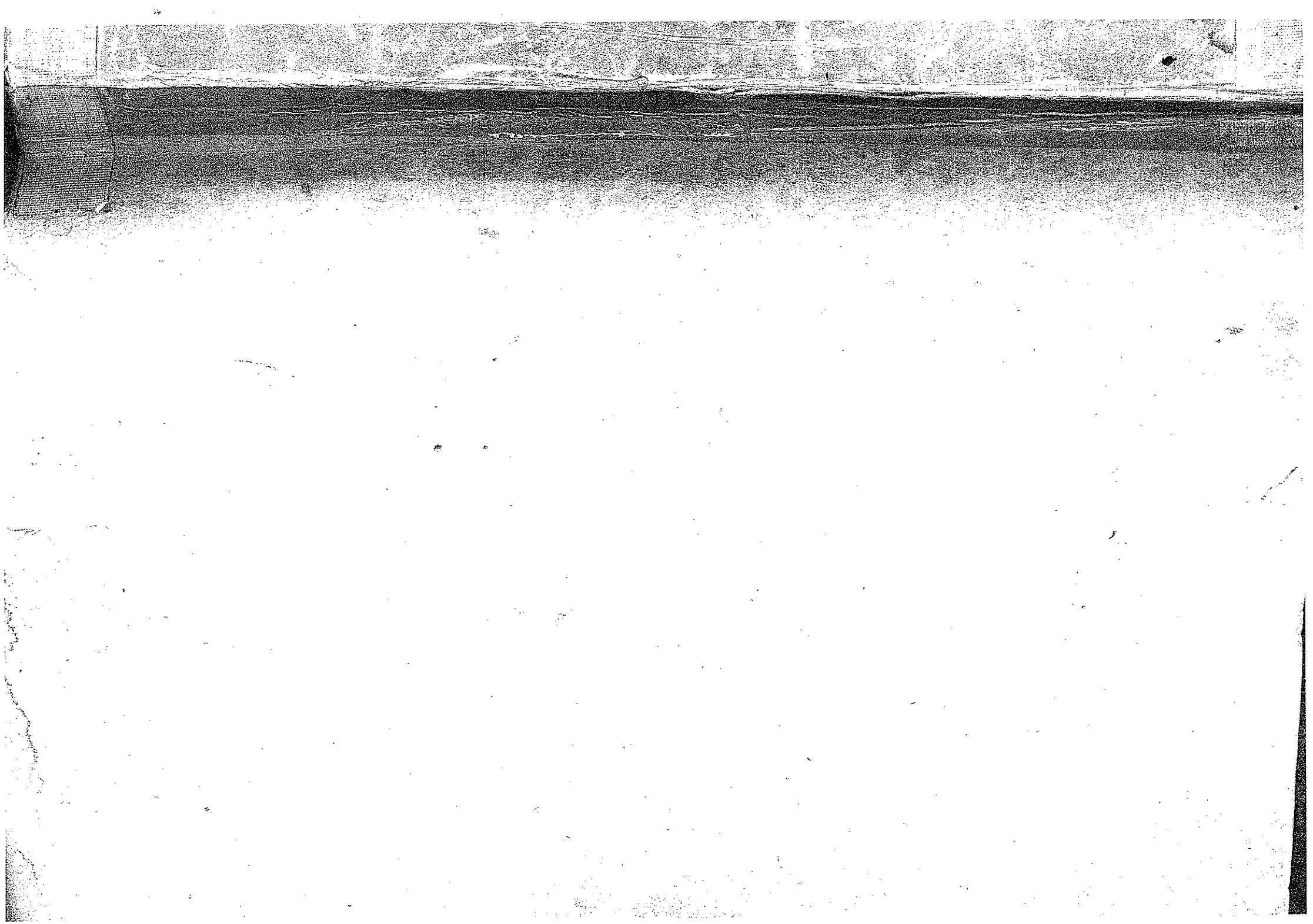
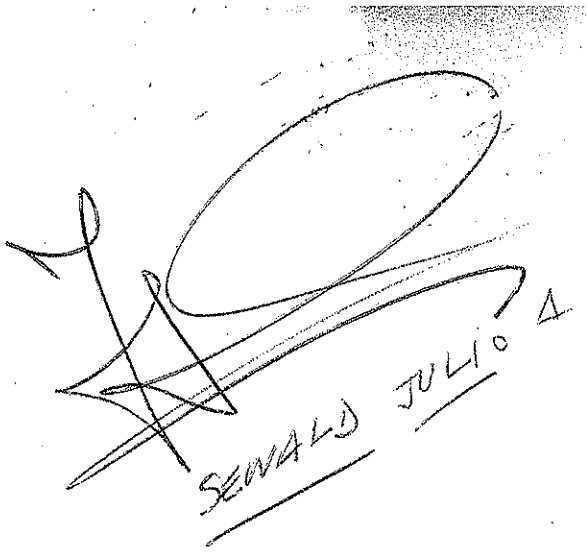
NOCIONES DE ALGEBRA

MARIA LUISA GASTAMINZA

Cooperadora de la Universidad Nacional del Sur





NOCIONES DE ALGEBRA

MARIA LUISA GASTAMINZA

Cooperadora de la Universidad Nacional del Sur

INDICE GENERAL

Estas notas tienen por objeto ofrecer una exposición ordenada y breve de los temas incluidos en el programa del primer curso de Algebra.

El alumno de primer año tiene en general dificultades para adaptarse al regimen cuatrimestral de estudios que exige un ritmo rápido y un intenso aprovechamiento del tiempo; le cuesta tomar apuntes completos y no está habituado todavía a consultar bibliografía. Este hecho , comprobado durante una larga actuación en la cátedra de Algebra I , nos ha llevado a redactar estas notas que pretenden ser una guía para el estudiante flamante.

Se ha procurado desarrollar los distintos temas en forma concisa a la par que clara, consignando el mínimo de resultados exigibles en cada caso.

Pero se recomienda decididamente al lector consultar los libros men - cionados en la bibliografía para complementar estas notas y tener un panorama más amplio y rico de los tópicos aquí tratados. Dicha biblio grafía la hemos confeccionado con algunos textos en castellano, cuya lectura no ofrece dificultades. Por supuesto que no pretendemos haber agotado con ella la lista de obras que pueden ser consultadas.

Agradecemos a la señorita Hilda Candel el excelente dactilografiado de estas páginas, y a la licenciada Olga Rueda su valiosa colaboración en la revisión de los originales.

Bahía Blanca, junio de 1970

CAPITULO I

1.1. NOTACIONES LOGICAS. Indicaremos rápidamente algunas notaciones lógicas que se usan en matemática.

Una proposición es una sentencia del lenguaje a la cual le corresponde uno y solo uno de los siguientes valores: verdad o falsedad. Una proposición P es verdadera o falsa pero no las dos cosas a la vez.

Implicación. Se dice que una proposición P implica o tiene por consecuencia otra proposición Q si toda vez que P es verdadera Q también lo es.

Simbólicamente se escribe $P \Longrightarrow Q$ y se lee de varias maneras diferentes: "P implica Q", "si P entonces Q", "P es condición suficiente para que Q", "Q es condición necesaria para que P".

Si además Q implica P, las proposiciones P y Q se dicen lógicamente equivalentes y se escribe $P \Longleftrightarrow Q$, pudiéndose reemplazaruna de ellas por la otra en cualquier razonamiento.

Si $P \Longleftrightarrow Q$ se dice que "P es condición necesaria y suficiente para que Q" o también "P si y solo si Q".

EJEMPLO. Sean P: El sol brilla en un cielo sin nubes.
Q: No está lloviendo.

Entonces $P \Longrightarrow Q$ pues si P es verdadera también lo es Q. En cambio Q no implica P pues puede ser Q verdadera y P falsa.

La implicación es transitiva, es decir

si $A \Longrightarrow B$ y $B \Longrightarrow C$ entonces $A \Longrightarrow C$.

Cuantificadores. Ciertas expresiones se representan por símbolos especiales. Así se escribe:

$\forall$ y se lee "Cualquiera que sea".
$\exists$ y se lee "Existe".

Negación. Dada una proposición P, la proposición "no P" se llama la negación de P. Por ejemplo, si P es la proposición: "2 es un múltiplo de 7", su negación es: "2 no es un múltiplo de 7". "no P" es verdadera si y solo si P es falsa.

Habitualmente, cuando una proposición P se representa por un signo, para representar a la negación de P se atraviesa el signo que la simboliza con un trazo. Por ejemplo, "x no es igual a y" se escribe $x \neq y$; "P no implica Q" se escribe $P \not\Longrightarrow Q$.

Conjunción, disyunción. En el lenguaje corriente el significado de la proposición "P y Q" es perfectamente claro: "P y Q" significa las dos cosas, P y Q al mismo tiempo. En términos más precisos, "P y Q" es verdadera si y solo si P y Q son ambas verdaderas. En cambio no sucede lo mismo con la disyunción pues la expresión "P o Q" puede tener dos significados: que se da P o Q y que una posibilidad excluye a la otra; o que se dan P o Q, o las dos a la vez. Por ejemplo, cuando se dice: "Esta tarde iré al cine o me queda

ré en casa estudiando" cada una de las posibilidades excluye a la otra. En cambio si se dice: "En la ciudad X siempre llueve o hace frío", se quiere significar que llueve, hace frío o las dos cosas a la vez.

En matemática la disyunción se utiliza siempre en el segundo sentido: "P o Q" significa P o Q o las dos a la vez. Más correctamente, la proposición "P o Q" es verdadera si y solo si por lo menos una de las dos, P o Q, es verdadera.

1.2. ALGEBRA DE CONJUNTOS. Consideraremos como nociones primitivas (no definidas) las de conjunto, objeto (o elemento), pertenencia e igualdad.

Intuitivamente un conjunto es una colección de objetos. Así se dice: "el conjunto de las páginas de un libro", "el conjunto de los puntos de una recta","el conjunto de las letras del alfabeto", etc. Usaremos los términos conjunto y colección como sinónimos.

Los elementos de un conjunto son los objetos que lo forman. Páginas, puntos y letras son respectivamente los elementos de los conjuntos citados en el ejemplo anterior.

En este curso notaremos los conjuntos con letras mayúsculas: A,B,C,D,..... y los elementos con letras minúsculas: a,b,c,d,.....

Para indicar que x es elemento del conjunto A se escribe $x \in A$ y se lee "x pertenece a A" o "A contiene a x".

La negación de $x \in A$ se escribe $x \notin A$ y se lee: "x no pertenece a A".

Por ejemplo, si Q es el conjunto de los números racionales, dados los números 1 , π , -1.85 , $\sqrt{2}$, $\ln 3$ se tiene: $1 \in Q$, $\pi \notin Q$, $-1.85 \in Q$, $\sqrt{2} \notin Q$, $\ln 3 \notin Q$.

Un conjunto está definido o determinado cuando se puede decidir si un objeto cualquiera pertenece o no al conjunto.

Una manera de definir un conjunto es enumerar todos sus elementos. Por ejemplo, si el conjunto A está formado por los elementos a,b,c escribiremos:

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

Cualquier otro objeto diferente de los que figuran en la llave no pertenece al conjunto.

Pero no siempre es posible dar un conjunto de esta manera, por ejemplo si el conjunto no es finito.

La forma general de definir un conjunto es indicar una condición o propiedad que verifican los elementos de ese conjunto y ningún otro objeto.

La notación

$$A = \{x : \underset{\text{(proposición sobre x)}}{.....}\}.$$

representa al conjunto A formado por todos los x para los cuales la proposición en cuestión es verdadera.

Notación: Las letras N , Z , Q , R y C representarán los conjuntos de los números naturales, enteros, racionales, reales y complejos respectivamente, salvo expresa aclaración.

EJEMPLOS:

1) $A = \{x : x \in Z \text{ y } x^2 = 1\}$ es el conjunto formado por los números -1 y 1. Así $A = \{-1,1\}$.

2) La notación $B = \{x : x \in Q \text{ y } 0 < x < 1/2\}$ quiere decir que B es el conjunto de todos los números racionales positivos menores que 1/2. Así por ejemplo, $0 \notin B$, $1/3 \in B$, $1/2 \notin B$.

3) El conjunto $M = \{y : y=5\}$ tiene un solo elemento: 5. Se puede escribir también $M = \{5\}$.

4) Los elementos del conjunto $S = \{x : x \in N \text{ y } x \text{ divide a } 30\}$ son los números 1,2,3,5, 6,10,15 y 30.

5) Cualquiera sea el conjunto A, se tiene:

$$A = \{x : x \in A\} .$$

Relación de igualdad. "x = y" significa intuitivamente que x e y simbolizan el mismo objeto.

La relación de igualdad es reflexiva, es decir para todo objeto x se tiene x = x; es simétrica, o sea si x = y entonces y = x; es transitiva, es decir si x = y e y = z entonces x = z.

Conjunto vacío. Consideremos la siguiente expresión:

$$\{x : x \in Z \text{ , } x = 0 \text{ y } x = 1\}$$

Es claro que este conjunto no tieneningún elemento ya que no existe ningún número igual a 0 y a 1 a la vez, pues $0 \neq 1$.

A fin de evitar excepciones en las definiciones, teoremas, etc. conviene introducir un conjunto sin elementos, llamado conjunto vacío y que se representa $\emptyset$ definido como sigue:

$$\emptyset = \{x : x \neq x\}$$

En general, cualquier propiedad que no sea verificada por ningún objeto puede usarse para definir el conjunto vacío.

Igualdad de conjuntos.

Definición. Un conjunto A es igual a otro B y se escribe A = B si todo elemento de A es elemento de B y todo elemento de B es elemento de A.

Es decir, dos conjuntos son iguales cuando están formados por los mismos objetos.

De la definición resulta en particular que, por ejemplo, los conjuntos {0,1,2,3} y {1,0,3,2} son iguales, es decir no interesa el orden de los elementos. Además cada elemento del conjunto debe figurar una sola vez: por ejemplo, los conjuntos {1,2,2} y {1,2} son iguales.

EJEMPLOS:

1) Consideremos los conjuntos $A = \{x : x \in N \text{ y } x \text{ divide a } 8\}$, $B = \{x : x=2^n \text{ para } n=0,1,2,3\}$. Entonces $A = \{1,2,4,8\}$ y $B = \{1,2,4,8\}$. Luego A = B.

2) Si $X = \{x : x \in N, x \text{ es primo y } 2 < x \leq 9\}$ e $Y = \{x : x \in Z, x \text{ no es divisible por}$

2 y $2 \leq x < 9\}$, se ve fácilmente que X = Y.

3) Sean $A = \{x : x \in R \text{ y } x < 1\}$, $B = \{x : x \in R \text{ y } 2x-1 < 1\}$. Veamos que A = B.

Hay que probar que todo elemento de A pertenece a B y recíprocamente, que todo elemento de B pertenece a A. Sea x un elemento cualquiera de A, $x \in A$. Luego $x < 1$, lo que implica $2x < 2$ y de aquí se deduce $2x-1 < 1$. Por lo tanto $x \in B$.

Recíprocamente, sea x un elemento cualquiera de B, $x \in B$. Entonces $2x-1 < 1$, lo que implica $2x < 2$ o sea $x < 1$. Luego $x \in A$.

Queda probado así que A = B.

4) Sea P = conjunto de todos los números naturales pares.
T = conjunto de todos números naturales cuyo cuadrado es par.

Demostrar que P = T.

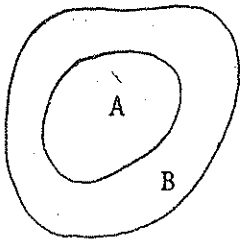
5) Si $X = \{x : x \text{ es un triángulo equilátero}\}$.
$Y = \{x : x \text{ es un triángulo con los tres ángulos iguales}\}$.

se demuestra en geometría que X = Y.

Relación de inclusión.

Definición. Se dice que un conjunto A está contenido en otro conjunto B, que A es una parte de B, que A es un subconjunto de B o que B contiene a A si todo elemento de A es elemento de B.

Se escribe $A \subseteq B$ ó $B \supseteq A$.



Notemos que esta relación no excluye la igualdad de los dos conjuntos.

Si $A \subseteq B$ y además $A \neq B$ se dice que A es un subconjunto propio de B y se escribe $A \subsetneq B$.

La notación $A \not\subset B$ significa que la relación $A \subset B$ no se verifica, es decir que existe por lo menos un elemento de A que no pertenece a B.

EJEMPLOS:

1) Dados los conjuntos $A = \{a,b,c,d\}$, $B = \{b,d,e\}$, $C = \{a,f,b,c,h,d,e\}$ se verifica:

$A \subset C$, $B \subset C$, $B \not\subset A$, $A \not\subset B$.

2) $N \subset Z \subset Q \subset R \subset C$.

3) Sean P = conjunto de los enteros pares.
T = conjunto de los enteros múltiplos de 3.
S = conjunto de los enteros múltiplos de 6.

Entonces $S \subset P$, $S \subset T$, $P \not\subset T$, $T \not\subset P$.

La relación de inclusión verifica las siguientes propiedades:

I) Propiedad reflexiva: $A \subset A$, cualquiera sea el conjunto A.

II) Propiedad antisimétrica: Si $A \subset B$ y $B \subset A$ entonces A = B.

III) Propiedad transitiva: Si $A \subset B$ y $B \subset C$ entonces $A \subset C$.

Su demostración queda a cargo del lector.

El conjunto $\emptyset$ está contenido en cualquier conjunto A:

$$\emptyset \subset A$$

ya que la proposición "si $x \in \emptyset$ entonces $x \in A$" se verifica vacuamente.

Conjunto de las partes de un conjunto dado.

Dado un conjunto A cualquiera se puede considerar siempre el conjunto de los subconjuntos de A, es decir el conjunto formado por todas las partes de A, que se representa $P(A)$.

El conjunto $P(A)$ no es nunca vacío pues $\emptyset$ y A son elementos de $P(A)$ ya que $\emptyset \subset A$ y $A \subset A$.

Por ejemplo, si $A = \{a,b,c\}$ los subconjuntos de A son:

$$\emptyset \ , \ \{a\} \ , \ \{b\} \ , \ \{c\} \ , \ \{a,b\} \ , \ \{a,c\} \ , \ \{b,c\} \ , \ A$$

y $$P(A) = \{\emptyset \ , \ \{a\} \ , \ \{b\} \ , \ \{c\} \ , \ \{a,b\} \ , \ \{a,c\} \ , \ \{b,c\} \ , \ A\} \ .$$

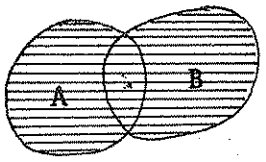
Si A es un conjunto finito con n elementos entonces $P(A)$ es un conjunto finito con 2^n elementos, como probaremos más adelante.

Se definen operaciones que a partir de conjuntos dados permiten obtener nuevos conjuntos: la reunión, la intersección y la complementación.

Reunión de conjuntos.

Definición. Dados dos conjuntos A y B se llama reunión de A y B y se representa $A \cup B$ al conjunto de todos los elementos que pertenecen a A ó a B.

$$A \cup B = \{x : x \in A \text{ ó } x \in B\}$$

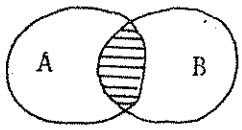


De la definición se deduce inmediatamente que: $A \subset A \cup B$ y $B \subset A \cup B$, para cualquier par de conjuntos A y B.

Intersección de conjuntos.

Definición. Dados dos conjuntos A y B se llama intersección de A y B y se representa $A \cap B$ al conjunto de todos los elementos que pertenecen simultáneamente a A y a B.

$$A \cap B = \{x : x \in A \text{ y } x \in B\}$$



De la definición resulta: $A \cap B \subset A$ y $A \cap B \subset B$, cualquiera sean los conjuntos A y B.

Si la intersección de dos conjuntos A y B es vacía, $A \cap B = \emptyset$, se dice que los conjuntos son disjuntos. La condición necesaria y suficiente para que dos conjuntos no sean disjuntos es que tengan por lo menos un elemento común.

EJEMPLOS:

1) Dados los conjuntos $A = \{a,c\}$, $B = \{a,b,d,e\}$, $C = \{1,a,c,d\}$, $D = \{0,1,2,b\}$ es:

$A \cup B = \{a,b,c,d,e\}$
$A \cap D = \emptyset$
$B \cup C = \{1,a,b,c,d,e\}$
$A \cap C = A$
$(B \cap C) \cup A = \{a,c,d\}$

2) Sean M = conjunto de los números enteros múltiplos de 6.
P = conjunto de los números enteros múltiplos de 2.
T = conjunto de los números enteros múltiplos de 3.

Entonces: $M \cup P = P$, $M \cup T = T$, $M \cap P = M$, $P \cap T = M$, $(M \cup P) \cap T = M$.

3) Sean A = conjunto de todas las personas de ojos negros.
B = conjunto de todas las personas que escriben a máquina.
C = conjunto de todas las personas zurdas.

$B \cap C$ es el conjunto de los zurdos que escriben a máquina.

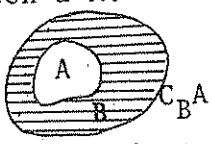
$(A \cap B) \cap C$ es el conjunto de las personas zurdas de ojos negros que escriben a máquina.

$B \cup (A \cap C)$ es el conjunto de las personas que escriben a máquina o son zurdas de ojos negros (incluyendo las que llenan las dos condiciones a la vez).

Complemento de un conjunto.

Definición. Dados dos conjuntos A y B , $A \subset B$, se llama complemento de A relativo a B al conjunto de todos los elementos de B que no pertenecen a A. Se lo representa $C_B A$.

$$C_B A = \{x : x \in B \text{ y } x \notin A\}$$



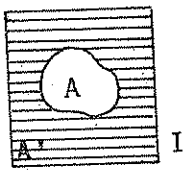
Por ejemplo, si B es el conjunto de todos los hombres y A el conjunto de todos los hombres casados, el complemento de A relativo a B es el conjunto de los hombres solteros. Si $P \subset Z$ es el conjunto de los enteros pares, el complemento de P relativo a Z es el conjunto de los enteros impares.

Notemos que si A = B entonces $C_B A = \emptyset$.

Cuando se desarrolla específicamente una cierta teoría, los razonamientos se aplican a una clase particular de objetos. Se estudian así las propiedades de los elementos de un conjunto fijo bien determinado y cualquier proposición se refiere siempre a esos objetos prescindiéndose de todos los demás.

Generalmente se considera entonces un conjunto fijo I y se trabaja siempre con subconjuntos de I y complementos relativos a I. A I se lo llama conjunto universal y el complemento de un subconjunto cualquiera de I se nota simplemente A' o -A.

$$A' = \{x : x \notin A\}$$



Por ejemplo, dado I = {1,2,3,a,b,c,d,e} y los subconjuntos A = {2,b,d,e} y B = {1,2,e} los respectivos complementos son:

$$A' = \{1,3,a,c\} \quad , \quad B' = \{3,a,b,c,d\}$$

Ejercicio: Demostrar las siguientes propiedades de la complementación:

1) $(A')' = A$

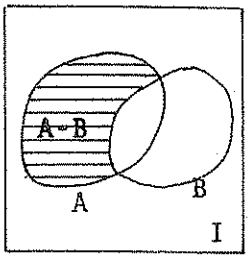
2) $A \cup A' = I$, $A \cap A' = \emptyset$

3) $I' = \emptyset$, $\emptyset' = I$

Diferencia de dos conjuntos.

Definición. Se llama diferencia de dos conjuntos A y B y se escribe A-B al conjunto d todos los elementos de A que no pertenecen a B.

$$A-B = \{x : x \in A \text{ y } x \notin B\}$$



Como $B' = \{x : x \notin B\}$ resulta $A-B = A \cap B'$

Propiedades de la reunión e intersección de conjuntos.

La reunión e intersección de conjuntos tienen las siguientes propiedades:

Propiedades de la reunión.

1. Propiedad idempotente: $A \cup A = A$
2. Propiedad conmutativa: $A \cup B = B \cup A$
3. Propiedad asociativa : $(A \cup B) \cup C = A \cup (B \cup C)$
4. Ley de absorción: $A \cup (A \cap B) = A$

Propiedades de la intersección.

1'. Propiedad idempotente: $A \cap A = A$
2'. Propiedad conmutativa: $A \cap B = B \cap A$
3'. Propiedad asociativa:: $(A \cap B) \cap C = A \cap (B \cap C)$
4'. Ley de absorción: $A \cap (A \cup B) = A$ ●

cualesquiera sean los conjuntos A, B, C.

La demostración queda a cargo del lector. Como ejemplo probaremos la propiedad asociativa y la ley de absorción de la reunión.

Para demostrar la igualdad de los conjuntos $(A \cup B) \cup C$ y $A \cup (B \cup C)$ hay que probar que todo elemento de $(A \cup B) \cup C$ pertenece a $A \cup (B \cup C)$ y recíprocamente, que todo elemento de $A \cup (B \cup C)$ pertenece a $(A \cup B) \cup C$, o sea hay que probar las dos inclusiones siguientes:

$$(A \cup B) \cup C \subset A \cup (B \cup C) \qquad \text{(i)}$$

$$A \cup (B \cup C) \subset (A \cup B) \cup C \qquad \text{(ii)}$$

Sea $x \in (A \cup B) \cup C$. Entonces $x \in A \cup B$ o $x \in C$, luego $x \in A$ o $x \in B$ o $x \in C$, es decir $x \in A$ o $x \in B \cup C$. Luego $x \in A \cup (B \cup C)$. Queda demostrada así la primera inclusión. La otra se prueba en forma análoga.

Probemos ahora la ley de absorción $A \cup (A \cap B) = A$.

Por definición de reunión se tiene $A \subset A \cup (A \cap B)$ (i).
Por otro lado, si $x \in A \cup (A \cap B)$ entonces $x \in A$ o $x \in A \cap B$, es decir $x \in A$ o $x \in A$ y $x \in B$. En cualquier caso $x \in A$. Luego $A \cup (A \cap B) \subset A$ (ii).

De (i) y (ii) sigue la igualdad.

Propiedades distributivas.

5. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
5'. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

cualesquiera sean los conjuntos A,B,C.

Demostración:

Probemos la primera. Hay que demostrar las dos inclusiones siguientes:

$$A \cup (B \cap C) \subset (A \cup B) \cap (A \cup C) \qquad \text{(i)}$$

$$(A \cup B) \cap (A \cup C) \subset A \cup (B \cap C) \qquad \text{(ii)}$$

a) Sea $x \in A \cup (B \cap C)$. Entonces $x \in A$ o $x \in B \cap C$.
Si $x \in A$, entonces $x \in A \cup B$ y $x \in A \cup C$; luego $x \in (A \cup B) \cap (A \cup C)$. Si $x \in B \cap C$ entonces $x \in B$ y $x \in C$;luego $x \in A \cup B$ y $x \in A \cup C$ o sea $x \in (A \cup B) \cap (A \cup C)$. Por lo tanto en cualquier caso $x \in (A \cup B) \cap (A \cup C)$, lo que prueba la inclusión (i).

b) Sea $x \in (A \cup B) \cap (A \cup C)$. Entonces $x \in A \cup B$ y $x \in A \cup C$. Luego $x \in A$ o $x \in B$ y $x \in A$ o $x \in C$. Si $x \in A$ entonces $x \in A \cup (B \cap C)$. Si $x \notin A$ entonces debe ser $x \in B$ y $x \in C$; luego $x \in B \cap C$ de donde $x \in A \cup (B \cap C)$. De modo que en cualquier caso $x \in A \cup (B \cap$
Queda demostrada así la inclusión (ii).

De (i) y (ii) sigue la igualdad.

La otra propiedad distributiva puede demostrarse por un razonamiento análogo, lo que queda propuesto como ejercicio, o también aplicando la propiedad distributiva recién demostrada y las propiedades 2,4,3',4' de la siguiente manera:

$$(A \cap B) \cup (A \cap C) \overset{5}{=} [(A \cap B) \cup A] \cap [(A \cap B) \cup C] \overset{2}{=} [A \cup (A \cap B)] \cap [C \cup (A \cap B)] \overset{4}{=}$$

$$A \cap [C \cup (A \cap B)] \overset{5}{=} A \cap [(C \cup A) \cap (C \cup B)] \overset{3'}{=} [A \cap (C \cup A)] \cap (C \cup B) \overset{2}{=}$$

$$\overset{2}{=} [A \cap (A \cup C)] \cap (B \cup C) \overset{4'}{=} A \cap (B \cup C).$$

Las propiedades de la $\cup$ y la $\cap$ hasta aquí mencionadas son similares a leyes de la aritmética. Notemos que si se reemplaza la $\cup$ por la suma y la $\cap$ por multiplicación de números esas propiedades corresponden a propiedades de la suma y la multiplicación, excepto 1,1',4,4' y 5 que no tienen equivalente en aritmética.

Leyes de Morgan.

6. $(A \cup B)' = A' \cap B'$
6'. $(A \cap B)' = A' \cup B'$

Demostración:

6. $x \in (A \cup B)' \iff x \notin A \cup B \iff x \notin A$ y $x \notin B \iff x \in A'$ y $x \in B' \iff x \in A' \cap B$

La otra se demuestra en forma análoga o también aplicando la recién demostrada como sigu

$$A' \cup B' = [(A' \cup B')']' = [(A')' \cap (B')']' = (A \cap B)'$$

TEOREMA 1.1. Las siguientes proposiciones son equivalentes, cualesquiera sean los conjuntos A y B:

(1) $A \subset B$
(2) $A \cap B = A$
(3) $A \cup B = B$
(4) $B' \subset A'$

Demostración: Se trata de demostrar que estas proposiciones son todas equivalentes entre sí, es decir que (1) $\Longleftrightarrow$ (2) , (1) $\Longleftrightarrow$ (3) , (1) $\Longleftrightarrow$ (4) , (2) $\Longleftrightarrow$ (3) , (2) $\Longleftrightarrow$ (4) y (3) $\Longleftrightarrow$ (4). Para ello es suficiente demostrar las implicaciones (1) $\Longrightarrow$ (2) , (2) $\Longrightarrow$ (3) , (3) $\Longrightarrow$ (4) y (4) $\Longrightarrow$ (1) pues de aquí, por transitividad de la implicación, resultan todas las equivalencias anteriores. Por ejemplo, veamos que (1) $\Longleftrightarrow$ (3). De (1) $\Longrightarrow$ (2) y (2) $\Longrightarrow$ (3) sigue por transitividad que (1) $\Longrightarrow$ (3). De (3) $\Longrightarrow$ (4) y (4) $\Longrightarrow$ (1) resulta que (3) $\Longrightarrow$ (1). Luego (1) $\Longleftrightarrow$ (3).

Probaremos entonces las implicaciones indicadas.

(1) $\Longrightarrow$ (2)

Tenemos que probar que si $A \subset B$ entonces $A \cap B = A$. Supongamos $A \subset B$. Sea x un elemento cualquiera de A, $x \in A$. Por la hipótesis hecha $x \in B$. Es decir $x \in A \cap B$. Luego $A \subset A \cap B$ (i).

Además sabemos, por definición de intersección, que $A \cap B \subset A$ (ii). De (i) y (ii) resulta $A \cap B = A$.

(2) $\Longrightarrow$ (3)

Hay que probar que si $A \cap B = A$ entonces $A \cup B = B$. Supongamos $A \cap B = A$. Entonces reemplazando se tiene

$$A \cup B = (A \cap B) \cup B = B \cup (B \cap A) = B$$

por la propiedad conmutativa de la reunión y la ley de absorción.

(3) $\Longrightarrow$ (4)

Hay que probar que si $A \cup B = B$ entonces $B' \subset A'$. Supongamos $A \cup B = B$. Sea x un elemento cualquiera de B', $x \in B' \Longrightarrow x \notin B \Longrightarrow x \notin A \cup B$, pues por hipótesis $B = A \cup B \Rightarrow$ $\Longrightarrow x \notin A \Longrightarrow x \in A'$. Luego $B' \subset A'$.

(4) $\Longrightarrow$ (1)

Hay que demostrar que si $B' \subset A'$ entonces $A \subset B$. Supongamos $B' \subset A'$. Sea $x \in A$. Luego $x \notin A'$ y como $B' \subset A'$ por hipótesis, $x \notin B'$. Entonces $x \in B$. Luego $A \subset B$.

EJERCICIO: Interpretar gráficamente el teorema anterior mediante un diagrama de Venn. (Se llaman así los diagramas que hemos utilizado para representar las operaciones con conjuntos).

Los conceptos de reunión e intersección de conjuntos se generalizan extendiéndolos a una colección cualquiera de conjuntos de la siguiente manera: Si C es una colección (finita o infinita) de conjuntos, entonces se llama:

a) Reunión de los conjuntos de C al conjunto formado por todos los elementos que pertenecen por lo menos a uno de los conjuntos de C. Se representa $\bigcup_{X \in C} X$.

   En símbolos

$$\bigcup_{X \in C} X = \{x : x \in X \text{ para algún } X \in C\}$$

b) Intersección de los conjuntos de C al conjunto formado por todos los elementos que pertenecen simultáneamente a todos los conjuntos de C. Se representa $\bigcap_{X \in C} X$.

$$\bigcap_{X \in C} X = \{x : x \in X \text{ para todo } X \in C\}$$

Por ejemplo, para cada número natural n sea X_n el conjunto de todos los números enteros menores que n. Así X_1 es el conjunto de todos los enteros negativos y el cero; X_2 está formado por los enteros negativos, 0 y 1; X_3 por los enteros negativos, 0,1 y 2, etc.

Considerando la colección $(X_n)_{n\in N}$ de todos estos conjuntos se tiene

$$\cup_{n\in N} X_n = Z$$

$$\cap_{n\in N} X_n = Z - N$$

Producto cartesiano de conjuntos.

Dados dos conjuntos A y B se pueden considerar los pares ordenados (a,b) donde $a \in A$ y $b \in B$. Dos pares ordenados (a,b) y (a',b') son iguales si y solo si $a = a'$ y $b = b'$.

Definición. Dados dos conjuntos A y B se llama producto cartesiano de A por B y se representa $A \times B$ al conjunto de todos los pares ordenados (a,b) donde $a \in A$ y $b \in B$.

$$A \times B = \{(a,b) : a \in A \text{ y } b \in B\}$$

A y B se llaman primer y segundo factor respectivamente. Si $A \neq B$ es claro que $A \times B \neq B \times A$.

Si A = B el producto cartesiano $A \times A$ se nota tambien A^2.

EJEMPLOS:

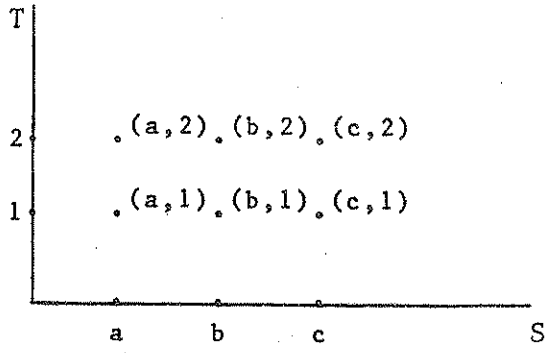
1) Sean $S = \{a,b,c\}$, $T = \{1,2\}$. Entonces

$$S \times T = \{(a,1),(a,2),(b,1),(b,2),(c,1),(c,2)\}$$

$$T \times S = \{(1,a),(1,b),(1,c),(2,a),(2,b),(2,c)\}$$

Para representar el producto cartesiano de A por B se suelen considerar dos rectas perpendiculares en el plano, los elementos de A se representan por puntos de la recta horizontal, los de B por puntos de la vertical y cada elemento (a,b) del producto cartesiano $A \times B$ por el punto del plano que es intersección de las perpendiculares trazadas a los ejes por los puntos que representan a a y a b.

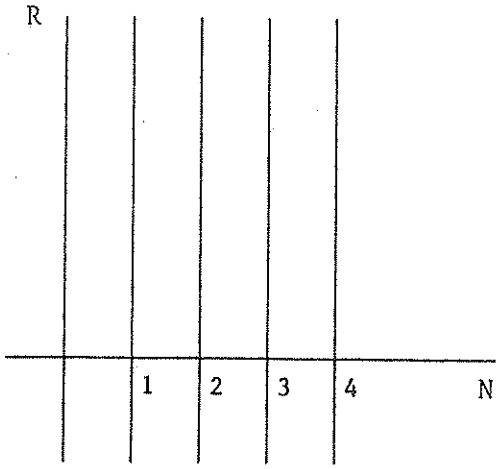
Por ejemplo, el gráfico siguiente representa al producto $S \times T$.



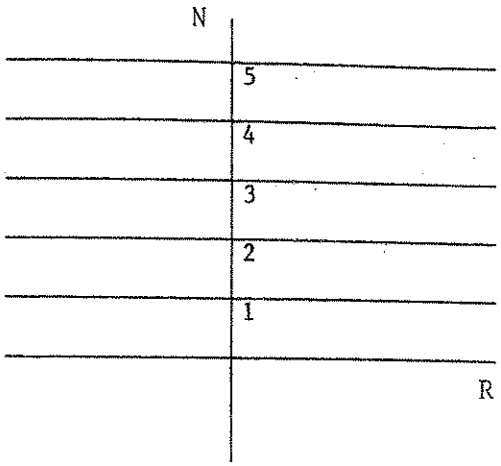
2) Si $A = \{a,b\}$ entonces
$A \times A = \{(a,a),(a,b),(b,a),(b,b)\}$.

3) El plano puede considerarse el producto cartesiano $R \times R = R^2$ de dos rectas.

4) Si N es el conjunto de los números naturales y R el de los números reales, los productos cartesianos N × R y R × N están representados en los gráficos siguientes:



N × R



R × N

La noción de producto cartesiano se puede generalizar para un número cualquiera de conjuntos. Por ejemplo, dados tres conjuntos A,B,C se llama producto cartesiano de A por B por C y se representa A × B × C al conjunto de todas las ternas ordenadas (a,b,c) donde a ε A, b ε B y c ε C.

Así por ejemplo, el espacio de tres dimensiones de la geometría euclídea puede considerarse el producto cartesiano de tres rectas $R \times R \times R = R^3$.

En general el producto cartesiano de n conjuntos $A_1, A_2, \ldots, A_n$ en ese orden es el conjunto de todas las n-uplas $(a_1, a_2, \ldots, a_n)$ donde $a_i \in A_i$ para i=1,2,...,n. Se escribe $A_1 \times A_2 \times \ldots \times A_n$.

NOTA: La importancia de la teoría de conjuntos radica en que todos los entes que se estudian y analizan en matemática (con muy raras excepciones) pueden definirse en términos de conjuntos. La teoría de conjuntos comenzó a desarrollarse alrededor de 1870, época en que George Cantor (1845-1918) inició su trabajo sobre conjuntos infinitos.

Pero la idea intuitiva de conjunto, elemento, pertenencia y la formulación ingenua de la teoría de conjuntos conduce a contradicciones. Por ejemplo, la existencia de un universo absoluto, es decir del conjunto de todos los conjuntos, era naturalmente admitida en los orígenes del desarrollo de la teoría de conjuntos, pero no se puede considerar que exista tal conjunto. Es bien conocida a este respecto la llamada paradoja de Bertrand Russell (1872-1970): Supongamos que A es el conjunto de todos los conjuntos y sea

$$B = \{x \in A : x \notin x\}$$

Una de las dos proposiciones siguientes es verdadera: $B \in B$ ó $B \notin B$.

Si $B \in B$, como $B \in A$ por la hipótesis sobre A, es $B \notin B$. Contradicción.

Si $B \notin B$, como $B \in A$ por la hipótesis sobre A, es $B \in B$. Contradicción.

Esto prueba que $B \in A$ es imposible, luego $B \notin A$ y por lo tanto hay un conjunto que no pertenece a A.

En consecuencia es necesario imponer ciertas restricciones para que cada proposición determine un conjunto.

Para precisar las nociones de conjunto y relación de pertenencia y evitar las contradicciones se formularon en este siglo teorías axiomáticas de conjuntos. El método axiomático consiste en establecer ciertas proposiciones de partida, llamadas axiomas o postulados, que se consideran verdaderas y que se refieren a entes dados sin definición, las llamadas nociones primitivas. A partir de los axiomas se deducen otras propiedades o teoremas de la teoría. En algunas teorías axiomáticas de conjuntos se consideran como nociones primitivas, por ejemplo, las de conjunto y pertenencia. Mediante los axiomas se trata de fijar su comportamiento de modo que la teoría resultante sea consistente, es decir que no haya lugar a contradicciones.

La idea clásica de que un axioma o postulado es "una verdad evidente" hace mucho que ha desaparecido de la matemática. Un axioma es simplemente una proposición inicial que se considera verdadera y que se usa para deducir nuevas propiedades o relaciones entre las nociones primitivas.

Por otra parte, es imposible elaborar una teoría en la que se definan todos los términos y se demuestren todas las proposiciones. Algunos términos se podrán definir a partir de otros cuyo significado ya ha sido dado pero en cualquier caso se llegará a términos que no serán susceptibles de definición, a conceptos primitivos. Análogamente, en general pueden deducirse nuevas proposiciones de propiedades ya establecidas, las que a su vez pueden haberse demostrado a partir de otras anteriores. Pero cualquiera sea la teoría que se considere, siempre se tienen proposiciones primeras que no pueden demostrarse a partir de ninguna precedente.

Cuando los hechos de una teoría se ordenan lógicamente de tal manera que todos sin excepción pueden ser deducidos a partir de algunos de ellos convenientemente elegidos, entonces se dice que se tiene una formulación axiomática de la teoría, en la que los axiomas correspondientes son esas proposiciones a partir de las cuales se deducen todas las demás.

Históricamente el método axiomático en matemática se remonta a Euclides (a su obra "Elementos" escrita hacia el año 300 a.C.) y desde entonces la geometría ha sido el prototipo de una disciplina axiomatizada. Se toman como nociones primitivas los conceptos geométricos de punto, recta, incidencia, etc. y los distintos postulados dan una definición implícita de estos conceptos al fijar su comportamiento.

Para ver un tratamiento más formal de la teoría de conjuntos se puede consultar, por ejemplo, el libro de Paul R. Halmos, Naive set theory.

EJERCICIOS.

1. Escribir si cada uno de los siguientes números: -2, $\sqrt{3}$, -1.35, π, $2i$, $1+\sqrt{2}$, $0.3333..$, $18/5i$ pertenece o no a los conjuntos: i) Q ; ii) R ; iii) C.

2. Dados el conjunto $I = \{a,b,c,d,e,f,0,1,2,3\}$ y los subconjuntos $A = \{a,b,c,d,e,f\}$, $B = \{0,1\}$, $C = \{1,a,c,d\}$, $D = \{0,1,2,a,c,d\}$:

   a) Relacionarlos dos a dos por la relación de inclusión " $\subset$ ".

   b) Hallar: $A \cap C$, $A \cap B$, $C \cap D$, $B \cup C$, $C \cup D$, $C \cup D \cup B$, $(A \cap B) \cup (B \cap D)$, A' , $(A \cap C)'$, $(A \cup C)'$, $(A \cup D)'$, $(A \cap B)'$, $A' \cup B$, $A - C$, $(A' \cup C') - B$.

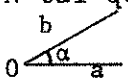
   c) Verificar que:

   i) $(A')' = A$
   ii) $(A \cap D)' = A' \cup D'$
   iii) $(A \cup D)' = A' \cap D'$
   iv) $A \cap (A \cup C) = A$
   v) $A \cup (A \cap C) = A$
   vi) $A \subset B'$
   vii) $D' \subset (A \cap C)'$
   viii) $(A \cap D)' = C' \cup B$

3. Representar gráficamente en cada caso mediante diagramas de Venn tres conjuntos A, B,C no vacíos tales que:

   a) $A \subset B$, $C \subset B$, $A \cap C \neq \emptyset$
   b) $A \subset B$, $A \cap C = \emptyset$, $B \cap C \neq \emptyset$
   c) $C \supset A$, $A \neq C$, $B \cap C = \emptyset$
   d) $A \subset B \cap C$, $C \neq B$, $A \neq B$
   e) $B \cup C \subset A$, $B \cap C = \emptyset$
   f) $A \cup B = A$, $A \cup C = B$, $B \neq C$

4. Decir qué elementos forman los siguientes conjuntos:

   a) $A_1 = \{x : x \in N \text{ y } x = 2n+1 , n \in N\}$
   b) $A_2 = \{x : x \in N \text{ y } x^2 = 1\}$
   c) $A_3 = \{x : x \in Z \text{ y } x^2 = 1\}$
   d) $A_4 = \{x : x \in N , x = 2n+1 \text{ y } n \leq 10\}$
   e) $A_5 = \{x : x \in Z \text{ y } x^2 = 2\}$
   f) $A_6 = \{x : x \in R \text{ y } -1 \leq x < 3\}$
   g) $A_7 = \{x : x \in Z \text{ y } x^2 < 30\}$
   h) $A_8 = \{x : x \in N \text{ y } x \text{ divide a } 24\}$
   i) $A_9 = \{x : x \in Z , x^2+1 > 4 \text{ y } x = 2n , n \in Z\}$
   j) $A_{10} = \{x : x \in Q \text{ y } 2x \in Z\}$
   k) $A_{11} = \{x : x \in N \text{ y } \exists a \in N \text{ tal que } x = a^2\}$
   l) Consideremos el ángulo 

   y sea P el conjunto de los puntos del plano.

   $A_{12} = \{x : x \in P \text{ y } d(x,a)=d(x,b)\}$ donde d significa distancia.

5. Definir simbólicamente los siguientes conjuntos:

   a) De los números naturales mayores que 20.
      De los números naturales múltiplos de 6.
   b) De los números reales positivos de cuadrado ≤ 2.
      De los números reales de valor absoluto menor que 10.

c) De los números enteros impares menores que 5.
De los números enteros cuadrados perfectos.

d) En el plano, el conjunto de puntos de una circunferencia de centro 0 y radio r.
En el plano, el conjunto de puntos de un círculo abierto de centro 0 y radio r.
En el plano, el conjunto de puntos de un círculo cerrado de centro 0 y radio r.

6. Consideremos los siguientes subconjuntos de números naturales:

$P = \{x : x \text{ par}\}$
$I = \{x : x \text{ impar}\}$
$S = \{x : x \text{ divisible por } 5\}$
$T = \{x : x \text{ múltiplo de } 10\}$

Hallar: $P \cap I$, $P' \cap I$, $P' \cup I$, $S \cap T$, $P \cap T$, $I \cap T$, $(P \cap T) \cup S$, $I' \cap S$, $(I' \cap S) \cup T$, $(T - P) \cup (S - P)$.

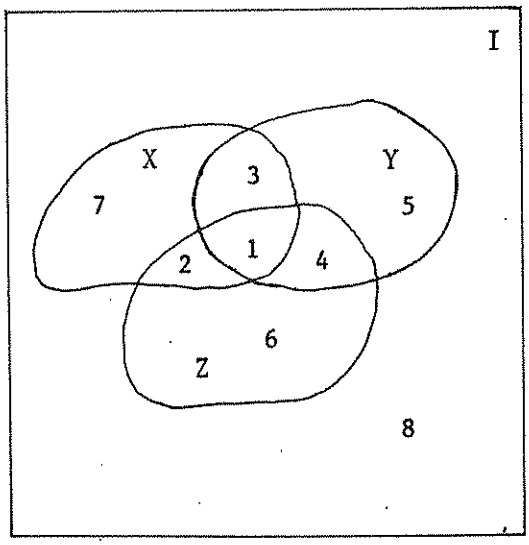
7. Sea X = conjunto de los números naturales pares.
Y = conjunto de los números naturales cuyo cuadrado es par.

Demostrar que X = Y.

8. Demostrar las siguientes propiedades interpretándolas previamente en forma gráfica:

a) Si $A \subset B$ y $C \subset D$ entonces $A \cap C \subset B \cap D$ y $A \cup C \subset B \cup D$
b) Aplicando a) deducir que:
si $C \subset A$ y $C \subset B$ entonces $C \subset A \cap B$
si $A \subset C$ y $B \subset C$ entonces $A \cup B \subset C$
c) $A \supset B$ si y solo si $B \cup (A - B) = A$
d) $A \subset B$ si y solo si $A \cap B' = \emptyset$
e) Si $A \neq B$ entonces $A-B \neq B-A$

9. Encontrar una expresión algebraica en X,Y,Z que determine exactamente cada uno de los subconjuntos que figuran en el diagrama siguiente:



Por ejemplo, el subconjunto 7 es $X - (Y \cup Z)$.

10. Verificar por cálculo directo las siguientes igualdades:

a) $B \cap (A - B) = \emptyset$
b) $(A - B) \cup (B - A) = (A \cup B) - (A \cap B)$

c) $X \cap (Y - Z) = (X \cap Y) - Z$

d) $(A \cap B')' \cup (A \cup B)' = A' \cup B$

e) $A \cap [(B \cap C) \cap A]' = (A - B) \cup (A - C)$

f) $A - B = A - (B \cap A) = (A \cup B) - B$

g) $(A \cup B) - (A \cap B) = [A - (A \cap B)] \cup [B - (A \cap B)]$

h) $[(A' \cup C)' \cap B]' \cup (B' \cup A)' = (A \cap B)' \cup C$

i) $[(A - B)' \cup C] \cap [(A \cap C) - B]' = A' \cup B$

j) $[(A \cup B)' \cup (A' - B)]' \cap (B - A)' = A$

11. En una encuesta realizada entre 100 estudiantes los resultados obtenidos fueron los siguientes:

25 alumnos estudiaban el idioma español.
32 " " " " alemán.
41 " " " " francés.
7 " " " " español y alemán.
12 " " " " español y francés.
6 " " " " alemán y francés.
4 " " los 3 idiomas.

a) ¿Cuántos alumnos estudiaban sólo francés?

b) ¿ " " no estudiaban ningún idioma?

c) ¿ " " estudiaban sólo español?

12. Dado un conjunto A qué condición debe cumplir B para que se verifique $A \cup (B-A)=A$?

13. (De P. SUPPES, Introduction to logic). Sean:

V = conjunto de todas las personas.
A = conjunto de todos los americanos.
C = conjunto de todas las personas que beben café.
F = conjunto de todos los franceses.
M = conjunto de todos los asesinos.
P = conjunto de todos los filósofos.
T = conjunto de todas las personas que beben té.
W = conjunto de todas las personas que beben vino.

Expresar simbólicamente las siguientes proposiciones mediante la igualdad, inclu - sión o desigualdad de dos conjuntos:

a) Algunos americanos que beben vino son filósofos.

b) Ningún francés es americano.

c) Personas que beben vino y café también beben té.

d) Todos los asesinos franceses beben café, té y vino.

e) Algunos asesinos americanos beben té y café, pero no beben vino.

f) Algunos asesinos franceses que beben vino no beben té ni café.

g) Un filósofo no bebe té ni café.

h) Algunos franceses son filósofos o asesinos.

i) Todos los bebedores de café beben té o vino.

Por ejemplo la frase e) puede representarse:

$$(A \cap M \cap C \cap T) - W \neq \emptyset$$

14. Conjunto de las partes de un conjunto. Dado $A = \{a,b,c\}$ hallar el conjunto $P(A)$ de las partes de A.

Idem para $A = \{1,2,3,4\}$

15. Producto cartesiano de conjuntos.

a) Dados los conjuntos $A = \{a,b\}$, $B = \{1,2,3\}$ hallar los conjuntos $A \times B$, $B \times A$, A^2 , B^2.

b) ¿Cuáles son los elementos del conjunto R^2?. ¿Cómo se puede representar gráficamente este conjunto?.

c) Si A es un conjunto con n elementos y B uno con m elementos, cuántos elementos tienen $A \times B$ y $B \times A$?.

d) Demostrar que cualesquiera sean A,B,C tres conjuntos no vacíos tales que $B \cap C \neq \emptyset$ se tiene:

$$A \times (B \cap C) = (A \times B) \cap (A \times C)$$

1.3. RELACIONES BINARIAS. La noción de relación es familiar a todos. No es exclusiva de la matemática, sino que se usa continuamente en la vida corriente. Se habla de relaciones cuando se dice: "Vicente es padre de Marcos", "Marcos es hermano de Inés", "Elena es la mujer de Vicente", "Mi auto es más rápido que el tuyo", "4 divide a 12", dos triángulos equiláteros son semejantes", etc.

Vamos a definir la noción de relación. Y esta definición se hace en términos de conjuntos como veremos enseguida.

Consideremos, por ejemplo, la relación "ser padre de". Cuando se dice: "Vicente es padre de Marcos", se está afirmando algo sobre el par de personas Vicente, Marcos e implica que existe un cierto vínculo entre ellos. Es claro que se trata de un par ordenado porque intercambiando los nombres se obtiene una proposición falsa. La relación "ser padre de" es entonces una propiedad que el par ordenado (Vicente, Marcos) tiene en común con otros pares (los formados por padre e hijo) y que no verifican todos los pares de hombres. Esta relación permite entonces distinguir algunos pares ordenados de otros es decir, determina un subconjunto R del producto cartesiano $A \times A$, (si con A representamos al conjunto de todos los hombres): el subconjunto R formado por todos los pares (a,b) tales que a es padre de b. Este subconjunto R proporciona un conocimiento total de la relación "ser padre de" porque dado un par de hombres (a,b), a es padre de b si y solo si $(a,b) \in R$.

Análogamente, la relación "ser la esposa de" puede pensarse como un subconjunto S del producto cartesiano $A \times B$, donde A es el conjunto de las mujeres y B el de los hombres. Afirmar que $a \in A$ es la esposa de $b \in B$ equivale a decir que $(a,b) \in S$.

La relación "divide" entre números naturales puede definirse como el subconjunto $R \subset N \times N$ tal que $R = \{(a,b) : a,b \in N \text{ y } \exists\, c \in N \text{ tal que } b = c.a\}$

Definición. Dados dos conjuntos A y B, una relación binaria R entre elementos de A y elementos de B es un subconjunto del producto cartesiano $A \times B$.

$$R \subset A \times B$$

Dados $a \in A$ y $b \in B$, se dice que a está en la relación R con b y se escribe aRb si y sólo si $(a,b) \in R$.

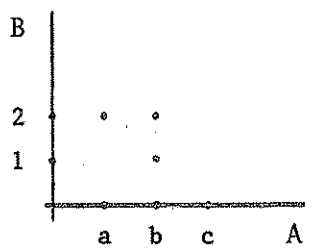
$$aRb \iff (a,b) \in R$$

Se escribe $a\not Rb$ si $(a,b) \notin R$.

Si A = B, todo subconjunto de $A \times A$ es una relación binaria en A.

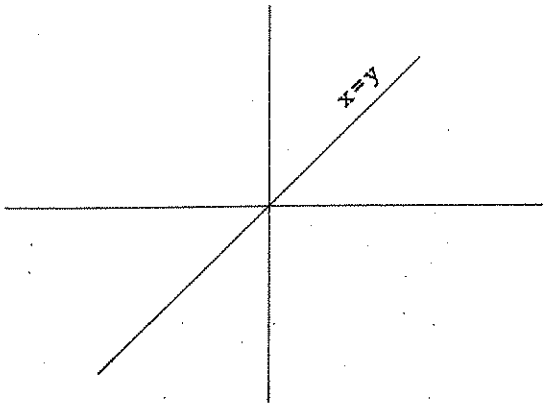
EJEMPLOS:

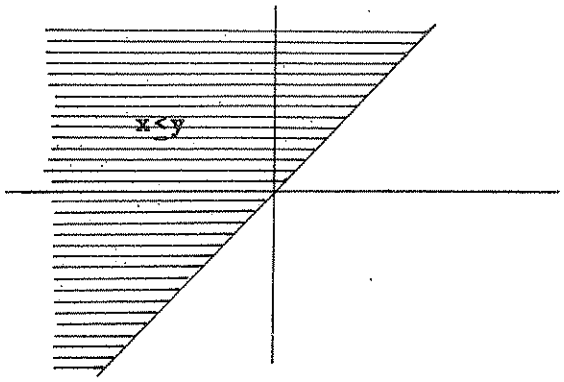
1) Dados los conjuntos $A = \{a,b,c\}$ y $B = \{1,2\}$ el conjunto $R = \{(a,2),(b,1),(b,2)\}$ es una relación binaria entre elementos de A y de B. Gráficamente se la puede representar como sigue:



Un gráfico de este tipo representa efectivamente a la relación R, pues indica exactamente qué pares del conjunto A × B están en ella.

2) La igualdad "x = y" y la relación de menor o igual "x ≤ y" son sendas relaciones binarias en el conjunto R de los números reales. Gráficamente la primera está representada por los puntos de la diagonal del 1^{er} y 3^{er} cuadrantes y la segunda por la región del plano que aparece sombreada.





3) "x divide a y" es una relación binaria definida en el conjunto N de los números naturales. Representarla gráficamente.

Definición. Una relación binaria R definida en un conjunto A ($R \subset A \times A$) se dice:

1) Reflexiva, si aRa cualquiera sea $a \in A$.
2) Simétrica, si aRb implica bRa.
3) Antisimétrica, si aRb y bRa implica a = b.
4) Transitiva, si aRb y bRc implica aRc.

EJEMPLOS:

1) El paralelismo entre las rectas del plano es una relación reflexiva, simétrica y transitiva. No es antisimétrica. Las mismas propiedades tienen la semejanza de polígonos y la equivalencia de polígonos.

2) La relación "menor o igual" entre números reales es reflexiva, antisimétrica y transitiva. No es simétrica.

3) La relación "menor que" entre números reales es transitiva y antisimétrica, no es reflexiva ni simétrica.

4) La relación "divide" entre números naturales es reflexiva, antisimétrica y transitiva. Si se considera esta misma relación definida en el conjunto Z de los enteros entonces sólo es reflexiva y transitiva.

5) La relación de igualdad "x = y" definida en cualquier conjunto verifica todas las propiedades.

6) Sea A un conjunto y P(A) el conjunto de las partes de A. La relación de inclusión entre conjuntos es una relación binaria en P(A) y ya vimos que es reflexiva, antisimétrica y transitiva.

7) Sea A = {1,2,3,4} y R = {(1,1),(2,2),(4,4),(1,2),(2,3),(1,3)}.

Esta relación no es reflexiva porque $3 \not R 3$; no es simétrica pues 1R2 y $2 \not R 1$; es antisimétrica y es transitiva.

Como ejercicio para el lector, si $R = \{(1,1),(3,3),(4,4),(1,3),(1,4),(3,1),(4,1)\}$ decir qué propiedades tiene R.

Las relaciones binarias importantes en matemática son de dos tipos: las relaciones de equivalencia y las relaciones de orden.

Definición. Una relación binaria definida en un conjunto se llama:

Una relación de equivalencia si es reflexiva, simétrica y transitiva.

Una relación de orden si es reflexiva, antisimétrica y transitiva.

De los ejemplos anteriores, son relaciones de equivalencia la igualdad, el paralelismo entre rectas, la semejanza de polígonos, la equivalencia de polígonos; son relaciones de orden la relación "menor o igual" entre números reales, la relación "divide" entre números naturales (y no entre enteros) y la relación de inclusión entre conjuntos.

Estos son unos pocos ejemplos de dos nociones que aparecen constantemente en matemática.

Relación de equivalencia.

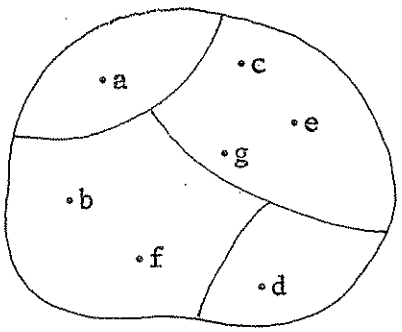
Si $R \subset A \times A$ es una relación de equivalencia, representaremos a R con el símbolo "$\sim$" escribiendo "$a \sim b$" en lugar de aRb. Entonces las propiedades características son:

Propiedad reflexiva: $a \sim a \quad \forall \quad a \in A$
Propiedad simétrica: si $a \sim b$ entonces $b \sim a$.
Propiedad transitiva: si $a \sim b$ y $b \sim c$ entonces $a \sim c$.

La propiedad fundamental que tiene una relación de equivalencia definida en un conjunto A es que determina una partición de A. Antes de seguir adelante precisemos lo que se entiende por partición de un conjunto.

Definición. Una colección de subconjuntos no vacíos de un conjunto A se dice una partición de A si son disjuntos dos a dos y su reunión da A.

Por ejemplo, si $A = \{a,b,c,d,e,f,g\}$, los subconjuntos $X_1 = \{a\}$, $X_2 = \{c,e,g\}$, $X_3 = \{b,f\}$ y $X_4 = \{d\}$ forman una partición de A.



El conjunto de los enteros positivos, el conjunto de los enteros negativos y {0} forman una partición de Z.

Vamos a ver ahora que toda relación de equivalencia definida en un conjunto A determina una partición de A en subconjuntos que se llaman clases de equivalencia, es decir, permite agrupar los elementos de A en subconjuntos disjuntos dos a dos.

Definición. Dada una relación de equivalencia en un conjunto A, se llama clase de equivalencia de un elemento $a \in A$ al conjunto C_a de todos los elementos de A equivalentes con a.

$$C_a = \{x \in A : x \sim a\}$$

Considerando las clases de equivalencia de todos los elementos de A se tiene una colección de subconjuntos de A.

Propiedades de las clases de equivalencia.

1) $a \in C_a$, $\forall\ a \in A$

2) $C_a = C_b$ si y solo si $a \sim b$

3) Si $C_a \neq C_b$ entonces $C_a \cap C_b = \emptyset$

Demostración:

1) $a \in C_a$ pues $a \sim a$, $\forall\ a \in A$ (propiedad reflexiva).

2) Hay que demostrar dos implicaciones:

   a) $C_a = C_b \Longrightarrow a \sim b$. Si $C_a = C_b$, como $a \in C_a$ es $a \in C_b$. Luego, por definición de clase de equivalencia es $a \sim b$.

   b) $a \sim b \Longrightarrow C_a = C_b$. Supongamos $a \sim b$.

   Probemos primero que $C_a \subset C_b$. Sea $x \in C_a$; luego $x \sim a$. Como $a \sim b$, por la propiedad transitiva resulta $x \sim b$ o sea $x \in C_b$. Luego $C_a \subset C_b$ (i).

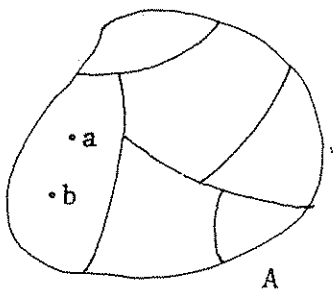
   Veamos que $C_b \subset C_a$. Sea $x \in C_b$. Luego $x \sim b$. Como $a \sim b$, por simetría se tiene $b \sim a$; aplicando la propiedad transitiva resulta $x \sim a$. Luego $x \in C_a$ y $C_b \subset C_a$ (ii)

   De (i) y (ii) sigue $C_a = C_b$.

3) Sea $C_a \neq C_b$ y supongamos que $C_a \cap C_b \neq \emptyset$. Entonces existe por lo menos un elemento $x \in C_a \cap C_b$. Luego $x \sim a$ y $x \sim b$. De aquí resulta, aplicando las propiedades simétrica y transitiva, $a \sim b$. Por 2) es entonces $C_a = C_b$. Esta contradicción provino de suponer $C_a \cap C_b \neq \emptyset$. Luego si $C_a \neq C_b$ es $C_a \cap C_b = \emptyset$.

De estas tres propiedades se deduce que las clases de equivalencia distintas forman una partición del conjunto A. En efecto, por 1) $a \in C_a$, $\forall\ a \in A$, lo que implica que las clases de equivalencia no son vacías y que su reunión da A; y por 3) las clases de equivalencia distintas son disjuntas dos a dos.

Recíprocamente, dada una partición de un conjunto A se puede definir una relación de equivalencia en A de modo que las clases de equivalencia correspondientes coincidan con los subconjuntos de la partición dada. Tratemos de hacerlo. Dado a $\in$ A, a pertenece a un subconjunto X de la partición. Como a debe pertenecer a su clase de equivalencia y las clases de equivalencia deben coincidir con los subconjuntos de la partición dada se ve que la clase de equivalencia de a tiene que ser X. Entonces la relación en cuestión se define como sigue:



Cualesquiera sean a,b ε A

a ∿ b si y solo si a y b pertenecen al mismo suboonjunto de la partición dada.

La verificación de que esta relación es de equivalencia corre por cuenta del lector.

Queda demostrado así el siguiente

TEOREMA 1.2. (Fundamental de las relaciones de equivalencia).

Toda relación de equivalencia en un conjunto A determina una partición de A formada por las distintas clases de equivalencia. Recíprocamente, toda partición de A determina una relación de equivalencia tal que las clases de equivalencia correspondientes son los subconjuntos de la partición dada.

El conjunto de las clases de equivalencia determinadas por una relación de equivalencia ∿ definida en un conjunto A se llama el conjunto cociente de A por la relación ∿ .

El concepto de relación de equivalencia es una generalización del de igualdad. Los elementos que figuran en una misma clase de equivalencia son equivalentes entre sí, intuitivamente iguales bajo un cierto aspecto, y cualquiera de ellos puede tomarse como representante de esa clase. Considerando, por ejemplo, el paralelismo entre las rectas de un plano, las clases de equivalencia correspondientes están formadas por las rectas paralelas entre sí y la noción de dirección de una recta es en realidad la de clase de equivalencia: una recta tiene la misma dirección que todas las paralelas a ella y una cualquiera de ellas representa a esa dirección. En este ejemplo el conjunto cociente es infinito. Eligiendo como representante de cada clase de equivalencia la recta que pasa por un punto fijo 0, el conjunto cociente puede representarse por todas las rectas que pasan por 0.

EJEMPLOS.

1) Veamos un ejemplo no matemático. Sea A un conjunto de 14 personas. Consideremos la relación "x tiene los mismos padres que y" entre los elementos de A. Se trata de una relación de equivalencia pues es reflexiva, simétrica y transitiva. Las clases de equivalencia correspondientes están formadas por las personas que tienen los mismos padres, es decir que son hermanas entre sí. Si entre las personas que figuran en A hay 5, 3 y 2 que son respectivamente hermanas entre sí, entonces A queda dividido en 7 clases de equivalencia: una formada por 5 personas, otra por 3, otra por 2 y cuatro clases de equivalencia formadas por cada una de las 4 personas restantes que no son hermanas de ningún integrante del grupo A. Así el conjunto cociente tiene 7 elementos.

2) Consideremos la siguiente relación en R:

   a ∿ b si y solo si a = ±b. Es una relación de equivalencia y cada número real x≠0 es equivalente consigo mismo y con -x. Luego la clase de equivalencia correspondiente es {x,-x}. Así cada clase de equivalencia está formada por dos números, excepto la formada por el cero. Si como representante de cada clase ≠ {0} se elige el número positivo, el conjunto cociente puede representarse por una semirrecta:

0

3) Consideremos en Z la siguiente relación:

a $\sim$ b si y solo si a.b > 0 ó a = b.

Se verifica fácilmente que es una relación de equivalencia y que un número a es equivalente a otro b si y solo si a y b tienen el mismo signo. Luego las clases de de equivalencia son: el conjunto N de todos los enteros positivos, el conjunto N' de los enteros negativos y {0}. Así Z = N $\cup$ {0} $\cup$ N' y el conjunto cociente de Z por esta relación tiene 3 elementos.

Para terminar vamos a ver un ejemplo más de relación de equivalencia: la congruencia aritmética de números enteros, fundamental en la teoría de números.

Definición. Sea m $\in$ Z un entero fijo. Dos números enteros a y b se dicen congruentes módulo m y se escribe a $\equiv$ b (mód. m) si y solo si a-b es múltiplo de m.

$$a \equiv b \text{ (mód. m)} \iff \exists\, k \in Z \text{ tal que } a-b = k.m$$

EJEMPLOS:

4 $\equiv$ 19 (mód. 3) ; -5 $\equiv$ 30 (mód. 7) ; 11 $\equiv$ -25 (mód. -2) ; -35 $\equiv$ 4 (mód. 13).

Como a $\equiv$ b (mód. m) si y solo si a $\equiv$ b (mód. -m) se considera m $\geq$ 0.

La relación de congruencia entre números enteros para un módulo fijo m es una relación de equivalencia. En efecto, es

1) Reflexiva: a $\equiv$ a (mód. m) $\forall$ a $\in$ Z.
pues a-a = 0.m

2) Simétrica: Si a $\equiv$ b (mód. m) entonces b $\equiv$ a (mód. m).
a $\equiv$ b (mód. m) $\Longrightarrow$ $\exists$ k $\in$ Z tal que a-b = k.m $\Longrightarrow$ b-a = (-k).m $\Longrightarrow$
$\Longrightarrow$ b $\equiv$ a (mód. m).

3) Transitiva: Si a $\equiv$ b (mód. m) y b $\equiv$ c (mód. m) entonces a $\equiv$ c (mód. m).
a $\equiv$ b (mód. m) $\Longrightarrow$ a-b = k.m con k $\in$ Z.
b $\equiv$ c (mód. m) $\Longrightarrow$ b-c = k'.m con k' $\in$ Z.
Sumando las dos igualdades resulta
a-c = (k+k').m o sea a $\equiv$ c (mód. m)

Por lo tanto la relación de congruencia módulo un entero m determina una partición de Z en clases de equivalencia. Por ejemplo, las clases de equivalencia determinadas por la congruencia módulo 3 son:

$$\bar{0} = \{.....,-9,-6,-3,0,3,6,9,12,.....\}$$

$$\bar{1} = \{.....,-8,-5,-2,1,4,7,10,13,....\}$$

$$\bar{2} = \{.....,-7,-4,-1,2,5,8,11,14,....\}$$

En la primera clase están todos los múltiplos de 3, en la segunda todos los múltiplos de 3 más 1 y en la última los múltiplos de 3 más 2. Quedan clasificados así todos los números enteros.

En general, la congruencia módulo m, m $\neq$ 0, determina m clases de equivalencia. Esto resulta de la siguiente propiedad:

PROPOSICION. Dos números enteros a y b son congruentes módulo m, m $\neq$ 0, si y solo si

dan el mismo resto al dividirlos por m.

Demostración:

1) Supongamos que $a \equiv b$ (mód. m) y demostremos que a y b dan restos iguales al dividirlos por m. Sean q y r el cociente y el resto de dividir b por m: $b = qm+r$.

   De $a \equiv b$ (mód. m) resulta $a-b = km$, $k \in Z$. Luego $a = km+b$. Reemplazando b resulta:

$$a = km + qm + r$$
$$a = (k + q)m + r$$

   y en virtud de la unicidad de cociente y resto, de esta última igualdad resulta que $k + q$ es el cociente de dividir a por m y r es el resto.

2) Supongamos que a y b dan el mismo resto:

$$a = qm + r$$
$$b = q'm + r$$

   Hay que probar que $a \equiv b$ (mód. m). Restando las igualdades anteriores resulta: $a-b = (q-q')m$. Luego $a \equiv b$ (mód. m).

Los restos posibles de dividir un número entero cualquiera por otro m son: 0,1,2,..,m-1. Considerando la congruencia módulo m, de la proposición anterior resulta que los números enteros que al ser divididos por m dan resto 0 forman una clase de equivalencia, los que dan resto 1 forman otra,....., los que dan resto m-1 forman otra, y no hay más que éstas. Luego el conjunto cociente de Z por esta relación de equivalencia tiene m elementos. Las clases de equivalencia se suelen representar: $\bar{0}$, $\bar{1}$, $\bar{2}$,....., $\overline{m-1}$ y el conjunto cociente Z_m. Así $Z_3 = \{\bar{0},\bar{1},\bar{2}\}$.

OBSERVACION: La congruencia módulo 0 es la identidad. En efecto, $a \equiv b$ (mód. 0) $\Longleftrightarrow$ $\Longleftrightarrow$ $a-b = k.0 \Longleftrightarrow a = b$. En este caso cada número entero forma una clase de equivallencia.

EJERCICIOS.

1. Hallar las clases de equivalencia determinadas en Z por la congruencia módulo 5. Decir a qué clase pertenecen los números 18, 152, -1347 y 28.474. Idem para la congruencia módulo 6.

2. Indicar un procedimiento para obtener una partición de Z en 19 subconjuntos disjuntos dos a dos.

3. Probar la siguiente propiedad de la congruencia aritmética: Si $a \equiv b$ (mód. m) y $c \equiv d$ (mód. m) entonces $a+c \equiv b+d$ (mód. m) y $a.c \equiv b.d$ (mód. m).

RELACION DE ORDEN. Ya dijimos que una relación binaria R definida en un conjunto A se llama de orden si es reflexiva, antisimétrica y transitiva.

Se dice que A es un conjunto ordenado por R o que R ordena al conjunto A.

Habitualmente una relación de orden se representa por el símbolo " $\leq$ ". Así, en lugar de aRb se escribe $a \leq b$. Con esta notación las propiedades características de una re-

lación de orden se escriben:

Propiedad reflexiva: $a \leq a$, $\forall\, a \in A$

Propiedad antisimétrica: Si $a \leq b$ y $b \leq a$ entonces $a = b$.

Propiedad transitiva: Si $a \leq b$ y $b \leq c$ entonces $a \leq c$.

y se conviene también en la notación siguiente:

$a \not\leq b$ si y solo si $a \leq b$ no se verifica

$a < b$ si y solo si $a \leq b$ y $a \neq b$

$a \geq b$ si y solo si $b \leq a$

Como ejemplos de relaciones de orden citamos la relación "menor o igual" entre números reales, la relación "divide" entre números naturales y la relación de inclusión entre conjuntos.

En general,una relación de orden definida en un conjunto A no permite ordenar todos los pares de elementos de A, es decir dados dos elementos $a,b \in A$ puede suceder que $a \not\leq b$ y $b \not\leq a$.

Un conjunto ordenado A se dice totalmente ordenado si cualesquiera sean $a,b \in A$ es $a \leq b$ o $b \leq a$.

Por ejemplo,

a) El conjunto de los números reales es totalmente ordenado con respecto a la relación menor o igual ordinaria.

b) El conjunto $P(A)$ de las partes de un conjunto A con más de un elemento está ordenado por la relación de inclusión, pero $P(A)$ no es totalmente ordenado porque existen subconjuntos X, Y tales que $X \not\subset Y$ y $Y \not\subset X$. Por ejemplo, sea $A = \{a,b\}$. Entonces $P(A) = \{\emptyset,\{a\},\{b\},\{a,b\}\}$ y $\{a\} \not\subset \{b\}$ y $\{b\} \not\subset \{a\}$.

c) Considerando el conjunto N de los naturales ordenado por la relación "divide", N no es totalmente ordenado.

Los conjuntos totalmente ordenados se llaman también linealmente ordenados o cadenas.

En un conjunto ordenado A se dice que un elemento $p \in A$ es primer elemento de A si $p \leq x$, cualquiera sea $x \in A$. Dualmente se define último elemento.

EJERCICIOS:

1. Demostrar que si un conjunto ordenado tiene primer elemento éste es único. Idem para último elemento.

2. Sea $A = \{x \in N : x \leq 9\}$ y la relación "divide" en A. Escribirla como subconjunto de $A \times A$ y representarla gráficamente.

Diagramas de HASSE: Sirven para visualizar gráficamente los conjuntos ordenados destacando el orden entre los elementos. Son especialmente útiles en el caso de conjuntos finitos.

Dado un conjunto ordenado el diagrama de Hasse correspondiente se construye de acuerdo con las siguientes reglas:

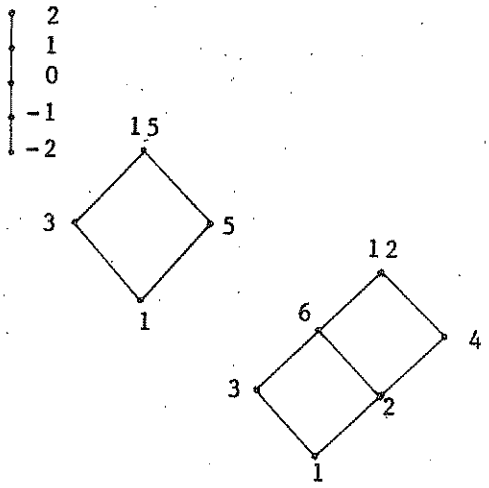
1) Se dice que un elemento b sigue a otro a si $a < b$ y no existe ningún x tal que $a < x < b$.
2) Cada elemento del conjunto se representa por un punto que se llama el afijo de ese elemento.
3) Si un elemento b sigue a otro a, el afijo de b se ubica encima del de a y se unen por un segmento.

EJEMPLOS:

1) Sea $A = \{x \in Z : |x| \leq 2\}$ ordenado por la relación menor o igual habitual.

2) Sea $A = \{x \in N : x \text{ divide a } 15\}$ ordenado por la relación "divide".

3) $A = \{x \in N : x \text{ divide a } 12\}$ ordenado por la relación "divide".



EJERCICIOS.

1. a) Dar cuatro ejemplos familiares de relaciones binarias.
   b) Sean $A = \{a,b\}$ y $B = \{1,2\}$. Indicar todas las relaciones binarias posibles entre A y B y entre B y A y representarlas gráficamente.
   c) Sea $A = \{x \in N : x \leq 10\}$ y consideremos la relación xRy si y solo si $|x-y| = 3$. Escribirla como subconjunto de $A \times A$ y representarla gráficamente.
   d) Idem para $A = \{x \in Z : |x| \leq 4 , x \neq 0\}$ y la relación "divide" en A.
   e) Representar gráficamente la relación "mayor o igual" entre números reales. Idem para la relación "mayor".

2. ¿Qué quiere decir que una relación $R \subset A \times A$ sea reflexiva, simétrica, transitiva o antisimétrica?. Indicar qué propiedades de las anteriores verifica cada una de las siguientes relaciones:

   a) A = conjunto de todas las personas, xRy si y solo si x es padre de y.
   b) A = " " " " " , xRy " x es compatriota de y.
   c) A = " " " " " , xRy " x es 1 cm. más alto que y.
   d) A = conjunto de las rectas de un plano, xRy si y solo si x es paralela a y.
   e) A = " " " " " " " , xRy " x es perpendicular a y.
   f) $A = Z$, xRy si y solo si $x \leq y$.
   g) $A = N$, xRy " x divide a y.
   h) $A = Z$, xRy " x divide a y.
   i) $A = Q$, xRy " $x^2 = y^2$.
   j) $A = Q$, xRy " $x^3 = y^3$.
   k) $A = N$, xRy " $m.c.d.(x,y) = 2$.
   l) $A = \{a,b,c,d,e,f\}$, $R \subset A \times A$ la siguiente relación:

$$R = \{(a,a),(c,c),(a,c),(c,a),(a,b),(b,a),(b,c)\}$$

3. Indicar qué relaciones del ejercicio anterior son relaciones de equivalencia y cuáles son de orden. En el caso de las relaciones de equivalencia hallar las clases

de equivalencia.

4. Decir cuáles de las siguientes relaciones son de equivalencia en el conjunto A y en caso afirmativo indicar las clases de equivalencia y el conjunto cociente.

   a) A = conjunto de todas las personas, $a \sim b$ si y solo si a y b tienen el mismo padre.

   b) A = conjunto de todas las personas, $a \sim b$ si y solo si a y b tienen el mismo grupo sanguíneo.

   c) A = conjunto de todas las personas, $a \sim b$ si y solo si a vive a no más de 100 km. de b.

   d) $A = R$, $a \sim b$ si y solo si $|a| = |b|$

   e) $A = R^2$, $(x_1,x_2) \sim (y_1,y_2)$ si y solo si $x_1 = y_1$

   f) $A = R^2$, $(x_1,x_2) \sim (y_1,y_2)$ si y solo si $x_1^2 + x_2^2 = y_1^2 + y_2^2$

   g) $A = Z$, $a \sim b$ si y solo si $a^3 = b^3$

   h) $A = Z$, $a \sim b$ " a-b es múltiplo de 8.

1.4. FUNCIONES. Hablando intuitivamente, una función o aplicación de un conjunto A en otro conjunto B es una correspondencia que a cada elemento de A le asocia un elemento de B. Si con f designamos a la función, el elemento $y \in B$ que corresponde al elemento $x \in A$ se llama la imagen de x por f o valor que toma f en x y se escribe $y = f(x)$.

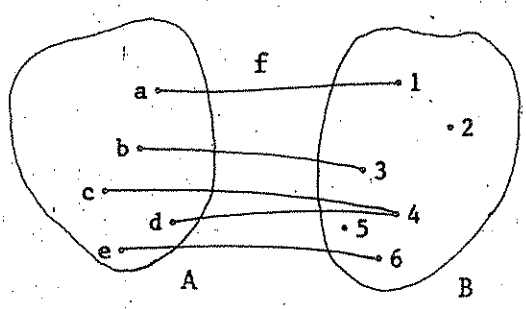
Claro que ésta no es una definición rigurosa puesto que no se ha precisado el significado del término correspondencia. La noción intuitiva de aplicación como una correspondencia o regla que permite pasar de un conjunto a otro se formaliza en la siguiente

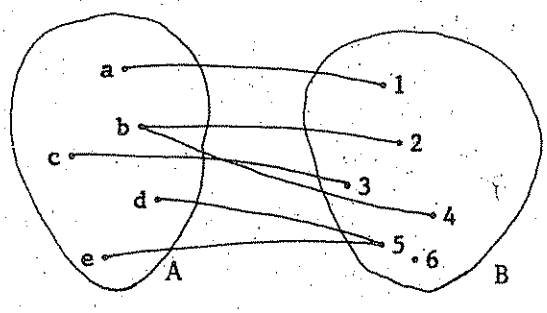
Definición. Dados dos conjuntos no vacíos A y B una aplicación de A en B es una relación f tal que para cada elemento $x \in A$ existe uno y solo un elemento $y \in B$ tal que x f y.

En lugar de x f y se escribe $f(x) = y$. Para indicar que f es una aplicación de A en B se escribe:

$$f: A \longrightarrow B \qquad \text{ó} \qquad A \xrightarrow{f} B$$

Por ejemplo, de las dos relaciones representadas en el dibujo, la primera es una aplicación de A en B porque a cada elemento de A le corresponde un único elemento en B, mientras que la segunda no lo es porque al elemento $b \in A$ le corresponden 2 y 4 $\in B$.





Se tiene: $f(a) = 1$; $f(b) = 3$; $f(c) = 4$; $f(d) = 4$; $f(e) = 6$.

Dada una función f: A $\longrightarrow$ B se llama imagen de f o rango de f al conjunto de todos los elementos de B que son imagen de los elementos de A.

$$\text{Im } f = \{y \in B : \exists\, x \in A \text{ tal que } y = f(x)\}$$

También se escribe Im f = f(A). El conjunto A se llama el dominio de f y B el codominio.

