenecer al subespacio S, y al mismo tiempo por pertenecer al subespacio T, deberá tener la forma

$$\vec{x} = d(2,5, -6, -5) + e(-1,2, -7, -3)$$

Por consiguiente, tendremos la igualdad

$a(1, 2, -1, 2) + b(3, 1, 1, 1) + c(-1, 0, 1, -1) = d(2, 5, -6, -5) + e(-1, 2, -7, -3)$ y de esta relación vectorial se desprende el sistema de ecuaciones simultáneas:

$$\begin{aligned} a + 3b - c - 2d + e &= 0 \\ 2a + b \quad\quad - 5d - 2e &= 0 \\ -a + b + c + 6d + 7e &= 0 \\ -2a + b - c + 5d + 3e &= 0 \end{aligned}$$

y aplicando el método de reducción, o de la forma escalonada, obtendremos

$$\left|\begin{array}{l} a + 3b - c - 2d + e = 0 \\ -5b + 2c - d - 4e = 0 \\ 4b \quad + 4d + 8e = 0 \\ 7b - 3c + d + 5e = 0 \end{array}\right.$$

$$\left|\begin{array}{l} a + 3b - c - 2d + e = 0 \\ b - \frac{2}{5}c + \frac{1}{5}d + \frac{4}{5}e = 0 \\ \quad \frac{8}{5}c + \frac{16}{5}d + \frac{24}{5}e = 0 \\ -\frac{1}{5}c - \frac{2}{5}d - \frac{3}{5}e = 0 \end{array}\right.$$

$$\left|\begin{array}{l} a + 3b - c - 2d + e = 0 \\ b - \frac{2}{5}c + \frac{1}{5}d + \frac{4}{5}e = 0 \\ \quad \frac{1}{5}c + \frac{2}{5}d + \frac{3}{5}e = 0 \\ \quad 0 \quad 0 \quad 0 \quad = \quad 0 \end{array}\right.$$

y que reducida a coeficientes enteros, nos da

$$\left|\begin{array}{l} a + 3b - c - 2d + e = 0 \\ 5b - 2c + d + 4e = 0 \\ \quad c + 2d + 3e = 0 \end{array}\right.$$

sistema que siendo compatible, nos permite hallar los valores de tres de las incógnitas en función arbitraria de las otras dos, y de esta manera se obtiene las infinitas soluciones, que sigue:

$$\begin{aligned} a &= 3d + 2e \\ b &= -d - 2e \\ c &= -2d - 3e \end{aligned}$$

Por lo tanto,

$\vec{x} = (3d + 2e)(1, 2, -1, -2) - (d + 2e)(3, 1, 1, 1) - (2d + 3e)(-1, 0, 1, -1) = d(2, 5, -6, -5) + e(-1, 2, -7, -3)$

o sea,

$3d(1, 2, -1, -2) - d(3, 1, 1, 1) - 2d(-1, 0, 1, -1) + 2e(1, 2, -1, -2) - 2e(3, 1, 1, 1) - 3e(-1, 0, 1, -1) = d(2, 5, -6, -5) + e(-1, 2, -7, -3)$

$d(2, 5, -6, -5) + e(-1, 2, -7, -3) = d(2, 5, -6, -5) + e(-1, 2, -7, -3)$

resultado que nos enseña que cualquier vector $\vec{x} \in S \cap T$ es combinación lineal de los vectores $(2, 5, -6, -5)$ y $(-1, 2, -7, -3)$, los cuales constituyen, por lo tanto, una base para este subespacio.

Luego,

$\dim (S \cap T) = 2$

Por otra parte, puede fácilmente verse que los sistemas dados X e Y constituyen una base para los subespacios S y T. Luego,

$\dim S = 3, \dim T = 2$

Este resultado y los anteriores, verifican la relación

$\dim S + \dim T = \dim (S + T) + \dim (S \cap T)$

Observación. Si en el ejercicio anterior se nos hubiera pedido sólo encontrar las dimensiones, sin especificar alguna base de los subespacios S + T y $S \cap T$, entonces el problema se resuelve como sigue: se determinan las dimensiones de los subespacios S, T y S + T por medio de la matriz de los vectores que los engendran, reduciendo en seguida cada una de estas tres matrices a la forma escalonada. Si así se procede, se encuentra:

$\dim S = 3, \dim T = 2, \dim (S + T) = 3$

Luego, la relación

$\dim S + \dim T = \dim (S + T) + \dim (S \cap T)$

nos da,

$3 + 2 = 3 + \dim (S \cap T)$

y de donde,

$$\dim (S \cap T) = 2$$

2) Supongamos que S y T son subespacios de un espacio vectorial V(K), tales que

$\dim S = 4, \dim T = 5, \dim (V(K) = 7.$

Hallar la posible dimensión del subespacio $S \cap T$.

Solución. Como S y T son diferentes, y como S + T contiene propiamente a S y a T, resulta entonces que

$\dim (S + T) \geq 5$

Pero, dim (S + T) no puede ser mayor que 7, puesto que dim V(K) = 7.

Luego, tenemos las siguientes alternativas:

$\dim (S + T) = 5, \dim (S + T) = 6, \dim (S + T) = 7$

y usando la relación

$$\dim S + \dim T = \dim (S + T) + \dim (S \cap T)$$

se encuentra:

$$\dim (S \cap T) = \dim S + \dim T - \dim (S + T)$$
$$\dim (S \cap T) = 4 + 5 - 5 = 4$$
$$\dim (S \cap T) = 4 + 5 - 6 = 3$$
$$\dim (S \cap T) = 4 + 5 - 7 = 2$$

Esto es, las posibles dimensiones del subespacio $S \cap T$ pueden ser 2, o 3, o 4.

10.16. Teorema de Grasmann-Steinitz

En muchos ejercicios que hemos resuelto en las secciones precedentes, se ha podido observar que en más de una ocasión se han reemplazado algunos vectores de una base de un subespacio por otros independientes de manera que el nuevo sistema constituye también una base para el citado subespacio. Nosotros formularemos de un modo general este hecho en el teorema que sigue.

Teorema de Grasmann-Steinitz (o teorema del canje). Sea $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$ una base de un espacio vectorial V(K) y sean $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_m$, vectores linealmente independientes en V(K). Entonces existen en la base dada, m vectores que pueden canjearse con los vectores $\vec{a}_i$, de modo que el sistema resultante después de tal canje siga constituyendo una base del espacio V(K).

Dem. Vamos a formar nuevos sistemas de generadores de V(K) en los cuales los vectores $\vec{b}_j$ se van a canjear uno a uno por algunos de los vectores $\vec{a}_i$, elegidos convenientemente, de modo de conservar siempre una base del espacio V(K). Esto supuesto, el sistema $\{\vec{b}_1, \vec{a}_1, \vec{a}_2, \ldots \vec{a}_n\}$ engendra también a V(K), pero es evidentemente dependiente, por el hecho de que $\vec{b}_1$, como cualquier otro vector del espacio V(K) es combinación lineal de los $\vec{a}_i$, en virtud de la definición de base. En la relación

$$\beta_1 \vec{b}_1 + \alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \ldots + \alpha_n \vec{a}_n = \vec{o}$$

no todos los escalares α_i son nulos, porque si así no fuese, se tendría $\beta_1 \neq 0$, y por consiguiente, $\vec{b}_1 = \vec{o}$, lo cual es imposible porque $\vec{b}_1$ pertenece a un sistema linealmente independiente del espacio V(K). Por lo tanto, hay pues un vector $\vec{a}_i$ que depende de los otros vectores, y como el orden no interesa, podemos suponer que este vector es el $\vec{a}_1$. Se abandona este generador y se obtiene el sistema $\{\vec{b}_1, \vec{a}_2, \vec{a}_3, \ldots, \vec{a}_n\}$ que engendra también al espacio V(K).

En consecuencia, como resultado de este razonamiento tenemos una nueva base de V(K) que es: $\vec{b}_1, \vec{a}_2, \ldots, \vec{a}_n$, obtenida de la base dada canjeando el vector $\vec{a}_1$ por $\vec{b}_1$.

Consideramos en seguida el sistema $\{\vec{b}_1, \vec{b}_2, \vec{a}_2, \ldots \vec{a}_n\}$ que es, como sabemos, dependiente. En la relación,

$$\beta_1 \vec{b}_1 + \beta_2 \vec{b}_2 + \alpha_2 \vec{a}_2 + \ldots + \alpha_n \vec{a}_n = \vec{o}$$

todos los escalares $\alpha_2, \alpha_3, \ldots \alpha_n$ no pueden ser nulos, porque en tal caso el subsistema $\{\vec{b}_1, \vec{b}_2\}$ detraído de un sistema linealmente independiente $\{\vec{b}_1, \ldots, \vec{b}_m\}$ no sería independiente. Luego, algunas de las $\vec{a}_i$ de la última igualdad es combinación lineal de los demás vectores. Sea ésta la $\vec{a}_2$, y suprimiéndola obtenemos un nuevo sistema de generadores $\{\vec{b}_1, \vec{b}_2, \vec{a}_3, \ldots, \vec{a}_n\}$ para el espacio V(K).

En consecuencia, como resultado de este segundo razonamiento tenemos otra nueva base de V(K) que es $\{\vec{b}_1, \vec{b}_2, \vec{a}_3, \ldots, \vec{a}_n\}$, obtenida de la base dada canjeando los vectores $\vec{a}_1$ y $\vec{a}_2$, respectivamente, por $\vec{b}_1$ y $\vec{b}_2$.

Es claro que el proceso de construcción puede repetirse. Los vectores $\vec{a}_i$ son sucesivamente canjeados por los vectores $\vec{b}_j$.

Si $m \leq n$ se obtiene finalmente el sistema de generadores $\{\vec{b}_1, \vec{b}_2, \ldots \vec{b}_m, \vec{a}_{m+1}, \ldots, \vec{a}_n\}$, donde figuran todos los vectores $\vec{b}_j$, que constituye una nueva base del espacio V(K).

Para $m = n$, no queda ningún vector $\vec{a}_i$.

Finalmente, nos resta por mostrar que la hipótesis $m > n$ es imposible.

En efecto, ya el sistema $\{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\}$ es una base para V(K), y por consiguiente, los vectores restantes $\vec{b}_{n+1}, \vec{b}_{n+2}, \ldots, \vec{b}_m$ serían combinaciones lineales de los vectores precedentes, lo cual es imposible en vista de la independencia del sistema $\{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_m\}$, y el teorema está demostrado.

Como consecuencias inmediatas de este importante teorema tenemos los corolarios que siguen.

Corolario. En un espacio vectorial, el número de vectores en un sistema independiente no puede superar al número de vectores de una base.

Corolario. Si un espacio vectorial tiene una base de n vectores, entonces $n + 1$ vectores cualesquiera son siempre dependientes.

Corolario. Dos bases cualesquiera de un mismo espacio vectorial tienen el mismo número de vectores. Este número constante, lo hemos anteriormente denominado la *dimensionalidad*, o simplemente, la *dimensión* del espacio.

Si el espacio V(K) es de dimensión n, lo llamaremos también un espacio *n-dimensional* y cuando sea necesario, se indicará su dimensión como subíndice, así, $V_n(K)$.

Corolario. En un espacio n-dimensional, n vectores linealmente independientes cualesquiera forman una base de dicho espacio.

Finalmente, nótese que mediante estos corolarios han quedado en particular demostrado nuevamente, de otra manera, muchos teoremas ya estudiados en varias de las secciones anteriores.

10.17. *Isomorfismo*

En la sección 9.4. dijimos que dos espacios vectoriales V(K) y U(K) sobre el mismo cuerpo K eran isomorfos, si existe una biyección f: V → U, tal que f verifica las condiciones siguientes:

$$\begin{cases} f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y}) \\ f(\alpha\, x) = \alpha\, f(x) \end{cases}$$

para todo $\vec{x}, \vec{y} \in V$ y $\alpha \in K$.

En otras palabras, dos espacios vectoriales sobre un mismo cuerpo de escalares se llaman isomorfos, si entre sus elementos se puede establecer una correspondencia biunívoca y recíproca tal que la suma de los vectores del primer espacio corresponderá a la suma de los vectores correspondientes del segundo espacio, y el producto de un escalar arbitrario por un vector arbitrario del primer espacio corresponderá al producto del mismo escalar por el vector correspondiente del segundo espacio.

Consideraremos las propiedades más simples de los isomorfismos de espacios vectoriales.

a) En una correspondencia isomórfica, el vector cero de un espacio se convierte en el vector cero del otro.

En efecto, sea $f: V \to U$ el isomorfismo que convierte al vector $\vec{x} \in V$ en el vector $f(\vec{x}) \in U$. Por lo tanto, de acuerdo con la definición del isomorfismo, el producto $0.\vec{x} = \vec{0}$ debe convertirse en el producto $\vec{0} \cdot f(\vec{x}) = \vec{0}'$ esto es, el vector cero del primer espacio debe convertirse en el vector cero del segundo.

b) Si en la relación $f(\alpha\vec{x}) = \alpha\, f(\vec{x})$ que verifica el isomorfismo f, hacemos $\alpha = -1$ (1 unidad del cuerpo K), tenemos

$$f(-\vec{x}) = -f(\vec{x})$$

Es decir, elementos opuestos del primer espacio deben convertirse en elementos opuestos del segundo.

c) Es claro que la relación $f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y})$ que verifica el isomorfismo $f: V \to U$, se extiende fácilmente a más de dos sumandos. Por inducción se prueba que

$$f\left(\sum_{i=1}^{n} \vec{x}_i\right) = \sum_{i=1}^{n} f(\vec{x}_i)$$

para todo $\vec{x}_i \in V$ $(i = 1, 2, \ldots, n)$.

Al combinar este resultado con la propiedad $f(\alpha\, \vec{x}) = \alpha\, f(\vec{x})$, obtenemos

$$f\left(\sum_{i=1}^{n} \alpha_i \vec{x}_i\right) = \sum_{i=1}^{n} \alpha_i\, f(\vec{x}_i)$$

para todo $\alpha_i \in K$, $\vec{x}_i \in V$.

d) En una correspondencia isomórfica, una base del primer espacio se convierte en alguna base del segundo.

En efecto, sea $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$ una base del espacio V(K). Entonces, un vector cualquiera $x \in V$ se escribe

$$\vec{x} = \alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \ldots + \alpha_n \vec{a}_n$$

y por el isomorfismo f, resulta

$$f(\vec{x}) = \alpha_1 f(\vec{a}_1) + \alpha_2 f(\vec{a}_2) + \ldots + \alpha_n\, f(\vec{a}_n)$$

Luego, el vector $f(\vec{x}) \in U$ es combinación lineal de los elementos $f(\vec{a}_1), f(\vec{a}_2), \ldots, f(\vec{a}_n)$ del espacio U(K).

Probaremos que estos últimos elementos (vectores) son linealmente independientes.

En efecto, supongamos que se tenga

$$\alpha_1\, f(\vec{a}_1) + \alpha_2\, f(\vec{a}_2) + \ldots + \alpha_n\, f(\vec{a}_n) = \vec{0}' \in U(K)$$

o sea,

$$f(\alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \ldots + \alpha_n \vec{a}_n) = \vec{0}' \in U(K)$$

lo que implica por la propiedad a):

$$\alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \ldots + \alpha_n \vec{a}_n = \vec{0} \in V(K)$$

Siendo los vectores $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$, linealmente independientes, por ser una base de V(K), resulta que

$$\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$$

Así hemos probado que los vectores $f(\vec{a}_1), f(\vec{a}_2), \ldots, f(\vec{a}_n)$ son también linealmente independientes, y por lo tanto, constituyen una base de U(K). Esto nos mueve a enunciar el teorema que sigue:

Teorema. Dos espacios vectoriales V(K) y U(K) de dimensión finita sobre el mismo cuerpo K son isomorfos si, y sólo si, $\dim V(K) = \dim U(K)$

Demostración. En la propiedad d) anterior, ya hemos demostrado que si los espacios son isomorfos entonces ellos tienen la misma dimensión.

Recíprocamente, probaremos que si los espacios tienen la misma dimensión, entonces ellos son isomorfos.

En efecto, supongamos que $\dim V(K) = \dim U(K) = n$, que $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$ sea una base de V(K) y $\{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\}$ una base de U(K).

Definamos una aplicación de V dentro de U como sigue:

$$f: V \to U$$

para cada $\vec{x} \in V$, el elemento:

$x = \alpha_1\vec{a_1}, \alpha_2\vec{a_2} + \ldots + \alpha_n\vec{a_n}, \alpha_i \in K$

define el elemento $f(\vec{x}) \in U$ de modo que:

$f(\vec{x}) = \alpha_1\vec{b_1} + \alpha_2\vec{b_2} + \ldots + \alpha_n\vec{b_n}$

Mostremos que esta aplicación f así definida es un isomorfismo entre los espacios vectoriales V(K) y U(K). En primer término, es claro que f es una biyección de V sobre U. En segundo lugar, demostraremos que ella preserva sumas y multiplicación escalar.

En efecto, sea $\vec{y} \in V$ otro elemento cualquiera, es decir,

$\vec{y} = \beta_1\vec{a_1} + \beta_2\vec{a_2} + \ldots + \beta_n\vec{a_n}$

entonces,

$\vec{x} + \vec{y} = (\alpha_1 + \beta_1)\vec{a_1} + (\alpha_2 + \beta_2)a_2 + \ldots + (\alpha_n + \beta_n)a_n$

y por la definición de la aplicación f, resulta

$f(\vec{x} + \vec{y}) = (\alpha_1 + \beta_1)\vec{b_1} + (\alpha_2 + \beta_2)\vec{b_2} + \ldots + (\alpha_n + \beta_n)b_n$

$f(\vec{x} + \vec{y}) = (\alpha_1\vec{b_1} + \alpha_2\vec{b_2} + \ldots + \alpha_n\vec{b_n}) + (\beta_1\vec{b_1} + \ldots + \beta_n b_n).$

$f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y})$

Esto prueba que f conserva la adición.

Semejantemente, se tiene

$f(\alpha\vec{x}) = \alpha f(\vec{x}) = \alpha(\alpha_1\vec{b_1} + \ldots + \alpha_n\vec{b_n})$

resultado que muestra que f conserva la multiplicación escalar. Por lo tanto, la aplicación f es un isomorfismo y el teorema queda demostrado.

Nota. Espacios vectoriales isomorfos son idénticos desde el punto de vista algebraico. Esto es, aquellas propiedades implicadas solamente por la adicción y multiplicación escalar, son exactamente las mismas en dos espacios vectoriales isomorfos.

Por consiguiente, dos espacios vectoriales que tienen la misma dimensión son idénticos algebraicamente, como se acaba de ver en el teorema recién probado.

En otras palabras, sin dim V(K) = dim U(K), entonces la única distinción que puede haber entre V(K) y U(K) está en la naturaleza de sus elementos.

En particular, si $V_n(K)$ es un espacio vectorial de dimensión n sobre un cuerpo K, entonces $V_n(K)$ es isomorfo a K^n, en vista de que K^n es también un espacio vectorial sobre K.

En efecto, si $V_n(K)$ es un espacio vectorial n-dimensional y $\{\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}\}$ es una base de este espacio, entonces todo elemento $\vec{x} \in V(K)$ admite, como sabemos, la única representación de la forma

(1) $\vec{x} = \alpha_1\vec{a_1} + \alpha_2\vec{a_2} + \ldots + \alpha_n\vec{a_n}$

en donde los escalares o coeficientes $\alpha_1, \alpha_2, \ldots, \alpha_n$ están determinados de manera única.

Definición. Los coeficientes $\alpha_1, \alpha_2, \ldots, \alpha_n$, unívocamente determinados de (1), los llamaremos *las componentes* del vector $\vec{x}$ en la base $\{\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}\}$.

La componente α_i se dirá *la componente según* el vector a_i de la base.

En base a esta definición, consideremos la aplicación

$$g : V_n(K) \to K^n$$

en que a cada vector $\vec{x} \in V_n(K)$ se le hace corresponder la sucesión $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in K^n$ de las componentes de $\vec{x}$ en la base $\{\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}\}$.

Probaremos que tal aplicación es un isomorfismo de $V_n(K)$ sobre K^n. En efecto, demostremos primeramente que g es biyectiva. Para ver que g es sobreyectiva, observemos que dado un $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in K^n$, el vector definido por

$\vec{x} = \alpha_1\vec{a_1} + \alpha_2\vec{a_2} + \ldots + \alpha_n\vec{a_n}$

es tal que $g(\vec{x}) = (\alpha_1, \alpha_2, \ldots, \alpha_n)$

Asimismo, para ver que g es inyectiva, observemos que si los vectores de $V_n(K)$:

$\vec{x} = \alpha_1\vec{a_1} + \alpha_2\vec{a_2} + \ldots + \alpha_n\vec{a_n}$

$\vec{y} = \beta_1\vec{a_1} + \beta_2\vec{a_2} + \ldots + \beta_n\vec{a_n}$

son distintos, $\vec{x} \neq \vec{y}$, entonces sus imágenes $g(\vec{x})$ y $g(\vec{y})$ son también distintas, $g(\vec{x}) \neq g(\vec{y})$, porque si no lo fuesen, es decir, $g(\vec{x}) = g(\vec{y})$, se tendría

$(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\beta_1, \beta_2, \ldots, \beta_n).$

y lo que implacaría.

$\alpha_1 = \beta_1, \alpha_2 = \beta_2, \ldots, \alpha_n = \beta_n$

y los vectores $\vec{x}$ e $\vec{y}$ serían iguales, lo cual es contrario a lo supuesto.

Luego, la aplicación g es biyectiva.

