LECCIONES

de

ALGEBRA LINEAL

por

María Jesús Iranzo Aznar

y

Francisco Pérez Monasor

Departamento de Algebra.

Facultad de Matemáticas. Universitat de València

PROGRAMA.

El Algebra Lineal estudia la estructura de los espacios vectoriales y las aplicaciones lineales entre ellos. Las lecciones que vamos a desarrollar constituyen una iniciación a dicho estudio. Una continuación natural de estas lecciones es la teoría del endomorfismo, caracterizando la semejanza de matrices, obteniendo las formas canónicas y la dimensión de los subespacios fundamentales. Estos temas se desarrollan en Lecciones de Algebra Multilineal.

Si a un espacio vectorial se le dota de un producto escalar, pasa a ser un espacio métrico y objeto central de la así llamada Algebra Geométrica por E. Artin, cuyo texto sigue siendo atractivo a lo largo de los años.

Concluidos estos estudios preliminares, puede abordarse el estudio del Espacio Afín, que puede verse como un conjunto de puntos sobre el que actúa un espacio vectorial. La estructura y transformaciones de dicho espacio subyacente, tienen implicaciones en las del espacio afín asociado. Parte importante de la Geometría Afín es el estudio del Espacio Afín Euclidiano, cuyo espacio vectorial subyacente posee una métrica euclidiana.

Desde un punto de vista más aplicado, podemos citar la Teoría de Codigos correctores de errores y especialmente de los códigos lineales de longitud dada n, que son los subespacios del espacio vectorial $\mathbf{F}^n$, donde $\mathbf{F}$ es un cuerpo finito. De acuerdo con la frase de R. Hill en su texto introductorio a la teoría de códigos, *la estructura de un cuerpo finito se encuentra entre las más bellas de la estructuras matemáticas.* Por otra parte los códigos cíclicos de longitud n sobre un cuerpo finito $\mathbf{F}$, son simplemente los ideales del anillo cociente $\mathbf{F}[x]/(x^n - 1)$. Estos temas se encuentran en las Lecciones de Elementos de Algebra. Aplicaciones, de forma que, mediante dicho desarrollo, se pueden realizar algunos de los conceptos básicos introducidos en Algebra Lineal.

Muchas más son las aplicaciones del Algebra Lineal, además de las ya citadas en el entorno matemático. Físicos, quíimicos. ingenieros.. la utilizan muy frecuentemente.

Lección 1. Preliminares

Consideraremos la noción de conjunto como primitiva, es decir no intentaremos dar una definición de éste concepto, nos contentaremos con la idea intuitiva que del mismo tenemos. Algo parecido sucede con los conceptos anejos fundamentales, elemento de un conjunto o pertenecer a un conjunto.

Si E es un conjunto, escribiremos $a \in E$ para indicar que a es un elemento del conjunto E. Para describir un conjunto, utilizaremos la notación $E = \{a, b, c\}$ para significar que E tiene exactamente los elementos a, b y c o $E = \{a, b, \ldots\}$ para indicar que E se compone de los elementos a, b y otros. Otra forma de describir un conjunto es $E = \{x | \text{proposición acerca de } x\}$, para indicar que E es el conjunto de los elementos x tales que la proposición acerca de x es correcta.

Aún cuando la noción de conjunto es primitiva, convendremos que conjunto y elemento están sometidos a las siguientes reglas:

a) Un conjunto E está bien definido o determinado cuando se posee un criterio que permita afirmar si un objeto a pertenece o no al conjunto E.

b) Un mismo ente matemático no puede ser a la vez un conjunto y un elemento de ese conjunto, es decir no es válido escribir $a \in a$.

c) La colección de todos los conjuntos no es un conjunto.

d) Igualdad de conjuntos. Se entiende que dos conjuntos son iguales si tienen los mismos elementos.

e) La clase $\emptyset$ que no contiene nigún elemento es un conjunto.

Aplicaciones

(1.1) **Definición**: Dados dos conjuntos A y B, una aplicación $f : A \longrightarrow B$ es un criterio que permite asociar a cada elemento de $a \in A$ un único elemento $b \in B$. Escribiremos $b = f(a)$ y diremos que b es la **imagen** por f de a y que a es una **antiimagen** por f de b. A se dice **conjunto inicial** de f y B conjunto final. Si $A_1 \subseteq A$, se denota $f(A_1) = \{f(a_1) | a_1 \in A_1\}$. Por convenio $f(\emptyset) = \emptyset$, donde $\emptyset$ representa al conjunto vacío. En particular $f(A)$ es llamado el **conjunto imagen** y es denotado frecuentemente por Im f.

Si $B_1 \subseteq B$, denotamos por $f^{-1}(B_1) = \{a \in A | f(a) \in B_1\}$. Por convenio, $f^{-1}(\emptyset) = \emptyset$. Al conjunto $f^{-1}(B_1)$ se le llama **imagen inversa** de B_1.

(1.2) **Definición**: Dos aplicaciones f, g se consideran **iguales** si tienen el mismo conjunto inicial, el mismo conjunto final y además $f(a) = g(a)$ para todo a del conjunto inicial.

(1.3) **Definición**: Una aplicación $f : A \longrightarrow B$ es **suprayectiva** si $f(A) = B$.

(1.4) **Definición**: Una aplicación $f : A \longrightarrow B$ es **inyectiva** si siempre que $a \neq a'$ se sigue que $f(a) \neq f(a')$.

(1.5) **Definición**: Una aplicación $f : A \longrightarrow B$ es **biyectiva** si es a la vez inyectiva y suprayectiva. Una aplicación biyectiva de A en sí se dice una **permutación** de A. Dada una aplicación $f : A \longrightarrow B$ biyectiva, la aplicación $g : B \longrightarrow A$ dada por $g(b) = a$ tal que $f(a) = b$ se dice **aplicación inversa** de f y se escribe f^{-1}

(1.6) **Definición**: Dada una aplicación $f : A \longrightarrow B$ y un subconjunto A_1 de A a la aplicación $f_1 : A_1 \longrightarrow B$, dada por $f_1(a) = f(a), \forall a \in A_1$, se le llama **restricción** de f a A_1 y la denotaremos por $f|_{A_1}$.

(1.7) **Definición**: Dadas dos aplicaciones $f : A \longrightarrow B$ y $g : B \longrightarrow C$, queda definida una aplicación $h : A \longrightarrow C$ mediante $h(a) = g(f(a)), \forall a \in A$. Esta aplicación h recibe el nombre de **compuesta o composición** de f y g y se denota por $g \circ f$ o bien fg.

Ejercicio: La composición de dos permutaciones del conjunto A es una permutación de A.

Llamaremos par ordenado (a, b) a una colección de dos objetos a y b donde a está señalado como primero y b como segundo. Es decir $(b, a) \neq (a, b)$ en general. Dados dos conjuntos A, B, el conjunto de pares ordenados (a, b) tales que $a \in A$ y $b \in B$ se llama **producto cartesiano** . Se denota por $A \times B$.

Relaciones.

(1.8) **Definición**: Una **relación binaria R** entre A y B es una ley o criterio que permite señalar ciertos pares ordenados en $A \times B$. Las propiedades más frecuentes que puede tener una relación binaria sobre A, es decir entre A y A, son:

i) **Reflexiva**: $a\mathbf{R}a, \forall a \in A$.

ii) **Simétrica**: Si $a\mathbf{R}b$, entonces $b\mathbf{R}a$.

iii) **Transitiva**: Si $a\mathbf{R}b$ y $b\mathbf{R}c$, entonces $a\mathbf{R}c$.

iv) **Antisimétrica**: Si $a\mathbf{R}b$ y $b\mathbf{R}a$, entonces $a = b$.

Una relación sobre un conjunto A se dice de **relación de equivalencia** si tiene las propiedades i),ii) y iii). Una relación sobre un conjunto A es una **relación de orden** si tiene las propiedades i), iii) y iv).

Dada una relación de equivalencia sobre un conjunto A y $a \in A$, la **clase de equivalencia** de a está dada por el conjunto $\{b \in A | b\mathbf{R}a\} := [a]$. Notar que el conjunto A puede expresarse como unión disjunta de sus distintas clases de equivalencia. Observar que $a\mathbf{R}b \Leftrightarrow b\mathbf{R}a$. Es importante comprender que $a\mathbf{R}b \Leftrightarrow [a] = [b]$.

Dado un conjunto A sobre el que hay definida una relación de equivalencia $\mathbf{R}$, llamaremos **conjunto cociente** de A por $\mathbf{R}$ y lo denotaremos por $A/\mathbf{R}$ al conjunto cuyos elementos son las clases de equivalencia de A bajo la relación $\mathbf{R}$.

Ejemplos. 1) Si $f : A \longrightarrow B$ es una aplicación de A en B, queda definida una relación de equivalencia sobre el conjunto A por el siguiente criterio: Si $a, a' \in A$, entonces $a\mathbf{R}\, a'$ si $f(a) = f(a')$

2) Definimos en $\mathbf{Z}$ la relación de **congruencia**, es decir: dos elementos $a, b \in \mathbf{Z}$ son congruentes módulo $n > 0$ y lo escribiremos $a \equiv b\,(n)$, cuando $a - b = \dot{n}$ (múltiplo de n). La congruencia es una relación de equivalencia. La clase de equivalencia definida por un entero cualquierea a es $[a] := \{b | a \equiv b\,(n)\} = \{b | b = a + zn$, para algún $z \in \mathbf{Z}\}$:$= a + n\mathbf{Z}$. Las clases de equivalencia son llamadas **clases de congruencia**. Se suele emplear la notación $\bar{a}$ para $[a]$.

(1.9) **Definición**: Diremos que un conjunto A es **ordenado** (**parcialmente ordenado**) si existe una relación de orden $\mathbf{R}$ sobre él.

Usualmente escribiremos $\leq$ para una tal relación. El conjunto A se dice **totalmente ordenado**, cuando A es ordenado y cumple además que dados dos elementos cualesquiera $a, b \in A$, se tiene que $a \leq b$ ó $b \leq a$. A un conjunto totalmente ordenado también se le llama **cadena**.

Dado un conjunto ordenado aparecen los siguientes elemento notables:

i) Un **mayorante** ó **cota superior** de $B \subseteq A$: es un elemento $a \in A \ni b \leq a, \forall b \in B$.

ii) Un **minorante ó cota inferior** de $B \subseteq A$ es un elemento $a \in A \ni a \leq b, \forall b \in B$.

iii) El **máximo** de A es un elemento $a \in A \ni x \leq a, \forall x \in A$.

iv) El **mínimo** de A es un elemento $a \in A \ni a \leq x, \forall x \in A$.

v) Un **máximal** de A es un elemento $a \in A$ tal que si existe $b \in A$ tal que $a \leq b$ entonces $a = b$.

vi) Un **minimal** de A es un elemento $a \in A$ tal que si existe $b \in A$ tal que $b \leq a$ entonces $a = b$.

vii) El **supremo** de B en A es la menor cota superior (si existe) de B en A.

viii) El **ínfimo** de B en A es la mayor cota inferior (si existe) de B en A.

(1.10) **Lema de Zorn**: Si A es un conjunto ordenado, no vacío, en el que cada cadena de A tiene una cota superior en A , entonces A posee al menos un elemento maximal.

(1.11) **Lema**: Sea $n \neq 0$ un entero positivo. Para cualquier $a \in \mathbf{N}$ existen números naturales q, r con $0 \leq r < n$ tales que $a = nq + r$.

Demostración: Lo probaremos por inducción sobre a.

Si $a = 0$ tomar $q = 0 = r$. Por la hipótesis inductiva, existen q_1, r_1 tales que $a - 1 = q_1.n + r_1$, con $0 \leq r_1 < n$. Por tanto, $a = q_1.n + (r_1 + 1)$ y $r_1 + 1 \leq n$. Si $r_1 + 1 < n$, tomar $q = q_1, r = r_1 + 1$. Si $r_1 + 1 = n$, tomar $q = q_1 + 1, r = 0$.

(1.12) **Teorema**: Sea $n \neq 0$ un entero positivo. Para cualquier $a \in \mathbf{Z}$ existen enteros únicos q, r con $0 \leq r < n$ tales que $a = n.q + r$.

Demostración: Probamos en primer lugar la existencia.

Si $a \geq 0$ está probado en el lema anterior. Si $a < 0$, entonces , por el lema $-a = q_1.n + r_1$, así que $a = -q_1.n - r_1$. Si $r_1 = 0$ tomar $q = -q_1, r = 0$. Si $0 < r_1$, entonces $a = -n - q_1.n + (n - r_1)$ y bastaría tomar $q = (-1 - q_1), r = n - r_1$.

Unicidad: Supongamos que existan dos descomposiciones : $a = n.q_1 + r_1$, y $a = n.q_2 + r_2$, con $0 \leq r_i < n$. Entonces se tiene que $0 = n(q_2 - q_1) + r_2 - r_1$. Supongamos ahora $r_1 \neq r_2$. Si $r_1 < r_2$, entonces $0 < r_2 - r_1 < n$, lo que contradice a la igualdad $n(q_1 - q_2) = r_2 - r_1$. De modo análogo encontraríamos una contradicción en el caso $r_2 < r_1$.

(1.13) **Corolario**: Sea $n \neq 0$ un entero . Para cualquier $a \in \mathbf{Z}$ existen enteros únicos q, r con $0 \leq r < |n|$ tales que $a = n.q + r$.

Demostración: Si $n > 0$, el resultado es cierto por (1.12).

Si $n < 0$, por (1.12) se tiene que $a = (-n)q + r$ con $r = 0$ ó $0 < r < -n = |n|$.

Supongamos que existan dos descomposiciones : $a = n.q_1 + r_1$, y $a = n.q_2 + r_2$, con $0 \leq r_i < |n|$. Supongamos ahora $r_1 \neq r_2$. Si $r_1 < r_2$, entonces $0 < r_2 - r_1 < |n|$, lo que contradice a la igualdad $n(q_1 - q_2) = r_2 - r_1$. De modo análogo encontraríamos una contradicción en el caso $r_2 < r_1$.

Luego debe ser $r_1 = r_2$ y se sigue que $0 = n(q_2 - q_1)$ y por tanto $q_1 = q_2$ como se quería demostrar.

(1.14) **Definición**: Dados dos enteros a, b un máximo común divisor de a y b es un entero d que cumple:

i) d divide a a y a b;

ii) si d' es otro entero que divide tanto a a como a b, entonces d' divide a d.

Observación : Sean a, b dos enteros no nulos, por (1.13) existe r_1 tal que $a = q_1 b + r_1$, con $0 \leq r_1 < |b|$. Si $r_1 \neq 0$ existe r_2 tal que $b = q_2 r_1 + r_2$ con $0 \leq r_2 < r_1$, ..., así si $r_{i-1} \neq 0$ existe r_i tal que $r_{i-2} = q_i.r_{i-1} + r_i$ con $0 \leq r_i < r_{i-1}$, por tanto debe existir un n tal que $r_{n-1} = q_{n+1} r_n + 0$. Hemos obtenido así una sucesión de números naturales r_i que verifican : $|b| = r_0 > r_1 > ... > r_n$ con $r_n | r_i$, $i = 0, 1, .., n-1$. En particular $r_n | b$ y por tanto a a.

(1.15) **Teorema**:(Identidad de Bezout): Sean a, b dos enteros no nulos, entonces existe d máximo común divisor de a y b. Además existen enteros s, t tales que $d = s.a + t.b$.

Demostración: Con la notación de la observación anterior, tenemos: $r_1 = 1.a - q_1 b$, suponiendo probado que hasta el índice i existen enteros s_i, t_i tales que $r_i = s_i.a + t_i.b$, de la expresión $r_{i+1} = r_{i-1} - r_i q_{i+1}$, se obtiene la existencia de elementos s_{i+1}, t_{i+1} tales que $r_{i+1} = s_{i+1}.a + t_{i+1}.b$. Por lo tanto, existen elementos s, t tales que $r_n = s.a + t.b$. Para concluir la demostración basta observar que r_n es máximo común divisor de a y b, ya que sabemos que r_n divide a a y a b; además, de la expresión $r_n = s.a + t.b$, es claro que si d' divide tanto a a como a b, tambien d' dividirá a r_n.

Recapitulando r_n es máximo común divisor de a y de b y se verifica la identidad enunciada. Usualmente escribiremos m.c.d(a, b) ó incluso simplemente (a, b)

Consecuencias:

i) Dado un entero positivo n hay exactamente n clases de congruencia.

Las clases de congruencia son las diferentes progresiones aritméticas de razón n ilimitadas en ambos sentidos, esto es:

$[0] = \{\ldots, -2n, -n, 0, n, 2n \ldots\}; [1] = \{\ldots, 1-2n, 1-n, 1, 1+n, 1+2n \ldots\},$

$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,$

$[n-1] = \{\ldots, -n-1, -1, n-1, n-1+n, n-1+2n \ldots\}.$

ii) En **N** la relación de divisibilidad es relación de orden, pero con ella **N** no es totalmente ordenado.

iii) En **Z** la divisibilidad no es relación de orden. No es simétrica.

iv) Si A es cualquier conjunto, en el conjunto $\mathcal{P}(A) = \{X | X \subseteq A\}$, llamado el conjunto de las **partes de** A, la relación de inclusión es una relación de orden.

Lección 2. Leyes de Composición.

Dado tres conjunto A,B y C una ley de composición binaria u operación binaria es una aplicación $f : A \times B \longrightarrow C$.

(2.1) **Definición**: Dado un conjuntos $A \neq \emptyset$, una operación binaria interna (o ley de composición interna) sobre A es una aplicación $f : A \times A \longrightarrow A$. La imagen $f(a, b)$ del par (a, b) se acostumbra a escribirla utilizando un símbolo (por ejemplo $*$), poniendo $a * b$ en vez de $f(a, b)$ y el elemento $a * b$ de A se dice resultado (o compuesto) de a por b mediante la operación $*$. **Ejemplos**: 1) $* : \mathbf{N} \times \mathbf{N} \longrightarrow \mathbf{N}$, dada por $a * b = a + b$.

2) Sea E un conjunto y $A = F(E, E) = \{$ aplicaciones de E en $E\}$, entonces $f * g = g \circ f$ es una operación binaria interna.

(2.2) **Definición**: Sea A un conjunto con una operación binaria interna $*$ y sean B y C dos subconjuntos no vacíos de A. Se llama compuesto de B y C al subconjunto $\{b * c | b \in B, c \in C\} = B * C$.

Queda así definida una operación binaria interna en $\mathcal{P}(A) - \{\emptyset\}$, y por ello distinta de la anterior, aunque también la sigamos denotando con el mismo signo.

(2.3) **Definición**: Una parte B de A se dice estable (o cerrada) para la operación $*$, si $B * B \subseteq B$, es decir, si siempre que: $a \in B$ y $b \in B$, se sigue $a * b \in B$. En tal caso B aparece dotado de una operación binaria interna que es la restricción de $*$ a B (ó a $B \times B$) y que no hay inconveniente en designarla con el signo $*$.

(2.4) **Definición**: Sea A un conjunto dotado de una operación binaria $*$ y B otro dotado de una $(.)$. Entonces en $A \times B$ queda definida una nueva operación interna α , escribiendo $(a, b)\alpha(a', b') = (a * a, b.b')$, que se dice operación producto de $*$ y ..

El caso más interesante es cuando $A = B$ y $(*) = (.)$, obteniéndose en este caso la extensión de $*$ a $A \times A$.

Un conjunto A con una operación binaria interna $*$ se denomina **grupoide**.

Una ley de composición interna $*$ sobre A puede tener una ó varias de las propiedades siguientes:

(2.5) **Definiciones**

a) **Asociativa**: $\forall a, b, c \in A : a * (b * c) = (a * b) * c$.

Sobre el conjunto de los números reales no nulos la operación producto es asociativa, sin embargo la operación cociente no es asociativa.

b) **Conmutativa**: $\forall a, b \in A : a * b = b * a$.

Sobre el conjunto de los números reales no nulos la operación producto es conmutativa, sin embargo la operación cociente no lo es.

(2.6) **Definición**: **Elementos notables respecto de una operación binaria interna**:

Un elemento $a \in A$: es **simplificable ó regular, a izquierda** si la igualdad $a * b = a * c$ implica que $b = c$

simplificable a derecha: si la igualdad $b * a = c * a$ implica que $b = c$

Un elemento $e \in A$ es **neutro a izquierda**, si $\forall a \in A$ es $e * a = a$. Análogamente se define **elemento neutro a derecha**. Un elemento $e \in A$ es **neutro**, si $\forall a \in A$ es $e * a = a * e = a$.

Si una operación interna $*$ posee un elemento neutro, este es único.

En una operación interna con elemento neutro, dos elementos a y a' se dicen **simétricos** si cumplen: $a * a' = a' * a = e$. Si la notación es multiplicativa, se dicen inversos.

Si en una operación asociativa un elemento a posee simétrico, este es único.

Si a es un elemento en un conjunto con una operación binaria interna asociativa, entonces se escribe $a * \ldots^{p} * a := a^p, p \in \mathbf{N} - \{0\}$. Se sigue que $\forall p, q \in \mathbf{N}$, $a^p * a^q = a^{p+q}, (a^p)^q = a^{pq}$. Si $(A, *)$ posee elemento neutro e, por convenio será $a^0 = e \forall a \in A$. Si a posee simétrico que denotamos por a^{-1}, entonces a^p tiene simétrico $(a^{-1})^p$ que escribiremos a^{-p}. Con los convenios anteriores, es fácil comprobar que $\forall p, q \in \mathbf{Z}$, $a^p * a^q = a^{p+q}, (a^p)^q = a^{pq}$. Si además la operación $*$ es conmutativa se tiene que $(a * b)^p = a^p * b^p, p \in \mathbf{Z}$

Si en $(A, *)$ existe elemento neutro e, un elemento $a \in A$ es **nilpotente** si existe un $n \in \mathbf{N} - \{0\}$ tal que $a^n = e$.

Un elemento $a \in A$ es **idempotente** si $a * a = a$.

(2.7) **Definición**: Dado un conjunto A con dos operaciones internas $*$ y ., se dice que $(.)$ es

distributiva respecto de $(*)$ a izquierda si cumple: $\forall a, b, c \in A$, $a.(b * c) = (a.b) * (a.c)$

distributiva respecto de $(*)$ a derecha si cumple: $\forall a, b, c \in A$, $(a * b).c = (a.c) * (b.c)$

distributiva respecto de $(*)$, si lo es a izquierda y a derecha.

(2.8) **Definición**: **Operación binaria externa** sobre un conjunto A, con dominio de operadores **a derecha** K es una aplicación $\bullet : A \times K \longrightarrow A$. La imagen por $\bullet$ del par (a, k) será usualmente denotada por ak.

Operación binaria externa sobre un conjunto A, con dominio de operadores a **izquierda** K es una aplicación $\bullet : K \times A \longrightarrow A$. La imagen por $\bullet$ del par (k, a) será usualmente denotada por ka.

(2.9) **Definición**: Una parte B de A es **estable** (ó **cerrada**) respecto de la operación externa $\bullet$, con dominio de operadores a izquierda K si es aplicación $\bullet|_{K\times B} : K \times B \longrightarrow B$, es decir, si cualesquiera que sean los elementos $t \in K$ y $b \in B$ se verifica que $t \bullet b \in B$

Análogamente definiremos parte estable para una operación externa con dominio de operadores por la derecha.

(2.10) **Definición**: Supongamos que sobre el conjunto A hay definida una relación de equivalencia $\mathbf{R}$ y una operación interna $*$. Diremos que $*$ es estable respecto de $\mathbf{R}$ si se cumple: $a\mathbf{R}a'$ y $b\mathbf{R}b' \Rightarrow (a * b)\mathbf{R}(a' * b')$.

(2.11) **Teorema**: Si la operación $*$ es estable respecto de la relación de equivalencia $\mathbf{R}$ la asignación $([a], [b]) \mapsto [a * b]$ es una aplicación de $A/\mathbf{R} \times A/\mathbf{R} \to A/\mathbf{R}$ y por tanto una operación binaria interna sobre $A/\mathbf{R}$. Dicha operación binaria hereda las propiedades de la inicialmente definida sobre A

Demostración: Basta probar que la clase $[a * b]$ es única, independientemente de los representantes de $[a]$ y de $[b]$. Si tomamos otros representantes $a' \in [a]$ y $b' \in [b]$, entonces, $(a' * b')\mathbf{R}(a * b)$ ya que la relación $\mathbf{R}$ es estable para $*$. Luego $[a * b] = [a' * b']$.

(2.12) **Definición**: La operación interna cuya existencia asegura el teorema anterior se llama **operación inducida** por $*$ en $A/\mathbf{R}$.

(2.13) **Definición**: Supongamos que sobre el conjunto A hay definida una relación de equivalencia $\mathbf{R}$ y una operación externa $\bullet$ con dominio de operadores a izquierda K. Diremos que $\bullet$ es estable respecto de $\mathbf{R}$ si se cumple: $\forall t \in K\ a\mathbf{R}a' \Rightarrow (t \bullet a)\mathbf{R}(t \bullet a')$. También se dice entonces que $\mathbf{R}$ es estable para $\bullet$. (La definición análoga a la derecha).

(2.14) **Teorema**: Si la operación $\bullet$ es estable respecto de la relación de equivalencia $\mathbf{R}$ la asignación $(t, [a]) \mapsto [t \bullet a]$ es una aplicación de $K \times A/\mathbf{R} \to A/\mathbf{R}$ y por tanto una

operación binaria externa sobre $A/\mathbf{R}$ con dominio de operadores a izquierda K que se dice **inducida** por la operación $\bullet$.

Demostración: Basta probar que la clase $[t \bullet a]$ es única, independientemente del representante de $[a]$: Si tomamos otro representantes $a' \in [a]$, entonces $(t \bullet a')\mathbf{R}(t \bullet a)$ ya que la relación $\mathbf{R}$ es estable para $\bullet$. Luego $[t \bullet a] = [t \bullet a']$.

(2.15) **Definición**: La operación interna cuya existencia asegura el teorema anterior se llama **operación inducida** por $\bullet$ en $A/\mathbf{R}$.

Ejemplos: Sea n un entero positivo. En el conjunto $\mathbf{Z}$ de los números enteros hay definidas dos operaciones + y (.). Ambas son estables respecto de la relación de congruencia módulo n. Por (2.11), en el conjunto cociente $\mathbf{Z}/n\mathbf{Z}$ hay dos operaciones internas inducidas por + y (.) respectivamente, definidas de la siguiente forma $[a]+[b] = [a+b]$ y $[a].[b] = [a.b]$.

Lección 3. Grupos

(3.1) **Definición**: Un **grupo** es un conjunto no vacío G dotado de una operación binaria interna $\mathbf{.}$ que es asociativa, de forma que existe $e \in G$ verificando:

i) para cada $x \in G : e.x = x.e = x$.

ii) Para cada $x \in G$ existe $y \in G$ tal que $x.y = y.x = e$.

Ejemplos:i) **Grupo simétrico** sobre un conjunto $\Omega \neq \emptyset$ ó grupo de todas las permutaciones de Ω

ii)$(\mathbf{Z}, +), (\mathbf{Q}, +), (\mathbf{R}, +), (\mathbf{Q}^*, .), (\mathbf{R}^*, .)$;

iii) El **cuatro-grupo de Klein**, V_4 o K_4 .

iv) El grupo de los movimientos de un cuadrado D_8, conocido como el **grupo diédrico** de orden 8;

v) El **grupo cuaternio** de orden 8, Q_8 es el conjunto $\{\pm 1, \pm i, \pm j, \pm k\}$ cuya tabla de multiplicación, basada en las reglas $i^2 = j^2 = k^2 = -1, i.j = k, j.k = i, k.i = j$.

Damos a continuación las tablas de los grupos Q_8, V_4, D_8:

	1	-1	i	j	k	-i	-j	-k
1	1	-1	i	j	k	-i	-j	-k
-1	-1	1	-i	-j	-k	i	j	k
i	i	-i	-1	k	-j	1	-k	j
j	j	-j	-k	-1	i	k	1	-i
k	k	-k	j	-i	-1	-j	i	1
-i	-i	i	1	-k	j	-1	k	-j
-j	-j	j	k	1	-i	-k	-1	i
-k	-k	k	-j	i	1	j	-i	-1

	1	(1,3)	(2,4)	(1,3)(2,4)
1	1	(1,3)	(2,4)	(1,3)(2,4)
(1,3)	(1,3)	1	(1,3)(2,4)	(2,4)
(2,4)	(2,4)	(1,3)(2,4)	1	(1,3)
(1,3)(2,4)	(1,3)(2,4)	(2,4)	(1,3)	1

	1	(1,3)	(2,4)	(1,3)(2,4)	(1,2)(3,4)	(1,4)(2,3)	(1,2,3,4)	(1,4,3,2)
1	1	(1,3)	(2,4)	(1,3)(2,4)	(1,2)(3,4)	(1,4)(2,3)	(1,2,3,4)	(1,4,3,2)
(1,3)	(1,3)	1	(1,3)(2,4)	(2,4)	(1,4,3,2)	(1,2,3,4)	(1,4)(2,3)	(1,2)(3,4)
(2,4)	(2,4)	(1,3)(2,4)	1	(1,3)	(1,2,3,4)	(1,4,3,2)	(1,2)(3,4)	(1,4)(2,3)
(1,3)(2,4)	(1,3)(2,4)	(2,4)	(1,3)	1	(1,4)(2,3)	(1,2)(3,4)	(1,4,3,2)	(1,2,3,4)
(1,2)(3,4)	(1,2)(3,4)	(1,2,3,4)	(1,4,3,2)	(1,4)(2,3)	1	(1,3)(2,4)	(1,3)	(2,4)
(1,4)(2,3)	(1,4)(2,3)	(1,4,3,2)	(1,2,3,4)	(1,2)(3,4)	(1,3)(2,4)	1	(2,4)	(1,3)
(1,2,3,4)	(1,2,3,4)	(1,2)(3,4)	(1,4)(2,3)	(1,4,3,2)	(2,4)	(1,3)	(1,3)(2,4)	1
(1,4,3,2)	(1,4,3,2)	(1,4)(2,3)	(1,2)((3,4)	(1,2,3,4)	(1,3)	(2,4)	1	(1,3)(2,4)

Ejercicio.: Si A y B son conjuntos biyectivos entonces los correspondientes grupos

simétricos son isomorfos.

Una vez probado el ejercicio anterior, denotaremos por Σ_n al grupo de todas las permutaciones de un conjunto de n elementos.