Por ejemplo, la imagen de la función representada en el dibujo anterior es

$$\text{Im } f = \{1,3,4,6\}$$

Se dice que una aplicación f: A $\longrightarrow$ B es:

1) EPIYECTIVA o que f es de A sobre B si Im f = B.

2) INYECTIVA o que f es biunívoca de A en B si elementos diferentes de A tienen imágenes diferentes en B, o sea si

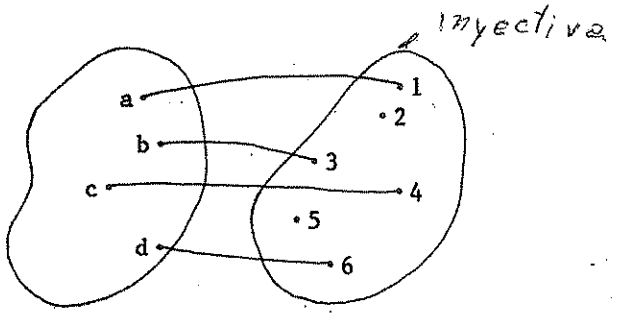
$$x \neq x' \Longrightarrow f(x) \neq f(x') \text{ , o lo que es equivalente}$$

$$f(x) = f(x') \Longrightarrow x = x'$$

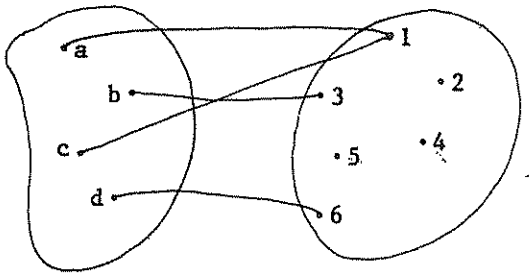
3) BIYECTIVA o que f es una correspondencia biunívoca de A sobre B si es epiyectiva e inyectiva a la vez.

EJEMPLOS:

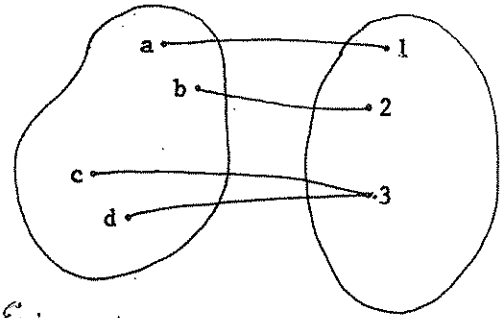
1) Consideremos las funciones representadas en los diagramas siguientes:



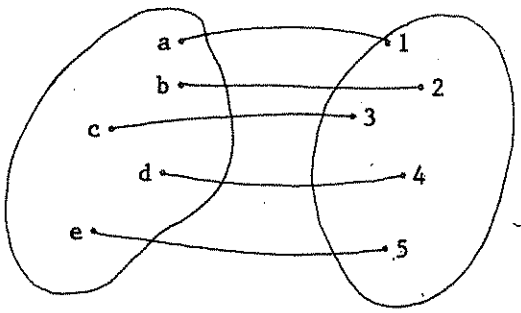
Inyectiva



ni Epiy. ni Iny.



Epiyectiva pero no inyetiva



Es biyectiva ↓ Epiyectiva Inyectiva

La primera es inyectiva pero no es epiyectiva, la segunda no es inyectiva ni epiyectiva la tercera es epiyectiva pero no es inyectiva y la cuarta es biyectiva.

2) Sea A un conjunto no vacío. La aplicación f: A $\longrightarrow$ A definida por f(x) = x, cualquiera sea x $\in$ A, se llama la aplicación idéntica de A y se representa I_A. Es claro que es biyectiva. Dibujar el gráfico de esta función cuando A = R.

3) Sean A y B dos conjuntos no vacíos, $b \in B$ un elemento fijo. Definiendo $f:A \longrightarrow B$ tal que $f(x) = b$ para todo $x \in A$ se tiene una aplicación que se llama constante (porque toma valor constante b en todos los puntos de A). Se tiene $\text{Im } f = \{b\}$. Se ve que no es inyectiva si A tiene más de un elemento y no es epiyectiva si B no se reduce a $\{b\}$. Dibujar el gráfico de una tal función cuando $A = B = R$.

4) Sean $A = \{a,b\}$ y $B = \{1,2,3\}$. Todas las funciones posibles de A en B son las indicadas en el cuadro siguiente:

	a	b
f_1	1	1
f_2	2	2
f_3	3	3
f_4	1	2
f_5	2	1
f_6	1	3
f_7	3	1
f_8	2	3
f_9	3	2

Ninguna de ellas es epiyectiva y son inyectivas f_4, f_5, f_6, f_7, f_8 y f_9.

5) Sea $f: R \longrightarrow R$ definida como sigue: $f(x) = x + 2$, $\forall x \in R$. $\text{Im } f = R$ pues un número real cualquiera y es imagen del número $x = y - 2$: $f(y-2) = y$. Luego f es epiyectiva. Además es inyectiva porque $x \neq x'$ implica $f(x) \neq f(x')$. Por lo tanto f es biyectiva. Dibujar el gráfico de f. En general, la aplicación $f: R \longrightarrow R$ definida por $f(x) = mx + n$, donde m y n son dos números reales fijos, es biyectiva si y solo si $m \neq 0$. ¿Cuál es la representación gráfica de una función de este tipo?

6) Sea $f: N \longrightarrow N$ tal que $f(x) = x + 2$, $\forall x \in N$. f es inyectiva pero no epiyectiva. Representar a f gráficamente.

7) Sea $g: Z \longrightarrow Z$ tal que $g(x) = x^2$ $\forall x \in Z$. No es epiyectiva porque su imagen es el conjunto de los enteros cuadrados perfectos y tampoco es inyectiva porque los enteros de igual valor absoluto tienen la misma imagen. Por ejemplo $g(2) = g(-2)$. Si consideramos $g: N \longrightarrow N$ definida como antes, g es inyectiva y no es epiyectiva.

8) Sea $f: Q^2 \longrightarrow Q$ definida por $f(x,y) = x + y$. f es epiyectiva pues $\text{Im } f = Q$ dado que todo número racional puede escribirse como la suma de otros dos, pero no es inyectiva pues esta descomposición no es única.

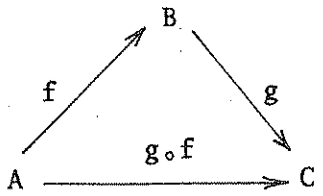
Igualdad de aplicaciones. De la definición de aplicación resulta:

Si f y g son dos aplicaciones de un conjunto A en otro B

$$f = g \text{ si y solo si } f(x) = g(x) \quad \forall x \in A$$

Composición de aplicaciones. Dados tres conjuntos A, B, C, una aplicación $f: A \longrightarrow B$ y otra $g: B \longrightarrow C$, se obtiene una aplicación de A en C haciendo corresponder a cada $x \in A$ el elemento $g(f(x)) \in C$. Esta aplicación se llama la aplicación compuesta de f y g y se nota $g \circ f$.

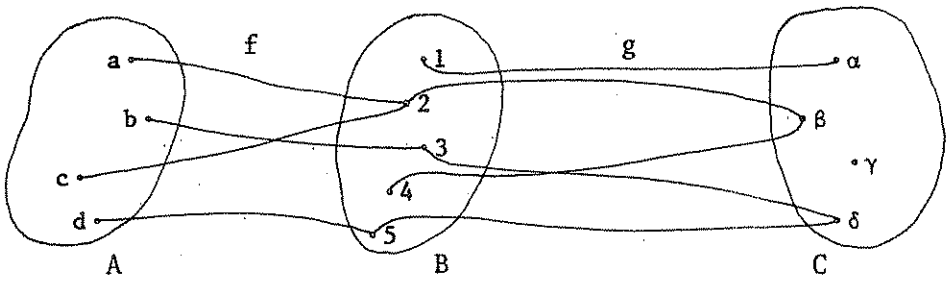
$$(g \circ f)(x) = g(f(x)) \ , \quad \forall \ x \in A$$



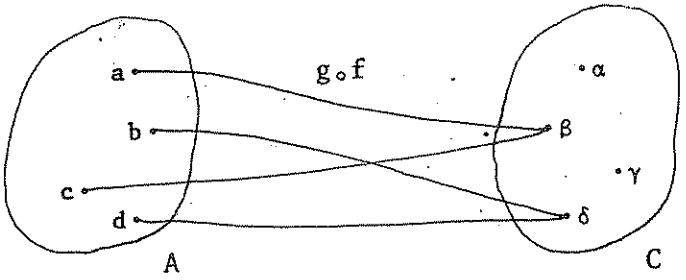
Así dadas dos funciones $A \xrightarrow{f} B \xrightarrow{g} C$ se obtiene una de A en C.

EJEMPLOS:

1) Sean $f: A \longrightarrow B$ y $g: B \longrightarrow C$ las aplicaciones representadas en el siguiente dibujo:



Entonces la aplicación compuesta $g \circ f$ es:



2) Sean $f: Z \longrightarrow Z$, $f(x) = x^3 \quad \forall \ x \in Z$

$g: Z \longrightarrow R$, $g(x) = 1 + \sqrt{|x|} \quad \forall \ x \in Z$

Entonces $g \circ f: Z \longrightarrow R$ es $(g \circ f)(x) = 1 + \sqrt{|x^3|}$

3) Si $f: N \longrightarrow Q$ es $f(x) = \frac{x}{x+1}$, $\forall \ x \in N$ y $g: Q \longrightarrow R$ es $g(x) = \text{sen } 2x$, $\forall \ x \in Q$, entonces $g \circ f: N \longrightarrow R$ está definida por $(g \circ f)(x) = \text{sen } \frac{2x}{x+1}$, $\forall \ x \in N$.

Asociatividad de la composición de aplicaciones.

TEOREMA 1.3. La composición de aplicaciones es asociativa. Es decir, dadas tres aplicaciones

$$A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D$$

es $\qquad h \circ (g \circ f) = (h \circ g) \circ f$

Demostración: Hay que probar que

$$[h \circ (g \circ f)](x) = [(h \circ g) \circ f](x) \quad \text{cualquiera sea } x \in A$$

Por definición de composición se tiene:

$$[h\circ(g\circ f)](x) = h[(g\circ f)(x)] = h[g(f(x))] = (h\circ g)(f(x)) = [(h\circ g)\circ f](x) \quad , \forall\ x \in A.$$

TEOREMA 1.4. Sean $f: A \longrightarrow B$, I_A la aplicación idéntica de A, I_B la aplicación idéntica de B. Entonces

$$I_B \circ f = f \quad , \quad f \circ I_A = f$$

Demostración: A cargo del lector.

Aplicación inversa de una biyectiva. Sea $f: A \longrightarrow B$ una función biyectiva. Entonces para cada $y \in B$ existe un único $x \in A$ tal que $f(x) = y$. Se puede definir así una función $g: B \longrightarrow A$ poniendo $g(y) = x$ si $f(x) = y$. Se ve sin dificultad que g también es biyectiva y que $g\circ f = I_A$ y $f\circ g = I_B$. g se llama la función inversa de f y se suele representar f^{-1}.

Por ejemplo, si $A = \{a,b,c,d\}$, $B = \{1,2,3,4\}$ y $f: A \longrightarrow B$ es tal que $f(a) = 3$, $f(b) = 1$, $f(c) = 4$, $f(d) = 2$, entonces $f^{-1}: B \longrightarrow A$ está definida como sigue:

$$f^{-1}(1) = b \ , \ f^{-1}(2) = d \ , \ f^{-1}(3) = a \ , \ f^{-1}(4) = c \ .$$

Si $f: R \longrightarrow R$ es $f(x) = 2x + 5 \ \forall\ x \in R$, entonces $f^{-1}: R \longrightarrow R$ está definida por

$$f^{-1}(x) = \frac{x-5}{2} \quad , \quad \forall\ x \in R \ .$$

Conjuntos coordinables, finitos, infinitos, numerables.

En matemática se trabaja continuamente con conjuntos infinitos. Para comparar "la magnitud o tamaño" de dos conjuntos finitos se cuentan sus elementos pero este sistema no se puede aplicar para comparar infinitos. Es natural querer extender la noción de "número de elementos de un conjunto finito" a los conjuntos infinitos.

Hasta fines del siglo pasado la noción de infinito y la aritmética del infinito era oscura y estaba repleta de paradojas. Fue Cantor (1845-1918) quien construyó la aritmética de los números transfinitos, generalizando las leyes y propiedades que valen en la aritmética finita.

Daremos tan sólo unas pocas definiciones y ejemplos. El lector puede ampliar este párrafo leyendo el capítulo XII de Algebra moderna, de Birkhoff y Mac Lane.

Para comparar "la magnitud" de dos conjuntos diferentes la idea fundamental es la coordinabilidad.

Definición. Un conjunto A se dice coordinable o equipotente con otro B si existe una aplicación biyectiva de A sobre B. Se escribe $A \approx B$.

La coordinabilidad de conjuntos es reflexiva simétrica y transitiva (1).

(1) No se puede considerar que existe el conjunto de todos los conjuntos porque esto conduce a contradicciones. Por eso que no se puede decir que " $\approx$ " es una relación de equivalencia, porque no es una relación, no está definida en ningún conjunto.

Se dice que dos conjuntos A y B tienen el mismo número cardinal si y solo si $A \approx B$, es decir, si y solo si los elementos de A pueden ponerse en correspondencia biunívoca con los de B. Esta definición coincide en el caso de los conjuntos finitos con la de igualdad del número de elementos.

EJEMPLOS.

1) Los conjuntos $A = \{a,b,c,d\}$ y $B = \{1,2,3,4\}$ son coordinables. Cuando se dice que un conjunto A tiene n elementos se está expresando que A tiene el mismo número cardinal que el conjunto $\{1,2,3,.....,n\}$.

2) El conjunto N de los números naturales tiene el mismo cardinal que el subconjunto 2N de los naturales pares. En efecto, la aplicación $f: N \longrightarrow 2N$ definida por $f(x) = 2x \quad \forall \; x \in N$ es biyectiva.

$$\begin{array}{ccccccccc} 1 & 2 & 3 & 4 & \ldots\ldots\ldots & n & \ldots\ldots\ldots\ldots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & & \updownarrow & \\ 2 & 4 & 6 & 8 & \ldots\ldots\ldots & 2n & \ldots\ldots\ldots\ldots \end{array}$$

3) El conjunto Z de los enteros es coordinable con N. Definamos $f: Z \longrightarrow N$ como sigue:

$$f(x) = \begin{cases} 2x + 1 & \text{si } x \geq 0 \\ 2|x| & \text{si } x < 0 \end{cases}$$

Se verifica sin dificultad que f es biyectiva.

Entonces, trabajando con conjuntos infinitos no podemos valernos de la intuición puesto que, por ejemplo, aunque 2N es un subconjunto propio de N y N es un subconjunto propio de Z, los números naturales pares, naturales y enteros están en correspondencia biunívoca.

Definición. Un conjunto es infinito si es coordinable con algún subconjunto propio. Un conjunto que no es infinito se dice finito.

Es decir, un conjunto es finito si y solo si no es coordinable con ninguno de sus subconjuntos propios.

Los conjuntos coordinables con N reciben un nombre especial.

Definición. Un conjunto se dice numerable si es coordinable con N.

Son numerables, por ejemplo, el conjunto Z de los enteros y el conjunto Q de los racionales. En el próximo capítulo veremos que Q tiene el mismo cardinal que N.

No todos los conjuntos infinitos son numerables. Por ejemplo, el conjunto R de los números reales, el conjunto de los puntos de una recta, el conjunto de los puntos de un segmento cualquiera de recta, el conjunto de los números irracionales, son conjuntos no numerables. Todos ellos son coordinables entre sí, es decir tienen el mismo número cardinal y se dice que tienen la potencia del continuo.

Pero no se crea que sólo hay estos dos tipos de conjuntos infinitos. Dada una colección de conjuntos infinitos no coordinables entre sí siempre existe otro conjunto infinito no

coordinable con ninguno de ellos, es decir que tiene un número cardinal diferente del de todos los dados.

EJERCICIOS.

1. a) Dada la función $f: R \longrightarrow R$ definida por $f(x) = x^3 - 2x + 1$ hallar $f(-1)$, $f(0)$ $f(1/2)$, $f(\sqrt[3]{2})$, $f(-2.3)$, $f(0.666....)$.

   b) Sea $f: Q \times Q \longrightarrow Q$ definida por $f(x,y) = \frac{x + y}{x^2 + 3}$. Hallar $f(2,-1)$, $f(-1/2,0)$, $f(1.5,-1/3)$, $f(0,-2/5)$.

2. Dados los conjuntos $A = \{a,b,c\}$ y $B = \{1,2\}$ ¿cuántas aplicaciones diferentes se pueden definir de A en B y cuáles son? ¿Cuántas son epiyectivas y cuántas inyectivas

3. Dadas las siguientes funciones, hallar sus imágenes y decir si son inyectivas, epiyectivas o biyectivas justificando la respuesta. En el caso de las biyectivas hallar la función inversa:

   a) $f: N \longrightarrow N$, $f(x) = 2x$
   b) $f: Q \longrightarrow Q$, $f(x) = 2x$
   c) $f: Z \longrightarrow Z$, $f(x) = |x|$
   d) $f: N \longrightarrow N$, $f(x) = x + 5$
   e) $f: Z \longrightarrow Z$, $f(x) = x + 5$
   f) $f: Z \longrightarrow N$, $f(x) = x^2$
   g) $f: R \longrightarrow R^+$, $f(x) = x^2$ siendo $R^+ = \{x \in R : x \geq 0\}$
   h) $f: N \longrightarrow N$, $f(x) = 3x - 1$
   i) $f: R \longrightarrow R$, $f(x) = 3x - 1$
   j) $f: R \longrightarrow R$, $f(x) = x^3$
   k) $f: N \longrightarrow Q$, $f(x) = \frac{x}{x+1}$
   l) $f: R^* \longrightarrow R^*$, $f(x) = \frac{1}{x}$ siendo $R^* = \{x \in R : x \neq 0\}$

4. a) Dadas las aplicaciones $f: A \longrightarrow B$ y $g: B \longrightarrow C$ definir la aplicación composición $g \circ f$. ¿Qué significa que la composición de aplicaciones es asociativa?.

   b) Sean $A = \{a,b,c,d,e\}$, $B = \{0,1,2,3\}$ y $C = \{v,w,x,y,z\}$ y f y g las siguientes funciones:

$f(a) = 0$	$g(0) = w$
$f(b) = 2$	$g(1) = x$
$f(c) = 3$	$g(2) = y$
$f(d) = 2$	$g(3) = y$
$f(e) = 3$	

   Hallar $g \circ f$ y representarla gráficamente.

   c) Dadas $f: N \longrightarrow Q$ definida por $f(x) = \frac{x}{x+2}$

   $g: Q \longrightarrow R$ " " $g(x) = x^2 + \sqrt{2}$
   $h: R \longrightarrow R$ " " $h(x) = |x| + 1$
   $t: R \longrightarrow R$ " " $t(x) = \text{sen } 3x$

Hallar $g \circ f$, $t \circ h$, $h \circ t$, $h \circ g$, $t \circ f$ y calcular $(g \circ f)(2)$, $(t \circ h)(-\pi)$, $(h \circ t)(-\pi)$, $(h \circ g)(-1/2)$, $(t \circ f)(1)$.

d) Verificar que $(h \circ g) \circ f = h \circ (g \circ f)$ y que $(t \circ h) \circ g = t \circ (h \circ g)$.

e) ¿Se pueden componer h con g, t con f, g con f?.

5. Si una persona recibe un peso el primer día del año, dos el segundo, cuatro el tercero, ocho el cuarto, y continúa así recibiendo cada día el doble de la cantidad obtenida el día anterior, escribir una fórmula para la función que a cada día le hace corresponder la suma percibida ese día. ¿En qué día recibirá aproximadamente 1.000.000 de pesos?.

6. a) Dar un ejemplo de una función definida de R en R que tenga la propiedad de que cualesquiera sean $a,b \in R$, si $a < b$ entonces $f(a) < f(b)$.

b) Idem de una función g de R en R tal que cualesquiera sean $a,b \in R$, si $a < b$ entonces $g(b) < g(a)$.

7. Sean $A \xrightarrow{f} B \xrightarrow{g} C$. Demostrar que:

1) Si f y g son epiyectivas , $g \circ f$ es epiyectiva.
2) Si f y g son inyectivas , $g \circ f$ es inyectiva.
3) Si f y g son biyectivas , $g \circ f$ es biyectiva.
4) Si $g \circ f$ es epiyectiva, g es epiyectiva.
5) Si $g \circ f$ es inyectiva, f es inyectiva.

Las propiedades recíprocas de las 1) , 2) , 3) , 4) y 5) no son en general verdaderas. Queda a cargo del lector encontrar ejemplos que lo prueben.

1.5. OPERACIONES BINARIAS.

Definición. Una operación binaria definida en un conjunto A es una aplicación de $A \times A$ en A.

Es decir, una operación binaria (o también ley de composición interna) en A es una aplicación que a cada par ordenado (x,y) de elementos de A hace corresponder un elemento de A.

En lugar de usar la notación indicada para las aplicaciones, habitualmente el elemento que corresponde al par (x,y) por una operación se representa escribiendo los elementos x,y uno a continuación del otro separados por un signo característico que simboliza a dicha operación. Por ejemplo, si el signo elegido es T con xTy se representa al elemento que corresponde al par (x,y):

$$(x,y) \longrightarrow x \, T \, y$$

Si $z \in A$ es el elemento que corresponde al par (x,y) se dice que z es el resultado de operar x con y y se escribe $xTy = z$.

De la definición de operación binaria resulta inmediatamente que:

1) Una operación binaria $\top$ en A está definida para todo par (x,y) de elementos de A.

2) Cualquiera sea el par (x,y), $x \top y \in A$. Esto se expresa diciendo que el conjunto A es cerrado con respecto a la operación.

3) Toda operación binaria $\top$ tiene la siguiente propiedad:

   Si $x = x'$ e $y = y'$ entonces $x \top y = x' \top y'$. (Propiedad uniforme).

   En efecto, si $x = x'$ e $y = y'$ entonces $(x,y) = (x',y')$.

Como por definición la operación es una aplicación de $A \times A$ en A es claro que $x \top y = x'$

EJEMPLOS.

1) La suma $(x,y) \longrightarrow x+y$ y la multiplicación $(x,y) \longrightarrow x.y$ son operaciones binarias e N , Z , Q , R y C.

2) Si E es un conjunto, la intersección $(X,Y) \longrightarrow X \cap Y$ y la reunión $(X,Y) \longrightarrow X \cup Y$ son operaciones binarias en el conjunto $A = P(E)$ de las partes de E.

3) La diferencia $(x,y) \longrightarrow x-y$ es una operación binaria en los conjuntos Z , Q , R y C En cambio no lo es en N puesto que no está definida para los pares de números naturales (x,y) tales que $x \leq y$.

4) La división $(x,y) \longrightarrow x/y$ no es una operación binaria en ninguno de los conjuntos numéricos porque la división por cero no está definida.

5) Si F es el conjunto de todas las aplicaciones de un conjunto A en sí mismo, la composición de aplicaciones $(f,g) \longrightarrow f \circ g$ es una operación binaria en F.

Operaciones asociativas, conmutativas, distributivas.

Notemos que en un conjunto en el que está definida una operación binaria $\top$ se puede hablar de la composición de dos elementos pero en general no se podrá hablar sin ambigüedad de la composición de tres elementos a,b,c pues el resultado puede variar según se operen a y b primero y luego este resultado con c o en cambio se opere a con el resultado de b y c. Esta ambigüedad se salva si ambos resultados son iguales:

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

Además observemos que en general $a \top b \neq b \top a$.

Definición. Una operación binaria $\top$ definida en un conjunto A se dice:

1) Asociativa si $(x \top y) \top z = x \top (y \top z)$ cualesquiera sean $x,y,z \in A$.
2) Conmutativa si $x \top y = y \top x$ cualesquiera sean $x,y \in A$.

Si una operación es asociativa se escribe simplemente $x \top y \top z$ puesto que el resultado no depende del orden en que se efectúen las operaciones. Más aún, se demuestra que la asociatividad de una operación binaria permite hablar sin ambigüedad del resultado de operar un número finito cualquiera de elementos de A dados en un cierto orden pues el resultado es independiente de la distribución de paréntesis que se considere y por eso se omite escribirlos. No ocurre así si la operación en cuestión no es asociativa.

Si en un conjunto A están definidas dos operaciones binarias $\top$ y $\circ$ se dice que:

$\top$ es distributiva a izquierda con respecto a $\circ$ si

$$x \top (y \circ z) = (x \top y) \circ (x \top z) \quad \text{cualesquiera sean } x,y,z \in A.$$

T es distributiva a derecha con respecto a o si

$$(x \circ y) \top z = (x \top z) \circ (y \top z) \quad \text{cualesquiera sean } x,y,z \in A.$$

T se dice distributiva con respecto a o si es distributiva a derecha e izquierda.

Una operación binaria puede ser o no conmutativa, asociativa o distributiva con respecto a otra operación.

EJEMPLOS.

1) La suma y la multiplicación en N , Z , Q , R o C son conmutativas y asociativas y la multiplicación es distributiva con respecto a la suma, no así la suma con respecto a la multiplicación. La diferencia en Z (Q,R o C) es una operación que no es asociativa ni conmutativa.

2) La reunión e intersección de conjuntos son operaciones asociativas y conmutativas y cada una de ellas es distributiva con respecto a la otra.

3) La composición de aplicaciones es asociativa pero no conmutativa.

4) Consideremos la operación binaria definida en el conjunto R de los números reales de la siguiente manera:

$$x \circ y = x^2 + y^2$$

o es conmutativa pero no asociativa.

5) Sea T la operación binaria definida en R como sigue:

$$x \top y = x$$

T es asociativa y no es conmutativa; es distributiva a la izquierda con respecto a la operación del ejemplo 4, pero no lo es a la derecha.

6) La operación definida en Z como sigue:

$$x \top y = x^2y + x + y$$

no es asociativa ni conmutativa.

Elemento neutro.

Definición. Dada una operación binaria T en un conjunto A, un elemento $e \in A$ se llama neutro con respecto a esa operación si

$$x \top e = e \top x = x \quad \text{cualquiera sea } x \in A.$$

Es claro que si existe, el neutro es único.

En efecto, si e y e' son dos elementos neutros para la operación T entonces:

$e \top e' = e'$ por ser e neutro.
$e \top e' = e$ " " e' neutro.

Luego $e = e'$.

EJEMPLOS.

1) El 0 es elemento neutro para la suma en Z , Q , R y C y el 1 es neutro para la multiplicación en N , Z , Q , R y C. En N la suma no tiene neutro.

2) Considerando las operaciones de reunión e intersección de conjuntos en $P(E)$, el conjunto $\emptyset$ es neutro para la reunión y el conjunto E lo es para la intersección.

3) Considerando la composición de aplicaciones en el conjunto F de todas las aplicaciones de un conjunto A no vacío en sí mismo, el neutro de esta operación es la aplicación idéntica de A.

4) La operación del ejemplo 4 anterior no admite neutro; la definida en el ejemplo 5 tampoco porque 0 es neutro a derecha pero no a izquierda; finalmente la del ejemplo 6 tiene neutro: el 0.

Inverso.

Definición. En un conjunto A con una operación binaria τ que admite elemento neutro, se dice que un elemento $b \in A$ es inverso de otro $a \in A$ si

$$a \tau b = b \tau a = e$$

Si un elemento a tiene inverso se dice que a es inversible.

EJEMPLOS.

1) En Z , Q , R y C todo número x tiene inverso con respecto a la suma: -x. En Q , R y C todo número x distinto de 0 tiene inverso con respecto a la multiplicación: x^{-1}

   En Z los únicos números inversibles con respecto a la multiplicación son 1 y -1.

2) En el conjunto F de todas las aplicaciones de un conjunto A en sí mismo, los elementos de F que tienen inverso con respecto a la composición de aplicaciones son las aplicaciones biyectivas.

3) En un conjunto A con una operación que tiene neutro e, éste es su propio inverso. Puede suceder que e sea el único elemento inversible de A. Por ejemplo, en el conjunto N de los naturales 1 es el neutro de la multiplicación y es el único natural inversible en N con respecto a esta operación.

PROPOSICION. En un conjunto A con una operación binaria que es asociativa y admite neutro, si un elemento $x \in A$ tiene inverso éste es único.

Demostración: A cargo del lector.

NOTA. Hemos visto algunos ejemplos de conjuntos en los que están definidas operaciones binarias e indicado algunas propiedades que una operación binaria puede o no verificar. En cada uno de los sistemas numéricos Z , Q , R y C están definidas dos operaciones, la

suma y la multiplicación, que son asociativas y conmutativas y tal que la multiplicación es distributiva con respecto a la suma. Además en todos estos sistemas la suma y la multiplicación admiten neutro, todo elemento es inversible con respecto a la suma en Z , Q , R o C y todo elemento no nulo es inversible con respecto a la multiplicación en Q , R o C. No sólo los sistemas numéricos presentan propiedades de este tipo, sino que son numerosos los ejemplos particulares de conjuntos en los que están definidas una o dos operaciones que verifican propiedades similares y que por lo tanto, dejando de lado la naturaleza de los elementos, presentan una semejanza formal en cuanto a las leyes de cálculo. La observación de este hecho llevó a estudiar las estructuras algebraicas fundamentales que intervienen en todas las ramas de la matemática, analizando las leyes y propiedades que verifican distintas operaciones, con abstracción de la naturaleza de los elementos sobre los que actúan. Este nuevo enfoque se inició a principios del siglo pasado con Ruffini (1765-1822), Abel (1802-1829) y Galois (1811-1832) quienes, estudiando el problema de la resolución de ecuaciones algebraicas, se valieron de la noción de grupo.

Del concepto general de operación binaria en un conjunto se pasa a los conceptos abstractos de grupo, anillo y cuerpo, que dependen exclusivamente de las propiedades de una o dos operaciones definidas en un conjunto no vacío, sin que interese la índole de sus elementos. Se establece que esas operaciones verifican ciertas propiedades, que se llaman axiomas o postulados, y luego se estudian las propiedades que se deducen a partir de ellas.

El carácter abstracto de estos conceptos hace que se los pueda aplicar a casos particulares (números, matrices, vectores, traslaciones en el plano o en el espacio, permutaciones, funciones numéricas, etc.) y una vez que se ha comprobado que se verifica la definición de grupo, anillo o cuerpo, se sigue inmediatamente que se verifican también todas las demás propiedades deducidas a partir de los axiomas, siendo innecesario hacer la demostración en cada caso. Esta teoría general permite entonces una importante economía de esfuerzos y ayuda a aclarar las ideas.

Este tema no tiene cabida en el programa de Algebra I. Daremos breves definiciones de lo que se entiende por grupo, anillo y cuerpo tan sólo a título informativo..

GRUPO. Se llama grupo a todo conjunto en el que está definida una operación binaria que es asociativa, admite neutro y tal que todo elemento es inversible. Si además la operación es conmutativa el grupo se llama abeliano.

Por ejemplo: a) Son grupos abelianos con respecto a la suma Z , Q , R , C; b) Son grupos abelianos con respecto a la multiplicación $Q - \{0\}$, $R - \{0\}$, $C - \{0\}$; c) Si S_n = conjunto de todas las aplicaciones biyectivas del conjunto $\{1,2,\ldots,n\}$ sobre sí mismo, S_n es un grupo no abeliano con respecto a la composición de aplicaciones, para $n \geq 3$.

ANILLO. Un anillo (unitario) es un grupo abeliano en el que está definida una segunda

operación binaria que es asociativa, admite neutro y es distributiva con respecto a la operación grupal. Si además esta segunda operación es conmutativa el anillo se dice conmutativo.

Por ejemplo, son anillos conmutativos con respecto a la suma y a la multiplicación $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$. Las matrices cuadradas de orden n de números reales forman un anillo no conmutativo ($n > 1$). Ya veremos más adelante este ejemplo.

En un anillo, la operación de grupo se suele llamar suma y la otra multiplicación; el neutro de la suma se representa 0 y el de la multiplicación 1. Con esta notación, la definición de cuerpo es la siguiente.

CUERPO. Un cuerpo es un anillo conmutativo en el que todo elemento distinto de 0 es inversible con respecto a la multiplicación.

$\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$ son cuerpos con respecto a la suma y a la multiplicación ordinarias de números.

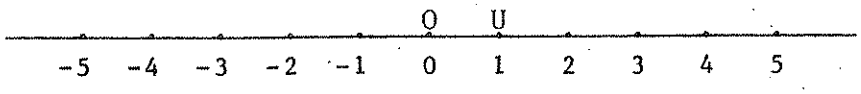
CAPITULO II

NUMEROS REALES

En el colegio secundario se estudian las propiedades de los números naturales, enteros, racionales y reales. Cada uno de estos sistemas numéricos es una ampliación del anterior y las sucesivas ampliaciones fueron motivadas por la necesidad de efectuar operaciones que no son posibles en el sistema numérico precedente. Históricamente el número natural es el primero en aparecer, originado en la operación de contar objetos. La necesidad de medir longitudes, superficies, pesos, etc. determinó la creación de los números fraccionarios positivos, de uso corriente en la antigüedad. En cambio los números negativos hacen su aparición muy tarde y sólo a principios del siglo XVII son aceptados por los matemáticos como entes similares a los números positivos. Los griegos nunca consideraron como solución de un problema una cantidad negativa. Finalmente, recién en el siglo pasado los números irracionales adquieren respetabilidad cuando en el período de revisión de los fundamentos y principios de la matemática Dedekind (1872), Cantor (1872) y Weierstrass (1869) elaboran teorías rigurosas para definir los números reales a partir de los racionales.

El alumno está familiarizado con la idea de que la imagen geométrica del conjunto R es una recta. Existe una correspondencia biunívoca entre R y los puntos de una recta, de modo que si a y b son números reales y a < b entonces el punto que corresponde a a está a la izquierda del que corresponde a b.

Los números racionales pueden representarse como sigue: Se elige una recta y sobre ella dos puntos distintos O y U, quedando así determinada una unidad de longitud, el segmento $\overline{OU}$, y un sentido positivo sobre la recta, el de la semirrecta $\overrightarrow{OU}$. El punto O representa al número 0 y el punto U al 1.



Los números enteros están representados por puntos equidistantes, como indica la figura

Para representar un número fraccionario positivo $\frac{a}{b}$ se divide al segmento unidad en "b" segmentos iguales y luego se transporta uno de ellos "a" veces a partir del origen sobre la semirrecta de la derecha. Si el número es negativo está representado por el punto simétrico del anterior con respecto al origen.

A los puntos de la recta que corresponden a números racionales se los llama puntos racionales.

En cualquier segmento AB de la recta existe un punto racional, o lo que es equivalente, entre dos puntos distintos A y B cualesquiera existen infinitos puntos racionales.

Sin embargo los puntos racionales, a pesar de estar tan próximos como se desee, no llenan la recta.

Hace 25 siglos los matemáticos pitagóricos lo descubrieron mediante una construcción bien sencilla.

En primer lugar no existe ningún número racional cuyo cuadrado sea 2. En efecto, supongamos que existen dos enteros a y b tales que

$$\left(\frac{a}{b}\right)^2 = 2$$

Podemos suponer $\frac{a}{b}$ irreducible, es decir a y b sin factores comunes propios.

De la igualdad anterior resulta

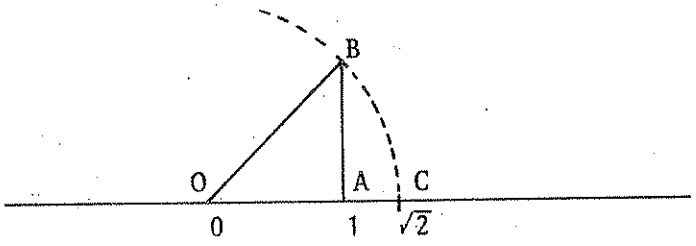
$$a^2 = 2b^2 \Longrightarrow a^2 \text{ es par} \Longrightarrow a \text{ es par.} \quad (1)$$

Luego podemos escribir a = 2c. Entonces

$$4c^2 = 2b^2 \Longrightarrow 2c^2 = b^2 \Longrightarrow b^2 \text{ es par} \Longrightarrow b \text{ es par.} \quad (2)$$

(1) y (2) contradicen la hipótesis de que $\frac{a}{b}$ es irreducible. Esta contradicción prueba que no existe ningún número racional cuyo cuadrado sea 2.

Es fácil ahora determinar un punto en la recta que no es racional.



En la figura, sea OA la unidad de longitud y $\overline{OA} = \overline{AB}$. Por el teorema de Pitágoras el cuadrado de la longitud de la hipotenusa $\overline{OB}$ es 2. Haciendo centro con el compás en O y trazando un arco de circunferencia con radio $\overline{OB}$ se obtiene el punto C que por lo recién demostrado no es racional.

Los griegos, que crearon los números racionales para poder medir las longitudes con números, descubrieron así que existían segmentos no susceptibles de ser medidos (con los números racionales), es decir inconmensurables. Esto provocó la ruina de la escuela pitagórica porque no acertaron a completar el sistema racional de modo que existiera una correspondencia biunívoca entre números y puntos de la recta. Para ello es necesario ampliar Q introduciendo los números irracionales.

Para estudiar los números reales se puede partir de los números naturales, definiéndolos en base a la teoría de conjuntos o introduciéndolos axiomáticamentes (por medio de los postulados de Peano (1889), por ejemplo), definir luego a partir de ellos los números enteros, construir a partir de éstos el sistema de los racionales y finalmente definir los números reales a partir de los racionales por el método de las cortaduras de Dedekind, o el de las sucesiones monótonas contiguas o también el de las sucesiones de Cauchy desarrollado por Méray y Cantor en 1872.

Una construcción rigurosa del sistema de los números reales escapa los límites de este curso de álgebra. Los lectores interesados pueden consultar la bibliografía indicada y

en especial el libro "Qué es la matemática" de Courant y Robbins.

El camino que seguiremos en cambio es enumerar las propiedades fundamentales del sistema R de los números reales, distinguiendo dentro de R el conjunto N de los números naturales, el conjunto Z de los enteros y el conjunto Q de los racionales.

Y al decir propiedades fundamentales de R queremos significar un conjunto de propiedades bien determinado que caracteriza al sistema R de los números reales, es decir tal que todas las propiedades de R sin excepción pueden deducirse a partir de ellas y todo sistema que verifica esas propiedades coincide necesariamente con R desde el punto de vista algebraico. Son las propiedades de cuerpo ordenado completo. El alumno está bien familiarizado con todas ellas, excepto tal vez con el llamado axioma de completidad, que dejaremos para el final.

2.1. NUMEROS REALES.

En el conjunto R de los números reales están definidas dos operaciones binarias, la suma + y la multiplicación . , y una relación binaria " < " de modo que se verifican las siguientes propiedades:

S1. Asociatividad de la suma:

$$(a + b) + c = a + (b + c) \qquad \forall\ a,b,c \in R$$

S2. Conmutatividad de la suma:

$$a + b = b + a \qquad \forall\ a,b \in R$$

S3. Existe $0 \in R$ que es neutro de la suma, es decir

$$a + 0 = a \qquad \forall\ a \in R$$

S4. Todo número real es inversible con respecto a la suma, es decir, para todo $a \in R$ existe $b \in R$, llamado simétrico de a, tal que $a + b = 0$.

M1. Asociatividad de la multiplicación:

$$(a.b).c = a.(b.c) \qquad \forall\ a,b,c \in R$$

M2. Conmutatividad de la multiplicación:

$$a.b = b.a \qquad \forall\ a,b \in R$$

M3. Existe $1 \in R$, $1 \neq 0$, que es neutro de la multiplicación, es decir $a.1 = a \quad \forall a \in R$

M4. Todo número real distinto de 0 es inversible con respecto a la multiplicación, es decir, para todo $a \in R$, $a \neq 0$, existe $b \in R$, llamado inverso de a, tal que $a.b = 1$

D . Distributividad de la . con respecto a la + :

$$a.(b + c) = a.b + a.c \qquad \forall\ a,b,c \in R$$

01. Ley de tricotomía: Cualesquiera sean $a,b \in R$ se verifica una y solo una de las siguientes relaciones:

$$a = b \quad , \quad a < b \quad \text{ó} \quad b < a$$

02. Ley transitiva: Si $a < b$ y $b < c$ entonces $a < c$.

03. Ley de monotonía de la suma: Si $a < b$ entonces $a + c < b + c$, $\forall\, c \in R$.

04. Ley de monotonía de la multiplicación: Si $a < b$ y $0 < c$ entonces $a.c < b.c$.

NOTACION: $a > b$ si y solo si $b < a$

$a \leq b$ si y solo si $a = b$ ó $a < b$

" $\leq$ " es una relación de orden.

Las propiedades anteriores se resumen diciendo que R es un cuerpo ordenado: cuerpo por que se verifican las propiedades S1 , S2 , S3 , S4 , M1 , M2 , M3 , M4 y D; ordenado porque está definida una relación de orden que verifica las propiedades 01,02,03 y 04.

A la lista de propiedades enunciadas falta agregarle una más, el llamado axioma de completidad, que veremos más adelante.

De las propiedades de cuerpo se deducen las siguientes, cuya demostración queda a cargo del lector:

1. Los elementos neutros de la suma y el producto son únicos.
2. Cada número real x tiene un único simétrico y un único inverso (si $x \neq 0$). El simétrico de x se representa $-x$ y su inverso x^{-1}.
3. Valen las leyes de cancelación de la suma y de la multiplicación, es decir:

$$x + a = x + b \text{ implica } a = b$$
$$x.a = x.b \text{ y } x \neq 0 \text{ implica } a = b$$

4. $a.0 = 0 \quad \forall\, a \in R$.
5. En R no hay divisores de cero, es decir

$$a.b = 0 \text{ implica } a = 0 \text{ ó } b = 0$$

6. Cualesquiera sean $a,b \in R$, la ecuación $b + x = a$ tiene solución única dada por

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

NOTACION: El elemento $a + (-b)$ se representa $a - b$ y se llama la diferencia de a y b:

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

7. Cualesquiera sean $a,b \in R$, $b \neq 0$, la ecuación $b.x = a$ tiene solución única dada por $x = a.b^{-1}$.

NOTACION: Dados $a,b \in R$, $b \neq 0$, el elemento $a.b^{-1}$ se llama el cociente de a por b y se representa $\frac{a}{b}$.

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

8. Valen las siguientes propiedades:

i) $-(-a) = a$

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

iii) $(a^{-1})^{-1} = a$ si $a \neq 0$

iv) $(a.b)^{-1} = a^{-1}.b^{-1}$ si $a \neq 0$, $b \neq 0$

v) $(-a)^{-1} = -(a^{-1})$ si $a \neq 0$

vi) $(-a).b = a.(-b) = -(a.b)$

vii) $(-a).(-b) = a.b$

viii) $a.(b-c) = a.b - a.c$

9. Los cocientes verifican las siguientes reglas:

i) $\frac{a}{b} = \frac{c}{d}$ si y solo si $a.d = b.c$

ii) $\frac{a}{b} \pm \frac{c}{d} = \frac{a.d \pm b.c}{b.d}$

iii) $\frac{a}{b} \cdot \frac{c}{d} = \frac{a.c}{b.d}$

iv) $-\frac{a}{b} = \frac{-a}{b} = \frac{a}{-b}$

v) $\frac{a}{b} = \frac{-a}{-b}$

vi) Si $\frac{a}{b} \neq 0$, $(\frac{a}{b})^{-1} = \frac{b}{a}$

Definición. Un número $a \in R$ se dice positivo si $0 < a$ y negativo si $a < 0$.

De la propiedad 01 resulta que todo número real es positivo, negativo o nulo, es decir R está dividido en tres subconjuntos disjuntos dos a dos: el de los números positivos, el de los negativos y el que tiene por único elemento a 0. De 02 sigue que todo número negativo es menor que cualquiera positivo.

De las propiedades de cuerpo ordenado resulta que en R valen también las siguientes:

10. La suma y el producto de dos números positivos es un número positivo.

11. $a < b$ si y solo si $b - a > 0$

12. $a > 0$ si y solo si $-a < 0$

13. El producto de dos números negativos es un número positivo y el de un positivo por un negativo es negativo.

14. El cuadrado de cualquier número no nulo es positivo. En particular $1 > 0$.

De aquí y de 12 sigue $-1 < 0$.

15. $a + c < b + c$ si y solo si $a < b$

16. Si $a \leq b$ y $c < d$ entonces $a + c < b + d$

17. Si $a < b$ y $c < 0$ entonces $b.c < a.c$

18. Si $a.c < b.c$ y $c > 0$ entonces $a < b$
    Si $a.c < b.c$ y $c < 0$ entonces $b < a$

19. $a > 0$ si y solo si $a^{-1} > 0$

20. Si a y b son ambos positivos o ambos negativos, $a < b$ si y solo si $b^{-1} < a^{-1}$

Valor absoluto.

Se llama valor absoluto de un número real a y se representa $|a|$ al número definido de la siguiente manera:

$$|a| = \begin{cases} a & \text{si } a \geq 0 \\ -a & \text{si } a < 0 \end{cases}$$

Propiedades.

1. $|a| \geq 0$; $|a| = 0$ si y solo si $a = 0$

2. $|a| \leq m$ si y solo si $-m \leq a \leq m$ $(m \geq 0)$

3. $|a + b| \leq |a| + |b|$

4. $||a| - |b|| \leq |a - b|$

5. $|a.b| = |a|.|b|$

6. $\left|\frac{a}{b}\right| = \frac{|a|}{|b|}$

Demostración:

1. Resulta de la definición de valor absoluto.

2. Supongamos $|a| \leq m$. Como $a \leq |a|$, por transitividad resulta $a \leq m$ (i).

   Como $-a \leq |a|$ por transitividad se tiene $-a \leq m$. Multiplicando por -1 resulta (pc propiedad 17) $-m \leq a$ (ii). De (i) y (ii) sigue $-m \leq a \leq m$.

   Recíprocamente, supongamos $-m \leq a \leq m$, $m \geq 0$. Si $a \geq 0$, $|a| = a \leq m$. Si $a < 0$ $|a| = -a$. De $-m \leq a$ sigue $-a \leq m$. Luego $|a| \leq m$.

3. Podemos escribir $|a| \leq |a|$. Aplicando la propiedad anterior con $m = |a|$, se tiene: $-|a| \leq a \leq |a|$. Análogamente $-|b| \leq b \leq |b|$. Sumando miembro a miembro se tiene: $-(|a| + |b|) \leq a + b \leq |a| + |b|$. Aplicando nuevamente 2 con $m = |a|+|b|$ resulta $|a + b| \leq |a| + |b|$.

4. Podemos escribir $a = b + (a - b)$. Aplicando la propiedad anterior es:

   $$|a| \leq |b| + |a - b|$$

   De aquí $|a| - |b| \leq |a - b|$ (i)

   Análogamente, de $b = a + (b - a)$ es $|b| \leq |a| + |b - a|$

   Luego $-|b - a| \leq |a| - |b|$ (ii)

   Pero $|b - a| = |a - b|$. De (i) y (ii) sigue

   $$-|a - b| \leq |a| - |b| \leq |a - b|$$

   Aplicando 2 con $m = |a - b|$ resulta $||a| - |b|| < |a - b|$.

5. Si $a = 0$ o $b = 0$ es obvio que $|a.b| = |a|.|b|$

Si $a \neq 0$ y $b \neq 0$ pueden ser ambos positivos, ambos negativos o uno positivo y otro negativo. Si $a > 0$ y $b > 0$ es $a.b > 0$ (propiedad 10). Luego $|a.b| = a.b = |a|.|b|$
Si $a < 0$ y $b < 0$ es $a.b > 0$ (propiedad 13). Luego $|a.b| = a.b$, $|a| = -a$, $|b| = -b$. Como $(-a).(-b) = a.b$ se tiene $|a.b| = |a|.|b|$.
Finalmente, si $a < 0$ y $b > 0$ es $a.b < 0$ (propiedad 13).

Luego $|a.b| = -(a.b) = (-a).b = |a|.|b|$.

6. $a = \frac{a}{b} . b$. Luego $|a| = |\frac{a}{b}|.|b|$ y de aquí resulta $|\frac{a}{b}| = \frac{|a|}{|b|}$

Vamos a distinguir ahora dentro de R los sistemas de los números naturales, de los enteros y de los racionales.

$$N \subset Z \subset Q \subset R$$

2.2. NUMEROS NATURALES

Para definir el conjunto N de los números naturales y teniendo en cuenta que $1 \in N$ y que si un número $x \in N$ también $x+1 \in N$, comenzaremos por considerar todos los subconjuntos de R que satisfacen estas propiedades.

Definición. Un subconjunto A de R se dice inductivo si satisface:

1. $1 \in A$
2. Si $x \in A$ entonces $x+1 \in A$

EJEMPLOS:

1) Es claro que R es inductivo.

2) El conjunto de todos los reales positivos es inductivo. (Por propiedades 14 y 10).

3) El conjunto $\{x \in R : x = 1 , x = 2 \text{ ó } x \geq 3\}$ es inductivo.

Ningún subconjunto finito de R es inductivo. En general, ningún subconjunto acotado superiormente lo es.

Es claro que los números naturales pertenecen a todos los conjuntos inductivos y ésta es la clave de la siguiente definición.

Definición. Se llama conjunto N de los números naturales a la intersección de todos los subconjuntos inductivos de R.

Resulta inmediatamente que N es inductivo y que todo conjunto inductivo contiene a N, es decir N es el menor de los subconjuntos inductivos de R (con respecto a la relación de inclusión). En particular, como los números positivos forman un conjunto inductivo, los números naturales son positivos.

NOTACION: Como $1 \in N$, también $1 + 1 \in N$. Se escribe $1 + 1 = 2$. Como $2 \in N$, $2 + 1 \in N$ Se escribe $2 + 1 = 3$, etc.

(1) Principio de inducción.

(2) Principio de buena ordenación.

(3) N es cerrado con respecto a la suma y a la multiplicación.

(4) Factorización única: Todo número natural distinto de 1 puede escribirse como un producto de números naturales primos y esta descomposición es única salvo el orde de los factores.

TEOREMA 2.1. (Principio de inducción).

Si S es un subconjunto de N tal que:

1) $1 \in S$

2) Si $n \in S$ entonces $n+1 \in S$

entonces $S = N$.

Demostración: Por 1) y 2) S es inductivo. Por lo tanto, en virtud de la definición d N es $N \subset S$. Como $S \subset N$, entonces $S = N$.

El principio de inducción puede enunciarse de otra manera. Supongamos que a cada número natural n se le asocia una proposición P(n) bien determinada, que puede ser verdadera o falsa, por ejemplo:

a) P(n): "n es un número mayor o igual que 1".

b) P(n): "La suma de los n primeros números naturales es igual a $\frac{n(n+1)}{2}$ ".

c) P(n): "n es mayor que 6".

d) P(n): "El duplo de n es menor que 6"

Las proposiciones a) y b) son verdaderas para todo $n \in N$; c) es falsa para n=1,2,3,4,5, y d) es falsa para todo $n \geq 3$.

El principio de inducción puede enunciarse equivalentemente como sigue:

TEOREMA 2.2. Si P(n) es una proposición relativa al número natural n y se verifica que

1) P(1) es verdadera.

2) Si P(k) es verdadera entonces P(k+1) es verdadera.

entonces P(n) es verdadera para todo $n \in N$.

Es claro que 2.1 $\Longrightarrow$ 2.2 porque si S es el conjunto de todos los números naturales para los que P(n) es verdadera, S verifica las dos condiciones del teorema 2.1 y por lo tanto coincide con N. Recíprocamente, 2.2 $\Longrightarrow$ 2.1 pues si S es un subconjunto de N que cumple las condiciones 1) y 2) de 2.1 , considerando la proposición P(n): "$n \in S$", se tiene en virtud de 2.2 que $n \in S$, $\forall\, n \in N$ o sea $S = N$.

El principio de inducción es una propiedad fundamental de los números naturales que proporciona un método de demostración llamado por inducción o recurrencia. Para demostrar que los números naturales verifican una cierta propiedad es suficiente:

1) Demostrar que es verdadera para 1.
2) Suponiéndola verdadera para n, probar que se verifica para n+1.

Por ejemplo, probemos que $1 \leq n$, $\forall\, n \in N$.

Para cada $n \in N$ consideremos la proposición P(n): "$1 \leq n$".

1) Es verdadera para $n = 1$ pues $1 \leq 1$.
2) Si es verdadera para k, entonces es verdadera para k+1.

   En efecto, como $0 < 1$ resulta por monotonía de la suma $k < k+1$; si $1 \leq k$, por transitividad se tiene $1 < k+1$.

   Queda así probado que $1 \leq n$, $\forall\, n \in N$.

EJERCICIO: Demostrar aplicando el principio de inducción que:

a) Para todo $n \in N$, $n \neq 1$ es $n-1 \in N$.

   Considerar el conjunto $S \subset N$ definido así: $S = \{x \in N : x-1 \in N\} \cup \{1\}$. Demostrar que S es inductivo.

b) Para todo $n \in N$ no existe ningún número natural entre n y n+1.

El conjunto de los números naturales tiene una propiedad que no comparte con ninguno de de los conjuntos Z , Q ó R. Es el llamado Principio de buena ordenación: Todo subconjunto no vacío del conjunto de números naturales tiene primer elemento.

En general, si en un conjunto ordenado se verifica la propiedad de que todo subconjunto no vacío tiene primer elemento, el conjunto se dice que está bien ordenado.

Para demostrar el principio de buena ordenación utilizaremos el principio de inducción. (en verdad son equivalentes).

TEOREMA 2.3. (Principio de buena ordenación).

Todo subconjunto no vacío de N tiene primer elemento.

Demostración: Consideremos la siguiente proposición: P(n): "Todo subconjunto de N que contiene un número menor o igual que n tiene primer elemento".

1) P(1) es verdadera pues si $A \subset N$ es un conjunto que contiene un número k menor o igual que 1, como $1 \leq n$, $\forall\, n \in N$, debe ser $k = 1$; luego $1 \in A$ y 1 es el primer elemento de A.
2) Supongamos que P(n) es verdadera y probemos que P(n+1) es verdadera. Sea $A \subset N$ un conjunto que contiene un número menor o igual que n+1. Si A no contiene ningún número menor que n+1 entonces debe ser $n+1 \in A$ y n+1 es el primer elemento de A. Si A contiene un número k menor que n+1, $k < n+1$, entonces como $n \in N$ y entre n y n+1 no hay ningún número natural según el ejercicio b), debe ser $k \leq n$. Luego, A contiene un número menor o igual que n y por la hipótesis de inducción A tiene primer elemento.

A menudo se utiliza en las demostraciones una segunda forma del principio de inducción.

TEOREMA 2.4. Sea P(n) una proposición relativa al número natural n. Si se verifica qu

1) P(1) es verdadera.
2) Para todo $n > 1$, la hipótesis de que P(k) es verdadera para todo $k < n$ implica que P(n) es verdadera.

entonces P(n) es verdadera para todo $n \in N$.

Demostración: Sea $S = \{n \in N : P(n) \text{ es verdadera}\}$. S verifica las dos condiciones siguientes:

1) $1 \in S$
2) Para todo $n > 1$, si $k \in S$ para todo $k < n$ entonces $n \in S$.

Se trata de probar que $S = N$. Sea $S' = N - S$. Si $S' \neq \emptyset$, por el principio de buena ordenación tiene primer elemento p y $p > 1$ ya que $1 \notin S'$.

Luego todos los números naturales menores que p no pertenecen a S', es decir, para todo $k < p$, $k \in S$.

Luego por 2) resulta $p \in S$. Como $p \in S'$, esta contradicción prueba que $S' = \emptyset$, o sea que $S = N$ lo que termina la demostración.

Se tiene así un segundo método de demostración por inducción. Para demostrar que los números naturales verifican una cierta propiedad es suficiente:

1) Demostrar que es verdadera para 1.
2) Suponiéndola verdadera para todos los números naturales menores que n, demostrarla para n ($n > 1$).

TEOREMA 2.5. N es cerrado con respecto a la suma y a la multiplicación.

Demostración: Probemos que N es cerrado con respecto a la suma. Se demuestra por inducción. Consideremos el conjunto S de todos los números naturales n que tienen la propiedad de que sumados a cualquier número natural dan por resultado un número natural:

$$S = \{n \in N : m+n \in N \ , \ \forall \ m \in N\}$$

1) $1 \in S$ pues $\forall\, m \in N$, $m+1 \in N$.
2) Supongamos que $n \in S$ y probemos que $n+1 \in S$.

   $\forall\, m \in N$, $m+(n+1) = (m+n)+1$. Como $n \in S$, es $m+n \in N$ lo que implica $(m+n)+1 \in N$. Luego $m+(n+1) \in N$, $\forall\, m \in N$, es decir $n+1 \in S$.

   Entonces $S = N$, lo que termina la demostración.

Análogamente se prueba que N es cerrado con respecto a la multiplicación, lo que queda a cargo del lector.

En N la suma y la multiplicación verifican las propiedades S1 , S2 , M1 , M2 , M3 , D y 3. (Leyes cancelativas de la suma y la multiplicación). Respecto a la relación de orden de R restringida a N es claro que verifica O1 , O2 , O3 y O4.

Como no se verifican las propiedades S3 y S4 , la diferencia no es una operación binaria en N y la ecuación $b + x = a$, $a,b \in N$ no tiene siempre solución en N.

Análogamente, como los inversos de los números naturales distintos de 1 no son natura-

les, ya que si $n > 1$ por la propiedad 20 es $n^{-1} < 1$, no se verifica la propiedad M4 y la ecuación $b.x = a$ con $a,b \in N$ no siempre tiene solución en N.

Con respecto a la diferencia N tiene la siguiente propiedad:

TEOREMA 2.6. Si $m,n \in N$ y $m < n$ entonces $n-m \in N$.

Su demostración queda propuesta como ejercicio.

En cuanto al teorema de factorización única o teorema fundamental de la aritmética puede demostrarse directamente por inducción sin introducir el algoritmo de la división entera ni la noción de máximo común divisor. Sin embargo creemos más conveniente desarrollar la teoría de divisibilidad en el conjunto Z de los enteros de esa manera porque, como veremos más adelante, una teoría enteramente análoga vale en el anillo de polinomios en una indeterminada con coeficientes racionales, reales o complejos.

EJERCICIOS:

1. Demostrar por inducción que $\forall\, n \in N$:

   a) $1+2+3+.....+n = \frac{n(n+1)}{2}$

   b) $2+4+6+....+2n = n(n+1)$

   c) $1+3+5+....+(2n-1) = n^2$

   d) $1^2+2^2+3^2+....+n^2 = \frac{1}{6}\, n(n+1)(2n+1)$

   e) $1+2+4+.....+2^{n-1} = 2^n - 1$

   f) $\frac{1}{1.2} + \frac{1}{2.3} + \frac{1}{3.4} +.....+ \frac{1}{n(n+1)} = \frac{n}{n+1}$

   g) $4.(1^3+2^3+3^3+.....+n^3) = n^2(n+1)^2$

   h) $\frac{1}{3} + \frac{2}{3^2} + \frac{2^2}{3^3} +.....+ \frac{2^{n-1}}{3^n} = 1 - (\frac{2}{3})^n$

   i) Se dice que una sucesión de números $a_1, a_2, ..., a_n, ...$ es una progresión aritmética si $a_{i+1} - a_i = d$ para $i=1,2,3,.....$, siendo d un número fijo. Demostrar que la suma de n términos consecutivos de una progresión aritmética es igual a la semisuma de los extremos por el número de términos. Señalar entre las fórmulas anteriores las que dan la suma de los términos de una progresión aritmética y verificar el resultado.

   j) Se dice que una sucesión de números $a_1, a_2, ..., a_n, ...$ es una progresión geométrica si $\frac{a_{i+1}}{a_i} = q$ para $i = 1,2,3.....$ siendo q un número fijo. Demostrar que la suma de n términos consecutivos de una progresión geométrica es:

   $$a_1 + a_2 + ...+ a_n = a_1 . \frac{q^n - 1}{q-1}$$

   Señalar de entre las primeras fórmulas las que dan la suma de los términos de una progresión geométrica y verificar el resultado.

   k) Demostrar que:

i) $2^n > n$ $\forall\ n \in N$

ii) 2 divide a n^2+n " "

iii) 3 " " n^3-n+3 " "

iv) 4 " " 7^n-3^n " " . (Sugerencia: $7^{n+1}-3^{n+1}=7^{n+1}-3.7^n+3.7^n-3^{n+1}$).

v) Si $b \in R$ y $1+b \geq 0$ demostrar que $(1+b)^n \geq 1+n.b$, $\forall\ n \in N$.

2. Tratar de probar cada una de las siguientes fórmulas por inducción e indicar qué hipótesis del principio de inducción falla. (Todas son falsas).

a) $3+5+7+.....+(2n+1) = n^2+2$ $\forall\ n \in N$

b) $4+5+6+.....+(n+3) = n^3+3$ " "

c) $5+10+15+.....+5n = \frac{n(n+1)}{2}$ " "

d) $3+6+9+.....+3n = \frac{3n(n+1)}{2} + 1$ " "

2.3. NUMEROS ENTEROS.

El conjunto Z de los números enteros es el subconjunto de R formado por los números naturales, sus simétricos y el 0.

$$Z = N \cup \{0\} \cup \{x \in R : -x \in N\}$$

Como para todo $x \in N$ es $x > 0$, resulta $-x < 0$.

Luego los enteros positivos son los naturales y los enteros negativos los simétricos de los naturales. Observemos que $m \in Z$ si y solo si $-m \in Z$.

Z es un subconjunto propio de R porque los inversos de los números enteros distintos de 1 y -1 no son enteros. Basta verificarlo para los enteros positivos: si $m \in Z$ y $m > 1$ por las propiedades 19 y 20 es $0 < m^{-1} < 1$ y por lo tanto $m^{-1} \notin Z$.

Propiedades de Z.

(1) Z es cerrado con respecto a la suma y a la multiplicación y Z verifica las propiedades S1 , S2 , S3 , S4 , M1 , M2 , M3 , D y 5. (También 1,2,3,4 y 6 que se deducen de ellas). Esto se expresa diciendo que Z es un anillo conmutativo sin divisores de cero. Es claro que la relación de orden restringida a Z verifica 01,02, 03 y 04.

Como vale la propiedad S4 , la ecuación $b + x = a$, con $a,b \in Z$ tiene siempre solución en Z, es decir Z es cerrado con respecto a la diferencia. En cambio Z no verifica M4 pues ya vimos que los inversos de los enteros distintos de 1 y -1 no pertenecen a Z. De modo que la ecuación $b.x = a$, con $a,b \in Z$, $b \neq 0$ no siempre tiene solución en Z, es decir no siempre es posible la división de un entero a por otro $b \neq 0$. Existe en cambio la llamada "división entera".

(2) Z es numerable.

(3) Algoritmo de la división entera. Para todo par de enteros a,b con $b \neq 0$, existen dos enteros q y r, llamados el cociente y el resto respectivamente de dividir a por b, unívocamente determinados, tales que: $a = q.b + r$ y $0 \leq r < |b|$.

(4) Factorización única. Todo número entero distinto de 0 , 1 , -1 se puede escribir como producto de ± 1 por enteros primos positivos y esta descomposición es única salvo el orden de los factores.

La propiedad (1) es fácil de verificar y la (2) ya fue demostrada. Probaremos la (3) y la (4) al final del capítulo para no cortar el hilo de la exposición.

2.4. NUMEROS RACIONALES.

El conjunto Q de los números racionales es el subconjunto de R formado por todos los cocientes de números enteros.

$$Q = \{a.b^{-1} : a,b \in Z , b \neq 0\}$$

Un número racional puede representarse como cociente de diferentes pares de enteros: si $a,b,c,d \in Z$, $b \neq 0$, $d \neq 0$ y $ad = bc$ entonces $\frac{a}{b} = \frac{c}{d}$. Se dice que la fracción $\frac{a}{b}$ representa al número racional x. Cada número racional $x \neq 0$ se puede representar por una única fracción irreducible de denominador positivo. (Se dice que $\frac{a}{b}$ es irreducible si a y b no tienen factores comunes distintos de ± 1).

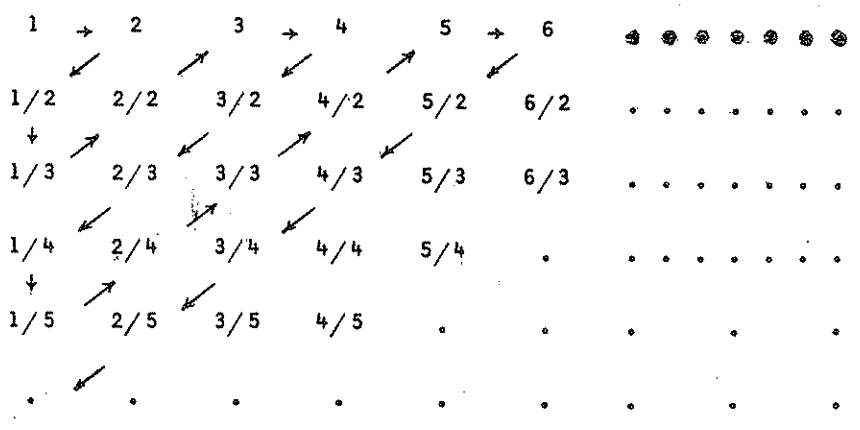
Z es un subconjunto propio de Q pues cualquiera sea $m \in Z$ es $m = \frac{m}{1} \in Q$; además si $m \neq 1,-1$ entonces $\frac{1}{m} = m^{-1} \in Q - Z$.

Propiedades de Q.

(1) Q es cerrado con respecto a la suma y a la multiplicación y en Q se verifican todas las propiedades de cuerpo ordenado, es decir S1 , S2 , S3 , S4 , M1 , M2 , M3, M4 , D , O1 , O2 , O3 y O4. (Considerando la restricción a Q de la relación de orden definida en R).

(2) Q es numerable.

Veamos primero que el conjunto Q^+ de los números racionales positivos es numerable. Se pueden disponer todos los cocientes de enteros positivos en un cuadro infinito como sigue:



Tachando los números que aparecen más de una vez, como $\frac{4}{2}$, $\frac{6}{3}$, $\frac{8}{4}$, etc., se pueden "co tar" todos los números racionales positivos recorriendo el cuadro ordenadamente como i dican las flechas. Es decir existe una correspondencia biunívoca entre Q^+ y N: $\frac{a}{b} \leftrightarrow$ r Pero esta correspondencia puede extenderse a una de Q sobre Z poniendo $\frac{a}{b} \leftrightarrow n$, $-\frac{a}{b} \leftrightarrow$ $0 \leftrightarrow 0$. Como Z es numerable tambien lo es Q.

(3) Q es denso en R. Entre dos números reales siempre existe un número racional, es decir, si $a,b \in R$ y $a < b$ existe $x \in Q$ tal que $a < x < b$.

Esto equivale a decir que entre dos números reales existen infinitos racionales.

Para demostrarlo es necesario usar la propiedad de R que nos falta enunciar: el axioma de completidad.

2.5. AXIOMA DE COMPLETIDAD.

Dijimos que íbamos a indicar una colección de propiedades que caracterizan al sistema de los reales. Es evidente que las de cuerpo ordenado no son suficientes porque el su sistema Q de los racionales es también un cuerpo ordenado.

Para poder enunciar la propiedad que falta necesitamos precisar algunos conceptos.

Definición. Sea $A \neq \emptyset$ un subconjunto de R.

1°) Se dice que un número $c \in R$ es una cota superior (inferior) de A si $x \leq c$ ($c \leq x$) cualquiera que sea $x \in A$.

2°) A se dice acotado superiormente (inferiormente) si existe alguna cota superior (i ferior) de A. Se dice acotado si es acotado superior e inferiormente.

3°) Un número $\ell \in R$ se dice extremo superior o supremo de A si verifica las dos sigui tes propiedades:

S1. ℓ es una cota superior de A.

S2. ℓ es la menor de las cotas superiores de A, es decir, si ℓ' es otra cota sup rior de A entonces $\ell \leq \ell'$.

Dualmente se define extremo inferior o ínfimo.

El extremo superior de un subconjunto, si existe, es único lo que resulta de inmediato de la propiedad S2. Análogamente para el extremo inferior.

Notemos que el extremo superior (inferior) de un subconjunto A puede pertenecer o no a Si pertenece entonces es el último (primer) elemento de A.

EJEMPLOS:

1) Sea $A = \{x \in R : x < 0\}$. Una cota superior de A es 0 o cualquier número positivo El extremo superior de A es 0 y $0 \notin A$. A no es acotado inferiormente.

2) $S = \{x \in Z : -2 < x < 6\}$. S es acotado superior e inferiormente. -1 es el extremo inferior de S y 5 el extremo superior.

3) $T = \{x \in Q : -2 < x \leq 6\}$. T es acotado. El extremo inferior es -2 y el superior 6.

$\mathbb{C}$) Axioma de Completidad. Todo subconjunto no vacío de R acotado superiormente tiene extremo superior en R.

Esta propiedad junto con las de cuerpo ordenado caracterizan al sistema de los números reales y se enuncian diciendo que R es un cuerpo ordenado completo. Se puede demostrar que cualquier sistema formado por un conjunto en el que están definidas dos operaciones binarias y una relación de orden que verifican las propiedades S1 , S2 , S3 , S4 , M1, M2 , M3 , M4 , D , 01 , 02 , 03 , 04 y el axioma de completidad es una copia de R desde el punto de vista algebraico.

Es interesante notar que el axioma de completidad puede enunciarse equivalentemente como sigue:

$\mathbb{C}'$) Todo subconjunto no vacío de R acotado inferiormente tiene extremo inferior en R.

En efecto, veamos que $\mathbb{C} \Longrightarrow \mathbb{C}'$.

Sea S un subconjunto no vacío de R y acotado inferiormente y consideremos el conjunto A de todas las cotas inferiores de S. Como S es acotado inferiormente, $A \neq \emptyset$ y A es acotado superiormente ya que todo elemento de S es cota superior de A.

Luego por $\mathbb{C}$, A tiene extremo superior $\ell \in R$.

Probemos que ℓ es el extremo inferior de S. En primer lugar $\ell \leq x$, cualquiera que sea $x \in S$, por ser ℓ la menor de las cotas superiores de A. Además si $\ell' \in R$ es una cota inferior de S entonces $\ell' \in A$ y por lo tanto $\ell' \leq \ell$, lo que termina la demostración.

En formá análoga se prueba que $\mathbb{C}' \Longrightarrow \mathbb{C}$ lo que queda propuesto como ejercicio.

Con el auxilio del axioma de completidad se demuestran las siguientes propiedades de R:

A. Propiedad arquimedeana. Cualesquiera sean $a,b \in R$, a y b positivos, existe un número natural n tal que $a < nb$.

   En particular, para todo $a \in R$ existe un $n \in N$ tal que $a < n$.

B. Existencia de raíces en R. Para todo número real positivo a y todo $n \in N$ la ecuación $x^n = a$ tiene una y solo una solución real positiva.

De aquí la siguiente definición:

Definición. Dados $a \in R$, $a > 0$ y $n \in N$ se llama raíz n-ésima aritmética de a al número real positivo b tal que $b^n = a$.

Se escribe $b = \sqrt[n]{a}$

Para ver la demostración de estas dos propiedades y de que Q es denso en R remitimos al lector al libro de Birkhoff y Mac Lane, Algebra Moderna, páginas 72 a 75.

La propiedad B asegura, en particular, que en R existe la raíz cuadrada de 2, es decir un número cuyo cuadrado es 2. Ya vimos que en cambio en Q no hay ningún número con esta propiedad. Lo que pasa es que

Q es un cuerpo ordenado pero no es completo,

es decir existen conjuntos de números racionales no vacíos y acotados superiormente (inferiormente) que no tienen extremo superior (inferior) en Q.

Para verlo consideremos, por ejemplo, el conjunto de los números racionales positivos de cuadrado menor que 2:

$$S = \{x \in Q : x > 0 \quad y \quad x^2 < 2\}$$

$S \neq \emptyset$ y es acotado superiormente pero su extremo superior no pertenece a Q. Probemos que $\sqrt{2}$ es el extremo superior de S.

1°) $x < \sqrt{2}$, $\forall$ $x \in S$, ya que si fuera $x \geq \sqrt{2}$ para algún $x \in S$ sería, por la ley de monotonía de la multiplicación, $x^2 \geq 2$. Luego $\sqrt{2}$ es una cota superior de S.

2°) Veamos que $\sqrt{2}$ es la menor de las cotas superiores. Si $a \in R$ es tal que $x \leq a$ $\forall$ $x \in S$ entonces $\sqrt{2} \leq a$. En efecto, si fuera $a < \sqrt{2}$, como Q es denso en R existe un $x \in Q$ tal que $a < x < \sqrt{2}$. Luego $x > 0$ y $x^2 < 2$ lo que implica $x \in S$ y contradice la hipótesis de que a es una cota superior de S. Esta contradicción prueba fue $\sqrt{2} \leq a$.

Por lo tanto $\sqrt{2}$ es el extremo superior de S y $\sqrt{2} \notin Q$.

Los números reales que no son racionales se llaman irracionales.

Se puede demostrar que R es no numerable. De aquí resulta que el conjunto de los números irracionales no es numerable, porque la reunión de dos conjuntos numerables es numerable y como Q es numerable, si el conjunto de los irracionales fuera numerable sería R numerable. Hablando informalmente, hay más números irracionales que racionales.

Pensando en la recta, a pesar de que existe una infinidad de puntos racionales en cualquier segmento de la recta y que los puntos racionales están tan próximos como se desee, los puntos que no corresponden a ningún número racional existen en mayor cantidad que los puntos racionales.

Son números irracionales, por ejemplo, $\sqrt{2}$, $\sqrt{3}$, $\sqrt{5}$, $\sqrt[3]{2}$, $\sqrt[4]{5}$, $\sqrt[7]{12}$,..... Pero no todos los números irracionales son de la forma $\sqrt[n]{a}$, con a entero. Más generalmente, hay números irracionales, como por ejemplo $\log 2$, π, $2^{\sqrt{2}}$, que no son solución de ninguna ecuación $x^n + ax^{n-1} + bx^{n-2} + \ldots + c = 0$ con coeficientes a,b,.....,c enteros. Y se demuestra que los irracionales de este último tipo forman un conjunto no numerable mientras que los que son raíz de alguna ecuación con coeficientes enteros (entre los que se encuentran los de la forma $\sqrt[n]{a}$) forman un conjunto numerable.

Para terminar, veamos ahora cómo a cada punto de la recta se puede hacer corresponder un número real. Establecida la correspondencia entre puntos de la recta y números racionales como indicamos al principio del capítulo, sea P un punto al que no corresponde ningún número racional. Sea I el conjunto de todos los números racionales que corresponden a puntos situados a la izquierda de P y D el conjunto de todos los números racionales que corresponden a puntos situados a la derecha de P. Quedan clasificados así todos los números racionales.

Como I es $\neq \emptyset$ y acotado superiormente tiene extremo superior en R. Análogamente D tiene extremo inferior en R. Sea a = extremo superior de I.

b = extremo inferior de D.

Debe ser $a < b$ ó $a = b$.

Supongamos $a < b$. Entonces, por ser Q denso en R existe un número racional q tal que $a < q < b$. Como q es racional, pertenece a I ó a D. Pero $q \in I$ es imposible pues $a<q$ y $q \in D$ es imposible pues $q < b$. Esta contradicción prueba que $a \neq b$.

Luego $a = b$ y éste es el número real que se hace corresponder al punto P. La correspondencia así establecida entre puntos de la recta y números reales es biunívoca.

2.6. POTENCIACION DE EXPONENTE ENTERO DE NUMEROS REALES.

Cualquiera sea $a \in R$ consideremos la siguiente definición por recurrencia:

$$a^1 = a$$

$$a^{n+1} = a^n . a$$

Queda definida así la potencia a^n de exponente natural $\forall\, a \in R$, $n \in N$. La definición se extiende a exponente entero cualquiera como sigue:

$$a^0 = 1$$

$$a^n = (a^{-1})^{-n} \quad , \quad a \in R \, , \, a \neq 0 \, , \, n \in Z \, , \, n < 0$$

Se verifica que

$$a^n = (a^{-1})^n = (a^n)^{-1} \quad , \; n \in N$$

Propiedades de la potenciación.

1) $a^n . a^m = a^{n+m}$

2) $(a^m)^n = a^{m.n}$

3) $(ab)^n = a^n . b^n$

4) $1^n = 1$

2.7. PROPIEDADES DE LA RAIZ ARITMETICA.

1) $\sqrt[n]{a.b} = \sqrt[n]{a} \, . \, \sqrt[n]{b}$

2) $\sqrt[m]{\sqrt[n]{a}} = \sqrt[mn]{a}$

3) $\sqrt[m]{a^n} = (\sqrt[m]{a})^n$ con $n \in Z$

4) $\sqrt[ms]{a^{n.s}} = \sqrt[m]{a^n}$ con $n \in Z$

2.8. POTENCIACION DE EXPONENTE RACIONAL.

Se define la potenciación de base real positiva y exponente racional cualquiera como sigue: Dados $a \in R$, $a > 0$, $m,n \in Z$, $n > 0$ por definición es:

$$a^{m/n} = \sqrt[n]{a^m}$$

Está bien definida pues si $\frac{m}{n} = \frac{p}{q}$ $(q > 0)$ entonces $a^{m/n} = a^{p/q}$. En efecto, si $\frac{m}{n} =$ es $mq = np$. Entonces aplicando la propiedad 4 de la raíz aritmética resulta:

$$a^{m/n} = \sqrt[n]{a^m} = \sqrt[nq]{a^{mq}} = \sqrt[nq]{a^{np}} = \sqrt[q]{a^p} = a^{p/q}$$

En particular

$$\sqrt[n]{a} = a^{1/n}$$

y además de acuerdo con la propiedad 3 de la raíz aritmética se tiene:

$$a^{m/n} = \sqrt[n]{a^m} = (\sqrt[n]{a})^m$$

Propiedades: La potenciación de exponente racional verifica propiedades análogas a la de exponente entero.

1) $a^{\alpha} \cdot a^{\beta} = a^{\alpha+\beta}$

2) $(a^{\alpha})^{\beta} = a^{\alpha.\beta}$

3) $(a.b)^{\alpha} = a^{\alpha} \cdot b^{\alpha}$

4) $1^{\alpha} = 1$

2.9. DIVISIBILIDAD DE ENTEROS.

Probaremos ahora las propiedades (3) y (4) de los números enteros enunciadas en el parágrafo 2.3.

TEOREMA 2.7. Para todo par de números enteros a y b , $b \neq 0$, existen dos enteros q y r , llamados el cociente y el resto respectivamente de dividir a por b , unívocamente determinados, tales que:

$$a = q.b + r \quad y \quad 0 \leq r < |b|$$

Demostración: Demostremos primero la existencia de un par q , r al menos.

Si $a = c.b$ para algún $c \in Z$ entonces haciendo $q = c$ y $r = 0$ se tiene un par en esas condiciones.

Supongamos que $a \neq xb \ \forall \ x \in Z$. Consideremos el conjunto S de todos los números positivos de la forma $a - xb$, $x \in Z$. Veamos que $S \neq \emptyset$.

Sean

$$\alpha = \begin{cases} 1 & \text{si } a > 0 \\ -1 & \text{si } a < 0 \end{cases} \qquad \beta = \begin{cases} 1 & \text{si } b > 0 \\ -1 & \text{si } b < 0 \end{cases}$$

$|a| = \alpha.a$ y $|b| = \beta.b$. Tomemos $x = -\alpha\beta a$. Luego $a - xb = a + \alpha\beta ab =$

$= \alpha|a| + |a|.|b| = |a|(\alpha + |b|) > 0$ pues $|b| > 1$ ya que $a \neq xb \ \forall\, x \in Z$.

Entonces $S \neq \emptyset$ y por el principio de buena ordenación tiene primer elemento r que se obtiene para un cierto valor $x = q : r = a - qb$. Luego $a = qb + r$.

Falta probar que $r < |b|$. (ya sabemos que $r > 0$).

Si fuera $r \geq |b|$ tendríamos

$$r > r - |b| \geq 0$$

$r - |b| = a - qb - |b| = a - (q + \beta)b > 0$ porque $a \neq xb \ \forall\, x \in Z$.

Entonces $r - |b| \in S$ pues es positivo y de la forma $a - xb$. Como $r > r - |b|$ esto contradice la hipótesis de que r es el primer elemento de S.

Esta contradicción prueba que $r < |b|$.

Unicidad: Sean $q,q',r,r' \in Z$ tales que:

$$a = qb + r \quad y \quad 0 \leq r < |b| \qquad (1)$$

$$a = q'b + r' \quad y \quad 0 \leq r' < |b| \qquad (2)$$

Luego

$$qb + r = q'b + r' \Longrightarrow (q - q')b = r' - r \Longrightarrow |q - q'|.|b| = |r' - r|$$

Por otro lado, de (1) y (2) resulta $|r' - r| < |b|$ (3)

Si $r \neq r'$, $|r' - r| \neq 0$ y $|q - q'| \neq 0$. Entonces

$$|b| \leq |q - q'|.|b| = |r' - r|$$

lo que contradice (3).

Luego $r = r'$. Entonces $(q - q')b = 0$ y como $b \neq 0$ es $q - q' = 0$ o sea $q = q'$.

Notemos que la demostración del teorema se basa en el principio de buena ordenación de los naturales.

EJEMPLOS:

Si $a = 11$ y $b = 2$ es $q = 5$ y $r = 1$: $11 = 5.2 + 1$

Si $a = -17$ y $b = 3$ es $q = -6$ y $r = 1$: $-17 = (-6).3 + 1$

Si $a = -10$ y $b = -6$ es $q = 2$ y $r = 2$: $-10 = 2.(-6) + 2$

Si $a = 3$ y $b = 8$ es $q = 0$ y $r = 3$: $3 = 0.8 + 3$

Relación divide.

Definición. Se dice que un número entero a divide a otro b si existe un entero c tal que $b = c.a$. Se escribe a/b.

Si a/b se dice también que a es un divisor de b o que b es un múltiplo de a.

Es claro que si $a \neq 0$, a divide a b si y solo si el resto de dividir b por a es cero.

Esta relación tiene las siguientes

Propiedades.