Probemos, por último, que la aplicación g preserva las sumas y la multiplicación escalar. Se tiene:

$\vec{x} + \vec{y} = (\alpha_1 + \beta_1)\vec{a_1} + (\alpha_2 + \beta_2)\vec{a_2} + \ldots + (\alpha_n + \beta_n)\vec{a_n}$

$g(\vec{x} + \vec{y}) = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \ldots, \alpha_n + \beta_n)$

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

$= g(\vec{x}) + g(\vec{y})$

Por otro lado,

$\alpha\vec{x} = \alpha(\alpha_1\vec{a_1} + \alpha_2\vec{a_2} + \ldots + \alpha_n\vec{a_n})$

$\alpha\vec{x} = (\alpha\,\alpha_1)\vec{a_1} + (\alpha\,\alpha_2)\vec{a_2} + \ldots + (\alpha\,\alpha_n)\vec{a_n}$

$g(\alpha\vec{x}) = (\alpha\alpha_1, \alpha\alpha_2, \ldots, \alpha\alpha_n)$

$= \alpha(\alpha_1, \alpha_2, \ldots, \alpha_n)$

$= \alpha g(\vec{x})$

Así hemos demostrado que nuestra aplicación g es un isomorfismo de $V_n(K)$ sobre K^n.

En consecuencia, podemos enunciar el siguiente teorema.

Teorema. Todo espacio n-dimensional $V_n(K)$ es isomorfo al espacio vectorial K^n. Esto es, desde el punto de vista algebraico los espacios $V_n(K)$ y K^n son »iguales« y tienen las mismas propiedades algebraicas; difieren solamente en las notaciones y naturaleza de sus elementos.

Es claro que:

Si $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$ es una base de $V_n(K)$, y g la aplicación isomorfismo, entonces

$$\vec{a}_1 = 1.\vec{a}_1 + 0.\vec{a}_2 + \ldots + 0.\vec{a}_n$$

$$\vec{a}_2 = 0 \cdot \vec{a}_1 + 1 \cdot \vec{a}_2 + \ldots + 0 \cdot \vec{a}_n$$

. . . . . .

. . . . . .

$$\vec{a}_n = 0.\vec{a}_1 + 0.\vec{a}_2 + \ldots + 1.\vec{a}_n$$

entonces, el sistema

$$g(\vec{a}_1) = (1, 0, \ldots, 0)$$
$$g(\vec{a}_2) = (0, 1, 0, \ldots, 0)$$

$$g(\vec{a}_n) = (0, 0, \ldots, 1)$$

es una base para el espacio K^n, en donde 1 es el elemento unidad del cuerpo conmutativo K.

Corolario. Todo el espacio vectorial real $V(\mathbb{R})$ de dimensión n es isomorfo al espacio $\mathbb{R}^n$.

Ejemplo. Si con la notación V_2 designamos el conjunto de todos los vectores geométricos del plano euclideano, entonces los espacios vectoriales V_2 y $\mathbb{R}^2$ tienen la misma dimensión 2 sobre el mismo cuerpo $\mathbb{R}$ de los números reales. Por consiguiente, V_2 y $\mathbb{R}^2$ son isomorfos.

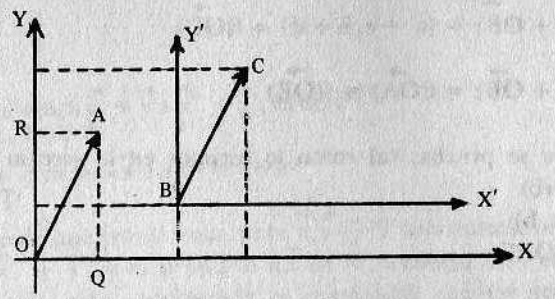
Así mismo, si V_3 denota el conjunto de todos los vectores geométricos del espacio ordinario, entonces también los espacios V_3 y $\mathbb{R}^3$ de dimensión 3 son isomorfos.

El isomorfismo entre V_2 y $\mathbb{R}^2$ puede realizarse, como lo vimos en la sección (9.1., como sigue:

Un vector geométrico en el plano será para nosotros un *segmento dirigido*, por ejemplo, $\vec{OA}$, que tiene su origen en el origen O del sistema de coordenadas cartesianas (XY) y su extremo A en cualquier punto del plano.

A un vector, tal como OA, lo llamaremos un *vector de posición*.

Cualquier otro vector, $\vec{BC}$, por ejemplo, puede compararse a un vector de posición seleccionando un nuevo sistema de ejes cartesianos (x'y'),



paralelos a los anteriores y teniendo al punto B como nuevo origen, tal como lo muestra la figura adjunta.

Es claro que el vector de posición OA puede ser identificado con el par ordenado (a, b) de números reales, donde *a* y *b* son las coordenadas cartesianas del punto A. Llamaremos a la pareja ordenada (a, b) *las componentes* del vector $\vec{OA}$.

Análogamente, si $B(a_1, b_1)$ y $C(a_2, b_2)$, el vector $\vec{AB}$ puede ser identificado con el par ordenado $(a_2 - a_1, b_2 - b_1)$ de números reales y que son sus componentes.

En consecuencia, existe, pues, una correspondencia biunívoca entre los vectores del plano y las parejas ordenadas de números reales.

En término de los elementos de los dos espacios V_2 y $\mathbb{R}^2$, se tiene:

$$\vec{OA} \rightarrow (\alpha_a, \alpha_b)$$

$\vec{OA} = \vec{BC}$ si, y sólo si, $(a, b) = (a_2 - a_1, b_2 - b_1)$

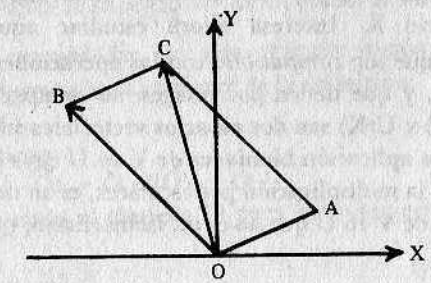
Es claro pues, que dos vectores del espacio V_2 son iguales (o equipolentes), sólamente si tienen las mismas componentes.

De esta manera tenemos ahora una forma natural de asociar los espacios vectoriales V_2 y $\mathbb{R}^2$:

$$f: V_2 \longrightarrow \mathbb{R}^2$$

donde, $f(\vec{OA}) = (a, b)$

Por lo tanto, si se tienen $f(\vec{OA}) = (a, b)$ y $f(OB) = (c, d)$. entonces se tendrá:



$$f(\vec{OA} + \vec{OB}) = (a + c, b + d) = f(\vec{OC})$$

o sea,

$$f(\vec{OA} + \vec{OB}) = f(\vec{OA}) + f(\vec{OB}) \qquad (1)$$

Asimismo se prueba, tal como lo hicimos en la sección 9.1., que $f(\alpha \vec{OA}) = \alpha_a, \alpha b)$

$$= \alpha(a, b)$$
$$= \alpha f(OA) \qquad (2)$$

Las relaciones (1) y (2) son las dos propiedades del isomorfismo entre V_2 y $\mathbb{R}^2$.

El procedimiento anterior puede transportarse intacto a V_3 y $\mathbb{R}^3$, para realizar el isomorfismo entre los espacios vectoriales V_3, de todos los vectores geométricos del espacio ordinario, y $\mathbb{R}^3$, de todas las ternas ordenadas de números reales.

Esto es:

$f : V_3 \longrightarrow \mathbb{R}$

definida por $f(\vec{OA}) = (a, b, c)$, donde a, b y c son las componentes del vector $\vec{OA}$, es decir, las coordenadas cartesianas del punto A en el sistema (XYZ) del espacio tridimensional.

B) *Transformaciones lineales*

10.18. Las aplicaciones de un espacio en otro, o también en sí mismo, las llamaremos *transformaciones del primero*, o *transformaciones del espacio* en el caso de la aplicación de un espacio en sí mismo.

Solamente estudiaremos un tipo especial de ellas, por ser las más simples y aún las más importantes que entran en el estudio de los espacios vectoriales. Las llamaremos: **TRANSFORMACIONES LINEALES** a causa de las condiciones de linealidad que contiene su definición, como luego se verá. Anticipamos que ellas, a pesar de su carácter especial o simple, tienen la máxima importancia teórica y práctica.

Sean pues, V(K) y U(K) dos espacios vectoriales sobre el mismo cuerpo conmutativo K. Interesa ahora estudiar aquellas aplicaciones de V dentro de U que son *compatibles* con las operaciones de adición y multiplicación escalar, y que tienen por imagen no siempre a todo U. Recordemos que si V(K) y U(K) son dos espacios vectoriales sobre el mismo cuerpo K, entonces una aplicación biunívoca de V en U cuya imagen es U y que preserva la suma y la multiplicación por escalares, es un isomorfismo.

Una aplicación de V en U que no es un isomorfismo, que lleva $\vec{x}, \vec{y}$ a $\vec{x}, \vec{y}$, es decir:

$$\vec{x} \rightarrow \vec{x}$$
$$\vec{y} \rightarrow \vec{y}$$

en general no llevará $\vec{x} + \vec{y}$ a $\vec{x} + \vec{y}$, es decir:

$$\vec{x} + \vec{y} \rightarrow \vec{x} + \vec{y}$$

y no llevará en general $\alpha\vec{x}$ a $\alpha\vec{x}$; es decir:

$$\alpha\vec{x} \rightarrow \alpha\vec{x}$$

Pero si sucede que transforma, para $\vec{x}, \vec{y} \in V$ arbitrarios, la suma $\vec{x} + \vec{y}$ en la suma $\vec{x} + \vec{y}$, y el producto $\alpha\vec{x}$ en el producto $\alpha\vec{x}$, entonces decimos que la aplicación considerada es *compatible* con las operaciones de adición y multiplicación escalar, o que es *lineal*.

Lo anterior lo formulamos, en síntesis, en la definición siguiente:

Definición. Sean V y U dos espacios vectoriales sobre el mismo cuerpo K.

Diremos que una aplicación $f: V \rightarrow U$ es una **TRANSFORMACION LINEAL**, o un **HOMOMORFISMO DE ESPACIO VECTORIAL**, si se verifican las dos condiciones siguientes:

1) $f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y})$

2) $f(\alpha\vec{x}) = \alpha f(\vec{x})$

cualesquiera sean $\vec{x}, \vec{y} \in V, \alpha \in K$.

En otras palabras, la aplicación $f: V \rightarrow U$ es lineal si preserva las dos operaciones básicas de un espacio vectorial: la adición de vectores y la multiplicación de vectores por un escalar.

Nótese que en esta definición de transformación lineal, no se ha exigido que los espacios V y U sean necesariamente diferentes; asimismo, la aplicación f puede ser, o no ser biunívoca. El isomorfismo de un espacio vectorial estudiado al final de la primera parte A) de este capítulo, es un caso especial de transformación lineal para la cual f es biunívoca.

Cuando V y U coinciden, una transformación lineal de V en V se convierte en un endomorfismo del espacio V, provisto de su estructura de espacio vectorial, en sí mismo.

Nota. Una transformación lineal $f: V \rightarrow U$ la llamaremos también un **OPERADOR LINEAL**.

10.19. Consecuencias de la definición de Operador lineal

Las siguientes propiedades de una transformación u operador lineal $f: V \rightarrow U$, se deducen inmediatamente de la definición:

a) $f(\vec{O}) = \vec{O}$.

En efecto, se tiene:

$$f(\vec{O}) = f(O \cdot \vec{O}) = O \cdot f(\vec{O}) = \vec{O}$$

Aquí hemos usado el símbolo $\vec{O}$ para el vector cero de ambos espacios V y U.

b) $f(-\vec{x}) = -f(\vec{x})$

En efecto, tenemos

$$f(-\vec{x}) = f[(-1)\vec{x}] = (-1)f(\vec{x}) = -f(\vec{x})$$

c) $f(\alpha\vec{x} + \beta\vec{y}) = \alpha f(\vec{x}) + \beta f(\vec{y})$

En efecto, de las condiciones 1) y 2) de la definición de transformación lineal, resulta

$$f(\alpha\vec{x} + \beta\vec{y}) = \underset{1)}{f(\alpha\vec{x})} + \underset{2)}{f(\beta\vec{y})} = \alpha f(\vec{x}) + \beta f(\vec{y})$$

Más generalmente, para todo escalar $\alpha_i \in K$ y todo vector $\vec{x}_i \in V$, $(i = 1, 2, \ldots, n)$, por aplicación reiterada de las condiciones 1) y 2), obtenemos la propiedad básica de las aplicaciones o transformaciones lineales:

$$f(\alpha_1\vec{x}_1 + \alpha_2\vec{x}_2 + \ldots + \alpha_n\vec{x}_n) = \alpha_1 f(\vec{x}_1) + \alpha_2 f(\vec{x}_2) + \ldots + \alpha_n f(\vec{x}_n)$$

Observación importante

La condición:

$$f(\alpha\vec{x} + \beta\vec{y}) = \alpha f(\vec{x}) + \beta f(\vec{y})$$

caracteriza completamente las transformaciones lineales y algunas veces se suele usar como definición de ellas.

10.20. Ejemplos de Operadores lineales

1) Sea la aplicación $f: \mathbb{R}^2 \longrightarrow \mathbb{R}$, definida por:

$$(x_1, x_2) \xrightarrow{f} (x_1)$$

Probaremos que f es una transformación lineal.

En efecto, tenemos, en primer lugar:

$$(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2) \xrightarrow{f} (x_1 + y_1) = (x_1) + (y_1)$$

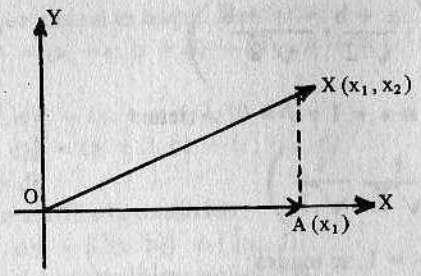
y en segundo lugar:

$$\alpha(x_1, x_2) = (\alpha x_1, \alpha_2) \xrightarrow{f} (\alpha x_1) = \alpha(x_1)$$

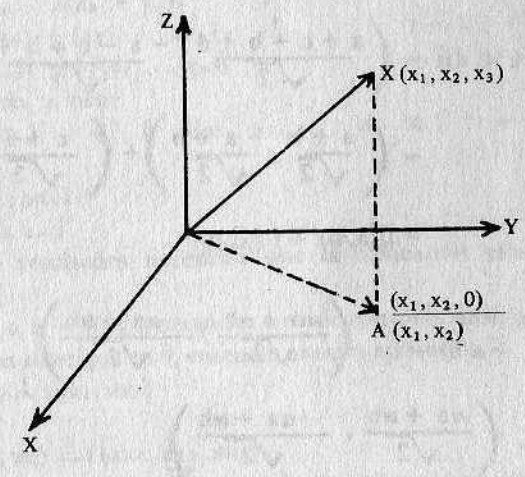
Por consiguiente, f es una transformación lineal de $\mathbb{R}^2$ en $\mathbb{R}$.

Esta transformación puede ser interpretada geométricamente, considerando un sistema de coordenadas cartesianas en el plano y asociando el vector (x_1, x_2) al trazo orientado $\overrightarrow{OX}$, cuyos extremos son el origen O (o, o) y el punto $X(x_1, x_2)$. En el eje de abscisas podemos asociar el vector (x_1) al segmento orientado $\overrightarrow{OA}$, cuyos extremos son el origen O(o) y el punto A (x_1).

La transformación lineal puede entonces ser representada por la proyección ortogonal de $\overrightarrow{OX}$ sobre el eje de las abscisas.



Esta misma ilustración geométrica puede utilizarse para la transformación $(x_1, x_2, x_3) \xrightarrow{f} (x_1, x_2)$ del espacio $\mathbb{R}^3$ en el espacio $\mathbb{R}^2$, quedando representada por la proyección ortogonal del vector geométrico $\overrightarrow{OX}$ del espacio tridimensional sobre el plano (XY) de coordenadas.



Por otro lado, esta misma interpretación geométrica, ilustra la transformación:

$$(x_1, x_2, x_2) \xrightarrow{g} (x_1, x_2, o)$$

del espacio $\mathbb{R}^3$ en sí mismo. Sin embargo, deberemos advertir que las aplicaciones anteriores f y g no son transformaciones idénticas, puesto que los codominios son espacios diferentes:

$f: \mathbb{R}^3 \to \mathbb{R}^2$

$g: \mathbb{R}^3 \to \mathbb{R}^3$

$\mathbb{R}^2$ para f y $\mathbb{R}^3$ para g.

2) Sea la aplicación $f: \mathbb{R}^2 \to \mathbb{R}^2$ definida por:

$$(a, b) \to f\left(\frac{a+b}{\sqrt{2}}, \frac{-a+b}{\sqrt{2}}\right)$$

Así, por ejemplo, si a = 1 y b = 0, se tiene:

$$(1, 0) \to f\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

Si fuese a = 1 y b = 1, se tendrá:

$$(1, 1) \to f\left(\frac{2}{\sqrt{2}}, \ o\right) = (\sqrt{2}, o)$$

Probaremos que esta aplicación es lineal. Tenemos:

$$(a, b) + (c, d) = (a + c, b + d) \to f\left(\frac{a+c+b+d}{\sqrt{2}}, \frac{-a-c+b+d}{\sqrt{2}}\right)$$

o sea, $f[(a, b) + (c, d)] = \left(\frac{a+c+b+d}{\sqrt{2}}, \frac{-a-c+b+d}{\sqrt{2}}\right)$

$$= \left(\frac{a+b}{\sqrt{2}}, \frac{-a+b}{\sqrt{2}}\right) + \left(\frac{c+d}{\sqrt{2}}, \frac{-c+d}{\sqrt{2}}\right)$$

$$= f[(a, b) + f[(c, d)]$$

Por otro lado, se tiene:
$\alpha(a, b) = (\alpha a, + \alpha b) \to f\left(\frac{\alpha a + \alpha b}{\sqrt{2}}, \frac{-\alpha a + \alpha b}{\sqrt{2}}\right)$

o bien,
$f[\alpha(a, b)] = \left(\frac{\alpha a + \alpha b}{\sqrt{2}}, \frac{-\alpha a + \alpha b}{\sqrt{2}}\right)$

$$= \alpha\left(\frac{a+b}{\sqrt{2}}, \frac{-a+b}{\sqrt{2}}\right)$$

$$= \alpha f[(a, b)].$$

Este resueltado y el anterior demuestran que nuestra aplicación es una transformación lineal.

En cambio, si la aplicación f: $\mathbb{R}^2 \to \mathbb{R}^2$ hubiera estado definida por,

$$(a, b) \to f(a + 2, b)$$

entonces, veremos que ella no es lineal. En efecto;

$$(a, b) + (c, d) = (a + c, b + d) \to (a + c + 2, b + d)$$

o sea,

$$f[(a, b) + (c, d)] = (a + c + 2, b + d)$$

y, $f[(a, b)] + f[(c, d)] = (a + 2, b) + (c + 2, d)$
$= (a + c + 4, b + d)$

y, por lo tanto, no se verifica la condición

$$f[(a, b) + (c, d)] = f[(a, b)] + f[(c, d)]$$

y por esto no es lineal nuestra última aplicación.

3) Sea la transformación f: $\mathbb{R}^n \to \mathbb{R}^n$ definida por

$$((x_1, x_2, \ldots, x_n) \xrightarrow{f} (ax_1, ax_2, \ldots, ax_n)$$

donde $a \in \mathbb{R}$, fijo arbitrario.

Probemos que es lineal. Tenemos:

$f[(x_1, \ldots, x_n) + (y_1 m, \ldots, y_n)] = f[(x_1 + y_1), \ldots,$
$(x_n + y_n)] =$
$= (a(x_1 + y_1), \ldots, a(x_n + y_n))$
$= (ax_1, \ldots, ax_n) + (ay_1, \ldots, ay_n)$
$= f[(x_1, \ldots, x_n)] + f[(y_1, \ldots, y_n)]$

Por otro lado, se tiene

$f[\alpha(x_1, \ldots, x_n)] = f[(\alpha x_1, \ldots, \alpha x_n)] = (\alpha ax_1, \ldots, \alpha ax_n)$
$= \alpha(ax_1, \ldots, ax_n)$
$= \alpha f[(x_1, \ldots, x_n)]$

Estos dos resultados muestran que la aplicación considerada es lineal.

Si n = 2, o n = 3, y asociando como de costumbre vectores y segmentos orientados con su origen en 0, entonces nuestra aplicación f

$$(x_1, x_2) \xrightarrow{f} (ax_1, ax_2)$$

o también,

$$(x_1, x_2, x_3) \xrightarrow{f} f(ax_1, ax_2, ax_3)$$

transforma el segmento $\overrightarrow{OA}$ en el segmento $\overrightarrow{OA'}$, siendo (x_1, x_2), o también (x_1, x_2, x_3) las coordenadas de A y (ax_1, ax_2), o también según el caso (ax_1, ax_2, ax_3) las coordenadas de A'.

¡Haga las respectivas ilustraciones geométricas!

A una transformación lineal de este tipo la llamaremos: *dilatación* si $a > 1$, y *contracción* si $0 < a < 1$.

4) La aplicación idéntica I: V(K) → V(K), la cual aplica cada $\vec{x} \in V$ en sí mismo, esto es,

$$I(\vec{x}) = \vec{x}$$

es lineal; puesto que se verifica la condición general:

$$I(\alpha\vec{x} + \beta\vec{y}) = \vec{\alpha x} + \beta\vec{y} = \alpha I(\vec{x}) + \beta I(\vec{y})$$

5) Sea f: V(K) → U(K) una aplicación biyectiva.

Si f es una transformación lineal, entonces, también la aplicación inversa $f^{-1}: U(K) \to V(K)$ es lineal.

En efecto, sean $\vec{y}, \vec{y}' \in U(K)$ arbitrarios. Como f es una biyección, existen entonces vectores únicos $\vec{x}, \vec{x}' \in V(K)$ tales que $f(\vec{x}) = \vec{y}$, $f(\vec{x}') = \vec{y}'$.