(3.2) **Lema** : Sea G un grupo. Entonces para $a, b \in G$ existen únicos elementos $x, y \in G$ tales que:

$$a.x = b, y.a = b.$$

Demostración: Sea $z \in G$ tal que $a.z = e = z.a$, entonces

$$a.(z.b) = (a.z).b = e.b = b,$$

así basta tomar $x = z.b$. Para la unicidad, si $a.x = a.x'$, tenemos $x = e.x = z.a.x = z.a.x' = e.x' = x'$. Para y tomar $y = b.z$ y realizar las mismas comprobaciones.

(3.3) **Teorema** : Sea G un conjunto con una operación binaria multiplicativa asociativa, y supongamos que existe $e \in G$ con las propiedades:

i) $x.e = x, \forall x \in G$,

ii) Para cada $x \in G$ existe $y \in G$ tal que $x.y = e$.

Entonces G es un grupo.

Demostración: Veamos que $e.x = x$ y que $y.x = e$. Usando ii) encontramos z con $yz = e$, y se tiene: $x = x.e = x.(y.z) = (x.y).z = e.z$ luego : $y.x = y.(e.z) = (y.e).z = y.z = e$. Ahora $e.x = (x.y).x = x.(y.x) = x.e = x$.

En lo sucesivo, si no hay confusión, denotaremos con 1 al elemento neutro del grupo. Dado un elemento $x \in G$ denotaremos con x^{-1} al inverso del elemento x.

(3.4) **Definición** : Un subconjunto $H \neq \emptyset$ de un grupo G es un **subgrupo** de G si es cerrado bajo la multiplicación de G y forma un grupo con respecto de dicha multiplicación.

(3.5) **Lema** : Sea G un grupo y $\emptyset \neq H \subseteq G$ un subconjunto. Suponer que $x.y^{-1} \in H$ para todo par de elementos $x, y \in H$. Entonces H es un subgrupo de G. En particular, cualquier subconjunto no vacío de G que sea cerrado para la multiplicación y la formación de inversos, es un subgrupo de G.

Demostración: Tomar $h \in H$. Entonces $1 = h.h^{-1} \in H$, por la hipótesis. Para $y \in H$, $y^{-1} = 1.y^{-1} \in H$ y si también $x \in H$, $x.y = x(y^{-1})^{-1} \in H$, así la multiplicación en G define una operación interna en H que será también asociativa y puesto que $1 \in H$ y que $y^{-1} \in H$ si $y \in H$, H pasa a ser un grupo con dicha operación.

(3.6) **Corolario**: Sea $\mathcal{H}$ una colección de subgrupos de un grupo G y sea $D = \bigcap_{H \in \mathcal{H}} H$, entonces D es un subgrupo de G.

Demostración: Como $1 \in H$, para cualquier $H \in \mathcal{H}$, se tiene que $1 \in D$ y así $D \neq \emptyset$. Ahora, si $x, y \in D$, entonces $x, y \in H, \forall H \in \mathcal{H}$, y así $x.y^{-1} \in H, \forall H \in \mathcal{H}$, luego $x.y^{-1} \in D$.

Si $X \subseteq G$, podemos considerar la familia $\mathcal{H}$ de subgrupos de G conteniendo a X, $(G \in \mathcal{H})$. El subgrupo $\bigcap_{H \in \mathcal{H}} H$ se llama **subgrupo generado** por X y se denota $< X >$. Se caracteriza por dos propiedades:

1. $X \subseteq < X >$,

2. Si $X \subseteq H$ y $H \leq G$, entonces $< X > \leq H$.

Así $< X >$ es el menor subgrupo de G conteniendo a X.

(3.7) **Lema** : Sea G un grupo y $X \subseteq G$. Entonces $< X >$ es el conjunto de todos los productos finitos $u_1 \dots u_n$ de elementos u_i en G tales que u_i ó u_i^{-1} está en X. (Entendiendo el 1 como el producto vacío ó con $n = 0$).

Demostración: Sea S el conjunto de dichos productos. Notar que $S \neq \emptyset$ (incluso si $X = \emptyset$). S es claramente cerrado por multiplicación y

$$(u_1 \dots u_n)^{-1} = u_n^{-1} \dots u_1^{-1} \in S,$$

luego $S \leq G$. Como $X \subseteq S$, tenemos que $< X > \leq S$ y como $< X > \leq G$ que contiene a X y $< X >$ es cerrado por multiplicación e inversos, se sigue que $S \leq < X >$. Luego coinciden.

Un caso particular muy importante es cuando $|X| = 1$. Un grupo G se dice **cíclico** si existe $g \in G$ tal que $< g >= G$. El siguiente resultado es inmediato a partir del lema anterior:

(3.8) **Corolario**: Sea $g \in G$, entonces $< g >= \{g^n | n \in \mathbf{Z}\}$.

(3.9) **Lema**: Sea $< g >= G$. Sea $H \leq G$ y sea n el entero positivo más pequeño tal que $g^n \in H$, entonces:

i) para $m \in \mathbf{Z}$, $g^m \in H \Leftrightarrow n|m$.

ii) $H =< g^n >$.

Demostración: i) Si $n|m$, $m = n.q, q \in \mathbf{Z}$ así $g^m = (g^n)^q \in H$. Recíprocamente, si $g^m \in H$, escribir $m = n.q + r$ con $0 \leq r < n$, entonces: $g^m = g^{n.q+r} = ((g)^n)^q.g^r$, de donde $g^r \in H$, así, debe ser $r = 0$.

ii) Es claro que $< g^n > \leq H$ y si $h \in H$, entonces $h = g^m$ con $n|m$ por i), así $g^m = g^{n.q} = (g^n)^q \in < g^n >$.

(3.10) **Corolario**: Cada subgrupo de un grupo cíclico es también cíclico.

(3.11) **Definición**: Sea G un grupo y $g \in G$. Llamamos **orden** de g al menor natural no nulo n tal que $g^n = 1$. Si no existe, se dirá que g es de **orden infinito**.

(3.12) **Lema**: Sea $g \in G$ con $o(g) = n < \infty$. Entonces:

i) $g^m = 1 \Leftrightarrow n|m$.

ii) $g^m = g^i \Leftrightarrow m \equiv i(n)$.

iii) $| < g > | = n$.

Demostración: i) Aplicar el lema anterior a $< g >$ con $H = 1$. ii) se sigue de i).

iii) Por ii), los elementos de $< g >$ están en correspondencia biunívoca con las clases de restos de los enteros módulo n y existen exactamente n clases.

(3.13) **Teorema**: Sea G un grupo finito cíclico de orden n. Entonces para cada divisor d de n, existe un único subgrupo de orden d.

Demostración: Sea $G =< g >$. Sea $e = n/d$, entonces $(g^e)^d = g^n = 1$, así $o(g^e)|d$, por otra parte si $o(g^e) = l$ es $g^{e.l} = 1$, así $n = e.d|e.l$ luego $d|l$, luego debe ser $o(g^e) = d$. Así por iii) del lema anterior $| < g^e > | = o(g^e) = d$. Sea $H \leq G$, con $|H| = d$, por el corolario (3.10), $H =< g^m >$, para algún entero m, que divide a cualquier otro entero s con $g^s \in H$. Como $g^n = 1 \in H, m|n$ y $n = m.l$. Luego $m = n/l$, pero hemos visto antes que $| < g^{n/l} > | = l$, luego $l = d$ y $m = n/d$, por tanto $H =< g^e >$.

Se define la **función de Euler** φ, asociando a cada entero positivo n,

$$\varphi(n) = |\{r \in \mathbf{Z} | 1 \leq r < n, (r,n) = 1\}.$$

(3.14) **Teorema**: Sea G un grupo cíclico de orden n, entonces G contiene exactamente $\varphi(n)$ elementos de orden n que son los de la forma g^r tales que $1 \leq r \leq n, (r,n) = 1$ siendo g cualquier elemento de orden n en G.

Demostración: Se trata de probar que $o(g^r) = n \Leftrightarrow (r,n) = 1$. Notar que en general $o(g^r) = n/(n,r)$, ya que si $o(g^r) = s$, como $(g^r)^{(n/(n,r))} = 1 \Rightarrow s|n/(n,r)$. Por otra parte, $(g^r)^s = 1$, así $n|r.s$ luego $n/(n,r)|(r/(n,r)).s$ y así $n/(n,r)|s$. Por tanto $s = n/(n,r)$.

(3.15) **Definición**: Un grupo G es **abeliano** si $x.y = y.x, \forall x, y \in G$.

Si $g \in G$, se definen $C_G(g) = \{x \in G | x.g = g.x\}$ y si $X \subseteq G$, $C_G(X) = \{y \in G | y.x = x.y, \forall x \in X\}$,.

Es sencillo comprobar que $C_G(g)$ y $C_G(X)$ son subgrupos de G. Denotaremos por $Z(G) := C_G(G)$ llamado **centro** de G.

Notar que $Z(G)$ es un grupo abeliano y que G es abeliano $\Leftrightarrow G = Z(G)$. Puede suceder que el centro de un grupo sea trivia, por ejemplo: $Z(\Sigma_3) = 1$.

Homomorfismos de Grupos .

(3.16) **Definición**: Una aplicación $\varphi : G \longrightarrow H$ donde G y H son grupos, se dice que es un **homomorfismo** de grupos si $\varphi(x.y) = \varphi(x).\varphi(y), \forall x, y \in G$.

Un homomorfismo inyectivo se dirá **monomorfismo**. Un homomorfismo suprayectivo se dirá **epimorfismo**. Un homomorfismo biyectivo se dirá **isomorfismo**.

Si $G = H$ el homorfismo se llamará **endomorfismo**. Un **automorfismo** es un endomorfismo biyectivo

Ejemplos: a) Sean $(\mathbf{C}^*, .)$ y $(\mathbf{R}^+, .)$, $\varphi : \mathbf{C}^* \longrightarrow R^+$ dada por $a \longrightarrow |a|$.

b) $(\mathbf{Z}, +)$ y $(\{1, -1\}, .)$, $\varphi : \mathbf{Z} \longrightarrow \{1, -1\}$ dada por $\varphi(n) = 1$ si n es par y $\varphi(n) = -1$ si n es impar.

c) El homomorfismo trivial $\varphi_o : G \longrightarrow H$ dado por $\varphi_o(g) = 1_H, \forall g \in G$.

d) Si G es un grupo abeliano, la aplicación $\varphi : G \longrightarrow G$ dada por $\varphi(x) = x^2, \forall x \in G$ es un endomorfismo, pero en general no es automorfismo.

Ejercicio.: Probar que G es abeliano si y sólo si la aplicación $\varphi : G \longrightarrow G$ dada por $\varphi(a) = a^{-1}$ es un homomorfismo.

(3.17) **Lema**: Sea $\varphi : G \longrightarrow H$ un homomorfismo, entonces:

i) $\varphi(1) = 1$, $\varphi(x^{-1}) = \varphi(x)^{-1}, \forall x \in G$.

ii) $N = \{g \in G | \varphi(g) = 1\}$ es un subgrupo de G. Además $N = 1$ si y solo si φ es un monomorfismo.

iii) Si S es un subgrupo de G, $\varphi(S)$ es un subgrupo de H.

iv) Si $V \leq H \Rightarrow \varphi^{-1}(V) \leq G$ y $N \leq \varphi^{-1}(V)$.

Demostración:i) $\varphi(1) = \varphi(1.1) = \varphi(1).\varphi(1)$, así que $\varphi(1) = 1$. Además $1 = \varphi(1) = \varphi(x).\varphi(x^{-1})$ luego $\varphi(x^{-1}) = \varphi(x)^{-1}$.

ii) Como $1 \in N$ es $N \neq \emptyset$. Si $x \in N, \varphi(x^{-1}) = \varphi(x)^{-1} = 1$, así $x^{-1} \in N$ y si $x, y \in N, \varphi(x.y) = \varphi(x).\varphi(y) = 1$, luego $x.y \in N$, por lo que $N \leq G$. Además si $g_1, g_2 \in G$, $\varphi(g_1) = \varphi(g_2)$ si y solo si $g_1 g_2^{-1} \in N$, y por tanto $N = 1$ si y solo si φ es un monomorfismo.

iii) Sean $h_1, h_2 \in \varphi(S)$, entonces existen $s_1, s_2 \in S$ tales que $\varphi(s_i) = h_i, i = 1, 2$. Por el apartado i) se tiene: $h_1.h_2^{-1} = \varphi(s_1).\varphi(s_2^{-1}) = \varphi(s_1.s_2^{-1}) \in \varphi(S)$.

iv) Sean $g_1, g_2 \in \varphi^{-1}(V)$. Entonces existen $v_1, v_2 \in V$ tales que $\varphi(g_i) = v_i, i = 1, 2$. Como V es subgrupo de H, $v_1.v_2^{-1} \in V$. Luego $\varphi(g_1)\varphi(g_2)^{-1} = \varphi(g_1.g_2^{-1}) \in V$. Por tanto $g_1.g_2^{-1} \in \varphi^{-1}(V)$.

Al subgrupo N se le llama **núcleo** del homomorfismo φ y se denota por Ker φ.

Para los ejemplos anteriores, se tiene:

a) Ker φ es el conjunto de los números complejos de módulo 1.

b) Ker $\varphi = 2\mathbf{Z}$.

c) Ker $\varphi = G$.

Si $\varphi : G_1 \longrightarrow G_2$ es un **isomorfismo** de grupos, es sencillo probar que $\varphi(Z(G_1)) = Z(G_2)$.

Un subgrupo H de G con la propiedad de que $\varphi(H) = H$ para cada automorfismo φ de G, se dice **característico** en G y se escribe H car G. Un ejemplo importante de

automorfismo de G es el **automorfismo interno** de G inducido por un elemento $g \in G$ dado por $\varphi_g(x) := x^g := g^{-1}.x.g$.

Si $H \leq G$, $H^g := \{h^g |\ h \in H\} \leq G$, pues $H^g = \varphi_g(H)$.

El conjunto Aut (G) de todos los automorfismos de G es un subgrupo de $\Sigma(G)$ y el de los automorfismos internos Int (G) es un subgrupo de Aut (G).

Vamos a introducir ahora uno de los conceptos más importantes en teoría de grupos:

(3.18) **Definición**: Un subgrupo N de G se dice **normal** en G y se escribe $N \trianglelefteq G$ si $N^g = N, \forall g \in G$. En otras palabras, los subgrupos normales de un grupo son exactamente, los subgrupos fijados por los automorfismos internos de dicho grupo.

Ej.: Probar que si $N \trianglelefteq G$ y $|N| = 2$ entonces $N \leq Z(G)$.

Notas: a) Si $\varphi : G \longrightarrow H$ es un homomorfismo, entonces $N = \operatorname{Ker} \varphi$ es un subgrupo normal de G.

b) N es un subgrupo normal de G si y solo si existe un grupo H y un homomorfismo π de G en H, tal que $N = \operatorname{Ker} \pi$

(3.19) **Lema** : Sea $H \leq G$. Suponer que $H^g \subseteq H, \forall g \in G$, entonces $H \trianglelefteq G$

Demostración: Debemos demostrar que $H^g = H, \forall g \in G$. Como $H^g \subseteq H, \forall g \in G$, tenemos $H = (H^g)^{g^{-1}} \subseteq H^{g^{-1}}, \forall g \in G$ y así, en particular $H \subseteq H^{(g^{-1})^{-1}} = H^g$ y por tanto $H = H^g$.

Ejercicio.: Si G es un grupo, probar que Int $(G) \trianglelefteq$ Aut (G).

Nota: Sea $H \leq G$, entonces $H = H^{-1}$ y

$$hH = Hh = H \Leftrightarrow h \in H$$

(3.20) **Lema**: i) Suponer $H, K \leq G$ entonces $HK := \{h.k \in H, k \in K\}$ es un subgrupo de G si y sólo si $HK = KH$.

ii) Sean $H \leq K \leq G$ y $L \leq G$, entonces:

$$K \cap HL = H(K \cap L) \text{ (identidad de Dedekind).}$$

Demostración: i) Suponer $HK \leq G$. Puesto que $1 \in H$, es $K \subseteq HK$ y de forma análoga $H \subseteq HK$, así $KH \subseteq HK$, pues HK es cerrado para la multiplicación. También, si $x \in HK$ entonces $x^{-1} \in HK$ así $x^{-1} = (h.k)^{-1} = k^{-1}h.^{-1}$ perteneciente a KH.

Supongamos ahora que $HK = KH$. Veamos que $HK \leq G$:

Sean $x, y \in HK$, entonces $x = h_1k_1, y = k_2h_2$ por tanto $x.y^{-1} = h_1k_1h_2^{-1}k_2^{-1}$ y $k_1h_2^{-1} = h_3k_3$ luego $x.y^{-1} = h_1h_3k_3k_2^{-1} \in HK$.

ii) $H(K \cap L) \subseteq K \cap HL$. Sea ahora $k \in K \cap HL$, entonces $k = h.l, h \in H, l \in L$, luego $l = h^{-1}.k \in K$ y así $k \in H(K \cap L)$.

Ejercicio.: Si $N \trianglelefteq G$ y $H \leq G$, entonces $NH \leq G$.

(3.21) **Definición 19**: Sea $H \leq G$. Si $g \in G$, entonces los conjuntos:

$$Hg = \{h.g | h \in H\},$$

$$gH = \{g.h | h \in H\}$$

son respectivamente la clase a derecha y a izquierda de H determinada por g.

Como $g \in gH$ y $g \in Hg$, es claro que G es unión de todas las clases a derecha (a izquierda) de cualquiera de sus subgrupos .

(3.22) **Lema**: Sea $H \leq G$:

i) $Hx \cap Hy \neq \emptyset \Rightarrow Hx = Hy$.

ii) $xH \cap yH \neq \emptyset \Rightarrow xH = yH$.

Demostración: Notemos que $Hh = H, \forall h \in H$. Así:

$$H(hx) = (Hh)x = Hx$$

por tanto, si $g \in Hx \cap Hy$ tenemos $Hg = Hx, Hg = Hy$ y en consecuencia $Hx = Hy$. Análogamente se prueba ii).

(3.23) **Corolario**: Sea $H \leq G$. Entonces G es unión disjunta de las distintas clases a derecha (a izquierda) de H.

(3.24) **Lema**: Sea $H \leq G$. Para cada $g \in G$, tenemos

$$|gH| = |H| = |Hg|$$

(3.25) **Lema**: Los conjuntos de clases a izquierda y a derecha de H en G son biyectivos.

Demostración: Establezcamos la asignación $gH \longrightarrow Hg^{-1}$ .entonces:

$$g_1H = g_2H \Leftrightarrow Hg_1^{-1} = (g_1H)^{-1} = (g_2H)^{-1} = Hg_2^{-1}$$

así la aplicación está bien definida y es inyectiva. Es claro que es suprayectiva.

Como consecuencia, puede definirse el **índice** de H en G como el cardinal de cualquiera de esos conjuntos y lo denotaremos por $|G : H|$.

Ejercicio: Si $H \leq G$ y $|G : H| = 2$, entonces $H \trianglelefteq G$.

(3.26) **Teorema** (Teorema de Lagrange): Sea $H \leq G$, entonces $|G| = |H||G : H|$. En particular si G es finito, $|H|$ divide a $|G|$ y $|G : H| = |G|/|H|$.

Demostración: G es la unión disjunta de $|G : H|$ clases a derecha de H en G, cada una de las cuales es de cardinal igual al de H.

(3.27) **Corolario**: Sea G un grupo finito y $g \in G$, entonces $o(g)$ divide a $|G|$ y $g^{|G|} = 1$.

Demostración: Sabemos que $o(g) = | < g > |$ y este divide al orden de G por el teorema de Lagrange.

Ejercicio.: Si $H \leq K \leq G$ probar que $|G : H| = |G : K||K : H|$.

Si $H \trianglelefteq G$ usamos la notación G/H para denotar $\{Hg|g \in G\}$. Recordar que $Hg = gH, \forall g \in G$.

(3.28) **Teorema**: Sea $H \trianglelefteq G$ se define en G/H:

$$(Hg_1).(Hg_2) = Hg_1g_2,$$

$g_1, g_2 \in G$. Con dicha operación G/H pasa a ser un grupo cuyo elemento neutro es H y $(Hx)^{-1} = Hx^{-1}$. A dicho grupo se le llama **grupo cociente** de G por H.

Demostración: Veamos que la operación está bien definida. Si $Hg_1 = H\tilde{g}_1$ y $Hg_2 = H\tilde{g}_2$, entonces $g_1\tilde{g}_1^{-1} \in H$ y $g_2\tilde{g}_2^{-1} \in H$. Veamos $Hg_1g_2 = H\tilde{g}_1\tilde{g}_2$ es decir que $g_1g_2(\tilde{g}_1\tilde{g}_2)^{-1} \in H$, pero ello es claro ya que

$$g_1g_2(\tilde{g}_1\tilde{g}_2)^{-1} = g_1g_2\tilde{g}_2^{-1}\tilde{g}_1^{-1} = g_1(g_2\tilde{g}_2^{-1})g_1^{-1}g_1\tilde{g}_1^{-1} \in H$$

pues $H \trianglelefteq G$.

El resto es sencillo de comprobar

(3.29) **Nota**: Si $H \leq (\mathbf{Z}, +)$ entonces H es cíclico y viene generado por un entero positivo, así $H = n\mathbf{Z}$ y aparece el grupo cociente $\mathbf{Z}/n\mathbf{Z} =< 1 + n\mathbf{Z} >$ con $o(1 + n\mathbf{Z}) = n$.

(3.30) **Definición**: Dado un subconjunto X de G, se llama normalizador en G de X a $N_G(X) := \{g \in G | X^g = X\}$.

(3.31) **Lema**: Dado un subconjunto X de G, el normalizador $N_G(X)$ es un subgrupo de G y si $X \leq G$, entonces $X \leq N_G(X)$.

b) $C_G(X) \trianglelefteq N_G(X)$

(3.32) **Definición**: Dado un grupo G, dos elementos $x, y \in G$ se dicen **conjugados** en G si existe un elemento $g \in G$ tal que $y = x^g = g^{-1}xg$.

La relación de conjugación es una relación de equivalencia sobre G. Las clases de equivalencia bajo esta relación son llamadas **clases de conjugación** de G. Por tanto el grupo G queda expresado como unión disjunta de sus clases de conjugación.

Si $N \trianglelefteq G$, se define $\pi : G \longrightarrow G/N$, por $\pi(g) = Ng$. Como $Ngh = Ng.Nh$, π es un homomorfismo que se llama **homomorfismo canónico**, y Ker $\pi = N$. Con esta observación obtenemos:

(3.33) **Teorema**: Sea $\varphi : G \longrightarrow H$ un homomorfismo suprayectivo y $N = \text{Ker}\,\varphi$, entonces $H \cong G/N$. De hecho, existe un isomorfismo canónico $\psi : G/N \longrightarrow H$ tal que $\pi\psi = \varphi$, donde π es el homomorfismo canónico de G sobre G/N.

Demostración: Puesto que ψ va a verificar $\pi\psi = \varphi$, definamos $\psi(Ng) = \varphi(g)$, $\forall g \in G$. Veamos que ψ está bien definida. Si $Nx = Ny$, sabemos que $\varphi(x) = \varphi(y)$. Por otra parte $\psi(Nx.Ny) = \psi(Nxy) = \varphi(x.y) = \varphi(x).\varphi(y) = \psi(Nx)\psi(Ny)$, así ψ es homomorfismo. Finalmente, si $\psi(Ng) = \varphi(g) = 1$ entonces $g \in N$, luego $Ng = N$ y Ker $\psi = 1$, luego por (3.17) ii), ψ es inyectiva. Adfemás, como φ es suprayectiva, así lo es ψ.

(3.34) **Corolario**: Salvo isomorfismo, los únicos grupos cíclicos son $\mathbf{Z}$ y $\mathbf{Z}/n\mathbf{Z}$ para enteros positivos n.

Demostración: Sea $G =< g >$, se define: $\varphi : \mathbf{Z} \longrightarrow G$ por $\varphi(n) = g^n$. Es homomorfismo de grupos pues $\varphi(n+m) = g^{n+m} = g^n.g^m$. Evidentemente es suprayectivo.

Por el teorema anterior, $G \cong \mathbf{Z}/\text{Ker}\varphi$, ahora bien, Ker φ es cíclico, luego Ker $\varphi =< n >$. Como $< n >=< -n >$, podemos tomar $n > 0$ si Ker φ no es trivial.

(3.35) **Nota**: a) Sea $N \trianglelefteq G$, entonces cada subgrupo de G/N es de la forma T/N para algún subgrupo T de G con $N \leq T$.

En efecto, si $\tilde{T}$ es un subgrupo de G/N y $\pi : G \longrightarrow G/N$ el epimorfismo canónico, $N \leq \pi^{-1}(\tilde{T}) = T \leq G$, así $\tilde{T} = \{aN | a \in T\} = T/N$.

b) Por otra parte, si $U \leq G$ entonces $\pi(U) \leq G/N$ y $\pi(U) = UN/N$ (notar que $UN \leq G$ pues $N \trianglelefteq G$).

Los grupos Simétrico y Alternado .

Recordemos que dado un conjunto Ω, el grupo de todas las aplicaciones biyectivas de Ω en sí (permutaciones de Ω) se denota $\Sigma(\Omega)$ ó Σ_Ω y se llama grupo simétrico sobre Ω y queda definido por $|\Omega|$. En lo que sigue Ω se considerará finito con n elementos $\Omega = \{1, 2 \ldots, n\}$ y al grupo simétrico correspondiente lo denotaremos por Σ_n que se llamará grupo simétrico de grado n.

(3.36) **Definición** : Se llama **ciclo** de longitud k a toda permutación de la forma:
$c = \begin{pmatrix} i_1\, i_2\, \ldots\, i_k\, i_{k+1}\, \ldots i_n \\ i_2\, i_3\, \ldots\, i_1\, i_{k+1}\, \ldots i_n \end{pmatrix}$,
que se expresará abreviadamente en lo sucesivo $c = (i_1, i_2, \ldots, i_k)$. Si $k = 2$ el ciclo se dice **trasposición**.

Ejercicios:1) Dos ciclos que operan (no trivialmente) sobre partes disjuntas de Ω son permutables.

2) Si c es un ciclo de longitud k, entonces $o(c) = k$.

3) Probar que si c y c' son ciclos disjuntos, se tiene:$< c > \cap < c' >= 1$.

Si α es una permutación, a veces denotaremos también la permutación α como la la n-tupla imagen $(\alpha(1), \ldots, \alpha(n))$.

(3.37) **Teorema** : Toda permutación $\alpha \neq 1$ se puede expresar como producto de ciclos disjuntos.

Demostración: Como $\alpha \neq 1$, sea i_1 un elemento de Ω tal que $\alpha(i_1) = i_2 \neq i_1$. Tomar $\alpha(i_2) = i_3$, ... En $i_1, i_2, \ldots, i_{r+1}, \ldots$, sea i_{r+1} el primer elemento que se repite. Entonces

$i_{r+1} = i_1$ pues si $i_{r+1} = i_3$ por ejemplo, $\alpha(i_2) = i_3 = \alpha(i_r) = i_{r+1}$, lo que implica $i_r = i_2$, contradicción.

Si $r = n$ o si los elementos de $\Omega - \{i_1, \ldots, i_r\}$ quedan fijos por α, se concluye que α es un ciclo. En otro caso, existe $j_1 \in \Omega - \{i_1, \ldots, i_r\}$ tal que $\alpha(j_1) = j_2 \neq j_1$ y volveríamos a hacer el proceso anterior. Así, como Ω es finito se llega a que $\alpha = c_1.c_2.\ldots.c_h$, donde los c_i son ciclos disjuntos.

(3.38) **Teorema**: La descomposición de una permutación $\alpha \neq 1$ como producto de ciclos disjuntos es única salvo el orden.

Demostración: Paso 1. Si c y c son ciclos que mueven la cifra i de Ω, tales que $c^t(i) = {c'}^t(i)\ \forall t$, entonces $c = c'$, pues si c es un ciclo de longitud r que mueve la cifra i, entonces $c = (i, c(i), \ldots, c^{r-1}(i))$

Paso 2. Si la permutación α no fija una letra i, y admite una descomposición como producto de ciclos disjuntos $\alpha = c_1.c_2 \ldots.c_r$ existe un único ciclo c_s tal que $\alpha^t(i) = c_s^t(i), \forall t \in \mathbf{N}$.

Paso 3. Conclusión. Si α tuviera dos descomposiciones como producto de ciclos disjuntos $\alpha = c_1.c_2.\ldots.c_h = \pi_1.\pi_2.\ldots.\pi_{h'}$ e i es una letra no fijada por c_1, entonces por el paso 2, $\alpha^t(i) = c_1^t(i), \forall t$. Otra aplicación del paso 2 nos permite afirmar que existe un ciclo de la segunda descomposición, que podemos suponer π_1 tal que $\alpha^t(i) = \pi_1^t(i), \forall t$, así que por el paso 1 se tiene que $c_1 = \pi_1$, y reiterando este proceso obtenemos la conclusión.

(3.39) **Nota**. Si $\alpha = c_1.c_2.\ldots.c_h$ y $l_1, l_2, \ldots, l_h$ son las longitudes respectivas de los ciclos, entonces $o(\alpha) = \text{m.c.m}(l_1, l_2, \ldots, l_h) = m$:

En efecto, $\alpha^m = (c_1 c_2 \ldots c_h)^m = c_1^m c_2^m \ldots c_h^m = 1$ y si $n = o(\alpha)$, entonces $n|m$, pero $\alpha^n = 1 = c_1^n c_2^n \ldots c_h^n = 1$, luego $c_i^n = 1, \forall i$ y así $l_i|n$ luego $m|n$ y por tanto m y n coinciden.

(3.40) **Corolario**: Toda permutación es producto de transposiciones.

Demostración: Observar que $(i_1, \ldots, i_r) = (i_r, i_{r-1})(i_{r-1}, i_{r-2}) \ldots (i_2, i_1)$.

(3.41) **Definición**: La **paridad** de una permutación π de Σ_n es la paridad de $n - c(\pi)$, donde $c(\pi)$ es el número de ciclos en la estructura de ciclos de π, incluyendo los ciclos de longitud 1. La **signatura** de π es $(-1)^{n-c(\pi)}$.

Notas: Una transposición es de clase impar y de signatura -1.

1_Ω tiene clase par y signatura 1.

(3.42) **Teorema**: Si τ es una trasposición y α una permutación, entonces $c(\alpha\tau) = c(\alpha) \pm 1$.