1. a/a , $\forall\ a \in Z$ (Propiedad reflexiva)
2. Si a/b y b/c entonces a/c (Propiedad transitiva)
3. $a/0$, $\forall\ a \in Z$
4. Si a/b y a/c entonces $a/xb+yc$, $\forall\ x,y \in Z$

La demostración es muy simple y queda a cargo del lector.

Los enteros que tienen inverso en Z con respecto a la multiplicación se llaman unitarios. Los enteros unitarios son 1 y -1. En términos de divisibilidad se caracterizan por la propiedad de dividir a todo número entero.

PROPOSICION 2.1. Cualesquiera sean $a,b \in Z$ las siguientes propiedades son equivalentes:

1) a y b tienen los mismos divisores.

2) a y b difieren en un factor unitario, es decir $a = \pm b$.

3) a/b y b/a.

Demostración: Seguiremos el siguiente esquema: 1) $\Longrightarrow$ 2) $\Longrightarrow$ 3) $\Longrightarrow$ 1).

1) $\Longrightarrow$ 2). Supongamos que a y b tienen los mismos divisores.

Como a/a entonces a/b. Luego $b = c.a$, $c \in Z$.
Como b/b entonces b/a. Luego $a = c'.b$, $c' \in Z$.

Reemplazando se tiene $a = c'.c.a$

Si $a \neq 0$, por la ley de cancelación de la multiplicación resulta $c'.c = 1$. Es decir c' y c son enteros unitarios, luego $c = \pm 1$, $c' = \pm 1$ y se verifica 2).

Si $a = 0$ entonces $b = 0$ y es claro que también se verifica 2).

2) $\Longrightarrow$ 3). Si $a = \pm b$ es claro que b/a. Por otro lado $b = \pm a$, luego a/b.

3) $\Longrightarrow$ 1). Supongamos que a/b y b/a. Hay que probar que a y b tienen los mismos divisores, es decir que si un número entero c es tal que c/a entonces c/b y recíprocamente, que si c/b entonces c/a, lo que resulta inmediatamente de la hipótesis y la transitividad de la relación divide.

Definición. Dos enteros a y b que verifican las condiciones de la proposición anterior se dicen asociados.

La relación "ser asociado" es una relación de equivalencia en Z. Los asociados de un entero a son a y -a.

Es claro que todo número a ε Z es divisible por 1 , -1 , a y -a. Estos se llaman los divisores triviales de a. Todo divisor distinto de ellos se dice un divisor propio de a.

EJERCICIO. Demostrar que si c es un divisor propio de a y a $\neq$ 0 entonces $|c| < |a|$.

Máximo común divisor.

Definición. Dados dos enteros a y b, un número entero d se dice un máximo común divisor de a y b si verifica las dos siguientes propiedades:

D1) d/a y d/b

D2) Si d' es un entero tal que d'/a y d'/b entonces d'/d.

Para indicar que d es un máximo común divisor de a y b escribiremos d = (a,b).

Vamos a probar que para todo par de enteros a y b existe un máximo común divisor d y que d se puede escribir como un múltiplo de a más un múltiplo de b, es decir, en la forma d = x.a + y.b con x,y ε Z.

Notemos que si d y d' son dos máximos comunes divisores de a y b, entonces por la propiedad D2 se tiene d/d' y d'/d. Es decir d y d' son asociados y por lo tanto d'=$\pm$d. Luego el m.c.d. de a y b es único salvo el signo.

Demostraremos que existe un m.c.d. de a y b calculándolo mediante el llamado algoritmo de Euclides. Para ello veamos la siguiente

PROPOSICION 2.2. Si a y b son dos enteros, b $\neq$ 0, y r es el resto de dividir a por b entonces a y b tienen los mismos divisores comunes que b y r.

Demostración: Sean q y r el cociente y el resto de dividir a por b: a = q.b + r (1)

Sea c un divisor común de a y b: c/a y c/b. De (1) resulta r = a - q.b . Entonces c/a y c/b $\Longrightarrow$ c/a+(-q).b $\Longrightarrow$ c/r. Luego todo divisor común de a y b es un divisor de b y r .

Recíprocamente, c/b y c/r $\Longrightarrow$ c/q.b+r $\Longrightarrow$ c/a. Luego todo divisor común de b y r divide a a y b .

COROLARIO. d = (a,b) $\Longleftrightarrow$ d = (b,r)

Como a y b y b y r tienen los mismos divisores comunes y un m.c.d. de dos números es un divisor común que verifica una cierta propiedad con respecto a los demás divisores comunes de esos dos números, es claro que si d = (a,b) entonces d = (b,r) y recíprocamente.

Algoritmo de Euclides.

Sean $a,b \in Z$, $a \neq 0$, $b \neq 0$.

.	q_1	q_2	q_3	. . . . .		q_{n-1}	q_n	q_{n+1}
a	b	r_1	r_2	. . . . .	r_{n-3}	r_{n-2}	r_{n-1}	r_n
r_1	r_2	r_3		. . . . .	r_{n-1}	r_n	0	

Se hacen divisiones sucesivas como indica el esquema. Como los restos sucesivos son todos positivos y estrictamente decrecientes este procedimiento no se puede reiterar más que un número finito de veces. El último paso será una división de resto cero.

PROPOSICION 2.3. Si $a,b \in Z$, $a \neq 0$, $b \neq 0$, el último resto no nulo que se obtiene en el algoritmo de Euclides es un máximo común divisor d de a y b. Además $d = xa+yb$, con $x,y \in Z$.

Demostración:

$$a = q_1 \cdot b + r_1 \quad , \quad 0 < r_1 < |b|$$

$$b = q_2 \cdot r_1 + r_2 \quad , \quad 0 < r_2 < r_1$$

$$r_1 = q_3 \cdot r_2 + r_3 \quad , \quad 0 < r_3 < r_2$$

. . . . . .

$$r_{n-3} = q_{n-1} \cdot r_{n-2} + r_{n-1} \quad , \quad 0 < r_{n-1} < r_{n-2}$$

$$r_{n-2} = q_n \cdot r_{n-1} + r_n \quad , \quad 0 < r_n < r_{n-1}$$

$$r_{n-1} = q_{n+1} \cdot r_n$$

De la última igualdad resulta r_n / r_{n-1}. Luego $r_n = (r_{n-1}, r_n)$. Por el corolario de la proposición anterior se tiene:

$$r_n = (r_{n-1}, r_n) = (r_{n-2}, r_{n-1}) = (r_{n-3}, r_{n-2}) = \ldots\ldots = (r_1, r_2) = (b, r_1) = (a,b) \ .$$

Queda probado así que existe un m.c.d. positivo $d = r_n$.

Veamos ahora que d puede escribirse en la forma $d = xa + yb$, con $x,y \in Z$.

De las igualdades anteriores resulta:

$$r_1 = a - q_1.b = a + (-q_1).b$$

$$r_2 = b - q_2.r_1 = b - q_2(a - q_1.b) = (-q_2)a + (1 + q_1.q_2).b$$

$$r_3 = r_1 - q_3.r_2 = a + (-q_1).b - q_3[(-q_2).a + (1 + q_1.q_2).b] =$$

$$= (1 + q_2.q_3).a + (-q_1 - q_3 - q_1.q_2.q_3).b$$

Se ve que de esta manera cualquier resto r_k se puede escribir como un múltiplo de a más un múltiplo de b , en particular $d = xa + yb$, con $x,y \in Z$.

EJERCICIO: En la demostración anterior se hace uso de la inducción en forma implícita. Demostrar por inducción que en el algoritmo de Euclides cada resto r_k puede escribirse

$$r_k = xa + yb \text{ , con } x,y \in Z.$$

Queda probado entonces que si $a \neq 0$ y $b \neq 0$ existe un m.c.d. positivo d y $d = xa + yb$, con $x,y \in Z$. Si a ó b fueran nulos, por ejemplo $a = 0$, entonces $(0,b) = b$ puesto que b/0 y b/b y la segunda condición que debe cumplir un m.c.d. se verifica obviamente. Además b se puede escribir $b = 1.0 + 1.b$.

Resulta así el siguiente

TEOREMA 2.8. Para todo par de números enteros a y b existe un m.c.d. d y $d = xa + yb$ con $x,y \in Z$. Además los únicos m.c.d. de a y b son d y -d.

Habitualmente se considera el m.c.d. positivo y dada su unicidad, se habla de el m.c.d.

OBSERVACIONES.

1. Los enteros x , y en la expresión $d = xa + yb$ no están unívocamente determinados. Hay infinitos pares de enteros que verifican esa relación.

2. Dado $a \in Z$, a y -a tienen los mismos divisores (Proposición 2.1.). Entonces de la definición de m.c.d. resulta:

$$(a,b) = (-a,b) = (a,-b) = (-a,-b)$$

De modo que para calcular el m.c.d. de dos enteros no nulos se puede suponer a ambos positivos.

EJEMPLO: Hallar (36,208)

	5	1	3	2
208	36	28	8	4
28	8	4	0	

El último resto no nulo es el m.c.d. : (36,208) = 4

Expresemos el m.c.d. en la forma $d = xa + yb$, $x,y \in Z$.

$$208 = 5.36 + 28$$

$$36 = 1.28 + 8$$

$$28 = 3.8 + 4$$

Entonces

$$28 = 208 - 5.36$$

$$8 = 36 - 1.28 = 36 - (208 - 5.36) = -208 + 6.36$$

$$4 = 28 - 3.8 = 208 - 5.36 - 3(-208 + 6.36) = 4.208 - 23.36$$

Por lo tanto el m.c.d. se puede escribir

$$4 = 4.208 + (-23).36$$

Números primos y relativamente primos.

Definición. Un entero p se dice primo si p es distinto de 0 , 1 , -1 y los únicos divisores de p son ±1 , ±p , es decir no tiene divisores propios.

Definición. Dos enteros a y b se dicen relativamente primos si (a,b) = 1.

TEOREMA 2.9. Si a/b.c y a y b son relativamente primos entonces a/c.

Demostración: Supongamos a/b.c y (a,b) = 1. Existen enteros x,y tales que:

$$1 = xa + yb$$

Entonces $$c = xac + ybc$$

a/a.c y a/b.c $\Longrightarrow$ a/xac+ybc $\Longrightarrow$ a/c.

Corolario 1. Si p/b.c y p es primo entonces p/b ó p/c.

En efecto, si p/b no hay nada que probar. Si p$\not$/b , entonces (p,b) = 1 pues p es primo. Aplicando el teorema resulta p/c.

Corolario 2. Si $p/a_1.a_2.....a_n$ y p es primo entonces p/a_i , para algún índice i , $1 \leq i \leq n$.

Demostrarlo por inducción sobre el número de factores.

TEOREMA 2.10. (Fundamental de la aritmética).

Todo número entero distinto de 0 , 1 , -1 se puede escribir como un producto de ±1 por

enteros primos positivos y esta descomposición es única salvo el orden de los factores.

Demostración: Basta demostrarlo para los enteros positivos.

Sea $a \in Z$, $a > 1$. Demostraremos la existencia de tal descomposición por inducción sobre a, usando la segunda forma del principio de inducción.

Si $a = 2$, a es un número primo y puede escribirse en la forma indicada. (Con un solo factor).

Sea $a > 2$ y supongamos que existe la descomposición en producto de factores primos positivos para todo número natural menor que a. Tenemos que probar que también a puede escribirse de esa manera.

Si a es un número primo, la propiedad se verifica. Si a no es número primo, entonces se puede escribir $a = b.c$ con $1 < b < a$, $1 < c < a$. Por la hipótesis de inducción b y c se pueden expresar como producto de factores primos positivos y por lo tanto lo mismo sucede con a.

Queda probado así que todo entero positivo $\neq 1$ puede escribirse como producto de primos positivos. Hay que demostrar ahora la unicidad de la descomposición salvo el orden de los factores.

Supongamos que

$$a = p_1 \cdot p_2 \cdot \ldots \cdot p_n$$

$$a = q_1 \cdot q_2 \cdot \ldots \cdot q_m$$

son dos descomposiciones de a en factores primos positivos.

Hay que demostrar que $n = m$ y que $p_i = q_i$, para $i = 1,2,\ldots,n$, ordenando convenientemente los factores.

Lo demostraremos por inducción sobre n.

Si $n = 1$, $a = p_1$ es un número primo y es claro que el producto $q_1 \cdot q_2 \cdot \ldots \cdot q_m$ solo puede tener un factor , $p_1 = q_1$.

Sea $n > 1$, supongamos que la unicidad de la descomposición vale cuando el número de factores es n-1 y probémoslo para n.

$$p_1 \cdot p_2 \cdot \ldots \cdot p_n = q_1 \cdot q_2 \cdot \ldots \cdot q_m$$

implica $p_1 / q_1 \cdot q_2 \cdot \ldots \cdot q_m$. Luego por el corolario 2 del teorema anterior p_1 / q_i para algún índice i, $1 \leq i \leq m$. Se puede suponer i=1, reordenando los factores si fuera necesario. Como p_1 y q_1 son números primos positivos, de p_1 / q_1 resulta $p_1 = q_1$. Luego, por la ley de cancelamiento de la multiplicación

$$p_2 \cdot \ldots \cdot p_n = q_2 \cdot \ldots \cdot q_m$$

De aquí, por la hipótesis de inducción, los factores de la izquierda son ordenadamente iguales a los de la derecha y como $p_1 = q_1$, resulta $n = m$ y $p_i = q_i$ para i=1,2,..,n.

Esto termina la demostración del teorema.

Por ejemplo, el número 16.500 se factoriza en primos como sigue:

$$16.500 = 2.2.3.5.5.5.11$$

Cuando aparecen primos repetidos se asocian los factores iguales escribiendo su producto en forma de potencia. Así

$$16.500 = 2^2.3.5^3.11$$

$a \neq 0, 1, -1$

En general, un número entero a se escribe:

$$a = \pm p_1^{e_1} \cdot p_2^{e_2} \cdots\cdots p_s^{e_s}$$

donde p_1 , p_2 ,....., p_s son primos positivos distintos y e_1 , e_2 ,....., e_s enteros positivos.

EJERCICIOS.

1. Hallar el cociente y el resto de dividir:

   a) 2.608 por 42
   b) -850 por 326
   c) 4.537 por -748
   d) -78 por -23
   e) 31 por 62
   f) -19 por 36
   g) 19 por -36
   h) -19 por -36

2. Calcular el m.c.d. de los siguientes pares de enteros:

   a) (27,129)
   b) (1560,125)
   c) (-10.324,-146)
   d) (360,-2.730)
   e) (2.275,792)
   f) (-71,-28)

   Expresar el m.c.d. en la forma $d = xa + yb$, con $x,y \in Z$ en los casos a) , b) , d) y f).

3. a) La definición de m.c.d. se puede generalizar como sigue: Dados enteros n_1 , n_2 ,...., n_k se llama un m.c.d. de n_1 , n_2 ,....., n_k a otro entero d tal que :

   D1) d divide a n_1 , n_2 ,....., n_k

   D2) Todo entero que divide a n_1 , n_2 ,....., n_k divide también a d.

   Demostrar que cualesquiera sean n_1 , n_2 ,...., n_k existe un m.c.d. d y que d se puede escribir

   $$d = x_1 n_1 + x_2 n_2 + \ldots + x_k n_k \quad , \quad \text{con} \quad x_1, x_2, \ldots, x_k \in Z \ .$$

   (Sugerencia: Por inducción sobre k). Indicar un método para calcularlo.

   b) Calcular:

   (1.585,15,645)

   (1460,122,55,12)

4. Mínimo común múltiplo.

   Definición. Dados dos enteros a y b , un entero m se dice un mínimo común múltiplo de a y b si verifica las dos propiedades siguientes:

   M1) a/m y b/m .

   M2) Si m' es un entero tal que a/m' y b/m' entonces m/m'.

   a) Demostrar que para todo par de enteros a,b existe un m.c.m. Es suficiente demostrar que si a y b no son simultáneamente nulos, el número entero $m = \frac{a.b}{(a,b)}$ es m.c.m. de a y b. Probar además que los mínimos comunes múltiplos de a y b son enteros asociados, es decir difieren en el signo.

Generalizar la definición de m.c.m. para más de dos enteros.

b) Calcular el m.c.m. de los pares de números del ejercicio 2.

5. Demostrar que:

a) $(0,a) = a \quad \forall \ a \in Z$.

b) Si $p/a_1.a_2....a_n$ y p es primo entonces p/a_i para algún índice i, $1 \leq i \leq n$.

c) Si a/m y b/m y $(a,b) = 1$ entonces $a.b/m$

d) $(ab,ac) = a.(b,c)$

e) Si $(a,m) = 1$ y $(b,m) = 1$ entonces $(ab,m) = 1$

6. a) Probar que $a + x \equiv b + x$ (mód. m) si y solo si $a \equiv b$ (mód. m).

b) Encontrar un contraejemplo que pruebe que la siguiente propiedad no se verifica: Si $a.x \equiv b.x$ (mód. m) entonces $a \equiv b$ (mód. m). (Sabemos que la recíproca es verdadera).

c) Siendo $m \neq 0$, demostrar que $a.x \equiv b.x$ (mód. m) implica $a \equiv b$ (mód. m), cualesquiera sean $a,b \in Z$ si y solo si $(x,m) = 1$.

7. a) Expresar como producto de enteros primos positivos los números:

880 , -9.180 , 16.758 , 14.703 .

b) Demostrar la imposibilidad de la existencia de números enteros a y b que satisfagan cualquiera de las relaciones siguientes:

i) $2a^2 = 3b^2$ iii) $a^2 = 2b^2$

ii) $a^2 = 18b^2$ iv) $a^3 = 4b^3$

c) Si $(m,n) = 1$ y $m.n$ es un cuadrado, entonces m y n son ambos cuadrados.

8. Dados dos enteros a y b de los que se conoce la descomposición en factores primos,

$$a = p_1^{n_1}.p_2^{n_2}.....p_s^{n_s} \quad , \quad b = p_1^{m_1}.p_2^{m_2}.....p_s^{m_s}$$

donde los exponentes $n_1,n_2,.....,n_s,m_1,m_2,.....,m_s$ son enteros mayores o iguales que cero y $p_1,p_2,.....,p_s$ son primos distintos, indicar cómo puede calcularse el m.c.d. y el m.c.m. de a y b y demostrar lo que se afirme.
Usando este método calcular el m.c.d. y el m.c.m. de los siguientes pares de enteros: a) 300 , 2.240 b) 2.420 , 441
¿Qué ventaja ofrece el algoritmo de Euclides sobre este método para calcular el m.c.d. de dos enteros?.

9. Demostrar que existen infinitos números primos.

(Sugerencia: Si p_1 , p_2 ,....., p_n son n primos, el número $p_1.p_2.....p_n + 1$ no es divisible por ninguno de esos primos).

10. Demostrar que:

a) Un entero es par si y solo si su cuadrado es par.

b) Un entero es múltiplo de 3 si y solo si su cuadrado es múltiplo de 3.

c) ¿Vale la misma propiedad para los múltiplos de 4? y de 9? y de 6? y de 12? ¿Porqué?.

d) La suma de los cuadrados de dos enteros no divisibles por 3 no es divisible por 3. ¿Vale la misma propiedad para 5?.

11. a) Si $a,b \in Z$, a/b y $b \neq 0$ entonces $|a| \leq |b|$.

b) Si $a,b \in Z$, a/b y $|b| < a$ entonces $b = 0$.

Números reales.

12. Ordenar en forma decreciente los siguientes números:

$$\frac{4}{3}\ ,\ \frac{1451}{758}\ ,\ -\frac{7}{8}\ ,\ \sqrt{3}\ ,\ -\frac{1}{5}\ ,\ \pi\ ,\ \sqrt[3]{29}$$

13. a) Demostrar que el cuadrado de todo número real no nulo es un número positivo.

b) ¿Existe un número real x tal que $x = \sqrt{2x} - 1$?.

c) Probar que si n es un número natural entonces $0 < \frac{1}{n^2} \leq \frac{1}{n}$. ¿Es cierta esta propiedad para todo $n \in R$?.

d) Demostrar que si $a,b \in R$:

$0 < a < b$ implica $a^2 < b^2$

$b < a < 0$ " $a^2 < b^2$

Dar un ejemplo de $a < 0 < b$ y $b^2 < a^2$

14. Demostrar que si $a,b \in R$, $a \neq 0$ y $b \neq 0$ entonces

i) $(a + b)^2 > 2ab$

ii) $\frac{1}{a} + \frac{1}{b} \neq \frac{1}{a+b}$

15. a) ¿Es verdadera la siguiente proposición?:

Si n es un número natural y a es un número real no natural entonces $n + a$ no es un natural.

b) ¿Se verifica siempre que si m es un entero y a un número real no entero entonces $m + a$ no es entero?.

c) Probar que si r es racional y q irracional entonces $r + q$ y $r.q$ son irracionales, si $r \neq 0$ en el segundo caso.

d) Encontrar dos números irracionales cuya suma sea irracional y dos cuya suma sea racional. Idem para la diferencia, el producto y el cociente. ¿Es cerrado el conjunto de los números irracionales con respecto a la suma y a la multiplicación?.

16. Probar que:

a) $\sqrt{2}$, $\sqrt{3}$ y $\sqrt{6}$ son irracionales.

b) Si $a,b \in Z$ y $\sqrt{a.b}$ es irracional entonces $\sqrt{a} + \sqrt{b}$ es irracional.

c) a irracional implica $-a$ y a^{-1} irracionales.

d)* a irracional y $a > 0$ implica $\sqrt{a}$ irracional.

e) Probar que los siguientes son números irracionales:

$$\sqrt{2} + \sqrt{3} \quad ; \quad \sqrt{2} + \frac{3}{2} \quad ; \quad (1 - \sqrt{2})^2 \quad ; \quad \frac{1 + \sqrt{3}}{1 - \sqrt{3}} \quad ; \quad \frac{1}{\sqrt{6}} \quad ; \quad \sqrt{2} + \sqrt{3} - 8 \ .$$

17. Hallar los valores de x que verifican las siguientes condiciones:

a) $|3x - 1| = 1$

b) $|3x - 1| < 1$

c) $|x - 2| < 1$ y $|x + 1| < 3$

d) $|2x + 1| > 2$ y $|x + 2| < 1$

18. a) Consideremos el conjunto ordenado Q^+ de los números racionales positivos. Demostrar que Q^+ no es bien ordenado.

b) Consideremos en Q^+ la siguiente definición: Un número $p \in Q^+$ se dice primo si no se puede escribir como producto de dos números racionales menores que p. Por ejemplo, $\frac{1}{2}$ es primo porque $\frac{1}{2} = a.b$, $a,b \in Q^+$, implica a ó b mayor que $\frac{1}{2}$. (Verificarlo). En cambio $\frac{3}{2}$ no es primo porque $\frac{3}{2} = \frac{5}{4}.\frac{6}{5}$ donde $\frac{5}{4} < \frac{3}{2}$ y $\frac{6}{5} < \frac{3}{2}$. Con esta definición de primo demostrar con un ejemplo que en Q^+ no vale la factorización única.

19. Para cada uno de los siguientes subconjuntos numéricos decir: i) Si es acotado superior o inferiormente y en tal caso indicar dos cotas superiores y dos inferiores. ii) Si tiene extremo superior o inferior en el conjunto numérico en cuestión. iii) Si tiene primer o último elemento.

$\{x \in R : x > 0\}$

$\{x \in R : x < 0\}$

$\{x \in R : x \leq 0\}$

$\{x \in Z : -1 < x\}$

$\{x \in Q : -1 < x\}$

$\{x \in R : x < -1 \text{ ó } x > 2\}$

$\{x \in Z : -3 < x < 10\}$

$\{x \in Q : -3 \leq x < 10\}$

$\{x \in N : x > 2\}$

$\{x \in Q : x > 2\}$

$\{x \in Z : x^2 > 0\}$

$\{x \in Q : x > 0 \text{ y } x^2 < 2\}$

$$\{x \in Q : x > 0 \text{ y } x^2 < 2\} \subset R$$

20. a) Es cierto que: Todo conjunto de enteros acotado superiormente tiene extremo

superior en Z? ¿qué sucede con los acotados inferiormente?.

b) Qué pasa si en lugar de Z se considera Q? ¿y considerando subconjuntos de R?.

c) Indicar dos conjuntos de números racionales acotados superiormente sin extremo superior en Q. Idem para inferiormente.

d) Dar dos conjuntos diferentes de números racionales que tengan el mismo extremo superior $\sqrt{2}$.

21. ¿Cuál es el extremo superior del conjunto de números racionales $\frac{a}{b}$ tales que $a^3 < 11\ b^3$?. Probar la respuesta.

22. a) Sea S un subconjunto de números reales con extremo superior p y extremo inferior q. Demostrar que:

i) El conjunto $S' = \{3x : x \in S\}$ tiene extremo superior 3p y extremo inferior 3q.

ii) El conjunto $S'' = \{x+7 : x \in S\}$ tiene extremo superior p+5 y extremo inferior q+5.

b) Sean S y T dos conjuntos acotados de números reales positivos, p y m sus extremos superiores y consideremos los conjuntos $S+T = \{x+y : x \in S\ ,\ y \in T\}$, $S.T = \{x.y : x \in S\ ,\ y \in T\}$. Demostrar que p+m es el extremo superior de S+T y p.m el extremo superior de S.T .

23. a) Demostrar las propiedades de la potenciación de números reales de exponente entero, las de la raíz aritmética y las de la potenciación de exponente racional enumeradas en los parágrafos 2.6. , 2.7. y 2.8.

b) Escribir bajo la forma de una potencia de x las siguientes expresiones:

a) $x \sqrt{x^{-1}\ \sqrt{x^{-1}}}$

b) $x^{-2}\ \sqrt[3]{x^2\ \sqrt{x^{-7}}}$

c) $\sqrt{x^3\ \sqrt{x^{-1}}} \ .\ \sqrt[3]{x^{-1}}$

d) $\dfrac{x^{-1/4}\ .\ \sqrt{x}}{\sqrt[3]{x\ \sqrt{x}}}$

e) $x^2\ \sqrt[4]{x^{-3}\ \sqrt{x}}$

f) $\sqrt{x^{-2}\ \sqrt{x\ \sqrt{x^{-1}}}} \ .\ \sqrt[3]{x^2} \ .\ \sqrt{x^{-1}}$

APENDICE DEL CAPITULO II

REPRESENTACION DECIMAL DE LOS NUMEROS REALES

Una "representación" de los números reales o "sistema de numeración" es una colección de reglas que se convienen para simbolizarlos mediante ciertos signos o símbolos, es decir es una manera de notar a los números. El lector conoce por lo menos dos sistemas de numeración diferentes para los números enteros: el decimal y el romano.

El sistema decimal data del siglo VI a.C., fué usado por primera vez en la India y se originó en la operación de contar con los diez dedos de las manos. En la Edad Media los árabes lo difundieron en Europa y actualmente es el sistema de numeración adoptado en todo el mundo. En el sistema decimal los números reales se representan por sucesiones formadas por los símbolos 0,1,2,3,4,5,6,7,8,9.

Es un hecho familiar al lector que todo número real puede representarse por una expresión decimal y que esta representación proporciona una distinción clara entre números racionales y números irracionales: los números racionales se representan por una expresión decimal periódica y los irracionales por una no periódica.

Comenzaremos por ver la

2.10. REPRESENTACION DECIMAL DE LOS NUMEROS ENTEROS.

El hombre descubrió prácticamente que para contar objetos es más sencillo separarlos en grupos con un número fijo de elementos, por ejemplo, tantos como los dedos de las dos manos. Y que para contar un número grande de objetos conviene separar estos grupos a su vez de a diez, agrupar los obtenidos de a diez, etc. Es decir, dividir el número de objetos dados por 10, luego dividir por 10 el número de grupos obtenidos con diez elementos y seguir así reiterando el procedimiento hasta que no se pueda aplicar más. Esta es la esencia del sistema decimal.

En lenguaje matemático, dado un número entero positivo a, aplicando el algoritmo de la división entera se pueden efectuar divisiones sucesivas por 10 como indica el esquema:

$$
\begin{array}{lll}
a & \vert\ 10 & \\
r_o & a_1 & \vert\ 10 \\
 & r_1 & a_2 \quad \vert\ 10 \\
 & & r_2 \quad a_3 \quad \vert \\
 & & \quad \cdots \\
 & & \quad \cdots \\
 & & \quad \cdots \\
 & & a_{n-1} \quad \vert\ 10 \\
 & & r_{n-1} \quad a_n \quad \vert\ 10 \\
 & & \quad r_n = a_n \quad 0
\end{array}
$$

Como $a > a_1 > a_2 > \ldots\ldots$ reiterando las divisiones se obtiene después de un número finito de pasos un cociente a_n tal que $0 < a_n < 10$ y entonces en el próximo paso el cociente es cero y $r_n = a_n$. Se verifican las siguientes relaciones:

$$a = a_1.10 + r_o \quad , \quad 0 \leq r_o < 10$$

$$a_1 = a_2.10 + r_1 \quad , \quad 0 \leq r_1 < 10$$

. . . . . .

$$a_{n-1} = r_n.10 + r_{n-1} \quad , \quad 0 \leq r_{n-1} < 10 \quad , \quad 0 < r_n < 10$$

Reemplazando se tiene

$$a = ((((r_n.10 + r_{n-1}).10 + r_{n-2}).10 + r_{n-3}).10 + \ldots\ldots + r_1).10 + r_o$$

Luego

$$a = r_n.10^n + r_{n-1}.10^{n-1} + \ldots\ldots + r_1.10 + r_o \quad \text{con} \quad 0 \leq r_i < 10 \quad . \qquad (1)$$

Por lo tanto todo número entero positivo a puede escribirse en la forma (1) y esta representación es única en virtud de la unicidad del cociente y el resto; a está completamente determinado por los números r_n , r_{n-1} , ... , r_1 , r_o.

Si $a = r_n.10^n + r_{n-1}.10^{n-1} + \ldots\ldots + r_1.10 + r_o$ con $0 \leq r_i < 10$ se escribe abreviadamente $a = r_n\, r_{n-1} \ldots\ldots r_1\, r_o$. Esta expresión se dice la representación decimal de a.

Eligiendo los símbolos 0,1,2,3,4,5,6,7,8,9 para notar ordenadamente a los enteros no negativos menores que 10, todo número entero se representa entonces por una sucesión finita de esos símbolos.

Por ejemplo, si

$$a = 3.10^3 + 4.10^2 + 0.10 + 9$$

su representación decimal es el símbolo 3409.

Si a es un entero negativo la representación decimal de a se obtiene escribiendo la de $|a|$ precedida por el signo - .

OBSERVACION.

Estamos tan habituados a usar el sistema de representación decimal que posiblemente el lector no haya reflexionado que la elección del número 10 es completamente arbitraria (si tuviéramos tres dedos en cada mano probablemente el número elegido sería 6) y el razonamiento es igualmente válido tomando en lugar de 10 un número entero $b > 1$ cualquiera. Así todo número entero positivo a se puede escribir de una única manera como

$$a = r_n.b^n + r_{n-1}.b^{n-1} + \ldots\ldots + r_1.b + r_o \quad \text{con} \quad 0 \leq r_i \leq b-1$$

Luego a está completamente determinado por los números r_o , r_1 ,..., r_n . Abreviadamente se escribe $a = r_n\ r_{n-1}.....r_1\ r_o$ y se dice que ésta es la representación de a en base b. Si se convienen ciertos símbolos para representar a los enteros no negativos menores que b, escribiendo los r_i de esa manera resulta que todo entero a se puede representar por una sucesión finita de dichos símbolos.

Por ejemplo, la representación binaria del entero 23 es 10111 (conviniendo en elegir 0,1 para simbolizar los enteros no negativos menores que 2), pues:

$$23 = 1.2^4 + 0.2^3 + 1.2^2 + 1.2 + 1$$

La representación de 23 en base 5 es 43 (eligiendo 0,1,2,3,4 para notar los enteros no negativos menores que 5) pues:

$$23 = 4.5 + 3$$

Si se elige a 12 como base de numeración, entonces para representar a los enteros no negativos menores que 12 se pueden elegir los símbolos 0,1,....,9 pero además se necesitan otros dos que pueden ser, por ejemplo, α y β. (No se pueden usar 10 y 11 porque daría lugar a confusiones).

Entonces la representación de 23 en base 12 es 1β pues

$$23 = 1.12 + 11 = 1.12 + \beta$$

Observemos que en un sistema de numeración de base b, el número b se representa 10.

Cuando se trabaja con sistemas de numeración de bases diferentes para evitar confusiones se escribe la base en sistema decimal como subíndice a la derecha de la representación. Así las representaciones de $23_{(10)}$ vistas más arriba se escriben:

$$10111_{(2)} \quad , \quad 43_{(5)} \quad , \quad 1\beta_{(12)}$$

EJERCICIO. Verifique el lector que las expresiones

$$111100110_{(2)} \quad , \quad 2130_{(6)} \quad , \quad 226_{(15)}$$

representan al mismo número entero.

Si en lugar del sistema decimal se usa la representación b-ádica de los enteros, con $b \neq 10$, las operaciones de suma, multiplicación, diferencia y división de números enteros escritos en la base b se efectúan en forma similar que en el sistema decimal pero es necesario conocer las tablas de la suma y la multiplicación para los enteros 0,1,2,.....,b-1 que desempeñan el papel de "dígitos" en la base b.

Por ejemplo, si b = 5 las tablas de la suma y la multiplicación son las siguientes:

+	0	1	2	3	4
0	0	1	2	3	4
1		2	3	4	10
2			4	10	11
3				11	12
4					13

·	0	1	2	3	4
0	0	0	0	0	0
1		1	2	3	4
2			4	11	13
3				14	22
4					31

EJERCICIO. Escribir 116 y 179 en base 5. Hallar su suma y producto.

116	5	
16	23	5
1	3	4

179	5		
29	35	5	
4	0	7	5
	2	1	

Luego

$116 = 4.5^2 + 3.5 + 1$; la representación en base 5 es 431.

$179 = 1.5^3 + 2.5^2 + 0.5 + 4$; " " " " " " 1204.

Suma: Disponiéndola en forma práctica

$$\begin{array}{r} 1204 \\ +\ 431 \\ \hline \end{array}$$

4 + 1 = 10 ; escribimos 0 y llevamos 1 ; 1 + 0 + 3 = 4 ;

2 + 4 = 11 ; escribimos 1 y llevamos 1 ; 1 + 1 = 2.

Luego la suma es 2140 en base 5 y 295 en base 10.

Producto:

$$\begin{array}{r} 1204 \\ 431 \\ \hline 1204 \\ 4122 \\ 10331 \\ \hline 1131024 \end{array}$$

El producto expresado en base 5 es 1131024 y en base 10 es 20764.

Se plantea naturalmente el siguiente problema: Conocida la representación de un en-

tero en una cierta base b, cómo obtener su representación en otra base distinta b'?.

i) Dada la representación decimal de un número entero, su representación en una base b $\neq$ 10 se obtiene, como ya dijimos, efectuando divisiones sucesivas por b en el sistema decimal hasta obtener un cociente nulo y escribiendo la sucesión de los restos en orden inverso a como fueron obtenidos, notándolos con los símbolos elegidos para representar en la base b a los enteros no negativos menores que b.

ii) Recíprocamente, dada la representación de un número entero a en base b $\neq$ 10 , $a = r_n\, r_{n-1} \ldots r_1\, r_o$ se escribe :

$$a = r_n.b^n + r_{n-1}.b^{n-1} + \ldots + r_1.b + r_o$$

se expresan los r_i y b en el sistema decimal y se efectúan las operaciones indicadas en el sistema decimal, obteniéndose así la representación decimal de a.

iii) Dada la representación de un entero en una base b $\neq$ 10, para encontrar su expresión en otra base b' $\neq$ 10 se pasa primero al sistema decimal y luego de aquí al de base b', según lo indicado en i) y ii).

Si se sabe calcular en el sistema de base b, entonces se puede encontrar directamente la representación en la base b' por un procedimiento similar al indicado en i) es decir efectuando divisiones sucesivas por b' en el sistema de base b según el esquema conocido. Si en cambio se sabe calcular en el sistema de base b', se puede encontrar la representación en base b' aplicando un método semejante al indicado en ii).

Por ejemplo, verifiquemos los resultados obtenidos para la suma y el producto en base 5 en el ejercicio anterior. La suma S en base 5 era 2140. Entonces

$$S = 2.5^3 + 1.5^2 + 4.5 + 0$$

y efectuando las operaciones en el sistema decimal se obtiene la expresión de S en base 10: S = 295 , resultado que coincide con el que se obtiene sumando $116_{(10)}$ más $179_{(10)}$ en el sistema decimal.

Análogamente, el producto P en base 5 era 1131024. Luego

$$P = 1.5^6 + 1.5^5 + 3.5^4 + 1.5^3 + 0.5^2 + 2.5 + 4$$

y efectuando los cálculos en el sistema decimal se obtiene la representación de P en base 10: P = 20764 , resultado que coincide con el que se obtiene efectuando el producto de $116_{(10)}$ por $179_{(10)}$ en el sistema decimal.

Dados los enteros $8_{(10)}$, $1011_{(2)}$ y $301_{(7)}$ encontrar sus respectivas expresiones en cada una de las tres bases:

Representación en base 10:	8	11	148
" " " 2:	1000	1011	10010100
" " " 7	11	14	301

2.11. REPRESENTACION DECIMAL DE LOS NUMEROS REALES.

Vamos a ver que todos los números reales tienen una representación decimal y que todo decimal representa a un número real. Consideraremos los números positivos porque la representación decimal de un número real negativo x se obtiene anteponiendo el signo $-$ al decimal que representa a $|x|$.

Para fijar ideas, consideremos, por ejemplo, la expresión 348.2305. El lector sabe que este símbolo representa al número

$$348 + \frac{2}{10} + \frac{3}{10^2} + \frac{0}{10^3} + \frac{5}{10^4}$$

donde 348 representa por su parte al entero $3.10^2 + 4.10 + 8$.

Llamando x al número representado por 348.2305 , la primera cifra decimal está diciendo que

$$348 + \frac{2}{10} \leq x < 348 + \frac{3}{10}$$

La segunda cifra decimal establece que

$$348 + \frac{2}{10} + \frac{3}{100} \leq x < 348 + \frac{2}{10} + \frac{4}{100}$$

La tercera establece que

$$348 + \frac{2}{10} + \frac{3}{100} + \frac{0}{1000} \leq x < 348 + \frac{2}{10} + \frac{3}{100} + \frac{1}{1000}$$

y finalmente la cuarta que

$$348 + \frac{2}{10} + \frac{3}{100} + \frac{0}{1000} + \frac{5}{10000} = x < 348 + \frac{2}{10} + \frac{3}{100} + \frac{0}{1000} + \frac{6}{10000}$$

De modo que x aparece como una suma de un número entero más fracciones de denominador igual a una potencia de 10 y numerador comprendido entre 0 y 9 .

Pero hay números cuya representación decimal tiene infinitas cifras. Por ejemplo, la expresión decimal de $\frac{1}{3}$ es 0,3333..... y en este caso no se puede escribir

$$\frac{1}{3} = \frac{3}{10} + \frac{3}{10^2} + \frac{3}{10^3} + \ldots\ldots\ldots$$

porque en R sólo está definida la suma de un número finito de sumandos.

Precisemos lo que se entiende por decimal y representación decimal.

Definición 1. Un decimal es una sucesión infinita de enteros $a_o.a_1\ a_2\ a_3\ldots\ldots$ tales que: $a_o \geq 0$ y $0 \leq a_i < 10$ para $i=1,2,3,\ldots\ldots$

Un decimal se dice finito si todos los a_i son cero a partir de un índice en adelante, $a_i = 0$ para $i > n$. Por ejemplo, $7.25000\ldots\ldots = 7.25$ es un decimal finito. Un decimal se dice periódico si $a_i = a_{i+p}$ para todo $i > n$, n y p enteros fijos.

Por ejemplo, $2.3465172172\ldots\ldots$ es un decimal periódico. En este ejemplo se puede tomar $n = 4$ y $p = 3$.

Observemos que un decimal finito es periódico.

Dado un decimal $a_o.a_1a_2a_3\ldots\ldots$ se pueden considerar los números racionales de la forma $a_o + \frac{a_1}{10} + \frac{a_2}{10^2} + \ldots\ldots + \frac{a_k}{10^k}$ para $k = 0,1,2,3,\ldots\ldots$

Definición 2. Se dice que un decimal $a_o.a_1a_2a_3\ldots\ldots$ es una representación decimal de un número real positivo x ó un desarrollo decimal de x si

$$a_o + \frac{a_1}{10} + \frac{a_2}{10^2} + \ldots\ldots + \frac{a_k}{10^k} \leq x \leq a_o + \frac{a_1}{10} + \frac{a_2}{10^2} + \ldots\ldots + \frac{a_k + 1}{10^k}$$

para todo $k = 1,2,3,\ldots\ldots$

Llamando $A_k = a_o + \frac{a_1}{10} + \frac{a_2}{10^2} + \ldots\ldots + \frac{a_k}{10^k}$ para todo entero $k \geq 0$, las desigualdades de arriba se escriben:

$$A_k \leq x \leq A_k + \frac{1}{10^k} \qquad \text{para } k = 1,2,3,\ldots\ldots$$

Un decimal no puede representar a dos números reales distintos. En efecto, supongamos que el decimal $a_o.a_1a_2a_3\ldots\ldots$ representa a los números x y x'. Entonces se verifican las desigualdades

$$A_k \leq x \leq A_k + \frac{1}{10^k}$$

$$A_k \leq x' \leq A_k + \frac{1}{10^k}$$

para $k = 1,2,3,\ldots\ldots$

Luego

$$|x - x'| \leq x - A_k \leq \frac{1}{10^k} \qquad \text{para } k = 1,2,3,\ldots\ldots$$

lo que implica $|x - x'| = 0$ o sea $x = x'$.

Recalquemos que un decimal no es un número sino una sucesión infinita de enteros comprendidos entre 0 y 9, excepto el primero, de acuerdo con la definición 1.

1°). Todo número racional tiene una representación decimal y ésta es periódica.

Representemos en forma decimal, por ejemplo, el número $\frac{4}{11}$. Se hacen divisiones sucesivas según el esquema siguiente:

```
4          | 11
40           0.3636.....
 70
  40
   70
    4
     .
      .
       .
        .
```

Los restos sucesivos son 4,7,4,7,..... y el proceso no termina. Se tiene el siguiente conjunto infinito de igualdades:

$$4.10 = 3.11 + 7$$
$$7.10 = 6.11 + 4$$
$$4.10 = 3.11 + 7$$
$$7.10 = 6.11 + 4$$
$$\vdots \qquad \vdots$$

De la primera resulta: $\frac{4}{11} = \frac{3}{10} + \frac{7}{11} \cdot \frac{1}{10}$ (1)

De la segunda se tiene: $\frac{7}{11} = \frac{6}{10} + \frac{4}{11} \cdot \frac{1}{10}$ valor que reemplazado en (1) da

$$\frac{4}{11} = \frac{3}{10} + \frac{6}{10^2} + \frac{4}{11} \cdot \frac{1}{10^2} \qquad (2)$$

De la tercera igualdad se tiene: $\frac{4}{11} = \frac{3}{10} + \frac{7}{11} \cdot \frac{1}{10}$ valor que reemplazado en (2) da

$$\frac{4}{11} = \frac{3}{10} + \frac{6}{10^2} + \frac{3}{10^3} + \frac{7}{11} \cdot \frac{1}{10^3}$$

De esta manera se deducen las infinitas igualdades:

$$\frac{4}{11} = \frac{3}{10} + \frac{7}{11} \cdot \frac{1}{10}$$

$$\frac{4}{11} = \frac{3}{10} + \frac{6}{10^2} + \frac{4}{11} \cdot \frac{1}{10^2}$$

$$\frac{4}{11} = \frac{3}{10} + \frac{6}{10^2} + \frac{3}{10^3} + \frac{7}{11} \cdot \frac{1}{10^3}$$

$$\frac{4}{11} = \frac{3}{10} + \frac{6}{10^2} + \frac{3}{10^3} + \frac{6}{10^4} + \frac{4}{11} \cdot \frac{1}{10^4}$$

$$\begin{array}{cccccc} \vdots & & \vdots \\ \vdots & & \vdots \\ \vdots & & \vdots \\ \vdots & & \vdots \\ \vdots & & \vdots \end{array}$$

de las que resultan las infinitas desigualdades siguientes:

$$\frac{3}{10} < \frac{4}{11} < \frac{4}{10}$$

$$\frac{3}{10} + \frac{6}{10^2} < \frac{4}{11} < \frac{3}{10} + \frac{7}{10^2}$$

$$\frac{3}{10} + \frac{6}{10^2} + \frac{3}{10^3} < \frac{4}{11} < \frac{3}{10} + \frac{6}{10^2} + \frac{4}{10^3}$$

$$\frac{3}{10} + \frac{6}{10^2} + \frac{3}{10^3} + \frac{6}{10^4} < \frac{4}{11} < \frac{3}{10} + \frac{6}{10^2} + \frac{3}{10^3} + \frac{7}{10^4}$$

$$\begin{array}{ccccc} \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots \end{array}$$

De estas desigualdades sigue que el decimal 0.3636..... representa al número $\frac{4}{11}$, de acuerdo con la Definición 2.

Supongamos que se trata ahora de representar en forma decimal el número $\frac{19}{8}$.

$$
\begin{array}{r|l}
19 & 8 \\
\cline{2-2}
30 & 2.37500..... \\
60 & \\
40 & \\
0 & \\
0 &
\end{array}
$$

Entonces

$$19 = 2.8 + 3 \quad ; \quad \frac{19}{8} = 2 + \frac{3}{8}$$

$$3.10 = 3.8 + 6 \quad ; \quad \frac{3}{8} = \frac{3}{10} + \frac{6}{8} \cdot \frac{1}{10}$$

$$6.10 = 7.8 + 4 \quad ; \quad \frac{6}{8} = \frac{7}{10} + \frac{4}{8} \cdot \frac{1}{10}$$

$$4.10 = 5.8 + 0 \quad ; \quad \frac{4}{8} = \frac{5}{10}$$

y de estas igualdades resultan las siguientes desigualdades:

$$2 + \frac{3}{10} < \frac{19}{8} < 2 + \frac{4}{10}$$

$$2 + \frac{3}{10} + \frac{7}{10^2} < \frac{19}{8} < 2 + \frac{3}{10} + \frac{8}{10^2}$$

$$2 + \frac{3}{10} + \frac{7}{10^2} + \frac{5}{10^3} = \frac{19}{8} < 2 + \frac{3}{10} + \frac{7}{10^2} + \frac{6}{10^3}$$

$$2 + \frac{3}{10} + \frac{7}{10^2} + \frac{5}{10^3} + \frac{0}{10^4} = \frac{19}{8} < 2 + \frac{3}{10} + \frac{7}{10^2} + \frac{5}{10^3} + \frac{1}{10^4}$$

Por la Definición 2, el decimal 2.37500..... representa al número $\frac{19}{8}$. En este ejemplo se obtiene un resto cero en el algoritmo de la división y el proceso termina en el sentido que de ahí en adelante todas las cifras que se obtienen son ceros. Abreviadamente el decimal se escribe 2.375 .

En general, todo número racional tiene una representación decimal. En efecto, por el procedimiento indicado cada número racional positivo $\frac{a}{b}$ se representa por un decimal periódico.

Se hacen divisiones como indica el esquema (se supone $b > 0$):

$$\begin{array}{l|l} a & b \\ \hline r_1 \cdot 10 & a_o . a_1 a_2 a_3 \ldots\ldots \\ r_2 \cdot 10 & \\ r_3 \cdot 10 & \\ \vdots & \end{array}$$

donde a_o y r_1 son el cociente y el resto de dividir a por b, a_i y r_{i+1} los de dividir $r_i \cdot 10$ por b, para $i=1,2,3,\ldots\ldots$ Como los restos posibles son $0,1,2,\ldots\ldots,b-1$ se ve que a lo sumo después de b pasos se obtiene un resto igual a uno de los ya obtenidos y a partir de ahí en adelante se repiten todos los cocientes y los restos en ciclos periódicos. Entonces $\frac{a}{b}$ está representado por el decimal periódico $a_o . a_1 a_2 a_3 \ldots\ldots$ pues se verifican las desigualdades:

$$a_o + \frac{a_1}{10} + \frac{a_2}{10^2} + \ldots\ldots + \frac{a_k}{10^k} \leq \frac{a}{b} < a_o + \frac{a_1}{10} + \frac{a_2}{10^2} + \ldots\ldots + \frac{a_k + 1}{10^k}$$

para $k = 1,2,3,\ldots\ldots$

lo que se demuestra en forma análoga a la vista en los ejemplos y queda a cargo del lector.

En particular, si $\frac{a}{b}$ es un número entero, es decir si $b = 1$, aplicando el procedimiento indicado resulta que el número entero a está representado por el decimal $a.0000\ldots\ldots\ldots\ldots$

Notemos que si en los b primeros pasos se obtiene un resto nulo entonces de ahí en adelante todas las cifras decimales son cero y la representación decimal de $\frac{a}{b}$ es un decimal finito. Si en cambio ninguno de esos restos es nulo entonces el decimal que representa a $\frac{a}{b}$ es periódico infinito. ¿Cuándo se presenta uno u otro caso?.

Veamos que un número racional $\frac{a}{b}$ admite una representación decimal finita si y solo si b es de la forma $b = 2^r.5^s$, con $r \geq 0$, $s \geq 0$.

Si b es de esa forma entonces se tiene

$$\frac{a}{b} = \frac{2^s.5^r.a}{10^{r+s}}$$

y todo número racional de la forma $\frac{m}{10^t}$ tiene una representacíón decimal finita. En efecto, expresando el entero m en base 10 es

$$m = a_n.10^n + a_{n-1}.10^{n-1} + \ldots\ldots + a_1.10 + a_o$$

y dividiendo por 10^t se ve que se obtiene una representación decimal finita de $\frac{m}{10^t}$.

Recíprocamente, supongamos que $\frac{a}{b}$ tiene una representación decimal finita $a_o.a_1 \ldots a_n$

Entonces $\frac{a}{b} = a_o + \frac{a_1}{10} + \frac{a_2}{10^2} + \ldots\ldots + \frac{a_n}{10^n}$. Sumando será $\frac{a}{b} = \frac{m}{10^n}$;

luego $10^n.\ a = b.m$

Como podemos suponer a y b primos entre sí de la última igualdad resulta que b divide a 10^n y en consecuencia los únicos factores primos que admite b son 2 y 5 , es decir $b = 2^r.5^s$ con $r \geq 0$, $s \geq 0$.

2°). <u>Para todo decimal periódico $A = a_o.a_1a_2\ldots\ldots\ldots$ existe un número racional positivo r tal que A es una representación decimal de r.</u>

Si $A = a_o.a_1a_2\ldots\ldots$ es un decimal periódico, $a_i = a_{i+p}$ para $i > n$, n y p enteros fijos, entonces A es un desarrollo decimal del número racional

$$r = \frac{10^p.\ A_{n+p} - A_n}{10^p - 1}$$

fórmula que es familiar al lector desde el colegio secundario y que en palabras se traduce como sigue: Un decimal periódico $a_o.a_1a_2\ldots\ldots\ldots$ representa al número racional r igual al entero a_o más una fracción cuyo numerador se forma escribiendo la parte no periódica seguida del período menos la parte no periódica y cuyo denominador es el número que se obtiene escribiendo tantos 9 como cifras tiene el pe -

ríodo seguidos de tantos ceros como cifras tiene la parte no periódica.

Por ejemplo, 2.91348348..... es la representación decimal del número racional

$$r = 2 + \frac{91348 - 91}{99900} = \frac{191057}{99900}$$

0.47222..... es la representación decimal del número racional

$$r = \frac{472 - 47}{900} = \frac{17}{36}$$

Demostremos la afirmación hecha más arriba.

Sea $A = a_o.a_1a_2a_3.....$ un decimal periódico, $a_i = a_{i+p}$ para $i > n$, n y p enteros fijos, y consideremos el número racional

$$r = \frac{10^p A_{n+p} - A_n}{10^p - 1}$$

se tiene

$$r = \frac{10^p A_{n+p} - A_n}{10^p - 1} = A_n + \frac{10^p(A_{n+p} - A_n)}{10^p - 1} =$$

$$= A_n + 10^p \cdot \frac{\frac{a_{n+1}}{10^{n+1}} + ... + \frac{a_{n+p}}{10^{n+p}}}{10^p - 1} =$$

$$= A_n + \frac{1}{10^n} \cdot \frac{a_{n+1}.10^{p-1} + ... + a_{n+p}}{10^p - 1}$$

Llamando $r_1 = a_{n+1}.10^{p-1} + ... + a_{n+p}$ es

$$r = A_n + \frac{1}{10^n} \cdot \frac{r_1}{10^p - 1}$$

Caso 1: Si $p = 1$ y $a_{n+1} = 9$, es decir si el decimal es de la forma

$A = a_o.a_1a_2.....a_n999.....$ entonces $\frac{r_1}{10^p - 1} = \frac{9}{9} = 1$ y

$$r = A_n + \frac{1}{10^n}$$

Para demostrar que el decimal A representa al número racional r hay que probar, de acuerdo con la Definición 2, que

$$A_k \leq r \leq A_k + \frac{1}{10^k} \quad , \quad \forall \ k = 1,2,.....$$

Se ve fácilmente que la doble desigualdad se verifica para $k = 1,2,.....,n$.

Si $k > n$ se tiene

$$A_k = A_n + \frac{9}{10^{n+1}} + ...+ \frac{9}{10^k} = A_n + \frac{10^{k-n} - 1}{10^k} \leq A_n + \frac{1}{10^n} = r$$

Por otro lado

$$A_k + \frac{1}{10^k} = A_n + \frac{9}{10^{n+1}} + ...+ \frac{9}{10^k} + \frac{1}{10^k} = A_n + \frac{1}{10^n} = r$$

Luego $$A_k \leq r \leq A_k + \frac{1}{10^k} \quad , \quad \forall \ k = 1,2,.....$$

Notemos que en este caso el número r está representado también por el decimal finito $a_o.a_1 a_2.....a_{n-1}(a_n + 1)$.

Caso 2: Supongamos el período del decimal dado distinto de 9. Como

$$r = A_n + \frac{1}{10^n} \cdot \frac{r_1}{10^p - 1}$$

veamos cuál es la representación decimal del número racional $\frac{r_1}{10^p - 1}$. Siguiendo el procedimiento indicado en 1°) podemos obtener una representación decimal de $\frac{r_1}{10^p - 1}$ efectuando divisiones sucesivas de acuerdo con el esquema ya visto.

$$\frac{10.r_1}{10^p - 1} = a_{n+1} + \frac{a_{n+2}.10^{p-1} + ...+ a_{n+p}.10 + a_{n+1}}{10^p - 1}$$

y como $r_2 = a_{n+2}.10^{p-1} + ...+ a_{n+p}.10 + a_{n+1} < 10^p - 1$

por ser el período distinto de 9, resulta que a_{n+1} es el cociente y r_2 es el resto de dividir el entero $10.r_1$ por $10^p - 1$.

$$\frac{10.r_2}{10^p - 1} = a_{n+2} + \frac{a_{n+3}.10^{p-1} + \ldots\ldots + a_{n+p}.10^2 + a_{n+1}.10 + a_{n+2}}{10^p - 1}$$

y $$r_3 = a_{n+3}.10^{p-1} + \ldots\ldots + a_{n+p}.10^2 + a_{n+1}.10 + a_{n+2} < 10^p - 1$$

Luego a_{n+2} y r_3 son el cociente y el resto respectivamente de dividir $10.r_2$ por $10^p - 1$. Reiterando el procedimiento en el paso p-ésimo se tiene

$$\frac{10.r_p}{10^p - 1} = a_{n+p} + \frac{a_{n+1}.10^{p-1} + \ldots\ldots + a_{n+p}}{10^p - 1}$$

y por lo tanto $r_{p+1} = r_1$.

De acuerdo con lo demostrado en el punto 1°), resulta que el número racional $\frac{r_1}{10^p - 1}$ está representado por el decimal periódico

$$0.a_{n+1}a_{n+2}\ldots\ldots a_{n+p}a_{n+1}a_{n+2}\ldots\ldots a_{n+p}\ldots\ldots\ldots\ldots\ldots$$

Entonces como el número r es

$$r = A_n + \frac{1}{10^n} \cdot \frac{r_1}{10^p - 1} = a_o + \frac{a_1}{10} + \ldots\ldots + \frac{a_n}{10^n} + \frac{1}{10^n} \cdot \frac{r_1}{10^p - 1}$$

sigue fácilmente que una representación decimal de r es el decimal dado

$$A = a_o.a_1\ldots\ldots a_n a_{n+1}\ldots\ldots a_{n+p}a_{n+1}\ldots\ldots a_{n+p}\ldots\ldots\ldots$$

como queríamos demostrar.

Observemos que del Caso 1 resulta que hay números racionales que admiten dos repre - sentaciones decimales distintas. Por ejemplo , $\frac{1}{2}$ puede escribirse 0.5 ó 0.4999...

3°). Veamos ahora que si $x \in R$ es un número irracional también tiene una representación decimal.

En R vale la siguiente propiedad: Dado un número real z cualquiera existe un único número entero n tal que

$$n \leq z < n+1$$

Aplicando esta propiedad, sea a_o el entero tal que

$$a_o < x < a_o + 1$$

(son desigualdades estrictas puesto que x es un número irracional).

Aplicando la misma propiedad al número $10(x - a_o)$, existe un entero a_1 tal que

$$a_1 < 10(x - a_o) < a_1 + 1$$

Como $x - a_o < 1$ es $0 \leq a_1 \leq 9$. Luego

$$a_o + \frac{a_1}{10} < x < a_o + \frac{a_1 + 1}{10} \quad \text{con} \quad 0 \leq a_1 \leq 9$$

En general, si $a_o + \frac{a_1}{10} + \ldots\ldots + \frac{a_k}{10^k} < x < a_o + \frac{a_1}{10} + \ldots\ldots + \frac{a_k + 1}{10^k}$ existe un entero a_{k+1} tal que

$$a_{k+1} < 10^{k+1}(x - a_o - \frac{a_1}{10} - \ldots\ldots - \frac{a_k}{10^k}) < a_{k+1} + 1$$

Luego

$$a_o + \frac{a_1}{10} + \ldots\ldots + \frac{a_{k+1}}{10^{k+1}} < x < a_o + \frac{a_1}{10} + \ldots\ldots + \frac{a_{k+1} + 1}{10^{k+1}} \quad \text{con} \quad 0 \leq a_{k+1} \leq 9$$

Por lo tanto existen enteros $a_o, a_1, a_2, a_3, \ldots\ldots$ con $0 \leq a_k \leq 9$ para $k = 1,2,3,\ldots$ y tales que

$$A_k < x < A_k + \frac{1}{10^k} \quad , \quad \forall \quad k = 1,2,3,\ldots\ldots\ldots$$

siendo $A_k = a_o + \frac{a_1}{10} + \ldots\ldots + \frac{a_k}{10^k}$.

De acuerdo con la Definición 2, el decimal infinito $a_o.a_1a_2a_3.....$ representa al número x.

Por ejemplo, consideremos el número irracional $\sqrt{2}$. Como primera aproximación se tiene $A_o = 1$ pues

$$1 < \sqrt{2} < 1 + 1 \quad \text{ya que} \quad 1^2 < 2 < 2^2$$

Se busca una mejor aproximación entre los números 1.1 , 1.2 , 1.3 ,....., 1.9 . Así $A_1 = 1.4$ pues $1.4 < \sqrt{2} < 1.5$ ya que $1.4^2 = 1.96$ y $1.5^2 = 2.25$. Buscando una mejor aproximación entre los números 1.41 , 1.42 ,....., 1.49 se obtiene $A_2 = 1.41$ pues $1.41 < \sqrt{2} < 1.42$ ya que $1.41^2 = 1.9881$ y $1.42^2 = 2.0164$. Continuando de esta manera se obtiene la sucesión de números racionales

$$1 \quad , \quad 1.4 \quad , \quad 1.41 \quad , \quad 1.414 \quad , \quad 1.4142 \quad , \ldots\ldots\ldots \text{ y}$$

este proceso determina un decimal no periódico 1.414213..... que representa a $\sqrt{2}$.

OBSERVACION. La demostración anterior prueba la existencia de una representación decimal para cada número irracional pero no proporciona un método para conocer las cifras del decimal en cuestión. Notemos que esta demostración sirve igualmente para probar que todo número racional tiene una representación decimal, pero la que vimos para números racionales es más útil por cuanto permite conocer todas las cifras del decimal correspondiente. Cuando se trata de calcular la expresión decimal de una raíz cuadrada, por ejemplo, se pueden obtener las primeras cifras mediante tanteos sucesivos como indicamos. Pero en general calcular el valor aproximado de la expresión decimal de un número irracional cualquiera, como por ejemplo π , e, ln 2, etc., no es tan sencillo y hay que desarrollar un método especial para cada caso.

Hemos visto entonces que todo número real tiene una representación decimal. Veamos ahora que

4°). Todo decimal representa a un número real.

Sea $A = a_o.a_1a_2.....$ un decimal. Consideremos el conjunto $\{A_o, A_1, A_2, ...\}$ donde

$$A_k = a_o + \frac{a_1}{10} + \ldots\ldots + \frac{a_k}{10^k} \qquad \text{para} \quad k = 0,1,2,.....$$

Este conjunto de números racionales es acotado superiormente y por lo tanto tiene extremo superior x en R. Se trata de probar que el decimal dado representa a x, es decir que se verifican las desigualdades:

$$A_1 \leq x \leq A_1 + \frac{1}{10}$$

$$\vdots \quad \vdots \quad \vdots$$

$$A_k \leq x \leq A_k + \frac{1}{10^k}$$

$$\vdots \quad \vdots \quad \vdots$$

para todo índice $k = 1,2,3,.....$

En virtud de la definición de extremo superior las desigualdades de la izquierda se verifican todas. Falta probar las de la derecha, es decir que

$$x \leq A_k + \frac{1}{10^k} \quad \text{para} \quad k = 1,2,3,.....$$

Para ello será suficiente demostrar que cada número $A_k + \frac{1}{10^k}$ es una cota superior del conjunto $\{A_o, A_1, A_2, ...\}$.

$$A_o < A_o + 1$$

$$A_1 < A_1 + \frac{1}{10}$$

$$A_2 < A_2 + \frac{1}{10^2}$$

$$\vdots \quad \vdots$$

(1)

$$A_k < A_k + \frac{1}{10^k}$$

$$\vdots \quad \vdots$$

Por otro lado, $(A_k + \frac{1}{10^k}) - A_{k-1} = \frac{a_k + 1}{10^k} \leq \frac{1}{10^{k-1}}$ puesto que $0 \leq a_k \leq 9$. Entonces

$$(2) \qquad A_k + \frac{1}{10^k} \leq A_{k-1} + \frac{1}{10^{k-1}} \quad \text{para} \quad k = 1,2,3,.....$$

Dado un índice k, de las desigualdades (1) y (2) se tiene para cada índice $i \geq k$:

$$A_i < A_i + \frac{1}{10^i} \leq A_{i-1} + \frac{1}{10^{i-1}} \leq \ldots\ldots\ldots \leq A_k + \frac{1}{10^k}$$

Si $i < k$

$$A_i \leq A_{i+1} \leq A_{i+2} \leq \ldots\ldots\ldots \leq A_k < A_k + \frac{1}{10^k}$$

Por lo tanto para todo $i = 0,1,2,\ldots\ldots\ldots$ se tiene:

$$A_i < A_k + \frac{1}{10^k}$$

Luego $A_k + \frac{1}{10^k}$ es una cota superior del conjunto $\{A_o, A_1, A_2, \ldots\ldots\ldots\}$, $\forall$ $k=1,2,\ldots$

Como x es el extremo superior de ese conjunto se tiene

$$x \leq A_k + \frac{1}{10^k} \qquad , \quad \forall \quad k = 1,2,\ldots\ldots$$

lo que termina la demostración.

Hemos probado que todo número real tiene una representación decimal y que todo decimal representa a un único número real. Además, de la última demostración resulta que un decimal $a_o.a_1a_2\ldots\ldots\ldots$ representa al número real x que es el extremo superior del conjunto de números racionales

$$\{a_o \ , \ a_o + \frac{a_1}{10} \ , \ a_o + \frac{a_1}{10} + \frac{a_2}{10^2} \ , \ \ldots\ldots\ldots\}$$

Con respecto a la unicidad de la representación vimos que hay números racionales que tienen dos representaciones decimales, por ejemplo $\frac{5}{4}$ se puede representar 1.25 ó 1.24999..... . Para terminar este apéndice vamos a demostrar que la representación decimal de un número real es única salvo en el caso de los números reales que admiten una representación decimal finita, los que tienen dos representaciones decimales distintas.

Probaremos que:

Si $A = a_o.a_1a_2a_3\ldots\ldots$ y $B = b_o.b_1b_2b_3\ldots\ldots$ son dos representaciones decimales

distintas de un número real x entonces una de ellas es finita, por ejemplo la primera, es decir existe un índice n tal que $a_n \neq 0$ y $a_i = 0$ para $i > n$ y en cuanto a la representación B es tal que $b_i = a_i$ para $0 \leq i < n$, $b_n = a_n - 1$ y $b_i = 9$ para $i > n$.

Como A y B se suponen decimales diferentes, sea n el menor índice tal que $a_n \neq b_n$ y supongamos $a_n > b_n$.

$$A_n = a_o + \frac{a_1}{10} + \ldots\ldots + \frac{a_n}{10^n}$$

$$B_n = b_o + \frac{b_1}{10} + \ldots\ldots + \frac{b_n}{10^n}$$

Por nuestra hipótesis sobre el índice n es

$$A_n - B_n = \frac{a_n - b_n}{10^n} \geq \frac{1}{10^n} \qquad (1)$$

Por otro lado, como A y B son representaciones decimales de x es

$$A_n \leq x \leq A_n + \frac{1}{10^n}$$

$$B_n \leq x \leq B_n + \frac{1}{10^n}$$

Luego $A_n - B_n \leq x - B_n \leq \frac{1}{10^n}$ (2). De (1) y (2) resulta $A_n - B_n = \frac{1}{10^n}$ o sea $\frac{a_n - b_n}{10^n} = \frac{1}{10^n}$, lo que implica $b_n = a_n - 1$, que es una de las cosas que queríamos demostrar.

Veamos ahora que A es un decimal finito.

Para ello probemos que $B_k \leq A_n$, $\forall\ k = 0,1,2,\ldots\ldots$

Si $k \leq n$ entonces $B_k \leq A_n$.

Si $k > n$

$$B_k = b_o + \frac{b_1}{10} + \ldots\ldots + \frac{b_n}{10^n} + \ldots\ldots + \frac{b_k}{10^k} \leq b_o + \frac{b_1}{10} + \ldots\ldots + \frac{b_n}{10^n} + \frac{9}{10^{n+1}} + \ldots\ldots + \frac{9}{10^k}$$

Pero $$\frac{9}{10^{n+1}} + \ldots\ldots + \frac{9}{10^k} = \frac{1}{10^n} \cdot \frac{10^k - 1}{10^k} < \frac{1}{10^n}$$

Luego

$$B_k < b_o + \frac{b_1}{10} + \ldots\ldots + \frac{b_n}{10^n} + \frac{1}{10^n} = A_n \qquad \text{pues} \qquad a_n = b_n + 1$$

Entonces

$$B_k \leq A_n \quad , \quad \forall \quad k = 0,1,2,\ldots\ldots$$

y A_n es una cota superior del conjunto de números racionales $\{B_o, B_1, B_2, \ldots\ldots\}$. Como el decimal B representa al número x, x es el extremo superior de dicho conjunto. Luego, por definición de extremo superior, debe ser

$$x \leq A_n$$

Por otro lado, como el decimal A representa al número x, es

$$A_n \leq x$$

De aquí resulta $x = A_n$ y por lo tanto A es un decimal finito, $a_i = 0$ para todo $i > n$, pues si existiera un $a_k \neq 0$ con $k > n$ se tendría

$$A_n < A_k \leq x$$

lo que contradice $x = A_n$.

Resta probar que $b_i = 9$ para $i > n$.

Supongamos que existe un índice $j > n$ tal que $b_j < 9$ y sea $k > j$.

$$B_k - B_n = \frac{b_{n+1}}{10^{n+1}} + \ldots\ldots + \frac{b_j}{10^j} + \ldots\ldots + \frac{b_k}{10^k}$$

Como $b_j < 9$ es

$$B_k - B_n \leq \frac{9}{10^{n+1}} + \ldots\ldots + \frac{9}{10^j} + \ldots\ldots + \frac{9}{10^k} - \frac{1}{10^j} = \frac{1}{10^n} \cdot \frac{10^{k-n} - 1}{10^{k-n}} - \frac{1}{10^j}$$

Luego

$$B_k - B_n < \frac{1}{10^n} - \frac{1}{10^j} \qquad (3)$$

Por otro lado $B_k - B_n = (x - B_n) - (x - B_k)$. Pero $x - B_n = A_n - B_n = \frac{1}{10^n}$ y $x - B_k \leq \frac{1}{10^k}$.

Luego

$$B_k - B_n \geq \frac{1}{10^n} - \frac{1}{10^k} \qquad (4)$$

De (3) y (4) sigue $\frac{1}{10^n} - \frac{1}{10^k} < \frac{1}{10^n} - \frac{1}{10^j}$

o sea $\frac{1}{10^k} > \frac{1}{10^j}$ lo que contradice nuestra hipótesis de que $k > j$. Esta contradicción prueba que $b_i = 9$ para todo $i > n$, lo que termina la demostración.

Resumiendo, todo número real tiene una representación decimal y todo decimal representa a un único número real. Un decimal $a_o.a_1a_2a_3\ldots\ldots$ representa al número real x que es el extremo superior del conjunto de números racionales $\{A_o, A_1, A_2, \ldots\ldots\}$. La representación decimal de un número real es periódica si y solo si el número es racional, es decir, un decimal representa a un número irracional si y solo si no es periódico. Además la representación decimal de un número real es única salvo en el caso de los decimales finitos. Como éstos representan a números racionales $\frac{a}{b}$ donde b es un producto de potencias de 2 y 5 se concluye que la expresión decimal de los números irracionales y la de los racionales que no son de ese tipo es única.

OBSERVACION. La elección del número 10 es arbitraria. En general, elegido un entero $b > 1$ cualquiera se demuestra análogamente que todo número real positivo se puede representar por una expresión infinita $a_o.a_1a_2a_3\ldots\ldots$ donde los a_i son enteros tales que $a_o \geq 0$ y $0 \leq a_i \leq b-1$, para $i = 1,2,3,\ldots\ldots$ y que, recíprocamente, toda expresión de este tipo representa a un número real. De aquí que la base del sistema de representación sea arbitraria, aunque la representación decimal es la más antigua

y la que se usa en la práctica.

Para terminar digamos que el sistema de numeración binario, es decir con base 2, tiene especial importancia en la vida moderna por cuanto es el sistema de numeración con el que trabajan las computadoras. Los datos que se suministran a una máquina computadora deben expresarse en base 2. Esta exigencia está determinada por la misma naturaleza de las computadoras, formadas por innumerables circuitos eléctricos cada uno de los cuales puede adoptar sólo dos estados: que pase corriente eléctrica o que no pase, y que corresponden a los valores 0 y 1.

Por ejemplo, expresemos $\frac{5}{3}$ en base 2. Las cifras de la representación correspondiente se obtienen por un procedimiento análogo al indicado en 1°), es decir efectuando divisiones sucesivas por 3 en el sistema de base 2. El número 5 en base 2 se escribe 101 y 3 se escribe 11. Luego

```
101        | 11
 100         1.1010.....
  010
   100
    010
      ..
       ..
        ..
         ..
          ..
```

y 1.1010..... es la representación binaria de $\frac{5}{3}$.

Hallemos ahora la representación binaria de $\frac{25}{8}$. El número 25 en base 2 se escribe 11001 y 8 se escribe 1000. Luego 11.001 es la representación binaria de $\frac{25}{8}$.

EJERCICIOS.

1. a) Teniendo en cuenta la expresión decimal de los enteros, demostrar los siguientes criterios de divisibilidad, familiares al lector:

. $\forall\, a \in Z$, si $a = r_n r_{n-1} \ldots\ldots r_o$ es su representación decimal entonces:

i) $2/a \iff 2/r_o$

ii) $5/a \iff 5/r_o$

iii) $3/a \iff 3/r_n + r_{n-1} + \ldots\ldots + r_o$

iv) $4/a \iff 4/r_o + 2r_1$

b) Hallar criterios de divisibilidad por 9 , 8 , 25 y 11.

2. a) Escribir en base 2 el entero $47_{(10)}$

b) " " " 5 " " $784_{(10)}$

c) " " " 3 " " $1021_{(10)}$

d) " " " 6 " " $3406_{(10)}$

3. Hallar la representación decimal de cada uno de los siguientes números enteros:

$1011011_{(2)}$, $231022_{(4)}$, $1002100_{(3)}$, $4152_{(6)}$, $1010_{(4)}$

4. a) Hallar la representación binaria del entero $241_{(6)}$

b) " " " en base 5 del entero $1201_{(3)}$

c) " " " " " 7 " " $110111_{(2)}$

5. Efectuar las siguientes operaciones en base 2:

$110111_{(2)} + 11011_{(2)}$ $\qquad$ $11011_{(2)} \times 101_{(2)}$

6. a) Hallar la representación decimal de los siguientes números racionales:

$\frac{11}{8}$, $\frac{348}{117}$, $-\frac{4035}{250}$

b) Hallar los números racionales representados por cada uno de los siguientes decimales periódicos:

$0.214175175\ldots\ldots$ $\qquad$ 6.351

$3.010202\ldots\ldots\ldots$ $\qquad$ $0.131999\ldots\ldots\ldots$

7. Hallar las primeras cuatro cifras de la representación decimal de los siguientes números reales:

$$\frac{17}{12} \quad , \quad \sqrt{3} \quad , \quad \sqrt[3]{2} \quad , \quad \frac{32}{33}$$

8. Probar que el siguiente decimal representa a un número irracional:

0.1212212221222212222221.....

9. a) Expresar $\frac{7}{15}$ en base 5, $\frac{1}{4}$ en base 5, $\frac{1}{2}$ en base 2, $\frac{11}{12}$ en base 2, $\frac{27}{32}$ en base 4.

b) Decir a qué números racionales representan cada una de las siguientes expresiones binarias periódicas:

0.1111..... 1.11010101.....

0.010101..... 11.001111.....

CAPITULO III

NUMEROS COMPLEJOS

En el sistema R de los números reales toda ecuación lineal

$$bX + a = 0 \quad , \quad a,b \in R \quad , \quad b \neq 0$$

tiene solución. (En general esto vale en cualquier cuerpo). En cambio no siempre es posible resolver una ecuación cuadrática en R. Por ejemplo, la sencilla ecuación

$$X^2 + 1 = 0$$

no tiene solución real puesto que $x^2 \neq -1$, $\forall$ $x \in R$, ya que el cuadrado de cualquier número real es positivo o nulo.

Análogamente la ecuación

$$X^4 + 2X^2 + 3 = 0$$

no tiene solución real pues $\forall\, x \in R$ es $x^2 \geq 0$, lo que implica $x^4 \geq 0$ y $2x^2 \geq 0$.

Luego $x^4 \geq 0$, $2x^2 \geq 0$ y $3 > 0 \Longrightarrow x^4 + 2x^2 + 3 > 0$, $\forall$ $x \in R$.

En general, una ecuación

$$a_n X^n + a_{n-1} X^{n-1} + \ldots\ldots + a_1 X + a_o = 0$$

donde a_n , a_{n-1} , ... , a_1 , a_o son números reales dados , $a_n \neq 0$, $n > 1$, no tiene siempre solución en R.

Para satisfacer la demanda de soluciones para todas las ecuaciones algebraicas se amplía el sistema R de los números reales al sistema C de los complejos y se verifica que toda ecuación algebraica con coeficientes reales o complejos tiene por lo menos una solución en C, propiedad que se conoce con el nombre de Teorema fundamental del álgebra y que analizaremos en el próximo capítulo. De modo que con C se cierra el problema de la existencia de raíces para las ecuaciones algebraicas con coeficientes numéricos.

DEFINICION. Sea $C = R \times R$ (es decir el conjunto de todos los pares ordenados (a,b) de números reales) y definidas en C dos operaciones binarias, la suma + y la

multiplicación $\cdot$, de la siguiente manera:

Suma: $(a,b)+(c,d) = (a+c,b+d)$

Multiplicación: $(a,b).(c,d) = (ac-bd,ad+bc)$

El conjunto C con estas dos operaciones se llama el sistema de los números complejos.

1) C es un cuerpo , es decir la suma y la multiplicación definidas en C tienen las siguientes propiedades:

S1. La suma es asociativa:

$$(z + z') + z'' = z + (z' + z'') \quad , \quad \forall\, z,z',z'' \in C$$

Se demuestra sin dificultad y es consecuencia de la asociatividad de la suma de números reales.

S2. La suma es conmutativa:

$$z + z' = z' + z \quad , \quad \forall\, z,z' \in C$$

Resulta inmediatamente de la conmutatividad de la suma de números reales.

S3. Existe neutro para la suma:

En efecto, el número complejo (0,0) es tal que

$(a,b)+(0,0) = (a,b) \quad , \quad \forall\, (a,b) \in C$

S4. Todo número complejo es inversible con respecto a la suma.

Dado el complejo (a,b),el número complejo (-a,-b) es su simétrico pues

$(a,b)+(-a,-b) = (0,0)$

M1. La multiplicación es asociativa:

$$(z.z').z'' = z.(z'.z'') \quad , \quad \forall\, z,z',z'' \in C$$

La demostración queda a cargo del lector.

M2. La multiplicación es conmutativa:

$$z.z' = z'.z \quad , \quad \forall\, z,z' \in C$$

Resulta de la conmutatividad de la suma y del producto de números reales.

M3. Existe neutro para la multiplicación.

El número complejo (1,0) es tal que

$$(a,b).(1,0) = (a,b) \quad , \quad \forall \ (a,b) \in C$$

M4. Todo número complejo diferente de (0,0) es inversible con respecto a la multiplicación.

Dado $z = (a,b) \in C$, $(a,b) \neq (0,0)$, el número complejo

$$\left(\frac{a}{a^2+b^2} \ , \ - \frac{b}{a^2+b^2} \right)$$ es inverso de z pues

$$(a,b) \cdot \left(\frac{a}{a^2+b^2} \ , - \frac{b}{a^2+b^2} \right) = (1,0)$$

D. La multiplicación es distributiva con respecto a la suma:

$$z.(z'+z'') = z.z' + z.z'' \quad , \quad \forall \ z,z',z'' \in C$$

La demostración a cargo del lector.

2) Sea R' el conjunto de los números complejos de segunda componente nula. La aplicación $f: R \longrightarrow R'$ que a cada número real x hace corresponder el número complejo (x,0) es una correspondencia biunívoca de R sobre R':

$$x \longleftrightarrow (x,0)$$

Esta correspondencia biunívoca conserva la suma y la multiplicación en el sentido que a la suma (multiplicación) de dos números reales x,y le corresponde la suma (multiplicación) de sus imágenes (x,0) , (y,0) :

$$x+y \longleftrightarrow (x,0) + (y,0)$$
$$x.y \longleftrightarrow (x,0) \cdot (y,0)$$

ya que

$$(x+y,0) = (x,0) + (y,0)$$
$$(x.y,0) = (x,0) \cdot (y,0)$$

En virtud de estas propiedades de la correspondencia biunívoca los sistemas R y R' tienen la misma estructura algebraica y se conviene en identificar el número real x con el número complejo (x,0). Se dice que R' es isomorfo a R. A través

de esta identificación R es un subconjunto de C.

Se escribe $(x,0) = x$

En particular , $(0,0) = 0$ y $(1,0) = 1$

3) C contiene una raíz de la ecuación $X^2 + 1 = 0$.

En efecto, el número complejo (0,1) es una raíz de esa ecuación pues:

$$(0,1)^2 + 1 = (0,1).(0,1) + 1 = (-1,0) + (1,0) = (0,0) = 0$$

De 1) , 2) y 3) resulta el siguiente

TEOREMA 3.1. C es un cuerpo que contiene un subcuerpo isomorfo a R y una raíz de la ecuación $X^2 + 1 = 0$.

Dado un número complejo $z = (a,b)$, a se llama la parte real de z y b la imaginaria y se escribe : $a = Re(z)$, $b = Im(z)$.

Los números de segunda componente nula se llaman complejos reales y los de primera componente nula imaginarios puros.

El número complejo (0,1) se llama unidad imaginaria y se representa $i = (0,1)$.

En 3) vimos que $i^2 = -1$.

Todo número complejo (a,b) puede escribirse de la siguiente manera:

$$(a,b) = (a,0) + (0,b) = (a,0) + (b,0).(0,1)$$

Notando: $(a,0) = a$, $(b,0) = b$, $(0,1) = i$ resulta: $(a,b) = a + bi$

Esta forma de escritura de los números complejos se llama forma binómica.

En particular, para los números imaginarios puros es:

$$(0,b) = bi$$

La forma binómica tiene la ventaja de que permite sumar y multiplicar números complejos como si fueran números reales, teniendo en cuenta tan sólo que $i^2 = -1$.

Por ejemplo, dados dos números complejos (a,b) y (c,d) su suma (a+c,b+d) se calcula en forma binómica sumando los términos semejantes:

$$(a+bi) + (c+di) = (a+c) + (b+d)i$$

Análogamente, el producto (ac-bd,ad+bc) se obtiene:

$$(a+bi).(c+di) = ac + adi + bci + bdi^2 = (ac - bd) + (ad + bc)i$$

multiplicando cada término del primer binomio por cada término del segundo, reemplazando i^2 por -1 y sumando los términos semejantes.

La diferencia se calcula así:

$$(a+bi) - (c+di) = a+bi-c-di = (a-c) + (b-d)i \quad /$$

Finalmente, para dividir un número complejo $a+bi$ por otro $c+di \neq 0$ se multiplican dividendo y divisor por el número $c-di$, llamado el conjugado de $c+di$, procediéndose como sigue:

$$\frac{a+bi}{c+di} = \frac{(a+bi)\cdot(c-di)}{(c+di)\cdot(c-di)} = \frac{ac - adi + bci - bdi^2}{c^2 + d^2} = \frac{(ac+bd) + (bc-ad)i}{c^2 + d^2} =$$

$$= \frac{ac+bd}{c^2+d^2} + \frac{bc-ad}{c^2+d^2}\, i$$

EJEMPLOS:

$$(3+i) + (1- \frac{1}{2}\, i) = 3 + i + 1 - \frac{1}{2}\, i = 4 + \frac{1}{2}\, i$$

$$(1-4i) - (-3+2i) = 1 - 4i + 3 -2i = 4 - 6i$$

$$(2-i) \cdot (1-3i) = 2 - 6i - i + 3i^2 = 2 - 6i - i - 3 = -1 - 7i$$

$$\frac{4+i}{2-3i} = \frac{(4+i)\cdot(2+3i)}{(2-3i).(2+3i)} = \frac{8 + 12i + 2i + 3i^2}{2^2 + 3^2} = \frac{5+14i}{13} = \frac{5}{13} + \frac{14}{13}\, i$$

Conjugado de un número complejo.