Ahora bien, como f es una transformación lineal se tiene:

$$\begin{cases} f(\vec{x} + \vec{x}') = f(\vec{x}) + f(\vec{x}') \\ f(\alpha\vec{x}) = \alpha f(\vec{x}) \end{cases}$$

Por otro lado, como f es biyectiva, entonces existe la inversa f^{-}; luego

$$f^{-1}(\vec{y}) = \vec{x}, f^{-1}(\vec{y}') = \vec{x}'$$
$$f^{-1}(\vec{y} + \vec{y}') = \vec{x} + \vec{x}', f^{-1}(\alpha\vec{y}) = \alpha\vec{x}$$

Por consiguiente,

$$\begin{cases} f^{-1}(\vec{y} + \vec{y}') = \vec{x} + \vec{x}' = f^{-1}(\vec{y}) + f^{-1}(\vec{y}') \\ f^{-1}(\alpha\vec{y}) = \alpha\vec{x} = \alpha f^{-1}(\vec{y}) \end{cases}$$

y por tanto, la aplicación inversa $f^{-1}: U(K) \to V(K)$ es lineal. (f^{-1} es el isomorfismo inverso).

6) Dados dos espacios vectoriales V(K) y U(K), podemos definir siempre una transformación lineal

$$f: V \to U$$

haciendo corresponder a todo vector $\vec{x} \in V(K)$ el vector nulo, $\vec{o} \in U(K)$.

En efecto, se tiene

$$f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y}) = \vec{o} + \vec{o} = \vec{o}$$
$$f(\alpha\vec{x}) = \alpha f(\vec{x}) = \alpha \cdot \vec{o} = \vec{o}$$

A esta aplicación lineal le daremos el nombre de TRANSFORMACION NULA.

7) Sean los espacios vectoriales reales $\mathbb{R}^3$ y $\mathbb{R}^2$. Consideremos además el siguiente cuadro o matriz de números reales

$$M = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix}$$

En relación con esta matriz definamos la siguiente aplicación

$$f: \mathbb{R}^3 \to \mathbb{R}^2$$

definida así

$$\vec{x} = (x_1, x_2, x_3) \xrightarrow{f} \vec{y} = (y_1, y_2)$$

de manera que:

$$\begin{cases} y_1 = a_1 x_1 + a_2 x_2 + a_3 x_3 \\ y_2 = b_1 x_1 + b_2 x_2 + b_3 x_3 \end{cases}$$

Probaremos que la aplicación f definida por $f(\vec{x}) = \vec{y}$, tal como se ha indicado, es lineal.

En efecto, sean $\vec{x} = (x_1, x_2, x_3)$, $\vec{x}' = (x'_1, x'_2, x'_3)$ vectores arbitrarios de $\mathbb{R}^3$; demostraremos, en primer lugar, que se verifica la relación

$$f(\vec{x} + \vec{x}') = f(\vec{x}) + f(\vec{x}')$$

Pongamos

$$f(\vec{x}) = (y_1, y_2), f(\vec{x}'_1) = (y'_1, y'_2)$$

donde, según la definición que hemos dado de la aplicación f, tenemos

$$(1) \begin{cases} y_1 = a_1 x_1 + a_2 x_2 + a_3 x_3 \\ y_2 = b_1 x_1 + b_2 x_2 + b_3 x_3 \end{cases}$$

$$(2) \begin{cases} y'_1 = a_1 x'_1 + a_2 x'_2 + a_3 x'_3 \\ y'_2 = b_1 x'_1 + b_2 x'_2 + b_3 x'_3 \end{cases}$$

Luego, podremos escribir

(3) $f(\vec{x}) + f(\vec{x}') = (y_1, y_2) + (y'_1, y'_2) = (y_1 + y'_1, y_2 + y'_2)$

Pero, de (1) y (2) resulta

$$(4) \begin{cases} y_1 + y'_1 = a_1(x_1 + x'_1) + a_2(x_2 + x'_2) + a_3(x_3 + x'_3) \\ y_2 + y'_2 = b_1(x_1 + x'_1) + b_2(x_2 + x'_2) + b_3(x_3 + x'_3) \end{cases}$$

y estas igualdades expresan, según la definición de nuestra aplicación, que

$$(5)\ f(\vec{x} + \vec{x}') = (y_1 + y'_1, y_2 + y'_2)$$

dado que

$$\vec{x} + \vec{x} = (x_1, x_2, x_3) + (x'_1, x'_2, x'_3)$$
$$= (x_1 + x'_2, x_2 + x_2, x_3 + x'_3)$$

De (3) y (5) se concluye que

$$f(\vec{x} + \vec{x}') = f(\vec{x}) + f(\vec{x}')$$

Por otro lado, tenemos ahora

$$(6)\ \alpha f(x) = \alpha(y_1, y_2) = (\alpha y_1, \alpha y_2)$$

Multiplicando las igualdades (1) por α, tenemos

$$(7) \begin{cases} \alpha y_1 = a_1 \cdot \alpha x_1 + a_2 \cdot \alpha x_2 + a_3 \cdot \alpha x_3 \\ \alpha y_2 = b_1 \cdot \alpha x_1 + b_2 \cdot \alpha x_2 + b_3 \cdot \alpha x_3 \end{cases}$$

Estas últimas igualdades expresan, según la definición de la aplicación f, que

(8) $f(\alpha\vec{x}) = (\alpha y_1, \alpha y_2)$

dado que

$$\alpha\vec{x} = \alpha(x_1, x_2, x_3) = (\alpha x_1, \alpha x_2, \alpha x_3)$$

De (6) y (8) resulta que

$$\alpha f(\vec{x}) = f(\alpha\vec{x})$$

Este resultado y el anterior demuestran que nuestra aplicación f es lineal; es decir, es una transformación lineal.

Observación. Este ejemplo que se acaba de ver nos ha demostrado que los seis escalares a_1, a_2, a_3, y b_1, b_2, b_3, determinan la aplicación lineal f; en otras palabras, la matriz

$$M = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix}$$

sobre el cuerpo $\mathbb{R}$ determina la transformación lineal f. Por consiguiente, diremos que la transformación lineal está representada por la matriz M.

A menudo una transformación lineal será descrita dando la matriz de la transformación, como se acaba de ver en el ejemplo anterior.

Generalizando, pues, esta idea, sean los espacios $\mathbb{R}^m$ y $\mathbb{R}^n$, y consideremos la matriz

$$M = \begin{cases} a_{11} & a_{12} \ldots\ldots a_{1m} \\ a_{21} & a_{22} \ldots\ldots a_{2m} \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1} & a_{n2} \ldots\ldots a_{nm} \end{cases}$$

En relación con esta matriz o cuadro M, definimos la transformación lineal

$$f: \mathbb{R}^m \to \mathbb{R}^n$$

en que a cada $\vec{x} = (x_1, x_2, \ldots, x_m) \in \mathbb{R}^m$ se hace corresponder el vector $\vec{y} = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$, de modo que:

$$* \quad \begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1m}x_m \\ y_2 = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2m}x_m \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ y_n = a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nm}x_m \end{cases}$$

Nota. Algunos autores emplean como definición de la transformación f, la transpuesta M^T de la matriz M.

A modo de ejercicio, pruebe, como en el ejemplo anterior $f: \mathbb{R}^m \to \mathbb{R}^n$ es una aplicación lineal.

Nota. En general, decimos que un cuadro tal como M es una matriz de n *filas o renglones* y de m *columnas,* o también, una *matriz de orden* $n \cdot m$

En el caso que $m = 1$, es $\mathbb{R}^m = \mathbb{R}$, y la matriz M se reduce a la siguiente

$$M = \begin{pmatrix} a_{11} \\ a_{21} \\ \cdot \\ \cdot \\ \cdot \\ a_{n1} \end{pmatrix}$$

y la llamamos una *matriz columna,* y las igualdades (*) que definen la transformación toman la forma:

$$\begin{cases} y_1 = a_{11}x_1 \\ x_2 = a_{21}x_1 \\ \ldots\ldots\ldots \\ \ldots\ldots\ldots \\ y_n = a_{n1}x_1 \end{cases}$$

Es claro que en este caso particular no hay razón para usar doble subíndice para indicar los elementos de la matriz M; y se escribirá más bien como sigue:

$$M = \begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{pmatrix}$$

Asimismo, en el caso que sea $n = 1$, $\mathbb{R}^n = \mathbb{R}$, la matriz M toma la forma

$$M = (a_{11}\ a_{12} \ldots a_{1m})$$

y la llamaremos una *matriz fila,* y las igualdades (*) se reducen a la única siguiente: $y_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1m}x_m$

y la matriz puede ser también escrita, como en el caso particular anterior, en la forma:

$$M = (a_1, a_2, \ldots, a_m)$$

Finalmente, en el caso $m = n = 1$, la matriz M asume la forma

$$M = (a_{11})$$

o bien, $M = (a)$

y la llamamos *matriz de un elemento,* y las igualdades (*) se reducen a

$$Y_1 = a_{11}x_1 = ax_1$$

20.21. *Determinación de una aplicación lineal.* Puede haberse observado que la definición que hemos dado de aplicación lineal ha sido establecida a priori. Sin embargo, mediante el empleo de una base podremos mostrar su existencia y precisar su determinación.

Por este motivo daremos ahora una descripción de las transformaciones lineales que resultarán muy útil para nuestros fines.

Teorema. Existe una aplicación lineal y una sola

$$f : V(K) \to U(K)$$

tal que los transformados de los vectores de una base del espacio V(K) sean vectores del espacio U(K) elegidos arbitrariamente.

Dem. Consideremos una base del espacio V(K); sea:

$\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$ esta base elegida. Elijamos arbitrariamente el transformado mediante f de cada vector a_i de la base considerada

$\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$

Sea $\{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\}$ el conjunto de estos transformados elegidos arbitrariamente y que pertenecen al espacio U(K); esto es:

$$f(\vec{a}_i) = \vec{b}_i, \qquad (i = 1, 2, \ldots, n)$$

Probaremos que se ha definido así una transformación lineal f del espacio V(K) dentro del espacio U(K). En efecto, cada vector $\vec{x} \in V(K)$ es combinación lineal de los vectores de la base $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$, es decir

$$\vec{x} = \alpha_1\vec{a}_1 + \alpha_2\vec{a}_2 + \ldots + \alpha_n\vec{a}_n = \sum_{i=1}^{n} \alpha_i\vec{a}_i$$

La propiedad fundamental exige que el transformado de $\vec{x}$ sea

$$f(\vec{x}) = \alpha_1\vec{b}_1 + \alpha_2\vec{b}_2 + \ldots\ldots + \alpha_n\vec{b}_n = \sum_{i=1}^{n} \alpha_i\vec{b}_i = \vec{y} \in U(K)$$

Mostremos que esta aplicacion así definida es lineal.

Si $\vec{x} = \alpha'_1\vec{a}_1 + \alpha'_2\vec{a}_2 + \ldots + \alpha'_n\vec{a}_n$ es otro vector cualquiera de V(K), entonces tendremos:

$$\vec{x} + \vec{x} = (\alpha_1 + \alpha'_1)\vec{a}_1 + (\alpha_2 + \alpha'_2)\vec{a}_2 + \ldots + (\alpha_n + \alpha'_n)\vec{a}_n$$

y la propiedad fundamental exige que el transformado de: $\vec{x} + \vec{x}'$ sea,

$$f(\vec{x} + \vec{x}') = (\alpha_1 + \alpha'_1)\vec{b}_1 + (\alpha_2 + \alpha'_2)\vec{b}_2 + \ldots + (\alpha_n + \alpha'_n)\vec{b}_n$$

$$= (\alpha_1\vec{b}_1) + (\alpha_2\vec{b}_2 + \ldots + \alpha_n\vec{b}_n) + (\alpha'_1\vec{b}_1 + \alpha'_2\vec{b}_2 + \ldots \alpha'_n\vec{b}_n)$$

$$= f(\vec{x}) + f(\vec{x}')$$

Por otro lado, tenemos

$$\alpha\vec{x} = (\alpha\alpha_1)\vec{a}_1 + (\alpha\alpha_2)\vec{a}_2 + \ldots + (\alpha\alpha_n)\vec{a}_n$$

y la propiedad fundamental exige que el transformado del vector $\alpha\,\vec{x}$ sea

$$f(\alpha\,\vec{x}) = (\alpha\alpha_1)\vec{b}_1 + (\alpha\alpha_2)\vec{b}_2 + \ldots + (\alpha\alpha_n)\vec{b}_n$$

$$= \alpha(\alpha_1\vec{b}_1 + \alpha_2\vec{b}_2 + \ldots + \alpha_n\vec{b}_n)$$

$$= \alpha\, f(\vec{x})$$

Este resultado y el anterior muestran que la aplicación f definida por $f(a_i) = \vec{b}_i$, $\vec{b}_i$ vector arbitrario de U(K), es lineal.

Así hemos demostrado que dicha transformación existe, y ella está bien determinada. Sólo nos resta probar que f es única.

En efecto, si g es una transformación lineal cualquiera, pero tal que

$$g(\vec{a}_i) = \vec{b}_i$$

entonces necesariamente se tendrá

$$g(\vec{x}) = \alpha_1\vec{b}_1 + \alpha_2\vec{b}_2 + \ldots + \alpha_n\vec{b}_n = f(\vec{x})$$

para todo $\vec{x} \in V(K)$.

Luego, en virtud de la definición de igualdad de funciones o aplicaciones resulta g = f (absurdo).

Esta contradicción demuestra la unicidad de la transformación lineal f, y el teorema está demostrado.

En resumen:

Una transformación lineal f de un espacio vectorial V(K) dentro de un espacio vectorial U(K) queda completamente determinada por las imágenes de un conjunto de vectores de base V(K). Esto es, dada una base cualquiera: $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$ del espacio V(K) y un conjunto arbitrario ordenado $\{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\}$ de vectores del espacio U(K), entonces existe una única transformación lineal f tal que $f(\vec{a}_i) = \vec{b}_i$, para $i = 1, 2, \ldots, n$.

En otros términos, una transformación lineal de V(K) en U(K) queda unívocamente determinada cuando se conocen los transformados de los vectores de una base del espacio V(K).

Señalamos nuevamente, y téngase bien presente, que el conjunto de vectores imágenes $\{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\}$ ha sido elegido con toda libertad. Por lo tanto, no es necesario que forme una base de U(K); de hecho puede no tener el número de vectores necesario para ello; tampoco es necesario que engendre el espacio U(K); podrían incluso ser todos los $\vec{b}_i$ iguales al vector nulo $\vec{0}$.

Como ilustración del teorema que se acaba de estudiar, consideremos el siguiente ejemplo.

Sea el espacio $\mathbb{R}^3$ con su base canónica $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$, siendo como sabemos

$$\vec{e}_1 = (1, 0, 0), \qquad \vec{e}_2 = (0, 1, 0), \qquad \vec{e}_3 = (0, 0, 1)$$

Entonces, los tres vectores

$$\vec{b}_1 = (0, 1, 1), \qquad \vec{b}_2 = (1, 1, 1), \qquad \vec{b}_3 = (1, 0, 0)$$

definen una aplicación lineal $f : \mathbb{R}^3 \to \mathbb{R}^3$, como sigue:

$f(\vec{e_1}) = (o, 1, 1)$, $f(\vec{e_2}) = (1, 1, 1)$, $f(\vec{e_3}) = (1, o, o)$

De esta manera, tenemos, por ejemplo

$$f[(5, -2, 1)] = f[5(1, o, o) - 2(o, 1, o) + (o, o, 1)]$$
$$= 5f[(1, o, o)] - 2f[(o, 1, o)] + f[(o, o, 1)]$$
$$= 5(o, 1, 1) - 2(1, 1, 1) + (1, o, o)$$
$$= (-1, 3, 3)$$

En general,

$$f[(a, b, c)] = a(0, 1, 1) + b(1, 1, 1) + c(1, o, o)$$
$$= (b + c, a + b, a + b)$$

Por lo tanto, se puede escribir, según este resultado, los transformados que siguen:

$$f[(5, -3, 2)] = (-3 + 2, 5 - 3, 5 - 3) = (-1, 2, 2)$$
$$f[(-7, 11, -5)] = (11, -5, -7 + 11, -7 + 11) = (6, 4, 4), \text{ etc.}$$

Veamos ahora otro ejemplo, pero en el cual una base del primer espacio no es la canónica. Sea pues

$f : \mathbb{R}^2 \to \mathbb{R}$

tal que $f[(1, 1)] = 5$ y $f[(0, 1)] = -3$

Ahora bien, como el sistema $\{1, 1), (o, 1)\}$ es una base para $\mathbb{R}^2$, entonces la aplicación considerada f es lineal, según el teorema anterior y por esto existe y es única.

Tomemos entonces un vector cualquiera $(a, b) \in \mathbb{R}^2$ y expresémoslo primeramente como la combinación lineal de los vectores de la base considerada; se tiene

$$(a, b) = \alpha(1, 1) + \beta(o, 1) = (\alpha, \alpha + \beta)$$

de donde resulta,

$$\left.\begin{array}{l} \alpha = a \\ \alpha + \beta = b \end{array}\right\} \Longrightarrow \quad \begin{array}{l} \alpha = a \\ \hline \beta = b - a \end{array}$$

Luego, el transformado $f[(a, b)]$ se escribirá según la definición de la aplicación f, en la forma

$$f[(a, b)] = \alpha f[(1, 1)] + \beta f[(o, 1)] = \alpha \cdot 5 + \beta \cdot (-3)$$
$$f[(a\ b)] = 5\alpha - 3\beta$$

o sea,

$$f[(a, b)] = 5a - 3(b - a) = 8a - 3b$$

Por lo tanto, pueden escribirse los transformados que siguen:

$$f[(5, -7)] = 8 \cdot 5 - 3 \cdot (-7) = 40 + 21 = 61$$
$$f[(-1, -3)] = 8 \cdot (-1) - 3 \cdot (-3) = -8 + 9 = 1, \text{ etc.}$$

Volviendo a las observaciones hechas al final del teorema anterior, dijimos que el conjunto de vectores imágenes $\{\vec{b_1}, \vec{b_2}, \ldots, \vec{b_n}]$ puede ser elegido con toda libertad.

Ahora, lo normal será que utilicemos una base $\{\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}\}$ de V(K) y una base $\{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_m}\}$ de U(K).

Entonces, los transformados $\vec{b_1}, \vec{b_2}, \ldots, \vec{b_n}$, del teorema anterior pueden expresarse como combinación lineal de los vectores de la base $\{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_m}]$, y en tal caso las transformaciones lineales

$f : V(K) \to U(K)$

vendrán representadas por el sistema de ecuaciones:

$$(*) \qquad \begin{array}{l} f(\vec{a_1}) = \vec{b_1} = a_{11}\vec{u_1} + a_{12}\vec{u_2} + \ldots + a_{1m}\vec{u_m} \\ f(\vec{a_2}) = \vec{b_2} = a_{21}\vec{u_1} + a_{22}\vec{u_2} + \ldots + a_{2m}\vec{u_m} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ f(\vec{a_n}) = \vec{b_n} = a_{n1}\vec{u_1} + a_{n2}\vec{u_2} + \ldots + a_{nm}\vec{u_m} \end{array}$$

Decimos que las ecuaciones del sistema (*) representan la transformación lineal f respecto a las bases

$\{\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}\}$ y $\{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_m}\}$.

Es claro que el número de vectores de una base del espacio V(K) no tiene por qué ser el mismo que el de una base del espacio U(K), salvo que ambos tengan la misma dimensión.

Ahora, si fuese V = U, entonces utilizaremos de ordinario la misma base para ambos espacios, y la aplicación lineal $f : V(K) \to V(K)$ será evidentemente una sobrevección.

10.22. *Algunas propiedades de las transformaciones lineales.*

a) Sea $f : V(K) \to U(K)$ una transformación lineal, y sea S(K) un subespacio del espacio V(K).

Probaremos que la imagen directa f(S) es un subespacio del espacio U(K).

En efecto, según el criterio que tenemos para los subespacios, bastará demostrar que f(S) está cerrado para la adición vectorial y la multiplicación escalar.

Tendremos, pues:

Si $\vec{x}, \vec{x'} \in S$, entonces también $\vec{x} + \vec{x'} \in S$, por ser S un subespacio. Por lo tanto, si $f(\vec{x}), f(\vec{x'}) \in f(S)$, entonces $f(\vec{x}) + f(\vec{x'}) = f(\vec{x} + \vec{x'}) \in f(S)$.

Así hemos probado que f (S) está cerrado respecto a la adición.

Por otra parte, si $\vec{x} \in S$, $\alpha \in K$, entonces $\alpha\vec{x} \in S$, por ser S un subespacio. Por consiguiente, si $f(\vec{x}) \in f(S)$, entonces $\alpha f(\vec{x}) = f(\alpha\vec{x}) \in f(S)$.

Así hemos probado que f(S) está cerrado respecto a la multiplicación escalar.

Este resultado y el anterior demuestran que la imagen f(S) es un subespacio de U(K).

En resumen: La imagen de un subespacio de V(K) es también un subespacio de U(K).

En particular, la imagen f(V) del espacio entero V(K) es un subespacio de U(K), que lo llamaremos *subespacio imagen* de la transformación lineal $f : V(K) \to U(K)$.

En particular, si $f : V(K) \to V(K)$, entonces se tiene evidentemente $f(V) = V$, si f es sobreyectiva.

b) Sea $f : V(K) \to U(K)$ una transformación lineal, y sea $S(K) \subset V(K)$ un subespacio.

Si el subespacio S(K) tiene una base $\{\vec{a_1}, \vec{a_2}, \ldots, \vec{a_m}\}$, entonces cada $\vec{x} \in S(K)$ tiene la forma

$$\vec{x} = \alpha_1\vec{a_1} + \alpha_2\vec{a_2} + \ldots + \alpha_m\vec{a_m}$$

luego, cada $f(x) \in f(S)$ tiene la forma

$$f(\vec{x}) = \alpha_1 f(\vec{a_1}) + \alpha_2 f(\vec{a_2}) + \ldots + \alpha_m f(\vec{a_m}).$$

Por lo tanto, se sigue que los vectores $f(\vec{a_1})$, $f(\vec{a_2})$, ..., $f(\vec{a_m})$ engendran el subespacio imagen $f(S) \subset U(K)$.