Demostración: Paso 1: Si $\pi = (i_1, \ldots, i_k, \ldots, i_t, \ldots, i_r)$ y $\tau = (i_k, i_t)$, entonces

$$\pi\tau = (i_1, \ldots, i_{k-1}, i_t, \ldots, i_r)(i_k, i_{k+1}, \ldots, i_{t-1}).$$

Paso 2: Si $\pi_1 = (i_1, \ldots, i_k, \ldots, i_{r_1})$, $\pi_2 = (j_1, \ldots, j_t, \ldots, j_{r_2})$ son dos ciclos disjuntos y $\tau = (i_k, j_t)$, entonces,

$$\pi_1\pi_2\tau = (i_1, \ldots, i_{k-1}, j_t, j_{t+1}, \ldots, j_{r_2}, j_1, \ldots, j_{t-1}, i_k, \ldots, i_{r_1}).$$

Paso 3: Conclusión. Por el teorema (3.38), la permutación α se descompone de modo único (salvo el orden), como producto de ciclos disjuntos. Si τ es la transposición (i, j), entonces existen uno o dos ciclos disjuntos de la descomposición de α en que intervengan las cifras i, j. Si existe un sólo ciclo π que las contiene a ambas, entonces $\alpha = \alpha'.\pi$, con α' el producto de los restantes ciclos y $c(\alpha\tau) = c(\alpha'(\pi\tau)) = c(\alpha) + 1$.

Si existen dos ciclos π_1 y π_2, cada uno de los cuales contiene una de las letras, entonces $\alpha = \alpha''.(\pi_1.\pi_2)$, y por el paso 2, $c(\alpha\tau) = c(\alpha''.(\pi_1.\pi_2.\tau)) = c(\alpha) - 1$.

(3.43) **Teorema**: Sea τ una transposición y α una permutación cualquiera, entonces sig $(\alpha\tau) = -$ sig (α).

Demostración: Por el teorema anterior,

sig $(\alpha\tau) = (-1)^{(n-(c(\alpha)\pm 1)} = (-1)^{(n-c(\alpha))}.(-1)^{\mp 1} =-$ sig (α).

(3.44) **Teorema**: Si α es un producto de r trasposiciones, entonces sig $(\alpha) = (-1)^r$

Demostración: Sea $\alpha = t_1.t_2 \ldots .t_r$, entonces

$$\text{sig}\,\alpha = -\text{sig}\,(t_1.t_2 \ldots .t_{r-1}) = (-1)(-1)\text{sig}\,(t_1.t_2 \ldots .t_{r-2}) = (-1)^r.$$

(3.45) **Corolario**: sig $(\alpha_1.\alpha_2) = \text{sig}(\alpha_1)\text{sig}(\alpha_2)$.

Así, si consideramos Σ_n y el grupo cíclico , $\{1, -1\}$, entonces φ: $\Sigma_n \longrightarrow \{1, -1\}$ dada por $\varphi(\alpha) = \text{sig}(\alpha)$ es para $n \geq 2$ un homomorfismo de grupos suprayectivo y a Ker φ le llamaremos **grupo alternado de grado** n, y será denotado A_n. Notar que por tanto $\Sigma_n/A_n \cong \{1, -1\} \cong C_2$. En consecuencia $|\Sigma_n/A_n| = 2$. .

Lección 4. Anillos.

Primeras propiedades.

(4.1) **Definición**: Un anillo es un conjunto $\mathbf{R}$ no vacío dotado de dos operaciones internas + y **.** de forma que:

i) $(\mathbf{R}, +)$ es un grupo abeliano; ii) (.) es asociativa y es distributiva respecto de (+) ;

iii) Si además **.** es conmutativa el anillo se dice **conmutativo**.

iv) Si en $\mathbf{R}$ hay un elemento neutro respecto de . el anillo se dice **unitario**, con **unidad** ó con **identidad** que denotaremos por 1 ó $1_{\mathbf{R}}$.

Un elemento $a \in \mathbf{R}$, anillo unitario, se dice una **unidad** ó **elemento inversible**, de $\mathbf{R}$ si existe $b \in \mathbf{R}$ tal que $a.b = b.a = 1_{\mathbf{R}}$

Ejercicio: El conjunto de la unidades en un anillo unitario forma un grupo con la operación (**.**).

(4.2) **Definición** : Un elemento $a \neq 0_{\mathbf{R}}, a \in \mathbf{R}$ se dice **divisor de cero** a izquierda (respectivamente a derecha) si existe $b \neq 0_{\mathbf{R}}, b \in \mathbf{R}$ tal que $a.b = 0_{\mathbf{R}}$ (respectivamente $b.a = 0_{\mathbf{R}}$).

(4.3) **Definición** : Un anillo $\mathbf{R}$ es un **dominio de integridad** si es un anillo conmutativo con identidad $1_{\mathbf{R}} \neq 0_{\mathbf{R}}$ y sin divisores de cero. Un **anillo de división** es un anillo con identidad $1_{\mathbf{R}} \neq 0_{\mathbf{R}}$ en el que cada elemento distinto de cero es una unidad. Un **cuerpo** es un anillo de división conmutativo.

(4.4) **Proposición**: Sea $\mathbf{R}$ un anillo, entonces se verifican:

i) $0.a = a.0 = 0, \forall a \in \mathbf{R}$; ii) $(-a).b = a.(-b) = -(a.b), \forall a, b \in \mathbf{R}$; iii)$(-a).(-b) = a.b, \forall a, b \in \mathbf{R}$.

(4.5) **Definición 5**: **Un homomorfismo de anillos** es una aplicación f:$\mathbf{R}_1 \longrightarrow \mathbf{R}_2$ verificando:

i) $f(a + b) = f(a) + f(b)$;

ii) $f(a.b) = f(a).f(b)$, $\forall a, b \in \mathbf{R}_1$.

(4.6) **NOTA**: Si R_1 y R_2 son anillos con identidad y $f : R_1 \longrightarrow R_2$ es un homomorfismo de anillos, no es cierto, en general que $f(1_{\mathbf{R}_1}) = 1_{\mathbf{R}_2}$, basta considerar el cuerpo $\mathbf{R}$

de los números reales y el homomorfismo $f : \mathbf{R} \longrightarrow M(2, \mathbf{R})$ que asocia a un elemento a de $\mathbf{R}$ la matriz 2x2 sobre $\mathbf{R}$ cuyo única entrada no nula es la (1,1) que es igual a a.

Sin embargo dicha propiedad es cierta bajo algunas condiciones:

(4.7) **Proposición** : i) Sea $f : \mathbf{R}_1 \longrightarrow \mathbf{R}_2$ un epimorfismo de anillos con identidad, entonces $f(1_{R_1}) = 1_{R_2}$.

ii) Si $f : \mathbf{R}_1 \longrightarrow \mathbf{R}_2$ es un homomorfismo de anillos con identidad y u es una unidad de $\mathbf{R}_1$ tal que $f(u)$ es unidad de $\mathbf{R}_2$, entonces $f(1_{\mathbf{R}_1}) = 1_{\mathbf{R}_2}$ y $f(u^{-1}) = f(u)^{-1}$.

Demostración: i) Dado $a_2 \in \mathbf{R}_2$ es de la forma $f(a_1)$ para algún $a_1 \in \mathbf{R}_1$, tendremos $f(1_{\mathbf{R}1}).a_2 = f(1_{\mathbf{R}1}).f(a_1) = f(1_{\mathbf{R}1}.a_1) = f(a_1) = a_2$, análogamente se obtiene que $a_2.f(1_{\mathbf{R}1}) = f(a_1).f(1_{\mathbf{R}1}) = f(a_1) = a_2$, y así $f(1_{\mathbf{R}1}) = 1_{\mathbf{R}_2}$.

ii) $u = u.1_{\mathbf{R}_1} = 1_{\mathbf{R}_1}.u$ y así $f(u) = f(u)f(1_{\mathbf{R}_1}) = f(1_{\mathbf{R}_1}).f(u)$ y como $f(u)$ es unidad, es $f(1_{\mathbf{R}_1}) = 1_{\mathbf{R}_2}$. Ahora $f(u).f(u^{-1}) = f(u^{-1}).f(u) = f(1_{\mathbf{R}_1}) = 1_{\mathbf{R}_2}$, por tanto $f(u^{-1}) = f(u)^{-1}$.

(4.8) **Definición** : Sea **(R, +, .)** un anillo y $\emptyset \neq \mathbf{S} \subseteq \mathbf{R}$, se dice que **S** es un **subanillo** de **R** si dados $a, b \in \mathbf{S}$ cualesquiera, se tiene que tanto $a + b$ como $a.b$ son elementos de **S** y **S** con ambas operaciones es un anillo.

Un subanillo **I** de **R** se dice **ideal** de **R** si dados $a \in \mathbf{R}$ y $x \in \mathbf{I}$, se tiene que $a.x \in \mathbf{I}$ y $x.a \in \mathbf{I}$.

(4.9) **Ejemplos**: i) $Z(\mathbf{R}) = \{a \in \mathbf{R} | a.x = x.a, \forall x \in \mathbf{R}\}$ es subanillo de **R** pero no necesariamente ideal.

ii) Si $f : \mathbf{R}_1 \longrightarrow \mathbf{R}_2$ un es un homomorfismo de anillos Ker $f = \{a \in \mathbf{R}_1 | f(a) = 0_{\mathbf{R}_2}\}$ es un ideal de $\mathbf{R}_1$.

iii) En $(\mathbf{Z}, +, .)$, $n.\mathbf{Z}$ es un ideal de **Z**.

(4.10) **Proposición**: Sea **R** un anillo y $\emptyset \neq \mathbf{S} \subseteq \mathbf{R}$, entonces **S** es un subanillo de **R** si y sólo si:

i) $a, b \in \mathbf{S}$ implica $a - b \in \mathbf{S}$ y

ii) $a, b \in \mathbf{S}$ implica $a.b \in \mathbf{S}$.

(4.11) **Proposición** : Un subconjunto $\emptyset \neq \mathbf{I} \subseteq \mathbf{R}$ es un ideal de **R** si sólo si

i) $a, b \in \mathbf{I}$ implica $a - b \in \mathbf{I}$ y

ii) $a \in \mathbf{I}, x \in \mathbf{R}$ implica $a.x \in \mathbf{I}$ y $x.a \in \mathbf{I}$.

(4.12) **Definición** : Sea $X \subseteq \mathbf{R}$, denotaremos con (X) a la intersección de todos los ideales de $\mathbf{R}$ que contienen a X. Dicha intersección es un ideal de $\mathbf{R}$ que se dice generado por X. Si $\mathbf{R}$ es un anillo conmutativo y con identidad y $a \in \mathbf{R}$, el ideal generado por $\{a\}$ es $(a) = \{x.a | x \in \mathbf{R}\} = \{a.y | y \in \mathbf{R}\}$, que se dice un **ideal principal**.

Si $\mathbf{I}$ es un ideal del anillo $\mathbf{R}$, considerar el conjunto $\mathbf{R/I}$ definido como $\{x + \mathbf{I} | x \in \mathbf{R}\}$, donde $x + \mathbf{I} = \{x + m | m \in \mathbf{I}\}$, llamaremos al conjunto anterior **cociente** de $\mathbf{R}$ módulo $\mathbf{I}$ o por $\mathbf{I}$

(4.13) **Teorema**: Sea $\mathbf{I}$ un ideal del anillo $\mathbf{R}$. Entonces el conjunto cociente $\mathbf{R/I}$ con las operaciones internas:

$(x + \mathbf{I}) + (y + \mathbf{I}) = (x + y) + \mathbf{I}$;

$(x + \mathbf{I}).(y + \mathbf{I}) = (x.y) + \mathbf{I}$

es un anillo que será llamado **anillo cociente** de $\mathbf{R}$ módulo $\mathbf{I}$ o por $\mathbf{I}$.

Demostración: Veamos que el producto está bien definido: Suponer $x + \mathbf{I} = x' + \mathbf{I}$ y $y + \mathbf{I} = y' + \mathbf{I}$,así $x' = x + a, a \in \mathbf{I}$, $y' = y + b, b \in \mathbf{I}$, en consecuencia $x'y' = (x + a).(y + b) = x.y + x.b + a.y + a.b$ por tanto, $x'y' + \mathbf{I} = xy + \mathbf{I}$. Es rutinario comprobar el resto de las condiciones para que $\mathbf{R/I}$ pase a ser anillo.

Observación: Como $n\mathbf{Z}$ es un ideal de $\mathbf{Z}$, se sigue que $\mathbf{Z}/n\mathbf{Z}$ es un anillo, que además es conmutativo y con identidad. Además,

(4.14) **Proposición** : El grupo de las unidades $\mathbf{U}_n$ de $\mathbf{Z}/n\mathbf{Z}$ es el $\{r + n\mathbf{Z} | 1 \leq r \leq n, (r, n) = 1\}$. Por tanto el orden de este grupo es $\varphi(n)$.

Demostración: Si $(r, n) = 1$, existen, por la identidad de Bezout, enteros α, β tales que $\alpha.r + \beta.n = 1$, por lo tanto, $1 + n\mathbf{Z} = (\alpha.r + n\mathbf{Z}) = (\alpha + n\mathbf{Z}).(r + n\mathbf{Z})$. Recíprocamente, si $r + n\mathbf{Z}$ es una unidad de $\mathbf{Z}/n\mathbf{Z}$, existe un entero s tal que $(r + n\mathbf{Z})(s + n\mathbf{Z}) = 1 + n\mathbf{Z}$, luego $r.s - 1 = \dot{n}$, por tanto, si $d = (r, n)$ debe ser $d = 1$.

(4.15) **Teorema** : Sea $f : \mathbf{R} \longrightarrow \mathbf{R}'$ un homomorfismo de anillos, entonces:

i) $f(\mathbf{R})$ es un subanillo de $\mathbf{R}'$.

ii) Existe un isomorfismo de anillos $\tilde{f} : \mathbf{R}/\mathrm{Ker}\, f \longrightarrow f(\mathbf{R})$ dado por : $\tilde{f}(a+\mathrm{Ker}(f)) = f(a)$

Demostración: i) Es sencillo; ii) Se comprueba que la aplicación anterior es un homomorfismo de anillos inyectivo y suprayectivo.

Ejercicios: i) Sean $\mathbf{I}$ y $\mathbf{J}$ ideales de un anillo $\mathbf{R}$, con $\mathbf{I} \subseteq \mathbf{J}$, entonces $\mathbf{J}/\mathbf{I}$ es un ideal de $\mathbf{R}/\mathbf{I}$ y $\mathbf{R}/\mathbf{I}/\mathbf{J}/\mathbf{I}$ es isomorfo a $\mathbf{R}/\mathbf{J}$.

ii) Sea $\mathbf{B}$ un subanillo de $\mathbf{R}$ e $\mathbf{I}$ ideal de $\mathbf{R}$. Se define

$$\mathbf{B} + \mathbf{I} := \{a + b | a \in \mathbf{B}, b \in \mathbf{I}\},$$

entonces $\mathbf{B} + \mathbf{I}$ es un subanillo de $\mathbf{R}$ y $\mathbf{B} + \mathbf{I}/\mathbf{I}$ es un anillo isomorfo a $\mathbf{B}/\mathbf{B} \cap \mathbf{I}$.

(4.16) **Definición** : Un ideal $\mathbf{M}$ de un anillo $\mathbf{R}$ se dice un ideal maximal de $\mathbf{R}$ si $\mathbf{M} \neq \mathbf{R}$ y cualquiera que sea el ideal $\mathbf{N}$ de $\mathbf{R}$ tal que $\mathbf{M} \subseteq \mathbf{N} \subseteq \mathbf{R}$ se tiene que $\mathbf{N} = \mathbf{M}$ ó $\mathbf{N} = \mathbf{R}$.

(4.17) **Teorema** : En un anillo $\mathbf{R} \neq \mathbf{0}$ con identidad, existen siempre ideales maximales. De hecho cada ideal $\mathbf{I}$ de $\mathbf{R}$, $\mathbf{I} \neq \mathbf{R}$, está contenido en un ideal maximal.

Demostración: Sea $\mathbf{I}$ ideal de $\mathbf{R}$ con $\mathbf{I} \neq \mathbf{R}$. Consideremos $\mathcal{S} = \{\mathbf{J}\,\text{ideal de}\,\mathbf{R} | \mathbf{J} \neq \mathbf{R}, \mathbf{I} \subseteq \mathbf{J}\}$, $S \neq \emptyset$ pues $\mathbf{I} \in S$. Se ordena $\mathcal{S}$ parcialmente por inclusión: $\mathbf{J}_1 \leq \mathbf{J}_2$ si $\mathbf{J}_1 \subseteq \mathbf{J}_2$.

Para aplicar el lema de Zorn debemos probar que cada cadena de $\mathcal{S}$ tiene cota superior en $\mathcal{S}$. Sea $\{\mathbf{J}_i\}_{i\in I}$ una cadena de $\mathcal{S}$ y $\mathbf{J} = \cup_{i\in I}\mathbf{J}_i$, entonces $\mathbf{J}$ es ideal de $\mathbf{R}$. En efecto, si $a, b \in \mathbf{J}$, existen i y j tales que $a \in \mathbf{J}_i, b \in \mathbf{J}_j$ y como $\mathbf{J}_i \subseteq \mathbf{J}_j$ ó $\mathbf{J}_j \subseteq \mathbf{J}_i$, a y b se localizarán a la vez en uno de ellos, sea $\mathbf{J}_i$. Como $\mathbf{J}_i$ es ideal de $\mathbf{R}$, $a - b \in \mathbf{J}_i \subseteq \mathbf{J}$. Además $a.r \in \mathbf{J}_i \subseteq \mathbf{J}, \forall r \in \mathbf{R}$ y análogamente $r.a \in \mathbf{J}_i \subseteq \mathbf{J}, \forall r \in \mathbf{R}$. Como $\mathbf{J}_i \neq \mathbf{R}$, se tiene que $1_{\mathbf{R}} \notin \mathbf{J}_i, \forall i$, luego $1_{\mathbf{R}} \notin \mathbf{J}$ y así $\mathbf{J} \neq \mathbf{R}$. Además $\mathbf{I} \subseteq \mathbf{J}$, luego $\mathbf{J} \in \mathcal{S}$ es cota superior de la cadena. Por el lema de Zorn existe al menos un elemento maximal en $\mathcal{S}$, que será un ideal maximal de $\mathbf{R}$ y contiene a $\mathbf{I}$.

(4.18) **Definición**: Un ideal $\mathbf{P}$ de $\mathbf{R}$ se dice primo si $\mathbf{P} \neq \mathbf{R}$ y cualesquiera que sean $a, b \in \mathbf{R}$, si $a.b \in \mathbf{P}$ es $a \in \mathbf{P}$ ó $b \in \mathbf{P}$.

(4.19) **Teorema**: Sea $\mathbf{R}$ anillo conmutativo y con unidad $1 \neq 0$ y $\mathbf{P}$ un ideal de $\mathbf{R}$. Entonces $\mathbf{P}$ es primo si y sólo si $\mathbf{R}/\mathbf{P}$ es un dominio de integridad.

Demostración: Sea $\mathbf{P}$ primo, entonces $\mathbf{R}/\mathbf{P}$ tiene identidad $1+\mathbf{P}$ y $1+\mathbf{P} \neq 0+\mathbf{P}$ pues $1 \notin \mathbf{P}$ ya que $\mathbf{P} \neq \mathbf{R}$, además $\mathbf{R}/\mathbf{P}$ es anillo conmutativo y no tiene divisores de cero, pues si $(a+\mathbf{P}).(b+\mathbf{P}) = \mathbf{P}$, entonces $a.b \in \mathbf{P}$ lo que implicaría $a \in \mathbf{P}$ ó $b \in \mathbf{P}$ luego $a+\mathbf{P} = \mathbf{P}$ ó $b+\mathbf{P} = \mathbf{P}$.

Recíprocamente, si $\mathbf{R}/\mathbf{P}$ es un dominio de integridad $1+\mathbf{P} \neq 0+\mathbf{P}$ así $\mathbf{P} \neq \mathbf{R}$. Además si $a.b \in \mathbf{P}$, sería $a.b+\mathbf{P} = \mathbf{P}$, luego $a+\mathbf{P} = \mathbf{P}$ ó $b+\mathbf{P} = \mathbf{P}$ y así $a \in \mathbf{P}$ ó $b \in \mathbf{P}$.

(4.20) **Teorema** : Sea $\mathbf{R}$ un anillo conmutativo con unidad $1 \neq 0$ y $\mathbf{M}$ un ideal de $\mathbf{R}$. Entonces $\mathbf{M}$ es un ideal maximal de $\mathbf{R}$ si y sólo si $\mathbf{R}/\mathbf{M}$ es un cuerpo.

Demostración:$\Rightarrow$):Sea $\mathbf{M}$ un ideal maximal, entonces $1+\mathbf{M} \neq \mathbf{0}+\mathbf{M}$, pues $1 \notin \mathbf{M}$. Además, si $a+\mathbf{M} \neq \mathbf{M}$, es decir, si $a \notin \mathbf{M}$ se tiene $\mathbf{R} = \mathbf{M}+(a)$, luego existen $m \in \mathbf{M}$ y $r \in \mathbf{R}$ tales que $1 = m + r.a$ y así $1+\mathbf{M} = r.a+\mathbf{M} = (r+\mathbf{M})(a+\mathbf{M})$.

$\Leftarrow$): Si $\mathbf{R}/\mathbf{M}$ es cuerpo, $1+\mathbf{M} \neq \mathbf{0}+\mathbf{M}$ luego $\mathbf{M} \neq \mathbf{R}$. Si $\mathbf{M} \subseteq \mathbf{N} \subseteq \mathbf{R}$ con $\mathbf{N}$ ideal de $\mathbf{R}$ y suponemos $\mathbf{M} \neq \mathbf{N}$, existe $a \in \mathbf{N}-\mathbf{M}$, luego $a+\mathbf{M}$ es una unidad, por tanto existe $b+\mathbf{M}$ tal que $(a+\mathbf{M}).(b+\mathbf{M}) = 1+\mathbf{M}$, de donde $a.b-1 \in \mathbf{M} \subseteq \mathbf{N}$, luego $1 \in \mathbf{N}$ y de ahí se sigue $\mathbf{N} = \mathbf{R}$.

Notar que en esta implicación sólo se utiliza que $\mathbf{R}/M$ es un anillo de división.

(4.21) **Definición** : Sea $\mathbf{F}$ un cuerpo. Consideremos el elemento $1_{\mathbf{F}}$ en el grupo aditivo $(\mathbf{F},+)$. Llamaremos **característica** de $\mathbf{F}$ al orden de $1_{\mathbf{F}}$. Si el orden de $1_{\mathbf{F}}$ es infinito, se dice que la característica de $\mathbf{F}$ es 0..

Ejemplos: 1) Son cuerpos de característica cero: $\mathbf{Q}, \mathbf{R}, \mathbf{C}$.

2) Si p es un primo, $\mathbf{Z}/p\mathbf{Z}$ es un cuerpo de característica p.

(4.22) **Proposición** : La característica de un cuerpo $\mathbf{F}$ es cero ó un primo p. Además, en el segundo caso $pa = 0$, $\forall a \in \mathbf{F}$.

Demostración: Sea $0 \neq n = \operatorname{car}\mathbf{F}$. Si n no es primo, entonces $n = p_1.p_2$, con $1 < p_i < n$, lo que conduce a $p_1.1_{\mathbf{F}} = 0$ ó $p_2.1_{\mathbf{F}} = 0$, contradicción.

Si la característica de $\mathbf{F} = p$, entonces $p.a = a+a+\ldots^{p}+a = a(1_{\mathbf{F}}+\ldots^{p}+1_{\mathbf{F}}) = a.(p.1_{\mathbf{F}}) = 0$.

En lo que sigue de la lección $\mathbf{R}$ será un dominio de integridad.

(4.23) **Definiciones** : i) Diremos que $a \neq 0$ divide a b y lo expresaremos $a|b$ si existe $c \in \mathbf{R}$ tal que $b = a.c$.

ii) a es **asociado** a b y lo expresaremos $a \sim b$ si simúltaneamente $a|b$ y$b|a$.

Las unidades de $\mathbf{R}$ son los asociados con $1_{\mathbf{R}}$.

iii) a es **irreducible** si :$a \neq 0, a \not\sim 1$ y si $b|a$ entonces $b \sim a$ ó $b \sim 1$.

iv) a es **primo** si: $a \neq 0, a \not\sim 1$ y si $a|b.c$ entonces $a|b$ ó $a|c$.

v) a es un **máximo común divisor** de $\emptyset \neq X \subseteq \mathbf{R}$, denotado $a = \text{m.c.d}(X)$ si: $a|x\,, \forall x \in X$ y si $c|x\,, \forall x \in X$ entonces $c|a$.

vi) a es un **mínimo común múltiplo** de $\emptyset \neq X \subseteq \mathbf{R}$, denotado $a = \text{m.c.m}(X)$ si: $x|a\,, \forall x \in X$ y si $x|c\,, \forall x \in X$ entonces $a|c$.

Son inmediatas las siguientes afirmaciones:

a) La divisibilidad es reflexiva y transitiva; b) La asociación es una relación de equivalencia

c) Si a y a' son dos m.c.d (X), $\emptyset \neq X \subseteq \mathbf{R}$ entonces $a \sim a'$.

d) $a|b \Leftrightarrow (b) \subseteq (a)$.

e) $a \sim b \Leftrightarrow (a) = (b)$.

f) u es unidad de $\mathbf{R} \Leftrightarrow (u) = \mathbf{R}$.

g) a primo $\Leftrightarrow (a)$ es un ideal primo no cero.

h) a es divisor propio de b (es decir $a|b$ pero $a \not\sim 1, a \not\sim b$) $\Leftrightarrow$(b)$\subset (a) \subset \mathbf{R}$.

i) $a \sim b \Leftrightarrow a = b.u$ con u unidad.

(4.24) **Proposición**: Sea $\mathbf{R}$ un D.I (abreviatura para dominio de integridad), entonces todo elemento primo es irreducible.

Demostración: Sea a primo, así $a \neq 0, a \not\sim 1$. Si $b|a$, entonces existe $c \in \mathbf{R}$ tal que $a = b.c$, por tanto $a|b.c \Rightarrow\ a|b$ *ó* $a|c$. Si $a|b \Rightarrow\ a \sim b$. Si se da lo segundo $b \sim 1$.

(4.25) **Proposición** : Sea $\mathbf{R}$ un D.I.P., es decir un dominio de integridad cuyos ideales son todos principales. Si $\mathbf{R}$ no es un cuerpo, son equivalentes:

i) a es irreducible .

ii) (a) es maximal .

iii) (a) es primo no cero .

iv) a es primo .

Demostración: Sabemos que ii) $\Rightarrow$ iii) (ya que 0 no es un ideal maximal salvo que el anillo sea un cuerpo); iii) $\Leftrightarrow$ iv) $\Rightarrow$ i). Así basta demostrar que i) $\Rightarrow$ ii).

Supongamos que a es un elemento irreducible, entonces $(a) \subset \mathbf{R}$. Suponer que (a) no es un ideal maximal, entonces existirá $\mathbf{I}$ ideal de $\mathbf{R}$, con $(a) \subset \mathbf{I} \subset \mathbf{R}$ y como $\mathbf{R}$ es un D.I.P., existirá $b \in \mathbf{R}$ tal que $\mathbf{I}= (b)$, luego $(a) \subset (b)$ y por tanto $b|a$ y como $(a) \neq (b)$, debe ser $b \sim 1$, luego $\mathbf{I}= (b) = \mathbf{R}$, contradicción.

En la lección primera, habíamos probado la existencia de máximo común divisor y la validez de la identidad de Bezout en el anillo $\mathbf{Z}$ de los números enteros. En la siguiente proposición obtendremos esos resultados para un dominio de ideales principales arbitrario.

(4.26) **Proposición** : Si $\mathbf{R}$ es un D.I.P. dos elementos $a, b \in \mathbf{R}$ tienen un máximo común divisor y un mínimo común múltiplo. En un D.I.P. vale siempre la identidad de Bezout.

Demostración: Considerar $(a) + (b)$ que es un ideal de $\mathbf{R}$ así, debe existir $d \in \mathbf{R}$ tal que $(a) + (b) = (d)$, entonces $(a), (b) \subseteq (d)$ y por tanto $d|a$ y $d|b$ y si $c|a$ y $c|b$ entonces $(a) \subseteq (c)$ y $(b) \subseteq (c)$ y así $(d) = (a) + (b) \subseteq (c)$ luego $c|d$.

En consecuencia d es un máximo común divisor de $\{a, b\}$.

Análogamente, considerar $(a) \cap (b) = (m)$, entonces $(m) \subseteq (a)$ y $(m) \subseteq (b)$, luego m es un múltiplo común de a y b y si m' es otro múltiplo común de a y de b, entonces $(m') \subseteq (a)$ y $(m') \subseteq (b)$ luego $(m') \subseteq (a) \cap (b) = (m)$, y así $m|m'$. Por tanto m es un mínimo común múltiplo de $\{a, b\}$.

Matrices sobre un anillo.

(4.27) **Definición**: Una **matriz** $n \times m$ con **entradas ó elementos** en un anillo R es un cuadro de elemetos de R dispuestos en n filas y m columnas

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nm} \end{pmatrix}.$$

Una matriz $1 \times m$ se dice una matriz **fila** y una $n \times 1$ se dice una matriz **columna**. Una matriz $n \times n$ se dice matriz **cuadrada** . La **diagonal principal** de la matriz $n \times n$ (a_{ij}) es la sucesión de los elementos $a_{ii}, i \in [1, n]$. Una matriz **triangular superior** es la que tiene nulos los elementos por debajo de la diagonal principal (i.e $a_{ij} = 0$, si $i > j$). Una matriz es **triangular inferior** si tiene nulos los elemento por encima de la diagonal principal. Una matriz **diagonal** tiene nulos todos los elementos que no están en la diagonal, a veces la escribiremos $A = \text{diag}\,[a_{11}, \ldots, a_{nn}]$. Si se señalan índices de filas $i_1, \ldots, i_p$ y otros de columnas $j_1, \ldots, j_q$, entonces la submatriz cuyas filas son $a_{ij_1}, a_{ij_2}, \ldots, a_{ij_q}$ con $i \in \{i_1, \ldots, i_p\}$ se le llama **submatriz** de A. Una submatriz cuyos índices de filas y de columnas son consecutivos se dice un **bloque** o **caja** de A. Los bloques más notables son los bloques fila A_i o bloques columna A^i.