Definición. Se llama conjugado de un número complejo $z = a+bi$ al número $\bar{z} = a-bi$.

Por ejemplo, el conjugado de $z = -1-2i$ es $\bar{z} = -1+2i$.

PROPIEDADES.

Se verifican las siguientes propiedades:

1) $\bar{\bar{z}} = z$

2) $\overline{z + z'} = \overline{z} + \overline{z}'$

3) $\overline{z \cdot z'} = \overline{z} \cdot \overline{z}'$

4) $z + \overline{z} = 2\,\mathrm{Re}(z)$; $z - \overline{z} = 2\,\mathrm{Im}(z)i$

5) $\overline{z} = z$ si y sólo si z es un complejo real.

6) $\overline{z} = -z$ si y sólo si z es un imaginario puro.

Demostración:

1) Inmediato.

2) Sean $z = a+bi$, $z' = a'+b'i$

$$\overline{z+z'} = \overline{(a+bi)+(a'+b'i)} = \overline{(a+a')+(b+b')i} = (a+a')-(b+b')i = (a-bi)+(a'-b'i) = \overline{z}+\overline{z}'$$

3) $$\overline{z.z'} = \overline{(a+bi).(a'+b'i)} = \overline{(aa'-bb')+(ab'+ba')i} = (aa'-bb')-(ab'+ba')i =$$

$$= (a-bi).(a'-b'i) = \overline{z}.\overline{z}'$$

4) $z+\overline{z} = (a+bi)+(a-bi) = 2a = 2\,\mathrm{Re}(z)$

$z-\overline{z} = (a+bi)-(a-bi) = 2bi = 2\,\mathrm{Im}(z)i$

5) a) $\overline{z} = z \Longrightarrow z$ real.

Si $\overline{z} = z$, $a-bi = a+bi$. Luego $-b = b$ o sea $2b = 0$, es decir $b = 0$.

b) z real $\Longrightarrow \overline{z} = z$.

Si z es real es de la forma $z = a+0i$. Luego $\overline{z} = a-0i = a+0i = z$.

6) Se demuestra en forma análoga a la propiedad anterior.

OBSERVACION.

1) De la propiedad 2 resulta que:

$$\overline{z-z'} = \overline{z} - \overline{z}'$$

En efecto, $z = z'+(z-z')$. Luego $\overline{z} = \overline{z'+(z-z')} = \overline{z}'+\overline{(z-z')}$. De aquí sigue que

$$\overline{z-z'} = \overline{z}-\overline{z}'$$

2) De la propiedad 3 se deduce que:

$$\overline{\frac{z}{w}} = \frac{\overline{z}}{\overline{w}} \qquad (w \neq 0)$$

Dados z y w , $w \neq 0$, podemos escribir

$z = w \cdot \frac{z}{w}$. Luego $\overline{z} = \overline{w \cdot \frac{z}{w}} = \overline{w} \cdot \overline{\frac{z}{w}}$

De aquí se tiene $\overline{\frac{z}{w}} = \frac{\overline{z}}{\overline{w}}$

Módulo de un número complejo.

Definición. Se llama módulo de un número complejo $z = a+bi$ y se representa $|z|$ al número real positivo o nulo $\sqrt{a^2+b^2}$.

$$|z| = \sqrt{a^2+b^2}$$

El módulo $|z|$ está bien definido pues si $z \neq 0$, a^2+b^2 es un número real positivo y $|z|$ es la raíz cuadrada aritmética de ese número. Si $z = 0$ entonces $|z| = 0$.

Por ejemplo, dado $z = 1-3i$ es $|z| = \sqrt{1^2+(-3)^2} = \sqrt{10}$.

Se ve que

$$|z| = |\overline{z}| = |-z|.$$

pues $\quad a^2+b^2 = a^2+(-b)^2 = (-a)^2+(-b)^2$

Además

$$|z|^2 = z.\overline{z}$$

En efecto, $z.\overline{z} = (a+bi).(a-bi) = a^2+b^2 = |z|^2$

Observemos que si z es un número real, el módulo coincide con el valor absoluto. El módulo de los números complejos es así una generalización del valor absoluto de los números reales y tiene propiedades enteramente análogas.

PROPIEDADES.

1) $|z| \geq 0$; $|z| = 0$ si y sólo si $z = 0$

2) $|z.z'| = |z|.|z'|$

3) $\left|\frac{z}{z'}\right| = \frac{|z|}{|z'|}$ $\qquad (z' \neq 0)$

4) $|z+z'| \leq |z| + |z'|$

5) $||z| - |z'|| \leq |z-z'|$

Demostración:

1) Sigue inmediatamente de la definición de módulo.

2) $|z.z'|^2 = (z.z').(\overline{z.z'}) = z.z'.\overline{z}.\overline{z'} = (z.\overline{z}).(z'.\overline{z'}) = |z|^2.|z'|^2 = (|z|.|z'|)^2$

Como $|z.z'|$ y $|z|.|z'|$ son dos números reales positivos o nulos, de la igualdad de sus cuadrados se deduce que $|z.z'| = |z|.|z'|$ c.q.d.

3) Sean z y z' , $z' \neq 0$. Podemos escribir:

$z = z' . \frac{z}{z'}$. Aplicando la propiedad anterior:

$$|z| = \left|z'. \frac{z}{z'}\right| = |z'|.\left|\frac{z}{z'}\right|$$

Luego $\left|\frac{z}{z'}\right| = \frac{|z|}{|z'|}$ c.q.d.

4) Para probar esta propiedad demostraremos primero el siguiente lema: "Si z y z' son dos números complejos tales que $z+z' = 1$ entonces $|z|+|z'| \geq 1$" .

Supongamos entonces $z+z' = 1$, $z = a+bi$, $z' = a'+b'i$.

De $|z|^2 = a^2+b^2 \geq a^2$ resulta $|z| \geq |a| \geq a$.

$|z'|^2 = a'^2+b'^2 \geq a'^2$ " $|z'| \geq |a'| \geq a'$.

Sumando las desigualdades $|z| \geq a$ y $|z'| \geq a'$ se tiene $|z| + |z'| \geq a + a'$

Pero por hipótesis $z+z' = 1$ o sea:

$(a+a')+(b+b')i = 1$. Luego debe ser $a + a' = 1$

Por lo tanto $|z| + |z'| \geq 1$ c.q.d.

Probemos ahora la propiedad 4.

Si $z+z' = 0$ entonces $|z+z'| = 0$ y como $|z| \geq 0$ y $|z'| \geq 0$ es $0 \leq |z| + |z'|$,

o sea $|z+z'| \leq |z| + |z'|$.

Si $z+z' \neq 0$ podemos escribir $\frac{z}{z+z'} + \frac{z'}{z+z'} = 1$.

Por el lema recién demostrado es entonces:

$$\left| \frac{z}{z+z'} \right| + \left| \frac{z'}{z+z'} \right| \geq 1$$

Aplicando la propiedad 3 se tiene: $\dfrac{|z|}{|z+z'|} + \dfrac{|z'|}{|z+z'|} \geq 1$ de donde resulta:

$$|z+z'| \leq |z| + |z'| \qquad \text{c.q.d.}$$

5) Podemos escribir $z = z'+(z-z')$.

Luego, por 4) se tiene:

$$|z| = |z'+(z-z')| \leq |z'| + |z-z'|$$

De aquí resulta: $|z| - |z'| \leq |z-z'|$ (i)

Análogamente $z' = z+(z'-z)$. Luego

$$|z'| \leq |z| + |z'-z| \quad \text{o sea} \quad -|z'-z| \leq |z| - |z'| \qquad \text{(ii)}$$

Como $|z-z'| = |z'-z|$, de (i) y (ii) resulta:

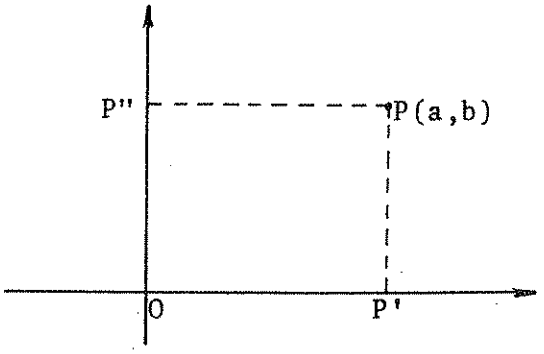
$$-|z-z'| \leq |z| - |z'| \leq |z-z'|$$

Esta es una desigualdad entre números reales; por la propiedad 2 del valor absoluto de los números reales vista en el capítulo II se tiene:

$$||z| - |z'|| \leq |z-z'| \qquad \text{c.q.d.}$$

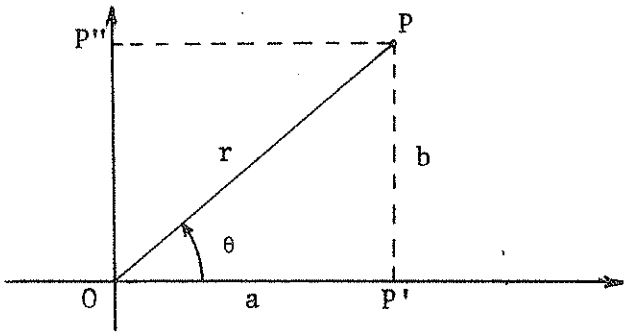
Representación geométrica y notación polar de los números complejos.

Así como los puntos de una recta representan al conjunto R de los números reales, los puntos de un plano representan al conjunto C de los complejos. Como $C = R \times R$, existe una correspondencia biunívoca entre los números complejos y los puntos de un plano: Dado un plano T en el que se ha fijado un sistema de coordenadas cartesianas mediante dos ejes ortogonales , la aplicación $C \longrightarrow T$ que a cada número complejo $z=(a,b)$ hace corresponder el punto P de coordenadas (a,b) es biyectiva. El punto P se dice el afijo de z. Los números complejos reales, es decir de segunda componente



nula, están representados por los puntos del eje de las abscisas, que por eso se llama eje real, y los números imaginarios puros por los puntos del eje de las ordenadas, que se llama eje imaginario. El plano así considerado como imagen de C se llama plano complejo.

Si en lugar de considerar coordenadas cartesianas se consideran coordenas polares entonces un punto P del plano queda determinado por la longitud del segmento $\overline{OP}$ o radio vector y el ángulo θ.



Dado un número complejo $z = (a,b)$, $z \neq (0,0)$, si P es el afijo de z, se llama argumento de z al ángulo θ formado por el semieje real positivo y la semirrecta OP, medido en radianes y a menos de un múltiplo entero de 2π , tomándose como sentido positivo de giro en el plano el que lleva al semieje real positivo sobre el semieje imaginario positivo recorriendo un ángulo de $\pi/2$.

$$\text{argumento de } z = \theta + 2k\pi \quad , \text{ con } k \in Z.$$

Los valores que toma entonces "argumento de z" son:

$$\ldots\ldots\ldots, \theta-3.2\pi \ , \ \theta-2.2\pi \ , \ \theta-2\pi \ , \ \theta \ , \ \theta+2\pi \ , \ \theta+2.2\pi \ , \ \theta+3.2\pi \ , \ldots\ldots\ldots$$

(Se trata de números reales puesto que los ángulos van medidos en radianes).

Se llama <u>valor principal del argumento</u> al comprendido entre 0 y 2π : $0 \leq \theta < 2\pi$.

Las coordenadas cartesianas y polares de un punto P están ligadas por las siguientes relaciones, que se deducen del triángulo rectángulo $\overset{\Delta}{OPP'}$:

$$r = \sqrt{a^2+b^2} = |z| \qquad (1)$$

$$\text{tg } \theta = \frac{b}{a} \qquad (2)$$

$$a = r \cos \theta \qquad (3)$$

$$b = r \text{ sen } \theta \qquad (4)$$

Si en la expresión binómica de un número complejo z, $z = a+bi$, se reemplazan a y b por los valores (3) y (4) se tiene

$$z = a+bi = r \cos \theta + i \, r \text{ sen } \theta$$

o sea

$$z = r(\cos \theta + i \text{ sen } \theta)$$

Observemos que en esta expresión puede figurar cualquier valor del argumento de z y que $r = |z|$. Luego

$$z = |z| \left[\cos(\theta+2k\pi)+i \operatorname{sen}(\theta+2k\pi)\right] \quad , \quad k \in Z \qquad (5)$$

Esta forma de escribir un número complejo se llama polar o trigonométrica. Por razones de comodidad se trabaja con el valor principal del argumento y se escribe abreviadamente:

$$z = |z|_{\theta}$$

Para cada número complejo $z = a+bi$ la expresión (5) está unívocamente determinada puesto que si $z = u(\cos v + i \operatorname{sen} v)$, con $u,v \in R$, $u \geq 0$, entonces se tiene $u = |z|$ y $v = \arg z$. En efecto, es $a = u \cos v$ y $b = u \operatorname{sen} v$, lo que implica

$$|z| = \sqrt{a^2+b^2} = u$$

Entonces de (1) , (3) , (4) y $u = |z|$ resulta

$$r \cos \theta = r \cos v \quad y \quad r \operatorname{sen} \theta = r \operatorname{sen} v \text{ , de donde sigue}$$

$$v = \theta + 2k\pi = \arg z$$

Luego:

$$z = z' \iff |z| = |z'| \quad y \quad \arg z = \arg z' \ .$$

Dado entonces un número complejo $z = a+bi$, para escribirlo en forma polar se aplican las fórmulas (1) y (2) que dan el módulo y el argumento en función de a y b. Observemos que como, en general, para cualquier ángulo α es

$$\operatorname{tg} \alpha = \operatorname{tg}(\alpha+\pi)$$

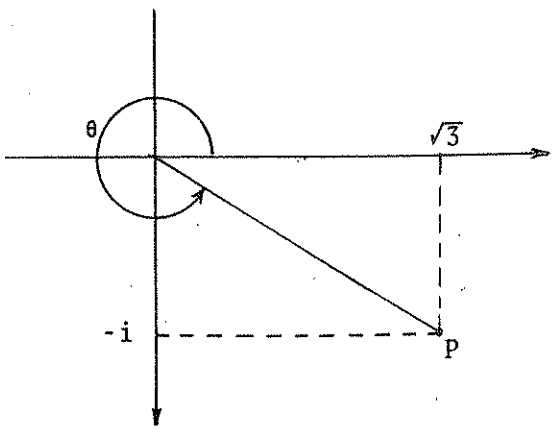
la fórmula (2) da el valor del argumento a menos de π. Para saber a qué cuadrante pertenece es necesario analizar el signo de $\operatorname{sen} \theta$ y $\cos \theta$. Como $|z| \geq 0$, de (3) y (4) resulta que el signo de a es el del $\cos \theta$ y el de b es el del $\operatorname{sen} \theta$.

Por ejemplo, escribamos en forma trigonométrica el número $z = \sqrt{3} - i$.

Hay que calcular el módulo $|z|$ y el argumento θ.

$$|z| = \sqrt{a^2+b^2} = \sqrt{(\sqrt{3})^2+(-1)^2} = 2$$

$$\operatorname{tg} \theta = \frac{b}{a} = \frac{-1}{\sqrt{3}} \text{ , luego } \theta = \operatorname{arc\,tg} \frac{-1}{\sqrt{3}}$$



Se obtienen dos valores posibles para el argumento, uno en el segundo cuadrante: $5/6\ \pi$, y otro en el cuarto cuadrante: $11/6\ \pi$. Pero como $b < 0$ y $a > 0$, el seno del argumento buscado es negativo y su coseno es positivo. Luego está en el cuarto cuadrante, es decir:

$$\theta = \frac{11}{6}\pi$$

Luego

$$z = |z|\ (\cos\theta + i \operatorname{sen}\theta) = 2\left(\cos\frac{11}{6}\pi + i \operatorname{sen}\frac{11}{6}\pi\right)$$

Abreviadamente

$$z = 2_{\frac{11}{6}\pi}$$

Análogamente

$$1 = \cos 0 + i \operatorname{sen} 0$$

$$-1 = \cos\pi + i \operatorname{sen}\pi$$

$$1+i = \sqrt{2}\left(\cos\frac{\pi}{4} + i \operatorname{sen}\frac{\pi}{4}\right)$$

$$-1-i = \sqrt{2}\left(\cos\frac{5\pi}{4} + i \operatorname{sen}\frac{5\pi}{4}\right)$$

$$3i = 3\left(\cos\frac{\pi}{2} + i \operatorname{sen}\frac{\pi}{2}\right)$$

$$-\frac{1}{2}i = \frac{1}{2}\left(\cos\frac{3\pi}{2} + i \operatorname{sen}\frac{3\pi}{2}\right)$$

Recíprocamente, dado un número complejo en forma polar, se pasa a la forma binómica aplicando las fórmulas (3) y (4). Por ejemplo, sea

$$z = 4_{\frac{2}{3}\pi}$$

Entonces

$$z = 4_{\frac{2}{3}\pi} = 4\left(\cos\frac{2}{3}\pi + i \operatorname{sen}\frac{2}{3}\pi\right) = 4\left(-\frac{1}{2} + i\frac{\sqrt{3}}{2}\right) = -2 + 2\sqrt{3}\,i$$

que es la forma binómica del número dado.

De la misma manera, dados $z_1 = 2_{\pi}$, $z_2 = \sqrt{3}_{\frac{7}{6}\pi}$, $z_3 = \sqrt{2}_{\frac{3}{4}\pi}$ en forma binómica se escriben:

$$z_1 = 2(\cos\pi + i \operatorname{sen}\pi) = -2$$

$$z_2 = \sqrt{3}(\cos \frac{7}{6}\pi + i \operatorname{sen} \frac{7}{6}\pi) = -\frac{3}{2} - \frac{\sqrt{3}}{2} i$$

$$z_3 = \sqrt{2}(\cos \frac{3}{4}\pi + i \operatorname{sen} \frac{3}{4}\pi) = -1 + i$$

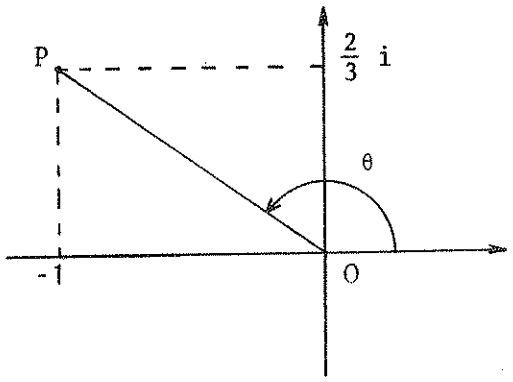
Pero los cálculos involucrados en el pasaje de una a otra forma no son siempre tan sencillos como en los ejemplos anteriores. Observemos que desde el punto de vista práctico, para pasar de la forma binómica a la trigonométrica la fórmula $\theta = \text{arc tg } \frac{b}{a}$ sólo permite calcular en la generalidad de los casos un valor aproximado del argumento θ, por cuanto dado el valor de la tangente de un ángulo no se conoce el valor exacto del ángulo, salvo pocas excepciones. Análogamente, cuando se trata de pasar de la forma polar a la binómica, generalmente no se puede calcular el valor exacto de cos θ y sen θ y en consecuencia tampoco se conocen con exactitud los valores de a y b.

EJEMPLOS.

1. Expresar en forma polar el número $z = -1 + \frac{2}{3} i$

$$|z| = \sqrt{a^2+b^2} = \sqrt{(-1)^2 + (\frac{2}{3})^2} = \frac{\sqrt{13}}{3}$$

$$\text{tg } \theta = \frac{b}{a} = \frac{2/3}{-1} = -0.666.....$$



Como $b > 0$ y $a < 0$, θ es un ángulo que tiene seno positivo y coseno negativo, es decir pertenece al segundo cuadrante. Buscando en la tabla de funciones trigonométricas, 0.666.... corresponde a la tangente de un ángulo en el primer cuadrante de aproximadamente 33° 41' 23" , o sea a un ángulo de 146° 18' 37" en el segundo cuadrante. Expresado este ángulo en radianes da un valor aproximado para el argumento θ de 2.55359 radianes.

Luego el número z dado se escribe en forma trigonométrica como sigue:

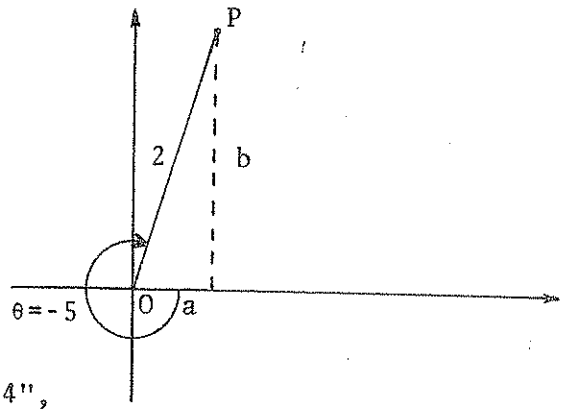
$$z = \frac{\sqrt{13}}{3} (\cos 2.55359 + i \operatorname{sen} 2.55359)$$

Abreviadamente

$$z = \frac{\sqrt{13}}{3}$$
$$2.55359$$

2. Escribir en forma binómica el número

$$z = 2_{-5} = 2\left[\cos(-5) + i \operatorname{sen}(-5)\right]$$



Para buscar en la tabla de funciones trigonométricas es necesario expresar el argumento de -5 radianes en el sistema sexagesimal. Haciendo la conversión se encuentra que corresponde aproximadamente a un ángulo de -286° 28' 44", o sea de 73° 31' 16". Buscando en la tabla se hallan para el seno y el coseno de este ángulo los valores aproximados de 0.95892 y 0.28367 respectivamente. Luego la forma binómica que se obtiene para el número dado es

$$z = 0.28367 + 0.95892\ i$$

Notemos que:

1) El argumento del número complejo cero no está definido. Pero cero queda determinado por la igualdad $|z| = 0$.

2) Si z es un número real positivo su argumento es 0 y recíprocamente, si un número complejo tiene argumento 0 es real positivo. Es decir, un número complejo es real positivo si y sólo si su argumento es 0. Análogamente, un número complejo es real negativo si y solo si su argumento es π .

3) Un número complejo es imaginario puro si y sólo si su argumento es $\frac{\pi}{2}$ ó $\frac{3}{2}\pi$.

Producto y cociente de números complejos en forma polar.

Sean $z = |z|_{\theta}$, $z' = |z'|_{\theta'}$. Multiplicando se tiene:

$$z.z' = \left[|z|\ (\cos\theta + i \operatorname{sen}\theta)\right] . \left[|z'|\ (\cos\theta' + i \operatorname{sen}\theta')\right] =$$

$$= |z|.|z'|\ (\cos\theta\ .\ \cos\theta' + i\cos\theta.\operatorname{sen}\theta' + i \operatorname{sen}\theta.\cos\theta' - \operatorname{sen}\theta.\operatorname{sen}\theta') =$$

$$= |z|.|z'|\ \left[(\cos\theta\ .\ \cos\theta' - \operatorname{sen}\theta.\operatorname{sen}\theta') + i(\operatorname{sen}\theta.\cos\theta' + \cos\theta.\operatorname{sen}\theta')\right] =$$

$$= |z|.|z'|\ \left[\cos(\theta+\theta') + i \operatorname{sen}(\theta+\theta')\right] =$$

$$= (|z|.|z'|)_{\theta+\theta'}$$

Luego

$$|z.z'| = |z|.|z'| \qquad y \qquad \arg(z.z') = \arg z + \arg z'$$

Es decir, el producto de dos números complejos tiene por módulo el producto de los módulos de los factores (propiedad ya demostrada) y por argumento la suma de sus argumentos.

Veamos ahora el cociente. Sean $z, z' \in C$, $z' \neq 0$.

$$\frac{z}{z'} . z' = z$$

Por lo recién demostrado se tiene:

$$\left| \frac{z}{z'} \right| . |z'| = |z| \qquad y \qquad \arg \frac{z}{z'} + \arg z' = \arg z$$

y de aquí

$$\left| \frac{z}{z'} \right| = \frac{|z|}{|z'|} \qquad y \qquad \arg \frac{z}{z'} = \arg z - \arg z'$$

Por lo tanto, el cociente de dos números complejos tiene módulo igual al cociente de los módulos (ya demostrado) y argumento igual a la diferencia de los argumentos.

En particular, si $z = |z|_\theta$ es distinto de cero, su inverso es $z^{-1} = (|z|^{-1})_{-\theta}$ pues :

$$z^{-1} = 1:z = 1_0 : |z|_\theta = (1:|z|)_{0-\theta} = (|z|^{-1})_{-\theta}$$

EJEMPLO. Hallar el producto y cociente de los números $z = 2_{\frac{2}{3}\pi}$ y $z' = 8_{\frac{7}{6}\pi}$ y expresar los resultados en forma binómica.

$$z.z' = (2.8)_{\frac{2}{3}\pi + \frac{7}{6}\pi} = 16_{\frac{11}{6}\pi} = 16(\cos \tfrac{11}{6}\pi + i \operatorname{sen} \tfrac{11}{6}\pi) =$$

$$= 16(\cos 330° + i \operatorname{sen} 330°) = 16(\tfrac{\sqrt{3}}{2} - \tfrac{1}{2} i) = 8\sqrt{3} - 8i$$

$$z:z' = (2:8)_{\frac{2}{3}\pi - \frac{7}{6}\pi} = \frac{1}{4}_{(-\frac{1}{2}\pi)} = \frac{1}{4}_{\frac{3}{2}\pi} = \frac{1}{4}(\cos \tfrac{3}{2}\pi + i \operatorname{sen} \tfrac{3}{2}\pi) =$$

$$= \frac{1}{4}(\cos 270° + i \operatorname{sen} 270°) = -\frac{1}{4} i .$$

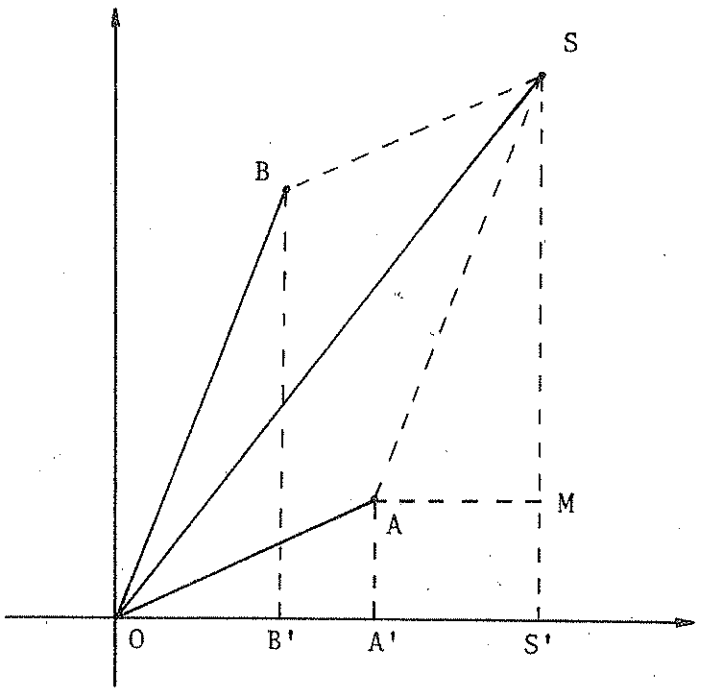
OBSERVACION. La definición de argumento de un número complejo que hemos dado se apoya en conceptos geométricos, pero la teoría de la forma polar de los números complejos puede desarrollarse independientemente de toda interpretación geométrica.

Representación geométrica de las operaciones con números complejos.

Sean $z = a+bi$, $z' = a'+b'i$ y A y B sus respectivos afijos. Como

$$z+z' = (a+a') + (b+b')i$$

el afijo S de la suma tiene abscisa igual a la suma de las abscisas de A y B y ordenada igual a la suma de las ordenadas de A y B. Entonces, en la figura, OB' = A'S' y B'B = MS. De la igualdad de los triángulos OB'B y AMS resulta que el segmento OB es igual y paralelo al AS. Luego OASB es un paralelogramo y el afijo de la suma se obtiene por la regla del paralelogramo.



La representación de la diferencia se obtiene por la misma regla, teniendo en cuenta que $z-z' = z+(-z')$ y que el afijo de $-z'$ es el punto simétrico del afijo de z' con respecto al origen de coordenadas.

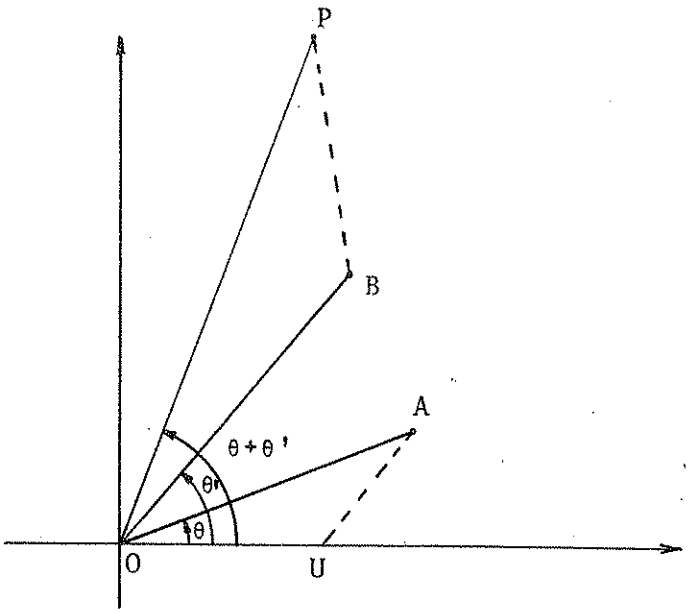
Para representar el producto $z.z'$, sabemos que el módulo es el producto de los módulos y el argumento la suma de los argumentos.

Si P es el afijo del producto

$$OP = OA.OB \quad , \quad \text{luego}$$

$$\frac{OP}{OB} = \frac{OA}{OU} \quad , \quad \text{donde OU es}$$

el segmento unidad. Entonces los triángulos OUA y OBP son semejantes, lo que permite obtener geométricamente el afijo P del producto construyendo sobre OB, considerado como lado homólogo del OU , un triángulo semejante al OUA.



Para representar el cociente se procede análogamente teniendo en cuenta que

$$z:z' = z.(z')^{-1} \quad \text{y que} \quad z'^{-1} = (|z'|^{-1})_{-\theta}$$

Potenciación de exponente entero de números complejos.

Se define la potenciación de números complejos de exponente natural por recurrencia, de la siguiente manera:

$$z^1 = z$$

$$z^{n+1} = z^n . z$$

La definición se extiende a exponente entero cualquiera como sigue:

$$z^0 = 1$$

$$z^n = (z^{-1})^{-n} \quad , \quad \text{siendo } z \neq 0 \ , \ n \in Z \ , \ n < 0 \ .$$

Se verifican propiedades análogas a las de la potenciación de números reales.

Fórmula de DE MOIVRE.

Consideremos el problema de calcular, por ejemplo, $(-1-i)^{20}$. En forma binómica el cálculo se hace interminable. La fórmula

$$\left[|z| \ (\cos \theta + i \operatorname{sen} \theta)\right]^n = |z|^n(\cos n\theta + i \operatorname{sen} n\theta) \qquad (6)$$

o abreviadamente

$$(|z|_\theta)^n = (|z|^n)_{n\theta}$$

llamada fórmula de De Moivre (1667-1754); permite calcular las potencias de exponente entero de los números complejos en forma polar. La demostraremos primero para exponente natural por inducción y luego para exponente nulo y entero negativo.

Si n = 1 la igualdad (6) se verifica trivialmente.

Supongamos que es verdadera para n y probémosla para n+1; es decir suponemos que

$(|z|_\theta)^n = (|z|^n)_{n\theta}$ y queremos probar que $(|z|_\theta)^{n+1} = (|z|^{n+1})_{(n+1)\theta}$

$$(|z|_\theta)^{n+1} = (|z|_\theta)^n . (|z|_\theta) = (|z|^n)_{n\theta} . (|z|)_\theta = (|z|^n . |z|)_{n\theta+\theta} = (|z|^{n+1})_{(n+1)\theta}$$

Queda demostrada así por inducción la fórmula (6) para exponente natural.

Si $n = 0$, también se verifica.

Resta probar entonces que también es verdadera siendo $z \neq 0$, $n \in Z$, $n < 0$. Aplicando la definición de potencia de exponente entero negativo, la relación $(|z|_\theta)^{-1} = (|z|^{-1})_{-\theta}$ ya demostrada , que $-n \in N$ y la validez de la fórmula de De Moivre para exponente natural, se tiene:

$$(|z|_\theta)^n = \left[(|z|_\theta)^{-1}\right]^{-n} = ((|z|^{-1})_{-\theta})^{-n} = ((|z|^{-1})^{-n})_{(-n)(-\theta)} = (|z|^n)_{n\theta}$$

EJERCICIO: Dado $z = -1-i$ calculemos z^{20}.

$$|z| = \sqrt{(-1)^2+(-1)^2} = \sqrt{2}$$

$$\theta = \text{arc tg } \frac{-1}{-1} = 225° = \frac{5}{4}\pi$$

$$z^{20} = \left[\sqrt{2}_{\frac{5}{4}\pi}\right]^{20} = (\sqrt{2})^{20}_{20.\frac{5}{4}\pi} = 2^{10}_{25\pi} = 2^{10}_{\pi} = 2^{10}(\cos\pi + i \operatorname{sen}\pi) = -1024.$$

Luego

$$(-1-i)^{20} = -1024$$

Radicación de números complejos.

Definición. Dado un número complejo z y un número natural n se llama raíz n-ésima de z a todo número complejo w tal que $w^n = z$.

Se trata de saber si dado un número complejo cualquiera $z = |z|_\theta$ existe alguna raíz n-ésima de z.

Si un número $w = |w|_\alpha$ es una raíz n-ésima de z, es decir si $w^n = z$, entonces

$$(|w|^n)_{n\alpha} = |z|_\theta$$

Luego

$$|w|^n = |z| \qquad y \qquad n\alpha = \theta + 2k\pi \qquad , \qquad \text{con } k \in Z$$

De aquí resulta

$$|w| = \sqrt[n]{|z|} \qquad (7) \qquad \text{y} \qquad \alpha = \frac{\theta+2k\pi}{n} \quad , \quad \text{con } k \in Z \qquad (8)$$

El módulo de la raíz buscada está bien determinado por (7) : es la raíz n-ésima aritmética del módulo de z. En cuanto al argumento, aparentemente existen infinitos puesto que en la expresión (8) que da α aparece la variable k que puede tomar cualquier valor entero.

Dando a k los valores 0,1,2,.....,n-1 se obtienen los siguientes argumentos:

$$\frac{\theta}{n} \quad , \quad \frac{\theta+2\pi}{n} \quad , \quad \frac{\theta+4\pi}{n} \quad , \quad \ldots\ldots \quad , \quad \frac{\theta+2(n-1)\pi}{n}$$

y en consecuencia los n números complejos w_1 , w_2 ,...., w_n que tienen módulo igual a $\sqrt[n]{|z|}$ y argumento igual a cada uno de los ángulos que acabamos de escribir son n raíces de z. Son todas distintas entre sí. En efecto, la diferencia de dos argumentos cualesquiera es:

$$\frac{\theta+2k\pi}{n} - \frac{\theta+2k'\pi}{n} = \frac{k-k'}{n}\ 2\pi$$

Como k y k' son enteros diferentes y están comprendidos entre 0 y n-1 su diferencia k-k' es no nula y menor en valor absoluto que n. Por lo tanto la diferencia de dos argumentos cualesquiera de los indicados no es un múltiplo entero de 2π , lo que prueba que w_1 , w_2 ,...., w_n son diferentes entre sí.

Veamos ahora que dándole a k cualquier valor entero en la fórmula (8) se obtiene un argumento que difiere de alguno de los recién indicados en un múltiplo entero de 2π y en consecuencia la raíz w que corresponde a ese argumento coincide con una de las n ya encontradas. Dado $k \in Z$, sean q y r el cociente y el resto de dividir k por n:

$$k = qn + r \quad , \quad 0 \leq r < n$$

Entonces

$$\frac{\theta+2k\pi}{n} = \frac{\theta+2(qn+r)\pi}{n} = 2q\pi + \frac{\theta+2r\pi}{n}$$

Luego el argumento $\frac{\theta+2k\pi}{n}$ difiere en un múltiplo entero de 2π del argumento $\frac{\theta+2r\pi}{n}$, donde r es uno de los números 0,1,2,...,n-1. Es decir, todos los argumentos que se obtienen dando a k valores enteros arbitrarios en la fórmula (8) difieren en un número entero de circunferencias de alguno de los n argumentos que se obtienen haciendo

$k = 0,1,2,.....,n-1$. Por lo tanto existen sólo n raíces n-ésimas diferentes de z.

Queda probado así el siguiente

TEOREMA 3.2. Todo número complejo $z \neq 0$ tiene n raíces n-ésimas distintas. Sus módulos son la raíz n-ésima aritmética del módulo de z y sus argumentos principales son respectivamente:

$$\frac{\theta}{n} \quad , \quad \frac{\theta+2\pi}{n} \quad , \quad \frac{\theta+4\pi}{n} \quad , \ldots\ldots , \quad \frac{\theta+2(n-1)\pi}{n}$$

donde θ = argumento principal de z.

Usaremos la notación $\sqrt[n]{((z))}$ para representar una cualquiera de las n raíces n-ésimas de z. La fórmula

$$\sqrt[n]{((z))} = \sqrt[n]{|z|}\ (\cos \frac{\theta+2k\pi}{n} + i \operatorname{sen} \frac{\theta+2k\pi}{n}) \qquad (9)$$

da entonces las n raíces de z haciendo $k = 0,1,2,....,n-1$.

Representación geométrica de las raíces.

Como las n raíces n-ésimas de un complejo $z = |z|_\theta \neq 0$ tienen todas módulo igual a $\sqrt[n]{|z|}$, sus afijos pertenecen a la circunferencia de centro O y radio $\sqrt[n]{|z|}$. El punto A_1 sobre esa circunferencia de argumento $\frac{\theta}{n}$ es el afijo de la raíz

$$w_1 = \sqrt[n]{|z|}\ (\cos \frac{\theta}{n} + i \operatorname{sen} \frac{\theta}{n})$$

Como los argumentos de las demás raíces se obtienen sumándole a $\frac{\theta}{n}$ los valores:

$$\frac{2\pi}{n} \quad , \quad 2\,\frac{2\pi}{n} \quad , \ldots\ldots , \quad (n-1)\,\frac{2\pi}{n}$$

resulta que los afijos A_1 , A_2 , ..., A_n de las n raíces son los vértices de un polígono regular de n lados inscripto en la circunferencia de centro O y radio $\sqrt[n]{|z|}$.

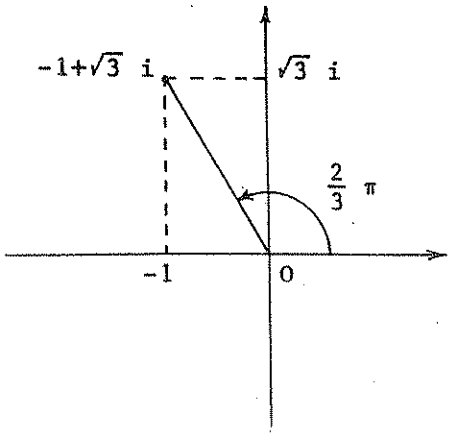
EJEMPLOS.

1. Dado $z = -1 + \sqrt{3}\,i$ hallar sus raíces de cuarto orden y expresarlas en forma binómica.

   Debemos aplicar la fórmula (9) para n = 4.

$|z| = \sqrt{(-1)^2+(\sqrt{3})^2} = 2$

$\theta = \text{arc tg } \frac{\sqrt{3}}{-1} = \frac{2}{3}\pi$



Luego se tiene:

$$\sqrt[4]{((z))} = \sqrt[4]{2}\ (\cos \frac{\frac{2}{3}\pi+2k\pi}{4} + i \text{ sen } \frac{\frac{2}{3}\pi+2k\pi}{4})$$

donde $k = 0,1,2,3$.

Las cuatro raíces son:

$$w_1 = \sqrt[4]{2}\ (\cos \frac{120°}{4} + i \text{ sen } \frac{120°}{4}) = \sqrt[4]{2}\ (\cos 30° + i \text{ sen } 30°) =$$

$$= \sqrt[4]{2}\ (\frac{\sqrt{3}}{2} + i\,\frac{1}{2}) = \frac{\sqrt[4]{18}}{2} + \frac{\sqrt[4]{2}}{2}\,i$$

$$w_2 = \sqrt[4]{2}\ (\cos \frac{120°+2\pi}{4} + i \text{ sen } \frac{120°+2\pi}{4}) = \sqrt[4]{2}\ (\cos 120° + i \text{ sen } 120°) =$$

$$= \sqrt[4]{2}\ (-\frac{1}{2} + \frac{\sqrt{3}}{2}\,i) = -\frac{\sqrt[4]{2}}{2} + \frac{\sqrt[4]{18}}{2}\,i$$

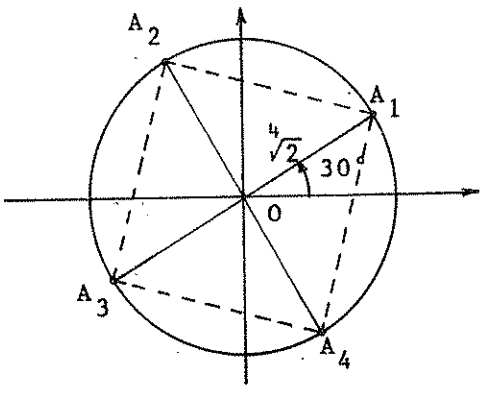
$$w_3 = \sqrt[4]{2}\ (\cos \frac{120°+4\pi}{4} + i \text{ sen } \frac{120°+4\pi}{4}) = \sqrt[4]{2}\ (\cos 210° + i \text{ sen } 210°) =$$

$$= \sqrt[4]{2}\ (-\frac{\sqrt{3}}{2} - \frac{1}{2}\,i) = -\frac{\sqrt[4]{18}}{2} - \frac{\sqrt[4]{2}}{2}\,i$$

$$w_4 = \sqrt[4]{2}\ (\cos \frac{120°+6\pi}{4} + i \text{ sen } \frac{120°+6\pi}{4}) = \sqrt[4]{2}\ (\cos 300° + i \text{ sen } 300°) =$$

$$= \sqrt[4]{2}\ (\frac{1}{2} - \frac{\sqrt{3}}{2}\,i) = \frac{\sqrt[4]{2}}{2} - \frac{\sqrt[4]{18}}{2}\,i$$

Los afijos de w_1 , w_2 , w_3 y w_4 son los vértices de un cuadrado inscripto en la circunferencia de centro 0 y radio $\sqrt[4]{2}$.

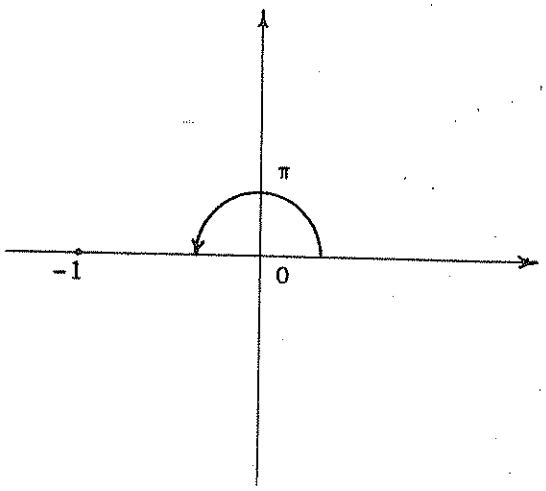


Hallar las raíces quintas del número z = -1.

Hay que aplicar la fórmula (9) para n = 5.

$|z| = \sqrt{(-1)^2+0^2} = 1$

$\theta = \arg z = \pi$



Luego la fórmula

$$\sqrt[5]{((-1))} = \sqrt[5]{1}\ (\cos \frac{\pi+2k\pi}{5} + i \operatorname{sen} \frac{\pi+2k\pi}{5}) =$$

$$= \cos \frac{\pi+2k\pi}{5} + i \operatorname{sen} \frac{\pi+2k\pi}{5} \quad \text{para}$$

k = 0,1,2,3,4 da las cinco raíces quintas de -1.

$w_1 = \cos \frac{\pi}{5} + i \operatorname{sen} \frac{\pi}{5} = \cos 36° + i \operatorname{sen} 36° = 0.80902 + i\ 0.58779$

$w_2 = \cos \frac{3\pi}{5} + i \operatorname{sen} \frac{3\pi}{5} = \cos 108° + i \operatorname{sen} 108° = -\cos 72° + i \operatorname{sen} 72° =$

$= -0.30902 + i\ 0.95106$

$w_3 = \cos \frac{5\pi}{5} + i \operatorname{sen} \frac{5\pi}{5} = \cos \pi + i \operatorname{sen} \pi = -1$

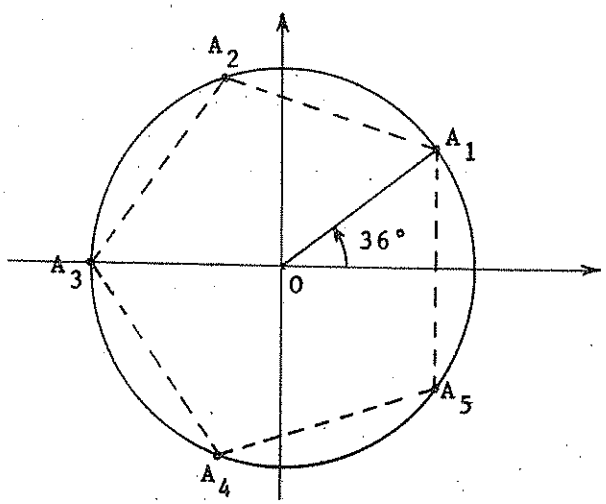
$w_4 = \cos \frac{7\pi}{5} + i \operatorname{sen} \frac{7\pi}{5} = \cos 252° + i \operatorname{sen} 252° = -\cos 72° - i \operatorname{sen} 72° =$

$= -0.30902 - i\ 0.95106$

$w_5 = \cos \frac{9\pi}{5} + i \operatorname{sen} \frac{9\pi}{5} = \cos 324° + i \operatorname{sen} 324° = \cos 36° - i \operatorname{sen} 36° =$

$= 0.80902 - i\ 0.58779$

Los afijos de w_1 , w_2 , w_3 , w_4 y w_5 son los vértices de un pentágono regular inscripto en la circunferencia de centro 0 y radio 1.



Raíces de la unidad.

Si consideramos el número 1, como su módulo es 1 y su argumento 0, las raíces n-ésimas de 1 están dadas por la fórmula:

$$\sqrt[n]{((1))} = \cos \frac{2k\pi}{n} + i \operatorname{sen} \frac{2k\pi}{n} \qquad , \quad k = 0,1,2,\ldots,n-1$$

Una de las raíces es 1, que corresponde al valor k = 0. Notemos que si n es par , n = 2t , hay otra raíz real , -1 , que corresponde al valor k = t. Si n es impar la única raíz real es 1.

En el plano complejo, las n raíces n-ésimas de la unidad tienen por afijos los vértices de un polígono regular de n lados inscripto en la circunferencia de centro 0 y radio 1. Como el número 1 es uno de los vértices, se deduce que las raíces que no son reales están situadas simétricamente con respecto al eje real, es decir son números conjugados y por lo tanto se presentan de a pares.

Las raíces n-ésimas de la unidad son particularmente importantes porque se verifica que:

Todas las raíces n-ésimas de un número complejo z se obtienen multiplicando una cualquiera de ellas por cada una de las n raíces n-ésimas de la unidad.

En efecto, sea r una raíz n-ésima de z y u_1 , u_2 ,...., u_n las n raíces n-ésimas de 1. (Suponemos $z \neq 0$). En primer lugar, si u es una raíz n-ésima de 1 se tiene

$$(ru)^n = r^n.u^n = z.1 = z$$

De modo que los n números ru_1 , ru_2 ,....., ru_n son todos raíces n-ésimas de z. Pero además son todos distintos entre sí pues $ru_i = ru_j$ implica $u_i = u_j$; luego son las n raíces n-ésimas de z.

Por lo tanto, conocida una raíz n-ésima de un número z, el cálculo de las n raíces n-ésimas de z se reduce al cálculo de las raíces n-ésimas de la unidad. En particular, si z es un número real positivo, como $|z| = z$ se tiene:

$$\sqrt[n]{((z))} = \sqrt[n]{z} \odot \sqrt[n]{((1))}$$

donde $\sqrt[n]{z}$ es la raíz n-ésima aritmética de z.

Por ejemplo, las raíces cuartas de 1 son: 1 , -1 , i , -i . Entonces las raíces cuartas de 16 son :

$$\sqrt[4]{((16))} = \sqrt[4]{16} \ , \ \sqrt[4]{((1))} = 2.\sqrt[4]{((1))} = \begin{cases} 2 \\ -2 \\ 2i \\ -2i \end{cases}$$

Las raíces n-ésimas de la unidad tienen las siguientes propiedades:

1) El producto de dos raíces n-ésimas de la unidad es una raíz n-ésima de la unidad

2) Si u es una raíz n-ésima de la unidad su inverso u^{-1} también lo es.

La demostración es muy simple y queda a cargo del lector. Entonces, si con G_n designamos al conjunto de las n raíces n-ésimas de 1, se puede decir que:

a) G_n es cerrado con respecto a la multiplicación, es decir la multiplicación es una operación en G_n y es asociativa.

b) En G_n hay neutro para la multiplicación pues $1 \in G_n$.

c) Todo número $u \in G_n$ tiene inverso $u^{-1} \in G_n$.

Estas tres propiedades se resumen diciendo que G_n es un grupo con respecto a la multiplicación de números complejos.

Por ejemplo,

$$G_1 = \{1\} \ ; \ G_2 = \{1,-1\} \ ; \ G_3 = \{1 \ , -\frac{1}{2} + \frac{\sqrt{3}}{2} i \ , \ -\frac{1}{2} - \frac{\sqrt{3}}{2} i\} \ ; \ G_4 = \{1,-1,i,-i\}$$

Raíces primitivas de la unidad.

Las n raíces n-ésimas de la unidad pueden clasificarse en dos clases: las que no son raíces de 1 de un orden inferior a n, que se llaman raíces primitivas de orden n , y las que no tienen esta propiedad, es decir las que aparecen como raíces de 1 de un orden menor que n.

En otras palabras:

Definición. Una raíz n-ésima ε de 1 se dice una raíz primitiva de orden n si el menor exponente natural k tal que $\varepsilon^k = 1$ es $k = n$.

Por ejemplo, considerando las raíces de la unidad de cuarto orden, 1,-1,i,-i , las raíces primitivas de este orden son i y -i. Las otras dos no lo son: 1 es raíz primitiva de orden 1 y -1 es raíz primitiva de orden 2.

A partir de una raíz primitiva de orden n de 1 pueden calcularse las demás raíces de ese orden de acuerdo con el siguiente

TEOREMA 3.3. Si ε es una raíz n-ésima de 1, ε es una raíz primitiva de orden n si y sólo si ε^0 , ε , ε^2 , ε^3 ,...., ε^{n-1} son las n raíces n-ésimas de 1.

Demostración: Supongamos que ε es una raíz primitiva de orden n y probemos que ε^0 , ε , ε^2 , ε^3 ,....., ε^{n-1} son las n raíces n-ésimas de 1. Como $\varepsilon^n = 1$, ε^k es raíz n-ésima de 1 cualquiera sea el exponente entero k pues

$$(\varepsilon^k)^n = (\varepsilon^n)^k = 1^k = 1$$

En particular, 1 , ε , ε^2 ,...., ε^{n-1} son raíces n-ésimas de la unidad.

Además estos n números son distintos entre sí. En efecto, si fuera $\varepsilon^a = \varepsilon^b$ con $0 \leq a \leq n-1$, $0 \leq b \leq n-1$ y $a \neq b$, suponiendo $a < b$ sería

$$\varepsilon^b - \varepsilon^a = 0$$

$$\varepsilon^a(\varepsilon^{b-a} - 1) = 0$$

Como $\varepsilon^a \neq 0$, $\varepsilon^{b-a} - 1 = 0$ o sea $\varepsilon^{b-a} = 1$. Pero $0 < b-a < n$ y esto contradice la hipótesis de que ε es raíz primitiva de orden n. Por lo tanto $1,\varepsilon,\varepsilon^2,\varepsilon^3,.....,\varepsilon^{n-1}$ son n raíces n-ésimas distintas de 1, es decir, todas las raíces n-ésimas de 1.

Recíprocamente, si ε es un número complejo tal que ε^0 , ε , ε^2 , ε^3 ,....., ε^{n-1} son las n raíces n-ésimas de 1 entonces ε es una raíz primitiva de orden n. En efecto , $\varepsilon^n = 1$ y como los números $\varepsilon^0 = 1$, ε , ε^2 , ε^3 ,....., ε^{n-1} son distintos entre sí , el menor exponente natural k tal que $\varepsilon^k = 1$ es $k = n$.

Por ejemplo, como i es una raíz primitiva de cuarto orden, las raíces cuartas de 1 son: i^0 , i , i^2 , i^3 o sea 1 , i , -1 , -i .

Veamos ahora cómo calcular las raíces primitivas.

TEOREMA 3.4. Las raíces primitivas de la unidad de orden n se obtienen dándole a k los valores primos con n y menores que n en la fórmula

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

Demostración: Sea u una raíz n-ésima de 1.

Luego

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

para un cierto valor $0 \leq k \leq n-1$. Entonces, por la fórmula de De Moivre se tiene para todo número natural h:

$$u^h = \cos \frac{2kh\pi}{n} + i \operatorname{sen} \frac{2kh\pi}{n}$$

Para que u sea una raíz de orden h de 1, es decir para que $u^h = 1$ es necesario y suficiente que el argumento $\frac{2kh\pi}{n}$ sea múltiplo de 2π, o sea que kh sea divisible por n

Si k es primo con n debe ser h divisible por n y por lo tanto el menor valor posible de h es n, es decir u es una raíz primitiva de orden n. En cambio, si k no es primo con n, simplificando en la expresión de u y llamando k' y n' a los cocientes de k y n por su m.c.d. se tiene

$$u = \cos \frac{2k'\pi}{n'} + i \operatorname{sen} \frac{2k'\pi}{n'}$$

y siendo k' primo con n', por lo recién demostrado resulta u una raíz primitiva de orden $n' < n$.

El teorema queda así probado.

EJERCICIO. Hallar las raíces primitivas de orden 12 de la unidad.

Se obtienen dándole al parámetro k en la fórmula:

$$\sqrt[12]{((1))} = \cos \frac{2k\pi}{12} + i \operatorname{sen} \frac{2k\pi}{12}$$

todos los valores primos con 12 y menores que 12, es decir: 1 , 5 , 7 , 11 . Por lo tanto, de las doce raíces de orden 12 sólo las cuatro siguientes son primitivas de ese orden:

$$\varepsilon_1 = \cos \frac{2\pi}{12} + i \operatorname{sen} \frac{2\pi}{12} = \cos 30° + i \operatorname{sen} 30° = \frac{\sqrt{3}}{2} + \frac{1}{2} i$$

$$\varepsilon_2 = \cos \frac{10\pi}{12} + i \operatorname{sen} \frac{10\pi}{12} = \cos 150° + i \operatorname{sen} 150° = -\frac{\sqrt{3}}{2} + \frac{1}{2} i$$

$$\varepsilon_3 = \cos \frac{14\pi}{12} + i \operatorname{sen} \frac{14\pi}{12} = \cos 210° + i \operatorname{sen} 210° = -\frac{\sqrt{3}}{2} - \frac{1}{2} i$$

$$\varepsilon_4 = \cos \frac{22\pi}{12} + i \operatorname{sen} \frac{22\pi}{12} = \cos 330° + i \operatorname{sen} 330° = \frac{\sqrt{3}}{2} - \frac{1}{2} i$$

Del teorema 3.4. resulta entonces que para un orden n determinado, hay tantas raíces primitivas de ese orden como números primos con n y menores que n. Este número se llama el indicador de Euler (1707-1783) de n y se representa $\varphi(n)$. Si p es un número primo, todas las raíces p-ésimas de la unidad son primitivas de ese orden excepto 1.

Dados dos números complejos z y z', si se multiplican las diferentes raíces n-ésimas de z por las de z' se obtienen todas las raíces n-ésimas de z.z'. Análogamente para la división.

Además calculando las raíces de orden m de cada una de las n raíces n-ésimas de un número complejo z se obtienen todas las raíces de orden m.n de z. Es decir, se verifican las siguientes reglas:

1) $\sqrt[n]{z} \,.\, \sqrt[n]{z'} = \sqrt[n]{z.z'}$

2) $\dfrac{\sqrt[n]{z}}{\sqrt[n]{z'}} = \sqrt[n]{\dfrac{z}{z'}}, \qquad (z' \neq 0)$

3) $\sqrt[m]{\sqrt[n]{z}} = \sqrt[mn]{z}$

Estas igualdades deben interpretarse en el sentido que el conjunto de valores del primer miembro es igual al conjunto de valores del segundo miembro. Su demostración queda propuesta como ejercicio. (Demostrarlas primero para $z = z' = 1$ y luego en general recordando que las n raíces n-ésimas de un número complejo z se pueden obtener multiplicando una de ellas por las raíces n-ésimas de la unidad).

Observemos que en cambio:

$$\sqrt[n]{z^m} \neq \left(\sqrt[n]{z}\right)^m \qquad (10)$$

Es decir no todos los valores de un miembro son valores del otro. Por ejemplo , si $z = 1$, $n = 4$ y $m = 2$ se tiene

$$\sqrt[4]{1^2} = \sqrt[4]{1} = \begin{cases} 1 \\ -1 \\ i \\ -i \end{cases} \qquad \left(\sqrt[4]{1}\right)^2 = \begin{cases} 1 \\ -1 \end{cases}$$

Luego el primer miembro de (10) tiene cuatro valores distintos mientras que el segundo miembro sólo tiene dos.

OBSERVACION. C es un cuerpo que contiene al sistema R de los números reales. Surge naturalmente la siguiente pregunta: ¿Es posible definir en C una relación de orden de modo tal que C sea un cuerpo ordenado, es decir, una relación de orden total que verifique las leyes de monotonía de la suma y la multiplicación?. La respuesta es negativa. En efecto, en todo cuerpo ordenado el cuadrado de cualquier elemento no nulo es positivo. En particular, $1 = 1^2 > 0$ lo que implica $-1 < 0$. Si C pudiera ordenarse totalmente de modo que fuera un cuerpo ordenado se tendría por un lado , $-1 < 0$ y por otro, $i^2 > 0$. Pero como $i^2 = -1$, sería $-1 > 0$.

Esto prueba la imposibilidad de definir en C una relación de orden total con las mismas propiedades que la definida en R.

Por otra parte, la observación del método seguido para la construcción del sistema C de los números complejos a partir de R lleva a preguntarse si no podría obtenerse mediante una construcción similar y a partir de C, un nuevo cuerpo que contuviera a C como subsistema, o por lo menos a R. La respuesta es negativa, pues se demuestra en general que si K es un cuerpo, entonces $K \times K$ es un cuerpo si y sólo si la ecuación $X^2 + 1 = 0$ no admite solución en K. Y sabemos que C no llena esta última condición

Siguiendo con la idea de generalizar el método de construcción de los números complejos, podríamos también preguntarnos si no es posible definir, no ya en $R \times R$, sino en el conjunto $R^n = R \times R \times \ldots\ldots \times R$, con $n > 2$, una suma y una multiplicación de modo que R^n resulte un cuerpo que contenga a R como subsistema. Aquí también la respuesta es negativa. Se demuestra que esto es sólo posible para n=2 y que C es el único sistema que llena esas condiciones. (Este resultado se conoce con el nombre de Teorema final de la Aritmética y fue demostrado en el siglo pasado por Hankel primero y luego por Weierstrass). Para n=4 existe un sistema en el que se verifican todas las propiedades de cuerpo, excepto la conmutatividad de la multiplicación: es el llamado sistema de los cuaterniones, que es una extensión de R y de C.

De modo que C es la última de las sucesivas ampliaciones de los sistemas numéricos.

En C son posibles todas las operaciones aritméticas y todas las ecuaciones algebraicas con coeficientes numéricos son resolubles en C, propiedad que se expresa diciendo que el cuerpo C es algebraicamente cerrado.

EJERCICIOS.

1. a) Escribir en la forma binómica a+bi los siguientes números complejos:

$(\frac{3}{4}, 2)$; $(-\frac{1}{2}, 1)$; $(\sqrt[3]{2}, -3)$; $(0,-1)$; $(-0.3,0)$

b) Escribir en forma de par (a,b) los siguientes complejos:

$4 + \frac{1}{2}i$; $-8i$; $-\sqrt{3}$; $-1-i$; $\frac{2}{3} - 0.8\,i$

c) Representar en el plano complejo los números de a) y b). ¿Qué distancia hay entre los puntos que representan a los números $(-\frac{1}{2}, 1)$ y $4 + \frac{1}{2}i$?.

2. Expresar los siguientes números complejos en forma binómica:

a) $(1-2i) + (\frac{1}{2} + 7i) + (-2i)$

b) $(-1-2i) - (-3 + \frac{1}{4}i) + 5i - (-\frac{1}{3} - i)$

c) $(1 + \frac{1}{2}i) \cdot (-2+i)$

d) $\frac{1-4i}{\sqrt{2}-i}$

e) $\frac{1}{i} + \frac{3}{1+i} - \frac{(1-i)(2+i)}{3-i} + (\frac{1}{3} - 2i)$

f) $i^{14} - i^9 + 3\,i^5 - i^3 + 1$

g) $(3i-2)^2 - \frac{1}{i^3} + \frac{1}{4}\,i\,(-\sqrt{2} + 5i) - 1$

3. Representar los siguientes números complejos en el plano complejo y expresarlos en forma trigonométrica:

a) $\sqrt{3} + i$

b) $\sqrt{3} - i$

c) $-1-4i$

d) $-1+i$

e) $-i$

f) -17

g) $4i$

h) $i^{15} - 2$

i) $(2+3i)^{-1}$

j) $\operatorname{sen}\frac{\pi}{6} + i \operatorname{sen}\frac{\pi}{4}$

k) $-18+7.4\,i$

l) $1 - i \operatorname{sen}\frac{\pi}{3}$

m) $\operatorname{sen}\alpha + i \operatorname{sen}\alpha$, $\pi \leq \alpha < 2\pi$

n) $(1-2i)^3$

4. Efectuar las siguientes operaciones en forma trigonométrica y expresar el resul-

tado en forma binómica:

a) $(-1+\sqrt{3}\,i)\ .\ (-3i)$

b) $(-\sqrt{3}-i)\ .\ (-\frac{\sqrt{3}}{2}+\frac{1}{2}\,i)$

c) $(-1+\sqrt{3}\,i)\ :\ (-3i)$

d) $(0.5-4i)\ :\ (-3+i)$

5. Aplicar la fórmula de De Moivre para calcular

a) $(1-i)^{47}$ b) $(-\sqrt{3}-i)^{100}$

dando el resultado en forma binómica.

6. a) Calcular y representar las siguientes raíces en el plano complejo y expresarlas en forma binómica:

$\sqrt[3]{-i}$; $\sqrt[4]{-1-\sqrt{3}i}$; $\sqrt[3]{-8}$; $\sqrt{-0.5+\sqrt{2}i}$; $\sqrt[5]{-1+i}$

b) Determinar todas las raíces de los siguientes polinomios complejos:

X^2+1 ; X^3-2i ; $(1+i)X^4+2i$

c) Hallar todos los z tales que:

i) $z^3=-z$, $z\neq 0$ iii) $z+\frac{1}{z}=2z^3$, $|z|=1$

ii) $z^{-2}=-z$ iv) $z^2=\bar{z}$

7. Dados los números complejos:

$z_1=-3(\cos\frac{3}{7}\pi+i\,\text{sen}\,\frac{3}{7}\pi)$ $z_3=\sqrt{3}(-\cos 2+i\,\text{sen}\,2)$

$z_2=2(\cos\frac{\pi}{3}-i\,\text{sen}\,\frac{\pi}{3})$ $z_4=\text{sen}\,1+i\cos 1$

demostrar que en forma polar se escriben como sigue:

$z_1=3(\cos\frac{10}{7}\pi+i\,\text{sen}\,\frac{10}{7}\pi)$ $z_3=\sqrt{3}(\cos(\pi-2)+i\,\text{sen}(\pi-2))$

$z_2=2(\cos\frac{5}{3}\pi+i\,\text{sen}\,\frac{5}{3}\pi)$ $z_4=\cos(\frac{\pi}{2}-1)+i\,\text{sen}(\frac{\pi}{2}-1)$

8. a) Dado $z = a+bi$, sea $Re(z) = a$, $Im(z) = b$. Probar las siguientes relaciones:

$z+\bar{z} = 2\ Re(z)$

$z-\bar{z} = 2\ Im(z)$

$z = \bar{z}$ sssi $Im(z) = 0$

$\bar{z} = -z$ sssi $Re(z) = 0$

$Re(z.\bar{z}' + \bar{z}.z') = z.\bar{z}' + \bar{z}.z'$

$Im(z.\bar{z}' - \bar{z}.z') = z.\bar{z}' - \bar{z}.z'$

b) Interpretar geométricamente en el plano complejo la conjugación. Describir en el plano complejo cada uno de los conjuntos de números complejos que satisfacen las siguientes condiciones:

i) $z = \bar{z}$

ii) $\bar{z} = -z$

iii) $z.\bar{z} = 1$

iv) $z^2 - z + 1 = 0$

v) $z^2 = \bar{z}$

9. Demostrar que:

a) $|z| \geq 0$, $|z| = 0$ sssi $z = 0$

b) $|z| = |-z| = |\bar{z}|$

c) $|z|^2 = z.\bar{z}$

d) $|z.z'| = |z|.|z'|$

e) $\left|\frac{z}{z'}\right| = \frac{|z|}{|z'|}$

f) $|z+z'| \leq |z| + |z'|$

g) $||z| - |z'|| \leq |z-z'|$

h) Si $z = a+0i = a \in R$, probar que el módulo de a coincide con el valor absoluto de a.

i) Si $z.z' = 0$ entonces $z = 0$ ó $z' = 0$

j) Sean $z.z' \in C$ no nulos. Probar aplicando propiedades anteriores que:

$$|z|^{-1} . |z-z'| . |z'|^{-1} = |z^{-1} - z'^{-1}|$$

k) Probar e interpretar geométricamente en el plano complejo la siguiente iden-

tidad (ley del paralelogramo):

$$|z-z'|^2 + |z+z'|^2 = 2(|z|^2 + |z'|^2)$$

(Sugerencia: usar c)).

10. i) Para cada una de las condiciones siguientes determinar la totalidad de complejos z que las satisfacen y dibujar en el plano complejo los recintos definidos en cada caso:

a) $|\mathrm{Re}(z)| < 1 \qquad y \qquad -1 \leq \mathrm{Im}(z) < 1$

b) $\mathrm{Re}(z) \in Z$

c) $\mathrm{Re}(z^2) = 0$

d) $z^{-1} = -z$

e) $|z-1| = |z-2|$

f) $z^2 \in R$

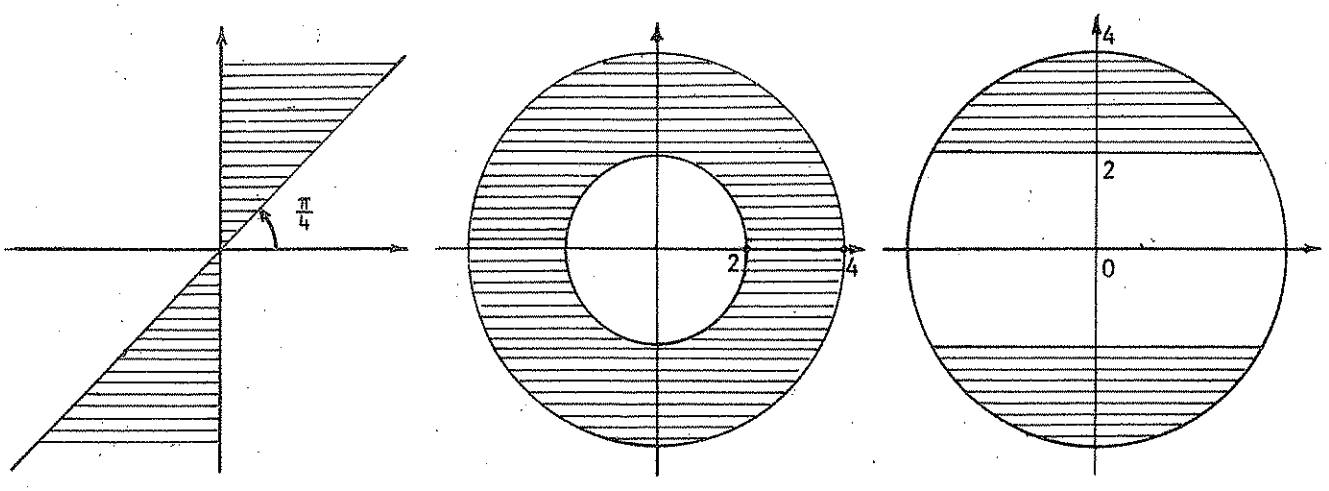
g) $z^2 = \bar{z}^2$

h) $4 < |z-i|^2$

i) $|z+i|^2 + |z-i|^2 = 3$

j) $\left| \frac{z+i}{z-i} \right|^2 = 1$

ii) Representar analíticamente cada uno de los siguientes recintos del plano complejo:



11. Dar los resultados en forma binómica y representarlos en el plano complejo:

a) $\sqrt[3]{(-1-i).\sqrt{3}\,i + (-\frac{1}{2}\,i) : i^9 - (0.5 - \sqrt{3}\,i)}$

b) $\sqrt[3]{\frac{(\sqrt{3} + \sqrt{3}\,i)^2}{1 + \sqrt{5}\,i} - (\sqrt{5} + 2i)}$

c) $\sqrt{\frac{(4 + 3i)(1 + i)}{-4 + 4i}}$

12. a) Demostrar que si r es una raíz de orden n del número complejo z y u_1 , u_2 ,....., u_n son las n raíces n-ésimas de 1 entonces ru_1 , ru_2 ,....., ru_n son las n raíces n-ésimas de z.

b) Hallar las raíces de la unidad de tercer orden, de cuarto orden y de quinto orden y representarlas en el plano complejo.

c) Utilizando los resultados de b) hallar todas las raíces de los polinomios:

$X^3 - 8$; $2X^4 - 81$; $X^5 + 32$

d) Hallar la suma y el producto de todas las raíces de la unidad de tercer orden. Idem para cuarto y quinto orden. ¿Qué propiedades se infieren?. Demostrarlas.

e) ¿Qué se puede decir entonces sobre la suma y el producto de las n raíces n-ésimas de un número complejo cualquiera $z \neq 0$?.

13. a) Hallar las raíces primitivas de la unidad de orden 18. Idem para 15.

b) Decir cuáles de los siguientes complejos son raíces de 1 y de qué orden son raíces primitivas.

$\cos \frac{\pi}{4} + i \operatorname{sen} \frac{\pi}{4}$

$\cos \sqrt{3}\,\pi + i \operatorname{sen} \sqrt{3}\,\pi$

$\cos \frac{8}{21}\,\pi + i \operatorname{sen} \frac{8}{21}\,\pi$

$\frac{1}{2} - \frac{\sqrt{3}}{2}\,i$

$1 - i$

c) Probar que si $t \in R$ entonces $\cos(t\pi) + i \operatorname{sen}(t\pi)$ es raíz de 1 sssi $t \in Q$.

14. Propiedades de las raíces n-ésimas de 1.

a) Demostrar que para cada $n \in N$, el conjunto G_n de las raíces n-ésimas de 1 unidad es un grupo conmutativo con respecto a la multiplicación de números complejos.

b) Si $u \in G_n$ entonces $\bar{u} \in G_n$

c) Si $\varepsilon \in G_n$ es una raíz primitiva de orden n entonces $G_n = \{1, \varepsilon, \varepsilon^2, \ldots, \varepsilon^{n-1}\}$

d) La suma de las n raíces n-ésimas de 1 es 0 y su producto es 1 ó -1 según sea n impar o par.

e) Hallar $G_2 \cap G_4$. Probar que en general $G_n \cap G_m = G_{(n,m)}$. Deducir que $G_n \subset G_m$ si y sólo si n/m.

15. Qué razones pueden darse para invalidar la siguiente "demostración"?:

$$1 = \sqrt{1} = \sqrt{(-1).(-1)} = \sqrt{-1} \; . \; \sqrt{-1} \equiv i.i = i^2 = -1$$

CAPITULO IV

POLINOMIOS Y SUS RAICES

Hay un tipo de problemas cuya resolución consiste en encontrar números que satisfagan ciertas condiciones. A menudo, al traducir estas condiciones en lenguaje matemático aparecen expresiones conocidas con el nombre de ecuaciones algebraicas.

El lector está familiarizado desde el colegio secundario con ecuaciones de primer y segundo grado en una incógnita:

$$a_1X + a_o = 0$$

$$a_2X^2 + a_1X + a_o = 0$$

y sabe resolverlas.

En general, se llama ecuación algebraica de grado n a una expresión del tipo:

$$a_nX^n + a_{n-1}X^{n-1} + \ldots\ldots + a_1X + a_o = 0 \qquad (1)$$

donde $a_n, a_{n-1}, \ldots, a_1, a_o$ son números reales o complejos dados, $a_n \neq 0$, n es un número natural y X un símbolo llamado incógnita o indeterminada.

Resolver la ecuación (1) significa hallar los números t tales que reemplazados en lugar de X verifican la igualdad, es decir tales que

$$a_nt^n + a_{n-1}t^{n-1} + \ldots\ldots + a_1t + a_o = 0$$

Tales números se dicen las raíces de la ecuación (1).

El cálculo de las raíces de las ecuaciones algebraicas es un capítulo muy importante del álgebra y uno de los más antiguos. Durante siglos fue el objetivo central del álgebra.

Para estudiar este problema es conveniente comenzar por estudiar las expresiones del tipo

$$a_nX^n + a_{n-1}X^{n-1} + \ldots\ldots + a_1X + a_o$$

que forman el primer miembro de (1) y que se llaman polinomios en X.

4.1. POLINOMIOS.

En lo que sigue K es el cuerpo de los números racionales, el de los números reales o

el de los números complejos . (K = Q , K = R ó K = C). Consideremos las sucesiones $(a_o , a_1 , a_2 , \ldots\ldots , a_n , \ldots\ldots\ldots)$ de elementos de K tales que a partir de u elemento en adelante son todos nulos, o sea, tales que $a_i = 0$ salvo un número finito de índices, y sea X un símbolo que llamaremos indeterminada. Se llama polinomi en X con coeficientes en K a la expresión (puramente formal):

$$P(X) = a_o + a_1X + a_2X^2 + \ldots\ldots + a_nX^n + \ldots\ldots\ldots$$

Dados dos polinomios $A(X) = a_o + a_1X + a_2X^2 + \ldots\ldots\ldots\ldots$,

$$B(X) = b_o + b_1X + b_2X^2 + \ldots\ldots\ldots\ldots$$

$A(X) = B(X)$ si y sólo si $a_i = b_i$ para $i = 0,1,2,\ldots\ldots\ldots$

Al conjunto de todos los polinomios en X con coeficientes en K se lo representa K[(Así Q[X] , R[X] , C[X] son los conjuntos de polinomios en X con coeficientes racionales, reales y complejos respectivamente. Es claro que $Q[X] \subset R[X] \subset C[X]$).

En K[X] se definen dos operaciones binarias, suma y multiplicación, como sigue:

Si $A(X) = a_o + a_1X + a_2X^2 + \ldots\ldots$, $B(X) = b_o + b_1X + b_2X^2 + \ldots\ldots \in K$

se llama:

Suma de A(X) y B(X) y se representa A(X)+B(X) al polinomio:

$A(X)+B(X) = s_o + s_1X + s_2X^2 + \ldots\ldots$ tal que $s_i = a_i + b_i$ para i = 0,1,2,.

es decir

$$A(X)+B(X) = (a_o+b_o) + (a_1+b_1)X + (a_2+b_2)X^2 + \ldots\ldots\ldots\ldots$$

Producto de A(X) por B(X) y se representa A(X).B(X) al polinomio:

$A(X).B(X) = p_o + p_1X + p_2X^2 + \ldots\ldots$ tal que $p_i = \sum_{k+j=i} a_k b_j$ para i=0,1,.

es decir:

$$p_o = a_o b_o$$
$$p_1 = a_o b_1 + a_1 b_o$$
$$p_2 = a_o b_2 + a_1 b_1 + a_2 b_o$$
$$p_3 = a_o b_3 + a_1 b_2 + a_2 b_1 + a_3 b_o$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

Dado un polinomio $P(X) = a_o + a_1X + a_2X^2 + \ldots\ldots$, si $a_i = 0$ para todo índice i entonces $P(X)$ se llama el polinomio nulo y se representa $P(X) = 0$. Si no todos los a_i son nulos y n es el mayor índice tal que $a_n \neq 0$, n se llama el grado del polinomio $P(X)$ y éste se escribe:

$$P(X) = a_o + a_1X + a_2X^2 + \ldots\ldots + a_nX^n$$

que es la forma habitual de escribir un polinomio, prescindiendo de los términos a_iX^i con $a_i = 0$. a_o , a_1 , a_2 ,...., a_n se llaman los coeficientes de $P(X)$, a_n se llama el coeficiente principal y a_o el término independiente o constante. Si $a_n = 1$ el polinomio se dice mónico.

Observemos que el grado del polinomio nulo no está definido. Los polinomios de grado cero, es decir de la forma $P(X) = a_o$, y el polinomio nulo se llaman polinomios constantes. Usualmente al escribir un polinomio se escribe X^i en lugar de $1X^i$.

EJEMPLOS.

1) Consideremos los siguientes polinomios de $R[X]$:

$A(X) = 3X^7 - X^3 + 2X^2 - \frac{1}{3}X - 7$

$B(X) = X^4 + X^3 + 5X^2 + X + 3$

$C(X) = X^2 - X$

$D(X) = -5$

gr $A(X) = 7$; gr $B(X) = 4$; gr $C(X) = 2$; gr $D(X) = 0$; $B(X)$ y $C(X)$ son mónicos, el coeficiente principal de $A(X)$ es 3 y el de $D(X)$ es -5 ; los términos independientes son respectivamente: -7 , 3 , 0 y -5.

$A(X)+B(X) = 3X^7 + X^4 + 7X^2 + \frac{2}{3}X - 4$

$C(X)+D(X) = X^2 - X - 5$

$A(X).C(X) = 3X^9 - 3X^8 - X^5 + 3X^4 - \frac{7}{3}X^3 - \frac{20}{3}X^2 + 7X$

$C(X).D(X) = -5X^2 + 5X$

2) Dados los polinomios de $C[X]$:

$A(X) = X^5 - iX^3 + (2-i)$; $B(X) = 3X^3 + 4$; $C(X) = (1+i)X^3 - 2i$ se tiene:

gr $A(X) = 5$, gr $B(X) = 3$, gr $C(X) = 3$; los coeficientes principales son 1 , 3 y 1+i respectivamente; los términos independientes son 2-i , 4 y -2i.

$A(X)+B(X) = X^5 + (3-i)X^3 + (6-i)$

$A(X)+C(X) = X^5 + X^3 + (2-3i)$

$A(X).C(X) = (1+i)X^8 + (1-i)X^6 - 2iX^5 + (1+i)X^3 - (2+4i)$

Con respecto al grado de una suma y de un producto vale la siguiente

PROPOSICION 4.1. Dados dos polinomios $A(X) \neq 0$, $B(X) \neq 0$ se verifican las siguientes relaciones:

1) Si $A + B \neq 0$, $gr(A + B) \leq \max(gr\ A\ ,\ gr\ B)$.

2) $A.B \neq 0$ y $gr(A.B) = gr\ A + gr\ B$.

Demostración:

1) Sigue claramente de la definición de suma de polinomios. (Observar que si $gr\ A = gr\ B = n$ y $a_n = -b_n$ entonces $gr(A + B) < n$).

2) Si $gr\ A = n$, $gr\ B = m$ y $A.B = p_o + p_1X + p_2X^2 + \ldots\ldots$ de la definición de producto resulta que $p_{n+m} = a_n.b_m$ y $p_i = 0$, $\forall\ i > n+m$.
Entonces $a_n \neq 0$, $b_m \neq 0 \Longrightarrow a_n.b_m \neq 0 \Longrightarrow A.B \neq 0$ y $gr(A.B) = gr\ A + gr\ B$.

TEOREMA 4.1. En $K[X]$ la suma y la multiplicación tienen las siguientes propiedades:

S1. Propiedad asociativa de la suma:
$(A + B) + C = A + (B + C)$, $\forall\ A\ ,\ B\ ,\ C \in K[X]$

S2. Propiedad conmutativa de la suma:
$A + B = B + A$, $\forall\ A\ ,\ B \in K[X]$

S3. Existe neutro para la suma: el polinomio nulo es tal que $A + 0 = A$, $\forall\ A \in K[X]$

S4. Todo polinomio de $K[X]$ tiene simétrico en $K[X]$, es decir cualquiera sea $A \in K[X]$ existe $B \in K[X]$ tal que $A + B = 0$.

M1. Propiedad asociativa de la multiplicación:
$(A.B).C = A.(B.C)$, $\forall\ A\ ,\ B\ ,\ C \in K[X]$

M2. Propiedad conmutativa de la multiplicación:
$A.B = B.A$, $\forall\ A\ ,\ B \in K[X]$

M3. Existe neutro para la multiplicación: el polinomio constante 1 es tal que $A.1 = A$, $\forall\ A \in K[X]$.