Pero estos vectores no son necesariamente una base para f(S); según el criterio fundamental de independencia, algún subconjunto de este conjunto es una base.

Así la dimensión del subespacio f(S) puede no exceder a m, la dimensión de S(K).

Por consiguiente, hemos demostrado que si S(K) es un subespacio de V(K) y si $f : V(K) \to U(K)$ es una transformación lineal, entonces f(S) es también un subespacio de U(K), verificándose.

$$\dim(f(S)) \leq \dim(S)$$

c) Sea $f: V(K) \to U(K)$ una transformación lineal. Llamaremos RANGO de la aplicación lineal f a la dimensión del subespacio imagen f(V).

Si denotamos por r(f) el rango de la aplicación lineal f, se puede entonces escribir la definición anterior en la forma:

$$r(f) = \dim(f(V))$$

Ahora, si fuese $r(f) = \dim(V)$, entonces diremos que f es una *transformación lineal no singular*, mientras que si fuese $r(f) < \dim(V)$, decimos que f es una *transformación lineal singular*.

Ejemplo

Sea el espacio $\mathbb{R}^3$, y sea $\{(1, o, o), (o, 1, o), (o, o, 1)\}$ su base canónica.

Definamos la siguiente aplicación lineal de $\mathbb{R}^3$ en $\mathbb{R}^3$:

$$f[(1, o, o)] = (o, 1, 1), f[(o, 1, o)] = (1, 1, 1), f[(o, o, 1)] = (1, o, o)$$

Luego el sistema

$$\{(o, 1, 1), (1, 1, 1), (1, o, o)\}$$

engendra el subespacio imagen $f(\mathbb{R}^3)$.

Al aplicar el criterio fundamental de independencia, vemos que este sistema de vectores es linealmente dependiente, puesto que se tiene

$$\begin{pmatrix} 1 & 1 & 1 \\ o & 1 & 1 \\ 1 & o & o \end{pmatrix}$$

.

$$\begin{pmatrix} 1 & 1 & 1 \\ o & 1 & 1 \\ o & -1 & -1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 & 1 \\ o & 1 & 1 \\ o & o & o \end{pmatrix}$$

y por tanto, sólo dos de ellos son linealmente independientes, esto es, el subespacio imagen es de $\dim(f(\mathbb{R}^3)) = 2$

Ahora bien, como $\dim(\mathbb{R}^3) = 3$, resulta entonces que nuestra aplicación f anteriormente definida es una transformación lineal singular.

En el ejemplo que dimos de ilustración del teorema que determina una aplicación lineal, vimos que el transformado $f[(a, b, c)]$ de un vector cualquiera de $\mathbb{R}^3$ en las condiciones del ejemplo presente, es el vector

$$f[(a, b, c)] = (b + c, a + b, a + b)$$

Luego, podremos encontrar vectores no nulos de $\mathbb{R}^3$ cuyo transformado sea el vector nulo $\vec{o}$. Para esto basta poner,

$$\begin{cases} b + c = o \\ a + b = o \\ a + = o \end{cases} \quad \text{, o sea} \quad \begin{cases} b + c = o \\ a + b = o \end{cases} \Longrightarrow \begin{cases} c = a \\ b = -a \end{cases}$$

y dando valores arbitrarios a la componente a, tendremos vectores no nulos de $\mathbb{R}^3$ cuyo transformado es el vector $\vec{o}$. Uno de estos vectores es $(2, -2, 2)$ que se obtiene tomando el valor $a = 2$.

d) *Aplicación lineal sobreyectiva*

Sea la aplicación lineal $f: V(K) \to U(K)$.

Sea X un subconjunto del espacio V, y sea $\bar{X}$ el subespacio de V(K) engendrado por el sistema X de vectores arbitrario de V(K).

Probaremos que el subespacio imagen $f(\overline{X})$ está engendrado por la imagen f(X) del subconjunto X.

En efecto, todo vector $\vec{y} \in f(\overline{X})$ es imagen de a lo menos un vector $\vec{x} \in \overline{X}$, el cual es a su vez combinación lineal de un número finito de vectores del sistema X

Por consiguiente, el vector y = f(X) es combinación lineal de las imagenes de estos vectores.

En virtud de este hecho, pasemos a estudiar el teorema que sigue.

Teorema Una condición necesaria y suficiente para que una aplicación lineal $f: V(K) \rightarrow U(K)$ sea sobreyectiva, es que los transformados de los vectores de una base de V(K) constituya un sistema de generadores del espacio U(K)

Dem. Sin hacer referencia alguna sobre las dimensiones de los espacios V(K) y U(K), consideremos una base B cualquiera del espacio V(K).

Por las consideraciones anteriormente expuestas, vemos que la imagen f(B) engendra el subespacio f(V). Ahora bien, si la aplicación lineal f es sobreyectiva, es decir f(V) = U, entonces f(B) engendra el espacio U(K).

Recíprocamente, si f(B) engendra el espacio U(K), entonces todo vector $\vec{y} \in U(K)$ es combinación lineal de un número finito de vectores de f(B), y es, por consiguiente, el transformado de un vector $\vec{x} \in V(K)$, y la aplicación lineal f será sobreyectiva.

Este último resultado y el anterior demuestran el teorema.

e) Núcleo e Imagen de una Transformación lineal

Sea la aplicación lineal $f: V(K) \rightarrow U(K)$.

Existen dos importantes subespacios, uno de V (K) y otro de U (K) que desempeñan un papel fundamental en la teoría de las transformaciones lineales. Pasemos a definirlos.

Definición 1) Llamaremos NUCLEO de f, que denotaremos por Nuc(f), al conjunto de todos los vectores $\vec{x} \in V(K)$ tales que $f(\vec{x}) = \vec{o}$.

Teorema. El Nuc(f) es un subespacio de V(K).

Dem. En primer lugar. nótese que Nuc(f) no es vacío, puesto que $\vec{o} \in$ Nuc(f), ya que $f(\vec{o}) = \vec{o}$.

Por lo tanto, sean $\vec{x}, \vec{x}' \in$ Nuc(f), y probemos que también $\vec{x} + \vec{x}' \in$ Nuc(f).

En efecto,

$$f(\vec{x} + \vec{x}') = f(\vec{x}) + f(\vec{x}'), \text{ por ser f lineal}$$
$$= \vec{o} + \vec{o}, \text{ porque } x, x' \in \text{Nuc(f)}$$
$$= \vec{o}$$

luego, $\vec{x} + x' \in$ Nuc(f).

Por otro lado, demostraremos que si $\vec{x} \in$ Nuc(f), entonces también $\alpha \vec{x} \in$ Nuc(f), cualquiera sea $\alpha \in K$.

En efecto,

$$f(\alpha \vec{x}) = \alpha f(\vec{x}), \text{ por ser f lineal}$$
$$= \alpha \cdot \vec{o}, \text{ porque } \vec{x} \in \text{Nuc(f)}$$
$$= \vec{o}$$

luego, $\alpha \vec{x} \in$ Nuc(f).

Por consiguiente, el conjunto Nuc(f) verifica las dos condiciones de la definición de subespacio.

Este subespacio será también llamado el *espacio nulo,* y es de importancia fundamental en el estudio del comportamiento de f en V(K).

Definición 2). Llamaremos IMAGEN o también RECORRIDO de la aplicación lineal f, que denotaremos por Im(f), al conjunto de todos los elementos de U(K) de la forma $f(\vec{x})$. Aunque este subconjunto de U(K) ha sido ya considerado anteriormente y que fué denotado por f(V), lo consideraremos de nuevo y también volveremos a probar que es un subespacio del espacio (U(K). Tenemos, pues, así el teorema que sigue.

Teorema. La imagen o el recorrido Im(f) es un subespacio de U(K).

Dem. Si $f(\vec{x}), f(\vec{x}') \in$ Im(f), entonces $f(\vec{x}) + f(\vec{x}') \in$ Im(f).

En efecto, se tiene, por una parte

$$f(\vec{x}) + f(\vec{x}') = f(\vec{x} + \vec{x}') \in \text{Im(f)}$$

y por otra, $\alpha f(\vec{x}) = f(\alpha \vec{x}) \in$ Im(f)

Por consiguiente, Im(f) verifica el criterio de los subespacios.

Ahora que hemos introducido los conceptos de espacio nulo o núcleo de f y de imagen o recorrido de una transformación lineal, nos propondremos dar un vistazo de cerca de aquellas transformaciones lineales $f: VK) \rightarrow U(K)$ para las que se tienen

$$\text{Nuc(f)} = \{\vec{o}\}, \text{ o, } \text{Im(f)} = U(K)$$

o ambas.

Tenemos así los teoremas que siguen a continuación.

Teorema. Una condición necesaria y suficiente para que una transformación lineal $f: V(K) \rightarrow U(K)$ sea inyectiva, es que el núcleo se reduzca al solo vector nulo de V(K).

Dem. Supongamos que f sea inyectiva; entonces cada elemento de la Im(f) = f(V) no tiene más que una sola preimagen. Luego, en particular, al vector cero, $\vec{o}$, de U(K) no debe corresponderle más que un solo elemento de V(K), y como $f(\vec{o}) = \vec{o}$ se concluye que:

$$\text{Nuc(f)} = f^{-1}(\vec{o}) = \{\vec{o}\}.$$

Así hemos probado que si la transformación lineal f es inyectiva, entonces el núcleo de f se reduce únicamente al vector nulo, $\vec{o}$, del espacio V(K).

Recíprocamente, supongamos ahora que $\text{Nuc}(f) = \{\vec{o}\}$; entonces, demostraremos que la aplicación lineal f es inyectiva.

En efecto, supongamos que se tuviese $f(\vec{x}) = f(\vec{x}')$; entonces se tendrá $f(\vec{x} - \vec{x}') = \vec{o}$; luego

$$\vec{x} - \vec{x}' \in \text{Nuc}(f)$$

Por consiguiente, si $\text{Nuc}(f) = \vec{o}$, entonces resulta $\vec{x} - \vec{x}' = \vec{o}$, o sea, $\vec{x} = \vec{x}'$.

Así hemos probado que si el núcleo de la transformación lineal f está reducido solamente al vector nulo del espacio V(K), entonces la aplicación lineal f es inyectiva.

Consecuencia Importante

Las transformaciones lineales que son tanto inyectivas como sobreyectivas, es decir biyectivas, son llamadas isomorfismos, como ya sabemos. Ellas son de particular importancia ya que, lo mismo que las aplicaciones ordinarias biyectivas, tienen inversas. Por lo tanto, podemos formular el teorema que sigue.

Teorema. Una condición necesaria y suficiente para que una transformación lineal $f: V(K) \to U(K)$ de V(K) sobre U(K), esto es $f(V) = \text{Im}(f) = U(K)$, sea biyectiva, es que el Nuc(f) se reduzca únicamente al vector nulo de V(K). La aplicación inversa f^{-1} es entonces una transformación lineal de U(K) sobre V(K).

Dem. Supongamos que f sea biyectiva; entonces, en particular, f es inyectiva; luego, por el teorema anterior, el Nuc(f) se reduce al solo elemento $\vec{o}$ de V(K).

Recíprocamente, supongamos que sea $f(V) = U(K)$ y, además, sea $\text{Nuc}(f) = \{\vec{o}\}$. Entonces, por una parte, la aplicación lineal f es sobreyectiva y, por otra parte, por el teorema anterior es inyectiva. Por lo tanto, f es biyectiva.

Por otro lado, sean ahora $\vec{y}, \vec{y}'$ dos elementos cualesquiera de $f(V) = U(K)$, y sean

$$\vec{x} \in f^{-1}(\vec{y}), \vec{x}' \in f^{-1}(\vec{y}')$$

Hemos visto que si $\text{Nuc}(f) = f^{-1}(\vec{o}) = \{\vec{o}\}$, entonces $\vec{x}$ y $\vec{x}'$ son únicos; pero en todos los casos se tiene

$$\vec{y} + \vec{y}' = f(\vec{x}) + f(\vec{x}') = f(\vec{x} + \vec{x}')$$

luego, $\vec{x} + \vec{x}' \in f^{-1}(\vec{y} + \vec{y}')$

y $\alpha y = \alpha f(\vec{x}) = f(\alpha x)$

luego, $\alpha \vec{x} \in f^{-1}(\alpha \vec{y})$

Si $\text{Nuc}(f) = f^{-1}(\vec{o}) = \{\vec{o}\}$, se ve así que f^{-1} es una aplicación lineal de U(K) sobre V(K). Las aplicaciones lineales f y f^{-1} establecen un isomorfismo entre V(K) y U(K). Si $\text{Nuc}(f) = f^{-1}(\vec{o}) \neq \{\vec{o}\}$, entonces la demostración precedente nos muestra que, si U'(K) es un subespacio vectorial de U(K), el subconjunto $V' = f^{-1}(U')$ de V(K) es un subespacio vectorial del espacio V(K).

f) Condición que debe cumplir una transformación lineal para que sea singular o no singular

Primeramente, recordemos que llamamos *rango u orden* de una aplicación lineal $f: V(K) \to U(K)$ a la dimensión del subespacio imagen f(V). Si con r(f) designamos el rango de f, se escribe

$$r(f) = \dim(f(V))$$

Ahora bien, vimos que:

1) Si $r(f) = \dim(f(V)) < \dim(V)$, entonces dijimos que la transformación lineal f es singular.

2) Si $r(f) = \dim(f(V)) = \dim(V)$, entonces dijimos que la transformación lineal f es no singular.

Para caracterizar a cada uno de estos dos casos, tenemos los dos teoremas siguientes.

Teorema 1. Una transformación lineal $f: V(K) \longrightarrow U(K)$ es singular sí, y sólo sí, existe al menos un $\vec{o} \neq \vec{x} \in V(K)$ tal que $f(\vec{x}) = \vec{o}$.

Dem. Sea $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$ una base de V(K).

Supongamos que la aplicación lineal f sea singular, esto es, $r(f) < \dim(V)$; entonces, los n generadores

$$f(\vec{a}_1), f(\vec{a}_2), \ldots, f(\vec{a}_n)$$

del subespacio imagen f(V), deben ser linealmente dependientes. Luego, existen escalares α_1 no todos nulos tales que $\alpha_1 f(\vec{a}_1) + \alpha_2 f(\vec{a}_2) + \ldots + \alpha_n f(\vec{a}_n) = \vec{o}$

o sea $f(\alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \ldots + \alpha_n \vec{a}_n) = f(\vec{x}) = \vec{o}$

y por tanto, resulta $\vec{x} \neq \vec{o}$.

Recíprocamente, supongamos ahora que se tenga

$$f(\vec{x}) = \vec{o}, \text{ para algún } \vec{x} \neq \vec{o}$$

es decir, $\vec{x} = \alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \ldots + \alpha_n \vec{a}_n \neq \vec{o}$,

Luego no todos los escalares α_i son nulos.

Por otra parte, según nuestra hipótesis supuesta, se tiene

$$f(\vec{x}) = \alpha_1 f(\vec{a}_1) + \alpha_2 f(\vec{a}_2) + \ldots + \alpha_n f(\vec{a}_n) = \vec{o}$$

y como todos los α_i son ceros, resulta que los vectores $f(\vec{a}_1), f(\vec{a}_2), \ldots, f(\vec{a}_n)$ son linealmente dependientes.

Por consiguiente, dim (f(V)) < dim (V) = n, y la aplicación lineal f es singular.

Este resultado y el anterior demuestran el teorema.

Teorema 2. Una transformación lineal $f: V(K) \longrightarrow U(K)$ es no singular sí, y sólo sí, f es una inyección.

Dem. Supongamos primeramente, por el contrario, que f no es inyectiva; luego, puede ocurrir que siendo $f(\vec{x}) = f(\vec{x})$ no sea $\vec{x} = \vec{x}$; es decir, $\vec{x} \neq \vec{x}$.

Pero, $f(\vec{x}) = f(\vec{x})$ implica $f(\vec{x}) - f(\vec{x}) = \vec{o}$, o sea

$$f(\vec{x} - \vec{x}) = \vec{o}, \quad \text{con } \vec{x} = \vec{x} \neq \vec{o}$$

y, por tanto, f sería una transformación lineal singular.

Este resultado contrario a la hipótesis, mestra que solamente en el caso de que f sea inyectiva, la aplicación lineal f es no singular.

Recíprocamente, supongamos ahora que, también por el contrario, f no sea no singular; vale decir que sea singular. Entonces, se podrá tener

$$f(\vec{x}) = \vec{o}, \quad \text{con algún } \vec{x} \neq \vec{o}$$

Por otra parte, en toda transformación lineal se tiene siempre

$$f(\vec{o}) = \vec{o}$$

De donde, concluímos que

$$f(\vec{x}) = f(\vec{o}) \quad \text{implica} \quad \vec{x} \neq \vec{o}$$

luego, f no es inyectiva.

Este resultado contrario a la hipótesis, muestra que solamente en el caso de que f sea no singular, la aplicación lineal f es inyectiva.

Este resultado y el anterior prueban el teorema.

De los dos teoremas que se acaban de demostrar, resulta que podremos dar otra definición de aplicación lineal singular y no singular.

Definición. Sea la aplicación lineal $f: V(K) \longrightarrow U(K)$.

Entonces:

1. Diremos que la aplicación lineal f es *singular*, si la imagen de algún vector no nulo mediante f es el vector $\vec{o}$; esto es, si existe $\vec{x} \in U(K)$ tal que

$$\vec{x} \neq \vec{o} \quad \text{y} \quad f(\vec{x}) = \vec{o}$$

2. De la definición 1) concluimos entonces que la aplicación lineal f es *no singular* si solamente el $\vec{o} \in V(K)$, es aplicado en el vector $o \in U(K)$, o, en forma equivalente, si su núcleo consta únicamente del vector $\vec{o}$ de V(K); esto es, si $\text{Nuc}(f) = \{\vec{o}\}$ entonces f es inyectiva.

Ahora si la aplicación lineal $f: V(K) \longrightarrow U(K)$ es biyectiva, entonces solamente $\vec{o} \in V(K)$ puede aplicarse en el $\vec{o} \in U(K)$ y, por tanto, f es no singular.

Por otra parte, tal como se vio anteriormente, la recíproca f^{-1} de una aplicación lineal f biyectiva es también una aplicación lineal biyectiva, resulta entonces que también f^{-1} es no singular.

Por consiguiente, hemos probado que:

Teorema. Una aplicación lineal $f: V(K) \longrightarrow U(K)$ es un isomorfismo sí, y sólo si, es no singular.

Por otro lado, haremos notar que las aplicaciones lineales no singulares pueden también ser caracterizados por el hecho de que ellas envían sistemas de vectores independientes en sistemas de vectores independientes. Tenemos pues, el teorema siguiente:

Teorema. La aplicación $f: V(K) \longrightarrow U(K)$ es no singular sí, y sólo sí la imagen de un sistema independiente de V(K) es un sistema independiente de U(K).

Dem. Supongamos que f sea no singular y que el sistema $\{\vec{a_1}, \vec{a_2}, \ldots, \vec{a_m}\}$ sea linealmente independiente en V(K).

Probaremos que también el conjunto imagen

$\{f(\vec{a_1}), f(\vec{a_2}), \ldots, f(\vec{a_m})\}$ es un sistema linealmente independiente en U(K).

En efecto, supongamos que se tuviese la combinación lineal nula siguiente:

$$\alpha_1 f(\vec{a_1}) + \alpha_2 f(\vec{a_2}) + \ldots + \alpha_m f(\vec{a_m}) = \vec{o}$$

Entonces, como f es aplicación lineal, escribimos la igualdad anterior en la forma

$$f(\alpha_1 \vec{a_1} + \alpha_2 \vec{a_2} + \ldots + \alpha_m \vec{a_m}) = \vec{o}.$$

Luego, el vector

$$\vec{x} = \alpha_1 \vec{a_2} + \alpha_2 \vec{a_2} + \ldots + \alpha_m \vec{a_m} \in \text{Nuc}(f)$$

Pero, como hemos supuesto f no singular, entonces $\text{Nuc}(f) = \{\vec{o}\}$; luego, resulta que

$$\vec{x} = \alpha_1 \vec{a_2} + \alpha_2 \vec{a_2} + \ldots + \alpha_m \vec{a_m} = \vec{o}$$

y como el sistema $\{a_1, a_2, \ldots, a_m\}$ es L.I., entonces todos los α_i son ceros.

Por consiguiente, el sistema $\{f(\vec{a_1}), f(\vec{a_2}), \ldots, f(\vec{a_m})\}$ es también L.I.

Recíprocamente, supongamos ahora que la imagen de cualquier sistema independiente es independiente. Probaremos entonces que, la aplicación lineal f es no singular.

En efecto, si $\vec{a_1} \in V(K)$ es distinto del vector $\vec{o}$, entonces el sistema $\{\vec{a_1}\}$ es independiente. Luego, el sistema $\{f(\vec{a_1})\}$ es independiente y, por tanto, $f(\vec{a_1}) \neq \vec{o}$.

Por consiguiente, la aplicación lineal f es no singular, y el teorema está demostrado.

g) Relación entre las dimensiones del núcleo y de la imagen de una transformación lineal

Teorema. Sea $f: V(K) \longrightarrow U(K)$ una transformación lineal. Entonces se tiene la relación:

$$\dim(V) = \dim(\text{Nuc}(f)) + \dim(\text{Im}(f))$$

Dem. Hay dos casos que contemplar:

$\text{Nuc}(f) = V(K)$ y $\text{Nuc}(f) \neq V(K)$

En el primer caso, resulta que para todo $\vec{x} \in V(K)$ es $f(\vec{x}) = \vec{o}$; luego, f es la aplicación lineal nula de V(K) en U(K). Por lo tanto, $\text{Im}(f) = f(V) = \{\vec{o}\}$ y $\dim(\text{Im}(f)) = \vec{o}$.

En este caso particular, la fórmula

$$\dim(V) = \dim(\text{Nuc}(f)) + \dim(\text{Im}(f))$$

es verificada.