Llamaremos descomposición de A en bloques a una partición de A en bloques:

$$A = \begin{pmatrix} A_{11} & \ldots & A_{1k} \\ \vdots & \ldots & \vdots \\ A_{h1} & \ldots & A_{hk} \end{pmatrix}$$

de forma que bloques en una misma fila tengan el mismo numero de filas y bloques en una columna tengan el mismo número de columnas. Si en la descomposición anterior $h = k$ y todos los bloque $A_{ij}, i \neq j$ son ceros, se dice que A es suma diagonal de submatrices A_{ii} y se escribe $A = [A_{11}, \ldots, A_{hh}]$.

(4.28) **Definición**: Si A es $n \times m$, la matriz $m \times n$, cuyo elemento (j, i) es el (i, j) de A se dice **traspuesta** de A y se escribe A^t ó A'. Es claro que $(\mathrm{A}^t)^t = A$. Si $A = A^t$, la matriz A se dice **simétrica** y en ella $a_{ij} = a_{ji}$. Si para todo i, j es $a_{ij} = -a_{ji}$ la matriz A se dice **antisimétrica**.

(4.29) **Proposición**: Sea R un anillo. El conjunto $\text{Mat}\,(n \times m, R)$ de las matrices $n \times m$ con entradas en R es un grupo abeliano con la operación suma definida por:

Si $A = (a_{ij})$ y $B = (b_{ij})$, entonces $A + B = (a_{ij} + b_{ij})$.

Demostración: Fácil.

(4.30) **Definición**: Sea R un anillo. Si $A = (a_{ij})$ es una matriz $n \times m$ y $B = (b_{ij})$ es una matriz $m \times k$ ambas sobre R, se llama matriz **producto** de A por B a la matriz $n \times k$ sobre R cuyo elemento (i, j) es $a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{im}b_{mj}$

(4.31) **Teorema**: Sea R un anillo. El conjunto $\mathrm{Mat}\,(n \times n, R)$ (denotado también como $\mathrm{M}\,(n, R)$) de las matrices $n \times n$ con entradas en R, es un anillo con la suma y producto definidos anteriormente. Si R tiene unidad, la unidad de $\mathrm{Mat}\,(n \times n, R)$ es la matriz **identidad** $I_n = \mathrm{diag}\,[1, \ldots, 1]$. Dicho anillo no es en general conmutativo.

Demostración: Por (4.29) $\mathrm{M}\,(n, R)$ es un grupo abeliano.

El producto es una operación binaria interna. Es asociativa:

Sean $A = (a_{ij}), B = (b_{ij}), C = (c_{ij})$, denotamos por $P = (p_{ij}) = AB$, $Q = (q_{ij}) = BC$, entonces $p_{ij} = \sum_{r=1}^{n} a_{ir} b_{rj}$ y $q_{ij} = \sum_{r=1}^{n} b_{ir} c_{rj}$. Ahora, el elemento (h, k) de PC es

$$\sum_{r=1}^{n} p_{hr} c_{rk} = \sum_{r=1}^{n} (\sum_{t=1}^{n} a_{ht} b_{tr}) c_{rk} = \sum_{r=1}^{n} (\sum_{t=1}^{n} a_{ht} b_{tr} c_{rk}) = \sum_{t=1}^{n} (\sum_{r=1}^{n} a_{ht} b_{tr} c_{rk}) =$$

$$\sum_{t=1}^{n} a_{ht} (\sum_{r=1}^{n} b_{tr} c_{rk}) = \sum_{t=1}^{n} a_{ht} q_{tk}$$

que es el elemento (h, k) de AQ. Así que el producto es asociativo.

En cuanto a las leyes distributivas:

a) El elemento (i, j) de $(A + B)C$ es

$$\sum_{r=1}^{n} (a_{ir} + b_{ir}) c_{rj} = \sum_{r=1}^{n} a_{ir} c_{rj} + \sum_{r=1}^{n} b_{ir} c_{rj},$$

que es el elemento (i, j) de $AC + BC$. Luego $(A + B)C = AC + BC$.

b) El elemeto (i, j) de $A(B + C)$ es

$$\sum_{r=1}^{n} a_{ir} (b_{rj} + c_{rj}) = \sum_{r=1}^{n} a_{ir} b_{rj} + \sum_{r=1}^{n} a_{ir} c_{rj}$$

que es el elemento (i, j) de $AB + AC$.

Las afirmaciones restantes son evidentes.

Lección 5. Polinomios sobre un anillo.

(5.1) **Definición**: Sea $\mathbf{R}$ un anillo, un **polinomio** con **coeficientes** en $\mathbf{R}$ es una sucesión $(a_n), n \geq 0$ con $a_n \in \mathbf{R}$ y sólo un número finito de elementos a_n son distintos de cero. El polinomio $(a_o, a_1, \ldots, a_n, 0, \ldots, 0, \ldots)$ será denotado como $a_o + a_1x + \ldots + a_nx^n = \sum_{i=o}^{n} a_ix^i$ donde las potencias de un misterioso x se usan de un modo meramente formal para señalar las posiciones que ocupa cada a_i en la sucesión. Los elementos a_i de $\mathbf{R}$ son los **coeficientes** del polinomio $\sum_{i=o}^{n} a_ix^i$. a_o es el **coeficiente** ó **término constante** y el último miembro no nulo de la sucesión a_n, si existe, es el **coeficiente director** y n es el **grado** del polinomio. Por convenio, el grado del **polinomio cero** $(0, 0, \ldots, 0, \ldots)$ es $-\infty$.

Denotaremos con $\mathbf{R}[x]$ al conjunto de todos las polinomios con coeficientes en $\mathbf{R}$.

En $\mathbf{R}[x]$ definimos las operaciones:

Suma: $(a_n) + (b_n) = (c_n)$, donde $c_n = a_n + b_n$, es decir
$\sum_{i=0}^{n} a_ix^i + \sum_{j=0}^{m} b_jx^j = \sum_{i=0}^{\max(n,m)} (a_i + b_i)x^i,$

Producto: $(a_n).(b_n) = (c_k)$, donde $c_k = \sum_{i+j=k} a_i.b_j$ es decir

$\sum_{i=0}^{n} a_ix^i . \sum_{j=0}^{m} b_jx^j = \sum_{k\geq 0} \sum_{i+j=k} (a_i.b_j)x^k.$

$(\mathbf{R}[x], +)$ es evidentemente un grupo abeliano cuyo elemento neutro es el polinomio cero. Es fácil comprobar las leyes distributivas. Se comprueba que el producto es asociativo utilizando inducción y las leyes distributivas. Así $(\mathbf{R}[x], +, .)$ es un anillo, llamado **anillo de polinomios** sobre el anillo $\mathbf{R}$

(5.2) **Teorema**. i) Si $\mathbf{R}$ tiene identidad también la posee $\mathbf{R}[x]$. Si $\mathbf{R}$ es conmutativo también lo es $\mathbf{R}[x]$. Si $\mathbf{R}$ no posee divisores de cero lo mismo ocurre con $\mathbf{R}[x]$. Por tanto si $\mathbf{R}$ es un D.I., lo mismo ocurrirá con $\mathbf{R}[x]$. Pero nunca $\mathbf{R}[x]$ es un anillo de división.

ii) La aplicación : $\mathbf{R} \longrightarrow \mathbf{R}[x]$ dada por $a \longrightarrow (a, 0, 0, \ldots)$ es un monomorfismo de anillos.

Demostración: i) Si 1 es la identidad de $\mathbf{R}$ entonces $(1, 0, \ldots, 0, \ldots)$ es la identidad de $\mathbf{R}[x]$.

Si $\mathbf{R}$ es conmutativo también lo es $\mathbf{R}[x]$ pues $(a_n).(b_n) = (c_k)$ donde $c_k = \sum_{i+j=k} a_i.b_j = \sum_{i+j=k} b_j.a_i$. Así $(a_n).(b_n) = (b_n).(a_n)$.

Si $\mathbf{R}$ no tiene divisores de cero y $(a_o, a_1, \ldots)$, $(b_o, b_1, \ldots)$ son dos polinomios distintos del polinomio cero en $\mathbf{R}[x]$, siendo r el menor índice tal que $a_r \neq 0$ y s el menor índice tal que $b_s \neq 0$, entonces se tiene :

$$(a_o, a_1, \ldots).(b_o, b_1, \ldots) = (0, \ldots, 0, a_r.b_s, a_r.b_{s+1} + a_{r+1}.b_s, \ldots)$$

y si este producto fuera el polinomio cero, tendríamos $a_r.b_s = 0$, lo que implicaría $a_r = 0$ ó $b_s = 0$, una contradicción. Finalmente, es claro que si $\mathbf{R}$ es un D.I. también lo es $\mathbf{R}[x]$. Sin embargo $\mathbf{R}[x]$ no es anillo de división, pues si lo fuera, necesariamente $\mathbf{R}[x]$ tendría identidad, pero el elemento $(0, 1, 0 \ldots, 0, \ldots)$ no tiene inverso en $\mathbf{R}[x]$.

ii) Es claro.

Vía dicho monomorfismo se puede identificar $\mathbf{R}$ con su imagen isomorfa en $\mathbf{R}[x]$, lo que haremos en lo sucesivo. Por ello, usando esta identificación, si $a \in \mathbf{R}$, cuando escribamos $a.(a_0, a_1, \ldots)$ entenderemos que es el elemento $(a, 0, \ldots)(a_o, a_1, \ldots) = (a.a_0, a.a_1, \ldots)$.

Si $\mathbf{R}$ tiene identidad 1 y x se identifica con el polinomio $(0, 1, 0, \ldots)$ de $\mathbf{R}[x]$, entonces $x^n = (0, 0, \ldots, 1^{(n+1}, 0, \ldots), \forall n \geq 0$.

En efecto: es cierto para $n = 0, 1$; supongámoslo cierto para n y comprobémoslo para $n + 1$:

$x^{n+1} = x^n.x = (0, \ldots, 1^{(n+1}, 0, \ldots).(0, 1, 0, \ldots) =$

$= (0, \ldots, 1^{(n+2}, 0, \ldots)$.

Nota: A partir de ahora supondremos que R tiene identidad, por tanto si $a \in \mathbf{R}$, $a.x^n = (0, \ldots, a^{(n+1)}, 0, \ldots) = x^n.a$.

Un polinomio se dice **mónico** si tiene coeficiente director igual a 1.

Nota: Si denotamos con $\delta(f)$ al grado del polinomio no nulo $f(x)$ se cumplen las afirmaciones: $\delta(f + g) \leq \max\{\delta(f), \delta(g)\}$ y $\delta(f.g) \leq \delta(f) + \delta(g)$

(5.3) **Teorema** : (Algoritmo de la división). Sea $\mathbf{R}$ un anillo f y g polinomios no nulos sobre $\mathbf{R}$. Supongamos que el coeficiente director de g es una unidad de $\mathbf{R}$. Entonces

existen dos únicos polinomios q y r tales que: $f = g.q + r$ con $r = 0$ ó $r \neq 0$ y grado $(r) <$ grado (g).

Demostración: Sean $f = a_0 + a_1x + \ldots + a_nx^n, a_n \neq 0$ y $g = b_0 + b_1x + \ldots + b_mx^m, b_m$ unidad.

Haremos la demostración por inducción sobre n. Si $n = 0$ y grado$(g) > 0$ tomamos $q = 0$ y $r = f$. Si $n = 0$ y $m = 0$ tomamos $r = 0$ y $q = {b_0}^{-1}a_0$. Así, podemos suponer $n > 0$ y grado$(g) \leq$ grado(f), pues si grado$(g) >$ grado(f), podemos tomar $q = 0$ y $r = f$.

Considerar $gx^{n-m}{b_m}^{-1}a_n$ que será un polinomio de grado n y coeficiente director a_n. Así $f = gx^{n-m}{b_m}^{-1}a_n + f_1$ con $f_1 = f - gx^{n-m}{b_m}^{-1}a_n$ y por tanto $f_1 = 0$ ó $f_1 \neq 0$ y grado$(f_1) < n$. En el segundo caso, por hipótesis de inducción existen q_1 y r_1 tales que $f_1 = g.q_1 + r_1$ con $r_1 = 0$ ó grado$(r_1) <$grado(g), por tanto $f = gx^{n-m}{b_m}^{-1}a_n + g.q_1 + r_1$, luego $f = g(x^{n-m}{b_m}^{-1}a_n + q_1) + r_1$.

Probemos ahora la unicidad: Suponer dos descomposiciones $f = g.q_1 + r_1 = g.q_2 + r_2$, con $r_1 = 0$ ó grado$(r_1) <$ grado(g) y $r_2 = 0$ ó grado$(r_2) <$ grado(g). Así $g(q_1 - q_2) = r_2 - r_1$.

Si $r_1 = r_2$ debe ser $q_1 = q_2$, pues en otro caso el coeficiente director de g que es una unidad por el coeficiente director de $(q_1 - q_2)$ darían producto cero, lo que no es posible.

Si $r_1 \neq r_2$, entonces $q_1 \neq q_2$ y grado$(g(q_1 - q_2)) =$ grado$(g) +$ grado$(q_1 - q_2) =$grado$(r_2 - r_1)$, pero grado$(r_2 - r_1) <$ grado(g), una contradicción.

Un caso importante del resultado anterior es cuando el anillo de coeficientes es un anillo de división. En este caso el único requerimiento para g es que éste sea no cero.

Recordar que un D.I.P es un dominio de integridad en el que cada ideal es principal.

(5.4) **Teorema** : Si $\mathbf{F}$ es un cuerpo entonces $\mathbf{F}[x]$ es un D.I.P..

Demostración: Sea $\mathbf{I}$ un ideal de $\mathbf{F}[x]$. Para demostrar que $\mathbf{I}$ es principal podemos suponer $\mathbf{I} \neq (0)$. Entonces $\mathbf{I}$ contiene elementos no cero. Elijamos $f \in \mathbf{I}$,$f \neq 0$, de grado el menor posible entre los elementos no nulos de $\mathbf{I}$. Es claro que $(f) \subseteq \mathbf{I}$ y afirmamos que $(f) = \mathbf{I}$. Sea $g \in \mathbf{I}$, g no nulo, entonces por el algoritmo de la división $g = f.q + r$ con $r = 0$ ó $r \neq 0$ y grado$(r) <$ grado(f). La única alternativa posible es que $r = 0$ y así $g \in (f)$.

(5.5) **Corolario**: Sea $\mathbf{F}$ un cuerpo y $f(x), g(x) \in \mathbf{F}[x]$ dos polinomios no nulos. Entonces existe máximo común divisor y mínimo común múltiplo de $f(x), g(x)$. Además se cumple la identidad de Bezout.

Demostración: Es consecuencia inmediata de (5.4) y (4.26).

Observar que dados dos polinomios no cero $f(x), g(x)$ existe un único máximo común divisor mónico. Nos referiremos a él cuando hablemos del máximo común divisor d y lo denotaremos $d = (f, g)$.

(5.7) **Teorema**: Sean $f, g, h \in \mathbf{F}[x]$, con $\mathbf{F}$ cuerpo. Supongamos que $(f, g) = 1$. Si $f|g.h$, entonces $f|h$.

Demostración: Por (5.5) existen polinomios $\alpha(x), \beta(x) \in \mathbf{F}[x]$, tales que $1 = \alpha(x)f(x) + \beta(x)g(x)$. Luego

$h(x) = 1.h(x) = (\alpha(x)f(x) + \beta(x)g(x))h(x)$,

que es un múltiplo de $f(x)$.

Nota: Si $\mathbf{F}$ es un cuerpo, las unidades de $\mathbf{F}[x]$, son los polinomios constantes distintos del cero. Por tanto los polinomios irreducibles sobre $\mathbf{F}$ (es decir los elementos irreducibles de $\mathbf{F}[x]$) tienen grado mayor ó igual que 1.

(5.8) **Teorema**: Si $\mathbf{F}$ es un cuerpo, todo polinomio de grado ≥ 1 se factoriza como producto de polinomios irreducibles.

Demostración: Por inducción sobre el grado n de $f(x)$. Si $n = 1$ la afirmación es trivial. Supongamos $n > 1$ y la afirmación cierta para polinomios de grado menor ó igual que $n - 1$. Si $f(x)$ es irreducible la afirmación es obvia. Si $f(x)$ no es irreducible existe un polinomio $g(x)$, no constante, que divide a $f(x)$ y es de grado inferior. Tenemos ahora $f(x) = g(x)h(x)$ conmo producto de dos polinomios de grado inferior. De la hipótesis inductiva se sigue la tesis.

Ejercicio. Probar que la descomposición de un polinomio de grado ≥ 1 como producto de polinomios irreducibles, de la que se habla en el teorema anterior, es única salvo el orden y la multiplicación por constantes.

(5.9) **Teorema** : Sean $\mathbf{R}$ y $\mathbf{S}$ anillos conmutativos $\varphi : R \longrightarrow S$ un homomorfismo de anillos con $\varphi(1_{\mathbf{R}}) = 1_{\mathbf{S}}$, entonces queda definido un único homorfismo $\tilde{\varphi} : \mathbf{R}[x] \longrightarrow \mathbf{S}$ tal que $\tilde{\varphi}|_{\mathbf{R}} = \varphi$ y $\tilde{\varphi}(x) = s \in \mathbf{S}$ previamente fijado.

Demostración: Dado $f = a_o + a_1x + \ldots + a_nx^n$, definimos $\tilde{\varphi}(f) = \varphi(a_o) + \varphi(a_1)s + \ldots + \varphi(a_n)s^n$. Si $g = b_0 + b_1x + \ldots + b_mx^m$, $\tilde{\varphi}(g) = \tilde{\varphi}(b_o + b_1x + \ldots + b_mx^m) = \varphi(b_o) + \varphi(b_1)s + \ldots + \varphi(b_m)s^m$, luego $\tilde{\varphi}(f+g) = \tilde{\varphi}(f) + \tilde{\varphi}(g)$ pues φ es un homomorfismo de anillos. Calculemos ahora $\tilde{\varphi}(fg)$. Para obtenerlo, recordar que $fg = \sum c_rx^r$, donde $c_r = \sum_{i+j=r} a_ib_j$, por tanto $\varphi(c_r) = \sum_{i+j=r} \varphi(a_i)\varphi(b_j)$, así la imagen del producto fg sería $\varphi(c_0) + \varphi(c_1)s + \ldots + \varphi(c_l)s^l = \varphi(a_0)\varphi(b_0) + (\varphi(a_1)\varphi(b_0) + \varphi(a_0)\varphi(b_1))s + \ldots$ que por la conmutatividad de $\mathbf{S}$ es el producto de las imágenes $\tilde{\varphi}(g).\tilde{\varphi}(f)$.

Es claro que $\tilde{\varphi}_{|\mathbf{R}} = \varphi$ y $\tilde{\varphi}(x) = s$.

Si $\psi : \mathbf{R}[x] \longrightarrow \mathbf{S}$ es otro homomorfismo de anillos verificando $\psi_{|\mathbf{R}} = \varphi$ y $\psi(x) = s$, entonces $\psi(a_0 + a_1x + \ldots + a_nx^n) = \psi(a_0) + \psi(a_1)\psi(x) + \ldots + \psi(a_n)\psi(x^n) = \varphi(a_0) + \varphi(a_1)s + \ldots + \varphi(a_n)s^n = \tilde{\varphi}(a_0 + a_1x + \ldots + a_nx^n)$.

El homomorfismo $\tilde{\varphi}$ se dice **evaluación ó sustitución** y se utilizará sin que a veces se cite expresamente.

(5.10) **Corolario**: Sea $\varphi : \mathbf{R} \longrightarrow \mathbf{S}$ un homomorfismo entre anillos conmutativos tal que $\varphi(1_R) = 1_S$. Entonces la aplicacción $\tilde{\varphi} : \mathbf{R}[x] \longrightarrow \mathbf{S}[x]$ dada por $\tilde{\varphi}(\sum_{i=0}^{n} a_ix^i) = \sum_{i=0}^{n} \varphi(a_i)x^i$ es un homomorfismo de anillos. Además Ker $\tilde{\varphi} = \{f(x) = \sum_{i=0}^{n} a_ix^i | \varphi(a_i) = 0, \forall i = 0, \ldots, n\}$. Si φ es un isomorfismo, entonces también lo es $\tilde{\varphi}$.

Demostración: Sea α el monomorfismo de $\mathbf{S}$ en $\mathbf{S}[x]$, dado por $\alpha(s) = (s, 0, \ldots, 0, \ldots)$ basta aplicar el teorema anterior al homomorfismo $\alpha \circ \varphi : \mathbf{R} \longrightarrow \mathbf{S}[x]$ tomando $s = x$. El resto de la demostración es pura comprobación.

(5.11) **Definición**: Si $f(x) = \sum_{i=0}^{n} a_ix^i \in \mathbf{R}[x]$ y $b \in \mathbf{R}$ escribiremos $f(b) := \sum_{i=0}^{n} a_ib^i$. Diremos que b es una **raíz** de $f(x)$ cuando $f(b) = 0$

(5.12) **Corolario**: Sea $\mathbf{R}$ es un anillo conmutativo y $a \in \mathbf{R}$. Entonces la aplicación $\varphi_a : \mathbf{R}[x] \longrightarrow \mathbf{R}$ dada por $\varphi_a(f(x)) = f(a)$ es un homomorfismo de anillos.

Demostración: Si $f(x) = a_o + a_1.x + \ldots + a_n.x^n$, entonces por (5.9) la aplicación φ_a dada por $\varphi_a(f) = a_o + a_1.a + \ldots + a_n.a^n$ es un homomorfismo

(5.13) **Teorema** : Sea $\mathbf{R}$ un anillo conmutativo y $f \in \mathbf{R}[x]$, $f \neq 0$. Entonces $a \in \mathbf{R}$ es raíz de f si y sólo si $x - a$ divide a f.

Demostración: Podemos hacer uso del algoritmo de la división. Así $f = (x-a)q+r$ con $r = 0$ ó $r \neq 0$ y grado$(r)<$ grado$(x - a)= 1$. Por el resultado anterior podemos hacer la sustitución de x por a de forma que $f(a) = 0.q(a) + r(a)$.

Si a es raíz de f, entonces $r(a) = 0$, así $r = 0$ y $x - a$ divide a f. Si recíprocamente $x - a$ divide a f, entonces $f = (x - a)q$ y así $f(a) = 0$.

(5.14) **Teorema**: Sean $\mathbf{R}$ y $\mathbf{S}$ dominios de integridad con $\mathbf{R} \subseteq \mathbf{S}$ y $f \in \mathbf{R}[x]$ un polinomio de grado $n \geq 1$. Entonces f tiene a lo sumo n raíces distintas en $\mathbf{S}$.

Demostración: Sea $c_1, c_2, \ldots$ raíces distintas de f en $\mathbf{S}$. Sabemos que $f(x) = (x - c_1)q_1(x)$; por el homomorfimos de sustitución es $f(c_2) = (c_2 - c_1)q_1(c_2) = 0$ así que $q_1(c_2) = 0$, luego $x - c_2$ divide a $q_1(x)$ y $f(x) = (x - c_1)(x - c_2)q_2(x)$. Si razonamos inductivamente, probaremos que si $c_1, \ldots, c_m$ son raíces distintas de f en $\mathbf{S}$, $g_m(x) = (x - c_1).(x - c_2).\ldots.(x - c_m)$ divide a $f(x)$, pero grado$(g_m(x)) = m$ y $m \leq n$, así el número de raíces distintas está acotado por n.

Ejercicio (5.14) es falso sin la condición de que S sea dominio de integridad. Tomar $\mathbf{R} = \mathbf{S} = \mathbf{Z}/8\mathbf{Z}$ y el polinomio $x^2 - 1$

Para obtener el siguiente corolario, que es de gran interés, utilizaremos:

Ejercicio: Sea G un grupo de orden n. Supongamos que para cada $d|n$ existen a lo sumo d elementos $g \in G$ verificando $g^d = 1$. Entonces G es un grupo cíclico.

(5.15) **Corolario**: Si $\mathbf{F}$ es un cuerpo y G es un subgrupo finito de $\mathbf{F} - \{0\}$, entonces G es cíclico.

Demostración: Sea $n = |G|$. Para cada $d|n$, existen a lo sumo d raices en $\mathbf{F}$ del polinomio $x^d - 1$. Por tanto, para cada $d|n$ existen a lo sumo d elementos g de G cumpliendo que $g^d = 1$. En consecuencia G es cíclico.

En particular, si p es un primo, $(\mathbf{Z}/p\mathbf{F} - \{[0]\}, .)$ es un grupo cíclico.

Notar que la irreducibilidad de polinomios depende del cuerpo, por ejemplo $x^2 + 1$ es irreducible sobre $\mathbf{R}$ pero no sobre $\mathbf{C}$.

El siguiente resultado, que damos sin demostración, es conocido como teorema fundamental del Algebra.

Teorema: Todo polinomio de grado mayor o igual que 1, con coeficientes en el cuerpo de los números complejos tiene al menos una raíz en $\mathbf{C}$.

(5.16) **Corolario** : En $\mathbf{C}[x]$ los únicos polinomios irreducibles son los de grado 1.

Demostración: Es evidente que los polinomios de la forma $x - a \in \mathbf{C}[x]$ son irreducibles. Recíprocamente, si $f(x) \in \mathbf{C}[x]$ es irreducible necesariamente tiene que ser de grado 1, ya que en otro caso, como tiene una raíz a en $\mathbf{C}$, $f(x)$ es divisible por $x - a$, y $f(x) = (x-a)g(x)$ con $g(x)$ de grado mayor o igual que 1, luego $f(x)$ no sería irreducible.

(5.17) **Proposición** : Los polinomios irreducibles sobre el cuerpo $\mathbf{R}$ de los números reales son de grado 1 ó 2.

Demostración: Sea $f(x) = a_0 + a_1x + \ldots + a_1x + a_nx^n$ un polinomio irreducible en $\mathbf{R}$. Si $f(x)$ tiene una raíz real a, entonces $f(x) = (x-a)g(x)$ y como $f(x)$ es irreducible sería $f(x) = a_1(x-a)$. Podemos suponer que $f(x)$ no tiene raices reales y que $n \geq 1$. Por el corolario (5.10), la conjugación compleja de $\mathbf{C}$ en sí se extiende a un homomorfismo de $\mathbf{C}[x]$ en $\mathbf{C}[x]$ definido por $\bar{g}(x) = \bar{b}_m x^m + \ldots + \bar{b}_1 x + \bar{b}_0$ indicando por $\bar{b}$ al conjugado complejo de b y siendo $g(x)$ un polinomio con coeficientes en $\mathbf{C}$. Visto el polinomio $f(x)$ en $\mathbf{C}[x]$ se puede descomponer, por el teorema (5.8) y el corolario (5.16), en la forma $f(x) = a_n(x-\alpha_1)\ldots(x-\alpha_n)$, por tanto $f(x) = \bar{f}(x) = \bar{a}_n(x-\bar{\alpha}_1)\ldots(x-\bar{\alpha}_n)$, luego si α_i es raíz de $f(x)$, entonces $\bar{\alpha}_i$ es también raíz de $f(x)$. Agrupando con cada factor $(x-\alpha_i)$ el factor $(x-\bar{\alpha}_i)$ obtenemos un factorización de $f(x)$ que, salvo constantes, es un producto de polinomios de la forma $x^2 + p_ix + q_i$, con p_i, q_i reales. Como $f(x)$ es irreducible, se sigue que, salvo constante, $f(x) = x^2 + p_1x + q_1$ es decir $f(x)$ tiene grado 2.

Algunos criterios de irreducibilidad

Observación : Si **F** es un cuerpo, un polinomio no cero $p(x) \in \mathbf{F}[x]$ es irreducible en $\mathbf{F}[x]$ (o **polinomio irreducible** sobre **F**) si grado$(p(x)) \geq 1$ y no existe una factorización $p(x) = f(x).g(x)$ en **F**[x] con grado$(f(x))$ < grado $(p(x))$ y grado$(g(x))$ < grado $(p(x))$.

Se trata en esta lección de encontrar criterios para que un polinomio sea irreducible, concretamente estamos interesados en criterios de irreducibilidad sobre **Q** de polinomios con coeficientes en **Z**.

(5.18) **Definición** : Un polinomio $(a_o, a_1, \ldots) \in \mathbf{Z}[x]$ se llama **primitivo** si el máximo común divisor de sus coeficientes es 1.

(5.19) **Lema** (Gauss): El producto de dos polinomios primitivos $f(x)$ y $g(x)$ es también primitivo.

Demostración: Suponer $f(x).g(x) = (\sum_i a_i x^i)(\sum_j b_j x^j) = \sum_k c_k x^k$ no primitivo, así que existe p primo divisor de cada c_k. Sea i el índice tal que p no divide a a_i pero p divide a a_k si $k < i$. Análogamente sea j el índice tal que p no divide a b_j, pero p divide a b_l si $l < j$. Entonces como $a_i.b_j = c_{i+j} - (a_o b_{i+j} + \ldots + a_{i-1}b_{j+1} + a_{i+1}b_{j-1} + \ldots + a_{i+j}b_o)$ y cada término de la derecha es divisible por p, se sigue que $p|a_i b_j$ luego $p|a_i$ ó $p|b_j$, contradicción.

(5.20) **Lema** : Cada polinomio no cero $f(x) \in \mathbf{Q}[x]$ tiene una única factorización $f(x) = c(f).f^*(x)$ donde $c(f) \in \mathbf{Q}$ es positivo y $f^* \in \mathbf{Z}[x]$ es primitivo.

Demostración: Sea $f(x) = (a_o/b_o) + (a_1/b_1)x + \ldots + (a_n/b_n)x^n \in \mathbf{Q}[x]$. Sea $b = b_o b_1 \ldots b_n$ así $b.f(x) \in \mathbf{Z}[x]$, $f(x) = (1/b)g(x), g(x) \in \mathbf{Z}[x]$. Sea ahora $b' = \pm$ m.c.d (coeficientes de $g(x)$) (el signo se elige para hacer (b'/b) positivo), así $f(x) = c(f)f^*(x)$ donde $c(f) = b'/b$ y $f^*(x) = (b/b')f(x)$.

Sea ahora $f(x) = d.h(x)$ una segunda factorización, así $f^*(x) = r.h(x)$ con $r = d/c(f)$ un racional positivo. Escribir $r = u/v$ con u, v positivos primos entre sí. Entonces $vf^*(x) = u.h(x)$ expresión en $\mathbf{Z}[x]$, los coeficientes de $u.h(x)$ tienen a v como divisor común y así v dividirá a los coeficientes de $h(x)$. Pero $h(x)$ es primitivo de ahí que $v = 1$. Un argumento similar lleva a $u = 1$, así $r = 1$ y $d = c(f)$, $f^*(x) = h(x)$.