D. Propiedad distributiva de la $\cdot$ con respecto a la $+$:
$A.(B + C) = A.B + A.C$, $\forall\ A\ ,\ B\ ,\ C \in K[X]$

S. En $K[X]$ no hay divisores de cero, es decir: Si $A.B = 0$ entonces $A = 0$ ó $B = 0$.

Demostración: A cargo del lector.

Las propiedades anteriores se resumen diciendo que $K[X]$ es un anillo conmutativo sin divisores de cero. $K[X]$ se llama el anillo de polinomios en una indeterminada con coeficientes en K.

OBSERVACION: $K[X]$ no es un cuerpo porque los únicos polinomios inversibles con respecto a la multiplicación son las constantes no nulas. En efecto, supongamos que un polinomio $A(X)$ tiene inverso $B(X)$ con respecto a la multiplicación:

$$A(X).B(X) = 1$$

Luego $$gr\ A + gr\ B = gr\ 1 = 0$$

lo que implica $gr\ A = gr\ B = 0$ y $A(X)$ y $B(X)$ son constantes.

Vamos a ver que en $K[X]$ existe una "división entera", analogamente a lo que sucede en el anillo Z de los enteros.

TEOREMA 4.2. Dados dos polinomios $A,B \in K[X]$, $B \neq 0$, existen dos polinomios $Q,R \in K[X]$, llamados el cociente y el resto respectivamente de dividir A por B, unívocamente determinados, tales que:

$$A = Q.B + R \quad y \quad R = 0 \ \text{ó} \ gr\ R < gr\ B$$

Demostración: Demostraremos primero la existencia de un par Q,R al menos.

Si $A = P.B$ para algún $P \in K[X]$, entonces haciendo $Q = P$ y $R = 0$ se tiene un par de polinomios en esas condiciones.

Si $A \neq P.B$ para todo $P \in K[X]$, consideremos el conjunto de todos los polinomios de la forma $A - P.B$, $P \in K[X]$. Los grados de estos polinomios forman un conjunto no vacío de enteros no negativos. Por el principio de buena ordenación existe un grado mínimo. Sea R un polinomio de grado mínimo entre los de la forma indicada , $R = A - Q.B$ para un cierto $Q \in K[X]$. Luego $A = Q.B + R$. Veamos que $gr\ R < gr\ B$.

Supongamos que $gr\ R \geq gr\ B$

$$R = r_o + r_1 X + r_2 X^2 + \ldots + r_s X^s$$

$$B = b_o + b_1 X + b_2 X^2 + \ldots + b_n X^n \quad , \quad s \geq n$$

Entonces el polinomio $R - \frac{r_s}{b_n} X^{s-n} . B$ es de grado $< s$. Además

$$R - \frac{r_s}{b_n} X^{s-n} . B = A - Q.B - \frac{r_s}{b_n} X^{s-n} . B = A - (Q + \frac{r_s}{b_n} X^{s-n}) . B$$

Es decir, el polinomio $R - \frac{r_s}{b_n} X^{s-n} . B$ es de la forma $A - P.B$ y de grado menor que

el grado de R, lo que contradice la elección de R. Esta contradicción prueba que gr R < gr B.

Unicidad. Sean $Q,Q',R,R' \in K[X]$ tales que:

$$A = Q.B + R \quad y \quad R = 0 \quad ó \quad gr\ R < gr\ B$$

$$A = Q'.B + R' \quad y \quad R' = 0 \quad ó \quad gr\ R' < gr\ B$$

Luego $Q.B + R = Q'.B + R'$

$$(Q - Q').B = R' - R$$

Si $R \neq R'$, $R' - R \neq 0$ y $gr\ (Q-Q') + gr\ B = gr\ (R'-R)$.

Pero de las condiciones que verifican R y R' resulta:

$$gr\ (R' - R) \leq máx\ (gr\ R\ ,\ gr\ R') < gr\ B$$

Luego $gr\ B \leq gr(Q-Q') + gr\ B = gr(R'-R) < gr\ B$ o sea $gr\ B < gr\ B$.

Esta contradicción prueba que $R = R'$.

Entonces

$$(Q - Q')\ .\ B = 0$$

Como $B \neq 0$ por hipótesis y en $K[X]$ no hay divisores de cero es $Q - Q' = 0$, es decir $Q = Q'$.

El teorema queda así demostrado.

El cociente y el resto de dividir un polinomio por otro no nulo se calculan mediante un algoritmo similar al de la división de números enteros.

Calculemos,por ejemplo, el cociente y el resto de dividir

$$A(X) = 6X^6 - 5X^5 - 5X^4 - \frac{17}{2} X^3 + 6X^2 - 2 \quad \text{por} \quad B(X) = 2X^3 - 3X^2 + 1 .$$

$$\begin{array}{l|l}
6X^6 - 5X^5 - 5X^4 - \frac{17}{2} X^3 + 6X^2 - 2 & 2X^3 - 3X^2 + 1 \\
\underline{6X^6 - 9X^5 \qquad + 3X^3} & \\
4X^5 - 5X^4 - \frac{23}{2} X^3 + 6X^2 - 2 & 3X^3 + 2X^2 + \frac{1}{2} X - 5 \\
\underline{4X^5 - 6X^4 \qquad + 2X^2} & \\
X^4 - \frac{23}{2} X^3 + 4X^2 - 2 & \\
\underline{X^4 - \frac{3}{2} X^3 + \frac{1}{2} X} & \\
- 10 X^3 + 4X^2 - \frac{1}{2} X - 2 & \\
\underline{- 10 X^3 + 15X^2 \qquad - 5} & \\
- 11X^2 - \frac{1}{2} X + 3 &
\end{array}$$

El cociente es $Q(X) = 3X^3 + 2X^2 + \frac{1}{2}X - 5$ y el resto $R(X) = -11X^2 - \frac{1}{2}X + 3$.

Los polinomios así obtenidos son efectivamente el cociente y el resto de dividir A por B pues el procedimiento en general es el siguiente:

Si $A(X) = a_n X^n + \ldots\ldots + a_1 X + a_o$

$B(X) = b_m X^m + \ldots\ldots + b_1 X + b_o$

y gr $A \geq$ gr B , se calcula $C_1 = A - \frac{a_n}{b_m} X^{n-m} \,.\, B$

Si gr $C_1 = n_1$ y su coeficiente principal es c_{n_1} , se calcula

$$C_2 = C_1 - \frac{c_{n_1}}{b_m} X^{n_1-m} \,.\, B = A - \left(\frac{a_n}{b_m} X^{n-m} + \frac{c_{n_1}}{b_m} X^{n_1-m} \right) . B$$

Si gr $C_2 = n_2$ y su coeficiente principal es c_{n_2} , se calcula

$$C_3 = C_2 - \frac{c_{n_2}}{b_m} X^{n_2-m} \,.\, B = A - \left(\frac{a_n}{b_m} X^{n-m} + \frac{c_{n_1}}{b_m} X^{n_1-m} + \frac{c_{n_2}}{b_m} X^{n_2-m} \right) . B$$

y así siguiendo se obtienen polinomios $C_1 = A - P_1.B$, $C_2 = A - P_2.B$, $C_3 = A - P_3.B$, ... de grados estrictamente decrecientes hasta llegar a uno $R = A - Q.B$ nulo ó de grado menor que B. En virtud de la unicidad del cociente y el resto, R es el resto y Q el cociente de dividir A por B.

Si gr A < gr B entonces $Q = 0$ y $R = A$

Regla de Ruffini. Un caso particular es el de la división de un polinomio

$A(X) = a_n X^n + \ldots\ldots + a_1 X + a_o$ por otro de la forma $X - c$.

Aplicando el algoritmo de la división, se ve que los coeficientes del cociente

$$Q(X) = q_{n-1} X^{n-1} + \ldots\ldots + q_1 X + q_o$$

verifican las siguientes relaciones:

$$q_{n-1} = a_n$$

$$q_{n-2} = a_{n-1} + q_{n-1}.c$$

$$q_{n-3} = a_{n-2} + q_{n-2}.c$$

. . . . . .

$$q_1 = a_2 + q_2.c$$

$$q_o = a_1 + q_1.c$$

y el resto de la división es $R(X) = a_o + q_o.c$

La regla que de aquí resulta para calcular los coeficientes del cociente y el resto de dividir un polinomio A(X) por otro de la forma X - c se llama la regla de Ruffini y el cálculo se dispone prácticamente como se ve en los siguientes ejemplos.

EJEMPLOS.

1) Dividir $A(X) = 5X^7 - 9X^6 - 3X^5 + X^3 + 8X + 1$ por $X - 2$

	5	-9	-3	0	1	0	8	1
2		10	2	-2	-4	-6	-12	-8
	5	1	-1	-2	-3	-6	-4	-7

El cociente y el resto son:

$Q(X) = 5X^6 + X^5 - X^4 - 2X^3 - 3X^2 - 6X - 4$; $R(X) = -7$

2) Dividir $A(X) = X^5 - 2X^4 + \frac{1}{2}X - 3$ por $X + 3$

	1	-2	0	0	$\frac{1}{2}$	-3
-3		-3	15	-45	135	$-\frac{813}{2}$
	1	-5	15	-45	$\frac{271}{2}$	$-\frac{819}{2}$

El cociente es $Q(X) = X^4 - 5X^3 + 15X^2 - 45X + \frac{271}{2}$ y el resto es $R(X) = -\frac{819}{2}$

4.2. DIVISIBILIDAD EN EL ANILLO DE POLINOMIOS K[X] .

En el anillo de polinomios K[X] se puede desarrollar una teoría de divisibilidad completamente semejante a la del anillo Z de los enteros y demostrar un teorema de factorización única en polinomios irreducibles, análogo al teorema fundamental de la aritmética. Es lo que haremos a continuación.

Relación divide.

Definición. Se dice que un polinomio A(X) divide a otro B(X) si existe un polinomio C(X) tal que B(X) = C(X).A(X). Se escribe A/B.

Si A/B también se dice que A(X) es un divisor de B(X) o que B(X) es un múltiplo de A(X).

Esta relación tiene las siguientes

PROPIEDADES.

1. A/A , $\forall$ $A \in K[X]$
2. Si A/B y B/C entonces A/C
3. $A/0$, $\forall$ $A \in K[X]$
4. Si A/B y A/C entonces $A/S.B+T.C$, $\forall$ $S,T \in K[X]$

Polinomios unitarios y asociados.

Definición. Un polinomio de $K[X]$ se dice unitario si es inversible con respecto a la multiplicación.

Hemos visto que los polinomios unitarios son las constantes no nulas.

En términos de divisibilidad, un polinomio es unitario si y sólo si divide a cualquier otro.

PROPOSICION 4.2. Dados dos polinomios $A(X)$, $B(X)$ las siguientes propiedades son equivalentes:

a) $A(X)$ y $B(X)$ tienen los mismos divisores.

b) $A(X)$ y $B(X)$ difieren en un factor unitario. (Es decir, en una constante no nula).

c) A/B y B/A

Demostración:

a) $\Longrightarrow$ b). Supongamos que $A(X)$ y $B(X)$ tienen los mismos divisores. Como A/A entonces A/B. Como B/B entonces B/A. Luego existen polinomios $C(X)$ y $C'(X)$ tales que $B(X) = C(X).A(X)$ (1) y $A(X) = C'(X).B(X)$ (2)

Entonces

$$A(X) = C'(X).C(X).A(X)$$

Si $A(X) \neq 0$ de $A(X).[1 - C'(X).C(X)] = 0$ sigue $C'(X).C(X) = 1$ pues en $K[X]$ no hay divisores de cero. Es decir $C(X)$ y $C'(X)$ son polinomios unitarios y de (1) y (2) sigue que se verifica b).

Si $A(X) = 0$ entonces $B(X) = 0$ y es claro que b) se verifica.

b) $\Longrightarrow$ c). Supongamos que $A(X)$ y $B(X)$ difieren en un factor unitario, es decir, por ejemplo, que $A(X) = k.B(X)$, con $k \in K$, $k \neq 0$. Luego B/A. Por otro lado, de la igualdad anterior resulta $B(X) = k^{-1}.A(X)$ o sea A/B.

c) $\Longrightarrow$ a). Supongamos que A/B y B/A. Hay que probar que $A(X)$ y $B(X)$ tienen los mis

mos divisores, es decir que si un polinomio D(X) es tal que D/A entonces D/B y recíprocamente, que si D/B entonces D/A, lo que resulta fácilmente de la hipótesis y la transitividad de la relación divide.

Observe el lector que la demostración de esta proposición es textualmente la misma que la de la proposición 2.1. correspondiente a Z. Así sucede con todas las demás demostraciones por lo que omitiremos escribirlas.

Queda a cargo del lector hacer su transcripción en términos de polinomios.

Definición. Dos polinomios A(X) y B(X) que verifican las condiciones de la proposición anterior se dicen asociados.

La relación "ser asociados" es una relación de equivalencia en K[X] y de la proposición resulta que los asociados de un polinomio $A(X) \in K[X]$ son todos los de la forma $k.A(X)$, con $k \in K$, $k \neq 0$.

Todo polinomio $A(X) \in K[X]$ es divisible por sus asociados y por los polinomios unitarios, es decir por los polinomios de la forma $k.A(X)$, con $k \in K$, $k \neq 0$ y por todas las constantes no nulas. Estos se llaman los divisores triviales de A(X). Un divisor de A(X) distinto de ellos se dice un divisor propio.

Notemos que las constantes no nulas desempeñan en K[X] el papel de 1 y -1 en Z.

Máximo común divisor.

Definición. Dados dos polinomios A(X) y B(X), un polinomio D(X) se llama un máximo común divisor de A(X) y B(X) si verifica las dos siguientes condiciones:

D1) D/A y D/B

D2) Si D'(X) es un polinomio tal que D'/A y D'/B entonces D'/D

Se escribe $D = (A,B)$.

Como en K[X] existe la división entera, es posible aplicar el algoritmo de Euclides para calcular un m.c.d. de dos polinomios A(X) y B(X), analogamente a como se procede en el anillo Z.

PROPOSICION 4.3. Si A(X) y B(X) son dos polinomios, $B(X) \neq 0$, y R(X) es el resto de dividir A(X) por B(X) entonces A(X) y B(X) tienen los mismos divisores comunes que B(X) y R(X).

COROLARIO. $D = (A,B) \iff D = (B,R)$.

Algoritmo de Euclides.

Dados dos polinomios no nulos A(X) , B(X) se hacen divisiones sucesivas de acuerdo con el esquema ya conocido. Como los grados de los restos que se van obteniendo son

estrictamente decrecientes, el procedimiento no se puede reiterar más que un número finito de veces. El último paso es una división de resto cero.

	$Q_1(X)$	$Q_2(X)$	. . . . .		$Q_n(X)$	$Q_{n+1}(X)$
$A(X)$	$B(X)$	$R_1(X)$	. . . . .	$R_{n-2}(X)$	$R_{n-1}(X)$	$R_n(X)$
$R_1(X)$	$R_2(X)$		. . . . .	$R_n(X)$	0	

PROPOSICION 4.4. Si $A(X)$, $B(X) \in K[X]$ son dos polinomios no nulos, el último resto no nulo que se obtiene en el algoritmo de Euclides es un máximo común divisor $D(X)$ de $A(X)$ y $B(X)$ y $D(X) = S(X).A(X) + T(X).B(X)$, con $S(X)$, $T(X) \in K[X]$.

Queda así probada la existencia de un m.c.d. para dos polinomios no nulos. Si alguno de ellos es nulo, por ejemplo $A(X) = 0$, entonces $(0,B(X)) = B(X)$ y $B(X) = 1.0 + 1.B(X)$.

Además de la definición de m.c.d. resulta que si $D(X)$ es un m.c.d. de $A(X)$ y $B(X)$, otro polinomio $D'(X)$ es un m.c.d. de $A(X)$ y $B(X)$ si y sólo si $D(X)$ y $D'(X)$ son asociados.

Vale entonces el siguiente

TEOREMA 4.3. Para todo par de polinomios $A(X)$, $B(X)$ de $K[X]$ existe un m.c.d. $D(X)$ en $K[X]$ y $D(X) = S(X).A(X) + T(X).B(X)$, con $S(X)$, $T(X) \in K[X]$. Además los m.c.d. de $A(X)$ y $B(X)$ son de la forma $k.D(X)$, con $k \in K$, $k \neq 0$.

Como el m.c.d. de dos polinomios es único salvo constantes no nulas, se considera el m.c.d. mónico y, dada su unicidad , se habla de el m.c.d.

EJEMPLOS.

1) Dados los polinomios de $R[X]$:

$A(X) = X^3 + 4X^2 + 3X + 12$ y $B(X) = X^4 - 3X^3 + 5X^2 - 9X + 6$

Hallar el m.c.d. y expresarlo en la forma $S(X).A(X) + T(X).B(X)$, $S(X),T(X) \in R[X]$

	$X - 7$	$X + 4$
$X^4 - 3X^3 + 5X^2 - 9X + 6$	$X^3 + 4X^2 + 3X + 12$	$X^2 + 3$
$X^4 + 4X^3 + 3X^2 + 12X$	$X^3 + 3X$	
$- 7X^3 + 2X^2 - 21X + 6$	$4X^2 + 12$	
$- 7X^3 - 28X^2 - 21X - 84$	$4X^2 + 12$	
$30X^2 + 90$	0	
$X^2 + 3$		

El m.c.d. es el último resto no nulo, es decir $X^2 + 3$. Para expresarlo como un múltiplo de A(X) más uno de B(X), del algoritmo de Euclides resulta:

$$A(X) = (X - 7) \cdot B(X) + 30(X^2 + 3)$$

Luego, $$X^2 + 3 = \frac{1}{30} A(X) + (-\frac{1}{30} X + \frac{7}{30}) \cdot B(X)$$

OBSERVACION. Como un polinomio A(X) tiene los mismos divisores que cualquier polinomio de la forma k.A(X) donde k es una constante no nula, dados dos polinomios A(X) y B(X) resulta que:

$$(A(X), B(X)) = (k.A(X) , k'.B(X))$$

cualesquiera sean $k, k' \in K$, $k \neq 0$, $k' \neq 0$.

Entonces para calcular el m.c.d. de dos polinomios se pueden simplificar los cálculos descartando los factores constantes de los polinomios dados y de los restos sucesivos obtenidos en el algoritmo de Euclides puesto que esto sólo varía el resultado a lo sumo en un factor constante, lo que no interesa en el caso del m.c.d.

2) Hallar el m.c.d. de los siguientes polinomios de Q[X] :

$A(X) = 6X^4 - 3X^3 + 11X^2 - 15X + 1$; $B(X) = 6X^3 + 14X - 8$

Podemos considerar $B'(X) = 3X^3 + 7X - 4$

	$2X - 1$	$3X$	$X + 1$	$X - 1$
$6X^4 - 3X^3 + 11X^2 - 15X + 1$	$3X^3 + 7X - 4$	$X^2 + 1$	$X - 1$	1
$6X^4 + 14X^2 - 8X$	$3X^3 + 3X$	$X^2 - X$	0	
$- 3X^3 - 3X^2 - 7X + 1$	$4X - 4$	$X + 1$		
$- 3X^3 - 7X + 4$	$4(X - 1)$	$X - 1$		
$- 3X^2 - 3$		2		
$-3(X^2 + 1)$				

El m.c.d. es 1. Para expresar el m.c.d. en la forma S(X).A(X) + T(X).B(X), del algoritmo de Euclides se deduce:

$A(X) = (2X - 1).B'(X) + (-3)(X^2 + 1)$

$B'(X) = 3X(X^2 + 1) + 4(X - 1)$

$X^2 + 1 = (X + 1)(X - 1) + 2$

Entonces

$$-3(X^2 + 1) = A(X) - (2X - 1).B'(X)$$

$$4(X - 1) = B'(X) - 3X(X^2 + 1) = B'(X) + X[A(X) - (2X - 1).B'(X)] =$$

$$= X.A(X) + (-2X^2 + X + 1).B'(X) .$$

$$2 = X^2 + 1 - (X + 1)(X - 1) = -\frac{1}{3} A(X) + \frac{1}{3} (2X - 1).B'(X) -$$

$$- (X + 1) \frac{1}{4} [X.A(X) + (-2X^2 + X + 1).B'(X)] .$$

$$2 = (-\frac{1}{4} X^2 - \frac{1}{4} X - \frac{1}{2}).A(X) + (\frac{1}{2} X^3 + \frac{1}{4} X^2 + \frac{1}{6} X - \frac{7}{12}).B'(X)$$

De modo que

$$1 = (-\frac{1}{8} X^2 - \frac{1}{8} X - \frac{1}{6}).A(X) + (\frac{1}{8} X^3 + \frac{1}{16} X^2 + \frac{1}{24} X - \frac{7}{48}).B(X)$$

Polinomios irreducibles y relativamente primos.

Definición. Un polinomio no constante $P(X) \in K[X]$ se dice irreducible o primo en $K[X]$ si no admite divisores distintos de los triviales, es decir si no se puede escribir como producto de dos polinomios no constantes de $K[X]$.

Por ejemplo, el polinomio $X^2 - 2 \in Q[X]$ es irreducible en $Q[X]$. En cambio si se lo considera polinomio de $R[X]$, no es irreducible en $R[X]$ pues se puede escribir:

$$X^2 - 2 = (X + \sqrt{2}).(X - \sqrt{2})$$

El polinomio $X^2 + 1$ es irreducible en $Q[X]$ y en $R[X]$ y es reducible en $C[X]$ pues

$$X^2 + 1 = (X + i)(X - i)$$

Está claro que todo polinomio de primer grado es irreducible puesto que el grado de un producto es la suma de los grados de los factores.

Definición. Dos polinomios de $K[X]$ se dicen relativamente primos si su m.c.d. es 1.

TEOREMA 4.4. Si $A/B.C$ y $(A,B) = 1$ entonces A/C.

La demostración es igual a la vista en Z.

COROLARIO 1. Si $P/B.C$ y P es un polinomio irreducible entonces P/B ó P/C.

COROLARIO 2. Si $P/A_1.A_2.....A_n$ y P es un polinomio irreducible entonces P/A_i para algún índice i, $1 \leq i \leq n$.

TEOREMA 4.5. (De factorización única).
Todo polinomio no constante de $K[X]$ se puede escribir como producto de una constante por polinomios irreducibles mónicos y esta descomposición es única salvo el orden de los factores.

Demostración: Probaremos la existencia de una descomposición de ese tipo por inducción sobre el grado del polinomio.

Sea $A(X) \in K[X]$, $gr\ A(X) \geq 1$.

Si $gr\ A(X) = 1$ entonces $A(X) = a_1X + a_o$. Luego $A(X) = a_1(X + \frac{a_o}{a_1})$ y $X + \frac{a_o}{a_1}$ es irreducible mónico.

Sea $gr\ A(X) = n > 1$ y supongamos que la descomposición existe para todos los polinomios no constantes de $K[X]$ de grado menor que n. Hay que probar que $A(X)$ también puede escribirse como producto de una constante por polinomios irreducibles mónicos.

Si $A(X)$ es irreducible, sea a_n su coeficiente principal.

Entonces $A(X) = a_n(\frac{1}{a_n} . A(X))$ y $\frac{1}{a_n} . A(X)$ es irreducible mónico.

Si $A(X)$ no es irreducible se puede descomponer en el producto de dos polinomios no constantes: $A(X) = B(X).C(X)$ con $gr\ B < gr\ A$, $gr\ C < gr\ A$.

Por la hipótesis de inducción $B(X)$ y $C(X)$ se pueden escribir como producto de una constante por polinomios irreducibles mónicos y entonces lo mismo sucede con $A(X)$.

Queda probada así la existencia de la descomposición para todo polinomio no constante de $K[X]$.

La unicidad se demuestra en forma análoga a la vista en el caso de los números enteros. (Teorema 2.10)

De acuerdo con este teorema, si $A(X) = a_nX^n + a_{n-1}X^{n-1} + \ldots\ldots + a_1X + a_o$ es un polinomio no constante de $K[X]$ se puede escribir:

$$A(X) = k.P_1(X).P_2(X).....P_s(X)$$

donde $P_1(X)$, $P_2(X)$, ... , $P_s(X)$ son polinomios irreducibles mónicos y $k \in K$.
Como los coeficientes principales de los $P_i(X)$ son todos 1 , es $k = a_n$.

Luego $A(X)$ se escribe:

$$A(X) = a_nP_1(X).P_2(X).....P_s(X)$$

Si aparecen factores repetidos se asocia los factores iguales escribiendo su producto en forma de potencia.

Por ejemplo, dado el polinomio

$A(X) = 3X^7 + 6X^6 + 6X^5 + 6X^4 - 15X^3 - 36X^2 - 18X \in Q[X]$ su descomposición en factores irreducibles en $Q[X]$ es la siguiente:

$$A(X) = 3X(X + 1)^2(X^2 - 2)(X^2 + 3)$$

Si se considera $A(X) \in R[X]$, el polinomio $X^2 - 2$ no es irreducible en $R[X]$. La descomposición de $A(X)$ en $R[X]$ es:

$$A(X) = 3X(X + 1)^2 (X + \sqrt{2})(X - \sqrt{2})(X^2 + 3)$$

Finalmente, la descomposición de $A(X)$ en factores irreducibles en $C[X]$ es:

$$A(X) = 3X(X + 1)^2 (X + \sqrt{2})(X - \sqrt{2})(X + \sqrt{3}\, i)(X - \sqrt{3}\, i)$$ *

NOTA. El lector ha comprobado la semejanza de las teorías de divisibilidad desarrolladas en Z y en $K[X]$ y la analogía de todas las demostraciones. ¿Podría señalar cuáles son las propiedades comunes a Z y a $K[X]$ que justifican este hecho?.

4.3. RAICES DE LOS POLINOMIOS.

Dado un polinomio $P(X) = a_n X^n + a_{n-1} X^{n-1} + \ldots\ldots + a_1 X + a_o$ de $K[X]$ y un elemento $c \in K$ se llama valor de $P(X)$ en c al elemento $P(c) \in K$ que se obtiene reemplazando la indeterminada X por c en el polinomio y efectuando las operaciones indicadas.

$$P(c) = a_n c^n + a_{n-1} c^{n-1} + \ldots\ldots + a_1 c + a_o$$

Por ejemplo, dado $P(X) = 4X^4 - \frac{1}{2} X^3 + X - 1 \in R[X]$ es

$$P(2) = 4.2^4 - \frac{1}{2} . 2^3 + 2 - 1 = 61$$

Si gr $P(X) = 0$, es decir si $P(X) = a_o$ entonces $P(c) = a_o$, $\forall\, c \in K$.

Definición. Dado un polinomio $P(X) = a_n X^n + a_{n-1} X^{n-1} + \ldots\ldots + a_1 X + a_o \in K[X]$, un elemento $c \in K$ se dice una raíz (ó un cero) de $P(X)$ si $P(c) = 0$, es decir si

$$a_n c^n + a_{n-1} c^{n-1} + \ldots\ldots + a_1 c + a_o = 0$$

TEOREMA 4.6. (Teorema del resto).

El resto de dividir un polinomio $P(X)$ por otro de la forma $X - c$ es $P(c)$.

Demostración: Sean $Q(X)$ y $R(X)$ el cociente y el resto de dividir $P(X)$ por $X - c$:

$$P(X) = Q(X).(X - c) + R(X) \quad y \quad R(X) = 0 \text{ ó } gr\, R < gr(X-c)$$

Como $gr(X-c) = 1$ resulta $gr\, R = 0$ o sea el resto es una constante.

$$P(c) = Q(c).(c - c) + R \text{ lo que implica } P(c) = R . \qquad c.q.d.$$

COROLARIO. Un elemento $c \in K$ es raíz de un polinomio $P(X)$ de $K[X]$ si y sólo si $P(X)$ es divisible por $X - c$.

En efecto ,

c raíz de $P(X) \iff P(c) = 0 \iff$ el resto de dividir $P(X)$ por $X - c$ es cero $\iff$

$\iff P(X)$ es divisible por $X - c$.

En virtud del teorema anterior se puede utilizar la regla de Ruffini para calcular el valor de un polinomio $P(X)$ para $X = c$ ya que el último número que se obtiene aplicándola es el resto de dividir $P(X)$ por $X - c$, o sea $P(c)$. También sirve entonces para verificar rápidamente si c es o no una raíz de $P(X)$. En general, calcular $P(c)$ por la regla de Ruffini requiere cálculos más sencillos que la sustitución directa en el polinomio.

EJEMPLOS.

1) Dado $P(X) = 3X^5 - 2X^3 + \frac{1}{3}X - 1 \in R[X]$ hallar $P(-1)$. Aplicando Ruffini se tiene:

	3	0	-2	0	$\frac{1}{3}$	-1
-1		-3	3	-1	1	$-\frac{4}{3}$
	3	-3	1	-1	$\frac{4}{3}$	$-\frac{7}{3}$

Luego $P(-1) = -\frac{7}{3}$

2) Dado $P(X) = 3X^3 - (1+6i)X^2 + X - (4+2i) \in C[X]$ hallar $P(2i)$.

	3	-1-6i	1	-4-2i
2i		6i	-2i	4+2i
	3	-1	1-2i	0

Luego $P(2i) = 0$ y 2i es una raíz de $P(X)$.

Raíces múltiples.

Hemos visto que si c es una raíz de $P(X)$ entonces $P(X)$ es divisible por $X - c$. Puede suceder que $P(X)$ sea no sólo divisible por $X - c$ sino por una potencia superior $(X - c)^k$. En cualquier caso siempre existe un número natural k tal que $P(X)$ es divisible por $(X - c)^k$ y no es divisible por $(X - c)^{k+1}$.

Definición. Si c es una raíz de un polinomio $P(X)$ se llama orden de multiplicidad de la raíz c al mayor número natural k tal que $P(X)$ es divisible por $(X - c)^k$ y no lo es por $(X - c)^{k+1}$.

Es decir

$P(X) = (X - c)^k . Q(X)$, donde $Q(X)$ no es divisible por $X - c$, o sea c no es raíz de $Q(X)$.

Las raíces de orden de multiplicidad 1 se dicen simples, las de orden 2 dobles, las

de orden 3 triples, etc. Las raíces de orden de multiplicidad > 1 reciben múltiples.

El orden de multiplicidad de una raíz se puede hallar aplicando la regla de Ruffini.

Por ejemplo, verificar que 2 es una raíz del polinomio

$P(X) = X^5 - 6X^4 + 11X^3 - 2X^2 - 12X + 8$ y hallar su orden de multiplicidad. Se calcula el cociente $Q_1(X)$ de dividir $P(X)$ por $X - 2$; después se divide $Q_1(X)$ por $X - 2$; si el resto de esta división es cero, se divide el cociente $Q_2(X)$ por $X - 2$; si el resto es cero, se divide el cociente $Q_3(X)$ por $X - 2$; y así siguiendo hasta obtener un resto R_{n+1} distinto de cero. Esto significará que Q_n no es divisible por $X - 2$. Entonces de

$$P(X) = (X - 2).Q_1(X)$$
$$Q_1(X) = (X - 2).Q_2(X)$$
$$Q_2(X) = (X - 2).Q_3(X)$$
$$\vdots \qquad \vdots$$
$$Q_{n-1}(X) = (X - 2).Q_n(X)$$

resulta

$P(X) = (X - 2)^n.Q_n(X)$, donde $Q_n(X)$ no es divisible por $X-2$.

Luego el orden de multiplicidad de la raíz 2 es n.

Haciendo los cálculos:

	1	-6	11	-2	-12	8
2		2	-8	6	8	-8
	1	-4	3	4	-4	| 0
2		2	-4	-2	4	
	1	-2	-1	2	| 0	
2		2	0	-2		
	1	0	-1	| 0		
2		2	4			
	1	2	| 3			

Luego

$$X^5 - 6X^4 + 11X^3 - 2X^2 - 12X + 8 = (X - 2)^3 . (X^2 - 1)$$

y 2 es una raíz triple del polinomio dado.

Pero hay otra forma de hallar la multiplicidad de una raíz de un polinomio, utilizando la noción de polinomio derivado.

Definición. Dado $P(X) = a_nX^n + \ldots\ldots + a_1X + a_o \in K[X]$ se llama derivado de $P(X)$ al polinomio

$$P'(X) = na_nX^{n-1} + (n-1)a_{n-1}X^{n-2} + \ldots\ldots + 2a_2X + a_1$$

Por ejemplo, si $P(X) = 3X^7 - \frac{2}{5}X^5 + X^3 - \sqrt{3}X^2 - X + 6$

es

$$P'(X) = 21X^6 - 2X^4 + 3X^2 - 2\sqrt{3}X - 1$$

PROPIEDADES.

1. $A' = 0 \iff A = 0$ ó $gr\ A = 0$
2. $(A+B)' = A' + B'$
3. $(A.B)' = A'.B + A.B'$

   cualesquiera sean $A,B \in K[X]$

Las propiedades 1 y 2 resultan fácilmente de la definición y la 3 se verifica sin dificultades. La demostración queda propuesta como ejercicio para el lector.

De estas propiedades se deducen las siguientes:

4. $(k.A)' = k.A' \quad , \forall\ k \in K \ , \ A \in K[X]$
5. $(A^n)' = n.A^{n-1}.A' \quad , \forall\ A \in K[X] \ , \ n \in N$

Dado un polinomio $P(X) \in K[X]$, se definen por recurrencia los polinomios $P^{(s)}(X)$, $s \in N$, como sigue:

$$P^{(1)}(X) = P'(X)$$

$$P^{(s+1)}(X) = (P^{(s)}(X))'$$

El polinomio $P^{(s)}(X)$ se llama el s-ésimo derivado de $P(X)$.

Por ejemplo, dado $P(X) = 5X^3 - X^2 + 4X + 7$ es:

$P'(X) = 15X^2 - 2X + 4$

$P''(X) = 30X - 2$

$P'''(X) = 30$

$P''''(X) = 0$

Vamos a convenir en decir que c es una raíz de multiplicidad cero de un polinomio $P(X)$ si y sólo si c no es raíz de $P(X)$.

TEOREMA 4.7. Si c es una raíz de un polinomio $P(X)$, c es una raíz k-múltiple de $P(X)$ si y sólo si c es una raíz k-1 - múltiple de $P'(X)$.

Demostración: Sea c una raíz k-múltiple de $P(X)$, $k \geq 1$.

Entonces

$$P(X) = (X - c)^k.Q(X) \quad , \quad \text{donde} \quad X - c \not| Q(X) \ .$$

Luego

$$P'(X) = k(X - c)^{k-1}.Q(X) + (X - c)^k.Q'(X)$$

Como $X - c \not| Q(X)$, $k-1$ es el mayor exponente tal que $(X - c)^{k-1}$ divide a $P'(X)$. Luego c es una raíz $k-1$- múltiple de $P'(X)$.

Recíprocamente, supongamos que c es una raíz de $P(X)$ tal que c es raíz $k-1$- múltiple de $P'(X)$.

Sea h el mayor exponente tal que $(X - c)^h / P(X)$. Luego $h \geq 1$ y

$$P(X) = (X - c)^h.Q(X) \quad , \quad \text{donde} \quad X - c \not| Q(X)$$

Entonces

$$P'(X) = h(X - c)^{h-1}.Q(X) + (X - c)^h \ .Q'(X)$$

Como $X - c \not| Q(X)$, $h-1$ es el mayor exponente t tal que $(X - c)^t / P'(X)$. Por nuestra hipótesis debe ser entonces $h-1 = k-1$, o sea $h = k$, lo que termina la demostración.

TEOREMA 4.8. Un elemento $c \in K$ es raíz k-múltiple de un polinomio $P(X) \in K[X]$, $k > 0$, si y sólo si

$$P(c) = P'(c) = \ldots\ldots = P^{(k-1)}(c) = 0 \quad y \quad P^{(k)}(c) \neq 0 \ .$$

Demostración: Por el teorema anterior, $c \in K$ es una raíz k-múltiple de $P(X)$, $k > 0$, si y sólo si $P(c) = 0$ y c es una raíz $k-1$-múltiple de $P'(X)$.

Aplicando reiteradamente este teorema a los polinomios

$$P(X) \ , \ P'(X) \ , \ P''(X) \ , \ \ldots\ldots \ , \ P^{(k-1)}(X) \quad \text{se tiene:}$$

c es raíz k-múltiple de $P(X)$, $k > 0 \iff P(c) = 0$ y c es raíz $k-1$- múltiple de $P'(X) \iff P(c) = P'(c) = 0$ y c es raíz $k-2$- múltiple de $P''(X) \iff P(c) = P'(c) = P''(c) = 0$ y c es raíz $k-3$- múltiple de $P'''(X) \iff \ldots\ldots \iff P(c) = P'(c) = P''(c) = \ldots\ldots = P^{(k-1)}(c) = 0$ y $P^{(k)}(c) \neq 0$.

De este teorema resulta que el orden de multiplicidad de una raíz c de un polinomio $P(X)$ es igual al orden de la derivada de menor orden de $P(X)$ que no se anula en c.

Por ejemplo, calculemos aplicando este resultado el orden de multiplicidad de la raíz 2 del polinomio $P(X) = X^5 - 6X^4 + 11X^3 - 2X^2 - 12X + 8$, ya calculado usando la regla de Ruffini.

$P'(X) = 5X^4 - 24X^3 + 33X^2 - 4X - 12$; $P'(2) = 0$

$P''(X) = 20X^3 - 72X^2 + 66X - 4$; $P''(2) = 0$

$P'''(X) = 60X^2 - 144X + 66$; $P'''(2) \neq 0$

Luego 2 es una raíz de tercer orden.

EJERCICIO.

Dado un polinomio $P(X) \in K[X]$, demostrar que:

a) $(P(X),P'(X)) = 1 \Longrightarrow P(X)$ no tiene raíces múltiples.

b) $P(X)$ irreducible $\Longrightarrow P(X)$ no tiene raíces múltiples.

Contando cada raíz de un polinomio $P(X)$ tantas veces como su orden de multiplicidad se tiene el siguiente

TEOREMA 4.9. Un polinomio $P(X) \in K[X]$ de grado $n > 0$ tiene a lo sumo n raíces en K.

Demostración: Por inducción sobre el grado del polinomio. Sea $P(X) \in K[X]$,gr $P>0$. Si gr $P = 1$, $P(X)$ es de la forma $P(X) = a_1X + a_o$ y su única raíz es $-\frac{a_o}{a_1}$. Luego el teorema se verifica.

Sea gr $P = n > 1$ y supongamos el teorema verdadero para todos los polinomios de $K[X]$ de grado $n-1$.

Si $P(X)$ no tiene ninguna raíz en K no hay nada que demostrar. Si $P(X)$ tiene una raíz $c \in K$ entonces

$$P(X) = (X - c).Q(X) \qquad (1)$$

con $Q(X) \in K[X]$ y gr $Q = n-1$.

De (1) resulta que toda raíz de $Q(X)$ es raíz de $P(X)$ y recíprocamente, que toda raíz de $P(X)$ es c ó una raíz de $Q(X)$ ya que $P(t) = 0$ implica $(t - c).Q(t) = 0$, y como en $K[X]$ no hay divisores de cero, debe ser $t = c$ ó $Q(t) = 0$.

Luego las raíces de $P(X)$ son c y las raíces de $Q(X)$. Por la hipótesis de inducción $Q(X)$ tiene a lo sumo $n-1$ raíces en K de modo que $P(X)$ tiene a lo sumo n raíces en K.

El teorema queda así demostrado.

EJEMPLOS.

1) Vimos que el polinomio $P(X) = X^5 - 6X^4 + 11X^3 - 2X^2 - 12X + 8 \in R[X]$ tiene una raíz triple igual a 2:

$$X^5 - 6X^4 + 11X^3 - 2X^2 - 12X + 8 = (X - 2)^3.(X^2 - 1)$$

Como las raíces de $X^2 - 1$ son 1 y -1 , las raíces del polinomio dado en R son 2 , 2 , 2 , 1 y -1 y $P(X)$ tiene 5 raíces en R.

2) El polinomio $X^2 + 1$ no tiene ninguna raíz en Q y tampoco tiene ninguna raíz en R. En cambio $X^2 + 1$ tiene 2 raíces en C: i y -i.

3) Sea $P(X) = X^6 - X^4 - 6X^2$. En $Q[X]$ se factoriza así:

$$P(X) = X^2(X^2 - 3)(X^2 + 2)$$

Entonces P(X) tiene en Q solamente una raíz doble: 0 .

Tiene 4 raíces en R: 0 , 0 , $\sqrt{3}$, $-\sqrt{3}$. Y tiene 6 raíces en C: 0 , 0 , $\sqrt{3}$, $-\sqrt{3}$, $\sqrt{2}\,i$, $-\sqrt{2}\,i$.

4.4. EXISTENCIA DE RAICES DE UN POLINOMIO.

Se plantea el problema de la existencia de raíces de un polinomio.

Hemos visto que hay polinomios en Q[X] y en R[X] que no tienen ninguna raíz en Q o en R. Podría pensarse que igualmente existen polinomios no constantes en C[X] que no tienen ninguna raíz en C. Pero no es así porque C tiene la siguiente propiedad:

TEOREMA 4.10. Todo polinomio no constante con coeficientes en C tiene por lo menos una raíz en C.

Este teorema, que se conoce con el nombre de Teorema fundamental del álgebra, fue formulado por primera vez por D'Alembert (1717-1783) pero su demostración, basada en resultados de la teoría de funciones analíticas, no era completa. La primera demostración correcta la dió Gauss (1777-1855) a principios del siglo pasado.

Se conocen diferentes demostraciones de este teorema pero ninguna es elemental y todas escapan al nivel de este curso. El lector interesado puede encontrar una en cualquier libro sobre funciones de variable compleja.

El teorema anterior equivale a decir que todo polinomio de grado n > 0 con coeficientes en C tiene exactamente n raíces en C. (contando las raíces múltiples tantas veces como su orden de multiplicidad).

En efecto, si $P(X) = a_n X^n + a_{n-1} X^{n-1} + \ldots\ldots + a_1 X + a_o$ es un polinomio no constante con coeficientes reales o complejos, por el teorema P(X) tiene una raíz $c_1 \in C$.

$$P(X) = (X - c_1).Q_1(X) \quad \text{donde} \quad gr\ Q_1(X) = n-1$$

Si $gr\ Q_1(X) > 0$, por el teorema $Q_1(X)$ tiene una raíz $c_2 \in C$.

$$Q_1(X) = (X - c_2).Q_2(X) \quad \text{donde} \quad gr\ Q_2(X) = n-2$$

Luego $$P(X) = (X - c_1).(X - c_2).Q_2(X)$$

Reiterando el procedimiento, al cabo de n pasos se tiene:

$$P(X) = (X - c_1).(X - c_2).\ldots\ldots.(X - c_n)\ .\ k$$

con $k = a_n$, es decir

$$P(X) = a_n(X - c_1).(X - c_2).\ldots\ldots.(X - c_n)$$

y P(X) tiene n raíces.

Resulta entonces que todo polinomio P(X) de grado n > 0 es igual al producto de su coeficiente principal por polinomios $X - c_1$, $X - c_2$, ... , $X - c_n$ donde c_1 , c_2 , ... , c_n son las raíces de P(X) en C. Como todo polinomio de primer grado es irreducible, ésta es la descomposición de P(X) en factores irreducibles en C[X]. De aquí se deduce que

Los únicos polinomios irreducibles de C[X] son los de primer grado.

Por ejemplo, ya vimos que las raíces del polinomio $P(X) = X^6 - X^4 + 6X^2$ son: 0 , 0 , $\sqrt{3}$, $-\sqrt{3}$, $\sqrt{2}\,i$, $-\sqrt{2}\,i$. Entonces

$$X^6 - X^4 + 6X^2 = X^2(X - \sqrt{3})(X + \sqrt{3})(X - \sqrt{2}\,i)(X + \sqrt{2}\,i)$$

OBSERVACION. Se ve que hay una estrecha relación entre las raíces de un polinomio P(X) de K[X] y la descomposición de P(X) en factores irreducibles en K[X], puesto que $c \in K$ es una raíz de P(X) si y sólo si P(X) es divisible por X - c, y éste es un polinomio irreducible de K[X]. En general entonces, un polinomio no constante $P(X) \in K[X]$ tiene tantas raíces en K como factores de primer grado aparecen en su de descomposición en factores irreducibles en K[X].

Raíces complejas de polinomios con coeficientes reales.

Vamos a ver que si $P(X) = a_nX^n + a_{n-1}X^{n-1} + \ldots\ldots + a_1X + a_o$ es un polinomio con coeficientes reales entonces un número complejo z es raíz de P(X) si y sólo si su conjugado $\bar{z}$ es raíz de P(X); y ambas raíces tienen el mismo orden de multiplicidad , es decir, las raíces complejas de P(X) se presentan de a pares conjugados.

Para ello probemos que $P(\bar{z}) = \overline{P(z)}$ cualquiera sea $z \in C$.

$$P(\bar{z}) = a_n\bar{z}^n + a_{n-1}\bar{z}^{n-1} + \ldots\ldots + a_1\bar{z} + a_o = \overline{a_nz^n} + \overline{a_{n-1}z^{n-1}} + \ldots\ldots + \overline{a_1z} + \overline{a_o} =$$

$$= \overline{a_nz^n + a_{n-1}z^{n-1} + \ldots\ldots + a_1z + a_o} = \overline{P(z)}$$

Entonces

$$P(z) = 0 \iff P(\bar{z}) = 0$$

ya que $\bar{0} = 0$.

Luego $z = a+bi$ es raíz de P(X) si y sólo si $\bar{z} = a-bi$ también lo es. Veamos que tienen el mismo orden de multiplicidad. P(X) es divisible por $(X - z)$ y por $(X - \bar{z})$, es decir por $(X - z).(X - \bar{z}) = (X-a-bi).(X-a+bi) = X^2 - 2aX + (a^2 + b^2) \in R[X]$.

Supongamos que el orden de multiplicidad de z es k, el de $\bar{z}$ es ℓ y que $k > \ell$. Entonces P(X) es divisible por $(X-z)^\ell$ y $(X-\bar{z})^\ell$, es decir por

$$[(X - z).(X - \bar{z})]^{\ell} = [X^2 - 2aX + (a^2 + b^2)]^{\ell} \in R[X]$$

Luego

$$P(X) = [X^2 - 2aX + (a^2 + b^2)]^{\ell}.Q(X)$$

Ahora bien, Q(X) tiene coeficientes reales, z es raíz de Q(X) de orden $k-\ell > 0$ y $\bar{z}$ no es raíz de Q(X), lo que es imposible. Esta contradicción prueba que $k = \ell$.

Queda demostrado así el siguiente

TEOREMA 4.11. Si z es una raíz compleja de un polinomio P(X) con coeficientes reales entonces $\bar{z}$ también es raíz de P(X) y las raíces z y $\bar{z}$ tienen el mismo orden de multiplicidad.

De la demostración anterior resulta que si $z = a+bi$ es una raíz compleja de un polinomio P(X) con coeficientes reales entonces P(X) es divisible por el polinomio $(X - z).(X - \bar{z}) = X^2 - 2aX + (a^2 + b^2)$, es decir por un polinomio con coeficientes reales de la forma $X^2 + pX + q$ con $p^2 - 4q < 0$, puesto que $(-2a)^2 - 4(a^2+b^2) = -4b^2 < 0$.

Entonces, cualquiera sea $P(X) \in R[X]$ de grado $n > 0$, si $c_1, c_2, \ldots, c_s$ son sus raíces reales y $z_1, \bar{z}_1, z_2, \bar{z}_2, \ldots, z_r, \bar{z}_r$ son sus raíces complejas ordenadas de a pares conjugados, figurando cada raíz tantas veces como su orden de multiplicidad (es decir, $s+2r = n$), P(X) queda descompuesto en un producto en R[X] como sigue:

$$P(X) = a_n(X-c_1)(X-c_2)\ldots.(X-c_s)(X^2+p_1X+q_1)(X^2+p_2X+q_2)\ldots.(X^2+p_rX+q_r)$$

donde a_n es el coeficiente principal de P(X) y todos los factores son irreducibles en R[X].

De aquí resulta inmediatamente que

COROLARIO 1. Los polinomios irreducibles de R[X] son los de primer grado y los de segundo grado de la forma $a(X^2 + pX + q)$ con $p^2 - 4q < 0$.

COROLARIO 2. Todo polinomio con coeficientes reales de grado impar tiene por lo menos una raíz real.

La propiedad que las raíces complejas de un polinomio con coeficientes reales se presentan de a pares conjugados no se verifica para los polinomios con coeficientes complejos. Por ejemplo las raíces del polinomio

$X^2 - (2+i)X + 2i$ son i y 2.

EJERCICIOS.

1. Hallar todas las raíces del polinomio $X^6 + 6X^5 + 9X^4 - X^2 - 6X - 9$ sabiendo que -3 es una raíz múltiple y expresarlo como producto de polinomios irreducibles en $C[X]$, $R[X]$ y $Q[X]$ sucesivamente.

   Calculemos el orden de multiplicidad de -3.

	1	6	9	0	-1	-6	-9
-3		-3	-9	0	0	3	9
	1	3	0	0	-1	-3	0
-3		-3	0	0	0	3	
	1	0	0	0	-1	0	
-3		-3	9	-27	81		
	1	-3	9	-27	80		

   Luego

$$X^6 + 6X^5 + 9X^4 - X^2 - 6X - 9 = (X + 3)^2 (X^4 - 1)$$

   y -3 es una raíz doble. Las cuatro raíces restantes son las raíces de $X^4 - 1$, es decir las raíces cuartas de la unidad: 1 , -1 , i , -i .

   Luego las seis raíces del polinomio dado son: -3 , -3 , 1 , -1 , i , -i.

   La factorización en polinomios irreducibles en $C[X]$ es :

$$X^6 + 6X^5 + 9X^4 - X^2 - 6X - 9 = (X + 3)^2 (X - 1) (X + 1) (X - i) (X + i)$$

   En $R[X]$ es :

$$X^6 + 6X^5 + 9X^4 - X^2 - 6X - 9 = (X + 3)^2 (X - 1) (X + 1) (X^2 + 1)$$

   y en $Q[X]$ coincide con la de $R[X]$.

2. Como en el ejercicio 1 , para el polinomio $X^4 - 2X^3 - 3X^2 + 10X - 10$ sabiendo 1+i es una raíz del mismo.

   Dado que los coeficientes del polinomio son reales , 1+i raíz del mismo implica que 1-i también es raíz. Luego el polinomio es divisible por $X - (1+i)$ y $X - (1-i)$, es decir por:

$$[X - (1+i)].[X - (1-i)] = X^2 - 2X + 2$$

   Entonces, dividiendo el polinomio dado por $X^2 - 2X + 2$ se tiene:

$$X^4 - 2X^3 - 3X^2 + 10X - 10 = (X^2 - 5)(X^2 - 2X + 2)$$

y las dos raíces restantes del polinomio dado son las raíces de $X^2 - 5$: $\sqrt{5}$ y $-\sqrt{5}$.

Luego las raíces son: $1+i$, $1-i$, $\sqrt{5}$, $-\sqrt{5}$.

La factorización en polinomios irreducibles en $C[X]$ es :

$$X^4 - 2X^3 - 3X^2 + 10X - 10 = (X - \sqrt{5})(X + \sqrt{5})(X - (1+i))(X - (1-i))$$

En $R[X]$ es :

$$X^4 - 2X^3 - 3X^2 + 10X - 10 = (X - \sqrt{5})(X + \sqrt{5})(X^2 - 2X + 2)$$

En $Q[X]$ es :

$$X^4 - 2X^3 - 3X^2 + 10X - 10 = (X^2 - 5)(X^2 - 2X + 2)$$

(Note el lector que este polinomio no tiene ninguna raíz en Q y es reducible en $Q[X]$) .

Relaciones entre las raíces de un polinomio y sus coeficientes.

Sea $P(X) = a_n X^n + a_{n-1} X^{n-1} + \ldots\ldots + a_1 X + a_o$ y $c_1, c_2, \ldots\ldots, c_n$ las n raíces de P(X).

La descomposición factorial de P(X) en $C[X]$ es :

$$P(X) = a_n(X - c_1)(X - c_2)\ldots\ldots(X - c_n)$$

Efectuando las multiplicaciones indicadas en el segundo miembro y sumando los términos semejantes se tiene:

$$P(X) = a_n X^n - a_n(c_1+c_2+\ldots\ldots+c_n)X^{n-1} + a_n(c_1c_2+c_1c_3+\ldots+c_1c_n+c_2c_3+\ldots+c_{n-1}c_n)X^{n-2} -$$

$$- a_n(c_1c_2c_3+c_1c_2c_4+\ldots+c_{n-2}c_{n-1}c_n)X^{n-3} + \ldots\ldots + (-1)^n a_n c_1 c_2 \ldots c_n$$

Comparando los coeficientes obtenidos en el 2° miembro con los de P(X) resultan las siguientes igualdades:

$$-\frac{a_{n-1}}{a_n} = c_1 + c_2 + \ldots + c_n$$

$$\frac{a_{n-2}}{a_n} = c_1c_2 + c_1c_3 + \ldots + c_{n-1}c_n$$

$$-\frac{a_{n-3}}{a_n} = c_1c_2c_3 + c_1c_2c_4 + \ldots + c_{n-2}c_{n-1}c_n$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$(-1)^n \frac{a_o}{a_n} = c_1c_2\ldots\ldots\ldots c_n$$

Es decir, la suma de las n raíces da el segundo coeficiente de P(X) dividido el primero con el signo cambiado; la suma de todos los productos de las raíces tomadas de a dos da el tercer coeficiente dividido el primero; en general, la suma de todos los productos posibles de h raíces da

$$(-1)^h \frac{a_{n-h}}{a_n} \quad , \quad \text{para} \quad h = 1,2,\ldots,n$$

En particular, si n = 2 entonces se obtienen las conocidas relaciones entre raíces y coeficientes del polinomio de 2° grado: $P(X) = a_2X^2 + a_1X + a_o$

$$c_1 + c_2 = -\frac{a_1}{a_2} \quad ; \quad c_1 \cdot c_2 = \frac{a_o}{a_2}$$

Mediante estas fórmulas es posible escribir un polinomio conocidas sus raíces.

EJERCICIOS.

1. Hallar un polinomio de grado mínimo con coeficiente principal 2 y cuyas raíces sean $0 , \frac{1}{2} , -1 , i$.

   El grado del polinomio debe ser 4: $a_4X^4 + a_3X^3 + a_2X^2 + a_1X + a_o$.

   Aplicando las relaciones recién demostradas se tiene:

   $$0 + \frac{1}{2} - 1 + i = -\frac{a_3}{a_4} \Longrightarrow -\frac{1}{2} + i = -\frac{a_3}{2} \Longrightarrow a_3 = 1 - 2i$$

   $$0.\frac{1}{2} + 0.(-1) + 0.i + \frac{1}{2}.(-1) + \frac{1}{2}.i + (-1).i = \frac{a_2}{a_4} \Longrightarrow -\frac{1}{2} - \frac{1}{2}i = \frac{a_2}{2} \Longrightarrow$$

   $$\Longrightarrow a_2 = -1-i .$$

   $$0.\frac{1}{2}.(-1) + 0.\frac{1}{2}.i + \frac{1}{2}.(-1).i = -\frac{a_1}{a_4} \Longrightarrow -\frac{1}{2}i = -\frac{a_1}{2} \Longrightarrow a_1 = i$$

   $$0.\frac{1}{2}.(-1).i = \frac{a_o}{a_4} \Longrightarrow a_o = 0$$

   Luego el polinomio es

   $$2X^4 + (1-2i)X^3 + (-1-i)X^2 + iX$$

   Otra forma de calcular el polinomio sería escribir

   $$P(X) = 2(X - \tfrac{1}{2})(X + 1)(X - i)$$

   y efectuar las operaciones indicadas.

2. Hallar un polinomio de grado mínimo de R[X] que llene las condiciones del ejerci

cio 1.

Como las raíces complejas de un polinomio con coeficientes reales se presentan de a pares , si i es una raíz del polinomio buscado, -i también debe ser raíz. Luego se trata de buscar un polinomio cuyas raíces sean 0 , $\frac{1}{2}$, -1 , i , $-i$ con coeficiente principal 2.

El polinomio es de quinto grado y procediendo como antes se encuentra que

$$-\frac{a_4}{a_5} = -\frac{1}{2} \Longrightarrow a_4 = 1$$

$$\frac{a_3}{a_5} = \frac{1}{2} \Longrightarrow a_3 = 1$$

$$-\frac{a_2}{a_5} = -\frac{1}{2} \Longrightarrow a_2 = 1$$

$$\frac{a_1}{a_5} = -\frac{1}{2} \Longrightarrow a_1 = -1$$

$$-\frac{a_o}{a_5} = 0 \Longrightarrow a_o = 0$$

El polinomio buscado es

$$2X^5 + X^4 + X^3 + X^2 - X$$

Puede calcularse también aplicando el teorema de factorización de polinomios:

$$P(X) = 2(X - \tfrac{1}{2})(X + 1)(X - i)(X + i)$$

y aprovechando el resultado del ejercicio 1, multiplicar directamente el polińomio allí obtenido por el factor $(X + i)$.

1. Dados en $R[X]$ los polinomios $A = 3X^3 - \frac{1}{2}$, $B = 6X^7 - 2X^6 - X^4 - X$,

$C = X^4 - X^3 + 3X - 5$ hallar :

a) $3A - B$, $B - 4C$, $6A - CA - \frac{1}{2}B + A^2$

b) El grado de CB^2 y el de $(A+B)^{23}$; el coeficiente principal de $AC - B$; el término independiente de $BC - AC + B^2$; el coeficiente de X^7 en BC.

c) El cociente y el resto de dividir B por A ; B por C ; A por C.

d) Idem para los polinomios de $C[X]$: $A = X^2 - i$, $B = (1+i)X^3 - 2iX + 1$, $C = X^4 + 3i$.

2. Hallar el m.c.d. (mónico) de los siguientes polinomios $\in R[X]$:

a) $X^3 - 1$ y $X^4 + X^3 + 2X^2 + X + 1$

b) $X^4 - 3X^2 - 2X$ y $X^2 - 4X + 4$

c) $X^3 - 5X + 6$ y $3X^4 + 9X^3 + 3X^2 - 9X - 6$

d) $16X^2 - 1$, $X - 4X^2$ y $16X^2 - 8X + 1$

y expresarlo en la forma $(A,B) = R.A + S.B$ en los casos a) y b). Indicar un método para hallar el m.c.m.

3. a) Aplicando la regla de Ruffini hallar el cociente y el resto de dividir:

$X^5 - 7X^4 + 2X^2 - X + 2$ por $X - 2$

$X^7 - 3X^3 + 2X^2 - 1$ " $X + 1$

$2X^4 - 5X^3 - iX^2 + \frac{1}{2}X - (3+i)$ por $X + i$

b) Dado $P(X) = 2X^5 - X^4 + 3X^3 - \frac{1}{3}$ hallar $P(-2)$

" $P(X) = 3X^7 - X^6 - 4X^4 + X^2 - 5$ hallar $P(\frac{1}{3})$

" $P(X) = -3X^3 + 6X^2 - (1-i)X + 1$ " $P(1-i)$

c) Decir en cada caso si los números indicados son raíces del polinomio respectivo y hallar su orden de multiplicidad:

$X^5 + 6X^4 + 11X^3 + 2X^2 - 12X - 8$, $a = -2$

$X^7 - 6X^4 - X^3 + 6$, $a = 3$, $b = i$

$X^5 + 7X^4 + 19X^3 + 25X^2 + 16X + 4$, $a = -1$, $b = -2$

$X^6 - 14X^5 + 89X^4 - 296X^3 + 496X^2 - 384X + 144$, $a = 2$, $b = 3$.

4. Aplicando el teorema del resto hallar las condiciones para que $X^n \pm a^n$ sea divisible por $X \pm a$ $(a \in R)$.

5. Determinar en cada caso los valores de los coeficientes $a,b \in Q$ de tal modo que:

   a) $X^5 + 3(X^2 - b)^2 - 6aX^2 + b = X^5 + 3X^4 + 2$

   b) $3X^3 + aX - 1$ sea divisible por $3X^2 + bX + 2$

   c) El resto de dividir $X^4 - a^2X + 3 - a$ por $X - 1$ sea 4.

   d) $X^2 + aX + 4$ dé el mismo resto al dividirlo por $X + 2$ y $X - 2$.

   e) $X^3 - aX^2 + 2X - 2a$ sea divisible por $X^2 + 2$.

   f) $X^4 + aX^3 - 2X^2 - 1$ sea divisible por $2X^2 + 1$.

6. Calcular las raíces de los siguientes polinomios:

   a) $2X^2 - X - 3$

   b) $X^2 - 2\sqrt{2}X + 3$

   c) $X^2 + (5+2i)X + (5+5i)$

   d) $iX^2 - X + i$

   e) $3X^4 - X^2 - 2$

   f) $X^4 + 2X^2 + 4$

7. a) Hallar todas las raíces del polinomio $X^5 + 6X^4 + 15X^3 + 26X^2 + 36X + 24$ sabiendo que -2 es una raíz múltiple.

   Idem para $8X^4 - 4X^3 - 10X^2 + 9X - 2$ y $\frac{1}{2}$

   " " $2X^5 + X^4 + X^2$ y -1

   b) Hallar las raíces de $2X^4 - X^3 - 17X^2 + 15X + 9$ sabiendo que $1+\sqrt{2}$ y $1-\sqrt{2}$ son raíces.

   c) Hallar las raíces de $X^3 - 2(1+i)X^2 - (1-2i)X + 2(1+2i)$ sabiendo que $1+2i$ es una raíz.

   d) Hallar las raíces de $X^4 - 3X^3 + 5X^2 - 27X - 36$ sabiendo que tiene una raíz imaginaria pura.

8. a) Demostrar que si $z = a+bi$ es raíz de un polinomio $P(X)$ con coeficientes reales entonces $\bar{z} = a-bi$ también es raíz del mismo. Concluir que todo polinomio con coeficientes reales de grado impar tiene por lo menos una raíz real.

   b) Demostrar que si un polinomio $P(X)$ con coeficientes racionales tiene una

raíz de la forma $a+\sqrt{b}$, con $a,b \in Q$ y $\sqrt{b}$ irracional entonces $a-\sqrt{b}$ también es raíz de $P(X)$.

(Sugerencia: Dividir el polinomio $P(X)$ por $[x-(a+\sqrt{b})].[x-(a-\sqrt{b})]$ y probar que el resto es 0).

c) Sabiendo que un polinomio $P(X)$ tiene raíces $2i$, $3-\sqrt{2}i$, $1+0.5\,i$ y que $P(X) \in R[X]$ indicar otras tantas raíces del mismo.

d) ¿Son válidas las siguientes afirmaciones?:

i) $X^3 + 7X - 6i$ tiene i como raíz; entonces -i es otra raíz.

ii) $X^3 + (1-2\sqrt{3})X^2 + (5-2\sqrt{3})X + 5$ tiene a $\sqrt{3} - \sqrt{2}\,i$ como raíz; entonces $\sqrt{3} + \sqrt{2}\,i$ es otra raíz.

iii) $X^4 + (1-2\sqrt{2})X^3 + (4-2\sqrt{2})X^2 + (3-4\sqrt{2})X + 1$ tiene $-1+\sqrt{2}$ como raíz; entonces $-1-\sqrt{2}$ es otra raíz.

9. a) Encontrar polinomios $\in Q[X]$ de grado mínimo que tengan las siguientes raíces:

i) -1 , 3 , $-\frac{1}{2}$

ii) 0 , $\frac{3}{4}$, $\frac{2}{3}$, -1 raíz doble.

iii) $\pm 3i$, $\pm\sqrt{2}$

iv) i , $1-i$, -2

v) 1 , $-\sqrt{3}$

vi) $-\frac{1}{2}$, $\frac{\sqrt{2}}{3}$, $1+\sqrt{2}\,i$

b) Hallar la suma y el producto de las raíces de los siguientes polinomios, sin calcularlas:

i) $X^2 + 3X - 1$

ii) $3X^3 - X^2 + 5X - 2$

iii) $7X^5 + X^3 - \sqrt{3}\,X^2$

iv) $2X^6 + (1-i)X^5 - 2X^3 + \sqrt{2}$

v) Si las raíces de la ecuación $2X^3 + 3X^2 + 4X + 2 = 0$ son a,b,c hallar: $a+b+c$, $ab+ac+bc$, abc , $\frac{1}{a}+\frac{1}{b}+\frac{1}{c}$ $\frac{1}{ab}+\frac{1}{bc}+\frac{1}{ca}$

vi) Dar una nueva demostración de la propiedad ya vista: "La suma de las n raíces n-ésimas de 1 es 0 y su producto es -1 ó +1 según que n sea par o impar".

10. Dada la ecuación $X^3 - 9X + c = 0$ hallar el valor de c en cada uno de los siguientes casos:

a) Una raíz es igual a otra cambiada de signo.

b) Hay una raíz doble.

c) Las tres raíces están en progresión geométrica.

d) " " " " " " aritmética.

e) Una raíz es $\sqrt[3]{9} + \sqrt[3]{3}$.

11. a) Si $w = \cos \frac{2\pi}{n} + i \operatorname{sen} \frac{2\pi}{n}$ demostrar que

$$X^{n-1} + X^{n-2} + \ldots\ldots + X + 1 = (X - w)(X - w^2)\ldots\ldots(X - w^{n-1}) .$$

b) Si ε es una raíz primitiva de 1 de orden n demostrar que

$$n = (1-\varepsilon)(1-\varepsilon^2)\ldots\ldots(1-\varepsilon^{n-1}) .$$

c) Hallar todas las raíces del polinomio $X^n + X^{n-1} + \ldots\ldots + X + 1$.

12. Descomponer los siguientes polinomios en producto de polinomios irreducibles en $Q[X]$, $R[X]$ y $C[X]$ sucesivamente:

a) $X^2 - 1$

b) $X^2 + 1$

c) $X^2 - 4X + 2$

d) $X^3 - 3X$

e) $X^4 - 4$

f) $3X^4 + 4X^2$

g) Los polinomios del ejercicio 7 , puntos a) , b) y d).

13. a) Sea $P(X) \in Q[X]$ un polinomio de 3^{er} grado. Probar que P es reducible en $Q[X]$ si y sólo si P posee una raíz en Q. Demostrar que todo polinomio de $R[X]$ de 3^{er} grado es reducible. Dar ejemplos de polinomios de 3^{er} grado de $Q[X]$ irreducibles.

Probar que el polinomio $X^4 + 2$ es irreducible en $Q[X]$.

b) La siguiente afirmación es falsa: "Si $P(X) \in Q[X]$ no posee ninguna raíz en Q entonces es irreducible". Dar un contraejemplo.

c) Analizar la validez de la siguiente proposición: "Si $P(X) \in R[X]$ no tiene ninguna raíz en R entonces es irreducible en $R[X]$ " .

d) Demostrar que los polinomios de $R[X]$ irreducibles en $R[X]$ son de 1° y 2° grado y que los polinomios irreducibles de $C[X]$ son los de 1^{er} grado.

4.5. CALCULO DE LAS RAICES DE UN POLINOMIO.

En el teorema fundamental del álgebra se demuestra que todo polinomio

$$P(X) = a_n X^n + \ldots\ldots + a_1 X + a_o$$

con coeficientes reales o complejos de grado $n > 0$ tiene n raíces en C, contando cada raíz tantas veces como su orden de multiplicidad. Pero se trata de una demostración de existencia pura que no proporciona ningún método para el cálculo efectivo de las raíces. Se escribe

$$a_n X^n + \ldots\ldots + a_1 X + a_o = 0$$

para indicar que se plantea el problema de calcular las raíces del polinomio que figura en el primer miembro. Ya dijimos que una expresión de este tipo se llama una ecuación algebraica de grado n con una incógnita.

El cálculo de las raíces de los polinomios fue el problema central del álgebra durante siglos. Hemos visto que es muy fácil calcular los coeficientes de un polinomio si se conocen sus raíces. En cambio, calcular las raíces a partir de los coeficientes es mucho más difícil y decididamente imposible en el caso de algunos polinomios de grado mayor que el cuarto, usando las operaciones de suma, multiplicación, diferencia, división y extracción de raíces.

Es claro que dado un polinomio de primer grado, $aX + b$, es inmediato calcular su única raíz $X = -\frac{b}{a}$; y desde la antigüedad se conocía una fórmula para calcular las dos raíces de un polinomio de segundo grado, $aX^2 + bX + c$, en función de los coeficientes a,b y c mediante operaciones racionales y radicales de segundo grado, fórmula que el lector conoce bien del colegio secundario. Pero durante muchísimos años los matemáticos trataron infructuosamente de encontrar fórmulas que permitieran hallar las raíces de los polinomios de grado superior al segundo en función de sus coeficientes, análogas a la conocida para los polinomios de segundo grado.

En general, se dice que una ecuación algebraica $a_n X^n + a_{n-1} X^{n-1} + \ldots\ldots + a_1 X + a_o = 0$ puede resolverse por radicales si sus raíces se pueden expresar en función de los coeficientes $a_n, a_{n-1}, \ldots\ldots, a_1, a_o$ mediante operaciones racionales (suma, resta, multiplicación y división) y radicales. Recién en el siglo XVI, Scipione del Ferro (1465-1526) y Nicolo Tartaglia (1499-1551) encontraron la fórmula que permite calcular las raíces de la ecuación de tercer grado y poco después, siguiendo un método similar, Girolamo Cardano (1501-1576) y Ludovico Ferrari (1523-1565) hicieron lo mismo para las de cuarto grado. Después de esto se sucedieron los esfuerzos para resolver por medio de radicales las ecuaciones de grado superior al cuarto. En los intentos se obtuvieron, es claro, fórmulas que permiten calcular las raíces de ciertos polinomios de grado > 4 de tipos especiales, pero no la solución general. Pasaron casi tres siglos más hasta que este problema fue totalmente aclarado. A principios del siglo pasado, Paolo Ruffini (1765-1822), Niels Abel (1802-1829) y finalmente Evariste Galois (1811-1832) demostraron que es imposible resolver por radicales las e-

cuaciones de grado superior al cuarto, es decir, que existen polinomios de grado ≥ 5 cuyas raíces no se pueden expresar en función de los coeficientes por operaciones racionales y radicales (de cualquier grado) por más complicada que sea la expresión.

Ruffini y Abel demostraron que para los polinomios de grado superior al cuarto no existe ninguna fórmula general del tipo indicado que dé las raíces, pero el trabajo de Galois fue más profundo y completo porque dió una condición necesaria y suficiente para que cada ecuación algebraica particular sea resoluble por radicales. Esta condición se refiere a las propiedades de un cierto grupo de transformaciones que se asocia a la ecuación, y para cada grado $n \geq 5$ existen ecuaciones que no la verifican. Galois dejó explicadas sus ideas en una carta que escribió el 30 de mayo de 1832 pues al día siguiente debía batirse en duelo por motivos galantes. Galois murió en este duelo y sus ideas, desarrolladas, constituyen lo que actualmente se llama Teoría de Galois. Su estudio, como el lector ya sospechará, supera los límites de un primer curso de Algebra.

Vamos a repasar la fórmula que da las raíces de la ecuación de segundo grado y a estudiar las que permiten resolver las de tercer y cuarto grado, aunque es poca la utilidad práctica de la última de ellas. Luego discutiremos cómo calcular las raíces de polinomios de grado superior al cuarto.

Un buen compendio de este tema, que los alumnos de primer año pueden leer sin dificultad es el libro de Uspensky, Teoría de ecuaciones.

Ecuaciones de segundo grado.

Consideremos la ecuación cuadrática

$$aX^2 + bX + c = 0$$

con coeficientes numéricos cualesquiera, $a \neq 0$.

Es claro que esta ecuación tiene las mismas raíces que la que se obtiene dividiendo los coeficientes por a:

$$X^2 + \frac{b}{a} X + \frac{c}{a} = 0$$

Se puede escribir

$$X^2 + \frac{b}{a} X + \frac{b^2}{4a^2} + \frac{c}{a} - \frac{b^2}{4a^2} = 0$$

y de aquí se tiene

$$\left(X + \frac{b}{2a}\right)^2 + \left(\frac{c}{a} - \frac{b^2}{4a^2}\right) = 0$$

Luego

$$(X + \frac{b}{2a})^2 = \frac{b^2}{4a^2} - \frac{c}{a}$$

Como el número $\frac{b^2}{4a^2} - \frac{c}{a}$ tiene en C dos raíces cuadradas y éstas difieren solamente en el signo , escribiendo $\pm\sqrt{\frac{b^2}{4a^2} - \frac{c}{a}}$ para representarlas se tiene

$$X = -\frac{b}{a} \pm \sqrt{\frac{b^2}{4a^2} - \frac{c}{a}}$$

o sea

$$X = \frac{-b \pm \sqrt{b^2-4ac}}{2a}$$

que es la fórmula ordinaria que permite calcular las raíces de la ecuación cuadrática.

El número $\Delta = b^2-4ac$ se llama el discriminante de la ecuación.

Se ve que si $\Delta = 0$ las dos raíces son iguales. Además en el caso en que la ecuación tiene coeficientes reales, las dos raíces son reales si y sólo si $\Delta \geq 0$ y son complejas conjugadas si y sólo si $\Delta < 0$.

De aquí resulta en particular una propiedad que ya conocemos: Un polinomio de 2° grado $aX^2 + bX + c$ de $R[X]$ es irreducible en $R[X]$ si y sólo si $b^2-4ac < 0$.

EJEMPLO.

Resolver la ecuación $4X^2 - 4iX + (-1+2i) = 0$. Aplicando la fórmula anterior resulta

$$X = \frac{4i \pm \sqrt{(4i)^2 - 4.4(-1+2i)}}{2.4} = \frac{4i \pm \sqrt{-32\ i}}{8} = \frac{i \pm \sqrt{-2i}}{2}$$

Calculando $\sqrt{-2i}$ según vimos en el capítulo III se encuentra que $\sqrt{-2i} = \pm(-1+i)$.

Luego las raíces de la ecuación dada son

$$x_1 = -\frac{1}{2} + i \qquad x_2 = -\frac{1}{2}$$

Ecuaciones de tercer grado.

Sea la ecuación

$$aX^3 + bX^2 + cX + d = 0 \qquad (1)$$

con coeficientes numéricos cualesquiera, $a \neq 0$.

Dividiendo por a, reemplazando X por $X = Y - \frac{b}{3a}$ y efectuando los cálculos se obtiene la ecuación:

$$Y^3 + pY + q = 0 \qquad (2)$$

donde $p = \frac{c}{a} - \frac{b^2}{3a^2}$, $q = \frac{d}{a} - \frac{bc}{3a^2} + \frac{2b^3}{27a^3}$

Es claro que las raíces de (1) difieren de las de (2) en $\frac{b}{3a}$. Bastará entonces estudiar las ecuaciones cúbicas del tipo (2), que se llaman reducidas porque no tienen término de 2° grado.

Sea t una raíz cualquiera de (2) y consideremos la ecuación auxiliar en la incógnita W:

$$W^2 - tW - \frac{p}{3} = 0$$

Es una ecuación de 2° grado y posee dos raíces u y v en C. Por la relación que existe entre raíces y coeficientes se verifica

$$\begin{cases} u+v = t & (3) \\ u.v = -\frac{p}{3} & (4) \end{cases}$$

Reemplazando Y por $t = u+v$ en (2) se tiene

$$(u+v)^3 + p(u+v) + q = 0$$

o sea $$u^3 + v^3 + (3uv+p)(u+v) + q = 0 \qquad (5)$$

Por otro lado de $u.v = -\frac{p}{3}$ se deduce $3uv + p = 0$.

Luego de (5) resulta

$$u^3 + v^3 = -q \qquad (6)$$

y de $u.v = -\frac{p}{3}$ resulta

$$u^3 . v^3 = -\frac{p^3}{27} \qquad (7)$$

(6) y (7) prueban que u^3 y v^3 son raíces de la siguiente ecuación cuadrática

$$Z^2 + qZ - \frac{p^3}{27} = 0$$

Resolviendo esta ecuación se tiene:

$$u^3 = -\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}} \quad , \quad v^3 = -\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}$$

Luego

$$u = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} \quad , \quad v = \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}$$

y entonces

$$t = u+v = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}$$

fórmula que da las raíces de la ecuación (2) en función de sus coeficientes mediante operaciones racionales y radicales de 2° y 3° grado y que se conoce con el nombre de fórmula de Cardano. Observemos que como hay tres raíces cúbicas, hay tres valores posibles para u y tres para v y combinándolos la fórmula de Cardano da aparentemente 9 valores para las raíces de (2). Pero no es así porque (3) y (4) se verifican simultáneamente y por lo tanto los valores de u y v deben elegirse en cada caso de tal manera que $u.v = -\frac{p}{3}$.

Para cada valor de u uno solo de los valores de v satisface esa condición. En efecto, sea u_1 un valor de u. Entonces, de acuerdo con lo visto sobre las raíces de un número complejo en el capítulo III, los otros dos valores del radical u son $u_1.\varepsilon$ y $u_1.\varepsilon^2$, donde ε es una raíz cúbica primitiva de la unidad. Sea v_1 el valor que corresponde a u_1 tal que $u_1 v_1 = -\frac{p}{3}$ según (4). Los otros dos valores de v son $v_1.\varepsilon$ y $v_1.\varepsilon^2$. Como $\varepsilon^3 = 1$ es claro que los valores de v que pueden combinarse con

$$u_1 \quad ; \quad u_1.\varepsilon \quad ; \quad u_1.\varepsilon^2$$

son respectivamente

$$v_1 \quad ; \quad v_1.\varepsilon^2 \quad ; \quad v_1.\varepsilon$$

De modo que las tres raíces de la ecuación cúbica (2) son:

$$t_1 = u_1 + v_1$$

$$t_2 = u_1.\varepsilon + v_1.\varepsilon^2$$

$$t_3 = u_1.\varepsilon^2 + v_1.\varepsilon$$

EJEMPLO. Resolver la ecuación $X^3 + 3X^2 - 3X - 14 = 0$.

Reemplazando X por $X = Y - 1$ y efectuando los cálculos se obtiene la ecuación reducida

$$Y^3 - 6Y - 9 = 0$$

donde $p = -6$ y $q = -9$.

Entonces $-\frac{q}{2} = \frac{9}{2}$ y $\frac{q^2}{4} + \frac{p^3}{27} = \frac{49}{4}$

Luego

$$u = \sqrt[3]{\frac{9}{2} + \frac{7}{2}} = \sqrt[3]{8}$$

$$v = \sqrt[3]{\frac{9}{2} - \frac{7}{2}} = \sqrt[3]{1}$$

Teniendo en cuenta que las raíces cúbicas de la unidad son 1 , $\varepsilon = -\frac{1}{2} + \frac{\sqrt{3}}{2} i$, $\varepsilon^2 = -\frac{1}{2} - \frac{\sqrt{3}}{2} i$ resulta que las raíces de la ecuación reducida son:

$$t_1 = 2+1 = 3$$

$$t_2 = 2(-\frac{1}{2} + \frac{\sqrt{3}}{2} i) + (-\frac{1}{2} - \frac{\sqrt{3}}{2} i) = -\frac{3}{2} + \frac{\sqrt{3}}{2} i$$

$$t_3 = 2(-\frac{1}{2} - \frac{\sqrt{3}}{2} i) + (-\frac{1}{2} + \frac{\sqrt{3}}{2} i) = -\frac{3}{2} - \frac{\sqrt{3}}{2} i$$

Entonces, de acuerdo con la sustitución hecha, las raíces de la ecuación dada son

$$t_1 - 1 \quad , \quad t_2 - 1 \quad , \quad t_3 - 1$$

es decir

$$2 \quad , \quad -\frac{5}{2} + \frac{\sqrt{3}}{2} i \quad , \quad -\frac{5}{2} - \frac{\sqrt{3}}{2} i$$

EJERCICIO: Calcular las raíces de los siguientes polinomios:

a) $X^3 + 21X + 342$

b) $X^3 - 18X - 35$

c) $2X^3 + 3X^2 + 3X + 1$

Ecuaciones de cuarto grado.

Sea la ecuación

$$aX^4 + bX^3 + cX^2 + dX + c = 0 \qquad (1)$$

con coeficientes numéricos arbitrarios , $a \neq 0$.

Dividiendo por a y reemplazando X por $X = Y - \frac{b}{4a}$ se obtiene la ecuación

$$Y^4 + pY^2 + qY + r = 0 \qquad (2)$$

cuyas raíces difieren de las de la ecuación dada en $\frac{b}{4a}$. Basta entonces resolver

las ecuaciones cuárticas de este último tipo.

La resolución de la ecuación de 4° grado se reduce a la resolución de una ecuación cúbica auxiliar de la siguiente manera: Se introduce un parámetro e sumando y restando las expresiones eY^2 y $\frac{e^2}{4}$. Así la ecuación (2) tiene las mismas raíces que

$$Y^4 + eY^2 + \frac{e^2}{4} - eY^2 - \frac{e^2}{4} + pY^2 + qY + r = 0$$

o sea

$$(Y^2 + \frac{e}{2})^2 - \left[(e-p)Y^2 - qY + (\frac{e^2}{4} - r)\right] = 0 \qquad (3)$$

El primer término es un cuadrado perfecto y el término entre corchetes lo será eligiendo e de tal modo que el polinomio $(e-p)Y^2 - qY + (\frac{e^2}{4} - r)$ tenga una raíz múltiple, es decir tal que su discriminante sea cero:

$$q^2 - 4(e-p)(\frac{e^2}{4} - r) = 0 \qquad (4)$$

La expresión (4) es una ecuación cúbica en la incógnita e. Resolviendo esta ecuación se obtienen tres raíces que, en virtud de la fórmula de Cardano, se pueden expresar mediante operaciones racionales y radicales de 2° y 3° grado en función de los coeficientes de la ecuación (4) o sea de los coeficientes de la ecuación (2). Sea e_o una de ellas. Eligiendo este valor para e la expresión entre corchetes de (3) es un cuadrado perfecto y la ecuación (3) queda:

$$(Y^2 + \frac{e_o}{2})^2 - (e_o-p)(Y-e_o)^2 = 0 \qquad (5)$$

que como es una diferencia de cuadrados puede escribirse

$$\left[Y^2 + \frac{e_o}{2} + \sqrt{e_o-p}\,(Y - e_o)\right].\left[Y^2 + \frac{e_o}{2} - \sqrt{e_o-p}\,(Y - e_o)\right] = 0$$

Así las raíces de (5) se calculan resolviendo las dos ecuaciones cuadráticas siguientes:

$$\begin{cases} Y^2 + \frac{e_o}{2} + \sqrt{e_o-p}\,(Y - e_o) = 0 \\ Y^2 + \frac{e_o}{2} - \sqrt{e_o-p}\,(Y - e_o) = 0 \end{cases} \qquad (6)$$

Las raíces de estas dos ecuaciones son las de (5) y por lo tanto las de la ecuación (2).