Excluyendo este caso trivial, consideremos el caso general en el cual $\text{Nuc}(f) \neq V(K)$.

Sea $r = \dim(\text{Nuc}(f))$ y $n = \dim(V)$, y probaremos que la $\dim(\text{Im}(f)) = n - r$.

En efecto, podemos determinar una base $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_r, a_{r+1}, \ldots, \vec{a}_n\}$ del espacio vectorial V(K) de manera que los r primeros vectores constituyan una base del Nuc(f).

Esto es evidente si $r = o$, pues ella puede elegirse en forma arbitraria.

Prescindiendo de este caso particular, supongamos entonces que sea $r \neq o$.

Esto supuesto, comenzaremos eligiendo una base $\{\vec{a}_1, \vec{a}_1, \ldots, \vec{a}_r$ del Nuc(f).

Ahora bien, estos r vectores $a_1, a_2, \ldots a_r$, independientes, no siendo una base del espacio V(K), pues $r < n$, pueden, en virtud del teorema de la base incompleta, ampliarse a una base de V(K). Sea pues, $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_r, \vec{a}_{r+1}, \ldots, \vec{a}_n\}$ el sistema ampliado y base de V(K).

Entonces, para probar que, $\dim(\text{Im}(f)) = n - r$, bastará mostrar que los vectores imágenes

$$f(\vec{a}_{r+1}), f(\vec{a}_{r+2}), \ldots, f(\vec{a}_n)$$

constituyen una base del subespacio $\text{Im}(f) = f(V)$.

En efecto, supongamos que se tuviese

$$\beta_1 f(\vec{a}_{r+1}) + \beta_2 f(\vec{a}_{r+2}) + \ldots + \beta_{n-r} f(\vec{a}_n) = \vec{o}$$

y como la aplicación f es lineal, la última igualdad se escribe también en la forma

$$f(\beta_1 \vec{a}_{r+1} + \beta_2 \vec{a}_{r+2} + \ldots + \beta_{n-r} \vec{a}_n) = \vec{o}$$

y lo que implica que el vector

$$\beta_1 \vec{a}_{r+1} + \beta_2 \vec{a}_{r+2} + \ldots + \beta_{n-r} a_n \in \text{Nuc}(f)$$

Si $\text{Nuc}(f) = \{\vec{o}\}$, entonces $r = o$, y tenemos inmediatamente

$$\beta_1 = \beta_2 = \ldots\ldots\ldots = \beta_{n-r} = o$$

Si no es este caso, será

$$\beta_1 \vec{a}_{r+1} + \beta_2 \vec{a}_{r+2} + \ldots + \beta_{n-r} \vec{a}_n = \alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \ldots + \alpha_r a_r$$

Luego,

$$(-\alpha_1)\vec{a}_1 + (-\alpha_2)\vec{a}_2 + \ldots + (-\alpha_r)\vec{a}_r + \beta_1 \vec{a}_{r+1} + \ldots + \beta_{n-r} \vec{a}_n = o$$

Como el sistema $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_r, \vec{a}_{r+1}, \ldots, \vec{a}_n\}$ es L.I., entonces todos los coeficientes de esta última combinación lineal nula son iguales a cero, en particular

$$\beta_1 = \beta_2 = \ldots\ldots = \beta_{n-r} = o$$

Así hemos probado que los vectores imágenes

$$f(\vec{a}_{r+1}), f(\vec{a}_{r+2}), \ldots, f(\vec{a}_n)$$

son linealmente independientes.

La misma demostración muestra que estos vectores imágenes son distintos dos a dos, y por lo tanto que son exactamente $n - r$ vectores.

Ahora nos resta probar que estos $n - r$ vectores imágenes engendran el subespacio $\text{Im}(f) = f(V)$.

En la sección 10.22. b), vimos que si los vectores: $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_r, \vec{a}_{r+1}, \ldots, \vec{a}_n$ engendran el espacio V(K), entonces los vectores $f(\vec{a}_1), f(\vec{a}_2), \ldots, f(\vec{a}_r), f(\vec{a}_{r+1}), \ldots, f(\vec{a}_n)$, engendran el subespacio imagen $\text{Im}(f) = f(V)$.

Pero, como $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_r \in \text{Nuc}(f)$, entonces los r primeros vectores imágenes anteriores coinciden con el vector nulo $\vec{o} \in U(K)$; luego, los vectores

$$\vec{o}, f(\vec{a}_{r+1}), f(\vec{a}_{r+2}), \ldots, f(\vec{a}_n)$$

engendran la $\text{Im}(f) = f(V)$.

Por consiguiente, los vectores $f(\vec{a}_{r+1}), f(\vec{a}_{r+2}), \ldots, f(\vec{a}_n)$ generan a $\text{Im}(f) = f(V)$.

Este resultado y el anterior muestran que el sistema

$$\{f(\vec{a}_{r+1}), f(\vec{a}_{r+2}), \ldots, f(\vec{a}_n)\}$$

constituye una base del subespacio Im(f).

Por consiguiente, resulta que

$$\dim(\text{Im}(f)) = n - r = \dim(V) - \dim(\text{Nuc}(f))$$

y de donde,

$$\dim(V) = \dim(\text{Nuc}(f)) + \dim(\text{Im}(f))$$

y el teorema está demostrado.

Observaciones. Mediante el teorema que se acaba de demostrar, podemos recobrar en forma bastante simple muchas de las propiedades de las transformaciones lineales que hemos venido estudiando. Veamos algunas de ellas.

Proposición 1. Si $f: V(K) \longrightarrow U(K)$ es una aplicación lineal biyectiva (isomorfismo), entonces $\dim(V) = \dim(U)$.

Dem. Por ser f, en particular, inyectiva, debe tenerse $\text{Nuc}(f) = \{\vec{o}\}$, o sea $\dim(\text{Nuc}(f)) = o$.

Por otra parte, por ser f, en particular, sobreyectiva, deberá tenerse $\text{Im}(f) = U$, o sea $\dim(\text{Im}(f)) = \dim(U)$.

Luego, $\dim(V) = \dim(\text{Nuc}(f)) + \dim(\text{Im}(f)) = \dim(U)$.

Proposición 2. Sea $f: V(K) \longrightarrow U(K)$ una aplicación lineal tal que $\dim(V) = \dim(U)$. Entonces, son equivalentes las siguientes proposiciones:

a) f es inyectiva
b) f es sobreyectiva
c) f es biyectiva

Dem. Sólo basta probar que a) $\Rightarrow$ b) y b $\Rightarrow$ a).

En efecto, supongamos f inyectiva, es decir $\text{Nuc}(f) = \{\vec{o}\}$; luego, $\dim(\text{Nuc}(f)) = o$.

Por consiguiente

$$\dim(V) = \dim(\text{Nuc}(f)) + \dim(\text{Im}(f)) = \dim(\text{Im}(f))$$

y como es $\dim(V) = \dim(U)$, resulta

$$\dim(U) = \dim(\text{Im}(f)) = n$$

siendo $\text{Im}(f) \subseteq U$.

Ahora bien, como es $\dim(\text{Im}(f)) = n$, entonces hay n vectores L.I.: $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n$ tales que

$$[\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n] = \text{Im}(f) \qquad (1)$$

Pero, siendo U también de dimensión n, los n vectores L.I.: $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n$ son una base de U y por lo tanto lo generan; es decir,

$$[\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n] = U \qquad (2)$$

De (1) y (2), resulta que $U = \text{Im}(f)$; y por tanto, f es sobreyectiva.

Así hemos probado que a) $\Rightarrow$ b)

Recíprocamente, supongamos ahora que sea f sobreyectiva, esto es, $\text{Im}(f) = U$; luego, $\dim(\text{Im}(f)) = \dim(U)$.

Por consiguiente,

$$\dim(V) = \dim(U) + \dim(\text{Nuc}(f))$$

y como es $\dim(V) = \dim(U)$, entonces la última igualdad implica $\dim(\text{Nuc}(f)) = o$, o sea $\text{Nuc}(f) = \{\vec{o}\}$.

Por consiguiente, f es inyectiva.

Así hemos demostrado que b $\Rightarrow$ a).

Este resultado y el anterior ponen de manifiesto la equivalencia de las proposiciones

a) $\Leftrightarrow$ b)

Luego, entre dos espacios de la misma dimensión, toda aplicación lineal inyectiva es también sobreyectiva y, recíprocamente. Por lo tanto, esta aplicación es una biyección.

Este resultado muestra entonces las equivalencias

$$\text{a)} \Leftrightarrow \text{b)} \Leftrightarrow \text{c)}$$

En consecuencia, una aplicación lineal biyectiva no puede establecerse más que entre dos espacios vectoriales de la misma dimensión.

Finalizaremos la presente sección, ilustrando a continuación con algunos ejemplos concretos los conceptos de núcleo e imagen o recorrido de una transformación lineal.

Ejemplo 1:

Sea $f: \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ la aplicación lineal definida por

$$f[(a, b, c)] = (a + 2b, b - c, a + 2c)$$

Hallar una base y la dimensión de la imagen y del núcleo de f.

Solución. Tomemos para el espacio $\mathbb{R}^3$ la base canónica

$$\vec{e}_1 = (1, 0, 0),\ \vec{e}_2 = (0, 1, 0),\ \vec{e}_3 = (0, 0, 1)$$

Entonces, en virtud de la definición de nuestra aplicación f, las imágenes de estos generadores de $\mathbb{R}^3$ engendran la imagen Im (f):

$$f(\vec{e}_1) = f[(1, 0, 0)] = (1 + 2\cdot 0,\ o - o,\ 1 + 2\cdot 0) = (1, 0, 1)$$
$$f(\vec{e}_2) = f[(0, 1, 0)] = (0 + 2\cdot 1,\ 1 - 0,\ 0 + 2\cdot 0) = (2, 1, 0)$$
$$f(\vec{e}_3) = f[(0, 0, 1)] = (0 + 2\cdot 0,\ 0 - 1,\ 0 + 2\cdot 1) = (0, -1, 2)$$

Formando la matriz de estos generados de la imagen y reduciéndola a la forma escalonada, se encuentra

$$\begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 0 \\ 0 & -1 & 2 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & -1 & 2 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & -0 \end{pmatrix}$$

Luego, el sistema $\{(1, 0, 1), (0, 1, -2)\}$ es una base de Im (f); por tanto, dim (Im (f)) = 2.

Ahora buscamos el conjunto de los vectores (a, b, c) tales que $f[(a,b,c,)] = \vec{0}$; esto es

$$f[(a, b, c)] = (a + 2b,\ b - c,\ a + 2c) = (0, 0, 0)$$

De donde el siguiente sistema lineal homogéneo cuyo espacio solución es el Nuc (f):

$$\left.\begin{array}{r} a + 2b = 0 \\ b - c = 0 \\ a + 2c = 0 \end{array}\right|$$

que reducido a la forma escalonada, se reduce a

$$\left.\begin{array}{r} a + 2b = 0 \\ b - c = 0 \\ -2b + 2c = 0 \end{array}\right|$$

$$\left.\begin{array}{r} a + 2b = 0 \\ b - c = 0 \\ 0 = 0 \end{array}\right|$$

$$\left.\begin{array}{r} a + 2b = 0 \\ b - c = 0 \end{array}\right|$$

y de donde la solución: a = − 2, b = c, c = c, por haber una sola incógnita arbitraria.

Haciendo c = − 1, obtenemos el vector (2, −1, −1) que genera el Nuc (f); por tanto dim (Nuc (f)) = 1.

Nótese que dim (Im (f)) + dim (Nuc (f)) = 2 + 1 = 3, que es la dimensión del dominio $\mathbb{R}^3$ de la aplicación lineal f.

Ejemplo 2:

Sea la aplicación lineal $f: \mathbb{R}^4 \longrightarrow \mathbb{R}^3$ definida por

$$f[(a, b, c, d)] = (a - b + c + d,\ a + 2c - d,\ a + b + 3c - 3d)$$

Encontrar, como en el ejemplo anterior, una base y la dimensión de la Im (f) y del Nuc (f).

Solución. Usando la base canónica de $\mathbb{R}^4$, tendremos los siguientes generadores de la Im (f), en virtud de la definición de f:

$$f(\vec{e_1}) = f[(1, 0, 0, 0)] = (1, 1, 1)$$
$$f(\vec{e_2}) = f[(0, 1, 0, 0)] = (-1, 0, 1)$$
$$f(\vec{e_3}) = f[(0, 0, 1, 0)] = (1, 2, 3)$$
$$f(\vec{e_4}) = f[(0, 0, 0, 1)] = (1, -1, -3)$$

Por consiguiente, la matriz de los generadores de la Im (f)

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

que reducida a la forma escalonada, dará los vectores independientes:

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & -2 & -4 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Por lo tanto, el sistema $\{(1, 1, 1), (0, 1, 2)\}$ es una base de Im(f); luego, dim (Im(f)) = 2.

Para los vectores del Nuc(f), es decir, para los que se tiene

$$f[(a, b, c, d)] = (a - b + c + d,\ a + 2c - d,\ a + b + 3c - 3d) = (0,0,0)$$

resulta el sistema homogéneo cuyo espacio solución es el Nuc(f):

$$\left.\begin{array}{r} a - b + c + d = 0 \\ a + 2c - d = 0 \\ a + b + 3c - 3d = 0 \end{array}\right|$$

o sea,

$$\left.\begin{array}{r} a - b + c + d = 0 \\ b + c - 2d = 0 \\ 2b + 2c - 4d = 0 \end{array}\right|$$

o bien,

$$\left.\begin{array}{r} a - b + c + d = 0 \\ b + c - 2d = 0 \\ 0 = 0 \end{array}\right|$$

Como hay solamente dos incógnitas arbitrarias c y d, entonces la dim (Nuc(f)) = 2.

Poniendo c = − 1, d = 0, obtendremos la primera solución

$$\begin{array}{rcl} a - b = -c - d &=& 1 \\ b = -c + 2d &=& 1 \end{array}$$

$$a = 2, \quad b = 1, \quad c = -1, \quad d = o$$

y poniendo ahora $c = o, d = 1$, tendremos la solución

$$\begin{array}{rcl} a - b &=& -1 \\ b &=& 2 \end{array}$$

$$a = 1, \quad b = 2, \quad c = o, \quad d = 1.$$

Luego, el sistema $\{(2, 1, -1, o), (1, 2, o, 1)\}$ es una base del Nuc(f).

Obsérvese que $\dim(\text{Im}(f)) + \dim(\text{Nuc}(f)) = 2 + 2 = 4$, que es la dimensión del dominio $\mathbb{R}^4$ de la aplicación f.

Ejemplo 3:

Los vectores $\vec{a_1} = (1, 1, -1)$, $\vec{a_2} = (1, o, 1)$, $\vec{a_3} = (2, 1, -1)$ forman una base de $\mathbb{R}^3$ y los vectores $\vec{b_1} = (1, o, 1, 0)$, $\vec{b_2} = (o, 1, 1, o)$, $\vec{b_3} = (1, o, o, 1)$, $\vec{b_4} = (1, 1, 1, o)$ forman una base de $\mathbb{R}^4$.

Se considera la aplicación lineal $f : \mathbb{R}^3 \to \mathbb{R}^4$ definida por:

$$\begin{aligned} f(\vec{a_1}) &= \vec{b_1} - \vec{b_4} &&= (o, -1, o, o) \\ f(\vec{a_2}) &= \vec{b_1} + \vec{b_2} + \vec{b_3} &&= (2, 1, 2, 1) \\ f(\vec{a_3}) &= \vec{b_1} + 2\vec{b_2} + 2\vec{b_3} + \vec{b_4} &&= (4, 3, 4, 2) \end{aligned}$$

Encontrar bases para la Im(f) y el Nuc(f).

Sol.: La matriz de los generadores del espacio Im(f) es

$$\begin{pmatrix} 2 & 1 & 2 & 1 \\ 0 & -1 & 0 & 0 \\ 4 & 3 & 4 & 2 \end{pmatrix}$$

que reducida a la forma escalonada da

$$\begin{pmatrix} 2 & 1 & 2 & 1 \\ o & -1 & o & o \\ o & 1 & o & o \end{pmatrix}$$

$$\begin{pmatrix} 2 & 1 & 2 & 1 \\ 0 & -1 & o & o \\ o & o & o & o \end{pmatrix}$$

Por consiguiente, el sistema libre

$$\{(2, 1, 2, 1), \quad (o, -1, o, o)\}$$

es una base para el espacio Im(f).

Buscamos en seguida el conjunto de los (a, b, c) tales que $f[(a, b, c)] = (o, o, o, o)$. Tenemos:

$$\begin{aligned} (a, b, c) &= \alpha\vec{a_1} + \beta\vec{a_2} + \gamma\vec{a_3} \\ f[(a, b, c)] &= \alpha f(\vec{a_1}) + \beta f(\vec{a_2}) + \gamma f(\vec{a_3}) \\ &= \alpha(o, -1, o, o) + \beta(2, 1, 1, 1) + \gamma(4, 3, 4, 2) \\ &= (2\beta + 4\gamma, -\alpha + \beta + 3\gamma, 2\beta + 4\gamma, \beta + 2\gamma) \end{aligned}$$

y de donde, el sistema lineal homogéneo

$$\begin{array}{rcl} -\alpha + \beta + 3\gamma &=& o \\ 2\beta + 4\gamma &=& o \\ 2\beta + 4\gamma &=& o \\ \beta + 2\gamma &=& o \end{array}$$

o sea,

$$\begin{array}{rcl} -\alpha + \beta + 3\gamma &=& o \\ \beta + 2\gamma &=& o \\ o &=& o \\ o &=& o \end{array} \quad \Rightarrow \begin{cases} \alpha = -\gamma \\ \beta = -2\gamma \end{cases}$$

y como no hay más una incógnita arbitraria, tenemos para $\gamma = 1$, el vector $(-1, -2, 1)$ que genera el Nuc(f) y es, por tanto, una base.

10.23. *Matrices y Transformaciones lineales de $\mathbb{R}^m$ en $\mathbb{R}^n$*

El estudio hecho hasta ahora de las transformaciones se ha aplicado, en general, a dos espacios vectoriales cualesquiera V(K) y U(K). Por razones de conveniencia, nos ocuparemos ahora específicamente con transformaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^m$. Además, estudiaremos con un poco de detalle las consecuencias de utilizar como bases de ambos espacios $\mathbb{R}^n$ en $\mathbb{R}^m$, sus bases naturales o canónicas $\{\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n}\}$ y $\{\vec{e'_1}, \vec{e'_2}, \ldots, \vec{e'_m}\}$, respectivamente. Esto significa que el sistema de ecuaciones (*) indicado en la sección 10.21., en lo referente a la determinación de una aplicación lineal, deberá ahora escribirse en la forma siguiente:

$$(*)' \quad \begin{array}{l} f(\vec{e_1}) = a_{11}\vec{e'_1} + a_{12}\vec{e'_2} + \ldots + a_{1m}\vec{e'_m} = (a_{11}, a_{12}, \ldots, a_{1m}) \\ f(\vec{e_2}) = a_{21}\vec{e'_1} + a_{22}\vec{e'_2} + \ldots + a_{2m}\vec{e'_m} = (a_{21}, a_{22}, \ldots, a_{2m}) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ f(\vec{e_n}) = a_{n1}\vec{e'_1} + a_{n}\vec{e'_2} + \ldots + a_{nm}\vec{e'_m} = (a_{n1}, a_{n2}, \ldots, a_{nm}) \end{array}$$

A la vista de los segundos miembros de las ecuaciones(*)', p. ece natural y lógico asociar a la transformación lineal, f, la ordenación rectangular n × m de números reales.

$$A = \begin{pmatrix} a_{11} & a_{12} \ldots a_{1k} \ldots a_{1m} \\ a_{21} & a_{22} \ldots a_{2k} \ldots a_{2m} \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{i1} & a_{i2} \ldots a_{ik} \ldots a_{im} \\ a_{n1} & a_{n2} \ldots a_{nk} \ldots a_{nm} \end{pmatrix}$$

Esto nos conduce a la generalización del concepto de matriz que hemos venido mencionando a lo largo del presente capítulo.

Definición: Llamaremos MATRIZ real de orden o de tipo n × m, a una ordenación de nm números reales dispuestos en n filas o renglones y en m columnas.

Los números individuales a_{ik} los llamaremos los *elementos* de la matriz.

Abreviadamente, una matriz de orden n × m será denotada en la siguiente forma $[a_{ik}]_{n,m}$
en donde (i = 1,2 . . . , n; k = 1,2, . . . , m).

En la presente sección no pretenderemos hacer un estudio demasiado detallado de la operatoria con matrices, sino solamente en forma muy breve. Un estudio más a fondo sobre el álgebra de matrices se hará al finalizar el TOMO III de estos apuntes.

La igualdad de matrices se define únicamente cuando ellas son del mismo tipo.

Definición: Decimos que dos matrices n × m, $[a_{ik}]_{n,m}$ y $[b_{ik}]_{n,m}$ son iguales si, y sólo si

$$a_{ik} = b_{ik}$$

para todo i y todo k.

Por consiguiente, dos matrices del mismo orden o tipo son iguales, cuando los elementos homólogos o correspondientes de cada una de las matrices sean iguales.

Ahora, si pensamos utilizar las matrices como una ayuda para expresar más adelante las operaciones entre transformaciones lineales, parece razonable entonces definir las operaciones algebraicas con matrices.

La haremos únicamente muy breve, porque su estudio detallado se hará, como lo hemos dicho más arriba, en el TOMO III.

La suma de matrices se define únicamente para matrices del mismo tipo de orden.

Definición: Sean $[a_{ik}]_{n,m}$ y $[b_{ik}]_{n,m}$ dos matrices del mismo orden n × m.

Definimos la *suma matricial* de las matrices $[a_{ik}]_{n,m}$ y $[b_{ik}]_{n,m}$ como la matriz $[c_{ik}]_{n,m}$, donde

$$c_{ik} = a_{ik} + b_{ik}$$

para todo i y todo k.