(5.21) **Lema** : Si $f(x) \in \mathbf{Q}[x]$ factoriza como $f(x) = g(x)h(x)$, entonces $c(f) = c(g)c(h)$ y $f^*(x) = g^*(x)h^*(x)$.

Demostración: $f(x) = g(x)h(x) = (c(g)g^*(x))(c(h)h^*(x)) = c(g)c(h)g^*(x)h^*(x)$. Como $c(g)c(h)$ es racional positivo y el producto de primitivos es primitivo, la unicidad de la factorización del lema anterior implica $c(f) = c(g)c(h)$ y $f^*(x) = g^*(x)h^*(x)$.

(5.22) **Teorema** : Sea $f(x) \in \mathbf{Z}[x]$ de grado positivo. Si $f(x)$ es irreducible en $\mathbf{Z}[x]$, entonces $f(x)$ es irreducible en $\mathbf{Q}[x]$. Si $f(x)$ es primitivo, también vale el recíproco.

Demostración: Suponer que $f(x)$ no es irreducible en $\mathbf{Q}[x]$, entonces existen polinomios $g(x)$ y $h(x)$ de grados positivos tales que $f(x) = g(x)h(x)$ en $\mathbf{Q}[x]$,por tanto

$$f(x) = c(g)c(h)g^*(x)h^*(x)$$

donde $g^*(x)$ y $h^*(x)$ en $\mathbf{Z}[x]$, pero $c(g)c(h) = c(f) \in \mathbf{Z}$ pues $f(x) \in \mathbf{Z}[x]$, así $f(x) = (c(f)g^*(x))h^*(x)$ es una factorización en $\mathbf{Z}[x]$, contradicción.

(5.23) **Teorema** : (Criterio de Eisenstein) Sea $f(x) = a_o + a_1x + \ldots + a_nx^n \in \mathbf{Z}[x], n \geq 1$. Si p es un primo en $\mathbf{Z}$ dividiendo a $a_i, \forall i \leq n$ pero no dividiendo a a_n y p^2 no divide a a_o, entonces $f(x)$ es irreducible sobre $\mathbf{Q}$.

Demostración: Como $f(x) = c(f)f^*(x)$, $f^*(x)$ primitivo en $\mathbf{Z}[x]$ y $c(f)$ es una unidad de $\mathbf{Q}[x]$, bastará probar que $f^*(x)$ es irreducible en $\mathbf{Q}[x]$. Por tanto podemos suponer quie $f(x)$ es primitivo. Veamos que $f(x)$ es irreducible en $\mathbf{Z}[x]$. Los únicos divisores constantes de $f(x)$ son 1 y -1. Suponer que $b_o + b_1x + \ldots + b_mx^m \in \mathbf{Z}[x]$, con $1 \leq m < n$, es un divisor de $f(x)$. Entonces existe $c_o + c_1x + \ldots + c_kx^k \in \mathbf{Z}[x]$, $k < n$, tal que: $f(x) = (b_o + b_1x + \ldots + b_mx^m)(c_o + c_1x + \ldots + c_kx^k)$. Como $p|a_o = b_oc_o$, entonces $p|b_o$ ó $p|c_o$, pero como $p^2 \not| a_o$, p sólo dividirá a uno de ellos. Supongamos $p|c_o$ pero $p \not| b_o$. El coeficiente director $a_n = b_mc_k$ no es divisible por p, así $p \not| c_k$. Sea c_r el primer coeficiente no divisible por p. Como $r < n$ entonces $p|a_r$ y $b_oc_r = a_r - (b_1c_{r-1} + \ldots + b_rc_o)$ divisible por p, luego $p|b_oc_r$, lo que es una contradicción. Así los únicos divisores de $f(x)$ en $\mathbf{Z}[x]$ son unidades o asociados a $f(x)$, y por tanto $f(x)$ es irreducible en $\mathbf{Z}[x]$. Por el teorema anterior concluimos que $f(x)$ es irreducible en $\mathbf{Q}[x]$.

Ejercicios: 1. Sea $p \in \mathbf{Z}$ un primo. Considerar el homomorfismo canónico $\varphi : \mathbf{Z} \longrightarrow \mathbf{Z}/p\mathbf{Z}$. Utilizando (5.10) probar que si $f(x) \in \mathbf{Z}[x]$ es mónico y $\tilde{\varphi}(f(x))$ es irreducible en

$\mathbf{Z}/p\mathbf{Z}[x]$, entonces $f(x)$ es irreducible en $\mathbf{Z}[x]$.

2. Sea $f(x) = a_o + a_1x + \ldots + a_nx^n$ un polinomio con coeficientes enteros. Sea r/s una raíz de $f(x)$ en el cuerpo de los números racionales con $(r, s) = 1$. Probar que $r|a_o$ y $s|a_n$.

Lección 6. Espacios vectoriales.

(6.1) **Definición**: Dado un cuerpo K se llama **espacio vectorial** sobre K ó K-**espacio vectorial** a un conjunto V, no vacío, cuyos elementos se llaman **vectores**, dotado de una operación binaria interna $+$ y una externa . con dominio de operadores a izquierda K de forma que:

i) $(V,+)$ es un grupo abeliano.

ii) La operación externa cumple:

$\forall t,s \in K, \forall a,b \in V$:

a) $t.(a+b) = t.a + t.b$; b) $(t+s).a = t.a + s.a$; c) $(ts).a = t.(s.a)$; d) $1_K.a = a$.

Los elementos de K se dicen **escalares**.

Consecuencias de la definición:

1ª) $0.a = \bar{0}, \forall a \in V$. En efecto: $0.a = (0+0)a = 0.a + 0.a, \Rightarrow 0.a = \bar{0}, \forall a \in V$.

2ª) $-a = (-1).a, \forall a \in V$, pues $a + (-1)a = (1 + (-1)a) = 0.a = \bar{0}$.

3ª) $t.\bar{0} = \bar{0}, \forall t \in K$, pues $t.\bar{0} = t.(\bar{0} + \bar{0}) = t.\bar{0} + t.\bar{0} \Rightarrow t.\bar{0} = \bar{0}$.

4ª) Si $t.a = \bar{0} \Rightarrow t = 0$ ó $a = \bar{0}$, pues $t \neq 0 \Rightarrow \bar{0} = t^{-1}.(t.a) = 1.a = a$.

Ejemplos: i) K con la suma y producto del cuerpo y dominio de operadores el propio K.

ii) El conjunto $K^n = \{(t_1, \ldots, t_n) | t_i \in K\}$, con las operaciones :

$(t_1, \ldots, t_n) + (s_1, \ldots, s_n) = (t_1 + s_1, \ldots, t_n + s_n)$, y $t.(t_1, \ldots, t_n) = (tt_1, \ldots, tt_n)$.

iii) $K[x]$ con la suma de polinomios y el producto $t.(a_o + a_1 x + \ldots + a_n x^n) = ta_o + ta_1 x + \ldots + ta_n x^n$.

En lo sucesivo V denotará un espacio vectorial sobre K. En general omitiremos la notación . tanto para la operación externa en V como para el producto interno en K.

(6.2) **Definición**: Se llama subespacio vectorial de V a un subconjunto no vacío S de V, que cumple: si $a, b \in S$ y si $t \in K$, entonces $a + b \in S$ y $t.a \in S$ y S es espacio vectorial con la restricción de ambas operaciones.

(6.3) **Proposición**: Una parte no vacía S de V es **subespacio vectorial** de V si y sólo si cumple:

i) Cualesquiera que sean a, b de S, $a - b \in S$.

ii) Cualesquiera que sean $a \in S$ y $t \in K$, $t.a \in S$.

Demostración: Si S es un subespacio vectorial de V, en particular $(S, +)$ es subgrupo de $(V, +)$, así se verifica i). Evidentemente se verifica ii).

Recíprocamente, supongamos que S es un subconjunto no vacío de V verificando i) y ii), entonces S es un subgrupo de $(V, +)$ y considerando la operación externa restringida a $K \times S$ se cumplirán todas las propiedades, para afirmar que S con las operaciones $+$ y $.$ es un subespacio vectorial de V.

(6.4) **Teorema**: Sea V un espacio vectorial y S un subespacio de V. Se define en V una relación $\mathbf{R}$ mediante: $a\,\mathbf{R}\,b$ si y sólo si $a - b \in S$. Dicha relación es de equivalencia. Ambas operaciones $+$ y $.$ son estables para $\mathbf{R}$. Denotamos al conjunto cociente por V/S. Por los teoremas (2.11) y (2.14), el conjunto cociente con las operaciones inducidas :

$+ : V/S \times V/S \longrightarrow V/S$ dada por $[a] + [b] = [a + b]$ y

$. : K \times V/S \longrightarrow V/S$ dada por $t.[a] = [t.a]$

pasa a ser un espacio vectorial sobre K llamado **espacio vectorial cociente** de V por S

Demostración: $\mathbf{R}$ es de equivalencia pues $a\,\mathbf{R}\,a, \forall a \in V$; si $a\,\mathbf{R}\,b$, es que $a - b \in S$, luego $-a + b \in S$, así que $b\,\mathbf{R}\,a$; finalmente, si $a\,\mathbf{R}\,b$ y $b\,\mathbf{R}\,c$, es que $a - b \in S$ y $b - c \in S$, luego $a - b + b - c = a - c \in S$ y por tanto $a\,\mathbf{R}\,c$.

Notar que $[a] = \{b | a\,\mathbf{R}\,b\} = \{b | a - b \in S\} = a + S$. Probemos que las operaciones son estables para la relación de equivalencia:

i) Si $[a] = [a']$ y $[b] = [b']$ es que $a - a' \in S$ y $b - b' \in S$. Por tanto $a - a' + b - b' \in S$, es decir $(a + b)\mathbf{R}(a' + b')$ y $[a + b] = [a' + b']$. Es sencillo probar que la suma de clases es asociativa, que $[0]$ es el elemento neutro y que $-[a] = [-a]$. Así V/S es un grupo y como $[a] + [b] = [b] + [a], \forall a, b \in V$, es un grupo abeliano.

ii) Si $[a] = [a']$, entonces $a - a' \in S$, luego $\forall t \in K$, $t(a - a') = ta - ta' \in S$, en consecuencia $[ta] = [ta']$.

Se cumplen las demás propiedades de espacio vectorial, pues $\forall s, t \in K$, y $, b \in V$:

$$t([a] + [b]) = t[a + b] = [t(a + b)] = [ta + tb] = [ta] + [tb] = t[a] + t[b];$$

$$(t + s)[a] = [(t + s)a] = [ta + sa] = [ta] + [sa] = t[a] + s[a];$$

$$(t(s[a]) = t[sa] = [t(sa)] = [(ts)a] = (ts)[a], \text{ además : } 1[a]) = [1.a] = [a].$$

(6.5) **Definición**: Una aplicación $f : V \longrightarrow W$ entre dos K-espacios vectoriales se dice **homomorfismo ó aplicación lineal** si $\forall a, b \in V$ y $\forall t \in K$:

$$f(a+b) = f(a) + f(b), \text{ y } f(ta) = tf(a).$$

Si además f es inyectiva, f se dirá **monomorfismo**. Si es suprayectiva se dirá **epimorfismo**. Si es biyectiva se dirá **isomorfismo**.

Suma e intersección de subespacios. Suma directa

(6.6) **Definición**: Dados r subespacios $S_1, \ldots, S_r$ de V, se llama **suma lineal** de ellos al conjunto $\{a_1 + \ldots + a_r | a_i \in S_i, i = 1, \ldots, r\}$, que será denotado $S_1 + \ldots + S_r$.

(6.7) **Proposición**: La suma lineal $S = S_1 + \ldots + S_r$ y la intersección $T = S_1 \cap \ldots \cap S_r$ son subespacios vectoriales de V.

Demostración: Notar que tanto S como T son no vacíos pues $\bar{0} \in S, T$. Sean $a, b \in S$, entonces existen elementos $a_i, b_i \in S_i$ tales que $a = a_1 + \ldots + a_r$ y $b = b_1 + \ldots + b_r$, en consecuencia $a - b = a_1 - b_1 + \ldots + a_r - b_r \in S$ puesto que $a_i - b_i \in S_i$ y cualquiera que sea $t \in K$ se tiene que $ta = ta_1 + \ldots + ta_r \in S$ pues $ta_i \in S_i$.

En cuanto a la intersección: Sean $a, b \in T$, entonces $a, b \in S_i, \forall i$ luego $a - b \in S_i, \forall i$, luego $a - b \in T$. Además, cualquiera que sea $t \in K$ se cumple que $ta \in S_i, \forall i$, luego $ta \in T$.

Nota: Con análoga demostración se prueba que la intersección de un conjunto no finito de subespacios de V es de nuevo un subespacio de V.

(6.8) **Definición**: La suma lineal S de los subespacios $S_1, \ldots, S_r$ se dice directa y se escribe $S = S_1 \oplus \ldots \oplus S_r$, cuando cada elemento a de S puede expresarse de una única forma como $a = a_1 + \ldots + a_r$, $a_i \in S_i$, $i = 1, \ldots, r$.

(6.9) **Proposición**: La suma lineal S de los subespacios $S_1, \ldots, S_r$ es directa si y sólo si $a_1 + \ldots + a_r = \bar{0}$ implica $a_1 = \ldots = a_r = \bar{0}$.

Demostración: Si la suma es directa de la expresiones $\bar{0} = a_1 + \ldots + a_r = \bar{0} + \ldots + \bar{0}$, se sigue $a_1 = \ldots = a_r = \bar{0}$.

Recíprocamente, si se cumple que $a_1 + \ldots + a_r = \bar{0} \Rightarrow a_1 = \ldots = a_r = \bar{0}$, supongamos $a = b_1 + \ldots + b_r = c_1 + \ldots + c_r, b_i, c_i \in S_i$, entonces $\bar{0} = (b_1 - c_1) + \ldots + (b_r - c_r)$, luego $b_i - c_i = \bar{0}, \forall i = 1, \ldots, r$ y así $b_i = c_i, \forall i = 1, \ldots, r$. Luego la suma es directa.

(6.10) **Definición**: Cuando la suma de los subespacios $S_1, \ldots, S_r$ es directa , se dice que los subespacios $S_1, \ldots, S_r$ son **subespacios independientes**.

(6.11) **Definición**: Dos subespacios S y T son **suplementarios** si $V = S \oplus T$.

Ejercicio : Demostrar que dos subespacios S y T de un espacio vectorial V son suplementarios si y sólo si $V = S + T$ y $S \cap T = \bar{0}$.

Clausura lineal. Dependencia e independencia lineal. Bases.

En este párrafo utilizaremos frecuentemente los términos de familia ó sistema de vectores para indicar una sucesión de vectores de V, en la que puede haber algún vector repetido.

(6.12) **Definición**: Sea $(a_1, \ldots, a_m)$ una familia finita de vectores de un espacio vectorial V. Se llama **combinación lineal** de dicha familia a cualquier vector a de V que pueda expresarse en la forma $a = t_1 a_1 + \ldots + t_m a_m$, $t_i \in K$.

(6.13) **Teorema**: El conjunto de todas las combinaciones lineales de $(a_1, \ldots, a_m)$ es un subespacio de V.

Demostración: Denotemos por S a dicho conjunto. Notar que $S \neq \emptyset$ pues $a_i \in S, \forall i$. Si $a = t_1 a_1 + \ldots + t_m a_m$, $t_i \in K$ y $b = s_1 a_1 + \ldots + s_m a_m$, $s_i \in K$, entonces $a - b = (t_1 - s_1)a_1 + \ldots + (t_m - s_m)a_m$ y si $t \in K$, $ta = (tt_1)a_1 + \ldots + (tt_m)a_m$.

(6.14) **Definición**: El subespacio S anterior se llama **clausura lineal** de $(a_i), i = 1, \ldots, m$. Tambien se dice que S está **generado** por dicha familia y que ésta es un **sistema generador** de S. Se escribe $S = K(a_1, \ldots, a_m)$ ó abreviadamente $S = K(a_i)$.

(6.15) **Definición**: Un espacio vectorial V se dice de **tipo finito** ó **finitamente generado** si posee un sistema generador finito.

(6.16) **Teorema**: Si una familia de vectores $(b_1, \ldots, b_r)$ está contenida en la clausura lineal de otra $K(a_i)$, entonces $K(b_j) \subseteq K(a_i)$.

Demostración: Sea $b \in K(b_j)$, entonces $b = t_1 b_1 + \ldots + t_r b_r$, pero como $b_j \in K(a_i)$, se tiene que $b_j = s_{j1}a_1 + \ldots + s_{jm}a_m$, luego

$b = t_1(s_{11}a_1 + \ldots + s_{1m}a_m) + t_2(s_{21}a_1 + \ldots + s_{2m}a_m) + \ldots + t_r(s_{r1}a_1 + \ldots + s_{rm}a_m) =$

$(t_1 s_{11} + t_2 s_{21} + \ldots + t_r s_{r1})a_1 + \ldots + (t_1 s_{1m} + t_2 s_{2m} + \ldots + t_r s_{rm})a_m \in K(a_i)$.

Esta es la propiedad transitiva de la dependencia lineal.

(6.17) **Definición**: Dos familias de vectores (a_i) y (b_j) se dicen **equivalentes** si $K(a_i) = K(b_j)$.

Por el teorema anterior, dos familias de vectores son equivalentes si y sólo si cada vector de una es combinación lineal de los vectores de la otra.

Ejercicios : 1) $K(a_1, \ldots, a_m)$ es el menor subespacio que contiene a $(a_1, \ldots, a_m)$.

2) La clausura lineal de una familia (a_i) coincide con la intersección de todos los subespacios que contienen a la familia (a_i).

3) $a \in K(a_1, \ldots, a_m) \iff K(a_1, \ldots, a_m, a) = K(a_1, \ldots, a_m)$.

(6.13), (6.14), (6.16) y (6.17) se pueden ampliar a una familia no finita $X = \{x_i | i \in I\}$ entendiendo que un elemento a de V es combinación lineal de X si es combinación lineal de un subconjunto finito de X.

Por tanto, la clausura lineal de la familia (a_i) es el conjunto S de todas las combinaciones lineales de cada subfamilia finita de (a_i).

Volviendo al caso finito

(6.18) **Definición**: Una familia o sistema de vectores de V, $(a_1, \ldots, a_m)$ se dice **familia ligada** (ó **sistema ligado**) si existe una combinación lineal $t_1 a_1 + \ldots + t_m a_m = \bar{0}$ con algún coeficiente $t_j \neq 0$. A los vectores de la familia se les dice **linealmente dependientes**.

En caso contrario, es decir si siempre que $t_1 a_1 + \ldots + t_m a_m = \bar{0} \Rightarrow t_1 = t_2 = \ldots = 0$, la familia ó sistema se dirá **libre** y los vectores **linealmente independientes**.

(6.19) **Propiedades**: 1) Cualquier familia que contenga al $\bar{0}$ es ligada.

2) Una familia $(a_1, \ldots, a_m)$ es ligada si y sólo si existe un a_i que es combinación lineal de los restantes.

3) Suponer que $(a_i) \subseteq (b_j)$. Entonces:

i) (a_i) ligada $\Rightarrow (b_j)$ ligada;

ii) (b_j) libre $\Rightarrow (a_i)$ libre.

Los conceptos de familia libre o ligada y las propiedades (6.19), pueden trasladarse al caso de familias no finitas.

(6.20) **Definición**: Una familia cualquiera de elementos de V es **ligada** si alguna subfamilia finita es ligada y es **libre** si toda subfamilia finita es libre.

(6.21) **Definición**: Una K-**base**, ó abreviadamente una **base**, de un espacio vectorial V es un sistema generador y libre de V.

Ejemplos: i) $K[x]$ es un espacio vectorial sobre K . Una base de este espacio vectorial es $\{1, x, x^2, \ldots\}$.

ii) K^n es un espacio vectorial sobre K. Una base de este espacio vectorial está formada por las n-tuplas $(1, 0, \ldots, 0), (0, 1, \ldots, 0), \ldots (0, 0, \ldots, 1)$.

iii) Si $V = K[x]$ y $S = \{x^3 f(x) | f(x) \in K[x]\}$, una base de V/S está formada por los vectores $\{1 + S, x + S, x^2 + S\}$.

Si $(a_1, \ldots, a_n)$ es una base de V y $a \in V$, existen $t_1, \ldots, t_n \in K$ de forma que $a = t_1 a_1 + \ldots + t_n a_n$, además dichos escalares son únicos, ya que si $a = s_1 a_1 + \ldots + s_n a_n$, entonces $\bar{0} = (t_1 - s_1)a_1 + \ldots + (t_n - s_n)a_n$, luego $t_1 = s_1, \ldots, t_n = s_n$. Por tanto tenemos definida un isomorfismo: $g : V \longrightarrow K^n$ dada por $a \longmapsto g(a) := (t_1, \ldots, t_n)$ si $a = t_1 a_1 + \ldots + t_n a_n$.

(6.22) **Definición**: Al isomorfismo g anterior se le llama **sistema coordenado** asociado a la base (a_i)de V. A la n-tupla imagen $g(a)$, se le llama n-**tupla coordenada** de a en la base $(a_1, \ldots, a_n)$ y a t_i la i-ésima **coordenada ó componente** de a en dicha base.

(6.23) **Proposición**: Una familia finita $(a_1, \ldots, a_m)$, donde $a_i \neq 0$, para algún i, contiene una subfamilia finita que es libre y genera el mismo subespacio que $(a_1, \ldots, a_m)$.

Demostración: Si $(a_1, \ldots, a_m)$ es libre ya está. En otro caso existe un a_i que es combinación lineal de los restantes y $K(a_1, \ldots, a_m) = K(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_m)$. Si

$(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_m)$ es libre estaría probado el resultado. En caso contrario se reitera el proceso obteniéndose la tesis tras un número finito de pasos.

(6.24) **Corolario**: Si $V \neq \bar{0}$ es de tipo finito, entonces V posee una base finita.

Demostración: Basta tomar un sistema generador finito $(a_1, \ldots, a_n)$ de V y aplicar la proposición anterior.

(6.25) **Proposición**: Los vectores $(a_1, \ldots, a_m)$, con $a_1 \neq \bar{0}$ son linealmente dependientes si y sólo si existe alguno de ellos que es combinación lineal de los anteriores.

Demostración: $\Leftarrow$) Es claro por (6.19)(2).

$\Rightarrow$) Si los vectores $(a_1, \ldots, a_m)$, con $a_1 \neq \bar{0}$ son linealmente dependientes existen $t_1, \ldots, t_m \in K$ con algún $t_i \neq 0$ tales que $t_1 a_1 + \ldots + t_m a_m = \bar{0}$. Sea $r = \text{max.}\{i | t_i \neq 0\}$. Notar que $r \neq 1$, pues si así lo fuera $t_1 a_1 = \bar{0}$ implicaría que $a_1 = \bar{0}$, contra la hipótesis. Así $t_1 a_1 + \ldots + t_r a_r = \bar{0}$ y como $t_r \neq 0$ existe $t_r^{-1} \in K$, luego $t_r^{-1} t_1 a_1 + \ldots + t_r^{-1} t_r a_r = \bar{0}$ y a_r se puede expresar como combinación lineal de los anteriores.

(6.26) **Teorema**: Si los vectores $(a_1, \ldots, a_n)$ forman un sistema generador de V y V contiene r vectores $b_1, \ldots, b_r$ que son linealmente independientes, entonces $r \leq n$.

Demostración: Considerar $(b_1, a_1, \ldots, a_n)$. Es un sistema generador de V y es ligado pues b_1 es combinación lineal de $(a_1, \ldots, a_n)$. Como $b_1 \neq \bar{0}$, sabemos que existe un vector combinación lineal de de los anteriores. Supongamos que es a_i. Entonces $V = K(b_1, a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n)$. Considerar $(b_2, b_1, a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n)$, es ligado pues b_2 es combinación lineal de los restantes, y como $b_2 \neq \bar{0}$ y (b_1, b_2) es libre, existe un vector que es combinación lineal de de los anteriores que debe ser un a_j, así se tiene que $V = K(b_2, b_1, a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n) = K(b_2, b_1, a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_{j-1}, a_{j+1}, \ldots, a_n)$. Si $r > n$, llegaríamos reiterando el proceso, a $V = K(b_1, \ldots, b_n)$. Pero en tal caso $(b_1, \ldots, b_n, b_{n+1})$ es ligado, lo que va contra la hipótesis. Luego $r \leq n$.

(6.27) **Corolario**: En un espacio vectorial $V \neq \bar{0}$ de tipo finito todas las bases tienen el mismo número de elementos. Dicho número n es **la dimensión** de V. Será denotado dim (V) (ó dim $_K(V)$ cuando sea necesario distinguir K). Por convenio, si $V = \bar{0}$, dim (V)=0.

Demostración: Por el teorema anterior dos bases cualesquiera de V son finitas.

Sean $(a_1, \ldots, a_n)$ y $(b_1, \ldots, b_m)$ dos bases de V. Aplicando el teorema anterior a (a_i) como sistema generado y a (b_j) como sistema libre, tenemos que $m \leq n$. Otra aplicación del teorema nos lleva a que $n \leq m$. Luego $m = n$.

(6.28) **Definición**: Un sistema de vectores es libre maximal si es libre y no está contenido propiamente en otro sistema libre.

(6.29) **Teorema**: Sea V un espacio vectorial de tipo finito. Una base de V es un sistema libre maximal y recíprocamente, todo sistema libre maximal es una base de V.

Demostración: $\Rightarrow$) Sea $(a_1, \ldots, a_n)$ una base de V. Si un sistema la contiene propiamente, será de la forma $(a_1, \ldots, a_n, b_1, \ldots, b_s)$ y no será un sistema libre pues b_1 es combinación lineal de $(a_1, \ldots, a_n)$.

$\Leftarrow$) Sea $(a_1, \ldots, a_m)$ un sistema libre maximal. Si añadimos a este sistema un vector b distinto de los a_i dicho sistema no puede ser libre, luego es ligado. Por lo tanto el último vector b es combinación lineal de los anteriores. Es decir (a_i) es un sistema generador libre de V. Luego base de V.

Nota: (6.29) se mantiene válido incluso cuando V no es de tipo finito

(6.30) **Corolario**: Sea V un espacio vectorial de dimensión n. Si S es un subespacio vectorial de la misma dimensión, entonces $S = V$.

Demostración: Una base de S será un sistema libre maximal de V.

(6.31) **Corolario**: Sea V un espacio vectorial de tipo finito. Si S es un subespacio de V, entonces S es de tipo finito, mas aún dim $(S) \leq$dim (V).

Demostración: Sea $n = \dim(V)$. Probaremos que S posee un sistema generador finito, pues en tal caso estará demostrado el resultado ya que S tendrá una base con $r \leq n$ elementos por (6.26):

Si $S = \bar{0}$, el resultado es obvio. Si $S \neq \bar{0}$, tomar $a_1 \in S - \{\bar{0}\}$. Si $S = K(a_1)$, S está engendrado por un único elemento. En otro caso existe $a_2 \in S - K(a_1)$. La familia formada por (a_1, a_2) es libre. Si $S = K(a_1, a_2)$, entonces S tiene un sistema generador (libre) de dos elementos. Si tras la reiteración de este proceso $n - 1$ veces no hemos obtenido que $K(a_1, \ldots, a_{n-1}) = S$, entonces $(a_1, \ldots, a_{n-1})$ es libre y existe $a_n \in S - K(a_1, \ldots, a_{n-1})$

así que el sistema $(a_1, \ldots, a_{n-1}, a_n)$ es libre y maximal, luego base de V. Por tanto $S = V$ y el corolario está demostrado.

(6.32) **Corolario**: (Teorema de completar base). Si $V \neq \bar{0}$ es un espacio vectorial finitamente generado, entonces cada sistema libre de vectores puede completarse hasta obtener una base.

Demostración: Se trata de encontrar un sistema libre maximal que lo contenga.

(6.33) **Teorema**: Sea V un espacio vectorial de dimensión n. Un sistema F de n vectores de V es base si y sólo si cumple una de las siguientes propiedades:

1) F es libre;

2) F es generador.

Demostración: $\Rightarrow$) Es claro.

$\Leftarrow$) Si se cumple 1), entonces F es libre maximal y por tanto base. Si se cumple 2) F no puede ser ligado, pues en ese caso, V tendría un sistema generador de menos de n elementos, lo que no es posible por (6.26).

(6.34) **Definición**: Se llama **rango** de un sistema de vectores $(a_1, \ldots, a_n)$ de V a la dimensión del subespacio $K(a_1, \ldots, a_n)$.

Corolario: Son equivalentes:

i) $\text{rango}(a_1, \ldots, a_m) = \text{rango}(a_1, \ldots, a_m, b)$;

ii) $K(a_1, \ldots, a_m) = K(a_1, \ldots, a_m, b)$;

iii) b es combinación lineal de $(a_1, \ldots, a_m)$.

Dimensiones de subespacios.

En lo que sigue supondremos que V es un espacio vectorial de dimensión n.

(6.35) **Teorema**: Si partimos una base de V en dos subconjuntos disjuntos $B = (a_1, \ldots, a_r)$, $B' = (a_{r+1}, \ldots, a_n)$ se tiene que $K(B) \oplus K(B') = V$.

Demostración: Es claro que $V = K(B) + K(B')$, ya que si $a \in V$, entonces $a = t_1a_1 + \ldots + t_ra_r + t_{r+1} + \ldots + t_na_n \in K(B) + K(B')$.

Además si $v \in K(B) \cap K(B')$, se tendrá que $v = s_1a_1 + \ldots + s_ra_r = s_{r+1}a_{r+1} + \ldots + s_na_n$, por tanto $s_1a_1 + \ldots + s_ra_r - s_{r+1}a_{r+1} - \ldots - s_na_n = \bar{0}$. Luego $s_i = 0, \forall i = 1, \ldots, n$. En consecuencia $v = \bar{0}$.

(6.36) **Teorema**: Todo subespacio S de V posee un suplementario.

Demostración: Si $S = \bar{0}$, es obvio. Supongamos $S \neq \{\bar{0}\}$. Sea B una base de S, sabemos que puede completarse con un subconjunto B' de V hasta obtener una base de V. Por el teorema anterior $V = K(B) \oplus K(B')$, pero como $S = K(B)$, entonces $T = K(B')$ es un subespacio suplementario de S.

Nota: Dado un subespacio $\bar{0} \neq S \neq V$ siempre existen varios subespacios suplementarios de S.