Resumiendo, la ecuación de 4° grado (2) se puede resolver calculando una raíz e_o de la ecuación cúbica (4) y resolviendo las ecuaciones cuadráticas (6). Como las raíces de las ecuaciones (6) se pueden escribir en función de sus coeficientes mediante ope

raciones racionales y radicales de 2° grado, se ve que en definitiva las raíces buscadas se pueden expresar en función de los coeficientes de la ecuación (2) mediante operaciones racionales y radicales de 2° y 3° grado. Las fórmulas finales son muy complicadas y poco útiles desde el punto de vista práctico.

Ecuaciones bicuadradas. Un caso particular de ecuaciones de cuarto grado son las de la forma

$$aX^4 + bX^2 + c = 0$$

llamadas bicuadradas, cuya resolución se reduce a una de segundo grado haciendo la sustitución $Y = X^2$. Las cuatro raíces están dadas por la fórmula

$$\pm \sqrt{\frac{-b \pm \sqrt{b^2-4ac}}{2a}}$$

Ecuaciones de grado superior.

Para calcular las raíces de polinomios de grado ≥ 5 (y también las de tercer y cuarto grado) se utilizan distintos métodos que permiten calcularlas con un grado de aproximación tan grande como se desee. La complejidad e importancia de este problema hizo que se desarrollaran métodos diversos para calcular las raíces racionales, las reales o las complejas de un polinomio dado, y según que éste tenga coeficientes racionales, reales o complejos. Ante la carencia de una fórmula para calcular las raíces, se trataba de deducir propiedades de las mismas, estudiando el número de raíces reales y de raíces complejas, el número de raíces reales positivas y negativas, la acotación de las raíces reales y complejas, la existencia de raíces múltiples, etc. .

Por ejemplo, es muy simple eliminar las raíces múltiples pues dado un polinomio P(X), el polinomio

$$P^*(X) = \frac{P(X)}{(P(X),P'(X))}$$

no tiene raíces múltiples y toda raíz de P(X) es raíz de P*(X) y recíprocamente. De modo que se pueden calcular directamente las raíces de P*(X).

El procedimiento a seguir para calcular las raíces de un polinomio depende del polinomio dado, de las raíces que se deseen calcular y del grado de aproximación que se exija.

Si interesa calcular, por ejemplo, las raíces reales de un polinomio con coeficientes reales el procedimiento a seguir consta de tres etapas:

1°) Acotar las raíces, es decir determinar un intervalo $[\ell, L]$ de modo que las raíces reales estén todas en ese intervalo.

2°) Separar las raíces, es decir determinar subintervalos $[a_i, b_i]$ de $[\ell, L]$ de modo que en cada uno de ellos haya una sola raíz.

3°) Calcular cada raíz con una aproximación deseada.

En este curso nos limitaremos a los polinomios con coeficientes reales y veremos tan sólo una regla para acotar las raíces reales, una regla para acotar el número de raíces reales positivas y negativas, un método para calcular las raíces racionales de un polinomio con coeficientes racionales y un método simple para mejorar la aproximación de una raíz. Esta es una parte mínima de los resultados obtenidos para el cálculo de raíces. Por otra parte, el lector debe saber que existe un método general que permite obtener simultáneamente todas las raíces, reales y complejas, con un grado de aproximación tan grande como se desee y que no exige la separación previa de las mismas. Es el método de Gräeffe (desarrollado en 1837) cuyo fundamento teórico es bastante sencillo y que es el único método práctico para calcular las raíces complejas. Su aplicación requiere cálculos laboriosos.

I. Acotación de las raíces reales de un polinomio con coeficientes reales.

Dado un polinomio P(X), un número $\ell \in R$ se dice una cota superior (inferior) de las raíces reales de P(X) si todas ellas son menores o iguales (mayores o iguales) que ℓ. La siguiente regla sirve para acotar las raíces reales de un polinomio con coeficientes reales.

TEOREMA 4.12. (Regla de Laguerre-Thibault).

Sea $P(X) = a_n X^n + a_{n-1} X^{n-1} + \ldots\ldots + a_1 X + a_o \in R[X]$, $a_n > 0$. Si al dividir P(X) por $X - c$, $c \geq 0$, todos los coeficientes del cociente y del resto son no negativos entonces c es una cota superior de las raíces reales de P(X).

Demostración: Por el teorema del resto se tiene:

$$P(X) = (X - c).Q(X) + P(c)$$

Si todos los coeficientes de Q(X) y P(c) son números no negativos, es claro que para todo $x \in R$, $x > c$, es $P(x) > 0$. Luego P(X) no tiene ninguna raíz mayor que c, y el teorema queda probado.

Para encontrar una cota inferior, observemos que las raíces del polinomio P(-X), es decir de la ecuación:

$$a_n X^n - a_{n-1} X^{n-1} + \ldots\ldots + (-1)^n a_o = 0$$

son las de P(X) cambiadas de signo. Entonces, si ℓ es una cota superior de las raíces de P(-X), como $t \leq \ell$ implica $-\ell \leq -t$ se tiene que $-\ell$ es una cota inferior de las raíces del polinomio dado P(X).

EJEMPLO. Acotar las raíces reales del polinomio $X^3 - 3X^2 + 5X + 4$.

Aplicando la regla de Ruffini se divide el polinomio dado por $X - 1$, $X - 2$, $X - 3$ obteniéndose en esta última división todos los coeficientes del cociente y el resto no negativos. Luego 3 es una cota superior de las raíces reales de dicho polinomio. Para determinar una cota inferior aplicamos el mismo procedimiento a la ecuación $X^3 + 3X^2 + 5X - 4 = 0$ y se ve que 1 es una cota superior de las raíces de la misma. Luego las raíces reales del polinomio dado están todas en el intervalo $[-1,3]$.

La regla de Laguerre-Thibault es uno de los métodos más simples de acotación de raíces reales. Existen otros más complicados que dan una aproximación mejor. Existen también métodos para acotar las raíces reales y complejas de polinomios con coefi - cientes complejos. En este caso lo que se acota es el módulo de las raíces, es decir se determina un número real $\ell > 0$ tal que todas las raíces quedan comprendidas en el círculo de centro 0 y radio ℓ del plano complejo.

EJERCICIOS.

Acotar las raíces reales de los siguientes polinomios:

a) $2X^4 - 7X^3 + X^2 + 10X - 5$

b) $X^3 - 4X^2 - 36$

c) $X^7 - 16X^5 - 5X^4 - 47X^3 - 6X + 1$

d) $2X^6 + X^4 + 11X^3 + 5X + 1$

e) $X^3 - 10X^2 + 1$

II. Regla de los signos de Descartes para acotar el número de raíces reales positivas y negativas de un polinomio con coeficientes reales.

Existen métodos para calcular el número exacto de raíces reales positivas y negativas que tiene un polinomio $P(X) \in R[X]$, más aún, que permiten saber el número de raíces que hay en cualquier intervalo (a,b). Estos métodos permiten así "separar" las raíces reales, es decir determinar intervalos (a_i, b_i) tales que en cada uno de ellos haya una sola raíz. Pero son en general bastante complicados, siendo el más sencillo de todos el de Sturm, desarrollado en 1829. No lo veremos en este curso y en cambio daremos la llamada regla de los signos de Descartes que proporciona solamente una acotación del número de raíces reales positivas y negativas.

Dado un polinomio $P(X) = a_n X^n + \ldots\ldots + a_1 X + a_o \in R[X]$ se consideran las variaciones de signo que presenta la sucesión de sus coeficientes. Por ejemplo, dado el polinomio $X^4 - 2X^3 - \sqrt{2}\,X + 3$, sus coeficientes presentan dos variaciones de sig

no; en $2X^7 - 4X^5 + 3X^4 + X^3 - 6X + 4$, hay cuatro variaciones de signo.

Observemos que si en un polinomio P(X) el coeficiente principal y el último coeficiente no nulo tienen el mismo signo entonces hay un número par de variaciones de signo.

TEOREMA 4.13. (Regla de Descartes).

El número de raíces positivas de un polinomio $P(X) \in R[X]$, contadas tantas veces como su orden de multiplicidad, es igual o menor que el número de variaciones de signo de los coeficientes de P(X) y difiere del mismo en un número par.

Teniendo en cuenta que las raíces del polinomio P(-X) son las de P(X) cambiadas de signo, de este teorema resulta inmediatamente que el número de raíces negativas de un polinomio $P(X) \in R[X]$, contadas tantas veces como su orden de multiplicidad, es igual o menor que el número de variaciones de signo de los coeficientes de P(-X) y difiere del mismo en un número par.

Para demostrar el teorema veremos primero el siguiente lema.

LEMA. Si se multiplica un polinomio no nulo $P(X) \in R[X]$ por otro de la forma $X - c$, $c > 0$, el número de variaciones de signo de los coeficientes de P(X) aumenta en un número impar.

Demostración: La haremos por inducción sobre el grado de P(X).

Si gr P(X) = 0 el lema se verifica trivialmente.

Sea gr P(X) = n > 0 y supongamos la propiedad verdadera para todos los polinomios no nulos de grado menor que n.

Si $P(X) = a_nX^n + \ldots\ldots + a_1X + a_o$, sea k el menor índice tal que a_n , a_{n-1} , $\ldots\ldots$, a_k son todos del mismo signo (se puede suponer $a_n > 0$) , y sea $Q(X) = a_{k-1}X^{k-1} + \ldots\ldots + a_1X + a_o$. Luego $P(X) = a_nX^n + \ldots + a_kX^k + Q(X)$ y por nuestra hipótesis sobre k se ve que el número de variaciones de signo de los coeficientes de P(X) es igual al de Q(X) más uno.

$$(X - c).P(X) = a_nX^{n+1} + (a_{n-1} - ca_n)X^n + \ldots + (a_k - ca_{k+1})X^{k+1} - ca_kX^k + (X - c).Q(X)$$

Los coeficientes de $(X - c).P(X)$ son:

$$a_n \ , \ a_{n-1} - ca_n \ , \ldots\ldots \ , \ a_k - ca_{k+1} \ , \ a_{k-1} - ca_k \ , \ a_{k-2} - ca_{k-1} \ , \ldots\ldots \ , \ -ca_o$$

Comparemos las variaciones de signo de estos coeficientes con las de P(X). Como es $a_n > 0$, $a_k > 0$, $a_{k-1} < 0$ y $a_{k-1} - ca_k < 0$, se tiene , por un lado , que la sucesión de coeficientes $a_{k-1} - ca_k$, $a_{k-2} - ca_{k-1}$, $\ldots\ldots$, $-ca_o$ tiene el mismo número

de variaciones de signo que la sucesión a_{k-1} , $a_{k-2} - ca_{k-1}$, ... , $-ca_o$ que son los coeficientes de $(X - c).Q(X)$; por la hipótesis de inducción este número es mayor que el de $Q(X)$ en un número impar. Por otro lado, la sucesión a_n , $a_{n-1} - ca_n$, ... , $a_{k-1} - ca_k$ presenta un número de variaciones impar, puesto que el primer y el último término tienen signos distintos, mientras que la sucesión a_n , a_{n-1} , ... , a_k de coeficientes de $P(X)$ no presenta ninguna variación de signo.

Luego el número total de variaciones de signo de los coeficientes de $(X - c).P(X)$ supera al de los coeficientes de $P(X)$ en un número impar.

Probaremos ahora la regla de Descartes.

Sea $P(X) \in R[X]$. Descomponiendo a $P(X)$ en producto de factores irreducibles en $R[X]$, $P(X)$ se escribirá en la forma:

$$P(X) = a_n X^k (X - c_1)(X - c_2) \ldots\ldots (X - c_s) \left[(X + t_1) \ldots\ldots (X + t_r) . (X^2 + p_1 X + q_1) \ldots\ldots (X^2 + p_h X + q_h)\right] \qquad (1)$$

donde c_1 , c_2 , ... , c_s son las raíces positivas de $P(X)$, $-t_1$, ... , $-t_r$ las raíces negativas y los polinomios de segundo grado son irreducibles, es decir $p_i^2 - 4q_i < 0$.

Luego t_1 , ... , t_r , q_1 , ... , q_h son números positivos y el polinomio entre corchetes tiene el primer y el último coeficientes del mismo signo lo que implica que el número de variaciones de signo de sus coeficientes es par. Al multiplicarlo por $(X - c_1)$, ... , $(X - c_s)$ el número de variaciones aumenta con cada multiplicación en un número impar. De (1) resulta entonces que el número de raíces positivas de $P(X)$ debe tener la misma paridad que el número de variaciones de signo de los coeficientes de $P(X)$ y no puede superar este número, lo que termina la demostración.

Por ejemplo, dado el polinomio $P(X) = X^9 - 3X^5 + 2X^2 - X + 12$ sus coeficientes presentan cuatro variaciones de signo. Esto significa que puede tener 4 , 2 ó ninguna raíz real positiva. Como $P(-X) = -X^9 + 3X^5 + 2X^2 + X + 12$ presenta una variación de signo, el polinomio dado $P(X)$ tiene una raíz real negativa. Resumiendo , $P(X)$ tiene una raíz real negativa , 4 , 2 ó ninguna raíz positiva y por lo menos 4 raíces complejas. (Tiene 4 , 6 ó 8 raíces complejas, que se presentan de a pares conjugadas). Este ejemplo muestra que la regla de Descartes no proporciona una información muy precisa sobre el número de raíces reales en algunos casos.

Veamos otro ejemplo. Consideremos el polinomio $P(X) = X^4 - X^3 - 2X - 1$. Como presenta una variación de signo, tiene una raíz real positiva.

Analizando $P(-X) = X^4 + X^3 + 2X - 1$, se ve que presenta una variación de signo.

Luego el polinomio dado tiene una raíz real positiva, una negativa y dos complejas conjugadas.

EJERCICIO. Analizar las raíces de los siguientes polinomios aplicando la regla de Descartes:

a) $2X^3 + 3X^2 - 13X + 6$

b) $X^2 - 2X + 7$

c) $X^4 - 2X^2 - 3X - 2$

d) $X^3 - 3X - 14$

e) $X^6 - 3X^2 - 4X + 1$

f) $X^6 + X^4 - X^3 - 2X - 1$

g) $X^6 + 2X^4 + X^2 - 3$

h) $X^3 + 10X + 1$

III. Cálculo de las raíces racionales de un polinomio con coeficientes racionales.

Si $P(X) \in Q[X]$, para calcular sus raíces podemos considerarlo de la forma

$$a_n X^n + a_{n-1} X^{n-1} + \ldots\ldots + a_1 X + a_o$$

con $a_i \in Z$, es decir con todos los coeficientes enteros, pues el polinomio $P(X)$ tiene las mismas raíces que el que se obtiene multiplicando a $P(X)$ por el mínimo común múltiplo de los denominadores de sus coeficientes.

Por ejemplo, dada la ecuación $3X^4 - \frac{1}{2} X^3 + \frac{2}{3} X^2 - \frac{3}{5} X + 1 = 0$ sus raíces coinciden con las de la ecuación $90X^4 - 15X^3 + 20X^2 - 18X + 30 = 0$.

El siguiente teorema da un método para saber si $P(X)$ tiene raíces racionales y calcularlas.

TEOREMA 4.14. Si un número racional $\frac{p}{q}$, p y q relativamente primos, es raíz de un polinomio $P(X) = a_n X^n + \ldots\ldots + a_1 X + a_o$ con coeficientes enteros entonces p/a_o y q/a_n .

Demostración: Multiplicando $P(\frac{p}{q})$ por q^n se tiene:

$$a_n p^n + a_{n-1} p^{n-1} q + \ldots\ldots + a_1 p q^{n-1} + a_o q^n = 0$$

Como los n primeros términos de la suma del primer miembro y el segundo miembro son múltiplos de p , debe ser $p/a_o q^n$. Siendo p y q relativamente primos, se tiene p/a_o . Análogamente se deduce que q/a_n .

Luego, para calcular las raíces racionales de $P(X)$ se determinan todos los divisores de a_o , todos los de a_n , se forman todos los números $\frac{p}{q}$ posibles y mediante sustitución directa o aplicando la regla de Ruffini se verifica cuáles de ellos son raíces del polinomio dado.

COROLARIO. Si $a_n = 1$, las raíces racionales de $P(X)$ sólo pueden ser enteras y se

encuentran entre los divisores de a_o .

Como aplicación inmediata de este corolario resulta, por ejemplo, que $\sqrt{2}$ es un número irracional.

En efecto, $\sqrt{2}$ es una raíz del polinomio $X^2 - 2$ y de acuerdo con el corolario, las raíces racionales de este polinomio sólo podrían ser ± 1 , ± 2 . En general, si p es un entero primo positivo entonces $\sqrt[n]{p}$ es un número irracional, si p_1 , p_2 ,..., p_k son enteros primos positivos distintos entonces $\sqrt[n]{p_1 \cdot p_2 \cdot \ldots \cdot p_k}$ es irracional, puesto que las ecuaciones

$$X^n - p = 0 \qquad y \qquad X^n - p_1 \cdot p_2 \cdot \ldots \cdot p_k = 0$$

no tienen raíces racionales.

EJEMPLO. Calcular las raíces racionales de $X^3 + \frac{1}{2} X^2 - \frac{7}{2} X - 3$. Equivale a calcular las del polinomio $2X^3 + X^2 - 7X - 6$.

Los divisores de -6 son ± 1 , ± 2 , ± 3 , ± 6 . Los de 2 son ± 1 , ± 2 . Entonces las posibles raíces racionales son: ± 1 , ± 2 , ± 3 , ± 6 , $\pm \frac{1}{2}$, $\pm \frac{3}{2}$. Reemplazando estos valores se ve que las raíces racionales son -1 , 2 y $-\frac{3}{2}$.

Cuando hay que hacer muchas verificaciones pueden ahorrarse cálculos analizando el número de raíces positivas y negativas mediante la regla de los signos de Descartes. Por ejemplo, el polinomio dado presenta una variación de signo; luego tiene una sola raíz positiva. De modo que una vez que se verificó que 2 es raíz, se pueden descartar todos los demás valores positivos.

EJERCICIOS.

1. Hallar todas las raíces racionales de los siguientes polinomios:

   a) $X^4 - 2X^2 - 3X - 2$

   b) $X^3 - X - 6$

   c) $X^5 - X^3 + 2$

   d) $X^5 - 10$

   e) $X^3 - \frac{9}{2} X^2 + 6X - \frac{5}{2}$

   f) $4X^4 - 11X^2 + 9X - 2$

   g) $2X^3 - X^2 + 1$

   h) $3X^4 + 7X^2 + 6$

2. a) Si $\frac{p}{q}$, p y q relativamente primos, es una raíz racional de un polinomio P(X) con coeficientes enteros, demostrar que $(p-tq)/P(t)$, $\forall\ t \in Z$.

   b) Demostrar que si un polinomio P(X) con coeficientes enteros es tal que P(0) y P(1) son números impares, entonces P(X) no tiene ninguna raíz entera.

IV. Cálculo aproximado de las raíces reales de un polinomio con coeficientes reales.

Supongamos que las raíces reales de un polinomio $P(X)=a_nX^n+a_{n-1}X^{n-1}+.....+a_o \in R[X]$ han sido separadas, es decir, que se conocen intervalos $[a_i,b_i]$ cada uno de los cuales contiene una única raíz. ¿Cómo hacer para calcularlas con un grado de aproximación deseado, digamos por ejemplo, con cuatro cifras decimales exactas?.

Existen diferentes métodos para aproximar una raíz: el de Horner, el de iteración, el de Newton-Fourier, etc. . La elección del método y la intensidad de su aplicación dependen de la ecuación dada y del grado de aproximación deseado, buscando realizar el menor número posible de cálculos.

Si sólo se necesitan aproximaciones pequeñas, el método gráfico puede resultar conveniente para localizar las raíces reales, realizando previamente un estudio del polinomio dado.

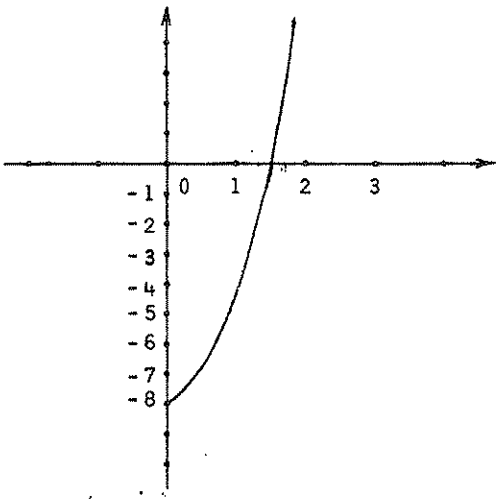
Sólo indicaremos este método, complementándolo con un recurso simple para aproximar raíces utilizando la regla de Ruffini, que se basa en el hecho de que la función real de variable real $f(x) = a_nx^n + a_{n-1}x^{n-1} + ... + a_1x + a_o$ es continua y en la siguiente propiedad de las funciones continuas:

TEOREMA. Si una función real continua f(x) en un intervalo [a,b] toma valores de signos opuestos en los extremos del intervalo, entonces se anula por lo menos una vez en dicho intervalo.

EJEMPLO. Sea la ecuación $X^3 + 3X - 8 = 0$

Por la regla de Descartes tiene una única raíz real positiva y dos complejas conjugadas. Se puede encontrar un valor aproximado de la raíz real t dibujando el gráfico de la función $f(x) = x^3 + 3x - 8$

x	0	1	2
f(x)	-8	-4	6



Como f(x) cambia de signo entre x=1 y x=2, t está en el intervalo [1,2] . Entonces se trata de mejorar la aproximación por tanteos sucesivos encontrando números a,b tales que $f(a) < 0$, $f(b) > 0$ y $1 < a < b < 2$.

Calculemos por ejemplo, f(1.5) :

	1	0	3	-8
1.5 *		1.5	2.25	7.875
	1	1.5	5.25	-0.125 = f(1.5)
1.6		1.6	2.56	8.896
	1	1.6	5.56	0.896 = f(1.6)

Como $f(1.5) < 0$ y $f(1.6) > 0$ es $1.5 < t < 1.6$.

Calculemos f(1.51) :

	1	0	3	-8
1.51		1.51	3.2501	9.43
	1	1.51	6.2501	1.43 = f(1.51)

Como $f(1.5) < 0$ y $f(1.51) > 0$ es $1.50 < t < 1.51$.

Luego el valor de t es 1.50 con las dos cifras decimales exactas. Calculemos una cifra decimal más

	1	0	3	-8
1.505		1.505	2.265	7.923
	1	1.505	5.265	-0.077 = f(1.505)
1.507		1.507	2.271	7.943
	1	1.507	5.271	-0.057 = f(1.507)
1.509		1.509	2.277	7.963
	1	1.509	5.277	-0.037 = f(1.509)

Luego t = 1.509 con las tres cifras decimales exactas.

EJERCICIOS.

a) Calcular las raíces positivas de la ecuación $X^3 - 2X^2 - 2X - 7 = 0$ con un error < 0.01 .

b) Idem para $X^3 + X - 3$.

c) Calcular la raíz negativa de $X^4 - X^3 - 2X^2 + 3X - 3$ con una cifra decimal exacta.

PROBLEMAS

1. Un recipiente de 750 litros puede ser llenado por un grifo en un cierto tiempo. Si se agrega otro que arroja 200 litros por hora, se necesita una hora menos.

   Se pide: a) Cuánto arroja por hora el primer grifo y cuánto tarda en llenar el recipiente.

   b) Dar una interpretación física de la solución extraña que aparece en la ecuación resolvente.

2. La longitud de una caja rectangular de 64 m^3 de volumen es el doble que su ancho y su profundidad es un metro mayor que el ancho. Hallar el ancho con un cifra decimal exacta.

3. Un señor compró 16 revistas de música, arqueología y deportes y pagó por cada revista tantos pesos como revistas compró de esa clase. Gastó en total $ 94.-

   ¿Cuántas revistas de cada clase compró si multiplicando esos números da 126 y aparecen enumeradas en orden decreciente de costo?.

4. Una caja sin tapa tiene la forma de un cubo de arista 10 cm. Si la capacidad de la caja es de 500 cm^3 , ¿cuál es el espesor de las paredes con dos cifras decimales exactas?.Se suponen de espesor uniforme.

NOTA.

Dado un cuerpo K, se puede generalizar la noción de polinomios en una indeterminada, y considerar polinomios en n indeterminadas X_1 , X_2 , ... , X_n con coeficientes en K.

Se llama monomio en X_1 , X_2 , ... , X_n con coeficiente en K a toda expresión de la forma:

$$a\, X_1^{t_1} X_2^{t_2} \ldots\ldots X_n^{t_n}$$

donde $a \in K$ y los exponentes t_i son enteros no negativos; a se llama coeficiente y si $a \neq 0$, el número $t_1+t_2+\ldots\ldots+t_n$ se dice el grado del monomio.

Un polinomio en las indeterminadas X_1 , X_2 , ... , X_n con coeficientes en K es una suma finita (puramente formal) de monomios.

Por ejemplo ,

$$P(X,Y) = 3X^3Y^2 + (-\tfrac{1}{2})XY^3 + 8X^2 + (-2)XY + Y + 1$$

$$P(X,Y,Z) = X^7YZ^3 + (-5)X^6Z^4 + 2Y^5 + Z^3 + XYZ + 6XY^2 + X + (-7)$$

son polinomios con coeficientes en Q, en dos indeterminadas X,Y el primero, y en tres indeterminadas X,Y,Z el segundo.

El conjunto de todos los polinomios en n indeterminadas X_1 , X_2 , ... , X_n con coeficientes en K se representa $K[X_1,X_2,\ldots\ldots,X_n]$. En este conjunto se definen una suma y una multiplicación en forma semejante a como se hace para los polinomios en una indeterminada y $K[X_1,X_2,\ldots\ldots,X_n]$ con estas operaciones es un anillo conmutativo sin divisores de cero.

Dado un polinomio $P \in K[X_1,\ldots\ldots,X_n]$ no nulo, se llama grado de P al grado del monomio de mayor grado que figura en P. Por ejemplo, los grados de los polinomios vistos en el ejemplo anterior son 5 y 11 respectivamente.

Un polinomio P se dice homogéneo de grado m si todos los monomios que lo forman son de grado m. Así por ejemplo

$$P(X,Y,Z) = 2X^3Y^2Z + (-3)XY^5 + 4Z^6 + 5X^2Z^4 + XYZ^4$$

es un polinomio homogéneo de 6° grado en tres indeterminadas.

Se demuestra que en $K[X_1,\ldots\ldots,X_n]$ todo polinomio no constante puede escribirse como producto de polinomios irreducibles y que esta factorización es única salvo constantes y el orden de los factores.

CAPITULO V

CALCULO COMBINATORIO

5.1. VARIACIONES. Se trata de poder resolver un problema como el siguiente: ¿Cuántos números de cuatro cifras se pueden escribir con los dígitos 1,3,4,6,7,8 y 9 sin que haya cifras repetidas en cada número? ¿Cuántos de ellos empiezan con 1? ¿Cuántos empiezan con 1 y terminan con 73? ¿Cuántos son pares?.

Definición. Dados m elementos distintos se llama una variación (sin repetición) de orden n de los m elementos a toda sucesión formada por n de esos elementos

$$a_1\ a_2 \cdot \cdot \cdot \cdot \cdot \cdot \cdot a_n$$

donde $a_i \neq a_j$ para $i \neq j$.

Es claro que una sucesión de este tipo existe si y sólo si $n \leq m$. Con sucesión de n términos queremos significar un conjunto con n elementos, totalmente ordenado. Dos variaciones de orden n son distintas si difieren en algún elemento o en el orden de los mismos.

El número de variaciones de orden n de m objetos se representa V_m^n .

Por ejemplo, consideremos el conjunto formado por cuatro objetos a,b,c,d. Las variaciones de estos 4 objetos tomados de a 2 son:

a b	b a	c a	d a
a c	b c	c b	d b
a d	b d	c d	d c

$$V_4^2 = 4.3 = 12$$

Las variaciones de orden 3 de esos cuatro objetos se pueden obtener agregando a la derecha de cada una de las variaciones de orden 2, cada uno de los dos objetos que no figuran en ella. Así el número de variaciones de orden 3 es

$$V_4^3 = V_4^2 \cdot 2 = 4.3.2$$

En general

TEOREMA 5.1. El número V_m^n de variaciones de orden n de m objetos ($n \leq m$) está dado por la fórmula

$$V_m^n = m(m-1)(m-2) \cdot \cdot \cdot \cdot \cdot \cdot \cdot (m-n+1)$$

Demostración: La fórmula se verifica para n = 1 pues obviamente el número de variaciones de m objetos tomados de a uno es m: $V_m^1 = m$.

Sea n > 1 , supongamos la fórmula verdadera para n-1 , es decir

$$V_m^{n-1} = m(m-1).....(m-n+2)$$

y probemos que también es válida para n.

Formadas las variaciones de orden n-1 de los m objetos dados, todas las variaciones de orden n pueden obtenerse agregando a la derecha de cada una de aquéllas, uno de los $m-(n-1) = m-n+1$ objetos que no figuran en dicha variación. Como de cada variación de orden n-1 se obtienen m-n+1 variaciones de orden n, se verifica que

$$V_m^n = V_m^{n-1}.(m-n+1) = m(m-1).....(m-n+2)(m-n+1) \qquad \text{c.q.d.}$$

De este teorema resulta entonces que el número de variaciones de orden n de m objetos es igual al producto de n factores decrecientes a partir de m.

Por ejemplo , $V_9^4 = 9.8.7.6 = 3024$

$$V_6^2 = 6.5 = 30$$

Resolvamos ahora el problema que planteamos al principio. La respuesta a la primera pregunta es el número de variaciones que se pueden formar con los 7 dígitos dados , tomados de 4 en 4 , es decir

$$V_7^4 = 7.6.5.4 = 840$$

Para contestar a la segunda pregunta, tengamos en cuenta que si la primera cifra es 1, entonces quedan tres lugares que pueden ser ocupados por los dígitos restantes 3, 4,6,7,8 y 9. Luego la respuesta es

$$V_6^3 = 6.5.4 = 120 \qquad 1 \_ \_ \_$$

Razonando análogamente, la respuesta a la tercera pregunta es

$$V_4^1 = 4 \qquad 1 \_ 7\ 3$$

Por último, son pares los que terminan en 4, 6 ó 8. El número de los mismos es

$$3.V_6^3 = 3.120 = 360$$

Resolvamos este otro problema: ¿Cuántas "palabras" de cinco letras se pueden formar con letras de la palabra CONJETURA sin repetir ninguna? ¿Cuántas empiezan con NO? ¿En cuántas figura la sílaba NO? ¿En cuántas figura la letra T?

Se trata de las variaciones de 9 elementos tomados de 5 en 5. Su número es

$$V_9^5 = 9.8.7.6.5 = 15120$$

Las que empiezan con NO son:

$$V_7^3 = 7.6.5 = 210$$

La sílaba NO figura en:

$$4.V_7^3 = 4.210 = 840$$

N O _ _ _

_ N O _ _

_ _ N O _

_ _ _ N O

El número de las que tienen la letra T es:

$$5.V_8^4 = 5.8.7.6.5 = 8400$$

Por razones de comodidad en la notación, se introduce la noción de factorial.

Definición. Dado un entero $n \geq 0$ se llama factorial de n y se representa $n!$ al número definido por recurrencia como sigue:

$$0! = 1$$

$$n! = n.(n-1)!$$

Es decir , $n! = n.(n-1).(n-2).....3.2.1$

Por ejemplo , $7! = 7.6.5.4.3.2.1 = 5040$

Con esta notación, la fórmula que da el número de variaciones de orden n de m elementos es

$$V_m^n = \frac{m!}{(m-n)!}$$

5.2. VARIACIONES CON REPETICION.

Supongamos ahora que se trata de resolver el problema que planteamos al principio del parágrafo anterior, pero pudiendo haber cifras repetidas en cada número, es decir admitiéndose números del tipo 3383.

Definición. Dado un conjunto finito A de m elementos distintos, se llama variación con repetición de orden n de esos m elementos a toda sucesión de n términos formada por elementos de A, no necesariamente distintos entre sí.

El número de variaciones con repetición de orden n de m objetos se representa V'^n_m .

Por ejemplo, las variaciones con repetición de orden 3 de dos elementos a,b son :

a a a , a a b , a b a , b a a , a b b , b a b , b b a , b b b

Luego $V'^{3}_{2} = 8$.

TEOREMA 5.2. El número V'^{n}_{m} de variaciones con repetición de orden n de m elementos está dado por la fórmula

$$V'^{n}_{m} = m^{n}$$

Demostración: Vamos a demostrarlo por inducción sobre n.

Dados m elementos, si n = 1 es inmediato que la fórmula se verifica: $V'^{1}_{m} = m$.

Sea n > 1 , supongamos la fórmula verificada para n-1, es decir $V'^{n-1}_{m} = m^{n-1}$, y probémosla para n. Razonando como lo hicimos en el teorema anterior, formadas las variaciones con repetición de orden n-1 de los m elementos dados, las de orden n se obtienen agregando a la derecha de cada una de las de orden n-1 un elemento más,siendo en este caso m los elementos disponibles para ello. Luego

$$V'^{n}_{m} = V'^{n-1}_{m}.m = m^{n-1}.m = m^{n}$$

Podemos resolver ahora el problema planteado. La primera respuesta se obtiene calculando el número de variaciones con repetición de orden 4 que se pueden formar con los 7 dígitos dados:

$$V'^{4}_{7} = 7^{4} = 2401$$

De estos números, los que empiezan con 1 son:

$$V'^{3}_{7} = 7^{3} = 343$$

Los que empiezan con 1 y terminan con 73 son:

$$V^{1}_{7} = 7$$

Por último, los números pares son:

$$3.V'^{3}_{7} = 3.7^{3} = 1029$$

Si consideramos el problema de las palabras de 5 letras que se pueden formar con las letras de la palabra CONJETURA antes enunciado y levantamos la restricción de que no haya letras repetidas en cada palabra, las respuestas a las sucesivas preguntas son:

$$V'^{5}_{9} = 9^5 = 59049$$

$$V'^{3}_{9} = 9^3 = 729$$

$$4.V'^{3}_{9} - 3.9 = 2889$$

$$V'^{5}_{9} - V'^{5}_{8} = 26281$$

5.3. PERMUTACIONES.

Definición. Dados m objetos distintos, se llama una permutación de los mismos a toda variación de orden m de esos m objetos. El número de permutaciones de m elementos se indica P_m y es

$$P_m = V^m_m = m!$$

Por ejemplo, dados tres elementos a,b,c , todas las permutaciones de los mismos son:

a b c , a c b , b a c , b c a , c a b , c b a

y $P_3 = 3! = 6$.

¿Cuántas palabras distintas pueden obtenerse reordenando las letras de la palabra DURAZNO?

La respuesta es el número de permutaciones de las siete letras que forman e s a pa labra .

$$P_7 = 7! = 5040$$

En cambio, si se trata de resolver el mismo problema con la palabra ABRACADABRA , la solución ya no es el número P_{11} de permutaciones de once elementos porque, como de las once letras que figuran en esa palabra hay iguales cinco A , dos B y dos R , si se intercambian estas letras entre sí en una permutación dada, las permutaciones que se obtienen no van a diferir de la de partida.

Por ejemplo, si se considera la permutación

B A R C D A A B A R A

intercambiando las cinco A entre sí en todas las formas posibles, se obtienen 5! permutaciones iguales a la dada; de cada una de éstas a su vez se obtienen 2! permutaciones iguales a la dada intercambiando las dos B; y de cada una de ellas se obtie-

nen 2! permutaciones iguales a las anteriores intercambiando las R entre sí. En total el número de permutaciones iguales que se obtienen es 5!2!2!. Como esto sucede con cada ordenación de esas once letras, si llamamos t al número de permutaciones distintas entre sí, es claro que se tiene $P_{11} = 5!2!2!t$. La solución del problema es entonces

$$\frac{P_{11}}{5!2!2!} = \frac{11!}{5!2!2!} = 83160$$

En general ,

PERMUTACIONES CON REPETICION.

Dados m elementos entre los que hay k_1 iguales entre sí, k_2 iguales entre sí,....,k_s iguales entre sí, el número de permutaciones distintas de los m elementos está dado por la fórmula

$$\frac{m!}{k_1!k_2!.....k_s!}$$

Se demuestra con un razonamiento análogo al que hicimos recién.

EJERCICIO.

El lector está en condiciones ahora de contestar las siguientes preguntas:

Dado un conjunto A con n elementos y uno B con m elementos

1) ¿Cuántas son las aplicaciones que se pueden definir de A en B?.

2) Si $n \leq m$, cuántas aplicaciones inyectivas hay de A en B?.

3) Si $n \geq m$, cuántas son las epiyectivas?.

4) Si $n = m$, cuántas aplicaciones biyectivas hay de A en B?.

Calcular esos números para $n = 8$ y $m = 13$.

5.4. COMBINACIONES.

Tratemos de resolver ahora el siguiente problema: En una oficina donde trabajan 15 empleados se desea elegir una delegación de 4 personas para viajar en comisión a otra ciudad. ¿De cuántas maneras es posible elegir dicha delegación?.

En general, se plantea así el problema siguiente: Dados m elementos distintos, de cuántas maneras se pueden elegir n entre esos m elementos, es decir, cuántos subconjuntos con n elementos se pueden formar con esos m elementos dados? ($n \leq m$).

Definición. Dado un conjunto A de m elementos, se llama combinación de orden n ($n \leq m$) de esos m elementos a todo subconjunto de A con n elementos.

El número de combinaciones de orden n de m objetos se indica C_m^n .

Por ejemplo, si $A = \{a,b,c,d\}$, las combinaciones de orden 3 de estos 4 elementos son:

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

Observemos que considerando todas las permutaciones de los elementos de cada uno de esos subconjuntos se obtienen las variaciones de orden 3 de los 4 elementos dados:

a b c	a b d	a c d	b c d
a c b	a d b	a d c	b d c
b a c	b a d	c a d	c b d
b c a	b d a	c d a	c d b
c a b	d a b	d a c	d b c
c b a	d b a	d c a	d c b

Entonces se ve que $V_4^3 = C_4^3 \cdot P_3$

En general

$$V_m^n = C_m^n \cdot P_n \qquad (1)$$

pues todas las variaciones de orden n de m elementos se pueden obtener permutando en todas las formas posibles los elementos de cada combinación de orden n de los m elementos dados ; como de cada combinación se obtienen P_n variaciones, sigue la fórmula (1).

Luego

$$C_m^n = \frac{V_m^n}{P_n}$$

De aquí resulta que

TEOREMA 5.3.

$$C_m^n = \frac{m!}{n!(m-n)!}$$

Resolvamos ahora el problema antes enunciado. La solución es $C_{15}^4 = \frac{15!}{4!11!} = 1365.$

Problema. La comisión directiva de un club está formada por 7 hombres y 5 mujeres.

a) ¿De cuántas maneras puede formarse una subcomisión de 4 miembros si se desea que en ella figure por lo menos un hombre?.

b) ¿En cuántas figurará el presidente del club?.

c) ¿En cuántas el presidente y el secretario?.

a) Para contestar a la primera pregunta, hay que tener en cuenta que la subcomisión de 4 miembros puede formarse con hombres solamente , con 3 hombres y una mujer, con 2 hombres y 2 mujeres o con 1 hombre y 3 mujeres. El número total de subcomisiones posibles es entonces

$$C_7^4 + C_7^3.C_5^1 + C_7^2.C_5^2 + C_7^1.C_5^3 = 490$$

b) El número de subcomisiones en que figura el presidente es:

$$C_6^3 + C_6^2.C_5^1 + C_6^1.C_5^2 + C_5^3 = 165$$

c) El número de subcomisiones en que figuran presidente y secretario es:

$$C_5^2 + C_5^1.C_5^1 + C_5^2 = 45$$

Problema. Dadas las letras A , T , R , o , p , m , e , s , 1 :

a) ¿Cuántas palabras de 5 letras pueden formarse de modo que comiencen con mayúscula y no haya letras repetidas? (ni mayúsculas en el medio).

b) ¿Cuántas empiezan con T?. ¿Cuántas de éstas terminan con s?.

c) ¿En cuántas figuran 2 vocales y 3 consonantes?.

a) Elegida una mayúscula para ocupar el primer lugar, hay que elegir cuatro minúsculas para llenar los lugares restantes y una vez elegidas éstas, hay que considerar todas sus ordenaciones posibles. Luego la respuesta es:

$$C_3^1.C_6^4.P_4 = 1080$$

b) Las respuestas a las preguntas de b) son:

$$V_6^4 = C_6^4.P_4 = 360 \qquad T\ \_\ \_\ \_\ \_$$

y

$$V_5^3 = C_5^3.P_3 = 60 \qquad T\ \_\ \_\ \_\ s$$

c) Entre las letras dadas hay una vocal mayúscula, dos consonantes mayúsculas, dos vocales minúsculas y cuatro consonantes minúsculas.

El número de palabras con 2 vocales y 3 consonantes que empiezan con A es:

$$C_2^1.C_4^3.P_4 \qquad\qquad A\ \_\ \_\ \_\ \_$$

El número de palabras con 2 vocales y 3 consonantes que empiezan con T es:

$$C_2^2.C_4^2.P_4 \qquad\qquad T\ \_\ \_\ \_\ \_$$

y lo mismo para las que empiezan con R. Luego la respuesta a la pregunta c) es:

$$C_2^1.C_4^3.P_4 + 2.C_2^2.C_4^2.P_4 = 480$$

5.5. NUMEROS COMBINATORIOS.

Los números de la forma $\frac{m!}{n!(m-n)!}$, con $m,n \in Z$, $0 \leq n \leq m$, se llaman números combinatorios y se representan $\binom{m}{n}$

$$\binom{m}{n} = \frac{m!}{n!(m-n)!}$$

Por ejemplo :

a) $\binom{10}{0} = \frac{10!}{0!10!} = 1$

b) $\binom{7}{4} = \frac{7!}{4!3!} = 35$

c) $\binom{7}{3} = \frac{7!}{3!4!} = 35$

d) $\binom{8}{4} = \frac{8!}{4!4!} = 70$

Dado un número combinatorio $\binom{m}{n}$, m se llama el numerador y n el denominador del mismo.

Dos números combinatorios se dicen complementarios cuando tienen igual numerador y sus denominadores sumados dan el numerador. Por ejemplo, los números combinatorios de b) y c) son complementarios y notemos que $\binom{7}{4} = \binom{7}{3}$. Además se ve también que

$$\binom{8}{4} = \binom{7}{3} + \binom{7}{4}$$

En general valen las siguientes

PROPIEDADES.

1. Dos números combinatorios complementarios son iguales.

$$\binom{m}{n} = \binom{m}{m-n}$$

2. $\binom{m}{n} = \binom{m-1}{n-1} + \binom{m-1}{n}$

3. $\sum_{k=0}^{m} \binom{m}{k} = 2^m$

Demostración:

1. $\binom{m}{m-n} = \dfrac{m!}{(m-n)!\left[m-(m-n)\right]!} = \dfrac{m!}{(m-n)!\,n!} = \binom{m}{n}$

2. $\binom{m-1}{n-1} + \binom{m-1}{n} = \dfrac{(m-1)!}{(n-1)!\left[m-1-(n-1)\right]!} + \dfrac{(m-1)!}{n!\,(m-1-n)!} = \dfrac{(m-1)!}{(n-1)!\,(m-n)!} + \dfrac{(m-1)!}{n!\,(m-n-1)!} =$

$$= \frac{n.(m-1)! + (m-n)(m-1)!}{n!\,(m-n)!} = \frac{m.(m-1)!}{n!\,(m-n)!} = \frac{m!}{n!\,(m-n)!} = \binom{m}{n}$$

La propiedad 3 se demuestra fácilmente aplicando la fórmula del binomio de Newton que veremos en el próximo parágrafo.

Aplicando la propiedad 2 se puede construir una tabla de los números combinatorios, llamada triángulo de Pascal o también triángulo de Tartaglia.

m=1					1		1					
m=2				1		2		1				
m=3			1		3		3		1			
m=4		1		4		6		4		1		
m=5	1		5		10		10		5		1	
m=6	1		6		15		20		15		6	1

. . . . . .

Los números que aparecen en la fila m-ésima de este triángulo, m = 1,2,3,....., son los números combinatorios $\binom{m}{0}$, $\binom{m}{1}$, $\binom{m}{2}$,....., $\binom{m}{m-1}$, $\binom{m}{m}$, en ese orden.

Por ejemplo, de la 4^{a} fila resulta que:

$\binom{4}{0} = 1$; $\binom{4}{1} = 4$; $\binom{4}{2} = 6$; $\binom{4}{3} = 4$; $\binom{4}{4} = 1$.

Observemos que por la propiedad 2 de los números combinatorios, cada uno de los elementos de la fila i-ésima del triángulo de Pascal , $i \geq 2$, es igual a la suma de los dos de la fila anterior situados encima suyo, excepto el primer y el último elementos que son 1. Luego, para construir la tabla basta escribir 1 , 1 en la primera fila y aplicar esa regla para escribir las demás. De la propiedad 1 de los números combinatorios resulta que el triángulo de Pascal es simétrico con respecto al eje vertical que pasa por el vértice.

COMBINACIONES CON REPETICION.

Definición. Dados m elementos distintos, se llama combinación con repetición de orden n de esos m elementos a todo conjunto formado por n elementos elegidos entre los m dados, distintos o repetidos, considerándose iguales a los formados por los mismos objetos, repetidos el mismo número de veces. El número de combinaciones con repetición de orden n de m elementos se representa C'^{n}_{m} .

Por ejemplo, dados los elementos a,b,c , las combinaciones con repetición de orden 3 de los mismos son:

a a a , b b b , c c c , a a b , a a c , a b b , a c c , a b c , b c c , b b c .

La fórmula que da C'^{n}_{m} es :

$$C'^{n}_{m} = \binom{m+n-1}{n}$$

y se demuestra por inducción.

EJERCICIO.

Hallar el número de términos de un polinomio homogéneo de 5° grado, completo, en tres variables X,Y,Z . Idem si no es homogéneo.

EJERCICIOS.

1. Calcular V_6^3 ; V_{10}^6 ; $V_8^4 - V_5^1$; V'^3_5 ; V'^4_3 ; P_4 ; P_{12} ; C_{10}^3 ; C_{20}^{15} .

2. a) Cuántos números de 5 cifras pueden escribirse con los dígitos 1,2,4,5,6 y 9 sin que haya cifras iguales en cada número?. ¿Cuántos son múltiplos de 5?. ¿Cuántos son menores que 60.000?. ¿En cuantos figuran cifras impares en los lugares impares y cifras pares en los pares?.

   R: 720 ; 120 ; 480 ; 36 .

   b) Resolver el mismo problema si se admiten números con cifras repetidas.

   R: 7776 ; 1296 ; 5184 ; 243 .

   c) Resolver el mismo problema con 0,1,2,5,6,7 y 8 , sin repetir cifras primero, y repitiéndolas después.

   R: 2160 ; 660 ; 1080 ; 72 .

   R: 14406 ; 2058 ; 7203 ; 432 .

3. Se tienen 5 discos de Mozart, 3 de Brahms y 2 de Palito Ortega.¿ De cuántas maneras distintas es posible ordenarlos en un estante si se desea que los discos de un mismo autor estén siempre juntos?.

   R: 8640 .

4. En una familia compuesta de padre, madre y seis hijas, se decide que las mujeres lavarán los platos una noche de la semana cada una. ¿De cuántas maneras distintas se pueden organizar los turnos?. ¿De cuántas si la hija menor no quiere lavar los platos los domingos?.

   R: 5040 ; 4320 .

5. ¿Cuántas manos distintas puede recibir un jugador de truco al comenzar el juego? ¿Cuántas estarán formadas por dos ases y un tres?. ¿En cuántas figurará el as de espadas?. ¿En cuántas las cartas serán todas del mismo palo?. (Las barajas españolas son 40, el jugador recibe 3 cartas y no interesa el orden en que las recibe).

   R: 9880 ; 24 ; 741 ; 480 .

6. ¿De cuántas maneras distintas pueden ordenarse las letras de la palabra RESUCITADO de modo que no haya dos consonantes juntas?. ¿Cuántas de esas ordenaciones empiezan con vocal?.

   R: 28800 ; 14400 .

7. ¿Cuántos polinomios distintos se pueden obtener reordenando los coeficientes del polinomio $3X^7 - 6X^6 + X^5 + 3X^4 + X^3 + X^2 + 3X - 6$?. ¿Cuántos de ellos son mónicos?.

   R: 560 ; 210 .

8. Averiguar m sabiendo que:

   a) $V_6^3 - V_m^2 = (P_4 + 3)m$; b) $35\ C_m^6 = 14\ C_m^4$

9. En una reunión, después que cada uno de los asistentes estrechó una vez la mano de cada uno de los demás, se averiguó que se habían intercambiado 45 apretones de manos en total. ¿Cuántos eran los asistentes?.

   R: 10 .

10. ¿Cuántos boletos capicúa hay en un rollo de boletos de ómnibus que empieza con el número 10000 y termina en 99999?.

    R: 900 .

11. Un bote de 8 remos va a ser tripulado por un grupo seleccionado de 11 hombres de los cuales 3 pueden llevar el timón pero no remar, y el resto puede remar pero no llevar el timón. ¿De cuántas maneras puede ordenarse el grupo si 2 de los hombres sólo pueden remar en uno de los lados?.

    R: 25920 .

12. ¿De cuántas maneras es posible alinear 14 signos + y 8 signos - sin que haya dos signos - juntos?.

    R: 6435

13. En un edificio de 9 pisos viajan en el ascensor 5 personas. ¿De cuántas maneras diferentes pueden bajarse esas 5 personas?. ¿De cuántas si no bajan dos en el mismo piso?.

    R: 59049 ; 15120 .

14. Dados en el plano 18 puntos:

    a) ¿Cuántos segmentos determinan?.

    b) Suponiendo que no hay tres que estén alineados, cuántos triángulos determinan?.

    c) Suponiendo que no hay tres que estén alineados, excepto cinco que lo están, cuántas rectas y cuántos triángulos determinan?.

    R: 153 ; 816 ; 144 ; 806 .

5.6. POTENCIA DE UN BINOMIO.

Vamos a ver una fórmula que permite calcular expresiones del tipo $(a+b)^n$, $n\in Z$, $n\geq 0$. Los desarrollos para $n=1$, $n=2$ y $n=3$, que el lector conoce del colegio secundario, se pueden escribir como sigue:

$$(a+b)^1 = a+b = \binom{1}{0} a + \binom{1}{1} b$$

$$(a+b)^2 = a^2+ 2ab + b^2 = \binom{2}{0} a^2 + \binom{2}{1} ab + \binom{2}{2} b^2$$

$$(a+b)^3 = a^3+ 3a^2b + 3ab^2 + b^3 = \binom{3}{0} a^3 + \binom{3}{1} a^2b + \binom{3}{2} ab^2 + \binom{3}{3} b^3$$

En general, vamos a demostrar que:

TEOREMA 5.4. Cualquiera sea el entero no negativo n ,

$$(a+b)^n = \binom{n}{0} a^n b^o + \binom{n}{1} a^{n-1} b^1 + \binom{n}{2} a^{n-2} b^2 + \ldots + \binom{n}{k} a^{n-k} b^k + \ldots \ldots \ldots$$

$$\ldots + \binom{n}{n} a^o b^n = \sum_{k=o}^{n} \binom{n}{k} a^{n-k} b^k \qquad (1)$$

Demostración: Lo demostraremos por inducción sobre n.

Para $n=0$ y $n=1$ la fórmula (1) se verifica.

Sea $n > 1$, supongámosla verdadera para $n-1$ y probémosla para n.

$$(a+b)^n = (a+b).(a+b)^{n-1} = (a+b).\sum_{k=o}^{n-1} \binom{n-1}{k} a^{n-k-1} b^k =$$

$$= \sum_{k=o}^{n-1} \binom{n-1}{k} a^{n-k} b^k + \sum_{k=o}^{n-1} \binom{n-1}{k} a^{n-k-1} b^{k+1} =$$

$$= \binom{n-1}{0} a^n b^o + \sum_{k=1}^{n-1} \binom{n-1}{k} a^{n-k} b^k + \sum_{k=o}^{n-2} \binom{n-1}{k} a^{n-k-1} b^{k+1} + \binom{n-1}{n-1} a^o b^n$$

Cambiando en la segunda sumatoria el índice variable k por $k-1$ queda

$$\sum_{k=1}^{n-1} \binom{n-1}{k-1} a^{n-k} b^k$$

y sumando los términos semejantes de las dos sumatorias se tiene:

$$(a+b)^n = a^n + \sum_{k=1}^{n-1} \left[\binom{n-1}{k} + \binom{n-1}{k-1} \right] a^{n-k}b^k + b^n = a^n + \sum_{k=1}^{n-1} \binom{n}{k} a^{n-k}b^k + b^n =$$

$$= \sum_{k=o}^{n} \binom{n}{k} a^{n-k}b^k \qquad \text{c.q.d.}$$

La fórmula (1) se llama fórmula del binomio de Newton.

EJEMPLOS:

1) Desarrollar $(x^2 + 3y^{-1})^4$.

Aplicando la fórmula se tiene:

$$(x^2 + 3y^{-1})^4 = \sum_{k=o}^{4} \binom{4}{k} (x^2)^{4-k}(3y^{-1})^k = \binom{4}{0} (x^2)^4 + \binom{4}{1} (x^2)^3 (3y^{-1}) +$$

$$+ \binom{4}{2} (x^2)^2 (3y^{-1})^2 + \binom{4}{3} x^2 (3y^{-1})^3 + \binom{4}{4} (3y^{-1})^4$$

La forma más fácil de calcular los coeficientes $\binom{4}{k}$ es usando el triángulo de Pascal. Haciendo cuentas resulta:

$$(x^2 + 3y^{-1})^4 = x^8 + 12x^6y^{-1} + 54x^4y^{-2} + 108\, x^2y^{-3} + 81y^{-4}$$

2) Desarrollar $(2a^{-3} - \frac{1}{2} at^2)^5$.

$$(2a^{-3} - \tfrac{1}{2} at^2)^5 = \sum_{k=o}^{5} \binom{5}{k} (2a^{-3})^{5-k} (-\tfrac{1}{2} at^2)^k =$$

$$= \binom{5}{0} (2a^{-3})^5 + \binom{5}{1} (2a^{-3})^4 (-\tfrac{1}{2} at^2) + \binom{5}{2} (2a^{-3})^3 (-\tfrac{1}{2} at^2)^2 +$$

$$+ \binom{5}{3} (2a^{-3})^2 (-\tfrac{1}{2} at^2)^3 + \binom{5}{4} (2a^{-3}) (-\tfrac{1}{2} at^2)^4 + \binom{5}{5} (-\tfrac{1}{2} at^2)^5 =$$

$$= 32\, a^{-15} - 40\, a^{-11}t^2 + 20\, a^{-7}t^4 - 5\, a^{-3}\, t^6 + \tfrac{5}{8}\, a\, t^8 - \tfrac{1}{32}\, a^5\, t^{10}$$

EJERCICIOS.

1. a) Desarrollar:

$$(3x+2y)^8 \qquad ; \qquad (1-3a^2)^6 \qquad ; \qquad (\tfrac{1}{2} x^2 - xy^{-3})^7$$

b) Calcular:

$(2+\sqrt{3})^5 + (2-\sqrt{3})^5$; $(1+3i)^3 - (1-3i)^3$

2. a) Escribir el 4° término del desarrollo de $(2x + \frac{1}{3})^{16}$

" " 5° " " " " $(x - \frac{1}{x})^{13}$

" " 13° " " " " $(\frac{a^2}{2} - 2b)^{17}$

" " 9° " " " " $(\frac{1}{3}x^2 - 3x^{-1})^{11}$

b) Hallar el coeficiente de:

x^{18} en el desarrollo de $(x^2 + \frac{3a}{x})^{15}$; de x^{18} en el de $(ax^4 - bx)^9$;

de t^{32} y t^{-17} en el de $(t^4 - \frac{1}{t^3})^{15}$.

c) Hallar el término independiente de x en el desarrollo de $(\sqrt{x} + \frac{1}{3x^2})^{10}$.

3. Demostrar que:

a) $\binom{n}{0} + \binom{n}{1} + \ldots\ldots + \binom{n}{n} = 2^n$

b) $\binom{n}{1} + \binom{n}{3} + \binom{n}{5} + \ldots\ldots = 2^{n-1}$

c) $\binom{n}{0} + \binom{n}{2} + \binom{n}{4} + \ldots\ldots = 2^{n-1}$

4. Demostrar que el número de subconjuntos de un conjunto con n elementos es 2^n . Si A es un conjunto con 10 elementos, cuántos elementos tiene el conjunto $P(A)$? .

5.7. CLASE DE UNA PERMUTACION.

Vamos a ver ahora un resultado que utilizaremos más adelante cuando estudiemos determinantes.

Sea A un conjunto finito de n elementos, dados en un cierto orden: $a_1, a_2, \ldots, a_n$

Consideremos las n! permutaciones de estos elementos. La permutación en que figuran en el orden dado se llama permutación principal.

Como lo único que nos va a interesar es el orden relativo de los elementos, se puede representar a los elementos dados $a_1, a_2, \ldots, a_n$ por sus subíndices y considerar directamente las permutaciones de los n primeros números naturales. La permutación principal es así: 1 2 . . . . . n .

En una permutación cualquiera, se dice que dos elementos forman sucesión o inversión según que aparezcan o no en el mismo orden, uno con respecto al otro, que en la permutación principal.

Por ejemplo, en la permutación

3 5 1 2 6 4 7

forman inversión 3 y 1 , 3 y 2 , 5 y 1 , 6 y 4. 5,2 ~ 5,4

El número total de inversiones que presenta una permutación se obtiene comparando cada elemento con los que lo siguen. En nuestro ejemplo, la permutación presenta ~~cuatro~~ 6 inversiones.

Definición. Una permutación se dice de clase par o impar según que el número total de inversiones que presenta sea par o impar.

Así , la permutación

1 2 3 . . . . . n

es par. La permutación del ejemplo anterior también es par. En cambio, la permutación

3 6 1 2 5 4 7

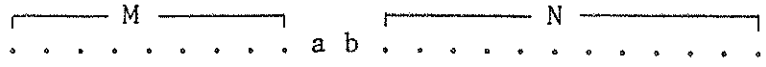
es impar, pues presenta en total siete inversiones.

Notemos que esta permutación se obtiene de la anterior intercambiando el 5 con el 6, y dejando fijos a los demás elementos. En general, si en una permutación se intercambian dos elementos cualesquiera, dejando fijos a los demás, se obtiene una nueva permutación. Se dice que ésta se obtiene de la primera mediante una trasposición y vale la siguiente propiedad:

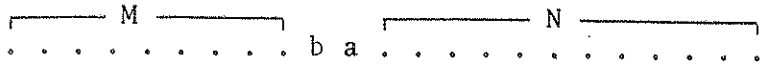
TEOREMA 5.5. Si en una permutación se trasponen dos elementos cualesquiera, cambia

la clase de la permutación.

Demostración: Supongamos primero que los elementos $\underline{a}$, $\underline{b}$ que se trasponen son consecutivos.

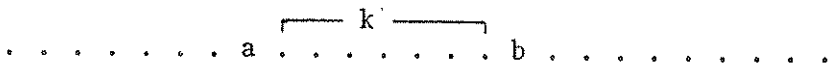


donde M es el conjunto de los elementos que preceden a $\underline{a}$ y N el de los que siguen a $\underline{b}$.

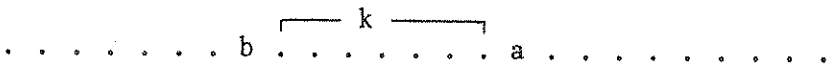


Al trasponer $\underline{a}$ con $\underline{b}$ se conservan todas las inversiones que forman los elementos de M con los que los siguen, las que forma $\underline{a}$ con los elementos de N, las que forma $\underline{b}$ con los elementos de N, las que forman los elementos de N entre sí. Lo único que ha cambiado es la situación relativa de $\underline{a}$ con $\underline{b}$: si antes formaban sucesión ahora forman inversión, y si antes formaban inversión, ahora forman sucesión. Por lo tanto el número total de inversiones aumenta o disminuye en 1, es decir, la permutación cambia de clase.

Supongamos ahora que entre los elementos $\underline{a}$ y $\underline{b}$ que se trasponen hay k elementos



Se puede llevar $\underline{a}$ al lugar anterior al de $\underline{b}$ trasponiendo $\underline{a}$ con los k elementos intermedios y luego, trasponiendo $\underline{b}$ con $\underline{a}$ y con esos k elementos, obtener la permutación deseada :



En total se han efectuado $k+k+1 = 2k+1$ trasposiciones, lo que significa un número impar de cambios de clase. Luego, la permutación que se obtiene es de clase distinta a la dada.

Para terminar este capítulo vamos a ver, a título informativo, el concepto de sustitución, que guarda una estrecha relación con el de permutación.

5.8. EL GRUPO SIMETRICO S_n .

Se llama una sustitución de orden n (también una permutación de orden n) a toda biyección del conjunto de los n primeros números naturales $\{1,2,....,n\}$ sobre sí mismo.

Para representar una sustitución f se escriben en una fila los números 1,2,.....,n y debajo de cada uno de ellos su imagen a_i por f.

$$f = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ a_1 & a_2 & a_3 & \ldots & a_n \end{pmatrix}$$

Por ejemplo, para notar la sustitución f de orden 6 tal que $f(1) = 3$, $f(2) = 1$, $f(3) = 5$, $f(4) = 4$, $f(5) = 6$ y $f(6) = 2$, se escribe:

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 1 & 5 & 4 & 6 & 2 \end{pmatrix}$$

Los números 1,2,....,n en la primera fila no tienen porqué estar en ese orden. Por ejemplo, la sustitución f anterior también puede escribirse:

$$\begin{pmatrix} 6 & 5 & 4 & 3 & 2 & 1 \\ 2 & 6 & 4 & 5 & 1 & 3 \end{pmatrix} \quad , \quad \begin{pmatrix} 3 & 6 & 5 & 1 & 4 & 2 \\ 5 & 2 & 6 & 3 & 4 & 5 \end{pmatrix} \quad , \quad \text{etc.}$$

Lo único que importa es que en la primera fila figuren todos los números de 1 a n y en la segunda sus respectivas imágenes.

La sustitución

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

se llama la sustitución idéntica de orden n.

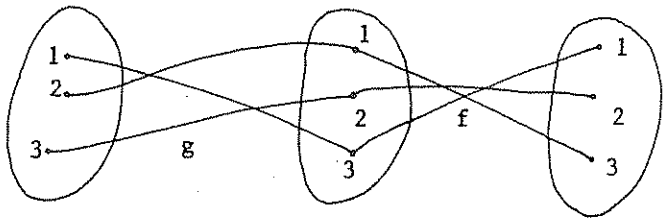
Consideremos el conjunto S_n de todas las sustituciones de orden n, es decir, la co - lección de todas las biyecciones del conjunto $\{1,2,....,n\}$ sobre sí mismo. Es claro que hay n! de ellas. Por ejemplo, S_3 está formado por las sustituciones:

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} , \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} , \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} , \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} , \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} , \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

Dadas dos sustituciones de orden n se puede hablar de su composición y se verifica que la composición de dos sustituciones de orden n es una sustitución de orden n.

Así, dadas $f,g \in S_3$, $f = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$, $g = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$ es

$f \circ g = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$



Entonces, la composición de aplicaciones define una operación binaria en S_n :

$$(f,g) \longrightarrow fg \quad , \quad \text{donde} \quad fg \in S_n \quad \text{es tal que}$$

$$(fg)(j) = f(g(j)) \quad , \quad \forall \; j=1,2,\ldots,n$$

Por ejemplo , si $f,g \in S_8$,

$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 6 & 1 & 5 & 8 & 2 & 7 & 3 \end{pmatrix}$, $g = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 7 & 2 & 6 & 3 & 5 & 1 & 8 & 4 \end{pmatrix}$, es:

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

Esta operación definida en S_n tiene las siguientes propiedades:

1. <u>Es asociativa</u>: $(fg)h = f(gh)$, $\forall$ $f,g,h \in S_n$, pues la composición de aplicaciones lo es.

2. <u>Existe elemento neutro</u>: es la sustitución $I = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ 1 & 2 & 3 & \ldots & n \end{pmatrix}$ pues

   $fI = If = f$, $\forall$ $f \in S_n$.

3. <u>Toda sustitución $f \in S_n$ es inversible</u> , es decir, existe $f^{-1} \in S_n$ tal que $ff^{-1} = f^{-1}f = I$.

   En efecto, si $f = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ a_1 & a_2 & a_3 & \ldots & a_n \end{pmatrix}$, la sustitución

$$f^{-1} = \begin{pmatrix} a_1 & a_2 & a_3 & \ldots & a_n \\ 1 & 2 & 3 & \ldots & n \end{pmatrix} \quad \text{es inversa de } f.$$

Por ejemplo , dada

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 5 & 4 & 7 & 2 & 6 & 1 \end{pmatrix} \in S_7 \quad \text{su inversa es} \quad f^{-1} = \begin{pmatrix} 3 & 5 & 4 & 7 & 2 & 6 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{pmatrix}$$

$$\text{pues} \quad ff^{-1} = f^{-1}f = I$$

S_n con esta operación se llama el grupo simétrico de orden n.

Observemos que el producto de sustituciones no es conmutativo si $n > 2$. En efecto,

sean por ejemplo, $f = \begin{pmatrix} 1 & 2 & 3 & 4 & \dots & n \\ 2 & 1 & 3 & 4 & \dots & n \end{pmatrix}$, $g = \begin{pmatrix} 1 & 2 & 3 & 4 & \dots & n \\ 3 & 2 & 1 & 4 & \dots & n \end{pmatrix}$

Entonces $fg = \begin{pmatrix} 1 & 2 & 3 & 4 & \dots & n \\ 3 & 1 & 2 & 4 & \dots & n \end{pmatrix}$, $gf = \begin{pmatrix} 1 & 2 & 3 & 4 & \dots & n \\ 2 & 3 & 1 & 4 & \dots & n \end{pmatrix}$

y $fg \neq gf$.

Luego S_n es un grupo no abeliano, si $n \geq 3$.

DESCOMPOSICION DE UNA SUSTITUCION EN CICLOS.

Dada una sustitución $f \in S_n$ diremos que f mueve al elemento $j \in \{1,2,\dots,n\}$ si $f(j) \neq j$, y que lo deja fijo si $f(j) = j$.

Definición. Una sustitución $f \in S_n$ se dice un ciclo de longitud k o un k-ciclo si deja fijos n-k elementos del conjunto $\{1,2,\dots,n\}$ y permuta circularmente a los k elementos restantes $a_1, a_2, \dots, a_k$, es decir, si f es de la forma

$$f = \begin{pmatrix} a_1 & a_2 & \dots & a_{k-1} & a_k & a_{k+1} & \dots & a_n \\ a_2 & a_3 & \dots & a_k & a_1 & a_{k+1} & \dots & a_n \end{pmatrix}$$

Para notar un ciclo se escribe simplemente $f = (a_1 a_2 \dots a_k)$. Observemos que en esta notación se omiten los elementos que quedan fijos.

Por ejemplo, las siguientes sustituciones de S_6 son ciclos:

$$\begin{pmatrix} 1 & 4 & 3 & 6 & 2 & 5 \\ 4 & 3 & 6 & 1 & 2 & 5 \end{pmatrix}, \begin{pmatrix} 3 & 5 & 4 & 1 & 2 & 6 \\ 5 & 4 & 3 & 1 & 2 & 6 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 6 & 4 & 5 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 5 & 3 & 4 & 2 & 6 \end{pmatrix}$$

Se pueden escribir respectivamente: (1436) , (354) , (134526) , (25) .

Definición. Dos sustituciones $f,g \in S_n$ se dicen disjuntas si el conjunto de los elementos que mueve f es disjunto del conjunto de elementos que mueve g.

Por ejemplo, los ciclos (3564) y (27) de S_7 son disjuntos.

Es claro que el producto de dos sustituciones disjuntas es conmutativo, en particular, el producto de dos ciclos disjuntos es conmutativo.

Se demuestra que: Toda sustitución $f \in S_n$, $f \neq I$, se puede escribir como producto de ciclos disjuntos de longitud ≥ 2 y esta descomposición es única salvo el orden de los factores.

No haremos la demostración, pero en líneas generales se razona como en el ejemplo siguiente:

Sea $f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 7 & 6 & 4 & 9 & 1 & 2 & 5 & 8 \end{pmatrix}$. Teniendo en cuenta los elementos que mueve f, se ve que $f(1) = 3$, $f(3) = 6$, $f(6) = 1$. Luego f transforma a los elementos 1,3,6 igual que el ciclo (136). Como $f(2) = 7$, $f(7) = 2$, con respecto a los elementos 2,7 la sustitución f actúa igual que el ciclo (27). Por último , de $f(5) = 9$, $f(9) = 8$ y $f(8) = 5$ se concluye que f transforma a estos elementos igual que el ciclo (598). Entonces se ve que f se obtiene componiendo estos ciclos, en cualquier orden puesto que son disjuntos.

$$f = (136)(27)(598)$$

El elemento 4 no figura porque f lo deja fijo.

EJEMPLOS.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 5 & 9 & 3 & 10 & 1 & 8 & 2 & 4 & 7 & 6 \end{pmatrix} = (15)(297)(4\ 10\ 6\ 8)$$

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 7 & 1 & 3 & 6 & 5 & 4 & 2 & 8 \end{pmatrix} = (172)(46)$$

Definición. Un ciclo de longitud 2 se llama una trasposición.

Observemos que $(ab)^{-1} = (ab)$ puesto que $(ab)^2 = I$

Todo ciclo puede escribirse como producto de trasposiciones, aunque esta descomposición no es única.

Por ejemplo , $(13264) = (14)(16)(12)(13)$

pero también se puede escribir $(13264) = (46)(42)(43)(41)$

En general

$$(a_1 a_2 \ldots\ldots a_k) = (a_1 a_k)(a_1 a_{k-1}) \ldots\ldots (a_1 a_2)$$

Como toda sustitución es producto de ciclos y todo ciclo es producto de trasposiciones resulta que:

Toda sustitución de S_n es producto de trasposiciones.

En particular, $I = (ab)(ab)$.

No sólo los factores de la descomposición en trasposiciones de una sustitución dada no están unívocamente determinados, sino que tampoco el número de trasposiciones en dos descomposiciones distintas es necesariamente el mismo. Por ejemplo, sea $f \in S_6$

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 3 & 6 & 4 & 1 & 2 \end{pmatrix} = (15)(236) = (15)(26)(23)$$

Se pueden agregar nuevas trasposiciones sin alterar el resultado. Por ejemplo

$$f = (15)(14)(14)(26)(23)$$

Si bien el número de trasposiciones puede variar, se conserva la paridad de ese número; es decir, para una sustitución dada el número de trasposiciones que figuran en cualquier descomposición es siempre par o impar. (Cauchy , 1789-1857).

De aquí la siguiente definición

Definición. Una sustitución $f \in S_n$ se dice de clase par o impar según que el número de trasposiciones que figuran en una descomposición cualquiera sea par o impar.

EJEMPLOS.

La sustitución idéntica es de clase par; toda trasposición es de clase impar; todo 3-ciclo es de clase par; la sustitución

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 2 & 8 & 6 & 7 & 1 & 5 & 3 \end{pmatrix}$$ es de clase par pues

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 2 & 8 & 6 & 7 & 1 & 5 & 3 \end{pmatrix} = (146)(38)(57) = (16)(14)(38)(57)$$

Observemos que el producto de dos sustituciones de la misma clase (par o impar) es una sustitución par y el producto de dos de distinta clase es una sustitución impar. Además f y f^{-1} son de la misma clase.

Es evidente que hay una estrecha relación entre las nociones de permutación y susti-

tución.

Dada una sustitución de orden n, si suponemos que en la notación adoptada los números de la fila superior aparecen en el orden natural 1,2,....,n , entonces es claro que cada sustitución queda determinada por la permutación que aparece en la segunda fila y existe así una correspondencia biunívoca entre las sustituciones de orden n y las permutaciones de los n primeros números naturales.

Además, la clase de una sustitución puede hallarse sin descomponerla en trasposiciones. Observemos en primer lugar que, dada una sustitución $f \in S_n$, las permutaciones que aparecen en la primera y en la segunda filas de f son siempre de la misma o de distinta clase, no importa la forma en que se escriba la sustitución, es decir , el orden en que aparezcan los números 1,2,....,n en la fila superior.

Por ejemplo,

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 5 & 6 & 1 & 2 & 4 \end{pmatrix} \qquad y \qquad \begin{pmatrix} 3 & 6 & 5 & 1 & 2 & 4 \\ 6 & 4 & 2 & 3 & 5 & 1 \end{pmatrix}$$

representan a la misma sustitución y en ambas notaciones las permutaciones superior e inferior son de la misma clase: ambas pares en el primer caso, y ambas impares en el segundo. Esto sucede porque dada una expresión arbitraria de una sustitución f , otra expresión cualquiera de f se puede obtener de la dada mediante unas cuantas trasposiciones en la fila superior y las correspondientes en la fila inferior. Como al trasponer dos elementos en una permutación, ésta cambia de clase, se produce el mismo número de cambios de clase en la permutación superior que en la inferior . De modo que, si eran de la misma clase siguen siéndolo, y si eran de distinta clase, también.

Vale la siguiente propiedad: Una sustitución $f \in S_n$ es de clase par o impar según que las permutaciones que aparecen en la primera y segunda filas de una expresión cualquiera de f, sean ambas de la misma o de distinta clase.

En virtud de nuestra observación, podemos suponer a las sustituciones escritas siempre en la forma

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ a_1 & a_2 & \cdots & a_n \end{pmatrix} \qquad (1)$$

Entonces la propiedad anterior se enuncia:

Una sustitución $f \in S_n$ es de clase par o impar según que, escrita en la forma (1) , el número de inversiones que presenta la permutación de la segunda fila sea par o impar.

Esta propiedad es consecuencia inmediata del siguiente:

TEOREMA 5.6. Dada una sustitución $f \in S_n$ escrita en la forma (1), el número de factores que aparece en una descomposición cualquiera de f en trasposiciones tiene

la misma paridad que el número de inversiones de la permutación de la segunda fila.

Demostración: Supongamos que ν es el número de inversiones que presenta la permutación de la segunda fila y sea $f = t_s \ldots\ldots t_2 t_1$ una descomposición de f en producto de trasposiciones. Podemos considerar entonces que f se obtiene de la sustitución idéntica I aplicándole la sucesión de trasposiciones $t_1, t_2, \ldots, t_s$. Cada vez que se aplica una trasposición t_i , se trasponen dos elementos en la segunda fila. Como la permutación que figura en la segunda fila de I es 1 2 .....n , al cabo del proceso se habrán producido s cambios de clase de esa permutación, que es par. Es claro entonces que la paridad de s debe coincidir con la de ν .

Por ejemplo, la sustitución

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

es de clase impar pues la permutación de la segunda fila tiene 13 inversiones.

OBSERVACION. Notemos que en el teorema anterior se demuestra que, para una sustitución dada, el número de trasposiciones que figuran en una descomposición cualquiera de esa sustitución en producto de trasposiciones es siempre par o impar, propiedad que habíamos mencionado sin demostración.

EJERCICIOS.

1. a) Escribir todos los elementos de S_2 , S_3 y S_4 . ¿Cuántos elementos tiene S_n ?.

   b) Dadas en S_7 las siguientes permutaciones:

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 6 & 1 & 7 & 3 & 4 & 5 \end{pmatrix}, \quad g = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 5 & 2 & 1 & 6 & 7 & 3 & 4 \end{pmatrix}, \quad h = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 5 & 7 & 3 & 1 & 5 & 4 & 6 \end{pmatrix}$$

   hallar fg , gf , hf , g^{-1} , hg^{-1} , h^{-1} , h^2 , f^3 , hfh^{-1} .

2. a) Descomponer en producto de ciclos disjuntos y en producto de trasposiciones las siguientes permutaciones y decir de qué clase son:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 6 & 7 & 5 & 9 & 2 & 1 & 4 & 8 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 4 & 3 & 6 & 1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 5 & 1 & 3 & 7 & 8 & 9 & 4 & 2 & 10 & 6 \end{pmatrix}$$

   b) Idem para las permutaciones:

$$f = (132)(23)(24) \in S_4$$
$$g = (1386)(361)(418) \in S_8$$
$$h = (13)(15)(234)(23) \in S_5$$

   c) Hallar f^{-1} , g^{-1} , h^{-1}

CAPITULO VI

SISTEMAS DE ECUACIONES LINEALES , MATRICES Y DETERMINANTES

En este capítulo K es el cuerpo de los números racionales, el de los números reales o el de los números complejos (K = Q , K = R ó K = C) y llamaremos escalares a los elementos de K.

6.1. SISTEMAS DE ECUACIONES LINEALES.

Una ecuación lineal o de primer grado en n incógnitas $x_1 , x_2 , \ldots , x_n$ es una expresión de la forma

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b \qquad (1)$$

donde $a_1 , a_2 , \ldots , a_n , b$ son números de K dados y $x_1 , x_2 , \ldots , x_n$ símbolos.

$a_1 , a_2 , \ldots , a_n$ se llaman los coeficientes de la ecuación y b el término independiente.

Se llama solución de la ecuación a toda n-upla $(k_1 , k_2 , \ldots , k_n)$ de números de K que reemplazados ordenadamente en lugar de las incógnitas $x_1 , x_2 , \ldots , x_n$ convierten a la expresión (1) en una identidad. Se dice que $k_1 , k_2 , \ldots , k_n$ satisfacen la ecuación.

Por ejemplo, dada la ecuación lineal

$$2x_1 - 3x_2 + \sqrt{2}x_3 - x_4 = 7$$

dos soluciones de la misma son $(1,-1,\sqrt{2},0)$ y $(-2, \frac{1}{3} ,0,12)$.

Vamos a estudiar los sistemas de ecuaciones lineales y su resolución.

El lector sabe resolver sistemas de dos ecuaciones lineales con dos incógnitas, y de tres ecuaciones lineales con tres incógnitas desde el colegio secundario, donde aprendió varios métodos para calcular las soluciones.

Ahora nos proponemos indicar un método para resolver en general sistemas de m ecuaciones lineales con n incógnitas, con m , n arbitrarios.

Para escribir un sistema de m ecuaciones lineales con n incógnitas adoptaremos la siguiente notación:

$$6.1 \qquad \begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots\ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots\ldots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots\ldots + a_{mn}x_n = b_m \end{cases}$$

Según esta notación, a_{ij} es el coeficiente de la incógnita x_j en la i-ésima ecuación, $1 \leq i \leq m$, $1 \leq j \leq n$; es decir, el primer subíndice indica la ecuación a la que pertenece a_{ij} y el segundo la incógnita de la que es coeficiente. b_1 , b_2 , ... , b_m son los términos independientes.

Si $b_1 = b_2 = \ldots\ldots = b_m = 0$ el sistema se llama <u>homogéneo</u>.

Por ejemplo ,

$$\begin{cases} 3x_1 - 2x_2 + x_3 + \frac{1}{2}x_4 + x_5 = -1 \\ x_1 + 8x_3 - x_4 = 0 \\ 5x_2 - x_3 + \frac{3}{4}x_4 = 7 \\ x_1 + x_2 + 3x_3 - x_4 + 2x_5 = 3 \end{cases} \quad , \quad \begin{cases} x_1 + x_2 - 2x_3 = 0 \\ 2x_1 + x_3 = 0 \\ 3x_1 + 5x_2 - x_3 = 0 \end{cases}$$

El primero es un sistema de 4 ecuaciones con 5 incógnitas y el segundo un sistema homogéneo con tantas ecuaciones como incógnitas.

Una solución del sistema 6.1. es una n-upla $(k_1, k_2, \ldots, k_n)$ de números de K tal que reemplazando x_1 por k_1 , x_2 por k_2 , ... , x_n por k_n se satisfacen <u>simultáneamente</u> todas las ecuaciones.

Puede suceder que un sistema de ecuaciones lineales no tenga ninguna solución, es decir, que no exista ningún conjunto de n números en esas condiciones, caso en que el sistema se dice <u>incompatible</u>. Si tiene solución se dice <u>compatible</u> , y <u>determinado</u> o <u>indeterminado</u> según que tenga una única solución o más de una.

<u>EJEMPLOS</u>.

1)

$$\begin{cases} 2x - y = -7 \\ x + 5y = 13 \end{cases} \quad , \quad \begin{cases} 2x - y = 0 \\ 2x - y = 8 \end{cases} \quad , \quad \begin{cases} x - 3y = 1 \\ 4x - 12y = 4 \end{cases}$$

El primer sistema es compatible determinado: su única solución es (-2,3). El segundo es claramente incompatible. El tercero es compatible indeterminado : $(k , \frac{1}{3}k + 1)$ es solución del sistema , $\forall\ k \in K$.

Geométricamente cada uno de estos tres sistemas puede interpretarse como dos rectas del plano, incidentes en el primer caso, paralelas en el segundo y coincidentes en el tercero. Se ve claramente que las soluciones son una, ninguna e infi-

nitas respectivamente.

2) Todo sistema lineal homogéneo es compatible puesto que $(0,0,\ldots,0)$ es una solución, que se llama la solución trivial.

La resolución de un sistema de ecuaciones lineales consiste en decidir si es compatible o no, y en caso de serlo, calcular todas las soluciones del sistema dado.

Vamos a indicar un método práctico para resolver un sistema de ecuaciones lineales, llamado el método de eliminación de Gauss, que consiste en eliminar sucesivamente las incógnitas mediante la aplicación reiterada de tres tipos de operaciones:

I. Intercambiar dos ecuaciones entre sí.

II. Reemplazar una ecuación por la que se obtiene multiplicándola por un escalar no nulo.

III. Reemplazar una ecuación por la que se obtiene sumando a dicha ecuación otra previamente multiplicada por un escalar.