Por consiguiente,

$$[a_{ik}]_{n,m} + [b_{ik}]_{n,m} = [a_{ik} + b_{ik}]_{n,m}$$

es decir, las matrices del mismo orden se suman sumando los elementos homólogos o correspondientes.

Ejemplo:

$$\begin{pmatrix} 3 & 5 \\ -2 & 1 \\ 0 & 4 \end{pmatrix} + \begin{pmatrix} 1 & -2 \\ 4 & 2 \\ 1 & -3 \end{pmatrix} = \begin{pmatrix} 3+1 & 5-2 \\ -2+4 & 1+2 \\ 0+1 & 4-3 \end{pmatrix} = \begin{pmatrix} 4 & 3 \\ 2 & 3 \\ 1 & 1 \end{pmatrix}$$

Para matrices de distinto orden no definiremos la adición. Llamaremos matriz nula de tipo n × m a una matriz n × m cuyos elementos son todos nulos.

Es evidente que si 0 es la matriz nula n × m y A cualquier matriz n × m, entónces se tiene

$$A + 0 = 0 + A = A$$

El producto de matrices exige todavía hipótesis más restrictivas, como lo indica la definición que sigue

Definición: El producto de la matriz $[a_{ij}]_{n,m}$ por la matriz $[b_{jk}]_{m,p}$, es la matriz $[c_{ik}]_{n,p}$, donde

$$c_{ik} = \sum_{j=1}^{m} a_{ij}\, b_{jk}$$

Es decir, para obtener el elemento que está en la fila i-ésima y la columna k-ésima de la matriz producto, sumamos todos los productos parciales

que resultan de multiplicar ordenadamente los elementos de la fila i-ésima del primer factor matricial $[a_{ij}]_{n,m}$ por los elementos de la columna k-ésima del segundo factor matricial $[b_{jk}]_{m,p}$.

Ejemplo:

$$\begin{pmatrix} 5 & -1 & 0 \\ 2 & 3 & 1 \\ 4 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 9 \\ -1 & 7 \\ 4 & 8 \end{pmatrix} = \begin{pmatrix} 6 & 38 \\ 3 & 47 \\ 12 & 52 \end{pmatrix}$$

Explicación: los elementos de la primera fila del producto son:

$$\begin{cases} 5 \cdot 1 + (-1) \cdot (-1) + 0 \cdot 4 = 5 + 1 = 6 \\ 5 \cdot 9 + (-1) \cdot 7 + 0 \cdot 8 = 45 - 7 = 38 \end{cases}$$

los elementos de la segunda fila del producto son:

$$\begin{cases} 2 \cdot 1 + 3 \cdot (-1) + 1 \cdot 4 = 2 - 3 + 4 = 3 \\ 2 \cdot 9 + 3 \cdot 7 + 1 \cdot 8 = 18 + 21 + 8 = 47 \end{cases}$$

y los elementos de la tercera fila del producto son:

$$\begin{cases} 4 \cdot 1 + 0 \cdot (-1) + 2 \cdot 4 = 4 + 8 = 12 \\ 4 \cdot 9 + 0 \cdot 7 + 2 \cdot 8 = 36 + 16 = 52 \end{cases}$$

Definición: La matriz cuadrada de de orden n × n de la forma

$$I = \begin{pmatrix} 1 & 0 & 0 & \ldots\ldots & 0 \\ 0 & 1 & 0 & \ldots\ldots & 0 \\ 0 & 0 & 1 & \ldots\ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots\ldots & 1 \end{pmatrix}$$

la llamaremos *matriz unidad* de orden n.

Definición: Dada una matriz $[a_{ij}]_{n,m}$ y un número real p, la matriz $p\,[a_{ij}]_{n,m}$, producto del número p por la matriz $[a_{ij}]_{n,m}$, será por definición la que tiene por elementos $p\,a_{ij}$, para todo i y todo j.

Ejemplo:

$$5 \cdot \begin{pmatrix} 2 & -1 & 3 \\ -6 & 5 & 1 \\ 4 & -2 & 1 \\ 1 & 6 & -3 \end{pmatrix} = \begin{pmatrix} 10 & -5 & 15 \\ -30 & 25 & 5 \\ 20 & -10 & 5 \\ 5 & 30 & -15 \end{pmatrix}$$

Propiedades. Unicamente nos limitaremos a enunciar las propiedades de las operaciones matriciales definidas anteriormente. Sus demostraciones serán dadas en el TOMO III.

ADICION: Esta operación goza de las propiedades siguientes:

a) $\{[a_{ij}]_{(n,m)} + [b_{ij}]_{(n,m)}\} + [c_{ij}]_{(n,m)}$
$[a_{ij}]_{(n,m)} + \{[b_{ij}]_{(n,m)} + [c_{ij}]_{(n,m)}\}$

b) $[a_{ij}]_{(n,m)} + [b_{ij}]_{(n,m)} = [b_{ij}]_{(n,m)} + [a_{ij}]_{(n,m)}$

c) Dada la matriz $[a_{ij}]_{(n,m)}$ existe la matriz $[-a_{ij}]_{(n,m)}$ tal que:

$[a_{ij}]_{(n,m)} + [-a_{ij}]_{(n,m)} = 0$ (matriz nula).

Además, sabemos que se tiene

$[a_{ij}]_{(n,m)} + 0 = [a_{ij}]_{(n,m)}$.

En consecuencia, el conjunto de todas las matrices del tipo u orden n × m, forman un grupo con respecto a la adición matricial (Grupo Aditivo).

MULTIPLICACION. Esta operación satisface las propiedades siguientes:

a) $\{[a_{ij}]_{(n,m)} \cdot [b_{jk}]_{(m,p)}\} \cdot [c_{kr}]_{(m,p)} =$
$[a_{ij}]_{(n,m)} \cdot \{[b_{jk}]_{(m,p)} \cdot [c_{kr}]_{(p,q)}\}$

b) $[a_{ij}]_{(n,m)} \cdot \{[b_{jk}]_{(m,p)} + [c_{jk}]_{(m,p)}\} =$
$[a_{ij}]_{(n,m)} \cdot [b_{jk}]_{(m,p)} + [a_{ij}]_{(n,m)} \cdot [c_{jk}]_{(m,p)}$

c) $\{[a_{ij}]_{(n,m)} + [b_{ij}]_{(n,m)}\} \cdot [c_{jk}]_{(m,p)} =$
$[a_{ij}]_{(n,m)} \cdot [c_{jk}]_{(m,p)} + [b_{ij}]_{(n,m)} \cdot [c_{jk}]_{(m,p)}$

De las propiedades que hemos indicado para la adición y la multiplicación entre matrices, resulta en particular que el conjunto de todas las matrices cuadradas de orden n × n forman un anillo con elemento unidad.

Multiplicación escalar. El producto de un número real por una matriz, llamado multiplicación escalar, goza de las propiedades que siguen:

a) $p\{[a_{ij}]_{(n,m)} + [b_{ij}]_{(n,m)}\} = p[a_{ij}]_{(n,m)} + p[b_{ij}]_{(n,m)}$

b) $(p+q)[a_{ij}]_{(n,m)} = p[a_{ij}]_{(n,m)} + q[a_{ij}]_{(n,m)}$

c) $(pq)[a_{ij}]_{(n,m)} = p\{q[a_{ij}]_{(n,m)}\}$

d) $1 \cdot [a_{ij}]_{(n,m)} = [a_{ij}]_{(n,m)}$

En resumen: de las propiedades dadas para la adición, multiplicación y multiplicación escalar, concluimos que el conjunto de todas las matrices de orden n × m es un espacio vectorial sobre el cuerpo $\mathbb{R}$ de los números reales.

En la observación hecha al ejemplo 7) de la sección 10.20., vimos que cada matriz $A = [a_{ij}]_{n,m}$ de orden n × m define una transformación lineal

$$f : \mathbb{R}^m \to \mathbb{R}^n$$

definida por las igualdades

$$(*)\quad \begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1m}x_m \\ y_2 = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2m}x_m \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_n = a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nm}x_m \end{cases}$$

en que a cada $\vec{x} = (x_1, x_2, \ldots, x_m) \in \mathbb{R}^m$ hace corresponder el vector $\vec{y} = f(\vec{x}) = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$.

Conveniendo en describir:

$$\vec{x} = (x_1, x_2, \ldots, x_m) = \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_m \end{pmatrix}, \quad \vec{y} = (y_1, y_2, \ldots, y_n) = \begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{pmatrix}$$

entonces las igualdades o ecuaciones (*) toman la forma matricial siguiente

$$\begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_m \end{pmatrix}$$

es decir,

$$\vec{y} = A \cdot \vec{x} = [a_{ij}]_{n,m} \cdot \vec{x}$$

como puede verificarse al aplicar la regla de multiplicación de matrices y la definición de igualdad de dos matrices.

En particular, poniendo

$$\vec{e_1} = (1, 0, 0, \ldots, 0)$$
$$\vec{e_2} = (0, 1, 0, \ldots, 0)$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\vec{e_m} = (0, 0, 0, \ldots, 1)$$

se verifica fácilmente por cálculo directo que el producto $A\,\vec{e_k} = [a_{ij}]_{n,m} \cdot e_k$ es el vector del espacio $\mathbb{R}^n$ que tiene por componentes los elementos de la k-ésima columna de la matriz $A = [a_{ij}]_{n,m}$; esto es

$$(1)\quad \begin{pmatrix} a_{11} & a_{12} & \ldots\ldots & a_{1m} \\ a_{21} & a_{22} & \ldots\ldots & a_{2m} \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ a_{n1} & a_{n2} & \ldots\ldots & a_{nm} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ 1_k \\ \cdot \\ \cdot \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1k} \\ a_{2k} \\ \cdot \\ \cdot \\ a_{nk} \end{pmatrix}$$

En resumen: a cada matriz $A = [a_{ij}]_{n,m}$ de tipo u orden n x m es posible asociar la transformación lineal que lleva el vector $\vec{x} \in \mathbb{R}^m$ a vector $A\vec{x} \in \mathbb{R}^n$.

Por otra parte, es fácil ver que si dos matrices:
$A = [a_{ij}]_{n,m}$ y $B = [b_{ij}]_{n,m}$ de tipo u orden n x m
son distintas, entonces ellas definen también transformaciones lineales distintas.

En efecto, si $A \neq B$, entonces tienen al menos alguna columna diferente, y sea ésta la k-ésima. Tenemos entonces, por la igualdad (1) anterior que,

$$A \cdot \vec{e_k} \neq B \cdot \vec{e_k}$$

es decir,

$$\begin{pmatrix} a_{1k} \\ a_{2k} \\ \cdot \\ \cdot \\ \cdot \\ a_{nk} \end{pmatrix} \neq \begin{pmatrix} b_{1k} \\ b_{2k} \\ \cdot \\ \cdot \\ \cdot \\ b_{nk} \end{pmatrix}$$

Luego, las transformaciones lineales definidas por A y B entre los espacios $\mathbb{R}^m$ y $\mathbb{R}^n$ son distintas, puesto que ellas llevan el vector $\vec{e_k} \in \mathbb{R}^m$ a dos vectores distintos de $\mathbb{R}^n$.

Así hemos probado que toda matriz de tipo n x m define una única transformación lineal del espacio $\mathbb{R}^m$ dentro del espacio $\mathbb{R}^n$.

Inversamente, demostraremos ahora que toda transformación lineal $f : \mathbb{R}^m \to \mathbb{R}^n$ es definida por alguna matriz de tipo n x m.

En efecto, probaremos que dada la transformación lineal

$$(2) \quad f : \mathbb{R}^m \to {}^n\mathbb{R}$$

existe una matriz A de tipo n x m tal que, para todo vector $\vec{x} \in \mathbb{R}^m$ se tiene:

$$(3) \quad f(\vec{x}) = A \cdot \vec{x}$$

Pues bien, sea $\vec{x} = (x_1, x_2, \ldots, x_m) \in \mathbb{R}^m$

es decir, $\vec{x} = x_1 \vec{e_1} + x_2 \vec{e_2} + \ldots + x_m \vec{e_m}$

luego,

$$(4) \quad f(\vec{x}) = x_1 f(\vec{e_1}) + x_2 f(\vec{e_2}) + \ldots + x_m f(\vec{e_m})$$

Denotemos los vectores de $\mathbb{R}^n$, transformados de los vectores $\vec{e_k} \in \mathbb{R}^m$ mediante la aplicación f, por

$$(5) \quad \begin{cases} f(\vec{e_1}) = (a_{11}, a_{21}, \ldots, a_{n1}) \\ f(\vec{e_2}) = (a_{12}, a_{22}, \ldots, a_{n2}) \\ f(\vec{e_3}) = (a_{13}, a_{23}, \ldots, a_{n3}) \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ f(\vec{e_m}) = (a_{1m}, a_{2m}, \ldots, a_{nm}) \end{cases}$$

entonces la igualdad (4) se escribirá en la forma

$$f(x) = x_1 (a_{11}, a_{21}, \ldots, a_{n1}) + x_2 (a_{12}, a_{22}, \ldots, a_{n2}) + \ldots + x_m (a_{1m}, a_{2m}, \ldots, a_{nm})$$

o bien, al tener en vista las reglas del producto de un escalar por un vector y la suma de vectores,

$$(6) \quad f(\vec{x}) = (a_{11}x_1 + a_{12}x_2 + \ldots + a_{1m}x_m, a_{21}x_1 + a_{22}x_2 + \ldots + a_{2m}x_m, \ldots, a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nm}x_m)$$

Ahora, si ponemos

$$A = \begin{pmatrix} a_{11}\ a_{12} \ldots\ldots a_{1m} \\ a_{21}\ a_{22} \ldots\ldots a_{2m} \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}\ a_{n2} \ldots\ldots a_{nm} \end{pmatrix}$$

entonces, por la regla de multiplicación de matrices, la igualdad (6) así

$$f(x) = A \cdot \vec{x}$$

esto es

$$f(x) = \begin{pmatrix} a_{11}\ a_{12} \ldots\ldots a_{1m} \\ a_{21}\ a_{22} \ldots\ldots a_{2m} \\ \ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots \\ a_{n1}\ a_{n2} \ldots\ldots a_{nm} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_m \end{pmatrix}$$

resultado que demuestra nuestra aseveración.

En resumen: se ha demostrado que existe una correspondencia biunívoca entre "matrices de tipo n x m" y "transformaciones lineales de $\mathbb{R}^m \to \mathbb{R}^n$", esto es, que a cada matriz de tipo n x m corresponde una aplicación $f : \mathbb{R} \to \mathbb{R}^n$ y, recíprocamente, a cada aplicación lineal $f : \mathbb{R}^m \to \mathbb{R}^n$ corresponde una matriz de orden n x m. Además, se vio antes, un poco más arriba, que a matrices diferentes corresponden transformaciones lineales diferentes. Luego, si M (n x m) denota el conjunto de todas las matrices de orden n x m y TL ($\mathbb{R}^m, \mathbb{R}^n$) el conjunto de todas las aplicaciones lineales de $\mathbb{R}^m$ dentro de $\mathbb{R}^n$, entonces la aplicación

$$M\ (n \times m) \to TL\ (\mathbb{R}^m, \mathbb{R}^n)$$

que a cada matriz A asocia una transformación lineal definida por las igualdades (5), es una biyección.

En virtud de esta correspondencia existente entre matrices y transformaciones lineales, diremos generalmente en vez de "transformación

lineal $f : \mathbb{R}^m \to \mathbb{R}^n$", la "transformación lineal $A : \mathbb{R}^m \to \mathbb{R}^n$ definida por la matriz A".

Esta identificación entre matrices y transformaciones lineales nos será muy útil en la sección que sigue.

10.24. *Operaciones con transformaciones lineales*

Podemos combinar las transformaciones lineales de varias maneras para obtener nuevas transformaciones lineales.

Pasemos a estudiarlas.

A) Sean:

$$V(K) \xrightarrow{f} U(K) \xrightarrow{g} W(K)$$

dos transformaciones lineales de los tres espacios vectoriales V(K), U(K) y W(K) sobre el mismo cuerpo K de escalares.

Recordamos que la función compuesta

$$g \circ f : V(K) \to W(K)$$

es la aplicación de V(K) en W(K) definida por,

$$(g \circ f)(\vec{x}) = g(f(\vec{x})), \forall \vec{x} \in V(K)$$

Probaremos que esta aplicación compuesta es lineal, si lo son f y g.

En efecto, sean x, x' $\in$ V(K), arbitrarios. Entonces,

$$(g \circ f)(\vec{x}+\vec{x}') = g[f(\vec{x}+\vec{x}')] = g[f(\vec{x}) + f(\vec{x}')]$$

por ser f una transformación lineal.

Por otro lado, se tiene

$$(g \circ f)(\vec{x}+\vec{x}') = g[f(\vec{x}) + f(\vec{x}') = g(f(\vec{x})) + g(f(\vec{x}'))$$

por ser g una transformación lineal.

Luego,

$$(g \circ f)(\vec{x}+\vec{x}') = (g \circ f)(\vec{x}) + (g \circ f)(\vec{x}')$$

Por otra parte, tenemos

$$(g \circ f)(\alpha\vec{x}) = g(f(\alpha\vec{x})) = g(\alpha f(\vec{x})) = \alpha\,[g(f(\vec{x}))]$$

por ser f y g transformaciones lineales. Luego,

$$(g \circ f)(\alpha\vec{x}) = \alpha\,(g \circ f)(\vec{x})$$

Este resultado y el anterior demuestran que la aplicación compuesta

$$g \circ f : V(K) \to W(K)$$

es una transformación lineal y que la llamaremos *la aplicación lineal producto.*

Naturalmente, como todo producto de aplicaciones en general, es asociativa; es decir, se verifica la igualdad

$$(h \circ g) \circ f = h \circ (g \circ f)$$

cualesquiera que sean las transformaciones lineales f, g, h.

En particular, sean A y B transformaciones lineales actuando como se indica en el diagrama siguiente:

$$\mathbb{R}^m \xrightarrow{A} \mathbb{R}^n \xrightarrow{B} \mathbb{R}^p$$

y donde A es matriz de orden n x m, B es matriz de orden p x n.

La composición

$$B \circ A : \mathbb{R}^m \to \mathbb{R}^p$$

puede ser identificada a una matriz de orden p x m.

Veremos que ella es, precisamente, la matriz producto de B por A; es decir, $B \circ A = BA$.

En efecto, sea la compuesta $C = B \circ A$; para mostrar que $C = BA$, obtendremos la columna k-ésima de la matriz $C (1 \le k \le m)$:

$$C \cdot \vec{e}_k = (B \circ A)\,\vec{e}_k = B(A\,\vec{e}_k)$$

pero,

$$A\,\vec{e}_k \quad \begin{pmatrix} a_{1k} \\ a_{2k} \\ \cdot \\ \cdot \\ \cdot \\ a_{nk} \end{pmatrix}$$

luego,

$$Ce_k = \begin{pmatrix} c_{1k} \\ c_{2k} \\ \cdot \\ \cdot \\ c_{pk} \end{pmatrix} = \begin{pmatrix} b_{11}\ b_{12}\ \ldots\ b_{1n} \\ b_{21}\ b_{22}\ \ldots\ b_{2n} \\ \ldots\ldots\ldots \\ \ldots\ldots\ldots \\ b_{p1}\ b_{p2}\ \ldots\ b_{pn} \end{pmatrix} \cdot \begin{pmatrix} a_{1k} \\ a_{2k} \\ \cdot \\ \cdot \\ a_{nk} \end{pmatrix}$$

en donde,

$$c_{jk} = b_{j1}\,a_{1k} + b_{j2}\,a_{2k} + \ldots + b_{jn}\,a_{nk}$$

Luego, C coincide con el producto de las matrices B y A; es decir, hemos probado que la compuesta $C = B \circ A$ de las transformaciones lineales representadas por las matrices A y B coincide con $C = BA$.

Observación: Del diagrama

$$\mathbb{R}^m \xrightarrow{A} \mathbb{R}^n \xrightarrow{I} \mathbb{R}^n$$

en donde A es matriz de tipo n x m, I es la matriz unidad de orden n x n, se deduce que

$$I\,A = A$$

que trivialmente de verificación directa.

Análogamente, se tiene

$$A\,I = A$$

Ejemplos:

1. Sean las aplicaciones

$$\mathbb{R}^2 \xrightarrow{f} \mathbb{R}^2 \xrightarrow{g} \mathbb{R}^2$$

definidas por, $f[(x_1,x_2)] = (x_2,2x_1)$ y $g[(x_1,x_2)] = (x_1 + x_2,x_2)$.

Hallar las aplicaciones compuestas g o f, f o g.

Tenemos:

se puede fácilmente comprobar que f y g son lineales.

Luego, podremos escribir

$(g \circ f)[(x_1,x_2)] = g[f(x_1,x_2)] = g[(x_2,2x_1)] = (x_2 + 2x_1,2x_1)$

$(f \circ g)[(x_1,x_2)] = f[g(x_1,x_2)] = f[(x_1 + x_2,x_2)] = (x_2,2x_1 + 2x_2)$

En particular, se tiene

$(g \circ f)[(1, 3)] = (3 + 2, 2) = (5, 2)$

$(f \circ g)[(1, 3)] = (3, 2 + 6) = (3, 8)$

Estos resultados nos muestran que el producto de transformaciones lineales en general no es conmutativo.

2. Sean las transformaciones lineales

$$\mathbb{R}^3 \xrightarrow{f} \mathbb{R}^3 \xrightarrow{g} \mathbb{R}^3$$

definidas o dadas por las matrices

$$f \to A = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \\ 0 & 2 & 1 \end{pmatrix}$$

$$g \to B = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 2 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

Hallar la matriz asociada a la transformación lineal g o f, y la imagen del vector $(1, -1, o)$ mediante g o f.