(6.37) **Teorema**: (Fórmula de Grassmann de las dimensiones) Sean S y T dos subespacios de V. Entonces

$$\dim S + \dim T = \dim (S + T) + \dim (S \cap T).$$

Demostración: La fórmula vale trivialmente cuando $S \cap T = \bar{0}$. Supongamos pues que $S \cap T \neq \bar{0}$ y sea $(c_1, \ldots, c_r)$ una base de $S \cap T$. La completamos por una parte hasta obtener una base $(c_1, \ldots, c_r, a_1, \ldots, a_t)$ de S y por otra hasta obtener una base $(c_1, \ldots, c_r, b_1, \ldots, b_m)$ de T. Entonces $(c_1, \ldots, c_r, a_1, \ldots, a_t, b_1, \ldots, b_m)$ es un sistema generador de $S + T$. Además es libre pues si tenemos una combinación lineal

(**1**) $t_1c_1 + \ldots + t_rc_r + s_1a_1 + \ldots + s_ta_t + p_1b_1 + \ldots + p_mb_m = \bar{0},\ t_i, s_j, p_k \in K$

se tendría que

(**2**) $t_1c_1 + \ldots + t_rc_r + s_1a_1 + \ldots + s_ta_t = -(p_1b_1 + \ldots + p_mb_m) \in S \cap T.$,

Luego, existirán elementos $l_1, \ldots, l_r \in K$ tales que $p_1 b_1 + \ldots + p_m b_m = l_1 c_1 + \ldots + l_r c_r$. Pero como $(c_1, \ldots, c_r, b_1, \ldots, b_m)$ es libre, se obtiene en particular que $p_1 = \ldots = p_m = 0$. Volviendo a la expresión **(2)**, obtenemos

$$t_1 c_1 + \ldots + t_r c_r + s_1 a_1 + \ldots + s_t a_t = \bar{0}.$$

La independencia lineal de $(c_1, \ldots, c_r, a_1, \ldots, a_t)$ origina ahora que $t_i = 0, \forall i$ y que $s_j = 0, \forall j$. Resumiendo: $t_1 c_1 + \ldots + t_r c_r + s_1 a_1 + \ldots + s_t a_t + p_1 b_1 + \ldots + p_m b_m = \bar{0}$, $\Rightarrow p_k = 0, \forall k, t_i = 0, \forall i, s_j = 0, \forall j$. Por lo tanto $(c_1, \ldots, c_r, a_1, \ldots, a_t, b_1, \ldots, b_m)$ es una base de $S + T$. Luego

$$\dim (S + T) = r + t + m = \dim S + \dim T - \dim (S \cap T).$$

(6.38) **Corolario**: La suma de dos subespacios S y T es directa si y sólo si $\dim (S + T) = \dim S + \dim T$.

Ejercicio : Si V es de tipo finito y S es un subespacio de V, entonces V/S es también de tipo finito.

(6.39) **Teorema**: Si $([a_1], \ldots, [a_r])$ es una base del espacio vectorial cociente V/S, entonces $K(a_1, \ldots, a_r) \oplus S = V$ y $(a_1, \ldots, a_r)$ es libre.

Demostración: Si $t_1 a_1 + \ldots + \ldots + t_r a_r = \bar{0}$, entonces $[\bar{0}] = [t_1 a_1 + \ldots + \ldots + t_r a_r] = t_1[a_1] + \ldots + \ldots + t_r[a_r] = [\bar{0}] \Rightarrow t_1 = \ldots = t_r = 0$, así $(a_1, \ldots, a_r)$ es libre.

Sea $a \in V$, entonces $[a] = t_1[a_1] + \ldots + \ldots + t_r[a_r]$ para ciertos $t_i \in K$, así $[a] = [t_1 a_1 + \ldots + \ldots + t_r a_r]$, luego $a - (t_1 a_1 + \ldots + \ldots + t_r a_r) = v \in S$, por tanto $a = (t_1 a_1 + \ldots + \ldots + t_r a_r) + v \in K(a_1, \ldots, a_r) + S$. En consecuencia $V = K(a_1, \ldots, a_r) + S$. Además la suma anterior es directa, pues si $t_1 a_1 + \ldots + \ldots + t_r a_r \in S$, entonces $[t_1 a_1 + \ldots + \ldots + t_r a_r] = [\bar{0}] = \bar{0}_{V/S}$, luego $t_1[a_1] + \ldots + \ldots + t_r[a_r] = [\bar{0}]$ y por tanto $t_1 = \ldots = t_r = 0$, luego $K(a_1, \ldots, a_r) \cap S = \bar{0}$.

(6.40) **Corolario**: $\dim V/S = \dim V - \dim S$.

Ejercicio : S y T son suplementarios $\iff \dim S + \dim T = \dim (S + T) = \dim V$.

Expresiones coordenadas. Cambio de coordenadas.

A menudo es interesante, al menos desde un punto de vista formal, utilizar matrices de vectores, es decir matrices cuyas entradas sean vectores de un espacio vectorial V sobre un cuerpo K y dar la siguiente definición:

(6.41) **Definición**: Llamaremos **producto** de una matriz $m \times n$ sobre K, $A = (a_{ij})$ por una matriz $n \times r$ sobre V, $E = (v_{hk})$, a la matriz $m \times r$ sobre V cuyo elemento (i, j) es el vector $w_{ij} = \sum_{r=1}^{n} a_{ir} v_{rj}$.

Notación Si X es la n-tupla $(x_1, \ldots, x_n)$ de elementos de K ó de V indicaremos con X^t para la matriz fila $(x_1, \ldots, x_n)$ y con X para la matriz columna $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. En particular, si $(v_1, \ldots v_n)$ es una base de V el vector $v = x_1 v_1 + \ldots + x_n v_n$ podrá ser escrito mediante $v = X^t(v_i)$

Cambio de coordenadas

Sean $(a_1, \ldots, a_n)$ y $(b_1, \ldots, b_n)$ dos bases del espacio vectorial V sobre el cuerpo K. Dado un vector $a \in V$ se podrá expresar de forma única como combinación lineal de los elementos de la base (a_i) es decir, existe una única n-tupla $(x_1, \ldots x_n)$ de elementos de K de modo que $a = x_1 a_1 + \ldots + x_n a_n$ y una única n-tupla $(\bar{x}_1, \ldots, \bar{x}_n)$ de modo que $a = \bar{x}_1 b_1 + \ldots + \bar{x}_n b_n$. Es decir, matricialmente $a = X^t(a_i) = \bar{X}^t(b_j)$.

Ahora bien, cada b_j tiene la expresión única $b_j = t_{j1} a_1 + \ldots + t_{jn} a_n, j = 1, \ldots, n$, luego matricialmente:

$$\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} t_{11} & t_{12} & \ldots & t_{1n} \\ \vdots & \vdots & \ldots & \vdots \\ t_{n1} & t_{n2} & \ldots & t_{nn} \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

Llamando

$$P = \begin{pmatrix} t_{11} & t_{12} & \ldots & t_{1n} \\ \vdots & \vdots & \ldots & \vdots \\ t_{n1} & t_{n2} & \ldots & t_{nn} \end{pmatrix}$$

podemos escribir la ecuación anterior como $(b_j) = P(a_i)$, se tiene $a = \bar{X}^t(b_j) = \bar{X}^t P(a_i)$ y $a = X^t(a_i)$ y como (a_i) es base se sigue que $X^t = \bar{X}^t P$ que es la **ecuación del cambio de coordenadas**. P es la **matriz de cambio de base** ó **matriz de cambio de coordenadas**.

Lección 7: Aplicaciones lineales.

Recordemos el concepto K-homomorfismo ó aplicación K-lineal entre espacios vectoriales:

(7.1) **Definición**: Una aplicación $f : V \longrightarrow W$ siendo V y W dos K-espacios vectoriales se dice una **aplicación lineal** u homomorfismo de espacios vectoriales si

$f(a+b) = f(a) + f(b)$ y $f(ta) = tf(a), \forall a, b \in V, \forall t \in K$

(7.2) **Ejemplos**.

a) Si S es un subespacio de V, la aplicación de V al espacio cociente V/S dada por:
$p : V \longrightarrow V/S$
$\quad a \mapsto [a]$
es una aplicación lineal llamada proyección canónica de V sobre V/S

b) Las aplicaciones de $V \times W$ en V, W dadas por: $(a,b) \mapsto a$ y $(a,b) \mapsto b$ respectivamente, son aplicaciones lineales.

c) La aplicación $f : K \longrightarrow K$ dada por $a \mapsto 5a := a+a+a+a+a$ es lineal.

d) Si V_1, V_2, V_3 son K espacios vectoriales y $f : V_1 \longrightarrow V_2$ y $g : V_2 \longrightarrow V_3$ son aplicaciones lineales, entonces la composición de ambas es una aplicación lineal.

(7.3) **Notas**: a) Si $f : V \longrightarrow W$ es una aplicacción lineal, entonces

i) $f(\bar{0}) = \bar{0}$, pues $f(\bar{0}) = f(\bar{0}+\bar{0}) = f(\bar{0}) + f(\bar{0})$, luego $f(\bar{0}) = \bar{0}$;

ii) $f(-a) = -f(a)$, pues $-f(a) = (-1)f(a) = f((-1)a) = f(-a)$.

b) Si $(a_1, \ldots, a_m)$ es ligado, entonces $(f(a_1), \ldots, f(a_m))$ es ligado, pues si existen $t_1, \ldots, t_m$, alguno distinto de 0 tales que $\bar{0} = t_1a_1 + \ldots + t_ma_m$, se tiene que $\bar{0} = f(\bar{0}) = f(t_1a_1 + \ldots + t_ma_m) = t_1f(a_1) + \ldots + t_mf(a_m)$.

c) La afirmación anterior es equivalente a:

$(f(a_1), \ldots, f(a_m))$ libre implica que $(a_1, \ldots, a_m)$ es libre.

Determinación y existencia.

(7.4) **Teorema**: Si $(a_1, \ldots, a_m)$ es un sistema generador de V y $(c_1, \ldots, c_m)$ es un sistema de vectores de W entonces existe a lo sumo una aplicación lineal $f : V \longrightarrow W$ tal que $f(a_i) = c_i, i = 1, \ldots, m$.

Demostración: Si existieran dos aplicaciones lineales f y g tales que $f(a_i) = c_i$ y $g(a_i) = c_i, i = 1, \ldots, m$, entonces dado un elemento cualquiera $a \in V$,existen elementos $t_1, \ldots, t_m \in K$, tales que $a = t_1a_1 + \ldots + t_ma_m$, luego $f(a) = t_1f(a_1) + \ldots + t_mf(a_m) = t_1c_1 + \ldots + t_mc_m = t_1g(a_1) + \ldots + t_mg(a_m = g(a)$, luego $f = g$.

(7.5) **Teorema**: Si $(a_1, \ldots, a_n)$ es base de V y $(c_1, \ldots, c_n)$ es un sistema de vectores de W, entonces existe una y solo una aplicación lineal $f : V \longrightarrow W$ tal que $f(a_i) = c_i, i = 1 \ldots, n$.

Demostración: Sabemos por (7.4) que si existe es única. Dado $a \in V$, a determina la n-tupla $(x_1, \ldots, x_n)$, tal que $a = x_1a_1 + \ldots + x_na_n$. Definimos $f(a) = x_1c_1 + \ldots + x_nc_n$. Entonces f es lineal, pues si $b = y_1a_1 + \ldots + y_na_n$, entonces $f(b) = y_1c_1 + \ldots + y_nc_n$, por otra parte, como $a+b = (x_1+y_1)a_1+\ldots+(x_n+y_n)a_n$, es $f(a+b) = (x_1+y_1)c_1+\ldots+(x_n+y_n)c_n = f(a) + f(b)$; además, para cualquier $t \in K$ es $f(t.a) = f(tx_1a_1 + \ldots + tx_na_n) = tx_1c_1 + \ldots + tx_nc_n = tf(a)$.

Expresión coordenada de una aplicación lineal

Sea $f : V \longrightarrow W$ una aplicación lineal, $(a_1, \ldots, a_n)$ una base de V y $(b_1, \ldots, b_m)$ una base de W. Sea $a \in V$, existe una única n-tupla $(x_1, \ldots, x_n)$ de elementos de K, tales que $a = x_1a_1 + \ldots + x_na_n$. Análogamente, como $f(a) \in W$, existe una única m-tupla $(y_1, \ldots, y_m)$ de elementos de K, tales que $f(a) = y_1b_1 \ldots + y_mb_m$. Por otra parte, como f es lineal, se tiene que $f(a) = x_1f(a_1) + \ldots + x_nf(a_n)$. Vamos a relacionar las coordenadas (x_i) de a con las (y_i) de $f(a)$. Para ello tengamos en cuanta que f queda definida por $f(a_1), \ldots, f(a_n)$ y que $f(a_i) = t_{i1}b_1 + \ldots + t_{im}b_m, i = 1, \ldots, n$, lo que matricialmente se expresa:

$$\begin{pmatrix} f(a_1) \\ \vdots \\ f(a_n) \end{pmatrix} = \begin{pmatrix} t_{11} & t_{12} & \ldots & t_{1m} \\ \vdots & \vdots & \ldots & \vdots \\ t_{n1} & t_{n2} & \ldots & t_{nm} \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} = A \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

O más brevemente, $(f(a_i)) = A(b_j)$.

En cuanto a la relación de coordenadas:

$$f(a) = X^t(f(a_i)) = X^tA(b_j) = Y^t(b_j),$$

luego como (b_j) es base de W, se tiene que $Y^t = X^t A$ (o equivalentemente $Y = A^t X$), es decir

$$(\, y_1 \qquad y_m \,) = (\, x_1 \quad \ldots \quad x_n \,) \begin{pmatrix} t_{11} & t_{12} & \ldots & t_{1m} \\ \vdots & \vdots & \ldots & \vdots \\ t_{n1} & t_{n2} & \ldots & t_{nm} \end{pmatrix}.$$

Así, A es la matriz cuya fila i-esima esta formada por las coordenadas de la imagen del i-ésimo vector de la base, es decir, las coordenadas de $f(a_i)$ expresadas en la base $(b_1, \ldots, b_m)$.

La matriz A es llamada la **matriz coordenada** de la aplicación lineal f referida a las bases (a_i) de V y (b_j) de W.

(7.6) **Nota**: Dada cualquier matriz $A = (t_{ij})$, $n \times m$, sobre el cuerpo K y dos espacios vectoriales sobre K, V de dimensión n, con base $(a_1, \ldots, a_n)$ y W de dimensión m y base $(b_1, \ldots, b_m)$, existe una única aplicación lineal $f : V \longrightarrow W$ que tiene a A como matriz coordenada en esas bases, a saber la aplicación definida por $f(a_i) = t_{i1}b_1 + \ldots + t_{im}b_m$.

Como consecuencia, tenemos definida una aplicación:

$$\psi : \mathrm{Hom}(V, W) \longrightarrow \mathrm{M}(n \times m; K)$$

dada por

$$f \mapsto A = \text{Matriz coordenada de } f \text{ en las bases } (a_i), (b_j).$$

Por el teorema anterior, la aplicación ψ es inyectiva y por la nota la aplicacion ψ es suprayectiva.

Imágenes y antiimágenes en una aplicación lineal.

Sea $f : V \longrightarrow W$ una aplicación lineal.

(7.7) **Teorema**: a) Si S es un subespacio de V, entonces $f(S)$ es un subespacio de W.

b) Si T es un subespacio de W, entonces $f^{-1}(T)$ es un subespacio de V.

Demostración: a) $f(S) \neq \emptyset$, pues $\bar{0} = f(\bar{0}) \in f(S)$. Sea $w_1, w_2 \in f(S)$, entonces existen elementos $v_1, v_2 \in S$ tales que $w_i = f(v_i)$, por tanto $w_1 - w_2 = f(v_1) - f(v_2) =$

$f(v_1 - v_2) \in f(S)$. Además $\forall t \in K$, $tw_1 = tf(v_1) = f(tv_1) \in f(S)$. Así $f(S)$ es un subespacio de W

b) Recordar que $f^{-1}(T) = \{v \in V | f(v) \in T\}$. Como $\bar{0} \in f^{-1}(T)$, es $f^{-1}(T) \neq \emptyset$. Sean $a, b \in f^{-1}(T)$, entonces $a - b \in f^{-1}(T)$ ya que $f(a - b) = f(a) - f(b) \in T$ y $\forall t \in K$, $ta \in f^{-1}(T)$ puesto que $f(ta) = tf(a) \in T$.

(7.8) **Definición**: Se llama núcleo de una aplicación lineal f de V en W al subespacio $f^{-1}\{\bar{0}\} = \{a \in V | f(a) = \bar{0}\}$. Usualmente denotado por Ker f.

Observar que $a - b \in$ Ker $f \Leftrightarrow f(a - b) = \bar{0} \Leftrightarrow f(a) = f(b)$.

Asímismo notar que f es inyectiva si y solo si Ker $f = \bar{0}$.

(7.9) **Definición**: Se llama **rango** de la aplicación lineal f y se escribe rang f a la dimensión del subespacio Im f.

(7.10) **Notas**: 1) Si $(a_1, \ldots, a_m)$ es una familia de vectores de V se tiene que

1) $f(K(a_1, \ldots, a_m)) = K(f(a_1), \ldots, f(a_m))$

2) f es suprayectiva si y sólo si rang f= dim W.

3) Si $(a_1, \ldots, a_m)$ es un sistema generador de V, rang f=rang $(f(a_1), \ldots, f(a_m)$, ya que, por la nota 1) $f(V) = K((f(a_1), \ldots, f(a_m))$.

(7.11) **Teorema**: Las siguientes afirmaciones son equivalentes:

i) Ker $f = \bar{0}$.

ii) Todo sistema libre de V tiene por imagen un sistema libre de W.

iii) Para cualquier subespacio S de V, se tiene que dim S= dim $f(S)$.

iv) Cualesquiera que sean $a_1, \ldots, a_m$ de V, se cumple que :

rang $(a_1, \ldots, a_m)$=rang $(f(a_1), \ldots, f(a_m))$.

Demostración: i)$\Rightarrow$ ii) Sea $(a_1, \ldots, a_m)$ un sistema libre de V, supongamos que $t_1 f(a_1) + \ldots + t_m f(a_m) = \bar{0}$, entonces $f(t_1 a_1 + \ldots + t_m a_m) = \bar{0}$, luego $t_1 a_1 + \ldots + t_m a_m \in$ Ker $f = \bar{0}$. Por tanto $t_1 = \ldots = t_m = 0$, es decir $(f(a_1), \ldots, f(a_m))$ es libre.

ii) $\Rightarrow$ iii) Es obvio.

iii $\Rightarrow$ iv) rang $(a_1, \ldots, a_m) =$ dim $K(a_1, \ldots, a_m) =$ dim $f(K(a_1, \ldots, a_m)) =$

$$= \text{dim } K(f(a_1), \ldots, f(a_m)) = \text{rang } (f(a_1), \ldots, f(a_m)).$$

iv) $\Rightarrow$ i) Si $a \neq \bar{0}$, entonces 1 =rang (a) = rang $(f(a))$, así $f(a) \neq \bar{0}$ luego Ker $f = \bar{0}$.

(7.12) **Corolario**: Dos K-espacios vectoriales V y W son isomorfos si y sólo si tiene la misma dimensión.

Demostración: Si existe $f : V \longrightarrow W$ isomorfismo, entonces $f(V) = W$ y Ker $f = \bar{0}$, así que dim (V) =dim $f(V)$ =dim W.

Recíprocamente, si dim (V) = dim $W = n$ y sean $(a_1, \ldots, a_n), (b_1, \ldots, b_n)$ son bases de V y W respectivamente, existe una única aplicación lineal $f : V \longrightarrow W$ cumpliendo que $f(a_i) = b_i, i \in [1, n]$, y se verifica i) de (7.11). Por lo tanto f es isomorfismo.

Teoremas de isomorfía.

(7.13) **Teorema**: Sea $f : V \longrightarrow W$ una aplicación lineal. Entonces $\tilde{f} : V/\text{Ker } f \longrightarrow f(V)$ dada por $[a] = a + \text{Ker } f \mapsto f(a)$ es un isomorfismo de espacios vectoriales.

Demostración: $[a] = [b] \iff a - b \in \text{Ker } f \iff f(a) = f(b)$. Así $\tilde{f}$ está bien definida y es inyectiva. Obviamente es suprayectiva, así es una aplicación biyectiva. Además, $\tilde{f}$ es lineal, pues:

$$\tilde{f}([a] + [b]) = \tilde{f}([a + b]) = f(a + b) = f(a) + f(b) = \tilde{f}([a]) + \tilde{f}([b]); \text{ y}$$
$$\tilde{f}(t[a]) = \tilde{f}([ta]) = f(ta) = tf(a) = t\tilde{f}([a]).$$

Luego $\tilde{f}$ es un isomorfismo de espacios vectoriales.

(7.14) **Corolario**: rang f = dim $f(V)$ = dim V − dim Ker$f$.

(7.15) **Teorema**: Sean S y T subespacios de un mismo K-espacio vectorial V. Entonces los espacios cocientes $S + T/S$ y $T/S \cap T$ son isomorfos.

Demostración: Como $S \subseteq S + T$ y ambos son subespacios de V, es S un subespacio de $S + T$ y podemos considerar el espacio vectorial cociente $S + T/S$. Definamos

$$\varphi : T \longrightarrow S + T/S \quad \text{mediante}$$
$$b \mapsto b + S$$

Notar que φ es una aplicación lineal, pues es la restricción a T de la proyección canónica $p : S + T \longrightarrow S + T/S$. Además es suprayectiva, pues si $a \in S$, $b \in T$, $a + b + S = b + S$, pues $(a + b) - b = a \in S$. Ker $\varphi = \{b \in T | b + S = \bar{0} + S = S\} = S \cap T$.

Aplicando ahora (7.13) se obtiene que $T/S \cap T$ es isomorfo a $S + T/S$.

El espacio vectorial de las aplicaciones lineales de un espacio en otro. El álgebra de los endomorfismos de un espacio vectorial.

Dados V y W dos K-espacios vectoriales, denotamos por $\mathrm{Hom}_K(V,W)$ para el conjunto de las aplicaciones lineales de V en W.

(7.16) **Teorema**: $\mathrm{Hom}_K(V,W)$ dotado de las operaciones:

$\varphi : \mathrm{Hom}_K(V,W) \times \mathrm{Hom}_K(V,W) \longrightarrow \mathrm{Hom}_K(V,W)$ mediante
$(f,g) \mapsto f+g$
donde $(f+g)(a) = f(a)+g(a), \forall a \in V$ y

$\psi : K \times \mathrm{Hom}_K(V,W) \longrightarrow \mathrm{Hom}_K(V,W)$ mediante
$(t,f) \mapsto tf$
donde $(tf)(a) = tf(a)$, pasa a ser un K-espacio vectorial.

Demostración: Observar en primer lugar que $f+g$ es aplicación lineal de V en W, pues $(f+g)(a+b) = f(a+b)+g(a+b) = f(a)+f(b)+g(a)+g(b) = (f+g)(a)+(f+g)(b)$ y $(f+g)(t.a) = f(t.a)+g(t.a) = tf(a)+tg(a) = t((f+g)(a))$.

Análogamente se comprueba que tf es lineal: $(tf)(a+b) = t(f(a+b)) = t(f(a)+f(b)) = tf(a)+tf(b) = (tf)(a)+(tf)(b)$ y $(tf)(s.a) = tf(s.a) = t(sf(a)) = ts(f(a)) = st(f(a) = s(tf)(a)$.

Notar que $+$ es asociativa, conmutativa, existe un neutro $h_0 : V \longrightarrow W$ dado por $h_0(v) = \bar{0}_W, \forall v \in V$, el homomorfismo nulo. Además, dada $f : V \longrightarrow W$ existe $-f$ tal que $f+(-f) = h_0$, dada por $(-f) := -f(a)$. El resto queda como ejercicio.

En particular, $\mathrm{Hom}_K(V,V)$ es un K-espacio vectorial que se designa por $\mathrm{End}_K(V)$. Pero en $\mathrm{End}_K(V)$ hay otra operación, la composición de aplicaciones de forma que

(7.17) **Teorema**: $\mathrm{End}_K(V)$ es un anillo con unidad. Los elementos inversibles de este anillo son los automorfismos de V.

Demostración: La demostración de la primera afirmación es rutinaria.

Los elementos inversibles de $\mathrm{End}_K(V)$ son los $f \in \mathrm{End}_K(V)$ tales que existe $g \in \mathrm{End}_K(V)$ verificando que $f \circ g = g \circ f = 1_V$. Si la composición de dos aplicaciones es inyectiva, la que primero actúa es inyectiva y si la composición es suprayectiva, la que segundo actúa es suprayectiva. Se sigue que los elementos inversibles de $\mathrm{End}_K(V)$ son automorfismos de V. Recíprocamente, los automorfismos de V son elementos inversibles

de $\mathrm{End}_K(V)$, pues si f es un automorfimo de V, f es biyectiva, existe f^{-1} la aplicación inversa que será también lineal y se verifica $f \circ f^{-1} = f^{-1} \circ f = 1_V$.

(7.18) **Definición**: Los automorfismo de un espacio vectorial forman un grupo que se conoce como el **grupo general lineal** de V y es denotado por $\mathrm{GL}(V)$.

(7.19) **Definición**: Un álgebra asociativa sobre un cuerpo K es una estructura $(A, +, ., *)$ de modo que $(A, +, .)$ es un K espacio vectorial, $(A, +, *)$ es un anillo y

$$t.(a * b) = (t.a) * b = a * t.b, \forall t \in K, a, b \in A.$$

(7.20) **Corolario**: $\mathrm{End}_K(V)$ es una K-álgebra asociativa.

Lección 8. Espacio vectorial dual de uno dado.

Seguimos considerando espacios vectoriales V de tipo finito sobre un cuerpo K.

Consideremos a K como K-espacio vectorial.

(8.1) **Definición**: Una aplicación lineal de V en K, se dice una **forma lineal** sobre V y al espacio vectorial $\mathrm{Hom}_K(V,K)$, se le llama **espacio dual** de V y se escribe V^*. Al espacio vectorial $(V^*)^* := V^{**}$, se le llama **espacio bidual** de V.

(8.2) **Teorema**: i) dim (V) =dim (V^*); ii) La aplicación $\varphi : V \longrightarrow V^{**}$ dada por $\varphi(a) : V^* \longrightarrow K$ donde $\varphi(a)(f) := f(a)$ es un isomorfismo de espacios vectoriales.

Demostración: Sea $(a_1,\ldots,a_n)$ una base de V. Definamos $f_i \in V^*$ mediante $f_i(a_j) = \delta_{ij}, i = 1,\ldots,n$. Veamos que $(f_1,\ldots,f_n)$ es una base de V^*. Una forma lineal $f \in V^*$, queda determinada por $f(a_1) = t_1,\ldots,f(a_n) = t_n, t_i \in K$. Considerar la forma lineal $t_1f_1 + \ldots + t_nf_n$, entonces $(t_1f_1 + \ldots + t_nf_n)(a_i) = t_i = f(a_i)$. Luego $f = t_1f_1 + \ldots + t_nf_n$, por tanto, $(f_1,\ldots,f_n)$ es un sistema generador de V^*. Además $(f_1,\ldots,f_n)$ es libre, pues si existiera una combinación lineal $x_1f_1 + \ldots + x_nf_n = \bar{0}_{V^*}$, se tendría $(x_1f_1 + \ldots + x_nf_n)(a_i) = x_i = \bar{0}_{V^*}(a_i) = 0, \forall i$. Así $(f_1,\ldots,f_n)$ es una base de V^* y por tanto dim (V) =dim (V^*). A la base $(f_1,\ldots,f_n)$ de V^* se le dice **base dual** de la base (a_i) de V.

ii) Definir:

$$\begin{array}{rccl} \varphi & : & V & \to & V^{**} \\ & & a & \mapsto & \varphi(a) \quad : V^* \to \quad K \\ & & & & f \mapsto \quad f(a) \end{array}$$

Notar que $\varphi(a) \in V^{**}$, pues:

$$\varphi(a)(f+g) = (f+g)(a) = f(a) + g(a) = \varphi(a)(f) + \varphi(a)(g)$$
$$\varphi(a)(tf) = (tf)(a) = tf(a) = t\varphi(a)(f).$$

Es fácil comprobar que φ es lineal. Veamos que φ es inyectiva. Para ello analicemos Ker (φ): si $a \in$ Ker φ, entonces $\varphi(a) = \bar{0}_{V^{**}}$, asi que $\varphi(a)(f) = 0, \forall f \in V^*$, es decir, $f(a) = 0, \forall f \in V^*$. Si a no fuese $\bar{0}$, podríamos formar una base de V con a como primer elemento . Considerando el primer elemento de su base dual, obtendríamos una forma lineal $f \in V^*$, tal que $f(a) \neq 0$, contradicción. Luego Ker $\varphi = \bar{0}$ luego $n =$ dim $(V) =$ dim $\varphi(V)$. Por otra parte $n =$ dim $(V^*) =$ dim (V^{**}), luego $\varphi(V) = V^{**}$ y φ es un isomorfismo.

Aplicación lineal dual de una dada.

Sean V y W K-espacios vectoriales con dim $(V) = n$ y dim $(W) = m$.

Dada una aplicación lineal $h : V \longrightarrow W$, queda definida $h^* : W^* \longrightarrow V^*$, mediante $h^*(g) = g \circ h$. Notar que $h^*(g_1 + g_2) = (g_1 + g_2) \circ h = g_1 \circ h + g_2 \circ h = h^*(g_1) + h^*(g_2)$ y $h^*(tg) = (tg) \circ h = t(g \circ h) = th^*(g)$. Luego h^* es una aplicación lineal, y se dice **aplicación dual** de h.

(8.3) **Lema** : Sean $(a_i)_1^n$ y $(b_j)_1^m$ bases respectivas de V y W. Sean $(f_i)_1^n$ y $(g_j)_1^m$ las correspondientes bases duales de V^* y W^*. Sea h una aplicación lineal entre V y W. Entonces, la matriz coordenada de h^* en las bases $(g_j)_1^m$ y $(f_i)_1^n$, es la traspuesta de la matriz coordenada de h en las bases $(a_i)_1^n$ y $(b_j)_1^m$.