Veamos un ejemplo para dar una idea del método. Consideremos el sistema

$$S \quad \begin{cases} 2x_1 + 3x_2 + x_3 = 1 \\ 4x_1 + 11x_2 - 5x_3 = -5 \\ 2x_1 + x_2 + 2x_3 = -7 \end{cases}$$

Podemos reducir a 1 el coeficiente de x_1 en la primera ecuación multiplicándola por $\frac{1}{2}$. (operación II).

$$\begin{cases} x_1 + \frac{3}{2}x_2 + \frac{1}{2}x_3 = \frac{1}{2} \\ 4x_1 + 11x_2 - 5x_3 = -5 \\ 2x_1 + x_2 + 2x_3 = -7 \end{cases}$$

Sumando ahora a la segunda ecuación la primera multiplicada por -4 y a la tercera la primera multiplicada por -2 (operación III), se obtiene:

$$\begin{cases} x_1 + \frac{3}{2}x_2 + \frac{1}{2}x_3 = \frac{1}{2} \\ 5x_2 - 7x_3 = -7 \\ -2x_2 + x_3 = -8 \end{cases}$$

Multiplicando ahora la segunda ecuación por $\frac{1}{5}$, y sumándole a la tercera la segunda así obtenida multiplicada por 2 se tiene:

$$S' \quad \begin{cases} x_1 + \frac{3}{2}x_2 + \frac{1}{2}x_3 = \frac{1}{2} \\ x_2 - \frac{7}{5}x_3 = -\frac{7}{5} \\ -\frac{9}{5}x_3 = -\frac{54}{5} \end{cases}$$

De aquí es muy sencillo despejar x_3 de la tercera ecuación, reemplazar este valor en la segunda para obtener x_2 y finalmente calcular x_1 de la primera, obteniéndose la solución (-13,7,6) , que evidentemente es única.

Si en el sistema dado, en lugar de la tercera ecuación figurara la siguiente : $x_1 - x_2 + 4x_3 = 4$, entonces, escribiendo en el sistema esta última ecuación en primer lugar (operación I) y siguiendo el procedimiento anterior se tiene:

$$S \quad \begin{cases} x_1 - x_2 + 4x_3 = 4 \\ 2x_1 + 3x_2 + x_3 = 1 \\ 4x_1 + 11x_2 - 5x_3 = -5 \end{cases} \longrightarrow \begin{cases} x_1 - x_2 + 4x_3 = 4 \\ 5x_2 - 7x_3 = -7 \\ 15x_2 - 21x_3 = -21 \end{cases} \longrightarrow$$

$$\begin{cases} x_1 - x_2 + 4x_3 = 4 \\ x_2 - \frac{7}{5} x_3 = -\frac{7}{5} \\ 0x_1 + 0x_2 + 0x_3 = 0 \end{cases} \longrightarrow \quad S' \quad \begin{cases} x_1 - x_2 + 4x_3 = 4 \\ x_2 - \frac{7}{5} x_3 = -\frac{7}{5} \end{cases}$$

Dándole valores numéricos a x_3 quedan determinados valores de x_1 y x_2 , es decir, para cada valor $k \in K$ de x_3 corresponden los valores: $x_1 = \frac{13}{5} - \frac{13}{5} k$, $x_2 = -\frac{7}{5} + \frac{7}{5} k$. Luego las soluciones del sistema S' son todas las ternas $(\frac{13}{5} - \frac{13}{5} k , -\frac{7}{5} + \frac{7}{5} k , k)$, con $k \in K$.

Estos ejemplos proporcionan una idea clara del procedimiento a seguir. En general , dado un sistema S de m ecuaciones lineales con n incógnitas como 6.1 , el método de Gauss consiste en pasar del sistema S a un sistema de ecuaciones lineales S' de más fácil resolución, mediante la aplicación de una sucesión finita de operaciones del tipo I , II y III , convenientemente elegidas. Se eliminan sucesivamente las incógnitas, obteniéndose sistemas intermedios, hasta llegar a un sistema S' de la forma :

$$S' \quad \begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots\ldots\ldots\ldots\ldots + a_{1n}x_n = b_1 \\ a'_{22}x_2 + a'_{23}x_3 + \ldots\ldots\ldots\ldots\ldots + a'_{2n}x_n = b'_2 \\ a''_{33}x_3 + \ldots\ldots\ldots\ldots\ldots + a''_{3n}x_n = b''_3 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{rr}^{(r-1)}x_r + \ldots\ldots + a_{rn}^{(r-1)}x_n = b_r^{(r-1)} \end{cases}$$

con $a_{11} \neq 0$, $a'_{22} \neq 0$, $a''_{33} \neq 0$, ... , $a_{rr}^{(r-1)} \neq 0$, $r \leq m$, y donde las incógnitas figuran eventualmente reordenadas.

Obtenido el sistema S' , <u>si r = n</u> , entonces la última ecuación de S' es de la forma $ax_n = b$ y de aquí resulta un único valor para x_n . Reemplazando este valor en la

penúltima ecuación, se obtiene un valor bien determinado para x_{n-1} . Siguiendo así, se ve que se obtiene un único sistema de valores numéricos para las incógnitas x_1 , x_2 , ... , x_n y en este caso el sistema es compatible determinado.

Si $r < n$, atribuyendo valores numéricos a las incógnitas x_{r+1} , x_{r+2} , . . . , x_n de la última ecuación se obtiene un valor unívocamente determinado para la incógnita x_r . Sustituyendo estos valores en la penúltima ecuación, se obtiene un valor bien determinado para la incógnita x_{r-1} . Así siguiendo, para cada sistema de valores numéricos atribuídos a las incógnitas x_{r+1} , x_{r+2} , ... , x_n , se obtienen valores unívocamente determinados para x_1 , x_2 , ... , x_r . El sistema S' tiene así infinitas soluciones, es decir, es compatible indeterminado.

Finalmente, observemos que si en el curso de las transformaciones que indicamos se obtiene alguna ecuación de la forma:

$$0x_k + 0x_{k+1} + ... + 0x_n = b \qquad \text{con} \quad b \neq 0$$

entonces el sistema correspondiente es claramente incompatible pues ningún conjunto de valores numéricos para las incógnitas satisfará esa ecuación.

Digamos además que si alguna ecuación se transforma en una del tipo:

$$0x_k + 0x_{k+1} + ... + 0x_n = 0$$

entonces puede eliminarse del sistema resultante puesto que es idénticamente verificada por cualquier sistema de valores de las incógnitas. (Observemos que este caso se presenta seguramente cuando el sistema de partida tiene más ecuaciones que incógnitas).

Pero cabe ahora plantear una pregunta. ¿Las soluciones así obtenidas son realmente todas las soluciones del sistema dado S?. Porque lo que se calcula en realidad son las soluciones de un sistema diferente S' obtenido del dado mediante la aplicación reiterada de un número finito de operaciones del tipo I , II y III.

Definición. Dos sistemas S_1 y S_2 de ecuaciones lineales se dicen equivalentes si tienen exactamente las mismas soluciones. (Eventualmente ninguna).

TEOREMA 6.1. Aplicando a un sistema S de ecuaciones lineales operaciones del tipo I , II ó III , se obtienen sistemas equivalentes al dado, y estas operaciones son inversibles en el sentido que si se pasa de un sistema S a uno S' aplicando una de ellas, se puede pasar de S' a S mediante una operación del mismo tipo.

Demostración: Si de un sistema S se obtiene uno S' mediante operaciones del tipo I, es decir, reordenando las ecuaciones, es claro que ambos sistemas tienen las mismas soluciones y que se puede obtener S de S' invirtiendo el reordenamiento. Si se aplica a un sistema S una operación del tipo II, es decir, si se reemplaza una ecuación

por la que se obtiene multiplicándola por un escalar no nulo k, es evidente que se puede pasar del sistema S' así obtenido al S multiplicando la ecuación correspondiente por k^{-1} , y se ve que toda solución de S es solución de S' y recíprocamente. Finalmente, supongamos que en un sistema S se reemplaza la h-ésima ecuación por la que se obtiene sumándole a la misma la j-ésima multiplicada por un escalar k, resultando así un nuevo sistema S' (operación III). Llamemos E_1 , E_2 , ... , E_m a los primeros miembros de las ecuaciones de S , b_1 , b_2 , ... , b_m a los segundos miembros. Los sistemas S y S' son:

$$S \begin{cases} E_1 = b_1 \\ E_2 = b_2 \\ \cdot \quad \cdot \\ \cdot \quad \cdot \\ \cdot \quad \cdot \\ E_h = b_h \\ \cdot \quad \cdot \\ \cdot \quad \cdot \\ \cdot \quad \cdot \\ E_m = b_m \end{cases} \qquad S' \begin{cases} E_1 = b_1 \\ E_2 = b_2 \\ \cdot \quad \cdot \\ \cdot \quad \cdot \\ \cdot \quad \cdot \\ E_h + kE_j = b_h + kb_j \\ \cdot \quad \cdot \\ \cdot \quad \cdot \\ \cdot \quad \cdot \\ E_m = b_m \end{cases}$$

Se ve claramente que toda solución de S es solución de S', pues si $(k_1, k_2, \ldots, k_n)$ es una solución del sistema S, es decir, si los k_i satisfacen todas las ecuaciones de S, entonces también satisfacen todas las ecuaciones de S'. Recíprocamente, toda solución del sistema S' es solución de S. Para verlo, llamemos E'_1 , E'_2 , ... , E'_m a los primeros miembros de las ecuaciones de S', b'_1 , b'_2 , ... , b'_m a los segundos miembros. Entonces S y S' se pueden escribir:

$$S \begin{cases} E'_1 = b'_1 \\ E'_2 = b'_2 \\ \cdot \quad \cdot \\ \cdot \quad \cdot \\ \cdot \quad \cdot \\ E'_h - kE'_j = b'_h - kb'_j \\ \cdot \quad \cdot \\ \cdot \quad \cdot \\ \cdot \quad \cdot \\ E'_m = b'_m \end{cases} \qquad S' \begin{cases} E'_1 = b'_1 \\ E'_2 = b'_2 \\ \cdot \quad \cdot \\ \cdot \quad \cdot \\ \cdot \quad \cdot \\ E'_h = b'_h \\ \cdot \quad \cdot \\ \cdot \quad \cdot \\ \cdot \quad \cdot \\ E'_m = b'_m \end{cases}$$

y es claro que si una n-upla $(k_1, k_2, \ldots, k_n)$ satisface las ecuaciones de S' también satisface las de S.

El teorema queda así demostrado.

Este teorema asegura la legitimidad del método de eliminación de Gauss para resolver un sistema de ecuaciones lineales. Como el método consiste en pasar de un sistema de ecuaciones lineales dado S a otro S' , de más fácil resolución, mediante la aplicación reiterada de operaciones del tipo I , II y III en número finito, de lo recién demostrado resulta que todos los sistemas que se obtienen en el curso del proceso son equivalentes entre sí, en particular, lo son S y S'.

En lo que sigue formalizaremos el procedimiento que antes indicamos a grandes rasgos, para demostrar que puede aplicarse a cualquier sistema de ecuaciones lineales. Pero antes de hacerlo, introduciremos la noción de combinación lineal porque es útil y práctico manejar este lenguaje.

COMBINACIONES LINEALES.

Resulta cómodo dar un nombre a las ecuaciones lineales que se pueden obtener de otras m ecuaciones dadas con n incógnitas multiplicando a la primera ecuación por un escalar k_1 , a la segunda por otro k_2 ,....., a la m-ésima por k_m y sumándolas todas. Se obtiene así una nueva ecuación lineal de la forma:

$$(k_1a_{11} + k_2a_{21} + \ldots\ldots + k_ma_{m1})x_1 + (k_1a_{12} + k_2a_{22} + \ldots\ldots + k_ma_{m2})x_2 + \ldots\ldots$$

$$\ldots\ldots + (k_1a_{1n} + k_2a_{2n} + \ldots\ldots + k_ma_{mn})x_n = k_1b_1 + k_2b_2 + \ldots\ldots + k_mb_m$$

y se dice que esta ecuación es una combinación lineal de las m dadas, con coeficientes k_1 , k_2 , ... , k_m .

Si llamamos E_1 , E_2 , ... , E_m a las ecuaciones dadas y E a la nueva ecuación, escribiremos:

$$E = k_1E_1 + k_2E_2 + \ldots\ldots + k_mE_m$$

Diremos que la combinación lineal es trivial si $k_1 = k_2 = \ldots\ldots = k_m = 0$, y que es no trivial en cualquier otro caso.

Observemos que una ecuación E es combinación lineal no trivial de otras E_1 ,...., E_m si y sólo si E puede obtenerse del sistema formado por E_1 ,...., E_m mediante una sucesión finita de operaciones del tipo II y III.

Queda a cargo del lector verificar que valen las siguientes propiedades:

1. Si una ecuación E es combinación lineal de otras m, entonces toda solución del sistema formado por esas m ecuaciones es solución de E.

2. Si una ecuación E es una combinación lineal de m ecuaciones, es también combinación lineal de esas m ecuaciones y otras q cualesquiera. En particular, dadas m ecuaciones, cualquiera de ellas puede escribirse como combinación lineal de las

m ecuaciones dadas.

3. Si una ecuación E es combinación lineal de las ecuaciones E_1 , E_2 ,....., E_m y cada una de éstas a su vez es una combinación lineal de las ecuaciones E'_1 , E'_2 ,....., E'_k , entonces E es una combinación lineal de E'_1 , E'_2 ,....., E'_k .

Observemos que de lo dicho resulta:

1) Al aplicar a un sistema de ecuaciones lineales dado S cualquier operación del tipo I , II ó III , ó una sucesión finita de las mismas, las ecuaciones del sistema S' resultante son todas combinaciones lineales de las del sistema dado , y recíprocamente (puesto que las operaciones de ese tipo son inversibles).

2) Si S y S' son dos sistemas de ecuaciones lineales tales que toda ecuación de S es una combinación lineal de las ecuaciones de S', y recíprocamente, toda ecuación de S' es una combinación lineal de las de S, entonces S y S' son sistemas equivalentes, como sigue claramente de la propiedad 1. (En particular, si en un sistema de ecuaciones S una de ellas es combinación lineal de otras ecuaciones del sistema, entonces el sistema S' que se obtiene eliminando dicha ecuación es equivalente al dado. Por lo tanto, cuando se va a resolver un sistema se pueden descartar las ecuaciones que son combinaciones lineales de otras).

MATRIZ DE UN SISTEMA.

Como para resolver un sistema por el método de Gauss, se trata de reordenar las ecuaciones, multiplicarlas por escalares y sumarlas, no hay necesidad de escribir el sistema completo y basta considerar el cuadro de los coeficientes y términos independientes:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix}$$

que se llama la matriz del sistema.

Por ejemplo, dado el sistema

$$\begin{cases} 2x_1 - x_2 + 3x_3 - 2x_4 = 1 \\ x_1 + \frac{1}{2}x_2 + \quad + x_4 = 0 \\ \quad x_2 + 5x_3 \quad\quad = -3 \end{cases}$$

la matriz correspondiente es:

$$\begin{pmatrix} 2 & -1 & 3 & -2 & 1 \\ 1 & \frac{1}{2} & 0 & 1 & 0 \\ 0 & 1 & 5 & 0 & -3 \end{pmatrix}$$

En general, dado un cuerpo K, una matriz m×n sobre K es un cuadro de m×n elementos de de K dispuestos en m filas y n columnas.

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix}$$

Los elementos a_{ij} , para $1 \leq i \leq m$, $1 \leq j \leq n$, se llaman los elementos de la matriz. Abreviadamente se escribe $A = (a_{ij})$.

Una matriz se dice cuadrada si tiene igual número de filas que de columnas, es decir si m = n. En caso contrario se dice rectangular.

EJEMPLOS.

$$\begin{pmatrix} 2 & -1 \\ 0 & 3 \\ \frac{1}{2} & 4 \end{pmatrix} \quad ; \quad \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad ; \quad \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \quad ; \quad \begin{pmatrix} 2 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 8 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad ; \quad \begin{pmatrix} 1 & 0 & 0 & \frac{1}{2} & 6 \\ 0 & 1 & 0 & -3 & 17 \\ 0 & 0 & 1 & 1 & -1 \end{pmatrix} \quad ; \quad \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}$$

A cada sistema de m ecuaciones lineales con n incógnitas se le puede asociar una matriz m×(n+1) formada por los coeficientes y los términos independientes. Recíprocamente, dada una matriz m×n , $A = (a_{ij})$, al sistema de m ecuaciones lineales con n-1 incógnitas que tiene a A por matriz le diremos el sistema asociado a la matriz A.

Por ejemplo , si $A = \begin{pmatrix} 1 & -2 & 3 & 4 & -3 \\ 2 & 0 & 7 & -5 & \frac{1}{2} \end{pmatrix}$, el sistema asociado a A es:

$$\begin{cases} x_1 - 2x_2 + 3x_3 + 4x_4 = -3 \\ 2x_1 + 7x_3 - 5x_4 = \frac{1}{2} \end{cases}$$

Las operaciones I , II y III se traducen en términos de la matriz del sistema como sigue:

I. Intercambiar dos filas de la matriz.

II. Reemplazar una fila por la que se obtiene multiplicándola por un escalar no nulo.

III. Reemplazar una fila por la que se obtiene sumando a sus elementos los de otra multiplicados por un escalar.

Llamaremos "elementales" a estas operaciones.

En lo que sigue, en lugar de trabajar con los sistemas de ecuaciones completos, lo haremos directamente con las matrices asociadas.

Sea $M_{m\times n}(K)$ el conjunto de todas las matrices m×n sobre K. Cada operación elemental puede interpretarse como una función e del conjunto $M_{m\times n}(K)$ en sí mismo, que a cada matriz $A \in M_{m\times n}(K)$ hace corresponder la matriz e(A) obtenida de A aplicándole la operación elemental e.

Expresando los resultados del teorema 6.1 en términos de matrices se tiene:

TEOREMA 6.2. Toda función (operación) elemental es inversible y su inversa es una operación elemental del mismo tipo que la dada.

Demostración: Si e es una operación elemental, razonando como en el teorema 6.1 , se ve que existe una operación elemental e' del mismo tipo que e tal que $e'(e(A)) = e(e'(A)) = A$, $\forall\ A \in M_{m\times n}(K)$. (1)

En efecto, si e es del tipo I , que intercambia, por ejemplo, la fila i-ésima con la j-ésima, sea e' = e.

Si e es del tipo II, que multiplica la fila i-ésima por el escalar $k \neq 0$, sea e' la operación que multiplica la fila i-ésima por k^{-1}

Finalmente, si e es del tipo III, que consiste en sumar a los elementos de la fila i-ésima los de la j-ésima previamente multiplicados por el escalar k, sea e' la operación tal que suma a los elementos de la i-ésima fila los de la j-ésima multiplicados por -k.

Se ve sin dificultad que en cada uno de los tres casos se verifica (1).

Definición. Si A y B son dos matrices m×n de elementos de K, se dice que A es equivalente por filas a B si B puede obtenerse de A mediante la aplicación de un número finito de operaciones elementales.

Esta es una relación de equivalencia en el conjunto $M_{m\times n}(K)$, como se verifica sin dificultad aplicando el teorema anterior.

Hablando en términos de sistemas de ecuaciones, si A y B son dos matrices equivalentes por filas, los sistemas de ecuaciones S y S' asociados a A y a B son sistemas equivalentes, como resulta del teorema 6.1.

MATRICES REDUCIDAS CANONICAS.

Para resolver un sistema de ecuaciones lineales S lo que debe hacerse entonces es considerar la matriz A del sistema dado, y mediante operaciones elementales convenientemente elegidas, llegar a una B cuya forma permita calcular rápidamente las soluciones del sistema asociado S' , que será equivalente al dado.

Veamos un ejemplo. Sea el sistema

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

$$A = \begin{pmatrix} 1 & 2 & 1 & 2 & 5 & 2 \\ 2 & 4 & 3 & 3 & -2 & 4 \\ -1 & -2 & 2 & 3 & 1 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 1 & 2 & 5 & 2 \\ 0 & 0 & 1 & -1 & -12 & 0 \\ 0 & 0 & 3 & 5 & 6 & 2 \end{pmatrix} \longrightarrow$$

$$\longrightarrow \begin{pmatrix} 1 & 2 & 0 & 3 & 17 & 2 \\ 0 & 0 & 1 & -1 & -12 & 0 \\ 0 & 0 & 0 & 8 & 42 & 2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 0 & 3 & 17 & 2 \\ 0 & 0 & 1 & -1 & -12 & 0 \\ 0 & 0 & 0 & 1 & \frac{21}{4} & \frac{1}{4} \end{pmatrix} \longrightarrow$$

$$\longrightarrow \begin{pmatrix} 1 & 2 & 0 & 0 & \frac{5}{4} & \frac{5}{4} \\ 0 & 0 & 1 & 0 & -\frac{27}{4} & \frac{1}{4} \\ 0 & 0 & 0 & 1 & \frac{21}{4} & \frac{1}{4} \end{pmatrix} = B$$

Como la 1^{a} , 3^{a} y 4^{a} columnas de B tienen un 1 y todos los demás elementos nulos, las incógnitas x_1 , x_3 y x_4 correspondientes a estas columnas aparecen en una sola ecuación cada una, y con coeficiente 1. Ordenando las incógnitas de modo que x_1 , x_3 y x_4 queden en primer término el sistema asociado a B es el siguiente:

$$\begin{cases} x_1 + 2x_2 + \frac{5}{4}x_5 = \frac{5}{4} & \text{ó} \quad x_1 = \frac{5}{4} - 2x_2 - \frac{5}{4}x_5 \\ x_3 - \frac{27}{4}x_5 = \frac{1}{4} & \text{ó} \quad x_3 = \frac{1}{4} + \frac{27}{4}x_5 \\ x_4 + \frac{21}{4}x_5 = \frac{1}{4} & \text{ó} \quad x_4 = \frac{1}{4} - \frac{21}{4}x_5 \end{cases}$$

Todas las soluciones del sistema se obtienen dándole valores arbitrarios a x_2 y x_5 y calculando los correspondientes valores de x_1 , x_3 y x_4 .

Luego, las soluciones del sistema son todas las quíntuplas

$$\left(\frac{5}{4} - 2a - \frac{5}{4} b \ , \ a \ , \ \frac{1}{4} + \frac{27}{4} b \ , \ \frac{1}{4} - \frac{21}{4} b \ , \ b \right) \quad \text{con} \quad a,b \in K.$$

Se ve que la matriz B tiene una forma muy conveniente para resolver el sistema. Este es el tipo de matriz al que se trata de llegar mediante operaciones elementales y que se llama matriz reducida canónica.

Formalizamos este concepto en la siguiente

Definición. Una matriz $m \times n$ C se dice reducida canónica si verifica las siguientes condiciones:

1) Toda fila nula de C aparece debajo de todas las filas no nulas.

2) El primer elemento no nulo de cada fila no nula es 1 y las columnas de C que contienen a esos 1 tienen todos los restantes elementos nulos.

3) Si las filas no nulas de C son la 1 , 2 , 3 , . . . , r y si el primer elemento no nulo de la fila i-ésima aparece en la columna j_i , i = 1,2,....,r , entonces $j_1 < j_2 < \ldots < j_r$.

EJEMPLOS.

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

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

son todas reducidas canónicas, excepto la tercera.

En una matriz reducida canónica llamaremos columnas independientes o principales a las columnas en que figuran los primeros elementos no nulos de las filas no nulas. En los ejemplos aparecen señaladas con flechas.

TEOREMA 6.3. Toda matriz $m \times n$ es equivalente a una reducida canónica.

Demostración: Sea A una matriz $m \times n$ no nula, pues en caso contrario el teorema se verifica. Podemos suponer que en A las filas no nulas son las r primeras, porque si hay filas nulas, intercambiando las filas entre sí (operación I) se puede obtener una matriz que verifique la condición 1) de la definición anterior. Si a_{1j_1} es el primer elemento no nulo de la $1^{\underline{a}}$ fila, multiplicando dicha fila por $a_{1j_1}^{-1}$ obtenemos una matriz $A_1 = (b_{ij})$ en la que el primer elemento no nulo de la $1^{\underline{a}}$ fila es 1 y está en la columna j_1 (operación II). Sumando a la fila i-ésima de A_1 la primera multiplicada por $-b_{ij_1}$, para i=2,3,...,r (operación III), se obtiene una matriz $A_2 = (c_{ij})$ donde la columna j_1 es $\begin{matrix}1\\0\\\vdots\\0\end{matrix}$. Si en A_2 todos los elementos de la $2^{\underline{a}}$ fila son nulos, no se hace nada. Si algún elemento no es nulo, sea c_{2j_2} el primer elemento no nulo de la $2^{\underline{a}}$ fila. Es claro que $j_2 \neq j_1$. Procediendo como antes se puede obtener, mediante operaciones elementales, una matriz A_3 que tiene la columna j_1-ésima igual a $\begin{matrix}1\\0\\\vdots\\0\end{matrix}$; la columna j_2-ésima igual a $\begin{matrix}0\\1\\\vdots\\0\end{matrix}$; y las columnas anteriores a la j_2-ésima iguales a las de la matriz A_2, puesto que como los elementos de la $2^{\underline{a}}$ fila que preceden al c_{2j_2} son todos nulos, las operaciones efectuadas no las afecta.

Aplicando este procedimiento fila por fila, al cabo de un número finito de pasos se llega a una matriz B que verifica la condición 2) de la definición de matriz reducida canónica. Y es claro que mediante adecuados intercambios de filas de B, se puede obtener una matriz reducida canónica C, que resulta ser equivalente a la dada A.

Observemos que la demostración anterior no sólo prueba la existencia de una matriz reducida canónica equivalente a una dada, sino que proporciona un método efectivo para calcularla.

De todo lo dicho resulta:

Para resolver un sistema de m ecuaciones lineales con n incógnitas se escribe la matriz A del sistema; se busca una matriz reducida canónica C equivalente a A. Los términos independientes del sistema asociado a C, o a cualquiera de las matrices intermedias, son los elementos de la última columna. Ahora bien, si al hacer el cálculo, alguna de las matrices intermedias o C tiene una fila con todos los elementos nulos excepto el último, es decir, una fila del tipo:

$$0 \quad 0 \quad . \quad . \quad . \quad . \quad 0 \quad b$$

con $b \neq 0$, entonces la ecuación correspondiente en el sistema asociado a esa matriz será

$$0x_1 + 0x_2 + \ldots\ldots + 0x_n = b$$

que obviamente no tiene ninguna solución. Luego dicho sistema no tiene solución y por lo tanto tampoco la tiene el dado, es decir se trata de un sistema incompatible.

Si en C no figura ninguna fila de ese tipo, el sistema es compatible. Se escribe el sistema asociado a C, y las incógnitas correspondientes a las columnas principales de C, digamos $x_{i_1}, x_{i_2}, \ldots, x_{i_r}$, figuran cada una en una sola ecuación y con coeficiente 1. Si r = n, el sistema tiene una única solución o sea es compatible determinado. Si r < n, se dan valores arbitrarios a las n-r incógnitas restantes y se calculan en función de ellos los valores de las incógnitas $x_{i_1}, x_{i_2}, \ldots, x_{i_r}$, que se llaman incógnitas principales. Se obtienen así todas las soluciones del sistema, que en este caso son infinitas. El sistema es compatible indeterminado.

EJEMPLOS.

1. $$\begin{cases} x_1 + 2x_2 - 3x_3 + x_4 = 1 \\ 2x_1 - x_2 + 2x_3 - x_4 = 1 \\ 4x_1 + 3x_2 - 4x_3 + x_4 = 2 \end{cases}$$

$$\begin{pmatrix} 1 & 2 & -3 & 1 & 1 \\ 2 & -1 & 2 & -1 & 1 \\ 4 & 3 & -4 & 1 & 2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & -3 & 1 & 1 \\ 0 & -5 & 8 & -3 & -1 \\ 0 & -5 & 8 & -3 & -2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & -3 & 1 & 1 \\ 0 & -5 & 8 & -3 & -1 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix}$$

Aunque todavía no se tiene una matriz reducida canónica no hace falta seguir los cálculos puesto que la última fila nos está diciendo que el sistema es incompatible.

2. $$\begin{cases} 3x_1 + 8x_2 + 2x_3 = 28 \\ x_1 + 2x_2 - x_3 = -1 \\ 4x_1 + 9x_2 - x_3 = 14 \end{cases}$$

$$\begin{pmatrix} 3 & 8 & 2 & 28 \\ 1 & 2 & -1 & -1 \\ 4 & 9 & -1 & 14 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & -1 & -1 \\ 3 & 8 & 2 & 28 \\ 4 & 9 & -1 & 14 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & -1 & -1 \\ 0 & 2 & 5 & 31 \\ 0 & 1 & 3 & 18 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & -1 & -1 \\ 0 & 1 & 3 & 18 \\ 0 & 2 & 5 & 31 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 0 & -7 & -37 \\ 0 & 1 & 3 & 18 \\ 0 & 0 & -1 & -5 \end{pmatrix} \longrightarrow \begin{pmatrix} \downarrow & \downarrow & \downarrow & \\ 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 5 \end{pmatrix}$$

El sistema asociado a la última matriz es:

$$\begin{cases} x_1 = -2 \\ x_2 = 3 \\ x_3 = 5 \end{cases}$$

Luego el sistema dado es compatible determinado y su única solución es (-2,3,5).

3. $$\begin{cases} 2x_1 + 5x_2 + 3x_3 + x_4 = 9 \\ 3x_1 + 8x_2 + 4x_3 + 2x_4 = 14 \\ x_1 + 2x_2 + 2x_3 = 4 \\ 2x_1 + 7x_2 + x_3 + 3x_4 = 11 \end{cases}$$

$$\begin{pmatrix} 2 & 5 & 3 & 1 & 9 \\ 3 & 8 & 4 & 2 & 14 \\ 1 & 2 & 2 & 0 & 4 \\ 2 & 7 & 1 & 3 & 11 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 2 & 0 & 4 \\ 2 & 5 & 3 & 1 & 9 \\ 3 & 8 & 4 & 2 & 14 \\ 2 & 7 & 1 & 3 & 11 \end{pmatrix} \longrightarrow$$

$$\begin{pmatrix} 1 & 2 & 2 & 0 & 4 \\ 0 & 1 & -1 & 1 & 1 \\ 0 & 2 & -2 & 2 & 2 \\ 0 & 3 & -3 & 3 & 3 \end{pmatrix} \longrightarrow \begin{pmatrix} \downarrow & \downarrow & & & \\ 1 & 0 & 4 & -2 & 2 \\ 0 & 1 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

La última matriz es reducida canónica y el sistema asociado es

$$\begin{cases} x_1 + 4x_3 - 2x_4 = 2 \\ x_2 - x_3 + x_4 = 1 \end{cases}$$

donde x_1 y x_2 son las incógnitas principales (pues corresponden a las columnas principales). Entonces el sistema se escribe:

$$\begin{cases} x_1 = 2 - 4x_3 + 2x_4 \\ x_2 = 1 + x_3 - x_4 \end{cases}$$

Luego es compatible indeterminado y las soluciones son todas las cuaternas

$$(2-4a+2b \ , \ 1+a-b \ , \ a \ , \ b) \quad \text{con} \quad a \ , \ b \in K \ .$$

4. Hallar las soluciones del siguiente sistema homogéneo:

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

$$\begin{pmatrix} 2 & 1 & 5 & 0 \\ 3 & 5 & 4 & 0 \\ 1 & -3 & 6 & 0 \\ 0 & 7 & -7 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -3 & 6 & 0 \\ 3 & 5 & 4 & 0 \\ 2 & 1 & 5 & 0 \\ 0 & 7 & -7 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -3 & 6 & 0 \\ 0 & 14 & -14 & 0 \\ 0 & 7 & -7 & 0 \\ 0 & 7 & -7 & 0 \end{pmatrix} \longrightarrow$$

$$\begin{pmatrix} 1 & -3 & 6 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} \overset{\downarrow}{1} & \overset{\downarrow}{0} & 3 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

El sistema asociado es

$$\begin{cases} x_1 + 3x_3 = 0 \\ x_2 - x_3 = 0 \end{cases}$$

y despejando las incógnitas principales se tiene:

$$\begin{cases} x_1 = -3x_3 \\ x_2 = x_3 \end{cases}$$

Luego las soluciones son todas las ternas $(-3a, a, a)$ con $a \in K$ y el sistema es indeterminado.

OBSERVACION.

Como los sistemas asociados a las matrices intermedias son todos equivalentes al dado, no es necesario aplicar el método hasta obtener una matriz reducida canónica, y se puede dar por terminado el proceso cuando se obtiene una matriz cuya forma posibilita el cálculo rápido de los valores de las incógnitas.

Por ejemplo, en el ejercicio anterior podríamos habernos detenido en la penúltima matriz y considerar el sistema asociado a la misma

$$\begin{cases} x_1 - 3x_2 + 6x_3 = 0 \\ x_2 - x_3 = 0 \end{cases}$$

del que se deduce rápidamente $x_2 = x_3$, $x_1 = -3x_3$.

Tampoco es necesario aplicar las operaciones elementales en el orden indicado en el teorema 6.3, pues aplicándolas de otra manera pueden resultar cálculos más sencillos. Por ejemplo, a veces conviene no dividir por el primer elemento no nulo de una fila para reducirlo a 1, evitándose así trabajar con fracciones. El procedimiento a se

guir para realizar cálculos lo más sencillos posibles depende de la matriz que se considere en cada caso y de la habilidad del calculista.

También se puede usar, para simplificar los cálculos, una observación ya hecha: Si en un sistema de ecuaciones lineales S una de ellas es combinación lineal de otras ecuaciones del sistema, entonces el sistema S' que se obtiene eliminando dicha ecuación es equivalente al dado.

5. Resolver el siguiente sistema:

$$\begin{cases} 2x_1 - x_2 + 3x_3 - x_4 + x_5 = 2 \\ 4x_1 - 2x_2 + 6x_3 - 2x_4 + 2x_5 = 4 \\ 2x_1 - x_2 - x_3 + 2x_4 = 0 \\ 2x_1 - x_2 + 7x_3 - 4x_4 + 2x_5 = 4 \\ 4x_3 - 3x_4 + x_5 = 2 \end{cases}$$

La matriz del sistema es:

$$\begin{pmatrix} 2 & -1 & 3 & -1 & 1 & 2 \\ 4 & -2 & 6 & -2 & 2 & 4 \\ 2 & -1 & -1 & 2 & 0 & 0 \\ 2 & -1 & 7 & -4 & 2 & 4 \\ 0 & 0 & 4 & -3 & 1 & 2 \end{pmatrix}$$

Se ve a simple vista que la segunda fila es igual a la primera multiplicada por 2, y que la quinta fila es igual a la primera menos la tercera. Entonces se pueden descartar del sistema dado las ecuaciones segunda y quinta, y considerar más simplemente la matriz:

$$\begin{pmatrix} 2 & -1 & 3 & -1 & 1 & 2 \\ 2 & -1 & -1 & 2 & 0 & 0 \\ 2 & -1 & 7 & -4 & 2 & 4 \end{pmatrix}$$

$$\begin{pmatrix} 2 & -1 & 3 & -1 & 1 & 2 \\ 2 & -1 & -1 & 2 & 0 & 0 \\ 2 & -1 & 7 & -4 & 2 & 4 \end{pmatrix} \longrightarrow \begin{pmatrix} 2 & -1 & 3 & -1 & 1 & 2 \\ 0 & 0 & -4 & 3 & -1 & -2 \\ 0 & 0 & 4 & -3 & 1 & 2 \end{pmatrix}$$

$$\begin{pmatrix} 2 & -1 & 3 & -1 & 1 & 2 \\ 0 & 0 & 4 & -3 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} \downarrow & & \downarrow & & & \\ 2 & -1 & 3 & -1 & 1 & 2 \\ 0 & 0 & 1 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Podemos detenernos en esta matriz y considerar el sistema asociado, en el que las incógnitas principales son x_1 y x_3 :

$$\begin{cases} 2x_1 - x_2 + 3x_3 - x_4 + x_5 = 2 \\ \qquad\quad x_3 - 2x_4 \qquad\quad = 0 \end{cases} \quad \text{o sea} \quad \begin{cases} 2x_1 + 3x_3 = 2 + x_2 + x_4 - x_5 \\ \qquad x_3 = 2x_4 \end{cases}$$

De aquí resulta $x_3 = 2x_4$ y $x_1 = \frac{1}{2}(2 + x_2 - 5x_4 - x_5)$ y el sistema es indeterminado. Las soluciones son todas las quintuplas

$$(1 + \tfrac{1}{2}a - \tfrac{5}{2}b - \tfrac{1}{2}c \ , \ a \ , \ 2b \ , \ b \ , \ c) \quad \text{con } a,b,c \in K$$

Conteste ahora el lector a las siguientes preguntas:

Dado un sistema lineal homogéneo de m ecuaciones con n incógnitas, cómo es el sistema si $m < n$? ¿Qué relación deben guardar m y n y qué forma debe tener la matriz reducida canónica equivalente a la matriz del sistema para que éste sea compatible?.

En general, dado un sistema cualquiera de m ecuaciones lineales con n incógnitas, qué relación guardan m y n y de qué forma es la matriz reducida canónica equivalente a la matriz del sistema si éste es compatible determinado? ¿Y si es indeterminado?. Y si es incompatible?.

NOTA. Como las operaciones que se efectúan en el método de Gauss son simples y es posible seguir un esquema bien definido (el visto en el teorema 6.3), este método es el más indicado para resolver sistemas de ecuaciones lineales cuando se utilizan máquinas de calcular.

Pero, desde el punto de vista teórico, no es satisfactorio y para estudiar muchos problemas, por ejemplo de geometría, es conveniente poder expresar de otra manera la compatibilidad o incompatibilidad de un sistema en función de los coeficientes y términos independientes.

Más adelante, en el parágrafo 6.4 de este capítulo veremos otro método para decidir si un sistema de ecuaciones lineales es compatible o no, y otra forma de calcular las soluciones.

EJERCICIOS.

1. a) Encontrar una matriz reducida canónica equivalente a la matriz dada en cada uno de los dos casos siguientes:

$$\begin{pmatrix} 2 & 4 & 1 & -8 & 7 \\ 1 & 2 & -2 & -4 & 3 \\ 3 & 6 & 1 & -12 & -2 \end{pmatrix} \qquad \begin{pmatrix} 1 & -i & 0 & 2 \\ 3i & 2 & 1 & -1 \\ i & -2 & 3 & 0 \end{pmatrix}$$

b) Hallar todas las soluciones de los siguientes sistemas de ecuaciones:

$$\begin{cases} 2x_1 + 4x_2 + x_3 - 8x_4 = 7 \\ x_1 + 2x_2 - 2x_3 - 4x_4 = 3 \\ 3x_1 + 6x_2 - x_3 - 12x_4 = -2 \end{cases} \qquad \begin{cases} x_1 - ix_2 = 2 \\ 3ix_1 + 2x_2 + x_3 = -1 \\ ix_1 - 2x_2 + 3x_3 = 0 \end{cases}$$

2. Hallar todas las soluciones de los siguientes sistemas de ecuaciones:

a) $$\begin{cases} x_1 + 2x_2 - x_3 = -1 \\ 4x_1 + 9x_2 - x_3 = 14 \\ 3x_1 + 8x_2 + 2x_3 = 28 \end{cases}$$

b) $$\begin{cases} x_1 + 2x_2 - 3x_3 + x_4 = 1 \\ 2x_1 - x_2 + 2x_3 - x_4 = 1 \\ 4x_1 + 3x_2 - 4x_3 + x_4 = 2 \end{cases}$$

c) $$\begin{cases} 2x_1 - x_2 + x_3 + x_4 = 1 \\ 4x_1 + 3x_2 - x_4 = 0 \\ x_1 - x_2 + x_3 - x_4 = -2 \end{cases}$$

d) $$\begin{cases} x_1 + 2x_2 - 3x_4 = 0 \\ 3x_1 + 6x_2 + 5x_3 - 5x_4 = 0 \\ 2x_1 + 4x_2 - 6x_4 = 0 \\ x_1 + 2x_2 + 5x_3 + x_4 = 0 \end{cases}$$

e) $$\begin{cases} 2x_1 + 5x_2 + 3x_3 + 10x_4 = 7 \\ x_1 + 2x_2 + 2x_3 + 4x_4 = 2 \\ 3x_1 + 5x_2 + 7x_3 + 10x_4 = 4 \end{cases}$$

f) $$\begin{cases} 3x_1 + 2x_2 + x_3 = 0 \\ 2x_1 + x_2 + 3x_3 = 0 \\ x_1 + 2x_2 + 3x_3 = 0 \end{cases}$$

g) $\begin{cases} x_1 + (1+i)x_2 + 2x_3 = 0 \\ -2x_1 + i\,x_2 = 3 \end{cases}$

h) $\begin{cases} 3x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5 + x_6 = 0 \\ 2x_1 - x_2 + 4x_3 + 5x_4 + x_5 = 0 \\ 4x_1 + 5x_2 + x_3 + 2x_4 - 3x_5 + 2x_6 = 0 \end{cases}$

i) $\begin{cases} x_1 + x_2 + x_3 = 1 \\ 2x_1 - x_3 = 0 \\ 3x_1 - 2x_2 + x_3 = -1 \\ 4x_1 + 5x_3 = 2 \end{cases}$

3. Dar un ejemplo de:

a) Un sistema de dos ecuaciones con dos incógnitas indeterminado.

b) " " " dos " " tres " incompatible.

c) Un sistema homogéneo con tres incógnitas determinado.

6.2. ALGEBRA DE MATRICES.

Vamos a estudiar más detenidamente las matrices sobre un cuerpo K.

Este concepto, que según acabamos de ver, surge naturalmente para simplificar la escritura cuando se trabaja con sistemas de ecuaciones lineales, desempeña un papel muy importante en álgebra lineal, y es un instrumento muy útil. La importancia y variedad de sus aplicaciones ha hecho que se desarrolle una teoría especial de matrices, de la que sólo veremos los rudimentos.

Si I_k representa al conjunto de los k primeros números naturales, una matriz m×n sobre K es una función de $I_m \times I_n$ en K, es decir, una aplicación que a cada par ordenado de números naturales (i,j) hace corresponder un elemento a_{ij} de K, $1 \leq i \leq m$, $1 \leq j \leq n$.

Usualmente una matriz se representa, como ya vimos, escribiendo los m×n elementos a_{ij} en un cuadro con m filas y n columnas, de modo que a_{ij} ocupa la i-ésima fila y la j-ésima columna. Abreviadamente se escribe (a_{ij}), especificándose el dominio de los índices: $1 \leq i \leq m$, $1 \leq j \leq n$.

Si m = n la matriz se dice cuadrada de orden n. En caso contrario se dice rectangular.

Notaremos a las matrices con letras mayúsculas A, B, C,. . . y al conjunto de todas las matrices m×n sobre K lo representaremos $M_{m\times n}(K)$. Si m = n escribiremos simplemente $M_n(K)$. (Así $M_n(Q)$, $M_n(R)$, $M_n(C)$ son los conjuntos de las matrices n×n de números racionales, reales y complejos respectivamente).

SUMA DE MATRICES.

Definición. Dadas dos matrices m×n , (a_{ij}) , (b_{ij}) , se llama suma de (a_{ij}) más (b_{ij}) a la matriz cuyo (i,j)-elemento es $a_{ij} + b_{ij}$.

$$(a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij})$$

EJEMPLO.

$$\begin{pmatrix} -3 & 0 & 1 \\ \frac{1}{2} & -1 & 5 \end{pmatrix} + \begin{pmatrix} -2 & 7 & 3 \\ -1 & 4 & -4 \end{pmatrix} = \begin{pmatrix} -5 & 7 & 4 \\ -\frac{1}{2} & 3 & 1 \end{pmatrix}$$

Se tiene así una operación binaria definida en el conjunto $M_{m\times n}(K)$ de todas las matrices m×n sobre K que tiene las siguientes propiedades:

PROPIEDADES.

S1. Es asociativa: $(A + B) + C = A + (B + C)$, $\forall$ $A,B,C \in M_{m\times n}(K)$

S2. Es conmutativa: $A + B = B + A$, $\forall$ $A,B \in M_{m\times n}(K)$

S3. Existe neutro para la suma: Es la matriz nula

$$0 = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

pues $A + 0 = A$, $\forall$ $A \in M_{m\times n}(K)$

S4. Toda matriz tiene simétrico: Dada $A = (a_{ij})$, la matriz $B = (-a_{ij})$ es tal que que $A + B = 0$.

Estas propiedades se resumen diciendo que $M_{m\times n}(K)$ es un grupo abeliano con respecto a la suma.

PRODUCTO DE MATRICES.

El producto de dos matrices está definido sólo cuando el número de columnas de la primera es igual al número de filas de la segunda.

Definición. Dadas una matriz m×n (a_{ij}) y una n×p (b_{ij}) se llama producto de la primera por la segunda a la matriz m×p (c_{ij}) tal que $c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$

$$(a_{ij}).(b_{ij}) = (\sum_{k=1}^{n} a_{ik}b_{kj})$$

Es decir, en la matriz producto el elemento ij es la suma de los productos de los elementos de la i-ésima fila de (a_{ij}) por los elementos de la j-ésima columna de (b_{ij}).

EJEMPLOS.

1. $$\begin{pmatrix} 2 & -1 & 0 \\ 3 & 1 & 4 \end{pmatrix} . \begin{pmatrix} 0 & 2 \\ 3 & 4 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 2.0+(-1).3+0.(-1) & 2.2+(-1).4+0.0 \\ 3.0+1.3+4.(-1) & 3.2+1.4+4.0 \end{pmatrix} = \begin{pmatrix} -3 & 0 \\ -1 & 10 \end{pmatrix}$$

2. $$\begin{pmatrix} 1 & -2 & 3 \\ 0 & 4 & -1 \\ 2 & -1 & 1 \end{pmatrix} . \begin{pmatrix} 0 & 5 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & -1 \\ -2 & 1 & 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} -8 & 8 & 2 & 3 & 3 \\ 6 & -1 & 3 & -1 & -4 \\ -3 & 11 & 2 & 1 & 3 \end{pmatrix}$$

3. $\begin{pmatrix} -2 & 3 & 1 & -1 \\ 0 & 2 & -4 & 1 \end{pmatrix} \cdot \begin{pmatrix} 7 \\ 3 \\ 0 \\ -5 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

PROPIEDADES.

La multiplicación de matrices es asociativa. Si $A = (a_{ij})$, $B = (b_{ij})$, $C = (c_{ij})$ son tres matrices tales que los productos A.B y B.C están definidos, entonces también están definidos los productos (A.B).C y A.(B.C) y

$$(A.B).C = A.(B.C)$$

Para demostrarlo, supongamos que A es de tipo m×n, B de tipo n×p y C de tipo p×s. Entonces

$$(A.B).C = ((a_{ij}).(b_{ij})).(c_{ij}) = (\sum_{k=1}^{n} a_{ik}b_{kj}).(c_{ij}) = (\sum_{h=1}^{p} (\sum_{k=1}^{n} a_{ik}b_{kh})c_{hj}) =$$

$$= (\sum_{k=1}^{n} a_{ik}(\sum_{h=1}^{p} b_{kh}c_{hj})) = (a_{ij}).(\sum_{h=1}^{p} b_{ih}c_{hj}) = (a_{ij}).((b_{ij}).(c_{ij})) = A.(B.C)$$

La multiplicación de matrices es distributiva con respecto a la suma. Si $A = (a_{ij})$, $B = (b_{ij})$ y $C = (c_{ij})$ son tres matrices tales que las operaciones en cuestión están definidas, entonces

$$A.(B + C) = A.B + A.C$$

y bajo las mismas condiciones es

$$(A + B).C = A.C + B.C$$

Supongamos que A es de tipo m×n, B y C de tipo n×p. Entonces

$$A.(B + C) = (a_{ij}).((b_{ij}) + (c_{ij})) = (a_{ij}).(b_{ij} + c_{ij}) = (\sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj})) =$$

$$= (\sum_{k=1}^{n} a_{ik}b_{kj} + \sum_{k=1}^{n} a_{ik}c_{kj}) = (\sum_{k=1}^{n} a_{ik}b_{kj}) + (\sum_{k=1}^{n} a_{ik}c_{kj}) =$$

$$= (a_{ij}).(b_{ij}) + (a_{ij}).(c_{ij}) = A.B + A.C$$

La otra se demuestra de igual forma.

PRODUCTO DE UN ESCALAR POR UNA MATRIZ.

Una matriz cuadrada (a_{ij}) se dice escalar si $a_{ij} = 0$ para $i \neq j$ y $a_{11} = a_{22} = \ldots = a_{nn}$, es decir, si es de la forma

$$\begin{pmatrix} k & 0 & 0 & \cdots & 0 \\ 0 & k & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & k \end{pmatrix}$$

Multiplicando una matriz escalar de orden m por otra cualquiera m×n se tiene:

$$\begin{pmatrix} k & 0 & \cdots & 0 \\ 0 & k & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & k \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} k\,a_{11} & k\,a_{12} & \cdots & k\,a_{1n} \\ k\,a_{21} & k\,a_{22} & \cdots & k\,a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ k\,a_{m1} & k\,a_{m2} & \cdots & k\,a_{mn} \end{pmatrix}$$

Por analogía con este resultado, se define el producto de un escalar k por una matriz cualquiera $A = (a_{ij})$ como sigue:

$$k.(a_{ij}) = (ka_{ij})$$

El lector puede verificar sin dificultad que la multiplicación de escalares por matrices tiene las siguientes

PROPIEDADES.

1. $k(A + B) = kA + kB$
2. $(k+k')A = kA + k'A$
3. $k(k'A) = (kk')A$
4. $1A = A$
5. $(kA)B = A(kB) = k(AB)$

cualesquiera sean $k, k' \in K$, A y B matrices.

NOTACION. Antes de seguir adelante, digamos que hay una forma cómoda de notar un sistema de m ecuaciones lineales con n incógnitas, usando la multiplicación de matrices.

Sea A la matriz m×n formada por los coeficientes de las m ecuaciones e Y la matriz m×1 cuyos elementos son los términos independientes:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \qquad Y = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Llamando

$$X = \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix}$$

vemos que el sistema dado se puede representar por la ecuación matricial:

$$AX = Y \qquad (1)$$

En efecto, el sistema dado equivale a la igualdad de las dos siguientes matrices:

$$\begin{pmatrix} a_{11}x_1+a_{12}x_2+ \ . \ . \ . \ . \ +a_{1n}x_n \\ a_{21}x_1+a_{22}x_2+ \ . \ . \ . \ . \ +a_{2n}x_n \\ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \\ a_{m1}x_1+a_{m2}x_2+ \ . \ . \ . \ . \ +a_{mn}x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ b_m \end{pmatrix}$$

y la matriz del primer miembro es el producto AX. Luego se tiene (1) .

ANILLO DE MATRICES CUADRADAS DE ORDEN n.

Consideremos el conjunto $M_n(K)$ de todas las matrices cuadradas de orden n. Se ve que este conjunto es cerrado con respecto a la suma y a la multiplicación de matrices y se tienen así dos operaciones binarias definidas en $M_n(K)$.

De acuerdo con lo visto, estas operaciones tienen en $M_n(K)$ las siguientes propiedades: La suma es asociativa , conmutativa , tiene elemento neutro (la matriz nula) y toda matriz tiene simétrico. La multiplicación es asociativa, distributiva a derecha e izquierda con respecto a la suma y además tiene neutro pues la matriz

$$I = \begin{pmatrix} 1 & 0 & 0 \ . \ . \ . \ . & 0 \\ 0 & 1 & 0 \ . \ . \ . \ . & 0 \\ 0 & 0 & 1 \ . \ . \ . \ . & 0 \\ . & . & . \ . \ . \ . \ . \ . & . \\ 0 & 0 & 0 \ . \ . \ . \ . & 1 \end{pmatrix}$$

es tal que $A.I = I.A = A$, $\forall \ A \in M_n(K)$.

Estas propiedades se resumen diciendo que $M_n(K)$ es un anillo con unidad con respecto a la suma y a la multiplicación de matrices, no conmutativo, pues la multiplicación no es conmutativa si $n \geq 2$.

Para ver esto último, sean por ejemplo

$$A = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & 0 \end{pmatrix} \quad , \quad B = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

Entonces

$$A.B = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & 0 \end{pmatrix} = 0 \quad , \quad B.A = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & 0 \end{pmatrix} = B$$

$$y \quad A.B \neq B.A$$

Observemos que este ejemplo muestra además que el producto de dos matrices puede ser 0 sin que ninguno de los factores sea 0, lo que se expresa diciendo que en $M_n(K)$ hay divisores de cero.

En una matriz cuadrada se llama diagonal principal a la que va del extremo superior izquierdo al extremo inferior derecho, es decir, a la formada por los elementos a_{11} , a_{22} , ... , a_{nn} .

Matrices inversibles en $M_n(K)$.

Definición. Una matriz $A \in M_n(K)$ se dice inversible (también regular o no singular) si existe una matriz $B \in M_n(K)$ tal que $A.B = B.A = I$.

Sabemos que si existe una matriz B en esas condiciones, es única (ver pag. 36). B se dice la matriz inversa de A y se nota A^{-1} .

Es importante saber, para las aplicaciones, cuáles son las matrices inversibles y cómo calcular la matriz inversa. A continuación vamos a indicar un procedimiento que se basa en la aplicación de operaciones elementales y en la noción de matrices equivalentes por filas, y en el próximo parágrafo, una vez que hayamos definido determinante de una matriz, daremos otra caracterización de las matrices inversibles y otra manera de calcular la matriz inversa.

En general, vale la siguiente propiedad para las operaciones elementales:

TEOREMA 6.4. Si A es una matriz m×n, B una n×p y e una operación elemental que actúa sobre matrices con m filas entonces

$$e(A.B) = e(A).B$$

Demostración: Es suficiente demostrar que $e(A) = e(I)A$, cualquiera sea la matriz A

con m filas, siendo I la matriz unidad m×m. En efecto, demostrada esa propiedad se tiene:

$$e(AB) = e(I)AB = (e(I)A)B = e(A)B$$

Probemos pues que $e(A) = e(I)A$.

Sea $A = (a_{ij})$ una matriz m×n. Si e es una operación elemental que intercambia la fila h-ésima con la fila q-ésima entonces se tiene

$$e(I) = \begin{array}{c} \\ \\ \\ \\ \underline{h} \\ \\ \underline{q} \\ \\ \\ \end{array} \begin{pmatrix} 1 & & & & & & & & & & \\ & 1 & & & & & & & & & \\ & & \ddots & & & & & & & & \\ & & & \ddots & & & & & & & \\ 0 & 0 & \cdots & 0 & \cdots & \cdots & 1 & \cdots & 0 & & \\ & & & & \ddots & & & & & & \\ & & & & & \ddots & & & & & \\ 0 & 0 & \cdots & 1 & \cdots & 0 & \cdots & \cdots & 0 & & \\ & & & & & & & \ddots & & & \\ & & & & & & & & 1 & & \\ & & & & & & & & & 1 & \end{pmatrix} = (e_{ij})$$

Multiplicando

$$e(I)A = (e_{ij})(a_{ij}) = \left(\sum_{k=1}^{m} e_{ik}a_{kj}\right)$$

se ve que e(I)A tiene todas las filas distintas de la h-ésima y la q-ésima iguales a las de A y que en la fila h aparecen los elementos de la fila q de A y viceversa. Luego e(I)A es precisamente la matriz que se obtiene de A aplicándole la operación e: $e(A) = e(I)A$.

Supongamos ahora que e es una operación elemental que multiplica la fila h por un escalar $k \neq 0$. Entonces

$$e(I) = \begin{array}{c} \\ \\ \\ \\ \underline{h} \\ \\ \\ \\ \end{array} \begin{pmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & \ddots & & & & \\ & & & \ddots & & & \\ & & & & k & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

y se ve que e(I)A es una matriz que tiene todas las filas iguales a las de A, excepto la h-ésima que aparece multiplicada por k. Luego e(A) = e(I)A.

Finalmente, si e es una operación elemental que suma a la fila h la fila q multiplicada por un escalar k, entonces la matriz e(I) tiene todas las filas iguales a las de I, excepto la fila h que es:

$$\underset{}{0 \ldots \underset{h}{1} \ldots \underset{q}{k} \ldots 0}$$

(suponiendo h < q ; si q < h entonces será: $0 \ldots \underset{q}{k} \ldots \underset{h}{1} \ldots 0$ y el razonamiento es el mismo).

Multiplicando, la matriz e(I)A tiene todas las filas distintas de la h-ésima iguales a las de A y la fila h-ésima formada por los elementos c_{h1} , c_{h2} ,...., c_{hn} tales que

$$c_{hj} = a_{hj} + k\, a_{qj} \quad , \quad \text{para} \quad j=1,2,\ldots,n$$

Entonces, también en este caso, se verifica que e(A) = e(I)A.

El teorema queda así demostrado.

COROLARIO. Si e_1 , . . . , e_t son operaciones elementales entonces

$$(e_t \ldots e_1(A))B = e_t \ldots e_1(AB)$$

Se demuestra sin dificultad por inducción sobre t.

El siguiente teorema proporciona varias caracterizaciones de las matrices inversibles.

TEOREMA 6.5. Si A es una matriz cuadrada de orden n, las siguientes propiedades son equivalentes:

a) A es inversible.

b) A tiene inverso a izquierda.

c) A tiene inverso a derecha.

d) El sistema homogéneo AX = 0 tiene una única solución (la trivial).

e) Todo sistema AX = Y tiene una única solución, cualquiera que sea la matriz nx1 Y .

f) Si C es una matriz reducida canónica equivalente por filas a A, entonces C = I.

g) A es equivalente por filas a I.

Demostración: Seguiremos el siguiente esquema:

$$a \Longrightarrow b \Longrightarrow d \Longrightarrow f \Longrightarrow g \Longrightarrow a \quad y \quad c \Longrightarrow a \Longrightarrow e \Longrightarrow c$$

a $\Longrightarrow$ b . Trivial.

b $\Longrightarrow$ d . Sea A' una matriz que es inversa a izquierda de A y consideremos el sistema homogéneo $AX = 0$.

Sabemos que tiene por lo menos una solución: la trivial. Es la única pues

$$AX = 0 \Longrightarrow A'AX = A'0 \Longrightarrow X = 0$$

d $\Longrightarrow$ f . Supongamos que el sistema $AX = 0$ no tiene más soluciones que la trivial y sea C una matriz reducida canónica equivalente por filas a A. Entonces los sistemas $AX = 0$ y $CX = 0$ son equivalentes. Luego el sistema $CX = 0$ tiene una única solución, lo que implica que C debe ser la matriz unidad n×n , dada la forma especial de C.

f $\Longrightarrow$ g . Trivial, teniendo en cuenta que existe una matriz reducida canónica equivalente a A. (Teorema 6.3).

g $\Longrightarrow$ a . Supongamos que A es equivalente por filas a la matriz unidad I. Luego existen operaciones elementales e_1 , e_2 , ... , e_t tales que

$$e_t \ . \ . \ . \ e_2 e_1(A) = I \qquad (1)$$

Vamos a probar que la matriz $A' = e_t . \ . \ .e_2 e_1(I)$ es inversa de A, o sea, que

$$AA' = A'A = I$$

De (1) se tiene $A = e_1^{-1} e_2^{-1} . \ . \ . e_t^{-1} (I)$, representando e_i^{-1} a la operación inversa de e_i , para i = 1,2,...,t.

Aplicando el corolario del teorema 6.4 se tiene:

$$AA' = (e_1^{-1}e_2^{-1} . \ . \ .e_t^{-1}(I))(e_t . \ . \ .e_2e_1(I)) = e_1^{-1}e_2^{-1} . \ . \ .e_t^{-1}[I.(e_t . \ . \ .e_2e_1(I))] =$$

$$= e_1^{-1}e_2^{-1} . \ . \ .e_t^{-1}(e_t . \ . \ .e_2e_1(I)) = I$$

$$A'A = (e_t . \ . \ . \ e_2e_1(I))A = e_t . \ . \ .e_2e_1(IA) = e_t . \ . \ .e_2e_1(A) = I$$

c $\Longrightarrow$ a . Supongamos que A tiene inversa a derecha B. Entonces, de $AB = I$ resulta que B tiene inversa a izquierda, lo que implica por lo recién demostrado que B es inversible y $B^{-1} = A$. Luego $AB = BA = I$ y A es inversible.

a $\Longrightarrow$ e . Sea A una matriz inversible. Entonces, dado un sistema $AX = Y$, una solución del mismo es $A^{-1}Y$ pues

$$A(A^{-1}Y) = (AA^{-1})Y = Y$$

Además ésta es la única solución pues si X es una solución del sistema se tiene:

$$AX = Y \Longrightarrow A^{-1}AX = A^{-1}Y \Longrightarrow X = A^{-1}Y$$

e $\Longrightarrow$ c . Supongamos que la matriz A es tal que el sistema AX = Y tiene una única solución, cualquiera que sea la matriz $n\times 1$ Y. Sean Y_1 , $Y_2,\ldots,$ Y_n las matrices $n\times 1$ formadas respectivamente por la primera, la segunda, ... , la n-ésima columna de la matriz unidad, y sea X_i la solución del sistema $AX = Y_i$, para $i=1,2,\ldots,n$.

Considerando la matriz $n\times n$ A' cuyas columnas son X_1, $X_2,\ldots,$ X_n se ve que:

$$AA' = I$$

Luego A tiene inverso a derecha.

El teorema queda así demostrado.

COROLARIO. Si A es una matriz inversible y $e_1,e_2,\ldots,e_t$ es una sucesión de operaciones elementales que permite pasar de A a la matriz unidad I, entonces la matriz inversa A^{-1} se obtiene aplicando la misma sucesión de operaciones elementales a la matriz I.

Resulta de la demostración de g $\Longrightarrow$ a .

EJEMPLOS.

1. Decir si la matriz A es inversible y en tal caso hallar su inversa.

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

De acuerdo con el teorema 6.5. hay que buscar una matriz reducida canónica equivalente por filas a A y ver si coincide con la matriz unidad.

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

$$\rightarrow \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & -3 \\ 0 & 4 & 1 \end{pmatrix} \xrightarrow{4} \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & -3 \\ 0 & 0 & 13 \end{pmatrix} \xrightarrow{5} \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{6}$$

$$\rightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I$$

Luego la matriz dada es inversible. La matriz inversa A^{-1} se obtiene aplicando

a I las mismas operaciones elementales que nos permitieron pasar de A a I, en el mismo orden: 1) Intercambiar la 1^{a} y la 2^{a} filas; 2) Restar de la 2^{a} fila la 1^{a} multiplicada por 3, y de la 3^{a} la 1^{a} multiplicada por 4; 3) Sumar a la 1^{a} fila la 2^{a} ; 4) Restar a la 3^{a} fila la 2^{a} multiplicada por 4; 5) Multiplicar la 3^{a} fila por $\frac{1}{13}$; 6) Sumar a la 1^{a} fila la 3^{a} multiplicada por 2, y a la 2^{a} fila la 3^{a} multiplicada por 3.

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{1} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{2} \begin{pmatrix} 0 & 1 & 0 \\ 1 & -3 & 0 \\ 0 & -4 & 1 \end{pmatrix} \xrightarrow{3}$$

$$\rightarrow \begin{pmatrix} 1 & -2 & 0 \\ 1 & -3 & 0 \\ 0 & -4 & 1 \end{pmatrix} \xrightarrow{4} \begin{pmatrix} 1 & -2 & 0 \\ 1 & -3 & 0 \\ -4 & 8 & 1 \end{pmatrix} \xrightarrow{5} \begin{pmatrix} 1 & -2 & 0 \\ 1 & -3 & 0 \\ -\frac{4}{13} & \frac{8}{13} & \frac{1}{13} \end{pmatrix} \xrightarrow{6}$$

$$\rightarrow \begin{pmatrix} \frac{5}{13} & -\frac{10}{13} & \frac{2}{13} \\ \frac{1}{13} & -\frac{15}{13} & \frac{3}{13} \\ -\frac{4}{13} & \frac{8}{13} & \frac{1}{13} \end{pmatrix} = A^{-1}$$

El lector puede verificar este resultado multiplicando: $AA^{-1} = I$

2. Idem para la matriz

$$A = \begin{pmatrix} -2 & 1 \\ 1 & -\frac{1}{2} \end{pmatrix}$$

Procediendo como antes:

$$A = \begin{pmatrix} -2 & 1 \\ 1 & -\frac{1}{2} \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -\frac{1}{2} \\ -2 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & 0 \end{pmatrix}$$

La última matriz es reducida canónica y no coincide con la matriz unidad. Por lo tanto A no es inversible.

3. $$A = \begin{pmatrix} 7 & 2 \\ 3 & 5 \end{pmatrix}$$

$$A = \begin{pmatrix} 7 & 2 \\ 3 & 5 \end{pmatrix} \xrightarrow{1} \begin{pmatrix} 1 & -8 \\ 3 & 5 \end{pmatrix} \xrightarrow{2} \begin{pmatrix} 1 & -8 \\ 0 & 29 \end{pmatrix} \xrightarrow{3} \begin{pmatrix} 1 & -8 \\ 0 & 1 \end{pmatrix} \xrightarrow{4} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

Luego A es inversible.

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \xrightarrow{1} \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \xrightarrow{2} \begin{pmatrix} 1 & -2 \\ -3 & 7 \end{pmatrix} \xrightarrow{3} \begin{pmatrix} 1 & -2 \\ -\frac{3}{29} & \frac{7}{29} \end{pmatrix} \xrightarrow{4} \begin{pmatrix} \frac{5}{29} & -\frac{2}{29} \\ -\frac{3}{29} & \frac{7}{29} \end{pmatrix} = A^{-1}$$

OBSERVACION.

En lugar de trabajar con operaciones elementales se suele introducir la noción de matriz elemental para describir el efecto de aplicar una operación elemental por medio de un producto de matrices. Por el teorema 6.4 sabemos que

$$e(A) = e(I).A$$

Luego la matriz e(A) es el resultado de multiplicar la matriz e(I) por la matriz A. Se da un nombre especial a las matrices de la forma e(I):

Definición. Se llaman matrices elementales a las matrices de $M_n(K)$ que se pueden obtener de la matriz unidad I_n aplicando una operación elemental.

Por ejemplo, en $M_2(K)$ las matrices elementales son las de la forma:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}; \begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}; \begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix}, k\neq 0; \begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix}, k\neq 0$$

De esta manera todo lo visto para matrices inversibles se traduce en términos de matrices y producto de matrices. Por ejemplo, se tiene que: "Una matriz n×n A es inversible si y sólo si A es un producto de matrices elementales, es decir, si y sólo si existen matrices elementales $E_1, E_2, \ldots, E_t$ tales que $A = E_t \ldots E_2 E_1$. Y en tal caso la matriz inversa es $A^{-1} = E_1^{-1} . E_2^{-1} \ldots . E_t^{-1}$".

EJERCICIOS.

1. Dadas las matrices

$$A = \begin{pmatrix} -1 & 0 & 2 & 1 \\ 3 & 4 & \frac{1}{2} & -1 \\ -5 & 2 & 0 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 5 & \frac{1}{4} & 0 & -3 \\ -2 & -1 & 1 & 0 \\ 6 & 0 & -3 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} -1 & 0 & 7 & -5 \\ 2 & -3 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{pmatrix},$$

$$D = \begin{pmatrix} 2 & -1 \\ 3 & 0 \\ 1 & -5 \\ 0 & 2 \end{pmatrix}, \quad E = \begin{pmatrix} 3 & -2 \\ 5 & -1 \end{pmatrix}, \quad F = (7 \quad -1 \quad 0 \quad 4), \quad G = \begin{pmatrix} 2 \\ 3 \\ -1 \\ 1 \end{pmatrix}$$

$$H = \begin{pmatrix} 1 & 0 \\ -1 & 3 \end{pmatrix}$$

a) Hallar $A+B$, $4A-2B+C$, $2(A+C)-B$, CD , DE , FD , $D-DE$, E^3-3E , FG , EH .

b) Verificar que: $EH \neq HE$; $FG \neq GF$; $(BD)E = B(DE)$; $(A+B)G = AG+BG$. ¿Está definido $AD-DA$?.

c) ¿Para qué pares de matrices de las anteriores está definido el producto?. Indicar en cada caso el tipo de la matriz producto.

2. Dadas las matrices de $M_2(C)$:

$$A = \begin{pmatrix} 1+i & -2i \\ 0 & i \end{pmatrix} \quad , \quad B = \begin{pmatrix} 1 & -3i \\ 1-i & 0 \end{pmatrix}$$

Hallar AB , BA , A^2 , B^2 , $(A+B)^2$, $A^2 + B^2$.

¿Es $(A+B)^2 = A^2 + 2AB + B^2$? . ¿Y $A^2 - B^2 = (A+B)(A-B)$? . ¿Cómo explica este hecho?.

3. a) Hallar en cada caso matrices de $M_2(R)$ que verifiquen las propiedades indicadas:

i) $AB = 0$, $A \neq 0$ y $B \neq 0$

ii) $AB = 0$ y $BA \neq 0$

iii) $A^2 = 0$, $A \neq 0$

iv) $A^2 = A$, $A \neq 0$, $A \neq I$ $\quad \left(I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right)$

v) $AB = I$, $A \neq I$

vi) $AB = AC$, $B \neq C$

b) Decir si las siguientes proposiciones son verdaderas, cualquiera sean las matrices A , B , C:

$AB = 0 \Longrightarrow A = 0$ ó $B = 0$ falso pues no verifica (a, i).

$AB = AC \Longrightarrow B = C$

$BA = CA \Longrightarrow B = C$

4. Una matriz cuadrada $A = (a_{ij})$ se dice:

a) Triangular superior si $a_{ij} = 0$ para $i > j$

b) Diagonal si $a_{ij} = 0$ para $i \neq j$.

c) Escalar si es diagonal y $a_{11} = a_{22} = \ldots = a_{nn}$.

Demostrar que el conjunto de las matrices triangulares superiores de orden n es cerrado con respecto a la suma y a la multiplicación de matrices, y a la multiplicación de matrices por escalares. Idem para las matrices diagonales y para las escalares.

Demostrar que en $M_n(K)$ las matrices escalares conmutan con cualquier otra.

5. Dada una matriz m×n, $A = (a_{ij})$ se llama traspuesta de A y se nota A^T a la matriz n×m (b_{ij}) tal que $b_{ij} = a_{ji}$ para todo par de índices i,j , $1 \leq i \leq n$, $1 \leq j \leq m$. Es decir, la traspuesta de A es la matriz que se obtiene intercambiando filas con columnas. Por ejemplo,

si $A = \begin{pmatrix} 2 & -1 \\ 0 & 3 \\ 4 & -8 \end{pmatrix}$, es $A^T = \begin{pmatrix} 2 & 0 & 4 \\ -1 & 3 & -8 \end{pmatrix}$

a) Dadas $A = \begin{pmatrix} 2 & 0 & 3 & 5 \\ 1 & -1 & 2 & 0 \end{pmatrix}$, $B = \begin{pmatrix} -2 & 0 \\ 1 & 2 \\ 3 & 5 \\ 0 & -1 \end{pmatrix}$.

Hallar A^T , B^T , $(AB)^T$, $B^T.A^T$, $A^T.B^T$

b) Demostrar que las siguientes propiedades valen en general:

1) $(A^T)^T = A$

2) $(A+B)^T = A^T + B^T$

3) $(AB)^T = B^T . A^T$

4) $(kA)^T = k.A^T$, k escalar

6. Aplicando operaciones elementales, averiguar si cada una de las siguientes matrices es inversible, y en tal caso hallar la matriz inversa:

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

7. a) Probar que si A y B son dos matrices n×n inversibles entonces AB también es inversible y $(AB)^{-1} = B^{-1}A^{-1}$.

b) Generalizarel resultado anterior y demostrar que si A_1 , A_2 ,..., A_s son matrices inversibles entonces el producto $A_1 A_2 \ldots\ldots A_s$ es inversible y $(A_1A_2\ldots\ldots A_s)^{-1} = A_s^{-1} \ldots\ldots A_2^{-1} A_1^{-1}$.

6.3. DETERMINANTES

Vamos a ver ahora que a cada matriz cuadrada de orden n sobre K se le puede asociar un elemento de K, llamado el determinante de la matriz, o sea, que existe una aplicación $M_n(K) \longrightarrow K$, que verifica ciertas propiedades particulares.

El concepto de determinante es muy útil y de gran aplicación en álgebra lineal.

Sea A una matriz cuadrada de orden n.

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

Consideremos todos los productos posibles de n elementos de la matriz, de modo que los n factores que figuran en cada producto pertenezcan a filas y a columnas diferentes, es decir, que aparezca un elemento de cada fila y uno de cada columna. A cada producto de este tipo se le adjunta el signo + ó - según que las permutaciones formadas por los índices que indican las filas y los que indican las columnas a las que pertenecen los n factores, sean de la misma o de distinta clase. El signo de cada producto no depende del orden de los factores puesto que al intercambiar dos factores entre sí, se trasponen dos índices en la permutación que forman los primeros índices y otros dos en la que forman los segundos índices, de modo que las permutaciones siguen siendo de la misma o de distinta clase. Entonces, suponiendo a los n factores ordenados de modo que los primeros índices formen la permutación 1 2 ..... n un producto del tipo indicado es de la forma:

$$(-1)^{\nu} a_{1\alpha_1} a_{2\alpha_2} \cdots\cdots a_{n\alpha_n} \qquad (1)$$

donde ν es el número de inversiones que presenta la permutación $\alpha_1 \alpha_2 \ldots\ldots \alpha_n$.

Por ejemplo, en una matriz de 4° orden el producto

$$a_{24} a_{33} a_{41} a_{12} = a_{12} a_{24} a_{33} a_{41}$$
$$\quad 2 \quad 4 \quad 3 \quad 1$$

tiene signo + pues la permutación 2 4 3 1 que forman los segundos índices presenta cuatro inversiones.

$$(-1)^4 \, a_{12} a_{24} a_{33} a_{41}$$

En una matriz de orden n, se pueden formar n! productos del tipo indicado, la mitad de los cuales tienen signo + y la otra mitad signo - .

Definición. Dada una matriz cuadrada $A = (a_{ij})$ de orden n, se llama determinante de

A a la suma de los n! productos de la forma (1).

$$\det(A) = \sum (-1)^{\nu} a_{1\alpha_1} a_{2\alpha_2} \cdots a_{n\alpha_n}$$

Los términos de esta suma se llaman los términos del determinante, los elementos a_{ij} de la matriz se dicen los elementos del determinante y det (A) se dice un determinante de orden n. Se escribe

$$\det(A) = |A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

De la definición resulta que los determinantes de segundo y tercer orden tienen la siguiente expresión:

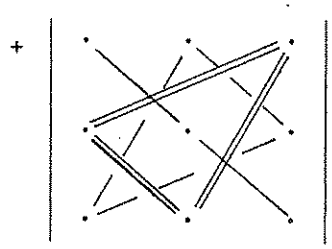
$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

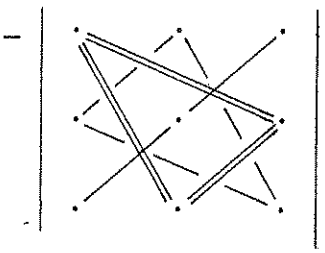
Es decir, el determinante de una matriz de segundo orden es igual al producto de los elementos de la diagonal principal menos el producto de los otros dos.

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$

El determinante de una matriz de tercer orden se obtiene aplicando la siguiente regla llamada regla de Sarrus: Uno de los términos positivos es el producto de los elementos de la diagonal principal, y cada uno de los otros dos se obtiene multiplicando los elementos situados sobre las paralelas a esa diagonal, por el elemento situado en el ángulo opuesto de la matriz. Los términos negativos se obtienen de la misma forma, con respecto a la otra diagonal.

Los diagramas siguientes visualizan la regla anterior:





EJEMPLOS.

1. $\begin{vmatrix} -1 & 2 \\ -5 & 7 \end{vmatrix} = (-1)7 - (-5)2 = 3$

2. $\begin{vmatrix} 3 & -2 & 5 \\ 1 & -4 & 0 \\ -7 & -1 & 3 \end{vmatrix} = 3(-4)3 + 1(-1)5 + (-2)0(-7) - (-7)(-4)5 - (-1)0.3 - 1(-2)3 =$

$= -36 - 5 - 140 + 6 = -175$

3. $\begin{vmatrix} 2 & 0 & 1 \\ -1 & 6 & 3 \\ 4 & 5 & 8 \end{vmatrix} = 2.6.8 + (-1).5.1 + 0.3.4 - 4.6.1 - 5.3.2 - (-1).0.8 =$

$= 96 - 5 - 24 - 30 = 37$

La expresión de un determinante de cuarto orden es más larga y no hay ya una regla sencilla para obtener sus 4! = 24 términos. Es imposible aplicar la definición para calcular determinantes de órdenes superiores. Reflexione el lector que, por ejemplo, un determinante de quinto orden tiene 120 términos, o que uno de décimo orden tiene 3.628.800 términos, que son productos de diez factores cada uno.

La computación efectiva de determinantes es necesaria y tiene gran importancia desde el punto de vista práctico.

En algunos casos de matrices con formas especiales es sencillo calcular el determinante aplicando la definición. Por ejemplo, si se trata de una matriz triangular (ver ejercicio 4 de la página 239) :

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

como según la definición de determinante, en cada término del mismo figura un elemento de cada fila y uno de cada columna, todos los términos se anulan excepto el formado por el producto de los elementos de la diagonal principal. Luego

$$\det(A) = a_{11}a_{22}\cdots\cdots a_{nn}$$

Existen varios métodos para calcular determinantes en general.

En primer lugar vamos a ver las propiedades más importantes de los determinantes y luego indicaremos algunos métodos para calcularlos.

TEOREMA 6.6. El determinante de una matriz coincide con el de su traspuesta.

$$\det(A) = \det(A^T)$$

Demostración: Dada una matriz $n \times n$, $A = (a_{ij})$, su traspuesta es $A^T = (b_{ij})$ donde $b_{ij} = a_{ji}$, para $1 \leq i \leq n$, $1 \leq j \leq n$. Todo término de det(A) es de la forma

$$(-1)^{\nu} a_{1\alpha_1} a_{2\alpha_2} \cdot \ldots \cdot a_{n\alpha_n}$$

Pero $(-1)^{\nu} a_{1\alpha_1} a_{2\alpha_2} \ldots\ldots a_{n\alpha_n} = (-1)^{\nu} b_{\alpha_1 1} b_{\alpha_2 2} \ldots\ldots b_{\alpha_n n}$

y este último es claramente un término de det (A^T) pues figuran n factores, uno de cada fila y uno de cada columna , y también el signo es el que corresponde, según la definición dada. Luego, todo término de det (A) es un término de det (A^T). Análogamente se ve que todo término de $\det(A^T)$ es un término de det (A). En consecuencia $\det(A) = \det(A^T)$.

De este teorema se deduce que: A toda propiedad de un determinante relativa a las columnas le corresponde una análoga para las filas, y viceversa, puesto que las columnas (filas) de una matriz son las filas (columnas) de la matriz traspuesta; de modo que se puede sustituir la palabra "columna" por la palabra "fila", y recíprocamente, en cualquier propiedad.

Las propiedades de los determinantes que veremos a continuación las enunciaremos y demostraremos para las columnas, pero el lector debe tener presente que para las filas valen propiedades análogas.

Se puede considerar que el determinante de una matriz A

$$\text{Det}(A) = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

es función de las columnas de A. Llamando

$$c_1 = (a_{11}, a_{21}, \ldots, a_{n1}) \ , \ c_2 = (a_{12}, a_{22}, \ldots, a_{n2}) \ , \ \ldots \ , \ c_n = (a_{1n}, a_{2n}, \ldots, a_{nn})$$

escribiremos:

$$\det(A) = \det(c_1, c_2, \ldots, c_n)$$

PROPIEDADES DE LOS DETERMINANTES.

1°) $\det(c_1, c_2, \ldots, c_i, \ldots, c_j, \ldots, c_n) = -\det(c_1, c_2, \ldots, c_j, \ldots, c_i, \ldots, c_n)$,

para $1 \leq i < j \leq n$.

O sea

Intercambiando dos columnas cualesquiera el determinante cambia de signo.

Demostración:

$$D = D(c_1,\ldots,c_i,\ldots,c_j,\ldots,c_n) = \begin{vmatrix} a_{11}\cdots & a_{1i}\cdots & a_{1j}\cdots & a_{1n} \\ a_{21}\cdots & a_{2i}\cdots & a_{2j}\cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1}\cdots & a_{ni}\cdots & a_{nj}\cdots & a_{nn} \end{vmatrix}$$

$$D' = D(c_1,\ldots,c_j,\ldots,c_i,\ldots,c_n) = \begin{vmatrix} a_{11}\cdots & a_{1j}\cdots & a_{1i}\cdots & a_{1n} \\ a_{21}\cdots & a_{2j}\cdots & a_{2i}\cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1}\cdots & a_{nj}\cdots & a_{ni}\cdots & a_{nn} \end{vmatrix}$$

Todo término del primer determinante es de la forma $(-1)^{\nu} a_{1\alpha_1} a_{2\alpha_2} \cdots a_{n\alpha_n}$, donde ν es el número de inversiones que presenta la permutación $\alpha_1\alpha_2\ldots\alpha_n$, en la que figuran los índices i,j en un cierto lugar. Se ve que en valor absoluto también es un término del segundo determinante pues figuran n factores, uno de cada fila y uno de cada columna. Pero pensado como un término de D', el número total de inversiones de la permutación $\alpha_1\alpha_2\ldots\alpha_n$ es mayor o menor que ν en una unidad, puesto que los índices i,j aparecen traspuestos en D', y el signo que le corresponde es entonces el opuesto al que tiene como término de D. Luego, todo término de D con el signo cambiado es un término de D', y recíprocamente. Entonces

$$D' = -D$$

2°) $\det(c_1,\ldots,c_{j-1},kc_j+k'c'_j, c_{j+1},\ldots, c_n) = k\det(c_1,\ldots,c_{j-1},c_j,c_{j+1},\ldots,c_n) +$

$$+ k'\det(c_1,\ldots,c_{j-1},c'_j,c_{j+1},\ldots,c_n)$$

con $k,k' \in K$, cualquiera que sea $j=1,2,\ldots,n$.

Es decir,

Si los elementos de la j-ésima columna de una matriz A son combinaciones lineales $k a_{ij} + k' a'_{ij}$, para $i=1,2,\ldots,n$, entonces

$$\det(A) = k\det(A_1) + k'\det(A_2)$$

donde A_1 y A_2 son matrices que tienen todos sus elementos iguales a los de A, excepto los de la j-ésima columna que son $a_{1j}, a_{2j}, \ldots, a_{nj}$ en A_1 y $a'_{1j}, a'_{2j}, \ldots, a'_{nj}$ en A_2.