Tenemos:

a) $$g \circ f \to BA = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 2 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \\ 0 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 4 & -1 \\ 3 & 2 & 1 \\ 2 & 4 & 1 \end{pmatrix}$$

b) $$(g \circ f)[(1, -1, 0)] = BA \cdot \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 4 & -1 \\ 3 & 2 & 1 \\ 2 & 4 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -4 \\ 1 \\ -2 \end{pmatrix}$$

B) Sean $f : V(K) \to U(K)$, $g : V(K) \to U(K)$, dos transformaciones lineales de los espacios V(K) y U(K) sobre el mismo cuerpo K. Entonces, una nueva transformación de la forma,

$f + g : V(K) \to U(K)$

puede definirse de una manera natural como sigue:

$$(f + g)(\vec{x}) = f(\vec{x}) + g(\vec{x}), \quad \forall \vec{x} \in V(K)$$

De esta manera, la compuesta f + g está definida en términos de la adición en el espacio U(K).

Ahora probaremos que siendo f, g transformaciones lineales, la compuesta f + g es también una transformación lineal.

En efecto, sean $\vec{x}, \vec{x'} \in V(K)$ arbitrarios; entonces

$$(f+g)[(\vec{x}+\vec{x'})] = f[(\vec{x}+\vec{x'})] + g[(\vec{x}+\vec{x'})]$$
$$= f(\vec{x}) + f(\vec{x'}) + g(\vec{x}) + g(\vec{x'})$$
$$= (f+g)(x) + (f+g)(x')$$

Por otra parte, se tiene

$$(f+g)(\alpha\vec{x}) = f(\alpha\vec{x}) + g(\alpha\vec{x})$$
$$= \alpha f(\vec{x}) + \alpha g(\vec{x})$$
$$= \alpha [f(\vec{x}) + g(\vec{x})]$$
$$= \alpha (f + g)(\vec{x})$$

este resultado y el anterior muestran que la compuesta f + g es una transformación lineal y que la llamaremos *la aplicación lineal suma.*

La conmutatividad y la asociatividad de la adición de aplicaciones resultan fácilmente de las propiedades análogas sobre V(K) y U(K). Luego,

$$(f + g) + h = f + (g + h)$$
$$f + g = g + f$$

cualquiera sea $\vec{x} \in V(K)$ y cualesquiera sean las aplicaciones f, g, h. Por otra parte, existe una aplicación neutra $\vec{0}$, que a todo vector $\vec{x} \in V(K)$ asocia el vector $0(\vec{x}) = \vec{o} \in U(K)$.

Luego,

$$\vec{0} + \vec{x} = \vec{x} + \vec{0} = \vec{x}, \quad \forall \vec{x} \in V(K)$$

Por último, la aplicación opuesta de f es $(-f)$ tal que

$$(-f)(\vec{x}) = -f(\vec{x}), \quad \forall \vec{x} \in V(K)$$

y se tiene,

$$f + (-f) = (-f) + f = 0$$

Por consiguiente, esta operación de adición de aplicaciones lineales hace del conjunto de todas las aplicaciones lineales de V(K) en U(K), un grupo aditivo. En particular, sean A y B transformaciones lineales actuando como se indica en el diagrama que sigue:

$$\mathbb{R}^m \underset{B}{\overset{A}{\rightrightarrows}} \mathbb{R}^n$$

en donde A y B son matrices de orden n x m.

Es de demostración inmediata que la suma de las transformaciones lineales A y B coincide con la suma de las matrices A y B.

Ejemplos:

1. Sean las transformaciones lineales

$$f : \mathbb{R}^2 \to \mathbb{R}^3$$
$$f : \mathbb{R}^2 \to \mathbb{R}^3$$

definidas por: $f[(x_1,x_2)] = (x_1 - x_2, x_2, x_1)$

$g[(x_1,x_2)] = (x_2, x_1 - x_2, x_1 + x_2)$

Encontrar la aplicación lineal suma $f + g$.

Tenemos:

$$(f+g)[(x_1,x_2)] = f[(x_1+x_2)] + g[(x_1+x_2)]$$
$$= (x_1 - x_2, x_2, x_1) + (x_2, x_1 - x_2, x_1 + x_2)$$
$$= (x_1, x_1, 2x_1 + x_2)$$

En particular, se tiene

$$(f+g)[(2,-3)] = (2,2,4-3) = (2,2,1)$$

2. Sean las transformaciones lineales

$$\mathbb{R}^3 \underset{g}{\overset{f}{\rightrightarrows}} \mathbb{R}^3$$

definidas por las matrices:

$$f \to A = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \\ 0 & 2 & 1 \end{pmatrix}$$

$$g \to B = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 2 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

Encontrar la matriz asociada a la transformación lineal $f + g$, y la imagen del vector $(2, 5, -1)$.

Tenemos:

a) $$f + g \to A + B = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \\ 0 & 2 & 1 \end{pmatrix} + \begin{pmatrix} 1 & - & 1 \\ 1 & 2 & 0 \\ 1 & 1 & 1 \end{pmatrix} \quad \begin{pmatrix} 2 & 1 & 0 \\ 2 & 2 & 1 \\ 1 & 3 & 2 \end{pmatrix}$$

b) $$(f+g)[(2,5,-1)] = (A + B) \cdot \begin{pmatrix} 2 \\ 5 \\ -1 \end{pmatrix}$$

$$= \begin{pmatrix} 2 & 1 & 0 \\ 2 & 2 & 1 \\ 1 & 3 & 2 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 5 \\ -1 \end{pmatrix} = \begin{pmatrix} 9 \\ 13 \\ 15 \end{pmatrix}$$

C) Dada una aplicación lineal $f : V(K) \to U(K)$ y un escalar $\rho \in K$, entonces la aplicación

$$\rho f : V(K) \to U(K)$$

definida por

$$(\rho f)(\vec{x}) = \rho \cdot f(\vec{x})$$

es una transformación lineal.

En efecto, se tienen

$$(\rho f)(\vec{x} + \vec{x}') = \rho \cdot f(\vec{x} + \vec{x}') = \rho\,[f(\vec{x}) + f(\vec{x}')]$$
$$= \rho \cdot f(\vec{x}) + \rho \cdot f(\vec{x}')$$
$$= (\rho f)(\vec{x}) + (\rho f)(\vec{x}')$$

y por otro lado,

$$(\rho f)(\alpha \vec{x}) = \rho \cdot f(\alpha \vec{x}) = \rho\,[\alpha f(\vec{x})]$$
$$= (\rho\alpha) f(\vec{x}) = (\alpha\rho) f(\vec{x})$$
$$= \alpha\,[\rho f(\vec{x})]$$
$$= \alpha\,[(\rho f)(\vec{x})]$$

resultados que muestran que la aplicación ρf es lineal, y la llamaremos *la aplicación lineal multiplicación escalar.*

Fácilmente se prueban las propiedades:

a) $\rho(f + g) = \rho f + \rho g$

b) $(\rho + \lambda) f = \rho f + \lambda f$

c) $(\rho\lambda) f = \rho(\lambda f)$

d) $\rho(f \circ g) = (\rho f) \circ g = f \circ (\rho g)$

e) $1 f = f$, siendo 1 la unidad del cuerpo K, que las dejamos como ejercicios al estudiante.

Las propiedades de la adición y de la multiplicación escalar dadas más arriba son precisamente las propiedades de un espacio vectorial. Por esto, tenemos el teorema siguiente:

Teorema: Si V(K) y U(K) son espacios vectoriales sobre un mismo cuerpo K, entonces el conjunto de todas las aplicaciones lineales de V(K) en U(K), denotado por T.L. (V,U), es también un espacio vectorial sobre el mismo cuerpo K de escalares.

Observamos finalmente que la multiplicación de un número por una matriz, corresponde a la multiplicación de un escalar por una transformación lineal, como es inmediato verificar.

D) Son consecuencias inmediatas de las definiciones de suma y producto de transformaciones lineales, las siguientes proposiciones:

Proposición 1: Sean

$f : V(K) \to U(K)$

$g : V(K) \to U(K)$

$h : U(K) \to W(K)$

aplicaciones lineales de los espacios vectoriales indicados.

Entonces se verifica la igualdad

$$h \circ (f + g) = h \circ f + h \circ g$$

Dem.: Debemos hacer ver que, para cada $\vec{x} \in V(K)$, se tienen

$$[h \circ (f + g)](\vec{x}) = [h \circ f + h \circ g](\vec{x}).$$

En efecto, recurriendo a las definiciones de suma y producto de aplicaciones lineales, se puede escribir

$$[h \circ (f + g)](\vec{x}) = h[(f + g)(\vec{x})]$$
$$[h \circ (f + g)](\vec{x}) = h[f(\vec{x}) + g(\vec{x})]$$

y como h es aplicación lineal, se tiene

$$[h \circ (f + g)](\vec{x}) = h(f(\vec{x})) + h(g(\vec{x}))$$
$$[h \circ (f + g)](\vec{x}) = (h \circ f)(\vec{x}) + (h \circ g)(\vec{x})$$
$$[h \circ (f + g)](\vec{x}) = [h \circ f + h \circ g](\vec{x})$$

o sea, $h \circ (f + g) = h \circ f + h \circ g, \quad \forall f, g, h$

Así hemos probado la distributividad a izquierda del producto de aplicaciones lineales con relación a la suma de aplicaciones lineales, si se cumple el siguiente esquema de aplicaciones:

$$V(K) \underset{g}{\overset{f}{\rightrightarrows}} U(K) \xrightarrow{h} W(K)$$

Proposición 2: Sean las siguientes aplicaciones lineales:

$h : W(K) \to V(K)$

$f : V(K) \to U(K)$

$g : V(K) \to U(K)$

Entonces se verifica la igualdad

$$(f + g) \circ h = f \circ h + g \circ h$$

Dem.: Debemos mostrar que para cada $\vec{z} \in W(K)$, se tiene

$$[(f + g) \circ h](\vec{z}) = [f \circ h + g \circ h](\vec{z})$$

En efecto, recurriendo a las definiciones de suma y producto de transformaciones lineales, se escribe

$$[(f + g) \circ h](\vec{z}) = (f + g)[h(\vec{z})]$$
$$= f(h(\vec{z})) + g(h(\vec{z})) = (f \circ h)(\vec{z}) + (g \circ h)(\vec{z}) =$$
$$= [f \circ h + g \circ h](z)$$

luego,

$$(f + g) \circ h = f \circ h + g \circ h, \quad \forall f, g, h.$$

Así hemos demostrado ahora la distributividad a derecha del producto de aplicaciones lineales con relación a la suma de aplicaciones lineales, si se cumple el siguiente esquema de aplicaciones:

$$W(K) \xrightarrow{h} V(K) \underset{f}{\overset{g}{\rightrightarrows}} U(K)$$

Proposición 3: La aplicación idéntica

$I : V(K) \to V(K)$

definida por $I(\vec{x}) = \vec{x}, \quad \forall \vec{x} \in V(K)$

es una transformación lineal del espacio $V(K)$ y, además se verifica la igualdad

$$I \circ f = f \circ I = f$$

para cada aplicación lineal de $V(K)$ en $V(K)$.

Dem.: Es inmediata. (¡Hágala como ejercicio!).

Nota: Puesto que espacios vectoriales diferentes tienen aplicaciones identidad diferentes, entonces, si queremos poner de relieve que I es la función identidad del espacio $V(K)$, la escribiremos en la forma: I_V. Nótese, además, que la aplicación identidad es el elemento neutro relativo a la operación de composición o de producto de aplicaciones; esto es

$$I_V \circ f = f, \quad f \circ I_U = f$$

cualquiera que sea la aplicación f en el conjunto T. L. (V, U) de todas las aplicaciones lineales del espacio V(K) en el espacio U(K).

E) En el caso especial cuando V(K) = U(K), el espacio vectorial TL. (V, U) se convierte en el espacio vectorial TL. (V) de todas las aplicaciones o transformaciones lineales de V(K) en V(K).

Tendremos así un conjunto de transformaciones lineales para las cuales están definidas una operación de adición, una de multiplicación y otra de multiplicación escalar, y las dos primeras de ellas con elemento neutro: el 0 para la adición y el I_V para la multiplicación. Si tenemos en cuenta las listas de propiedades o leyes que indicamos en los puntos A), B) y C), se desprende que todas ellas son también válidas para el conjunto T.L. (V), ahora en cuestión.

Por consiguiente, las transformaciones lineales de un espacio vectorial V(K) en sí mismo forman, por una parte, un espacio vectorial sobre el mismo cuerpo K y, por otra, forman un anillo con unidad. Se trata en general de un anillo no conmutativo, puesto que

$$g \circ f \neq f \circ g$$

Este anillo lo llamamos el *Anillo de los Endomorfismos* del espacio vectorial V(K).

F) *Ejercicios Variados:*

1) Dada la aplicación lineal $f : \mathbb{R}^2 \to \mathbb{R}^2$ definida por

$$f[(x,y)] = (3x - 4y, x + 5y)$$

Encontrar la matriz asociada a esta aplicación con respecto a la base canónica $\{e_1 = (1,0), \vec{e_2} = (0,1)\}$.

Sol.: Tenemos:

$$f(\vec{e_1}) = f[(1,0)] = (3,1) = 3\vec{e_1} + \vec{e_2}$$
$$f(\vec{e_2}) = f[(0,1)] = (-4,5) = -4\vec{e_1} + 5\vec{e_2}$$

luego, la matriz asociada es

$$A = \begin{pmatrix} 3 & -4 \\ 1 & 5 \end{pmatrix}$$

Obsérvese, que por la definición de la aplicación f, la imagen del vector (5, 2) es:

$$f[(5,2)] = (15 - 8, 5 + 10) = (7, 15)$$

y si utilizamos la matriz hallada, se encuentra

$$f[(5,2)] = \begin{pmatrix} 3 & -4 \\ 1 & 5 \end{pmatrix} \cdot \begin{pmatrix} 5 \\ 2 \end{pmatrix} = \begin{pmatrix} 7 \\ 15 \end{pmatrix}$$

igual al resultado anterior.

Ahora, si la base que usamos no es la canónica, sino otra, como la $\{\vec{a_1} = (3,2), \vec{a_2} = (1,5)\}$, entonces lo primero que haremos será hallar las componentes de un vector arbitrario $(a,b) \in \mathbb{R}^2$ con respecto a la base dada $\{\vec{a_1} = (3,2), \vec{a_2} = (1,5)\}$, y se tendrá

$$(a,b) = \alpha a_1 + \alpha a_2 = \alpha(3,2) + \beta(1,5) = (3\alpha + \beta, 2\alpha + 5\beta)$$
$$= (3\alpha + \beta, 2\alpha + 5\beta)$$

de donde el sistema lineal

$$\begin{aligned} 3\alpha + \beta &= a \\ 2\alpha + 5\beta &= b \end{aligned}$$

de donde se obtienen los valores

$$\alpha = \frac{5a - b}{13} \quad , \quad \beta = \frac{-2a + 3b}{13}$$

y por tanto, el vector (a, b) se expresa en la forma

$$(a,b) = \frac{5-13}{13}\vec{a_1} + \frac{-2+39}{13}\vec{a_2} \qquad (1)$$

Por otra parte, según la definición de la aplicación f, es decir,

$$f[(x,y)] = (3x - 4y, x + 5y)$$

tendremos para las imágenes de los vectores de la base, las expresiones

$$f(\vec{a_1}) = f[(3,2)] = (1,13) = \frac{5-13}{13}\vec{a_1} + \frac{-2+39}{13}a_2, =$$
$$\overset{(1)}{=} -\frac{8}{13}\vec{a_1} + \frac{37}{13}\vec{a_2}$$

$$f(a_2) = f[(1,5)] = (-17,26) = \frac{-85-26}{13}\vec{a_1} + \frac{-34+78}{13}\vec{a_2} =$$
$$= -\frac{111}{13}\vec{a_1} + \frac{44}{13}\vec{a_2} \qquad (1)$$

Luego, la matriz asociada será:

$$A = \begin{pmatrix} -\frac{8}{13} & -\frac{111}{13} \\ \frac{37}{13} & \frac{44}{13} \end{pmatrix} = \frac{1}{13}\begin{pmatrix} -8 & -111 \\ 37 & 44 \end{pmatrix}$$

2) Sean f y g dos aplicaciones lineales de $\mathbb{R}^2$ en $\mathbb{R}^2$ definidas por:

$$f[(x,y)] = (x+y,o),\ g[(x,y)] = (-y,x)$$

Encontrar las fórmulas que definan las aplicaciones:

$$f+g,\quad 5f-3g,\quad g\,o\,f,\quad f\,o\,g$$

Soluciones:

$$(f+g)[(x,y)] = f[(x,y)] + g[(x,y)] = (x+y,o)+(-y,x) = (x,x)$$

$$(5f-3g)[(x,y)] = 5f[(x,y)] - 3g[(x,y)] = 5(x+y,0) - 3(-y,x) = (5x+5y+3y,-3x) = (5x+8y,-3x)$$

$$(g\,o\,f)[(x,y)] = g(f[(x,y)]) = g(x+y,o) = (o,x+y)$$
$$(f\,o\,g)[(x,y)] = f(g[(x,y)]) = f(-y,x) = (x-y,o)$$

3) Sea la aplicación lineal $f: \mathbb{R}^3 \to \mathbb{R}^2$ definida por la matriz

$$f \to A = \begin{pmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{pmatrix}$$

Encontrar la representación matricial de la aplicación lineal f con respecto a las siguientes bases de $\mathbb{R}^3$ y de $\mathbb{R}^2$:

$$\{\vec{a_1} = (1,1,1),\ \vec{a_2} = (1,1,0),\ \vec{a_3} = (1,0,0)\}$$
$$\{\vec{b_1} = (1,3),\ \vec{b_2} = (2,5)\}$$

Sol.: Primeramente hallaremos las componentes de un vector arbitrario $(a,b) \in \mathbb{R}^2$ con respecto a la base $\{\vec{b_1}, \vec{b_2}\}$.

Tenemos.

$$(a,b) = \alpha(1,3) + \beta(2,5) = (\alpha + 2\beta, 3\alpha + 5\beta)$$

de donde, el sistema

$$\begin{cases} \alpha + 2\beta = a \\ 3\alpha + 5\beta = b \end{cases}$$
$$\begin{cases} \alpha = 2b - 5a \\ \beta = 3a - b \end{cases}$$

luego,

$$(a,b) = (2b - 5a)\vec{b_1} + (3a - b)\vec{b_2}.$$

Enseguida hallaremos los transformados de los vectores de la base $\{\vec{a_1}, \vec{a_2}, \vec{a_3}\}$ de $\mathbb{R}^3$ mediante la matriz A como la aplicación lineal de $\mathbb{R}^3$ en $\mathbb{R}^2$, y se tiene

$$f(\vec{a_1}) = A \cdot \vec{a_1} = \begin{pmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix}$$

$$f(\vec{a_2}) = A \cdot \vec{a_2} = \begin{pmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 7 \\ -3 \end{pmatrix}$$

$$f(\vec{a_3}) = A \cdot \vec{a_3} = \begin{pmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Pero, cualquier vector $(a,b) \in \mathbb{R}^2$ es de la forma

$$(a,b) = (2b - 5a)\vec{b_1} + (3a - b)\vec{b_2}$$

por lo tanto, tendremos:

$$f(\vec{a_1}) = \begin{pmatrix} 4 \\ 4 \end{pmatrix} = -12\vec{b_1} + 8\vec{b_2}$$

$$f(\vec{a_2}) = \begin{pmatrix} 7 \\ -3 \end{pmatrix} = -41\,b_1 + 24\vec{b_2}$$

$$f(\vec{a_3}) = \begin{pmatrix} 2 \\ 1 \end{pmatrix} = -8\,\vec{b_1} + 5\,\vec{b_2}$$

Luego, la matriz asociada a la aplicación lineal

$$f : \mathbb{R}^3 \to \mathbb{R}^2$$

con respecto a las bases $\{\vec{a_1}, \vec{a_2}, \vec{a_3}\}$ de $\mathbb{R}^3$ y $\{\vec{b_1}, \vec{b_2}\}$ de $\mathbb{R}^2$, es

$$B = \begin{pmatrix} -12 & -41 & -8 \\ 8 & 24 & 5 \end{pmatrix}$$

10.25. *Cambio de Base en Espacios Vectoriales*

En el estudio de las secciones precedentes se ha visto que, dado un espacio vectorial, hay muchas posibilidades de obtener bases para él.

También se ha visto que, dado un vector $\vec{\mu}$ de un espacio vectorial V(K) y fijada una base ordenada $\{\vec{a_1}, \vec{a_2}, \vec{a_n}\}$ de V(K), dicho vector queda completamente determinado por el conocimiento del conjunto ordenado de sus coordenadas $(\alpha_1, \alpha_2, \ldots, \alpha_n)$.

$$\vec{\mu} = \alpha_1 \vec{a_1} + \alpha_2 \vec{a_2} + \ldots + \alpha_n \vec{a_n}.$$

El conjunto ordenado $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ de escalares recibe el nombre de *conjunto de coordenadas* de $\vec{\mu}$ respecto a la *base ordenada* $\{\vec{a_1}, \vec{a_2}, \ldots, \vec{a_3}\}$.

Finalmente, también se vio que, dada una transformación lineal

$$f : V(K) \to U(K),$$

y dada una base $\{\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}\}$ de V(K) y una base $\{\vec{b_1}, \vec{b_2}, \ldots, \vec{b_m}\}$ de U(K), entonces se deduce una matriz

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

que caracteriza la transformación lineal f de modo que:

$$\begin{cases} f(\vec{a_1}) = a_{11}\vec{b_1} + a_{21}\vec{b_2} + \ldots + a_{m1}\vec{b_m} \\ f(\vec{a_2}) = a_{12}\vec{b_1} + a_{22}\vec{b_2} + \ldots + a_{m2}\vec{b_m} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ f(\vec{a_n}) = a_{1n}\vec{b_1} + a_{2n}\vec{b_2} + \ldots + a_{mn}\vec{b_m} \end{cases}$$

Nótese que la matriz A se construye escribiendo como primera columna las coordenadas del vector transformado $f(\vec{a_1})$ en la base

$\{\vec{b_1}, \vec{b_2}, \ldots, \vec{b_m}\}$,

como segunda columna las coordenadas de $f(\vec{a_2})$ en la base

$\{\vec{b_1}, \vec{b_2}, \ldots, \vec{b_m}\}$, etc.