Demostración: Sea $A = (t_{ij})$ la matriz coordenada de h en las bases (a_i) y (b_j). Esto significa que $h(a_i) = t_{i1}b_1 + \ldots + t_{im}b_m, i = 1, \ldots, n$. Supongamos que $h^*(g_j) = s_{j1}f_1 + \ldots + s_{ji}f_i + \ldots + s_{jn}f_n$, evaluando $h^*(g_j)(a_i)$ según esta expresión, tenemos $h^*(g_j)(a_i) = s_{ji}$. Por otra parte $h^*(g_j)(a_i) = (g_j \circ h)(a_i) = g_j(h(a_i)) = g_j(t_{i1}b_1 + \ldots + t_{im}b_m) = t_{ij}$. Luego $s_{ji} = t_{ij}$. Es decir, $h^*(g_j) = t_{1j}f_1 + \ldots + t_{nj}f_n$, así que la j-ésima fila de la matriz coordenada de h^* es la j-ésima columna de la matriz A. Por tanto, la matriz coordenada de h^* en las bases (g_j) y (f_i), es A^t.

(8.4) **Lema**: rang h = rang h^*.

Demostración: rang $h^* = \dim(W^*) - \dim(\mathrm{Ker}(h^*))$. Por otra parte,

$$\mathrm{Ker}(h^*) = \{g \in W^* | g \circ h = \bar{0}_{V^*}\} = \{g \in W^* | g(h(v)) = 0, \forall v \in V\}.$$

Si $h(V) = \{\bar{0}_W\}$, entonces $\mathrm{Ker}(h^*) = W^*$ y en este caso rang h = rang $h^* = 0$. Si $h(V) \neq \{\bar{0}_W\}$, sea $(c_1, \ldots, c_r)$ una base de $h(V)$, que completamos hasta obtener una base $(c_1; \ldots, c_r, c_{r+1}, \ldots, c_m)$ de W. Sea $(g_j)_1^m$ su dual en W^*. Es claro que $g_{r+1}, \ldots . g_m$ pertenecen a $\mathrm{Ker}(h^*)$. Sea $g \in \mathrm{Ker}(h^*)$, entonces si $g = t_1g_1 + \ldots + t_mg_m$ y $g(c_i) = t_i = 0,\ i = 1, \ldots, r$. Se sigue que $g = t_{r+1}g_{r+1} + \ldots + t_mg_m$. Luego $\dim(\mathrm{Ker}(h^*)) = m - r = \dim(W^*) - \dim(h(V))$. Es decir, rang h = rang h^*.

(8.5) **Definición**: Sea A una matriz $n \times m$ sobre K. Se define el **rango de filas** de A como rang $(A_1, \ldots, A_n)$, donde A_i representa la fila i-ésima de A considerada como vector

de K^m. Análogamente se define el **rango de columnas** de A como rang $(A^1, \ldots, A^m)$ donde A^i es la columna i-ésima de A considerada como vector de K^n.

(8.6) **Teorema**: Sean $(a_i)_1^n$ y $(b_j)_1^m$ bases respectivas de los K-espacios vectoriales V y W. Sea $f : V \longrightarrow W$ una aplicación lineal de expresión coordenada $Y^t = X^t A$ en esas bases. Entonces: rango de filas de A = rang f

Demostración: Por (7.10) (3), rang f = rang $(f(a_1), \ldots, f(a_n))$. Sea $g : W \longrightarrow K^m$ el sistema coordenado de W asociado a la base (b_j) de W. Recordar, (6.22) , que es un isomorfimo de W en K^m, que asocia a cada vector $w \in W$ su m-tupla coordenada, por tanto $g(f(a_i)) = A_i$. En consecuencia, por (7.11) se tiene que rang $(f(a_1), \ldots, f(a_n))$ = rang $(A_1, \ldots, A_n)$.

(8.7) **Nota**: Como consecuencia de (8.6), si A y B son matrices coordenadas de una misma $f : V \longrightarrow W$, entonces rango de filas de A =rango de filas de B = rang f.

(8.8) **Corolario**: Para cualquier matriz A sobre un cuerpo K, se tiene que rango de filas de A= rango de columnas de A. A dicho número le llamaremos **rango** de A.

Demostración: Sea $A \in \text{Mat}(n \times m, K)$. Sean V y W K-espacios vectoriales de dimensiones n y m respectivamente (considerar por ejemplo K^n y K^m). Elegir bases (a_i) y (b_j) de V y W respectivamente. A define una aplicación lineal, $h : V \longrightarrow W$ de expresión coordenada $Y^t = X^t A$ en esas bases. Por (8.6) rang h = rango de filas de A. Tomar (f_i) y (g_j) bases de V^* y W^*, duales de las (a_i) y (b_j). Por (8.3) la matriz coordenada de $h^* : W^* \longrightarrow V^*$ en esas bases es A^t, cuyas filas son $(A^1, \ldots, A^m)$ las columnas de A. Por (8.6) es rang h^* = rang $(A^1, \ldots, A^m)$= rango de columnas de A. Finalmente, por (8.4) es rang h = rang h^*. Luego rango de filas de A= rango de columnas de A.

Lección 9. Matrices

En la lección 7 hemos probado que si V y W son dos K-espacios vectoriales de dimensiones n y m respectivamente, entonces los espacios vectoriales $\mathrm{Hom}_K(V,W)$ y $\mathrm{M}(n\times m;K)$ son biyectivos mediante la aplicación

$$\psi\ :\ \mathrm{Hom}_K(V,W)\longrightarrow \mathrm{M}(n\times m,K)$$

dada por

$f\mapsto A=$ Matriz coordenada de f en las bases $(a_i),(b_j):=\mathrm{M}(f,(a_i),(b_j))$.

Es más :

(9.1) **Teorema**. La aplicación anterior es un isomorfismo de espacios vectoriales.

Demostración: Sean $f_1\in\mathrm{Hom}_K(V,W)$, $A_1=\mathrm{M}(f_1,(a_i),(b_j))=(t_{ij})$, ello significa que $f(a_i)=t_{i1}b_1+\ldots+t_{im}b_m$, $f_2\in\mathrm{Hom}_K(V,W)$ con $A_2=\mathrm{M}(f_2,(a_i),(b_j))=(s_{ij})$, y $t\in K$, entonces $(f_1+f_2)(a_i)=(t_{i1}+s_{i1})b_{i1}+\ldots+(t_{im}+s_{im})b_{im}$ y $tf_1(a_i)=tt_{i1}b_1+\ldots+tt_{im}b_m$,$\forall i=1,\ldots,n$, por tanto

$\mathrm{M}((f_1+f_2),(a_i),(b_j))=(t_{ij}+s_{ij})=A_1+A_2$ y $\mathrm{M}((tf_1),(a_i),(b_j))=(t.t_{ij})=tA_1$.

(9.2) **Proposición**: Sean V,W y Z espacios vectoriales de dimensiones n,m y t y bases $(a_i),(b_j)$ y (c_k) respectivamente. Sean $f\in\mathrm{Hom}_K(V,W)$ y $g\in\mathrm{Hom}_K(W,Z)$ con $A=\mathrm{M}(f,(a_i),(b_j))$ y $B=\mathrm{M}(g,(b_j),(c_k))$, entonces es $AB=\mathrm{M}(g\circ f,(a_i),(c_k))$.

Demostración: Si $A=(t_{ij})\in\mathrm{M}(n\times m,K)$ y $B=(s_{ij})\in\mathrm{M}(m\times t,K)$, entonces:

$f(a_i)=t_{i1}b_1+\ldots+t_{im}b_m$ y $g(b_j)=s_{j1}c_1+\ldots+s_{jt}c_t$, luego

$g(f(a_i)=t_{i1}g(b_1)+\ldots+t_{im}g(b_m)=t_{i1}(s_{11}c_1+\ldots+s_{1t}c_t)+\ldots+t_{im}(s_{m1}c_1+\ldots+s_{mt}c_t)=(t_{i1}s_{11}+\ldots+t_{im}s_{m1})c_1+\ldots+(t_{i1}s_{1t}+\ldots+t_{im}s_{mt})c_t$, así que la i-ésima fila de la matriz coordenada de $g\circ f$ en las bases (a_i) y (c_k) es

$(t_{i1}s_{11}+\ldots+t_{im}s_{m1}),\ldots,(t_{i1}s_{1t}+\ldots+t_{im}s_{mt})$,

por tanto la entrada ij de $M(g\circ f,(a_i),(c_k))$ es $t_{i1}s_{1j}+\ldots+t_{im}s_{mj}$, que es la entrada (i,j) de AB.

Nota: Si $\mathbf{R}$ es un anillo conmutativo y $A\in\mathrm{M}(n\times m,\mathbf{R})$ y $B\in\mathrm{M}(m\times t,\mathbf{R})$, entonces $(AB)^t=B^tA^t$. (En esta relación es fundamental el que $\mathbf{R}$ sea conmutativo)

Operaciones elementales con matrices. Matrices elementales

Recordar que si $(a_1, \ldots, a_m)$ es un sistema de vectores en un espacio vectorial V, se obtiene un sistema equivalente mediante cualquiera de las operaciones siguientes:

1) permutación de dos vectores del sistema.

2) sustitución de un vector a_i por $a_i + ta_j$, $t \in K$.

3) sustitución de a_i por sa_i siendo $0 \neq s \in K$.

Dada una matriz $A \in \mathrm{M}(n \times m, K)$ se pueden considerar el sistema de vectores $(A_1, \ldots, A_n)$ de K^m de las filas de A y el sistema de vectores $(A^1, \ldots, A^m)$ de K^n de las columnas de A.

(9.3) **Definición**: Las operaciones 1,2 y 3 aplicadas a las filas de A ó a las columnas de A reciben el nombre de **operaciones elementales** realizadas en A.

Notar que las operaciones elementales no alteran el rango de una matriz.

La razón de definir aquí estas operaciones se debe a que la matriz obtenida realizando en A una tal operación, se obtiene también multiplicando A por una matriz cuadrada. **Notación**: Con E_{ij} denotaremos a la matriz $n \times n$ con entrada ij igual a 1 y las restantes entradas iguales a cero.

(9.4) **Definición**: Se llama **matriz elemental** a una matriz cuadrada del tipo:

$$P_{ij} = \begin{matrix} \\ i \\ j \end{matrix} \begin{pmatrix} 1 & & & & & & \\ & \ddots & \vdots & \vdots & & & \\ \ldots & \ldots & 0 & 1 & \ldots & \ldots & \\ \ldots & \ldots & 1 & 0 & \ldots & \ldots & \\ & & \vdots & \vdots & \ddots & & \\ & & & & & 1 & \end{pmatrix}; P_{ij}(t) = \begin{matrix} \\ i \end{matrix} \begin{pmatrix} 1 & & & \vdots & & \\ & \ddots & & \vdots & & \\ & \ldots\ldots & 1 & \ldots t & \ldots & \\ & & \vdots & \ddots & & \\ & & & & 1 & \end{pmatrix};$$

$$Q_i(s) = \begin{matrix} \\ i \end{matrix} \begin{pmatrix} 1 & & & \vdots & & \\ & \ddots & & \vdots & & \\ & & 1 & \vdots & & \\ \ldots & \ldots & \ldots & s & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix}, s \neq 0$$

Escritas de otra forma: $P_{ij} = I_n - E_{ii} - E_{jj} + E_{ij} + E_{ji}$; $P_{ij}(t) = I_n + tE_{ij}$; $Q_i(s) = I_n - E_{ii} + s.E_{ii}$.

(9.5) **Teorema**: $\mathrm{Mat}(n \times n, K)$ es una K-álgebra asociativa isomorfa a $\mathrm{End}_K(V)$, siendo V un K-espacio vectorial de dimensión n.

(9.6) **Definición**: Las **matrices regulares** ó elementos inversibles del anillo $\mathrm{Mat}(n \times n, K)$ forman un grupo llamado **grupo general lineal** $\mathrm{GL}(n, K)$. Dicho grupo es isomorfo con $\mathrm{GL}(V)$, siendo V un espacio vectorial de dimensión n sobre K.

(9.7) **Proposición**: Sea $(a_i)_1^n$ una base de un K-espacio vectorial V y $P \in \mathrm{Mat}(n \times n, K)$. Sea $(b_j) = P(a_i)$ y $f : V \longrightarrow V$ el endomorfismo de V definido por $f(a_i) = b_i, i = 1, \ldots, n$. Entonces son equivalentes:

i) f es automorfismo

ii) (b_j) es base de V

iii) P es regular.

Demostración: *i)* $\Rightarrow$ *ii)* Cómo f es un automorfismo, es en particular una aplicación inyectiva, luego $(\mathrm{b}_1, \ldots, b_n)$ es libre por (7.11) ii), por tanto es base de V.

ii) $\Rightarrow$ *iii)* Existe $Q \in \mathrm{Mat}(n \times n, K)$ tal que $(a_i) = Q(b_j)$. Por tanto se tiene

$$(a_i) = Q(b_j) = QP(a_i) \Rightarrow I_n = QP$$
$$(b_j) = P(a_i) = PQ(b_j) \Rightarrow I_n = PQ.$$

iii) $\Rightarrow$ *i)* Si P es regular, existe $Q \in \mathrm{Mat}(n \times n, K)$ tal que $PQ = I_n = QP$. Sea g el endomorfismo que tiene por matriz coordenada Q en la base (a_i) y ψ la aplicación que hace corresponder a un endomorfismo su matriz coordenada en la base (a_i). Entonces se tiene:

$$\psi(f).\psi(g) = PQ = QP = \psi(g).\psi(f) = I_n$$

Luego $\psi(g \circ f) = \psi(f \circ g) = \psi(1_V)$ y se sigue que $g \circ f = f \circ g = 1_V$, así f es una unidad de $\mathrm{End}_K(V)$, es decir, es un automorfismo de V.

(9.8) **Corolario**: Las matrices elementales son regulares.

Demostración: Basta observar que el efecto que producen al multiplicar por ellas los elementos de una base es: permutar dos vectores de la base; ó sustituir un vector a_i de la base por $a_i + t.a_j$; ó sustituir a_i por $s.a_i, s \neq 0$.

Ejercicio : Sea $A \in \mathrm{Mat}(n \times m, K)$. Probar:

1) El producto $P_{ij}A$ es la matriz resultante de intercambiar las filas i y j de A. El producto AP_{ij} es la matriz resultante de intercambiar las columnas i y j de A.

2) El producto $P_{ij}(t)A$, es el resultado de sustituir la fila i de A por la suma de la fila i y t por la fila j de A. El producto $AP_{ij}(t)$, es el resultado de sustituir la columna j, por

la suma de t por la columna i a la columna j.

3) El producto $Q_i(s)A$ (el producto $AQ_i(s)$), es la matriz resultante de multiplicar la fila (columna) i por el producto de s por la i.

Equivalencia de matrices.

Sean V y W K-espacios vectoriales de dimensiones n y m respectivamente. Definiremos una relación de equivalencia sobre el conjunto $\mathrm{Mat}(n \times m, K)$. Probaremos que dos matrices son equivalentes si y sólo si son matrices coordenadas de un mismo homomorfismo de V en W (naturalmente, en distintas bases). Probaremos que dos matrices son equivalentes si y sólo si tiene el mismo rango.

(9.9) **Teorema**: Sea $f : V \longrightarrow W$ una aplicación lineal. Sea $A \in \mathrm{Mat}(n \times m, K)$ la matriz coordenada de f en las bases $(a_i)_1^n$ y $(b_j)_1^m$ de V y W respectivamente. Entonces, el conjunto de todas las matrices coordenadas de f es $\{PAQ | P \in \mathrm{GL}(n, K), Q \in \mathrm{GL}(m, K)\}$.

Demostración: Sea B la matriz coordenada de f en las bases $(\bar{a}_i)$ de V y $(\bar{b}_j)$ de W. Esto significa que: (1) $(f(\bar{a}_i)) = B(\bar{b}_j)$.

Supongamos que el cambio de bases en V está dado por $(\bar{a}_i) = P(a_i)$ y que el cambio en W es $(b_j) = Q(\bar{b}_j)$, entonces,

(2) $(f(\bar{a}_i)) = P(f(a_i)) = PA(b_j) = PAQ(\bar{b}_j)$,

por tanto, como $(\bar{b}_j)$ es base de W, se tiene que $B = PAQ$.

Recíprocamente, si $P \in \mathrm{GL}(n, K)$ y $Q \in \mathrm{GL}(m, K)$, por (9.7) podemos asegurar que $(\tilde{a}_i) = P(a_i)$ y $(\tilde{b}_j) = Q^{-1}(b_j)$ son bases de V y W respectivamente y así PAQ es matriz coordenada de f.

Nota Si $f : V \longrightarrow V$ es un endomorfismo y A es la matriz coordenada de f en la base (a_i) de V, entonces $\{PAP^{-1} | P \in \mathrm{GL}(n, K)\}$ es el conjunto de todas las matrices coordenadas de f, en las diferentes bases de V. En efecto, si $(\tilde{a}_i)$ es una nueva base de V, será $(\tilde{a}_i) = P(a_i)$, con P regular, así $(f(\tilde{a}_i)) = P(f(a_i)) = PA(a_i) = PAP^{-1}(\tilde{a}_i)$, y recíprocamente.

(9.10) **Teorema**: Sean $A \in \mathrm{Mat}(n \times m, K)$, $P \in \mathrm{GL}(n, K)$ y $Q \in \mathrm{GL}(m, K)$. Entonces $\mathrm{rang}\ A = \mathrm{rang}\ PAQ$.

Demostración: Sean V un espacio vectorial con $\dim_K(V) = n$ y base (a_i) y W otro con $\dim_K(W) = m$ y base (b_j) y sea $f : V \longrightarrow W$ la aplicación lineal de ecuación $(f(a_i)) = A(b_j)$. Por (9.9) A y PAQ son matrices coordenadas de f. Por (8.6) es $\mathrm{rang}\ (A) = \mathrm{rang}\ (f) = \mathrm{rang}\ (PAQ)$.

(9.11) **Definición**: Dos matrices $A, B \in \mathrm{Mat}(n \times m, K)$ son **equivalentes** si existen $P \in \mathrm{GL}(n, K)$ y $Q \in \mathrm{GL}(m, K)$ tales que $B = PAQ$.

Observación: Por el teorema (9.9) dos matrices son equivalentes si y sólo si son matrices coordenadas de un mismo homomorfismo. Por el teorema (9.10) si dos matrices $n \times m$ sobre K son equivalentes, entonces tienen el mismo rango.

(9.12) **Teorema**: Si $A \in \mathrm{Mat}(n \times m, K)$ tiene rango $r > 0$, entonces A es equivalente a la matriz $n \times m$ suma diagonal

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$

donde I_r indica la matriz identidad $r \times r$.

Demostración: Sea $A = (a_{ij}) \neq (0)$. Por cambio de filas y columnas, podemos suponer que $a_{11} \neq 0$, incluso usando las matrices $Q_i(s)$, podemos obtener una matriz equivalente con la A cuya entrada $(1, 1)$ sea 1. A continuación, usando matrices del tipo $P_{ij}(t)$ se puede obtener una matriz equivalente con A de la forma:

$$\begin{pmatrix} 1 & \dots & \dots & \dots & 0 \\ 0 & * & \dots & * & 0 \\ \vdots & \dots & B & \dots & \vdots \\ 0 & * & \dots & * & 0 \end{pmatrix}.$$

Si la submatriz B es nula habríamos llegado al resultado. Las operaciones elementales hechas sobre B no alteran la primera fila ni la primera columna. Se llegaría con este proceso a obtener un 1 en el lugar $(2, 2)$, así hasta obtener matrices $P_1, \dots, P_s, Q_1, \dots Q_t$ elementales tales que

$$P_s \dots . P_1 . A . Q_1 \dots . Q_t = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Como las matrices elementales son regulares, se obtiene la tesis. Notar que el rango se mantiene a lo largo de todo el proceso.

Ahora podemos dar

(9.13) **Corolario**: Dos matrices $A, B \in \mathrm{Mat}(n \times m, K)$ son equivalentes si y sólo si rang $(A) =$ rang $(B) = r$ (y ambas son equivalentes a $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$.)

(9.14) **Corolario**: Sea $A \in \mathrm{Mat}(n \times n, K)$. Son equivalentes las afirmaciones :

i) A es regular;

ii) rang $(A) = n$;

iii) A es producto de matrices elementales.

Demostración: i) $\Rightarrow$ ii) Como $A = A.I_n.I_n$, si A es regular A es equivalente a I_n y por tanto tiene rango n.

ii) $\Rightarrow$ iii) Si rang $(A) = n$, por el teorema anterior existen matrices elementales $P_1, \ldots, P_s, Q_1, \ldots Q_t$ tales que $P_s. \ldots .P_1.A.Q_1. \ldots .Q_t = I_n$. Luego

$$A = P_1^{-1} \ldots . P_s^{-1}.Q_t^{-1} \ldots . Q_1^{-1}.$$

Notar que las matrices inversas de matrices elementales son también elementales.

iii) $\Rightarrow$ i) Como las matrices elementales son regulares se sigue que A es regular.

(9.14) **Teorema**: El rang (A) es igual al máximo orden de las submatrices regulares de A.

Demostración: Supongamos $r =$ rang (A) y $s = \max\{t | \exists$ una submatriz regular $t \times t$ de $A\}$.Sea B una submatriz $s \times s$ regular de A. Supongamos que $i_1, \ldots, i_s$ son los índices de las filas de B. Considerar la submatriz C de A formada por las filas $A_{i_1}, \ldots, A_{i_s}$ de A. Tiene s columnas independientes (las s columnas de B), así las s filas de C son linealmente indepedientes. Luego $r \geq s$.

Recíprocamente, como $r =$ rang (A), existen r filas de A linealmente independientes, así que la submatriz D de A formada por esas filas tiene rango r, luego D tiene r columnas linealmente independientes. La submatriz $r \times r$ de A formada por esas filas y esas columnas tiene rango r, luego $s \geq r$.

Lección 10. Formas multilineales. Determinantes

(10.1) **Definición**: Dados $r+1$ espacios vectoriales $V_1, \ldots, V_r, W$ sobre un mismo cuerpo K, se llama **aplicación r-lineal** de $V_1 \times \ldots \times V_r$ en W a una aplicación

$f : V_1 \times \ldots \times V_r \longrightarrow W$ que para cada i cumple:

a) $f(v_1, \ldots, v_i + \bar{v}_i, \ldots, v_r) = f(v_1, \ldots, v_i, \ldots, v_r) + f(v_1, \ldots, \bar{v}_i, \ldots, v_r)$,

b) $f(v_1, \ldots, tv_i, \ldots, v_r) = tf(v_1, \ldots, v_i, \ldots, v_r)$.

Nota: Ambas propiedades equivalen a

$f(v_1, \ldots, tv_i + s\bar{v}_i, \ldots, v_r) = tf(v_1, \ldots, v_i, \ldots, v_r) + sf(v_1, \ldots, \bar{v}_i, \ldots, v_r)$.

El concepto de aplicación lineal es un caso particular para $r = 1$.

(10.2) **Definición**: Se llama **función (o forma) r-lineal** sobre V a una aplicación r-lineal de V^r en K.

(10.3) **Teorema**: El conjunto $\mathrm{M}(V^r, K)$de las formas r-lineales sobre V, es un K-espacio vectorial con las operaciones:

$(f+g)(v_1, \ldots, v_r) = f(v_1, \ldots, v_r) + g(v_1, \ldots, v_r); (tf)(v_1, \ldots, v_r) = tf(v_1, \ldots, v_r)$.

Expresión coordenada de una forma r-lineal

Comenzaremos estudiando el caso $r = 2$ y $\dim_K(V) = 3$. Si (a_1, a_2, a_3) es una base de V, $v_i = x_{i1}a_1 + x_{i2}a_2 + x_{i3}a_3, i = 1, 2$, entonces $f(v_1, v_2) = f(x_{11}a_1 + x_{12}a_2 + x_{13}a_3, x_{21}a_1 + x_{22}a_2 + x_{23}a_3) = f(x_{11}a_1, x_{21}a_1 + x_{22}a_2 + x_{23}a_3) + f(x_{12}a_2, x_{21}a_1 + x_{22}a_2 + x_{23}a_3) + f(x_{13}a_3, x_{21}a_1 + x_{22}a_2 + x_{23}a_3) = f(x_{11}a_1, x_{21}a_1) + f(x_{11}a_1, x_{22}a_2) + f(x_{11}a_1, x_{23}a_3) + f(x_{12}a_2, x_{21}a_1) + f(x_{12}a_2, x_{22}a_2) + f(x_{12}a_2, x_{23}a_3) + f(x_{13}a_3, x_{21}a_1) + f(x_{13}a_3, x_{22}a_2) + f(x_{13}a_3, x_{23}a_3) = \sum_{(i,j)} x_{1i}x_{2j}f(a_i, a_j)$, donde (i, j) recorre todos los pares posibles, dando a i valores $1, 2, 3$ y lo mismo a j.

Si $\dim_K(V) = n$, y $v_i = x_{i1}a_1 + \ldots + x_{in}a_n$ y si $f : V^r \longrightarrow K$, será $f(v_1, \ldots, v_r) = \sum_{j_1, \ldots, j_r} x_{1j_1}x_{2j_2} \ldots x_{rj_r}f(a_{j_1}, a_{j_2}, \ldots, a_{j_r})$, donde $(j_1, \ldots, j_r) \in E^r$, con $E = \{1, \ldots, n\}$. Es claro que f queda determinada por las imágenes $f(a_{j_1}, a_{j_2}, \ldots, a_{j_r})$, donde $(j_1, \ldots, j_r)$ recorren $\{1 \ldots, n\}$.

Para lo que sigue es conveniente recordar:

1) Si $\alpha \in \Sigma_r$, entonces α puede expresarse como un producto de trasposiciones.

2) Si α es un producto de m trasposiciones, entonces $\text{sig}(\alpha) = (-1)^m$.

3) Si t es una trasposición, entonces sig $(t.\alpha) = -$ sig (α)

Aplicación trasformada por una permutación.

Sea $f : V^r \longrightarrow K$ arbitraria y $\alpha \in \Sigma_r$, se llama aplicación transformada de f por α a $g : V^r \longrightarrow K$ dada por $g(v_1, \ldots, v_r) = f(v_{\alpha(1)}, \ldots, v_{\alpha(r)})$ y se escribe $g = \alpha f$.

Propiedades:

1. Si f es r-lineal, entonces αf también es r-lineal:

Supongamos que $v_i = v + \bar{v}$ y que $\alpha(j) = i$, entonces

$$\alpha f(v_1, \ldots, \underset{i}{v+\bar{v}}, \ldots, v_r) = f(v_{\alpha(1)}, \ \ldots, \underset{j}{v+\bar{v}}, \ldots, v_{\alpha(r)}) =$$
$$= f(v_{\alpha(1)}, \ldots, \underset{j}{v}, \ldots, v_{\alpha(r)}) + f(v_{\alpha(1)}, \ldots, \underset{j}{\bar{v}}, \ldots, v_{\alpha(r)}) =$$
$$= \alpha f(v_1, \ldots, \underset{i}{v}, \ldots, v_r) + \alpha f(v_1, \ \ldots, \underset{i}{\bar{v}}, \ldots, v_r)$$

Análogamente se demuestra la segunda propiedad.Λ

2. La aplicación $\text{M}(V^r, K) \longrightarrow M(\text{V}^r, K)$, dada por $f \mapsto \alpha f$ es una aplicación lineal.

3. Si $\alpha, \beta \in \Sigma_r$, se tiene: $\alpha(\beta f) = (\alpha \circ \beta) f$, pues

$\alpha(\beta f)(v_1, \ldots, v_r) = \beta f(v_{\alpha(1)}, \ldots, v_{\alpha(i)}, \ldots, v_{\alpha(r)}) = (\beta f)(w_1, \ldots, w_i, \ldots, w_r)$, donde $w_1 = v_{\alpha(1)}, \ldots, w_i = v_{\alpha(i)}, \ldots, w_r = v_{\alpha(r)}$. Por tanto,

$(\beta f)(w_1, \ldots, w_i, \ldots, w_r) = f(w_{\beta(1)}, \ldots, \ldots, w_{\beta(r)})$, pero $w_{\beta(i)} = v_{\alpha\circ\beta(i)}$, luego

$\alpha(\beta f)(v_1, \ldots, v_r) = f(v_{\alpha\circ\beta(1)}, \ldots, v_{\alpha\circ\beta(i)}, \ldots, v_{\alpha\circ\beta(r)}) = (\alpha \circ \beta) f(v_1, \ldots, v_r)$.

(10.4) **Definición**: La aplicación $f : V^r \longrightarrow K$ se dice **simétrica** si $\forall \alpha \in \Sigma_r$, es $\alpha f = f$. Una aplicación $f : V^r \longrightarrow K$ es **antisimétrica** si $\forall \alpha \in \Sigma_r$, es $\alpha f = e_\alpha f$, donde e_α es la signatura de α.

(10.5) **Teorema**: La aplicación f es antisimétrica si y sólo si para cualquier trasposición $t \in \Sigma_r$ es $tf = -f$.

Demostración: $\Rightarrow$) Inmediata.

$\Leftarrow$) Sea $\alpha \in \Sigma_r$, recordar que $\alpha = t_1 \circ \ldots \circ t_m$ y sig $\alpha = (-1)^m$. Ahora bien, $\alpha f = (t_1 \circ \ldots \circ t_m) f = (t_1 \circ \ldots \circ t_{m-1})(-f)$ y así hasta $\alpha f = (-1)^m f = e_\alpha f$.

(10.6) **Definición**: La aplicación $f : V^r \longrightarrow K$ se dice **alternada** si siempre que $v_i = v_j$ con $i \neq j$, se tiene que $f(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_r) = 0$.

(10.7) **Teorema**: Sea K un cuerpo con con car $K \neq 2$. Una forma r-lineal $f : V^r \longrightarrow K$, es antisimétrica si y sólo si es alternada.

Demostración: $\Rightarrow$) Se tiene $f(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_r) =$

$(i,j)f(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_r) = -f(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_r)$, luego si $v_i = v_j$ con $i \neq j$, entonces $2f(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_r) = 0$, luego $f(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_r) = 0$.

$\Leftarrow$) Supongamos que f es alternada. Situando en los lugares i, j la misma suma $v + \bar{v}$, sería $f(v_1, \ldots, v + \bar{v}, \ldots, v + \bar{v}, \ldots, v_r) = 0$. Pero, $f(v_1, \ldots, v + \bar{v}, \ldots, v + \bar{v}, \ldots, v_r) = f(v_1, \ldots, v, \ldots, v, \ldots, v_r) + f(v_1, \ldots, v, \ldots, \bar{v}, \ldots, v_r) +$

$+f(v_1, \ldots, \bar{v}, \ldots, v, \ldots, v_r) + f(v_1, \ldots, \bar{v}, \ldots, \bar{v}, \ldots, v_r) = 0$, luego

$$f(v_1, \ldots, v, \ldots, \bar{v}, \ldots, v_r) + f(v_1, \ldots, \bar{v}, \ldots, v, \ldots, v_r) = 0, \text{ es decir}$$

$$f(v_1, \ldots, v, \ldots, \bar{v}, \ldots, v_r) = -f(v_1, \ldots, \bar{v}, \ldots, v, \ldots, v_r).$$

Nota: Si la familia $(v_1, \ldots, v_r)$ es ligada y f es una forma r-lineal alternada, entonces: $f(v_1, \ldots, v_r) = 0$.