Demostración: Sea

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & \underset{\underset{j}{\downarrow}}{ka_{1j}+k'a'_{1j}} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & ka_{2j}+k'a'_{2j} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & ka_{nj}+k'a'_{nj} & \cdots & a_{nn} \end{pmatrix}$$

Queremos probar que:

$$\det(A) = k \begin{vmatrix} a_{11} & a_{12} & \cdots & \underset{\underset{j}{\downarrow}}{a_{1j}} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix} + k' \begin{vmatrix} a_{11} & a_{12} & \cdots & \underset{\underset{j}{\downarrow}}{a'_{1j}} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a'_{2j} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a'_{nj} & \cdots & a_{nn} \end{vmatrix}$$

Cada término de det(A) puede escribirse en la forma

$$(-1)^{\nu} a_{\alpha_1 1} a_{\alpha_2 2} \cdots (ka_{\alpha_j j}+k'a'_{\alpha_j j}) \cdots a_{\alpha_n n}$$

ordenando los factores de modo que los índices que indican las columnas formen la permutación $1\ 2\ \ldots\ n$, y donde ν es el número de inversiones de la permutación $\alpha_1\ \alpha_2\ \ldots\ \alpha_n$ correspondiente a los primeros índices.

Aplicando las propiedades distributiva, conmutativa y asociativa de la suma y la multiplicación en K, se tiene:

$$\det(A) = \sum (-1)^{\nu} a_{\alpha_1 1} a_{\alpha_2 2} \cdots (ka_{\alpha_j j}+k'a'_{\alpha_j j}) \cdots a_{\alpha_n n} =$$

$$= \sum (-1)^{\nu} (a_{\alpha_1 1} a_{\alpha_2 2} \cdots ka_{\alpha_j j} \cdots a_{\alpha_n n} + a_{\alpha_j 1} a_{\alpha_2 2} \cdots k'a'_{\alpha_j j} \cdots a_{\alpha_n n}) =$$

$$= k \sum (-1)^{\nu} a_{\alpha_1 1} a_{\alpha_2 2} \cdots a_{\alpha_j j} \cdots a_{\alpha_n n} + k' \sum (-1)^{\nu} a_{\alpha_1 1} a_{\alpha_2 2} \cdots a'_{\alpha_j j} \cdots a_{\alpha_n n}$$

Cada sumatoria de esta última expresión es el determinante de una matriz que tiene todas las columnas iguales a las de A excepto la j-ésima formada por los elementos $a_{1j}, a_{2j}, \ldots, a_{nj}$ en el primer caso, y $a'_{1j}, a'_{2j}, \ldots, a'_{nj}$ en el segundo, lo que termina la demostración.

3°) $\det(e_1, e_2, \ldots, e_n) = 1$

donde $e_i = (0, 0, \ldots, \underset{\underset{i}{\downarrow}}{1}, \ldots, 0)$ para $i = 1, 2, \ldots, n$

Es decir

El determinante de la matriz unidad es 1.

Demostración: Como en la matriz unidad $I = (a_{ij})$ es $a_{ij} = 0$ si $i \neq j$ y $a_{ii} = 1$ para $i = 1, 2, \ldots, n$, de la definición de determinante resulta que el único término no nu

lo es el de la forma

$$a_{11}a_{22}\cdot\ \ldots\ \cdot a_{nn}$$

Luego $\det(I) = 1$.

OBSERVACION. Se demuestra que las propiedades 1° , 2° y 3° caracterizan la noción de determinante de una matriz, es decir, que si una aplicación $f: M_n(K) \longrightarrow K$ verifica esas propiedades entonces necesariamente es $f(A) = \det(A)$, $\forall\ A \in M_n(K)$. Luego , existe una y sólo una aplicación $M_n(K) \longrightarrow K$ con las propiedades 1° , 2° y 3°: la de determinante que hemos definido.

De las propiedades anteriores se deducen las siguientes:

4°) Si en un determinante hay dos columnas iguales el determinante es cero.

En efecto, intercambiando en D las dos columnas iguales, por un lado D no varía, y por otro cambia de signo por la propiedad 1°. Luego

$$D = -D \Longrightarrow 2D = 0 \Longrightarrow D = 0$$

5°) Si todos los elementos de una columna de un determinante se multiplican por un escalar $k \in K$, el determinante queda multiplicado por k. En símbolos

$$\det(c_1,\ldots,kc_i,\ldots,c_n) = k\cdot\det(c_1,\ldots,c_i,\ldots,c_n) \quad , \quad k \in K$$

Resulta facilmente de la propiedad 2°, haciendo $k'=0$.

6°) Si una columna de un determinante es combinación lineal de otras entonces el determinante es cero.

Sea el determinante $\det(c_1,c_2,\ldots,c_n)$ y supongamos que la columna c_j es combinación lineal de las columnas c_{j_1} , c_{j_2} , . . . , c_{j_s} :

$$c_j = k_1c_{j_1} + k_2c_{j_2} + \ldots\ + k_sc_{j_s} \quad , \ k_1 ,k_2 ,\ldots,k_s \in K .$$

Entonces, aplicando la propiedad 2° s-1 veces se tiene:

$$\det(c_1,\ldots,c_j,\ldots,c_n) = k_1\ \det(c_1,\ldots,\underset{j}{c_{j_1}},\ldots,c_n) + k_2\ \det(c_1,\ldots,\underset{j}{c_{j_2}},\ldots,c_n) +$$
$$\ldots\ldots + k_s\ \det(c_1,\ldots,\underset{j}{c_{j_s}},\ldots,c_n)$$

En cada uno de los determinantes del segundo miembro hay dos columnas iguales: la que ocupa el lugar j es igual a la que ocupa el lugar j_i , $i = 1,2,\ldots,s$. Entonces, por la propiedad 4°, resultan todos iguales a cero, lo que termina la demostración.

7°) Sumándole a una columna de un determinante una combinación lineal de otras, el valor del determinante no varía.

Sea el determinante $\det(c_1,\ldots,c_n)$ y supongamos que a la columna c_j se le suma una combinación lineal de otras columnas:

$$k_1c_{j_1} + k_2c_{j_2} + \ldots + k_sc_{j_s} \quad , \quad k_1,k_2,\ldots,k_s \in K$$

Entonces

$$\det(c_1,\ldots, c_j + \sum_{i=1}^{s} k_i c_{j_i},\ldots,c_n) = \det(c_1,\ldots,c_j,\ldots,c_n) + \det(c_1,\ldots, \sum_{i=1}^{s} k_i c_{j_i},\ldots,c_n)$$

y el último determinante del segundo miembro es cero por la propiedad 6°, de donde resulta lo que queríamos demostrar.

CALCULO DE DETERMINANTES.

De las propiedades 1°, 5° y 7° recién vistas resulta que, dada una matriz A, pasando de A a otra matriz por una operación elemental, el determinante cambia de signo si la operación aplicada es del tipo I, queda multiplicado por un escalar no nulo si la operación es del tipo II, y no varía si se aplica una operación del tipo III. Entonces, como es muy fácil calcular el determinante de una matriz triangular, para calcular el determinante de una matriz A se puede pasar mediante operaciones elementales a una matriz triangular, cuyo determinante diferirá del de la dada a lo sumo en un escalar. Este es un método general y práctico para calcular determinantes.

Vamos a indicar otro método, que reduce el cálculo de un determinante de orden n al cálculo de determinantes de orden n-1, y luego daremos algunos ejemplos.

Desarrollo por los elementos de una línea.

Si en una matriz cuadrada de orden n se suprimen la fila i-ésima y la columna j-ésima se obtiene una submatriz de orden n-1 cuyo determinante M_{ij} se llama el menor complementario del elemento a_{ij} común a la fila y a la columna suprimidas.

Se llama complemento algebraico o cofactor del elemento a_{ij} a

$$(-1)^{i+j} M_{ij}$$

Se trata de demostrar que el valor del determinante de una matriz cuadrada es igual a la suma de los elementos de una línea (fila o columna) multiplicados por sus respectivos complementos algebraicos.

Si en el desarrollo de un determinante D sacamos factor común a_{ij} en todos los términos en que aparece, a_{ij} quedará multiplicando a una cierta suma A_{ij} :

$$D = a_{ij} A_{ij} + \ldots\ldots$$

Veamos a qué es igual A_{ij} .

Consideremos primero el elemento a_{11} en un determinante cualquiera. Todo término en que él figura es de la forma $(-1)^{\nu} a_{11} a_{2\alpha_2} \ldots a_{n\alpha_n}$, donde ν es el número de inversiones de la permutación $1\alpha_2 \ldots\ldots \alpha_n$. Separando el factor a_{11} queda $(-1)^{\nu} a_{2\alpha_2} \ldots a_{n\alpha_n}$, que es un término del determinante

$$M_{11} = \begin{vmatrix} a_{22} & \ldots\ldots & a_{2n} \\ a_{32} & \ldots\ldots & a_{3n} \\ \ldots & \ldots\ldots & \ldots \\ a_{n2} & \ldots\ldots & a_{nn} \end{vmatrix}$$

En efecto, en dicho producto figuran n-1 factores, uno de cada fila y uno de cada columna de M_{11}, y el signo $(-1)^{\nu}$ es el que corresponde porque el número de inversiones de la permutación $\alpha_2 \ldots\ldots \alpha_n$ es igual al de la permutación $1\alpha_2 \ldots\ldots \alpha_n$, ya que 1 no forma inversión con ningún índice.

Recíprocamente, todo término de M_{11} multiplicado por a_{11} es un término del determinante dado. Luego, sacando a_{11} factor común, lo que multiplica a_{11} es M_{11} .

Consideremos ahora el elemento a_{ij} en un determinante D. Para reducir este caso al anterior llevamos a a_{ij} mediante i-1 intercambios de filas y j-1 intercambios de columnas al lugar de a_{11}.

$$D' = \begin{vmatrix} a_{ij} & a_{i1} & a_{i2} & \ldots\ldots & a_{in} \\ a_{1j} & a_{11} & a_{12} & \ldots\ldots & a_{1n} \\ a_{2j} & a_{21} & a_{22} & \ldots\ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots\ldots & \ldots \\ a_{nj} & a_{n1} & a_{n2} & \ldots\ldots & a_{nn} \end{vmatrix}$$

Los menores complementarios de a_{ij} en D y en D' coinciden. Se tiene que:

$$D = (-1)^{(i-1)+(j-1)} D' = (-1)^{i+j} D' = (-1)^{i+j} \left[a_{ij} M_{ij} + \ldots\ldots \right] =$$

$$= a_{ij} \left[(-1)^{i+j} M_{ij} \right] + \ldots\ldots$$

Luego $A_{ij} = (-1)^{i+j} M_{ij}$. Queda probado así que A_{ij} es el complemento algebraico de a_{ij}.

TEOREMA 6.7. El determinante de una matriz es igual a la suma de los elementos de una línea cualquiera multiplicados por sus respectivos complementos algebraicos.

Demostración: Sea A una matriz de orden n y elijamos una línea cualquiera, por ejemplo, la fila i. Los n! términos del determinante det(A) se pueden agrupar de la siguiente manera: todos aquéllos en que figura a_{i1}, cuya suma es $a_{i1}A_{i1}$; todos en los que figura a_{i2}, cuya suma es $a_{i2}A_{i2}$; . . . . . ; todos en los que aparece a_{in}, que sumados dan $a_{in}A_{in}$. Así quedan clasificados todos los términos del determinante, porque en cualquier término figura un y sólo un factor de la fila i. Luego

$$\det(A) = a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in}$$

La expresión anterior se llama el desarrollo del determinante por los elementos de una línea.

COROLARIO. Dada una matriz A, la suma de los elementos de una línea multiplicados por los complementos algebraicos correspondientes a los elementos de otra línea paralela es cero.

$$a_{i1}A_{j1} + a_{i2}A_{j2} + \ldots + a_{in}A_{jn} = 0 \qquad \text{si} \quad i \neq j$$

En efecto, el primer miembro es el desarrollo del determinante de la matriz que se obtiene reemplazando en A la fila j por la fila i. Luego tiene dos filas iguales y por lo tanto es cero.

El teorema anterior reduce el cálculo de un determinante de orden n al cálculo de varios determinantes de orden n-1. Es claro que si en la línea elegida hay elementos nulos entonces se abrevian los cálculos. Conviene entonces elegir una línea con el mayor número de ceros, de ser ello posible. Además, para calcular un determinante, también se puede aplicar previamente la propiedad 7°, sumando a ciertas filas o columnas combinaciones lineales convenientes de otras líneas paralelas, de modo de reducir a cero todos los elementos de una línea, excepto uno. Entonces el cálculo del determinante de orden n dado se reduce al cálculo de un solo determinante de orden n-1.

EJEMPLOS.

1) Desarrollar el siguiente determinante por los elementos de la segunda fila.

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

$$= 4.13 + 3.(-1) + (-1).(-5) = 54$$

2) Calcular el siguiente determinante de dos maneras, desarrollándolo por los elementos de una línea, y pasando a la forma triangular.

$$D = \begin{vmatrix} 3 & 2 & 5 & 1 \\ -1 & 1 & 0 & 4 \\ 6 & 0 & -2 & 1 \\ 1 & -1 & 0 & 3 \end{vmatrix}$$

a) Conviene desarrollarlo por los elementos de la tercera columna.

$$D = (-1)^{1+3}5 \begin{vmatrix} -1 & 1 & 4 \\ 6 & 0 & 1 \\ 1 & -1 & 3 \end{vmatrix} + (-1)^{3+3}(-2) \begin{vmatrix} 3 & 2 & 1 \\ -1 & 1 & 4 \\ 1 & -1 & 3 \end{vmatrix} = 5(-42)+(-2)35 = -280$$

b) Pasando a la forma triangular:

$$D = -\begin{vmatrix} -1 & 1 & 0 & 4 \\ 3 & 2 & 5 & 1 \\ 6 & 0 & -2 & 1 \\ 1 & -1 & 0 & 3 \end{vmatrix} = -\begin{vmatrix} -1 & 1 & 0 & 4 \\ 0 & 5 & 5 & 13 \\ 0 & 6 & -2 & 25 \\ 0 & 0 & 0 & 7 \end{vmatrix} = -5\begin{vmatrix} -1 & 1 & 0 & 4 \\ 0 & 1 & 1 & 13/5 \\ 0 & 0 & -8 & 47/5 \\ 0 & 0 & 0 & 7 \end{vmatrix} =$$

$$= (-5)(-1).1.(-8).7 = -280$$

3) Calcular el siguiente determinante de quinto orden:

$$D = \begin{vmatrix} 4 & 2 & 0 & -1 & 3 \\ 3 & 1 & -2 & 3 & -7 \\ -2 & 6 & 0 & 2 & 1 \\ 1 & 2 & -1 & 1 & -2 \\ -1 & 3 & 4 & -2 & 5 \end{vmatrix}$$

Nos conviene reducir todos los elementos de la tercera columna a cero, excepto uno. Restando a la segunda fila la cuarta multiplicada por 2 y sumando a la quinta fila la cuarta multiplicada por 4, obtenemos:

$$D = \begin{vmatrix} 4 & 2 & 0 & -1 & 3 \\ 1 & -3 & 0 & 1 & -3 \\ -2 & 6 & 0 & 2 & 1 \\ 1 & 2 & -1 & 1 & -2 \\ 3 & 11 & 0 & 2 & 1 \end{vmatrix}$$

Desarrollando este determinante por los elementos de la tercera columna se tiene:

$$D = (-1)^{3+4}\ (-1)\begin{vmatrix} 4 & 2 & -1 & 3 \\ 1 & -3 & 1 & -3 \\ -2 & 6 & 2 & 1 \\ 3 & 11 & 2 & 1 \end{vmatrix}$$

Es fácil reducir a cero todos los elementos de la primera columna, excepto el segundo, restando a la primera fila la segunda multiplicada por 4; sumando a la tercera fila la segunda multiplicada por 2; y restando a la cuarta fila la segunda multiplicada por 3:

$$D = \begin{vmatrix} 0 & 14 & -5 & 15 \\ 1 & -3 & 1 & -3 \\ 0 & 0 & 4 & -5 \\ 0 & 20 & -1 & 10 \end{vmatrix}$$

Luego, desarrollando por la primera columna es:

$$D = -\begin{vmatrix} 14 & -5 & 15 \\ 0 & 4 & -5 \\ 20 & -1 & 10 \end{vmatrix} = -(560 + 500 - 1200 - 70) = 210$$

El desarrollo de un determinante por los elementos de una línea es un caso particular de una regla más general para resolver un determinante de orden n en función de determinantes de menor orden, que se conoce con el nombre de REGLA DE LAPLACE.

En general, dada una matriz A de orden n y elegidas k filas y k columnas, $1 \leq k \leq n-1$, los elementos comunes a esas filas y columnas forman una submatriz de orden k, cuyo determinante se llama un menor de orden k de det(A). Suprimiendo en la matriz A esas k filas y k columnas, se obtiene una submatriz cuadrada de orden n-k, cuyo determinante se llama el menor complementario del anterior.

Si un menor de orden k M está formado por las filas $i_1, i_2, \ldots, i_k$ y las columnas $j_1, j_2, \ldots, j_k$, se adjunta a su menor complementario M* el signo + ó - según que la suma de los índices de las filas y las columnas, $\epsilon = i_1+i_2+\ldots+i_k+j_1+j_2+\ldots+j_k$, sea par o impar.

$(-1)^{\epsilon}$ M* se llama el cofactor o complemento algebraico del menor M.

REGLA DE LAPLACE. Elegidas k filas (o k columnas) arbitrarias en un determinante de orden n, $1 \leq k \leq n-1$, el determinante es igual a la suma de los productos de todos los menores de orden k que se pueden formar con esas k filas por sus respectivos complementos algebraicos.

La regla de Laplace reduce el cálculo de un determinante de orden n al cálculo de determinantes de órdenes k y n-k. Cuando se toma k = 1 se tiene el desarrollo del determinante por los elementos de una línea ya visto.

EJEMPLOS.

1) Calculemos el determinante

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

Eligiendo la primera y la tercera filas todos los menores que con ellas se pueden formar son:

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

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

De ellos siete son nulos. Aplicando la regla de Laplace se tiene:

$$D = (-1)^{1+3+1+2} \begin{vmatrix} 2 & -1 \\ 2 & -3 \end{vmatrix} \begin{vmatrix} 3 & 1 & 4 \\ 0 & 2 & -1 \\ 3 & 1 & 0 \end{vmatrix} + (-1)^{1+3+1+4} \begin{vmatrix} 2 & -2 \\ 2 & 1 \end{vmatrix} \begin{vmatrix} 1 & 3 & 4 \\ 4 & 0 & -1 \\ -2 & 3 & 0 \end{vmatrix} +$$

$$+ (-1)^{1+3+2+4} \begin{vmatrix} -1 & -2 \\ -3 & 1 \end{vmatrix} \begin{vmatrix} -1 & 3 & 4 \\ 0 & 0 & -1 \\ 1 & 3 & 0 \end{vmatrix}$$

Haciendo cálculos

$$D = (-4)(-24) - 6.57 + (-7)(-6) = -204$$

Aplicando la regla de Laplace se deduce que un determinante de la forma:

$$D = \begin{vmatrix} a_{11} & \cdots & a_{1k} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{k1} & \cdots & a_{kk} & 0 & \cdots & 0 \\ a_{k+1,1} & \cdots & a_{k+1,k} & a_{k+1,k+1} & \cdots & a_{k+1,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nk} & a_{n,k+1} & \cdots & a_{nn} \end{vmatrix}$$

es igual al producto de dos menores:

$$D = \begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \cdots & \cdots & \cdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix} \cdot \begin{vmatrix} a_{k+1,k+1} & \cdots & a_{k+1,n} \\ \cdots & \cdots & \cdots \\ a_{n,k+1} & \cdots & a_{nn} \end{vmatrix}$$

ya que la suma de los índices de las k primeras filas y columnas es par.

DETERMINANTE DE UN PRODUCTO DE MATRICES.

TEOREMA 6.8. El determinante de un producto de dos matrices cuadradas de orden n es igual al producto de sus determinantes.

$$\det(A.B) = \det(A).\ \det(B)$$

Demostración: Sean $A = (a_{ij})$, $B = (b_{ij})$ dos matrices cuadradas de orden n. Consideremos el siguiente determinante auxiliar:

$$D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots & 0 \\ -1 & 0 & \cdots & 0 & b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & -1 & \cdots & 0 & b_{21} & b_{22} & \cdots & b_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & -1 & b_{n1} & b_{n2} & \cdots & b_{nn} \end{vmatrix}$$

Calculando este determinante por la regla de Laplace y eligiendo las n primeras filas para aplicarla, se tiene:

$$D = \det(A).\det(B) \qquad (1)$$

Vamos a calcular el mismo determinante por otro procedimiento. Se trata de transformar el determinante dado sin alterar su valor de modo que los lugares ocupados por los elementos b_{ij} queden todos ocupados por ceros.

Para esto se le suma a la columna n+1 la primera columna multiplicada por b_{11}, la segunda multiplicada por b_{21}, . . . , la n-ésima multiplicada por b_{n1}. A la columna n+2 se le suma la primera columna multiplicada por b_{12}, la segunda multiplicada por b_{22}, . . . , la n-ésima multiplicada por b_{n2}. Procediendo análogamente con todas

las columnas restantes, finalmente se obtiene:

$$D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & c_{11} & c_{12} & \cdots & c_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} & c_{21} & c_{22} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & c_{n1} & c_{n2} & \cdots & c_{nn} \\ -1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & -1 & 0 & 0 & \cdots & 0 \end{vmatrix}$$

Desarrollando ahora este determinante por la regla de Laplace por los menores de las n últimas filas se tiene:

$$D = (-1)^{\varepsilon} \begin{vmatrix} -1 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & -1 \end{vmatrix} \begin{vmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{vmatrix}$$

siendo $\varepsilon = (n+1)+(n+2)+ \ldots + 2n + 1 + 2 + \ldots + n = 2n^2 + n$.

Como el primer determinante es igual a $(-1)^n$ y $\varepsilon+n = 2n^2 + 2n$ es par, resulta:

$$D = \begin{vmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{vmatrix} \qquad (2)$$

Ahora bien, por las transformaciones efectuadas se ve que el elemento c_{ij} es:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}$$

para todo par de índices i,j , $1 \leq i \leq n$, $1 \leq j \leq n$

Luego $(c_{ij}) = A.B$

En consecuencia, (2) puede escribirse:

$$D = \det(A.B) \qquad (3)$$

De (1) y (3) sigue entonces

$$\det(A.B) = \det(A).\det(B) \qquad \text{c.q.d.}$$

MATRICES INVERSIBLES.

Estamos ahora en condiciones de dar otra caracterización de las matrices inversibles de orden n e indicar otro método para calcular la matriz inversa usando determinantes.

Introduciremos la noción de matriz adjunta.

Definición. Dada una matriz $n \times n$ $A = (a_{ij})$ se llama adjunta de A a la matriz

$$\text{adj } A = \begin{pmatrix} A_{11} & A_{21} & A_{31} & \dots & A_{n1} \\ A_{12} & A_{22} & A_{32} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & A_{3n} & \dots & A_{nn} \end{pmatrix}$$

donde A_{ij} = complemento algebraico del elemento a_{ij} en A , $\forall\ 1 \leq i \leq n$, $1 \leq j \leq n$.

Notemos que en adj A el lugar i,j está ocupado por A_{ji} .

EJEMPLO.

La adjunta de la matriz

$$A = \begin{pmatrix} -1 & 0 & 2 \\ 3 & 1 & 5 \\ 0 & -2 & -1 \end{pmatrix} \qquad \text{es} \qquad \text{adj } A = \begin{pmatrix} 9 & -4 & -2 \\ 3 & 1 & 11 \\ -6 & -2 & -1 \end{pmatrix}$$

Multiplicando una matriz por su adjunta, y teniendo en cuenta el teorema 6.7 , que da el desarrollo de un determinante por los elementos de una línea, y su corolario, resulta:

$$A.\text{adj } A = \text{adj } A \,.\, A = \begin{pmatrix} \det(A) & 0 & 0 \\ 0 & \det(A) & 0 \\ \dots & \dots & \dots \\ 0 & 0 & \det(A) \end{pmatrix}$$

Como

$$\begin{pmatrix} \det(A) & 0 & 0 \\ 0 & \det(A) & 0 \\ \dots & \dots & \dots \\ 0 & 0 & \det(A) \end{pmatrix} = \det(A) \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

se tiene que:

$$A.\text{adj } A = \text{adj } A \,.\, A = \det(A).I \qquad (1)$$

Si $\det(A) \neq 0$, de (1) resulta que la matriz $A^{-1} = \frac{1}{\det(A)} \text{adj } A$ es tal que:

$$A.A^{-1} = A^{-1}.A = I$$

En consecuencia A es inversible y su inversa es:

$$A^{-1} = \begin{pmatrix} \frac{A_{11}}{\det(A)} & \frac{A_{21}}{\det(A)} & \cdots & \frac{A_{n1}}{\det(A)} \\ \frac{A_{12}}{\det(A)} & \frac{A_{22}}{\det(A)} & \cdots & \frac{A_{n2}}{\det(A)} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{A_{1n}}{\det(A)} & \frac{A_{2n}}{\det(A)} & \cdots & \frac{A_{nn}}{\det(A)} \end{pmatrix}$$

TEOREMA 6.9. Si A es una matriz de orden n, A inversible $\Longleftrightarrow$ $\det(A) \neq 0$.

Demostración: Probemos que A inversible implica $\det(A) \neq 0$. Si A es inversible, existe una matriz A^{-1} tal que

$$A.A^{-1} = I$$

Por el teorema 6.8 es entonces

$$\det(A).\det(A^{-1}) = \det(I)$$

Pero $\det(I) = 1$. Luego $\det(A).\det(A^{-1}) \neq 0 \Longrightarrow \det(A) \neq 0$.

La otra implicación ya la demostramos: si $\det(A) \neq 0$, la matriz $A^{-1} = \frac{1}{\det(A)} \text{adj } A$ es inversa de A .

EJEMPLOS.

1) Decir si la siguiente matriz es inversible aplicando el teorema anterior y en tal caso hallar la inversa.

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

Calculando el determinante de A por la regla de Sarrus se encuentra que $|A| = 23$. Luego A es inversible. Su inversa es

$$A^{-1} = \begin{pmatrix} \frac{4}{23} & \frac{3}{23} & \frac{6}{23} \\ \frac{9}{23} & \frac{1}{23} & \frac{2}{23} \\ -\frac{1}{23} & \frac{5}{23} & -\frac{13}{23} \end{pmatrix}$$

2) Idem para la matriz

$$A = \begin{pmatrix} -1 & 3 \\ 7 & -21 \end{pmatrix}$$

Como $|A| = 0$, A no es inversible.

Como aplicación del teorema anterior podemos concluir que un sistema de n ecuaciones lineales con n incógnitas AX = Y tiene una única solución si y sólo si $\det(A) \neq 0$. En efecto, si A es una matriz n×n , aplicando el teorema recién demostrado y el teorema 6.5 se tiene:

$\det(A) \neq 0 \iff$ A inversible $\iff$ El sistema AX = Y tiene una única solución para todo Y .

La única solución de un sistema de este tipo se puede calcular de acuerdo con el siguiente

TEOREMA 6.10. Si AX = Y es un sistema de n ecuaciones lineales con n incógnitas tal que $\det(A) \neq 0$, la única solución del sistema está dada por la fórmula

$$x_i = \frac{\det(c_1,\ldots,c_{i-1},Y,c_{i+1},\ldots,c_n)}{\det(A)} \quad , \quad i = 1,2,\ldots,n$$

donde c_1 , c_2 , ... , c_n son las columnas de A.

Demostración: Sea AX = Y un sistema de n ecuaciones con n incógnitas tal que $\det(A) \neq 0$. Probemos que la fórmula indicada da el valor de las incógnitas.

$$AX = Y \Longrightarrow A^{-1}AX = A^{-1}Y \Longrightarrow X = A^{-1}Y$$

Como $A^{-1} = \frac{1}{\det(A)} \text{ adj } A$ es $X = \frac{1}{\det(A)} \text{ adj } A.Y$

Luego

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \frac{1}{\det(A)} \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \frac{1}{\det(A)} \begin{pmatrix} A_{11}b_1+A_{21}b_2+\ldots+A_{n1}b_n \\ A_{12}b_1+A_{22}b_2+\ldots+A_{n2}b_n \\ \cdots\cdots\cdots\cdots \\ A_{1n}b_1+A_{2n}b_2+\ldots+A_{nn}b_n \end{pmatrix}$$

Notemos que el elemento que figura en la primera fila de la matriz n×1 que aparece en último término es el desarrollo por los elementos de la primera columna del determinante de la matriz que se obtiene reemplazando en A la primera columna por $b_1, b_2, \ldots, b_n$, o sea por Y; el elemento que aparece en la segunda fila es el desarrollo por los elementos de la segunda columna del determinante de la matriz que se obtiene reemplazando en A la segunda columna por $b_1, b_2, \ldots, b_n$; etc., etc. ..... Es decir

$$A_{1i}b_1 + A_{2i}b_2 + \cdots\cdots + A_{ni}b_n = \det(c_1, \ldots, c_{i-1}, Y, c_{i+1}, \ldots, c_n)$$

para $i = 1, 2, \ldots, n$

Luego

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \frac{1}{\det(A)} \begin{pmatrix} \det(Y, c_2, \ldots, c_n) \\ \det(c_1, Y, \ldots, c_n) \\ \cdots\cdots\cdots \\ \det(c_1, c_2, \ldots, Y) \end{pmatrix}$$

lo que termina la demostración.

Este teorema se conoce con el nombre de REGLA DE CRAMER: Un sistema de n ecuaciones lineales con n incógnitas donde el determinante de la matriz de los coeficientes no es nulo tiene una única solución; el valor de cada incógnita se obtiene dividiendo por el determinante de la matriz de los coeficientes el determinante que se obtiene sustituyendo en él la columna de los coeficientes de dicha incógnita por los términos independientes.

EJEMPLO.

Verificar que el siguiente sistema puede resolverse por la regla de Cramer y calcular su solución.

$$\begin{cases} 2x_1 + 3x_2 - x_3 = -1 \\ x_1 + 4x_3 = 2 \\ x_1 - x_2 - x_3 = 7 \end{cases}$$

$$\begin{vmatrix} 2 & 3 & -1 \\ 1 & 0 & 4 \\ 1 & -1 & -1 \end{vmatrix} = 12 + 1 + 8 + 3 = 24 \neq 0$$

Como el determinante de la matriz de los coeficientes no es nulo puede aplicarse la regla de Cramer. Los valores de las incógnitas son:

$$x_1 = \frac{\begin{vmatrix} -1 & 3 & -1 \\ 2 & 0 & 4 \\ 7 & -1 & -1 \end{vmatrix}}{24} \quad ; \quad x_2 = \frac{\begin{vmatrix} 2 & -1 & -1 \\ 1 & 2 & 4 \\ 1 & 7 & -1 \end{vmatrix}}{24} \quad ; \quad x_3 = \frac{\begin{vmatrix} 2 & 3 & -1 \\ 1 & 0 & 2 \\ 1 & -1 & 7 \end{vmatrix}}{24}$$

Luego $\left(\frac{11}{3} , -\frac{35}{12} , -\frac{5}{12} \right)$ es la solución del sistema. (El lector puede verificarlo reemplazando en las ecuaciones).

EJERCICIOS.

1. Calcular el determinante de las siguientes matrices:

$$\begin{pmatrix} -1 & 2 \\ 3 & 4 \end{pmatrix} \qquad \begin{pmatrix} -3i & 1+i \\ 2-i & -2 \end{pmatrix} \qquad \begin{pmatrix} \frac{1}{2} & -1 & 0 \\ 3 & 2 & 5 \\ -1 & 2 & -3 \end{pmatrix} \qquad \begin{pmatrix} -4 & 1 & -2 \\ 3 & -1 & 5 \\ -6 & 1 & 2 \end{pmatrix}$$

$$\begin{pmatrix} \sqrt{2} & 0 & \frac{1}{4} \\ -3 & -1 & -\sqrt{2} \\ 4 & -5 & 2 \end{pmatrix}$$

2. Hallar los complementos algebraicos de los elementos a_{12} , a_{24} , a_{33} y a_{43} de la matriz:

$$\begin{pmatrix} 2 & -1 & 3 & 0 \\ 5 & -3 & 1 & 4 \\ 0 & -2 & 1 & -1 \\ 2 & -1 & 6 & 7 \end{pmatrix}$$

3. Calcular el determinante de las matrices dadas en cada una de las tres formas siguientes:

   a) Desarrollando por los elementos de una fila.

   b) " " " " " " columna.

   c) Pasando previamente a una matriz triangular.

$$\begin{pmatrix} 1 & -1 & 2 & 1 \\ 1 & -1 & 4 & -1 \\ 1 & 1 & -2 & 3 \\ 1 & 1 & 1 & 1 \end{pmatrix} , \quad \begin{pmatrix} 1 & -2 & 3 & 0 \\ -5 & 2 & 1 & -2 \\ -1 & 6 & 2 & 4 \\ 3 & -2 & 1 & -1 \end{pmatrix} , \quad \begin{pmatrix} -2 & \frac{1}{2} & 0 & -7 \\ 0 & 3 & 4 & 0 \\ 0 & 5 & 1 & -1 \\ -3 & 2 & 0 & -2 \end{pmatrix}$$

4. Hallar el determinante de cada una de las siguientes matrices pasando previamente a otra matriz con una línea (fila o columna) con todos los elementos nulos excepto uno.

$$\begin{vmatrix} 2 & 3 & -1 \\ -3 & -1 & 1 \\ 5 & 2 & -2 \end{vmatrix} \quad , \quad \begin{vmatrix} -7 & 5 & 4 \\ -2 & 3 & -3 \\ 9 & -10 & 6 \end{vmatrix} \quad , \quad \begin{vmatrix} -3 & 2 & -5 & 1 \\ 2 & -1 & 1 & -2 \\ -2 & 3 & 0 & -1 \\ 4 & -6 & -2 & 3 \end{vmatrix}$$

5. a) Decir porqué son nulos los determinantes de las siguientes matrices sin calcularlos:

$$\begin{vmatrix} 5 & -1 & 0 & 2 & -3 \\ 3 & 4 & 1 & -5 & 10 \\ -10 & 2 & 0 & -4 & 6 \\ 2 & 1 & 2 & 0 & 1 \\ 0 & -3 & 1 & 1 & -1 \end{vmatrix} \quad , \quad \begin{vmatrix} -1 & 0 & -1 & -2 \\ 3 & 2 & 5 & 0 \\ 4 & -1 & 3 & -4 \\ -2 & 3 & 1 & 7 \end{vmatrix}$$

b) Sin calcular los determinantes explicar porqué valen las siguientes igualdades:

$$\begin{vmatrix} a & b & c \\ p & q & r \\ x & y & z \end{vmatrix} = \begin{vmatrix} a & p & x \\ b & q & y \\ c & r & z \end{vmatrix} = \begin{vmatrix} b & q & y \\ c & r & z \\ a & p & x \end{vmatrix} = \begin{vmatrix} p & r & q \\ a & c & b \\ x & z & y \end{vmatrix}$$

c) Igual que en b) siendo:

$$\begin{vmatrix} 3 & 24 & -9 \\ -6 & -24 & -4 \\ 18 & 0 & 1 \end{vmatrix} = 3.12.(-2) \begin{vmatrix} 1 & 2 & -9 \\ 1 & 1 & 2 \\ 6 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 4 & 2 & 54 \\ 12 & 3 & -36 \\ 24 & 0 & -6 \end{vmatrix}$$

6. Calcular aplicando la regla de Laplace los siguientes determinantes:

$$\begin{vmatrix} 2 & 0 & -1 & 5 & 0 \\ -1 & 0 & 1 & 3 & 0 \\ 4 & 1 & -2 & 1 & 1 \\ 0 & 6 & 3 & 2 & -1 \\ 1 & 2 & -1 & 5 & 2 \end{vmatrix} \quad , \quad \begin{vmatrix} 1 & 1 & 3 & 2 & 1 & 5 \\ 1 & 0 & 4 & 0 & 0 & -1 \\ 3 & 0 & -1 & 0 & 0 & 3 \\ 2 & 0 & -2 & 0 & 0 & -4 \\ 7 & -3 & 1 & 2 & -1 & 3 \\ 4 & 1 & -1 & 1 & 1 & 1 \end{vmatrix}$$

7. Un determinante de la forma:

$$D_n = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ \cdots & \cdots & \cdots & \cdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{vmatrix}$$

se llama determinante de VANDERMONDE. Se verifica que

$$D_n = \prod_{i>j}(x_i - x_j) \qquad \text{o sea}$$

$$D_n = (x_2-x_1)(x_3-x_1) \ldots (x_n-x_1)(x_3-x_2)(x_4-x_2) \ldots (x_n-x_2) \ldots (x_n-x_{n-1})$$

a) Demostrar la fórmula para n = 3 y n = 4 (Desarrollando por los elementos de la primera fila).

b) Calcular

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & -1 & 3 & 1 \\ 4 & 1 & 9 & 1 \\ 8 & -1 & 27 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & -2 & 4 & -8 & 16 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & 2 & 4 & 8 & 16 \\ 1 & -3 & 9 & -27 & 81 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \qquad \begin{pmatrix} 2 & -4 & 8 & -16 \\ 3 & 9 & 27 & 81 \\ 2 & -6 & 18 & -54 \\ 3 & -3 & 3 & -3 \end{pmatrix}$$

8. Dadas dos matrices n×n , A y B, qué relación hay entre sus determinantes y det(A.B)?. Verificar la fórmula en el caso

$$A = \begin{pmatrix} -1 & 3 & 2 \\ -5 & 1 & -3 \\ 2 & 4 & 1 \end{pmatrix} \quad , \qquad B = \begin{pmatrix} 0 & -1 & 3 \\ 4 & -1 & 10 \\ 2 & 1 & 5 \end{pmatrix}$$

9. En cada uno de los casos siguientes hallar adj A , decir si A es inversible y en tal caso calcular la matriz inversa (utilizando determinantes):

a) $A = \begin{pmatrix} -1 & -2 \\ 3 & 5 \end{pmatrix}$

b) $A = \begin{pmatrix} 1 & 0 & 2 \\ -1 & 3 & -1 \\ 0 & 2 & 4 \end{pmatrix}$

c) $A = \begin{pmatrix} 0 & -4 & -3 \\ 1 & -2 & 0 \\ 1 & 2 & 3 \end{pmatrix}$

d) $A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ \cos\alpha & 0 & 0 & -\operatorname{sen}\alpha \\ 0 & 1 & 0 & 0 \\ \operatorname{sen}\alpha & 0 & 0 & \cos\alpha \end{pmatrix}$

10. Verificar si los siguientes sistemas de ecuaciones lineales son determinados y en tal caso hallar la solución aplicando la regla de Cramer:

a) $\begin{cases} 2x_1 - 6x_2 = 3 \\ x_1 - 3x_2 + x_3 = 2 \\ 2x_2 - 3x_3 = -1 \end{cases}$

b) $\begin{cases} x_1 - x_2 + x_4 = 0 \\ 2x_1 - x_3 + 3x_4 = 4 \\ 5x_2 - x_3 - x_4 = 1 \\ 3x_1 + 4x_2 - 2x_3 + 3x_4 = 5 \end{cases}$

c) $\begin{cases} 5x_1 + 2x_2 - x_3 = -1 \\ 2x_1 - x_2 + 2x_3 = 7 \\ x_1 - 3x_2 - x_3 = 2 \end{cases}$

6.4. RANGO DE UNA MATRIZ.

OTRO METODO PARA RESOLVER SISTEMAS DE ECUACIONES LINEALES.

Dada una matriz arbitraria $m \times n$, $A = (a_{ij})$, las filas de A son n-uplas de números de K. (Y las columnas son m-uplas).

Vamos a considerar las filas de A y para evitar confusiones las representaremos con letras griegas: $\alpha, \beta, \gamma, \ldots\ldots$

Sabemos que una fila β se dice una combinación lineal de otras filas $\alpha_1, \alpha_2, \ldots, \alpha_r$ si existen escalares $k_1, k_2, \ldots, k_r$ tales que:

$$\beta = k_1\alpha_1 + k_2\alpha_2 + \ldots + k_r\alpha_r$$

Los escalares $k_1, k_2, \ldots, k_r$ se llaman los coeficientes de la combinación lineal.

Por ejemplo, en la matriz

$$A = \begin{pmatrix} 2 & -1 & 0 & 7 & 1 \\ 4 & 1 & -2 & 5 & 2 \\ 1 & 1 & 3 & -6 & 0 \\ 11 & 2 & 15 & -9 & 3 \end{pmatrix}$$

la cuarta fila es combinación lineal de la primera y la tercera pues:

$$\alpha_4 = 3\alpha_1 + 5\alpha_3$$

En particular, observemos que una fila nula β (es decir, formada por ceros) se puede escribir como combinación lineal de cualquier sistema de filas $\alpha_1, \alpha_2, \ldots, \alpha_r$ ya que:

$$\beta = 0\alpha_1 + 0\alpha_2 + \ldots + 0\alpha_r$$

A una fila nula la notaremos 0.

Definición. Se dice que r filas

$$\alpha_1 , \alpha_2 , \ldots\ldots , \alpha_r \qquad , \qquad r \geq 2$$

de una matriz forman un sistema linealmente independiente si ninguna de ellas es combinación lineal de las demás. (1)
En caso contrario el sistema se dice linealmente dependiente.

Es decir, un sistema de filas $\alpha_1, \alpha_2, \ldots, \alpha_r$, con $r \geq 2$, es linealmente dependiente si y sólo si por lo menos una de ellas es combinación lineal de las demás.

Por ejemplo, cualquier sistema de filas en que figure una fila nula, o dos filas iguales, o dos filas proporcionales, es linealmente dependiente.

La definición anterior puede darse de otra manera equivalente, teniendo en cuenta que se verifica la siguiente propiedad:

Dadas r filas $\alpha_1, \alpha_2, \ldots, \alpha_r$ $(r \geq 2)$ ninguna de ellas es combinación lineal de las demás si y sólo si hay una única manera de obtener una fila nula como combinación lineal de las filas $\alpha_1, \alpha_2, \ldots, \alpha_r$, es decir, si y sólo si:

$$k_1\alpha_1 + k_2\alpha_2 + \ldots + k_r\alpha_r = 0 \quad \text{implica} \quad k_1 = k_2 = \ldots = k_r = 0 \qquad (2)$$

Vamos a demostrarlo.

Sean $\alpha_1, \alpha_2, \ldots, \alpha_r$ tales que ninguna es combinación lineal de las demás y suponga-mos que existe una combinación lineal de ellas que da una fila nula:

$$k_1\alpha_1 + k_2\alpha_2 + \ldots + k_r\alpha_r = 0$$

con no todos los coeficientes nulos. Si $k_i \neq 0$ se tiene:

$$\alpha_i = -\frac{k_1}{k_i}\alpha_1 - \frac{k_2}{k_i}\alpha_2 - \ldots - \frac{k_{i-1}}{k_i}\alpha_{i-1} - \frac{k_{i+1}}{k_i}\alpha_{i+1} - \ldots - \frac{k_r}{k_i}\alpha_r$$

lo que significa que la fila α_i es una combinación lineal de $\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_r$ y contradice nuestra hipótesis. Esta contradicción proviene de haber supuesto que no todos los coeficientes son nulos. Luego $k_1 = k_2 = \ldots = k_r = 0$ y queda probado que la propiedad (1) implica la propiedad (2).

Recíprocamente supongamos que $\alpha_1, \alpha_2, \ldots, \alpha_r$ verifican la condición (2), y que α_i es combinación lineal de las demás, para algún índice i, $1 \leq i \leq r$:

$$\alpha_i = k_1\alpha_1 + \ldots + k_{i-1}\alpha_{i-1} + k_{i+1}\alpha_{i+1} + \ldots + k_r\alpha_r$$

Luego se tiene:

$$k_1\alpha_1 + \ldots + k_{i-1}\alpha_{i-1} + (-1)\alpha_i + k_{i+1}\alpha_{i+1} + \ldots + k_r\alpha_r = 0$$

y no todos los coeficientes de esta combinación lineal son nulos, lo que contradice la hipótesis de que $\alpha_1, \alpha_2, \ldots, \alpha_r$ verifican la propiedad (2). Esta contradicción prueba que ninguna fila α_i es combinación lineal de las demás.

Queda demostrado así que las propiedades (1) y (2) son equivalentes. Por lo tanto, la siguiente definición es equivalente a la anterior:

Se dice que r filas $\alpha_1, \alpha_2, \ldots, \alpha_r$ de una matriz forman un sistema linealmente inde-

pendiente si

$$k_1\alpha_1 + k_2\alpha_2 + \ldots + k_r\alpha_r = 0 \quad \text{implica} \quad k_1 = k_2 = \ldots = k_r = 0$$

Como esta segunda definición se puede aplicar al caso r = 1, se conviene en que un sistema formado por una sola fila α es linealmente independiente si y sólo si α no es una fila nula. (Pues $k\alpha = 0$ implica $k = 0$ si y sólo si $\alpha \neq 0$).

Se verifican las muy importantes propiedades siguientes, cuya demostración queda a cargo del lector:

PI. Reordenando las filas de un sistema linealmente independiente $\{\alpha_1,\alpha_2,\ldots,\alpha_r\}$ el sistema sigue siendo linealmente independiente.

PII. Si en un sistema de r filas linealmente independiente se reemplaza una de e - llas por la que se obtiene multiplicándola por un escalar no nulo, se obtiene un sistema linealmente independiente, es decir:

$$\{\alpha_1,\ldots,\alpha_r\} \text{ L.I.} \Longrightarrow \{\alpha_1,\ldots,k\alpha_i,\ldots,\alpha_r\} \text{ L.I.}, \ \forall\, k \in K,\ k \neq 0\ ,\ 1 \leq i \leq r\ .$$

PIII. Si en un sistema de filas linealmente independiente se reemplaza una de ellas por la que se obtiene sumándole a esa fila otra distinta multiplicada por un escalar, se obtiene un sistema linealmente independiente, es decir:

$$\{\alpha_1,\ldots,\alpha_r\} \text{ L.I.} \Longrightarrow \{\alpha_1,\ldots,\alpha_{i-1},\alpha_i+k\alpha_j,\alpha_{i+1},\ldots,\alpha_r\} \text{ L.I.}\ ,\ \forall\, k \in K,\ i \neq j\ ,$$

$$1 \leq i \leq r\ ,\quad 1 \leq j \leq r\ .$$

RANGO DE UNA MATRIZ.

Dada una matriz m×n A interesa saber cuál es el mayor número de filas linealmente independientes de A. (Decimos que r filas son linealmente independientes si ellas forman un sistema linealmente independiente).

Definición. Se llama rango (por filas) de una matriz al mayor número de filas linealmente independientes.

Por ejemplo, dadas las matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\ ,\quad \begin{pmatrix} 1 & 2 & -1 & 3 \\ -1 & -2 & 1 & -3 \\ 3 & 6 & -3 & 9 \end{pmatrix}\ ,\quad \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}\ ,\quad \begin{pmatrix} 3 & 1 & 0 \\ 5 & -2 & -4 \\ 1 & -2 & 1 \\ 5 & -9 & -1 \end{pmatrix}$$

es inmediato que el rango de la primera y la segunda es uno, y que el de la tercera es cero. En cambio no es fácil saber a simple vista cuál es el rango de la cuarta matriz.

A continuación indicaremos dos métodos distintos para calcular el rango.

Notemos que si C es una matriz reducida canónica m×n entonces el rango de C es igual al número de filas no nulas de C. En efecto, si r es el número de filas no nulas y $j_1, j_2, \ldots, j_r$ son las columnas principales, C es de la forma:

$$C = \begin{array}{c} \begin{array}{ccccccccccccc} & & \underset{\sim}{j_1} & & & & \underset{\sim}{j_2} & & & & \underset{\sim}{j_r} & & \end{array} \\ \left(\begin{array}{ccccccccccccc} 0 & \ldots & 1 & a_{1,j_1+1} & \ldots & 0 & a_{1,j_2+1} & \ldots & 0 & a_{1,j_r+1} & \ldots & a_{1n} \\ 0 & \ldots & 0 & 0 & \ldots & 1 & a_{2,j_2+1} & \ldots & 0 & a_{2,j_r+1} & \ldots & a_{2n} \\ & & & & & & & & & & & \\ 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & 1 & a_{r,j_r+1} & \ldots & a_{rn} \\ 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ & & & & & & & & & & & \\ 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \end{array} \right) \end{array}$$

Entonces las r primeras filas $\alpha_1, \alpha_2, \ldots, \alpha_r$ son linealmente independientes pues si

$$k_1\alpha_1 + k_2\alpha_2 + \ldots + k_r\alpha_r = 0 \quad , \quad \text{se tiene:}$$

$$(\ldots, \underset{\sim}{j_1}\ k_1, \ldots, \underset{\sim}{j_2}\ k_2, \ldots, \underset{\sim}{j_r}\ k_r, \ldots) = (0, 0, \ldots, 0)$$

y de aquí resulta

$$k_1 = k_2 = \ldots = k_r = 0 \ .$$

Además está claro que r es el máximo número de filas de C linealmente independientes.

Por otra parte , dada una matriz m×n arbitraria A, cuando se pasa de A a otra matriz B mediante una operación elemental, la dependencia lineal de las filas no varía y por lo tanto se conserva el rango. En efecto, si las filas $\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_k}$ de A son linealmente independientes, por una de las propiedades PI , PII ó PIII , se ve que las filas correspondientes $\beta_{j_1}, \beta_{j_2}, \ldots, \beta_{j_k}$ de B también lo son. En particular, si el rango de A es r y $\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_r}$ son r filas linealmente independientes de A

entonces las filas correspondientes de B también son linealmente independientes. Luego rango A $\leq$ rango B. Como a su vez se puede pasar de B a A invirtiendo la operación elemental, por el mismo razonamiento resulta rango B $\leq$ rango A. Luego ambos rangos son iguales.

Se tiene entonces el siguiente

TEOREMA 6.11. Si A y B son dos matrices equivalentes por filas, entonces rango A = rango B. Si $\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_k}$ son filas linealmente independientes de A, las filas correspondientes $\beta_{j_1}, \beta_{j_2}, \ldots, \beta_{j_k}$ de B también lo son. (Teniendo en cuenta para establecer la correspondencia los reordenamientos de filas que se efectúan al pasar de A a B por operaciones elementales).

Demostración: A cargo del lector.

En particular, dada una matriz A, si C es una matriz reducida canónica equivalente por filas a A, es rango A = rango C.

Esto proporciona un método para calcular el rango de una matriz A: Se busca una matriz reducida canónica C equivalente a A y se cuentan las filas no nulas de C. Este número r da el rango de A. Además las r filas de A correspondientes a las filas no nulas de C (teniendo en cuenta los reordenamientos de filas efectuados al pasar de A a C) son linealmente independientes, y todas las demás son combinaciones lineales de ellas.

EJEMPLO.

Hallar el rango de la siguiente matriz e indicar un sistema de filas linealmente independiente maximal.

$$A = \begin{pmatrix} 3 & 1 & 0 \\ 5 & -2 & -4 \\ 1 & -2 & 1 \\ 5 & -9 & -1 \end{pmatrix}$$

$$A = \begin{pmatrix} 3 & 1 & 0 \\ 5 & -2 & -4 \\ 1 & -2 & 1 \\ 5 & -9 & -1 \end{pmatrix} \xrightarrow{1} \begin{matrix} 3^{\underline{a}} \\ 1^{\underline{a}} \\ 2^{\underline{a}} \\ 4^{\underline{a}} \end{matrix} \begin{pmatrix} 1 & -2 & 1 \\ 3 & 1 & 0 \\ 5 & -2 & -4 \\ 5 & -9 & -1 \end{pmatrix} \xrightarrow{2} \begin{pmatrix} 1 & -2 & 1 \\ 0 & 7 & -3 \\ 0 & 8 & -9 \\ 0 & 1 & -6 \end{pmatrix} \xrightarrow{3}$$

$$\xrightarrow{3} \begin{matrix} 3^{\underline{a}} \\ 4^{\underline{a}} \\ 1^{\underline{a}} \\ 2^{\underline{a}} \end{matrix} \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & -6 \\ 0 & 7 & -3 \\ 0 & 8 & -9 \end{pmatrix} \xrightarrow{4} \begin{pmatrix} 1 & 0 & -11 \\ 0 & 1 & -6 \\ 0 & 0 & 39 \\ 0 & 0 & 39 \end{pmatrix} \xrightarrow{5} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = C$$

Luego rango A = rango C = 3. Las filas primera, tercera y cuarta de A (que son las que corresponden a las filas no nulas de C) forman un sistema linealmente independiente, y la segunda fila es combinación de ellas.

OBSERVACION. Notemos que éste no es el único sistema de tres filas linealmente independientes en la matriz A. Por ejemplo, si en el tercer paso, en lugar de intercambiar filas, restamos a la segunda fila la tercera, obtenemos la fila: 0,-1 , 6 . Multiplicándola por (-1) y continuando el proceso se puede llegar a una matriz reducida canónica C sin necesidad de ulteriores reordenamientos de filas. El único intercambio de filas efectuado es entonces el del primer paso, y las filas no nulas de C corresponden en este caso a la primera, segunda y tercera filas de A.Luego éstas forman un sistema linealmente independiente y la cuarta fila es combinación lineal de ellas.

Todo lo dicho sobre dependencia lineal de las filas de una matriz A puede decirse también para las columnas.

Así, se dice que r columnas de A forman un sistema linealmente independiente si ninguna de ellas es combinación lineal de las demás.

Se llama rango por columnas de A al máximo número de columnas linealmente independientes.

Queremos probar que el rango por filas de una matriz es igual al rango por columnas, propiedad bastante sorprendente que autoriza a hablar simplemente de rango de una matriz.

Para ello veremos otra forma de calcular el rango de una matriz, usando determinantes.

Consideremos todos los menores que se pueden formar con las filas y columnas de una matriz A, es decir, los determinantes de todas las submatrices cuadradas de A, y entre ellos los que son distintos de cero. Diremos que un menor D es principal si $D \neq 0$ y D es de orden máximo entre todos los menores no nulos de A.

En lo que sigue, rango significa rango por filas.

TEOREMA 6.12. El orden de un menor principal de una matriz A es igual al rango de A. Las filas de A cuyos elementos figuran en un menor principal cualquiera son lineal-

mente independientes y todas las demás filas son combinación lineal de ellas.

Demostración: Sea r el mayor de los órdenes de los menores no nulos de A y sea D un menor principal. Podemos suponer que D está situado en el ángulo superior izquierdo de A, lo que no quita generalidad a la demostración.

$$A = \left(\begin{array}{ccc|ccc} \multicolumn{3}{c|}{\multirow{3}{*}{$\begin{matrix} a_{11} & \cdots & a_{1r} \\ \cdots & D & \cdots \\ a_{r1} & \cdots & a_{rr} \end{matrix}$}} & a_{1,r+1} & \cdots & a_{1n} \\ & & & \cdots & \cdots & \cdots \\ & & & a_{r,r+1} & \cdots & a_{rn} \\ \hline a_{r+1,1} & \cdots & a_{r+1,r} & a_{r+1,r+1} & \cdots & a_{r+1,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mr} & a_{m,r+1} & \cdots & a_{mn} \end{array} \right)$$

Como $D \neq 0$ las r primeras filas de A son linealmente independientes pues si alguna fuera combinación lineal de las demás, D sería cero. Luego rango $A \geq r$.

Para probar que el rango de la matriz es r tenemos que demostrar que no hay r+1 filas linealmente independientes.

Supongamos por el absurdo que hay r+1 filas $\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_{r+1}}$ que son linealmente independientes. Consideremos la submatriz B de A formada por esas filas y pasemos de B a una matriz reducida canónica C equivalente por filas a B. De la hipótesis sobre las filas de B resulta rango C = r+1 y todas las filas de C son no nulas. Luego C tiene r+1 columnas principales. Considerando el menor Δ formado por esas columnas es

$$\Delta = \begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{vmatrix} \neq 0$$

Sea D' el menor formado por las columnas correspondientes de B. Como las filas de Δ se obtienen de las de D' mediante operaciones elementales, Δ difiere de D' a lo sumo en una constante no nula. Luego $D' \neq 0$ y se tiene así un menor de A de orden r+1 no nulo, lo que contradice la hipótesis de que r es el mayor orden de los menores no nulos de A. Esta contradicción prueba que r es el máximo número de filas linealmente independientes de A.

De aquí resulta en particular que todas las filas de A son combinación lineal de las r primeras, lo que termina la demostración.

COROLARIO 1. El rango por filas de una matriz coincide con el rango por columnas.

Demostración: En el teorema anterior se puede reemplazar la palabra fila por columna. En efecto, las columnas de A son las filas de la matriz traspuesta A^T. Aplicando el teorema a A^T resulta que el número máximo de filas linealmente independientes es igual al orden de un menor principal de A^T. Pero el determinante de una matriz cualquiera coincide con el de la traspuesta, de modo que los menores de A^T coinciden con los menores de A. Luego el número máximo de columnas linealmente independientes de A es igual al orden de un menor principal de A. Y de aquí sigue el corolario.

COROLARIO 2. Si A es una matriz cuadrada de orden n, det(A) = 0 si y sólo si alguna fila (columna) es combinación lineal de las demás.

Demostración: A cargo del lector.

El teorema anterior proporciona otro método para calcular el rango de una matriz, sin recurrir a la matriz reducida canónica: consiste en encontrar el orden máximo de los menores no nulos de la matriz dada. Pero el número de todos los menores que se pueden formar en una matriz es en general muy grande y el cálculo se hace interminable. Existe un procedimiento metódico que ahorra cálculos, que explicaremos a continuación.

Digamos que, dada una submatriz $s \times s$ B de una matriz A, orlar B significa formar submatrices de tipo $(s+1)\times(s+1)$ agregándole a B una fila y una columna más en forma ordenada.

Regla práctica para calcular el rango de una matriz: Se busca una submatriz de segundo orden cuyo determinante sea distinto de cero. Se orla esta submatriz con las demás filas y columnas formando submatrices de tercer orden hasta encontrar una cuyo determinante sea distinto de cero. Una vez hallada una de este tipo, se la orla con las demás filas y columnas hasta encontrar una de cuarto orden de determinante no nulo, etc., ... Habiendo obtenido una submatriz de orden k de determinante distinto de cero, si al orlarla con las restantes filas y columnas todas las submatrices que se obtienen tienen determinante nulo, entonces k es el rango de la matriz dada.

Para demostrar que k es efectivamente el rango buscado, supongamos que el menor D de orden k distinto de cero encontrado está formado por elementos de las primeras k filas y las primeras k columnas. Orlando D con la i-ésima fila, todos los menores de orden k+1 que se obtienen son nulos, de acuerdo con nuestra hipótesis. Luego, por el teorema recién demostrado, el rango de la submatriz B de A formada por las k primeras filas y la i-ésima fila es k, y la i-ésima fila es combinación lineal de las k primeras. Haciendo este razonamiento para $i = k+1,\ldots,m$ se tiene que las k primeras filas son linealmente independientes y las m-k últimas son combinaciones linea

les de aquéllas. Entonces se puede pasar por operaciones elementales de A a una matriz C con las m-k últimas filas nulas y las k primeras linealmente independientes, lo que implica claramente que rango C = k . Luego rango A = k c.q.d.

EJEMPLOS.

1) Hallar el rango de la matriz

$$\begin{pmatrix} 2 & -6 & 2 & 0 & -1 \\ 1 & -3 & 1 & 4 & 1 \\ 0 & 0 & -2 & 1 & 5 \\ 1 & -3 & -3 & 6 & 11 \end{pmatrix}$$

Aplicando la regla, buscamos un menor de segundo orden no nulo. El que está situado en el ángulo superior de la izquierda es igual a cero, pero hay otros no nulos, por ejemplo

$$D = \begin{vmatrix} 2 & 0 \\ 1 & 4 \end{vmatrix} \neq 0$$

Orlamos este determinante con las demás filas y columnas hasta encontrar uno de tercer orden distinto de cero.

$$\begin{vmatrix} 2 & 0 & -6 \\ 1 & 4 & -3 \\ 0 & 1 & 0 \end{vmatrix} = 0 \quad , \quad D' = \begin{vmatrix} 2 & 0 & 2 \\ 1 & 4 & 1 \\ 0 & 1 & -2 \end{vmatrix} \neq 0$$

Orlamos ahora este último determinante con la cuarta fila:

$$\begin{vmatrix} 2 & 0 & 2 & -6 \\ 1 & 4 & 1 & -3 \\ 0 & 1 & -2 & 0 \\ 1 & 6 & -3 & -3 \end{vmatrix} = 0 \quad , \quad \begin{vmatrix} 2 & 0 & 2 & -1 \\ 1 & 4 & 1 & 1 \\ 0 & 1 & -2 & 5 \\ 1 & 6 & -3 & 11 \end{vmatrix} = 0$$

Como todos los menores de cuarto orden que así se obtienen son nulos, el rango de la matriz es 3. D' es un menor principal. Las filas y columnas de la matriz cuyos coeficientes figuran en D' son linealmente independientes y las demás son combinaciones lineales de ellas.

2) Hallar el rango de la siguiente matriz e indicar un sistema de filas y uno de columnas linealmente independiente maximal.

$$\begin{pmatrix} 4 & -1 & 0 & 1 & 4 \\ -8 & 2 & 0 & -2 & -8 \\ 3 & 5 & -1 & 4 & 2 \\ -1 & 6 & -1 & 3 & -2 \\ -5 & 7 & -1 & 2 & -6 \end{pmatrix}$$

Antes de empezar los cálculos conviene observar atentamente la matriz para ver si a simple vista hay alguna fila o columna que sea combinación lineal de las demás. Si es así, se las puede descartar.

Por ejemplo, en la matriz dada vemos que la segunda fila es igual a la primera multiplicada por -2, y que la última columna es igual a la suma de la primera y la tercera columnas. Luego podemos descartar la segunda fila y la última columna. Como no vemos ninguna otra combinación lineal, procedemos a buscar un menor principal.

Un menor de segundo orden no nulo es, por ejemplo:

$$\begin{vmatrix} 4 & -1 \\ 3 & 5 \end{vmatrix} \neq 0$$

Orlando este menor con las restantes filas y columnas se tiene:

$$\begin{vmatrix} 4 & -1 & 0 \\ 3 & 5 & -1 \\ -1 & 6 & -1 \end{vmatrix} = 0 \quad , \quad \begin{vmatrix} 4 & -1 & 1 \\ 3 & 5 & 4 \\ -1 & 6 & 3 \end{vmatrix} = 0$$

$$\begin{vmatrix} 4 & -1 & 0 \\ 3 & 5 & -1 \\ -5 & 7 & -1 \end{vmatrix} = 0 \quad , \quad \begin{vmatrix} 4 & -1 & 1 \\ 3 & 5 & 4 \\ -5 & 7 & 2 \end{vmatrix} = 0$$

Luego el rango de la matriz es 2. La primera y la tercera filas forman un sistema linealmente independiente y todas las demás son combinaciones lineales de ellas. Análogamente, las dos primeras columnas son linealmente independientes y todas las demás son combinaciones lineales de ellas.

Indique el lector otros sistemas linealmente independientes maximales de filas y columnas en la matriz dada.

COMPATIBILIDAD DE UN SISTEMA DE ECUACIONES LINEALES.

Vamos a ver ahora otra forma de decidir si un sistema de ecuaciones lineales es compatible o no, sin utilizar operaciones elementales.

Dado un sistema

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ \ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = b_m \end{cases}$$

Consideremos la matriz de los coeficientes A y la matriz del sistema A', que también se llama la matriz ampliada (porque se obtiene de A agregándole la columna de los términos independientes):

$$A = \begin{pmatrix} a_{11} \ldots a_{1n} \\ \ldots\ldots\ldots \\ a_{m1} \ldots a_{mn} \end{pmatrix} \qquad A' = \begin{pmatrix} a_{11} \ldots a_{1n} & b_1 \\ \ldots\ldots\ldots & \ldots \\ a_{m1} \ldots a_{mn} & b_m \end{pmatrix}$$

El siguiente se conoce con el nombre de teorema de Rouché-Frobenius, o también, de Kronecker-Cappelli:

TEOREMA 6.13. Un sistema de ecuaciones lineales es compatible si y sólo si el rango de la matriz de las coeficientes es igual al rango de la matriz ampliada.

Demostración: Si el sistema es compatible, sea $(k_1,k_2,\ldots,k_n)$ una solución del mismo. Reemplazando los k_i en lugar de las incógnitas se ve que la columna de los términos independientes es una combinación lineal de las columnas de los coeficientes. Luego la última columna de A' es combinación lineal de las demás, que coinciden con las de A. Entonces rango A = rango A' , pues está claro que el rango de una matriz no varía cuando se agrega una columna que es combinación lineal de las otras.

Recíprocamente, supongamos que rango A = rango A' = r. Sean $\alpha_1,\alpha_2,\ldots,\alpha_n$ las columnas de A y β la última columna de A'. Si $\alpha_{i_1},\alpha_{i_2},\ldots,\alpha_{i_r}$ es un sistema de columnas de A linealmente independiente, pensadas como columnas de A' son r columnas linealmente independientes, lo que implica que β es una combinación lineal de ellas. Entonces β se puede escribir como una combinación lineal de todas las demás columnas de A', es decir, existen escalares $k_1,k_2,\ldots,k_n$ tales que:

$$k_1\alpha_1 + k_2\alpha_2 + \ldots + k_n\alpha_n = \beta$$

De aquí resulta que $(k_1,k_2,\ldots,k_n)$ es una solución del sistema dado.

El teorema queda así demostrado.

Otro método para resolver un sistema de m ecuaciones lineales con n incógnitas, utilizando la regla de Cramer.

Sea S un sistema de m ecuaciones lineales con n incógnitas.

Para resolver el sistema S hay que averiguar en primer lugar si es compatible o no. Aplicando el teorema recién demostrado, se calculan los rangos de la matriz de los coeficientes A y de la matriz ampliada A'. Si son distintos el sistema es incompatible. Si rango A = rango A' = r el sistema es compatible. Veamos cómo calcular las soluciones. Al hallar el rango de A se ha encontrado un menor principal $D \neq 0$ de orden r, que es también menor principal de A'. Podemos suponer que en él figuran las r primeras filas y las r primeras columnas. Entonces el sistema dado S es equivalente al que se obtiene eliminando las últimas m-r ecuaciones, que son combinaciones lineales de las r primeras:

$$S' \begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rn}x_n = b_r \end{cases}$$

El sistema S' se puede escribir de modo que en el primer miembro de cada ecuación figuren solamente las r incógnitas cuyos coeficientes forman las columnas de D, pasando los n-r términos restantes al segundo miembro. Podemos suponer que se trata de $x_1, x_2, \ldots, x_r$, renumerando las incógnitas si fuera necesario:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1r}x_r = b_1 - a_{1,r+1}x_{r+1} - \ldots - a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2r}x_r = b_2 - a_{2,r+1}x_{r+1} - \ldots - a_{2n}x_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rr}x_r = b_r - a_{r,r+1}x_{r+1} - \ldots - a_{rn}x_n \end{cases}$$

Si se atribuye a las incógnitas $x_{r+1}, \ldots, x_n$ valores numéricos $k_{r+1}, \ldots, k_n$ se tiene un sistema:

$$S'' \begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1r}x_r = b_1 - a_{1,r+1}k_{r+1} - \ldots - a_{1n}k_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2r}x_r = b_2 - a_{2,r+1}k_{r+1} - \ldots - a_{2n}k_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rr}x_r = b_r - a_{r,r+1}k_{r+1} - \ldots - a_{rn}k_n \end{cases}$$

donde el determinante de la matriz de los coeficientes es el menor principal $D \neq 0$. Luego S'' tiene una única solución que puede calcularse por la regla de Cramer. (Teorema 6.10). Si $(k_1, k_2, \ldots, k_r)$ es una solución de S'', es claro que $(k_1, k_2, \ldots, k_r, k_{r+1}, \ldots, k_n)$ es una solución de S'.

Recíprocamente, si $(k_1, \ldots, k_n)$ es una solución cualquiera de S', entonces haciendo $x_{r+1} = k_{r+1}, \ldots, x_n = k_n$, la única solución del sistema S'' correspondiente será $x_1 = k_1, \ldots, x_r = k_r$.

De todo lo dicho resulta la siguiente regla:

Para resolver un sistema de m ecuaciones lineales con n incógnitas se determinan los rangos r y r' de la matriz de los coeficientes y de la matriz ampliada. Si $r \neq r'$ el sistema es incompatible. Si $r = r'$ es compatible. En este último caso las soluciones se calculan como sigue: Sea r el rango común de las dos matrices y D un menor principal de la matriz de los coeficientes. En el sistema dado se eligen las ecuaciones cuyos coeficientes aparecen en las filas de D. Se obtiene así un sistema de r ecuaciones. Se eligen las r incógnitas cuyos coeficientes figuran en las columnas de D y se pasan las n-r restantes al segundo miembro. Resolviendo el sistema así obtenido por la regla de Cramer se calcula el valor de las incógnitas $x_1, x_2, \ldots, x_r$ en función de $x_{r+1}, x_{r+2}, \ldots, x_n$. Dándole valores numéricos arbitrarios a estas últimas, se calcula el valor correspondiente de $x_1, x_2, \ldots, x_r$. Se obtienen así todas las soluciones del sistema.

El número n-r se llama el grado de indeterminación del sistema.

Podemos hacer el siguiente esquema:

Si $r \neq r'$ el sistema es incompatible.

Si $r = r'$ el sistema es compatible $\begin{cases} \text{determinado si } r = n. \\ \text{indeterminado si } r < n. \end{cases}$

(Es claro que r es siempre menor o igual que n).

En particular, un sistema lineal homogéneo de m ecuaciones con n incógnitas tiene una solución diferente de la trivial si y sólo si $r < n$.

EJEMPLOS.

1. Resolver el sistema

$$\begin{cases} 2x_1 - x_2 + x_3 - x_4 = 0 \\ 2x_1 - x_2 + x_3 + 2x_4 = 1 \\ 2x_1 - x_2 + x_3 - 4x_4 = -2 \end{cases}$$

Formamos la matriz de los coeficientes y la matriz ampliada:

$$A = \begin{pmatrix} 2 & -1 & 1 & -1 \\ 2 & -1 & 1 & 2 \\ 2 & -1 & 1 & -4 \end{pmatrix} \quad , \quad A' = \begin{pmatrix} 2 & -1 & 1 & -1 & 0 \\ 2 & -1 & 1 & 2 & 1 \\ 2 & -1 & 1 & -4 & -2 \end{pmatrix}$$

Para calcular el rango de A, se ve que el menor de segundo orden del extremo superior derecho es distinto de cero; al orlar este determinante con la tercera fila, los dos determinantes de tercer orden que se obtienen son nulos. Luego rango A = 2.

Para calcular el rango de A', podemos elegir el mismo menor de segundo orden no nulo, pero al orlarlo con la tercera fila de A' se obtiene el menor:

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

Por lo tanto rango A' = 3. Como los rangos de A y A' son diferentes el sistema es incompatible.

2. Resolver el siguiente sistema:

$$\begin{cases} 2x_1 + x_2 - x_3 = 4 \\ 4x_1 + 2x_2 + x_3 = 7 \\ 2x_1 + x_2 - 4x_3 = 5 \end{cases}$$

Calculemos el rango de la matriz de los coeficientes y el rango de la matriz ampliada.

$$\left(\begin{array}{ccc|c} 2 & 1 & -1 & 4 \\ 4 & 2 & 1 & 7 \\ 2 & 1 & -4 & 5 \end{array}\right)$$

Encontramos

$$D = \begin{vmatrix} 2 & -1 \\ 4 & 1 \end{vmatrix} = 6 \neq 0$$

Al orlar este determinante con la tercera fila en la matriz de los coeficientes A, y luego en la matriz ampliada A', los menores de tercer orden que se obtienen son todos nulos. Entonces D es un menor principal y rango A = rango A' = 2.

En consecuencia el sistema es compatible e indeterminado, pues el rango es menor que el número de incógnitas.

Como D es un menor principal, deducimos que las dos primeras ecuaciones son linealmente independientes y la tercera es una combinación lineal de las dos primeras, por lo que puede descartarse. Elegimos las incógnitas x_1 y x_3 cuyos coeficientes figuran en las columnas de D y pasamos x_2 al segundo miembro. Se obtiene así el siguiente sistema:

$$\begin{cases} 2x_1 - x_3 = 4 - x_2 \\ 4x_1 + x_3 = 7 - 2x_2 \end{cases}$$

El determinante de la matriz de los coeficientes de este sistema es el menor D.

Suponiendo que a x_2 se le ha atribuído un valor numérico, se resuelve el sistema por la regla de Cramer y se calculan las incógnitas x_1 y x_3 en función de x_2 :

$$x_1 = \frac{\begin{vmatrix} 4-x_2 & -1 \\ 7-2x_2 & 1 \end{vmatrix}}{\begin{vmatrix} 2 & -1 \\ 4 & 1 \end{vmatrix}} = \frac{11-3x_2}{6} = \frac{11}{6} - \frac{1}{2}x_2$$

$$x_3 = \frac{\begin{vmatrix} 2 & 4-x_2 \\ 4 & 7-2x_2 \end{vmatrix}}{\begin{vmatrix} 2 & -1 \\ 4 & 1 \end{vmatrix}} = \frac{-2}{6} = -\frac{1}{3}$$

Para cada valor numérico arbitrario de x_2 se obtiene una solución del sistema . Luego las soluciones del sistema son todas las ternas de la forma

$$\left(\frac{11}{6} - \frac{1}{2}k \, , \, k \, , \, -\frac{1}{3} \right) \quad , \quad \text{con } k \in K \, .$$

Por ejemplo, tomando $k = 0$ se tiene la solución particular $\left(\frac{11}{6} , 0 , -\frac{1}{3} \right)$; haciendo $k = -2$ se obtiene la solución $\left(\frac{5}{6} , -2 , -\frac{1}{3} \right)$; etc. . .

3. Resolver el siguiente sistema:

$$\begin{cases} 3x_1 + 3x_2 - x_3 = -4 \\ x_1 + 5x_3 = 9 \\ x_1 - 6x_2 + 4x_3 = 5 \end{cases}$$

Al calcular el rango de la matriz de los coeficientes se tiene:

$$\begin{vmatrix} 3 & 3 & -1 \\ 1 & 0 & 5 \\ 1 & -6 & 4 \end{vmatrix} = 99 \neq 0$$

Entonces el rango de la matriz de los coeficientes es 3 y el rango de la matriz ampliada también es 3. Luego el sistema es compatible y como el rango es igual al número de incógnitas, es determinado.

Aplicando la regla de Cramer se calcula el valor de las incógnitas:

$$x_1 = \frac{\begin{vmatrix} -4 & 3 & -1 \\ 9 & 0 & 5 \\ 5 & -6 & 4 \end{vmatrix}}{99} = \frac{-99}{99} = -1$$

$$x_2 = \frac{\begin{vmatrix} 3 & -4 & -1 \\ 1 & 9 & 5 \\ 1 & 5 & 4 \end{vmatrix}}{99} = \frac{33}{99} = \frac{1}{3}$$

$$x_3 = \frac{\begin{vmatrix} 3 & 3 & -4 \\ 1 & 0 & 9 \\ 1 & -6 & 5 \end{vmatrix}}{99} = \frac{198}{99} = 2$$

La única solución del sistema es $\left(-1 , \frac{1}{3} , 2\right)$.

4. Resolver el sistema:

$$\begin{cases} x_1 - 2x_2 + 3x_3 + x_4 + 3x_5 = 0 \\ 3x_1 - 6x_2 + 9x_3 - 2x_4 + 4x_5 = 0 \\ 5x_1 - 10x_2 + x_3 - x_4 + x_5 = 0 \end{cases}$$

Como el sistema es homogéneo, es compatible y es claramente indeterminado pues el número de ecuaciones es menor que el número de incógnitas. Calculemos el rango de la matriz de los coeficientes para saber el grado de indeterminación del sistema y hallar las soluciones.

$$\begin{pmatrix} 1 & -2 & 3 & 1 & 3 \\ 3 & -6 & 9 & -2 & 4 \\ 5 & -10 & 1 & -1 & 1 \end{pmatrix}$$

Un menor no nulo de segundo orden es:

$$\begin{vmatrix} 1 & 1 \\ 3 & -2 \end{vmatrix} = -5 \neq 0$$

Al orlar este determinante con la tercera fila encontramos un menor de tercer orden no nulo:

$$D = \begin{vmatrix} 1 & 1 & 3 \\ 3 & -2 & 4 \\ 5 & -1 & 1 \end{vmatrix} = 40 \neq 0$$

Luego el rango de la matriz de los coeficientes (y el de la matriz ampliada) es 3.

Las tres ecuaciones dadas son linealmente independientes. Elegimos las incógnitas x_1 , x_4 y x_5 cuyos coeficientes forman las columnas de D y pasamos x_2 y x_3 al segundo miembro, obteniendo el sistema:

$$\begin{cases} x_1 + x_4 + 3x_5 = 2x_2 - 3x_3 \\ 3x_1 - 2x_4 + 4x_5 = 6x_2 - 9x_3 \\ 5x_1 - x_4 + x_5 = 10x_2 - x_3 \end{cases}$$

Resolviendo este sistema por la regla de Cramer obtenemos el valor de las incógnitas x_1, x_4 y x_5 en función de x_2 y x_3 .

$$x_1 = \frac{\begin{vmatrix} 2x_2-3x_3 & 1 & 3 \\ 6x_2-9x_3 & -2 & 4 \\ 10x_2-x_3 & -1 & 1 \end{vmatrix}}{40} = \frac{80x_2+20x_3}{40} = 2x_2+\frac{1}{2}x_3$$

$$x_4 = \frac{\begin{vmatrix} 1 & 2x_2-3x_3 & 3 \\ 3 & 6x_2-9x_3 & 4 \\ 5 & 10x_2-x_3 & 1 \end{vmatrix}}{40} = \frac{70x_3}{40} = \frac{7}{4}x_3$$

$$x_5 = \frac{\begin{vmatrix} 1 & 1 & 2x_2-3x_3 \\ 3 & -2 & 6x_2-9x_3 \\ 5 & -1 & 10x_2-x_3 \end{vmatrix}}{40} = \frac{-70x_3}{40} = -\frac{7}{4}x_3$$

Luego las soluciones del sistema son todas las quintuplas de la forma

$$(2k + \frac{1}{2}k' \ , \ k \ , \ k' \ , \ \frac{7}{4}k' \ , \ -\frac{7}{4}k' \ ) \ , \ \text{con } k,k' \in K$$

Por ejemplo, para $k = k' = 0$ se obtiene la solución trivial $(0,0,0,0,0)$; tomando $k = 1$, $k' = -1$ se obtiene la solución particular $(\frac{3}{2} , 1 , -1 , -\frac{7}{4} , \frac{7}{4})$; etc...

EJERCICIOS.

1. Resolver todos los sistemas propuestos al final del parágrafo 6.1 por este método.

2. Rango de un producto de matrices. Es natural preguntarse qué relación hay entre el rango de un producto de matrices y los rangos de los factores.

   a) Demostrar que el rango de un producto de matrices no supera el rango mínimo de los factores.

   b) Dar un ejemplo en que el rango del producto coincida con el rango mínimo de los factores y otro en que sea estrictamente menor.

   c) Demostrar que si una matriz cualquiera A se multiplica a derecha o izquierda por una matriz inversible B entonces

$$\text{rango } BA = \text{rango } AB = \text{rango } A$$

INDICE ALFABETICO

BIBLIOGRAFIA

(1) - ALBERT, A.A. Algebra Superior, U.T.E.H.A., México, 1961.

(2) - BIRKHOFF, G. y MAC LANE, S. Algebra Moderna, Teide, Barcelona, 1954.

(3) - COHN, P.M. Ecuaciones lineales, U.T.E.H.A., México, 1966.

(4) - COURANT, R. y ROBBINS, H. ¿Qué es la matemática?, Aeda, Bs. As., 1954.

(5) - GENTILE, E. Algebra de conjuntos, U.N.B.A., Bs. As., 1970.

(6) - GENTILE, E. Notas de álgebra, C.E.F.M. y M., Bs. As., 1964.

(7) - GODEMENT, R. Algebra, Tecnos, Madrid, 1967.

(8) - KUROSH, A.G. Curso de Algebra Superior, Editorial Mir, Moscú, 1968.

(9) - REY PASTOR, J., PI CALLEJA, P. y TREJO, C. Análisis Matemático, Vol. I, Kapelusz, Bs. As., 1958.

(10) - USPENSKY, J. Teoría de ecuaciones, Centro de Estudiantes de Ingeniería, Bs. As., 1958.

A continuación se detalla la bibliografía correspondiente a cada capítulo:

Capítulo I : (2), (5), (6) y (7)
Capítulo II : (1), (2), (4), (6) y (9)
Capítulo III : (1), (2), (6), (8) y (9)
Capítulo IV : (1), (2), (6), (8), (9) y (10)
Capítulo V : (1), (6), (8) y (9)
Capítulo VI : (1), (3) y (8)

ERRATAS ADVERTIDAS

Página 147: En la primera línea falta: Las raíces de orden de multiplicidad mayor que uno se dicen múltiples.

Página 152: El ejercicio que aparece en la página 150, línea 2, debe figurar en la página 152, línea 10, enunciado como sigue:

EJERCICIO. Dado $P(X) \in K[X]$, demostrar que:

a) $(P(X),P'(X)) = 1 \iff$ todas las raíces de $P(X)$ son simples.

b) $P(X)$ irreducible $\Longrightarrow$ todas las raíces de $P(X)$ son simples.

La propiedad a) proporciona un criterio útil para averiguar si un polinomio dado tiene o no raíces múltiples: basta calcular el m.c.d. del polinomio y su derivado.

Página 156: En la tercera línea empezando de abajo debe decir:

$$P(X) = 2X\ (X - \tfrac{1}{2})\ (X + 1)\ (X - i)$$

Página 157: En la tercera línea empezando de abajo debe decir:

$$P(X) = 2X\ (X - \tfrac{1}{2})\ (X + 1)\ (X - i)\ (X + i)$$

Página 197: En la línea 16 falta: 5 y 2, 5 y 4.

En la línea 18 debe decir seis inversiones en lugar de cuatro.

H. Hirigoyen 633 - 7mo A
Bahia Blanca
Tina Venning Nielsen