En particular, sólo consideraremos un espacio vectorial V(K) y una base ordenada

$$\{\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}\};$$

entonces, si

$$\{\vec{b_1}, \vec{b_2}, \ldots, \vec{b_n}\}$$

es una nueva base, por ser estos últimos vectores de V(K), podemos escribir cada uno de ellos en función de $\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}$, es decir, sean

$$\begin{cases} \vec{b_1} = a_{11}\vec{a_1} + a_{21}\vec{a_2} + \ldots + a_{n1}\vec{a_n} \\ \vec{b_2} = a_{12}\vec{a_1} + a_{22}\vec{a_2} + \ldots + a_{n2}\vec{a_n} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \vec{b_n} = a_{1n}\vec{a_1} + a_{2n}\vec{a_2} + \ldots + a_{nn}\vec{a_n} \end{cases}$$

y la matriz A que define la transformación lineal

$$F : V(K) \to V(K)$$

es ahora la matriz cuadrada de orden n, que sigue

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

Luego, la aplicación lineal $f : V(K) \to V(K)$, representada por la matriz cuadrada A en la base $\{\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}\}$ está determinada por $A\vec{a_1} = \vec{b_1}$, $A\vec{a_2} = \vec{b_2}, \ldots, A\vec{a_n} = \vec{b_n}$

Ahora bien, como el sistema $\{\vec{b_1}, \vec{b_2}, \ldots, \vec{b_n}\}$ es también una base del espacio V(K), debe existir entonces una transformación B (es otra matriz cuadrada de orden n) definida por

$$B\vec{b_1} = \vec{a_1}, \quad B\vec{b_2} = \vec{a_2}, \ldots, B\vec{b_n} = \vec{a_n}$$

Luego, resulta que

$$BA\,\vec{a_1} = \vec{a_1}, \quad BA\,\vec{a_2} = \vec{a_2}, \ldots, BA\,\vec{a_n} = \vec{a_n}$$

y, por tanto, la aplicación lineal BA es la identidad, es decir, la matriz unidad I_n :

$$I_n = \begin{pmatrix} 1 & 0 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & \ldots & 1 \end{pmatrix}$$

En resumen: supongamos que $\{\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}\}$ es una base de V(K) y que $\{\vec{b_1}, \vec{b_2}, \ldots, \vec{b_n}\}$ es otra. Entonces, de la definición de base se desprende que los $\vec{b_i}$ pueden ser expresados como combinaciones lineales de los a_i, y viceversa:

$$(*)\quad \begin{cases} \vec{b_1} = a_{11}\vec{a_1} + a_{21}\vec{a_2} + \ldots + a_{n1}\vec{a_n} \\ \vec{b_2} = a_{12}\vec{a_1} + a_{22}\vec{a_2} + \ldots + a_{n2}\vec{a_n} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \vec{b_n} = a_{1n}\vec{a_1} + a_{2n}\vec{a_2} + \ldots + a_{nn}\vec{a_n} \end{cases}$$

y

$$(**)\quad \begin{cases} \vec{a_1} = b_{11}\vec{b_1} + b_{21}\vec{b_2} + \ldots + b_{n1}\vec{b_n} \\ \vec{a_2} = b_{12}\vec{b_1} + b_{22}\vec{b_2} + \ldots + b_{n2}\vec{b_n} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \vec{a_n} = b_{1n}\vec{b_1} + b_{2n}\vec{b_2} + \ldots + b_{nn}\vec{b_n} \end{cases}$$

Simbolizando por $A = [a_{ij}]_{(n,n)}$ la matriz cuadrada de orden n, formada por los coeficientes a_{ij} de las ecuaciones (*), diremos que A es *la matriz del cambio de base de*

$$\{\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}\} \text{ a } \{\vec{b_1}, \ldots, \vec{b_n}\}.$$

Análogamente, $B = [b_{ij}]_{(n,n)}$ es la matriz del cambio de base de

$$\{\vec{b_1}, \vec{b_2}, \ldots, \vec{b_n}\} \text{ a } \{\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}\},$$

dada esta última por los coeficientes de las ecuaciones (**).

Además, entre ambas matrices se tiene para el producto de ellas la relación,

$$BA = I_n$$

Diremos que en tales casos, la matriz B que satisface la ecuación

$$BA = I_n$$

es la recíproca o la inversa de la matriz A, y se la denota por A^{-1}.

En el Capítulo XVII del Tomo III se estudiarán las condiciones para que una matriz cuadrada A tenga reciproca, y cumplida ella, se determinará la recíproca A^{-1}.

Ahora bien, si la matriz cuadrada A tiene recíproca A^{-1}, entonces ambas satisfacen la relación

$$A^{-1}A = AA^{-1} = I.$$

En efecto, sea

$$A^{-1}A = I$$

y supongamos que A' es una matriz que cumple la condición

$$AA' = I$$

Probaremos que $A' = A^{-1}$

Pues bien, multiplicando la igualdad $AA' = I$ a la izquierda por A^{-1}, resulta

$$A^{-1}(AA') = A^{-1}I$$

o sea,

$$(A^{-1}A)A' = A^{-1}$$
$$IA' = A^{-1}$$
$$A' = A^{-1}$$

Así hemos demostrado que cuando una matriz cuadrada A admite recíproca A^{-1}, esta última verifica la identidad

$$A^{-1}A = AA^{-1} = I$$

Daremos en seguida, para nuestros fines, un método sencillo y de uso frecuente en la operatoria, para hallar la recíproca de una matriz cuadrada, cuando ella existe.

Sea la matriz cuadrada

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

Escribamos el siguiente cuadro

$$\left(\begin{array}{ccc|ccc} 1 & 4 & 3 & 1 & 0 & 0 \\ 2 & 5 & 4 & 0 & 1 & 0 \\ 1 & -3 & -2 & 0 & 0 & 1 \end{array}\right)$$

Efectuando transformaciones elementales de filas, obtendremos las siguientes series de cuadros:

$$\left(\begin{array}{ccc|ccc} 1 & 4 & 3 & 1 & 0 & 0 \\ 0 & -3 & -2 & -2 & 1 & 0 \\ 0 & -7 & -5 & -1 & 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & 4 & 3 & 1 & 0 & 0 \\ 0 & 1 & \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} & 0 \\ 0 & -7 & -5 & -1 & 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & \frac{1}{3} & -\frac{5}{3} & \frac{4}{3} & 0 \\ 0 & 1 & \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} & 0 \\ 0 & 0 & -\frac{1}{3} & \frac{11}{3} & -\frac{7}{3} & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & \frac{1}{3} & -\frac{5}{3} & \frac{4}{3} & 0 \\ 0 & 1 & \frac{2}{3} & \frac{2}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 1 & -11 & 7 & -3 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{6}{3} & -\frac{3}{3} & 1 \\ 0 & 1 & 0 & \frac{24}{3} & -\frac{15}{3} & 2 \\ 0 & 0 & 1 & -11 & 7 & -3 \end{array}\right)$$

o sea,

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & -1 & 1 \\ 0 & 1 & 0 & 8 & -5 & 2 \\ 0 & 0 & 1 & -11 & 7 & -3 \end{array}\right)$$

La matriz inversa o recíproca de la dada A es entonces

$$A^{-1} = \begin{pmatrix} 2 & -1 & 1 \\ 8 & -5 & 2 \\ -11 & 7 & -3 \end{pmatrix}$$

(Ver demostración en Capítulo XVII, Tomo III).

Veamos otro ejemplo. Sea ahora la matriz

$$A = \begin{pmatrix} 4 & 0 & 8 \\ 0 & 1 & -6 \\ 2 & 0 & 4 \end{pmatrix}$$

Escribiendo como antes el cuadro inicial

$$\left(\begin{array}{ccc|ccc} 4 & 0 & 8 & 1 & 0 & 0 \\ 0 & 1 & -6 & 0 & 1 & 0 \\ 2 & 0 & 4 & 0 & 0 & 1 \end{array}\right)$$

y efectuando las transformaciones elementales de filas que se hicieron en el ejemplo anterior, obtenemos el cuadro que sigue:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 2 & \frac{1}{4} & 0 & 0 \\ 0 & 1 & -6 & 0 & 1 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} & 0 & 1 \end{array}\right)$$

Es claro que no podemos seguir operando, puesto que hay una fila de ceros a la izquierda de la línea vertical del cuadro. Cuando esto ocurre, decimos que la matriz considerada A no tiene recíproca.

En resumen: una matriz cuadrada A tiene recíproca A^{-1} sí, y sólo sí, el cuadro

$$(A \mid I)$$

puede cambiarse por transformaciones elementales de filas en el cuadro

$$(I \mid A^{-1})$$

Sentado esto, los sistemas (*) y (**) pueden escribirse en forma matricial del modo siguiente:

(1) $(\vec{b_1}, \vec{b_2}, \ldots, \vec{b_n}) = (\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}) \cdot A$

(2) $(\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}) = (\vec{b_1}, \vec{b_2}, \ldots, \vec{b_n}) \cdot A^{-1}$

donde A es la matriz de paso de la base $\{\vec{a_1}, \ldots, \vec{a_n}\}$ a la base $\{\vec{b_1}, \ldots, \vec{b_n}\}$, y A^{-1} es la matriz de paso de la base $\{\vec{b_1}, \ldots, \vec{b_n}\}$ a la base $\{\vec{a_1}, \ldots, \vec{a_n}\}$, siendo $\{\vec{a_1}, \ldots, \vec{a_n}\}$ y $\{\vec{b_1}, \ldots, \vec{b_n}\}$ dos bases cualesquiera del espacio vectorial V(K) de dimension n.

Sea ahora, $\vec{\mu} = \alpha_1 \vec{a_1} + \alpha_2 \vec{a_2} + \alpha_n \vec{a_n}$, un vector cualquiera del espacio V(K), expresado en la base $\{\vec{a_1}, \ldots, \vec{a_n}\}$. Nos interesa ahora su expresión en la nueva base $\{\vec{b_1}, \ldots, \vec{b_n}\}$. Sea

$$\vec{\mu} = \beta_1 \vec{b_1} + \beta_2 \vec{b_2} + \ldots + \beta_n \vec{b_n}$$

su expresión en esta base.

Probaremos que el conjunto de coordenadas $(\beta_1, \beta_2, \ldots, \beta_n)$ viene definido por

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \beta_n \end{pmatrix} = A^{-1} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_n \end{pmatrix}$$

En efecto, se puede escribir

$$\vec{\mu} = \alpha_1 \vec{a_1} + \alpha_2 \vec{a_2} + \ldots + \alpha_n \vec{a_n} = (\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}) \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_n \end{pmatrix}$$

y por la relación (2), resulta:

(3) $\vec{\mu} = (b_1, \vec{b_2}, \ldots, b_n) A^{-1} \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_n \end{pmatrix}$

Por otra parte, se tiene

(4) $\vec{\mu} = \beta_1 \vec{b_1} + \beta_2 \vec{b_2} + \ldots + \beta_n \vec{b_n} = (\vec{b_1}, \vec{b_2}, \ldots, \vec{b_n}) \cdot \begin{pmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \beta_n \end{pmatrix}$

De (3) y (4) se concluye en virtud de la definición de igualdad de vectores que,

(5) $$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \beta_n \end{pmatrix} = A^{-1} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_n \end{pmatrix}$$

lo que prueba nuestra aseveración.

Si multiplicamos (5) por A a la izquierda, resulta

(6) $$A \cdot \begin{pmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \beta_n \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_n \end{pmatrix}$$

Las fórmulas (5) y (6) permiten efectuar cambios de coordenadas al pasar de una base a otra.

En consecuencia, conociendo la matriz de paso o de cambio de base, podemos expresar las coordenadas primitivas de cualquier vector $\vec{\mu} \in V(K)$ en función de las nuevas coordenadas, y recíprocamente.

Ejemplos:

Se consideran las siguientes bases de $\mathbb{R}^3$:

$A = \{\vec{a_1} = (1, 1, -1),\ \vec{a_2} = (1, -1, o),\ \vec{a_3} = (o, 1, 1)\}$

$B = \{\vec{b_1} = (1, 2, 1),\ \vec{b_2} = (2, o, -1),\ \vec{b_3} = (-2, 3, 2)\}$

Hallar las matrices de paso o de cambio de base del sistema A al sistema B y recíprocamente.

Sol.: Un vector cualquiera $\vec{\mu} \in \mathbb{R}^3$ se expresa, respecto a la base A, en la forma:

$\vec{\mu} = x\vec{a_1} + y\vec{a_2} + z\vec{a_3} = x(1, 1, -1) + y(1, -1, o) + z(o, 1, 2)$

$\vec{\mu} = (x + y,\ x - y + z,\ -x + 2z)$

En particular, para los vectores de la base B, se tiene

$\vec{b_1}$	$\vec{b_2}$	$\vec{b_3}$
$x + y = 1$	$= 2$	$= -2$
$x - y + z = 2$	$= o$	$= 3$
$-x + 2z = 1$	$= -1$	$= 2$

esquema que representa un conjunto de tres sistemas lineales que se diferencian únicamente en sus segundos miembros.

Ellos tiene respectivamente las soluciones siguientes:

$$\begin{cases} x = 1 \\ y = o \\ z = 1 \end{cases} ; \quad \begin{cases} x = 1 \\ y = 1 \\ z = o \end{cases} ; \quad \begin{cases} x = o \\ y = -2 \\ z = 1 \end{cases}$$

Luego, las siguientes combinaciones lineales

$\vec{b_1} = 1 \cdot \vec{a_1} + o \cdot \vec{a_2} + 1 \cdot \vec{a_3}$

$\vec{b_2} = 1 \cdot \vec{a_1} + 1 \cdot \vec{a_2} + o \cdot \vec{a_3}$

$\vec{b_3} = o \cdot \vec{a_1} + (-2) \cdot \vec{a_2} + 1 \cdot \vec{a_3}$

Por consiguiente, la matriz del cambio de base de

$A = \{\vec{a_1}, \vec{a_2}, \vec{a_3}\}$ a $B = \{\vec{b_1}, \vec{b_2}, \vec{b_3}\}$ es

$$M = \begin{pmatrix} 1 & 1 & o \\ o & 1 & -2 \\ 1 & o & 1 \end{pmatrix}$$

Veamos ahora si esta matriz tiene recíproca. Tenemos

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & o & o \\ o & 1 & -2 & o & 1 & o \\ 1 & o & 1 & o & o & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & 1 & o & 1 & o & o \\ o & 1 & -2 & o & 1 & o \\ o & -1 & 1 & -1 & o & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & o & 2 & 1 & -1 & o \\ o & 1 & -2 & o & 1 & o \\ o & o & -1 & -1 & 1 & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & o & 2 & 1 & -1 & o \\ o & 1 & -2 & o & 1 & o \\ o & o & 1 & 1 & -1 & -1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & o & o & -1 & 1 & 2 \\ o & 1 & o & 2 & -1 & -2 \\ o & o & 1 & 1 & -1 & -1 \end{array}\right)$$

Como existe la inversa, entonces la matriz de paso de la base $B = \{\vec{b_1}, \vec{b_2}, \vec{b_3}\}$ a la base $A = \{\vec{a_1}, \vec{a_2}, \vec{a_3}\}$ es

$$M^{-1} = \begin{pmatrix} -1 & 1 & 2 \\ 2 & -1 & -2 \\ 1 & -1 & -1 \end{pmatrix}$$

Luego, las siguientes combinaciones lineales:

$\vec{a_1} = (-1)\vec{b_1} + 2\vec{b_2} + 1\vec{b_3}$

$\vec{a_2} = 1 \cdot \vec{b_1} + (-1)\vec{b_2} + (-1)\vec{b_3}$

$\vec{a_3} = 2 \cdot \vec{b_1} + (-2)\vec{b_2} + (-1)\vec{b_3}$

Sea ahora $\vec{v} = (3,2,5) \in \mathbb{R}^3$; entonces sus coordenadas en la base $A = \{\vec{a_1}, \vec{a_2}, \vec{a_3}\}$ se obtienen como sigue

$(3, 2, 5) = x(1, 1, -1) + y(1, -1, o) + z(o, 1, 2) =$

$= (x + y,\ x - y + z,\ -x + 2z)$

de donde el sistema lineal

$$\begin{array}{r} x + y = 3 \\ x - y + z = 2 \\ -x + 2z = 5 \end{array}$$

que tiene por solución, $x = 1,\ y = 2,\ z = 3,$; luego

$\vec{v} = (3, 2, 5) = 1 \cdot \vec{a_1} + 2 \cdot \vec{a_2} + 3\vec{a_3}$

Por lo tanto, las nuevas coordenadas del vector $\vec{v}$ en la nueva base $B = \{\vec{b_1}, \vec{b_2}, \vec{b_3}\}$ se obtienen por la fórmula (5); es decir:

$$\begin{pmatrix}\beta_1\\ \beta_2\\ \beta_3\end{pmatrix} = M^{-1} \cdot \begin{pmatrix}x\\ y\\ z\end{pmatrix}$$

$$= \begin{pmatrix}-1 & 1 & 2\\ 2 & -1 & -2\\ 1 & -1 & -1\end{pmatrix}\begin{pmatrix}1\\ 2\\ 3\end{pmatrix} = \begin{pmatrix}7\\ -6\\ -4\end{pmatrix}$$

y de donde resulta para las nuevas coordenadas de $\vec{v}$ los valores:

$\beta_1 = 7, \quad \beta_2 = -6, \quad \beta_3 = -4$

Por consiguiente, el vector $\vec{v}$ expresado en la base B tiene por expresión:

$\vec{v} = 7\vec{b_1} + (-6)\vec{b_2} + (-4)\vec{b_3}$

(FIN DEL SEGUNDO TOMO)

BIBLIOGRAFIA

En la confección del presente volumen, como en el primero, se utilizó en gran parte la siguiente fuente de información:

0.1. LOGICA

1. *Lógica Matemática.* José Ferrater Mora y H. Leblanc.
2. *Introducción a la Lógica y a la Metodología de las Ciencias Deductivas.* Alfredo Tarski.
3. *Introducción a la Lógica Simbólica.* Gerold Stahl.
4. *Introducción a la Epistemología y Fundamentación de la Matemática.* Fausto J. Toranzos.
5. *Introducción a la Filosofía Matemática.* Bertrand Russell.
6. *Introducción a la Lógica Matemática.* P. Suppes y S. Hill.
7. *Lecciones de Lógica y Teoría del Conocimiento.* Gregorio Fingermann.
8. *Lógica.* D. P. Gorski y P. V. Tavants.
9. *El sentido de la nueva lógica.* W. Quine.
10. *Fundamentos de Matemáticas Universitarias.* Allendoerfer y Oakley.

0.2. ALGEBRA

1. *Matemática Elemental Moderna* (estructura y método). César A. Trejo.
2. *Matemática Moderna.* César A. Trejo y J. Bosch.
3. *Introducción al Algebra.* Mischa Cotlar y Cora Ratto de Sadosky.
4. *Lecciones de Algebra.* J. Rey Pastor.
5. *Elementos de Análisis Algebraico.* J. Rey Pastor.
6. *Algebra Superior.* H. S. Hall y S. R. Knight.
7. *Elementos de Algebra Superior.* P. Miquel y Merino.
8. *Set Theory and Related Topics.* S. Lipschutz.
9. *Matemáticas e Imaginación.* E. Kasner y J. Newman.
10. *Les nombres et les Espaces.* G. Verniest.
11. *Introducción a la Matemática Superior.* Alberto Sagasume Berra.
12. *Introducción à la Teoría de Conjuntos.* Lía Oubiña.
13. *Introducción a la Teoría de Conjuntos y a la Topología.* Kazimierz Kuratowski.
14. *Teoría Intuitiva de los Conjuntos.* P. R. Halmos.

15. *Theorie des Ensembles.* N. Bourbaki.
16. *Algebra Superior.* A. Adrian Albert.
17. *Introduction to Algebraic Theories.* A. Adrian Albert.
18. *Algebra.* Roger Godement.
19. *Algebra Moderna.* Garret Birkhoff y Saunders Mac Lane.
20. *Algebra Moderna.* M. Queysanne et A. Delachit.
21. *Estructuras Algebraicas* (Departamento de Asuntos Científicos de la Unión Panamericana). Enzo R. Gentile.
22. *An Introduction to Abstract Algebra.* Cyrus Colton Mac Duffee.
23. *First Course in Abstract Algebra.* Richard E. Johnson.
24. *Lecons d'Algebre Moderne.* A. Lentin et J. Rivaud.
25. *Lecciones de Algebra Moderna.* Alberto E. Sagastume Berra.
26. *Algebre* (Tomo 1). Paul Dubreil.
27. *Introducción al Algebra y Análisis Moderno.* Marc Zamansky.
28. *Introducción a la teoría de los Grupos.* P. S. Alexandroff.
29. *Grupos Finitos.* W. Ledermann.
30. *Group Theory.* B. Baumslag and B. Chandler.
31. *Rings and Ideals.* Neal H. McCoy.
32. *Modern Algebra.* Frank Ayres.
33. *Abstract Algebra.* Ivong Fang.
34. *Algebre.* N. Bourbaki.
35. *Elementos de Algebra Lineal.* Lowell J. Paige y J. Dean Swift.
36. *Introducción al Estudio de Matrices y Vectores.* Jacob T. Schwartz.
37. *Elementary Matrix Algebra.* Franz E. Hohn.
38. *Introduction to Modern Algebra and Matrix Theory.* Ross A. Beaumont.
39. *Lineal Algebra.* Seymour Lipschutz.
40. *Determinantes y Matrices.* A. C. Aitken.
41. *Teoría de Ecuaciones.* J. V. Uspensky.