Función determinante.

(10.8) **Definición**: Dado un espacio vectorial V con $\dim_K(V) = n$ se llama **función determinante** sobre V a una forma n-lineal y alternada sobre V.

(10.9) **Teorema**: La expresión coordenada de una función determinante D, respecto de la base $(a_1, \ldots, a_n)$ de V es : (siendo $v_i = \sum_{j=1}^{n} x_{ij} a_j$)

$$D(v_1, \ldots v_n) = D(a_1, \ldots, a_n). \sum_{\alpha \in \Sigma_n} e_\alpha x_{1\alpha(1)} \ldots . x_{n\alpha(n)}.$$

Demostración: Veamos en primer lugar el caso $n = 3$. Sabemos que

$$D(v_1, v_2, v_3) = \sum_{j_1, j_2, j_3} x_{1j_1} x_{2j_2} x_{3j_3} D(a_{j_1}, a_{j_2}, a_{j_3}),$$

donde j_1, j_2, j_3 van recorriendo de 1 a 3. Pero como D es alternada, cada vez que coincidan dos de los índices j_1, j_2, j_3 será $D(a_{j_1}, a_{j_2}, a_{j_3}) = 0$. Suprimiendo las ternas con

un elemento repetido, sólo nos quedarán las ternas $123, 132, 213, 231, 312, 321$, es decir, las permutaciones de $\{1, 2, 3\}$. Por tanto se tiene:

$$D(v_1, v_2, v_3) = \sum_{\alpha \in \Sigma_3} x_{1\alpha(1)} x_{2\alpha(2)} x_{3\alpha(3)} D(a_{\alpha(1)}, a_{\alpha(2)}, a_{\alpha(3)}) =$$

$$= D(a_1, a_2, a_3) \sum_{\alpha \in \Sigma_3} x_{1\alpha(1)} x_{2\alpha(2)} x_{3\alpha(3)} e_\alpha.$$

En el caso general,

$$D(v_1, \ldots, v_n) = \sum_{j_1, \ldots, j_n} x_{1j_1} \ldots . x_{nj_n} D(a_{j_1}, \ldots, a_{j_n}),$$

donde $j_1, \ldots, j_n$ van recorriendo de 1 a n. Pero como D es alternada, cada vez que coincidan dos de los índices $j_1, \ldots, j_n$ será $D(a_{j_1}, \ldots, a_{j_n}) = 0$. Suprimiendo las n-tuplas con un elemento repetido, sólo nos quedarán las permutaciones de $\{1, \ldots, n\}$. Por tanto se tiene:

$$D(v_1, \ldots, v_n) = \sum_{\alpha \in \Sigma_n} x_{1\alpha(1)} \ldots x_{n\alpha(n)} D(a_{\alpha(1)}, \ldots, a_{\alpha(n)}) =$$

$$= D(a_1, \ldots, a_n) \sum_{\alpha \in \Sigma_n} x_{1\alpha(1)} \ldots x_{n\alpha(n)} e_\alpha.$$

(10.10) **Corolario**: Si para una base (a_i) de V es $D(a_1, \ldots a_n) = 0$, entonces para cualquier elección de $(v_1, \ldots, v_n)$ es $D(v_1, \ldots, v_n) = 0$. Es decir D es la función determinante nula si y sólo si para una base (a_i) de V es $D(a_1, \ldots a_n) = 0$.

(10.11) **Corolario**: Si D y $\bar{D}$ son dos funciones determinantes, no nulas, sobre V, entonces son proporcionales, ya que si $(a_i)_1^n$ es una base previamente fijada de V, $\forall (v_1, \ldots, v_n)$, se tiene,

$$D(v_1, \ldots, v_n)/D(a_1, \ldots, a_n) = \bar{D}(v_1, \ldots, v_n)/\bar{D}(a_1, \ldots, a_n).$$

(10.12) **Teorema**: Dada una base (a_i) de V, la aplicación $D : V^n \longrightarrow K$ dada por $D(v_1, \ldots, v_n) = c \sum_\alpha e_\alpha x_{1\alpha(1)} \ldots x_{n\alpha(n)}$, donde $v_i = x_{i1}a_1 + \ldots + x_{in}a_n$ y c un escalar fijo, es una función determinante.

Demostración: Para probar que D es n-lineal basta observar que en cada sumando de la expresión $\sum_\alpha e_\alpha x_{1\alpha(1)} \ldots x_{n\alpha(n)}$, aparece una coordenada y sólo una de v_i.

Veamos que D es alternada. Sea σ una trasposición. Entonces $\sigma D(v_1,\ldots,v_n) = D(v_{\sigma(1)},\ldots,v_{\sigma(n)}) = c\sum_{\alpha} e_\alpha x_{\sigma(1),\alpha(1)}\ldots x_{\sigma(n),\alpha(n)}$. Reordenemos:

el primero $x_{1,\alpha(i)}$ aparece cuando $\sigma(i) = 1$, luego $i = \sigma^{-1}(1) = \sigma(1)$ y $x_{1,\alpha(i)} = x_{1,\alpha\circ\sigma(1)}$, el segundo $x_{2,\alpha(j)} = x_{2,\alpha\circ\sigma(2)}$, ..., y así queda,

$$\sigma D(v_1,\ldots,v_n) = c\sum_{\alpha} e_\alpha x_{1,\alpha\circ\sigma(1)}\ldots, x_{n,\alpha\circ\sigma(n)}.$$

Pero $e_\alpha = -e_{\alpha\circ\sigma}$ y cuando α recorre Σ_n también $\alpha\circ\sigma$ recorre Σ_n , así:

$$\sigma D(v_1,\ldots,v_n) = c\sum_{\alpha\circ\sigma}(-e_{\alpha\circ\sigma})x_{1,\alpha\circ\sigma(1)}\ldots, x_{n,\alpha\circ\sigma(n)} =$$

$$= -c\sum_{\delta=\alpha\circ\sigma} e_\delta x_{1,\delta(1)}\ldots, x_{n,\delta(n)} = -D(v_1,\ldots,v_n).$$

Notar que $D(a_1,\ldots,a_n) = c$.

(10.13) **Corolario**: (existencia y unicidad) Dada una base (a_i) de V y un escalar $c \in K$ existe una y sólo una función determinante D tal que $D(a_1,\ldots,a_n) = c$.

Demostración: Por el teorema anterior, es claro existe una. Que sólo existe una, se sigue de que la expresión coordenada en la base (a_i) está unívocamente determinada por $D(a_1,\ldots,a_n)$.

(10.14) **Corolario**: Una familia $(v_1,\ldots,v_n)$ de vectores de V es ligada si y sólo si $D(v_1,\ldots,v_n) = 0$, para alguna función determinante $D \neq 0$.

Demostración: $\Rightarrow$) Basta tener en cuenta la nota siguiente a (10.7).

$\Leftarrow$) Si la familia (v_i) no es ligada, entonces es base de V y como $D(v_1,\ldots,v_n) = 0$, D sería nula. Contradicción.

Lección 11. Determinante de una matriz cuadrada. Propiedades

Sea K un cuerpo. Aplicaremos los resultados de la lección anterior y en especial el teorema (10.12) para definir una función n-lineal alternada sobre K^n, D tal que $D(e_1, \ldots, e_n) = 1$, donde (e_i) es la base canónica de K^n.

(11.1) **Definición**: Dada una matriz $A = (t_{ij}) \in \mathrm{Mat}(n \times n, K)$, se llama **determinante** de A y se escribe $\det A$, $D(A)$ ó $|A|$ al escalar $\sum_{\alpha \in \Sigma_n} e_\alpha t_{1\alpha(1)} \ldots . t_{n\alpha(n)}$.

(11.2) **Teorema**: $\det(A) = \det(A^t)$.

Demostración: Cómo el elemento (i, j) de la matriz A^t es el elemento t_{ji} de la matriz A, se tiene:

$\det(A^t) = \sum_{\alpha \in \Sigma_n} e_\alpha t_{\alpha(1)1} . t_{\alpha(2)2} \ldots . t_{\alpha(n)n} = \sum_{\alpha \in \Sigma_n} e_\alpha t_{1\alpha^{-1}(1)} . t_{2\alpha^{-1}(2)} \ldots . t_{n\alpha^{-1}(n)}$.

Ahora bien, como sig (α) =sig (α^{-1}), se tiene:

$\det(A^t) = \sum_{\alpha \in \Sigma_n} e_{\alpha^{-1}} t_{1\alpha^{-1}(1)} . t_{2\alpha^{-1}(2)} \ldots . t_{n\alpha^{-1}(n)}$ y puesto que si α recorre Σ_n también lo recorre α^{-1}, queda: $\det(A^t) = \sum_{\sigma \in \Sigma_n} e_\sigma t_{1\sigma(1)} \ldots t_{n\sigma(n)} = \det(A)$.

Nota: Observar que $\det(A) = \det(A^t) = \sum_{\alpha \in \Sigma_n} e_\alpha t_{\alpha(1)1} . t_{\alpha(2)2} \ldots . t_{\alpha(n)n}$.

Las propiedades principales de los determinantes, se siguen del hecho siguiente:

(11.3) **Teorema**: La aplicación $D : (K^n)^n \longrightarrow K$ dada por $D(A_1, \ldots, A_n) = \det(A)$, donde A es la matriz cuyas filas son $A_1, \ldots, A_n$, es una función determinante, no nula, sobre K^n.

Demostración: Considerar la base canónica o natural de K^n, $e_1, \ldots, e_n$. Sabemos que los elementos de una n-tupla A_i son sus propias coordenadas en dicha base. Así la función determinante $\bar{D}$ sobre K^n tal que $\bar{D}(e_1, \ldots, e_n) = 1$ es precisamente la del enunciado.

Propiedades de los determinantes de matrices.

Sea $A \in \mathrm{Mat}(n \times n, K)$:

1ª) Si las filas de A son $(A_1, \ldots, A_i + \bar{A}_i, \ldots, A_n)$, entonces:

$$\det(A) = D(A_1, \ldots, A_i, \ldots, A_n) + D(A_1, \ldots, \bar{A}_i, \ldots, A_n).$$

Análogamente con columnas.

2ª) Si una fila (o columna) de A se multiplica por un escalar t, el determinante de la nueva matriz es $t.\det(A)$.

3ª) Si dos filas (o dos columnas) se permutan entre sí, el determinante de la nueva matriz es $-|A|$.

4ª) Si dos filas (o columnas) de A son proporcionales, entonces $\det(A) = 0$. En general, la familia de las filas (o la de las columnas) es ligada si y sólo si $|A| = 0$.

5ª) Si una fila (o columna) de A se sustituye por la que resulta de sumarle una combinación lineal de las demás, entonces el determinante de la nueva matriz es tambien $|A|$.

Demostración: Las propiedades 1ª) y 2ª) se siguen de que la función determinante es n-lineal. La 3ª) se sigue de que D es alternada. La 4ª) es consecuencia de (10.14) . La 5ª) es consecuencia de las propiedades 1ª),2ª y 4ª).

Como consecuencia de la propiedad 4ª), se sigue:

Nota: Una matriz A es regular si y sólo si $\det(A) \neq 0$.

(11.4) **Teorema**: Sean $A, B \in \mathrm{Mat}(n \times n, K)$. Entonces $\det AB = \det(A)\det(B)$.

Demostración: Sean $A = (a_{ij})$ y $B_1, \ldots, B_n$ las filas de B. Sabemos que la entrada ij der la matriz AB es $a_{i1}b_{1j} + \ldots + a_{in}b_{nj}$. Luego la fila i de AB es :

$$(a_{i1}b_{11} + \ldots + a_{in}b_{n1}, a_{i1}b_{12} + \ldots + a_{in}b_{n2}, , \ldots, a_{i1}b_{1n} + \ldots + a_{in}b_{nn}),$$

que puede ser reescrita como:

$$a_{i1}(b_{11} \;\; b_{12} \ldots \;\; b_{1n}) + a_{i2}(b_{21} \;\; b_{22} \ldots \;\; b_{2n}) + \ldots + a_{in}(b_{n1} \;\; b_{n2} \ldots \;\; b_{nn}) =$$
$$= a_{i1}B_1 + a_{i2}B_2 + \ldots + a_{in}B_n.$$

Así, las filas de la matriz producto AB son

$$(\sum_{j_1=1}^{n} a_{1j_1}B_{j_1}, \ldots, \sum_{j_n=1}^{n} a_{nj_n}B_{j_n}).$$

$$\text{Luego } \det(AB) = D(\sum_{j_1=1}^{n} a_{1j_1}B_{j_1}, \ldots, \sum_{j_n=1}^{n} a_{nj_n}B_{j_n}) =$$

$$= \sum_{j_1,\ldots,j_n} a_{1j_1}a_{2j_2}\ldots a_{nj_n}D(B_{j_1}, \ldots, B_{j_n}) = D(B_1, \ldots, B_n)\sum_{\alpha} e_\alpha a_{1\alpha(1)}\ldots a_{n\alpha(n)} =$$

$= \det(B)\det(A) = \det(A)\det(B)$, pues K es conmutativo.

Desarrollo por los elementos de una línea.

Sea $A = (t_{ij})$ y **(1)** $|A| = \sum e_\alpha t_{1\alpha(1)} \dots t_{n\alpha(n)}$.

Fijemos un i. Cada sumando contiene una y sólo una entrada, la $t_{i\alpha(i)}$ en la fila i. Agrupamos los términos de esta suma que contengan t_{i1}, expresamos la suma de estos sumandos sacando t_{i1} cómo factor común como $t_{i1}A_{i1}$. Hacemos lo mismo con los sumandos que contienen $t_{i2}, \dots, t_{in}$. De esta forma, tenemos

$|A| = t_{i1}A_{i1} + t_{i2}A_{i2} + \dots + t_{in}A_{in}$.

(11.5) **Definición**: El escalar A_{ij} que aparece en la expresión anterior es el **adjunto** ó **cofactor** de t_{ij} en el determinante de A. La expresión anterior se dice **desarrollo de** $|A|$ **por los elementos de la fila** i-ésima. Análogamente se obtiene el **desarrollo por los elementos de una columna**.

(11.6) **Definición**: Dado un elemento t_{ij} de A, se llama **menor complementario** de t_{ij} al determinante de la submatriz de A que resulta de suprimir de ella la fila i y la columna j. Se escribe D_{ij}.

(11.7) **Teorema**: $A_{ij} = (-1)^{i+j} D_{ij}$.

Demostración: Si $i = j = 1$, entonces los términos de **(1)** en que entra t_{11} son aquellos para los cuales $\alpha(1) = 1$. Así $A_{11} = \sum_{\alpha \in \Sigma_n, \alpha(1)=1} e_\alpha t_{2\alpha(2)} \dots t_{n\alpha(n)}$. Ahora bien, una permutación $\alpha \in \Sigma_n$ tal que $\alpha(1) = 1$ da origen a una permutación $\beta \in \Sigma_{n-1}$, mediante $\beta(2) = \alpha(2), \dots, \beta(n) = \alpha(n)$, es decir $\beta = \alpha_{|\{2,\dots,n\}}$, además $e_\beta = e_\alpha$, ya que la descomposición de α como producto de trasposiciones es la misma que la de β. Por otra parte, toda permutación de $\{2, \dots, n\}$ puede describirse como una permutación $\{1, \dots, n\}$ que fija el 1. Así $A_{11} = \sum_{\beta \in \Sigma_{n-1}} e_\beta t_{2\beta(2)} \dots t_{n\beta(n)}$. Luedo $A_{11} = D_{11}$.

Caso general: Dado t_{ij}, escribamos

$$A^* = \begin{pmatrix} t_{ij} & t_{i1} & \dots & \not{j} & \dots & t_{in} \\ t_{1j} & t_{11} & \dots & \dots & \dots & t_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \not{i} & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ t_{nj} & t_{n1} & \dots & \dots & \dots & t_{nn} \end{pmatrix}$$

donde $\not{i}, \not{j}$ indica que faltan la fila i y la columna j (las cuales están en primer lugar). Un modo de obtener A^* es aplicar $i - 1$ trasposiciones de filas a la matriz A (la i-ésima

con cada una que la precede), seguida de $j-1$ trasposiciones de columnas (la j-ésima con cada precedente), así $|A^*|$=(-1)$^{i+j}|A|$. Ahora bien, desarrollando $|A|$ y $|A^*|$, el coeficiente de t_{ij} en el primer miembro es D_{ij} y en el segundo es $(-1)^{i+j}A_{ij}$.

(11.8) **Teorema**: La suma de los productos de los elementos de una línea por los adjuntos de una paralela es igual a cero.

Demostración: Consideremos las filas i y k con $i \neq k$. Sea A^* la matriz que se obtiene sustituyendo la fila k por la i. Cómo A^* tiene dos filas iguales, es $|A^*| = 0$. Pero desarrollando $|A^*|$ por los elementos de la fila k-ésima, se tendrá: $|A^*| = t_{i1}A_{k1} + t_{i2}A_{k2} + \ldots + t_{in}A_{kn} = 0$.

(11.9) **Corolario**: $t_{i1}A_{j1}+t_{i2}A_{j2}+\ldots+t_{in}A_{jn} = \delta_{ij}|A|$. Análogamente por columnas: $t_{1i}A_{1j} + t_{2i}A_{2j} + \ldots + t_{ni}A_{nj} = \delta_{ij}|A|$

Nota: Llamando matriz **adjunta** de A a la matriz que tiene por entrada (i, j) al adjunto de t_{ij}, entrada (i, j) de A, el corolario anterior puede enunciarse escribiendo

(11.9 bis) **Corolario**: $A(\text{adj}(A))^t = (\det(A))I$.

Por tanto, si $\det(A) \neq 0$, se obtiene que la inversa de A es $A^{-1} = (1/\det(A))(\text{adj}(A))^t$.

Nota: Podemos definir el determinante de una matriz $A \in \text{Mat}(n \times n, \mathbf{R})$, donde $\mathbf{R}$ es un anillo conmutativo y con identidad, como se hizo en (11.1). En tal caso conservan su validez (11.2) y las propiedades generales. Para matrices sobre $\mathbf{R}$ no es cierto que una matriz sea regular si su determinante es distinto de cero. Se mantendrán válidos:

(11.4), el desarrrollo de un determiante por los elementos de una línea, (11.5), (11.6), (11.7), (11.8),(11.9), así como la relación $A(\text{adj}(A))^t = (\det(A))I$.

La expresión para la inversa será válida en el caso en que $\det(A)$ sea una unidad de $\mathbf{R}$:

Lección 12. Sistemas de ecuaciones

Deseamos encontrar las soluciones a un sistema de ecuaciones lineales con coeficientes en un cuerpo K. Dado el sistema:

$$(\mathbf{1}) \quad \begin{array}{c} t_{11}x_1 + \ldots + t_{1m}x_m = b_1 \\ t_{21}x_1 + \ldots + t_{2m}x_m = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ t_{n1}x_1 + \ldots + t_{nm}x_m = b_n, \end{array}$$

se puede considerar la aplicación lineal $f : K^m \longrightarrow K^n$ dada por

$$(x_1, \ldots, x_m) \mapsto (t_{11}x_1 + \ldots + t_{1m}x_m, \ldots, t_{n1}x_1 + \ldots + t_{nm}x_m)$$

y así encontrar las soluciones al sistema equivale a encontrar $f^{-1}(b_1, \ldots, b_n)$.

Si $(b_1, b_2, \ldots, b_n) = (0, 0, \ldots, 0)$, entonces el conjunto de soluciones es $\mathrm{Ker}(f)$. Si $(b_1, b_2, \ldots, b_n) \notin \mathrm{Im}(f)$, entonces no hay solución al sistema. Cuando $(b_1, b_2, \ldots, b_n) \in \mathrm{Im}(f)$, existe solución y si $(a_1, a_2, \ldots, a_m)$ es una de ellas , entonces el conjunto de todas las soluciones al sistema está formado por la clase $(a_1, a_2, \ldots, a_m)+\mathrm{Ker}(f)$ (por la observación que sigue a (7.8)).

El sistema anterior se puede expresar matricialmente como $AX = B$ donde

$$A = \begin{pmatrix} t_{11} & \ldots & t_{1m} \\ t_{21} & \ldots & t_{2m} \\ \ldots & \ldots & \ldots \\ t_{n1} & \ldots & t_{nm} \end{pmatrix}; \; B = \begin{pmatrix} b_1 \\ \vdots \\ \vdots \\ b_n \end{pmatrix} \; X = \begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ x_m \end{pmatrix}.$$

A es la **matriz de coeficientes** y $(A|B)$ es la **matriz ampliada.**

Observar que la matriz coordenada de la aplicación f en las bases canónicas de K^m y de K^n es A^t.

(12.1) **Teorema**: (Rouché-Frobenius) Si $AX = B$ es un sistema de n ecuaciones con m incógnitas, existe alguna solución si y sólo si rang (A) = rang $(A|B)$.

Demostración: El que exista una solución $(a_1, \ldots, a_m)$ equivale a que

$$\begin{pmatrix} t_{11} \\ t_{21} \\ \vdots \\ t_{n1} \end{pmatrix} a_1 + \begin{pmatrix} t_{12} \\ t_{22} \\ \vdots \\ t_{n2} \end{pmatrix} a_2 + \ldots + \begin{pmatrix} t_{1m} \\ t_{2m} \\ \vdots \\ t_{nm} \end{pmatrix} a_m = \begin{pmatrix} b_1 \\ . \\ \vdots \\ b_n \end{pmatrix},$$

es decir, existe solucion si y sólo si

$$B \in K(A^1, \ldots, A^m) \iff K(A^1, \ldots, A^m) = K(A^1, \ldots, A^m, B) \iff$$

$$\iff \text{rang } (A) = \text{rang } (A|B).$$

Notas: i) Observar que Ker(f) es un subespacio de K^m de dimensión $m - \text{rang } (A)$.

ii) Un sistema homógeneo con $AX = (0)$ con n ecuaciones y n incógnitas tiene solución no trivial si y sólo si $\det(A) = 0$.

Regla de Cramer

(12.2) **Teorema**: **Un sistema de Cramer** es un sistema $AX = B$ de n ecuaciones y n incógnitas, con $\det(A) \neq 0$.

$$\begin{array}{ccccccc} t_{11}x_1 & + & \ldots & + & t_{1n}x_n & = & b_1 \\ t_{21}x_1 & + & \ldots & + & t_{2n}x_n & = & b_2 \\ \vdots & & & & \vdots & & \vdots \\ t_{n1}x_1 & + & \ldots & + & t_{nn}x_n & = & b_n \end{array}$$

Este sistema posee una solución única dada por:

$$x_i = 1/|A| \begin{vmatrix} & & & i) & & \\ t_{11} & \ldots & b_1 & \ldots & t_{1n} \\ t_{21} & \ldots & b_2 & \ldots & t_{2n} \\ \vdots & & & \vdots & \vdots \\ t_{n1} & \ldots & b_n & \ldots & t_{nn} \end{vmatrix}$$

donde se toma el determinante de la matriz que resulta de sustituir la columna i-ésima de A por la columna B.

Demostración: Multiplicar la j-ésima ecuación por A_{ji}:

$$\begin{array}{ccccccc} A_{1i}t_{11}x_1 & + & \ldots & + & A_{1i}t_{1n}x_n & = & A_{1i}b_1 \\ A_{2i}t_{21}x_1 & + & \ldots & + & A_{2i}t_{2n}x_n & = & A_{2i}b_2 \\ \vdots & & & & \vdots & & \vdots \\ A_{ni}t_{n1}x_1 & + & \ldots & + & A_{ni}t_{nn}x_n & = & A_{ni}b_n \end{array}$$

recordando que $t_{1i}A_{1j} + t_{2i}A_{2j} + \ldots + t_{ni}A_{nj} = \delta_{ij}|A|$, obtendremos al sumar dichas ecuaciones:

$$\delta_{i1}|A|x_1 + \ldots + |A|x_i + \ldots + \delta_{in}|A|x_n = A_{1i}b_1 + A_{2i}b_2 + \ldots + A_{ni}b_n =$$

$$= \begin{vmatrix} & & i) & & \\ t_{11} & \dots & b_1 & \dots & t_{1n} \\ t_{21} & \dots & b_2 & \dots & t_{2n} \\ \vdots & & & \vdots & \vdots \\ t_{n1} & \dots & b_n & \dots & t_{nn} \end{vmatrix}$$

Nota: A la hora de calcular las raíces de un sistema de ecuaciones lineales, se pueden intercambiar dos ecuaciones ó puede sustituirse una ecuación por ella más t por otra, con $t \in K$, ó puede sustituirse una ecuación por ella multiplicada por $t \in K, t \neq 0$.

Por tanto se le pueden aplicar a la matriz ampliada operaciones elementales sobre las filas hasta obtener una matriz triangular.

Lección 13. Valores y vectores propios de un endomorfismo.

Sea h un endomorfismo de un K espacio vectorial de dimensión n. Si $t \in K$, entonces $(h - t.1_V) : V \longrightarrow V$ dado por $(h - t.1_V)(v) = h(v) - t.v$ es un nuevo endomorfismo de V. Si la matriz coordenada de h en una cierta base es A, entonces la matriz coordenada de $(h - t.1_V)$ es $A - t.I_n$. Recordar que B es otra matriz coordenada de h si y sólo si existe $P \in \mathrm{GL}(n, K)$ tal que $B = PAP^{-1}$. Por tanto $\det(B) = \det(P)\det(A)\det(P^{-1}) = \det(A)$. Así tiene sentido definir el **determinante de un endomorfismo** h como el determinante de cualquiera de sus matrices coordenadas.

(13.1) **Definición**: a) Un elemento $t \in K$ es un **valor propio** de h si rang $(h - t.1_V) < n$. Un elemento $t \in K$ es un **valor propio** de $A \in \mathrm{Mat}(n \times n, K)$ si rang $(A - t.I_n) < n$. b) Un vector $v \neq \bar{0}$ (respectivamente una n-tupla no nula $(x_1, \ldots, x_n)$ de elementos de K) es un **vector propio** para h (respectivamente para $A \in \mathrm{Mat}(n \times n, K)$ si existe un $t \in K$ tal que $h(v) = t.v$ (respectivamente, $X^t A = tX^t$).

Nota: Un elemento $t \in K$ es valor propio de h si y sólo si es valor propio de alguna de sus matrices coordenadas. En efecto, si A es la matriz coordenada de un endomorfismo h en una base (a_i) de V, entonces la matriz coordenada de $(h - t.1_V)$ es $(A - t.I_n)$ y por (8.6) rang $(h - t.1_V) =$ rang $(A - t.I_n)$.

(13.2) **Proposición**: Son equivalentes:

1) t es un valor propio de h.

2) $\mathrm{Ker}(h - t.1_V) \neq \bar{0}$.

3) Existe $v \in V$, $v \neq \bar{0}$ tal que $h(v) = t.v$.

Demostración: 1)$\Rightarrow$ 2) Como $h - t.1_V \in \mathrm{End}_K(V)$, rang $(h - t.1_V) + \dim(\mathrm{Ker}(h - t.1_V)) = \dim(V)$, luego si rang $(h - t.1_V) < n$, entonces $\dim(\mathrm{Ker}(h - t.1_V)) > 0$ y $\mathrm{Ker}(h - t.1_V) \neq \bar{0}$.

2) $\Rightarrow$ 3) Si $\mathrm{Ker}(h - t.1_V) \neq \bar{0}$, existe $\bar{0} \neq v \in \mathrm{Ker}(h - t.1_V)$, luego $h(v) = t.v$.

3)$\Rightarrow$ 1) Cómo $\bar{0} \neq v$ y $h(v) = t.v$, se tiene que $\mathrm{Ker}(h - t.1_V) \neq \bar{0}$, luego rang $(h - t.1_V) < n$.

Nota: Cuando se cumpla $h(v) = t.v.$ para un $v \neq \bar{0}$ en V, diremos que v es **vector propio asociado al valor propio** t ó que el valor propio t es asociado al vector propio

v .

(13.3) **Proposición**: Son equivalentes:

1) t es un valor propio de la matriz A.

2) $\det(A - t.I_n) = 0$ ó equivalentemente $\det((A^t - t.I_n)) = 0$

3) Existe una n-tupla no nula $(x_1, \ldots, x_n)$ de elementos de K, tal que $X^t A = tX^t$.

Demostración: 1) $\Rightarrow$ 2) Trivialmente, si rang $(A - t.I_n) < n$, entonces $A - tI_n$ no es regular, luego $\det(A - tI_n) = 0$.

2) $\Rightarrow$ 3) Si $\det(A - tI_n) = 0$, entonces $\det((A - tI_n)^t) = 0 = \det(A^t - t.I_n)$. Por la nota ii) al teorema de Rouché-Frobenius, existe una n-tupla X no nula tal que $(A^t - t.I_n)X = (0)$, luego $A^t X = tX$ y por tanto trasponiendo se tiene $X^t A = tX^t$.

3) $\Rightarrow$ 1) Por la nota ii) al teorema de Rouché-Frobenius, se tiene que $\det((A - t.I_n)^t) = 0$, luego $\det(A - t.I_n) = 0$ ó equivalentemente, rang $(A - t.I_n) < n$.

(13.4) **Definición**: Llamaremos **polinomio característico** de $A \in \mathrm{Mat}(n \times n, K)$ al polinomio $\det(x.I_n - A)$.

Observaciones: 1) El polinomio característico de una matriz $A \in \mathrm{Mat}(n \times n, K)$, es el determinante de una matriz cuyas entradas están en el anillo de polinomios $K[x]$.

2) Es un polinomio de grado n, ya que sólo un sumando, en la expresión del determinante, tiene grado n, precisamente el sumando que aparece cuando tomamos todos los términos de la diagonal principal. Notar además que es un polinomio mónico.

3) Si $B = PAP^{-1}$, con $P \in \mathrm{GL}(n, K)$, entonces $\det(xI - B) = \det(xI - PAP^{-1}) = \det(P.xI.P^{-1} - PAP^{-1}) = \det(P(xI - A)P^{-1}) = \det(xI - A)$. Por tanto, dos matrices semejantes tiene el mismo polinomio característico.

Como consecuencia de la observación 3) podemos dar la siguiente

(13.5) **Definición**: Si $h \in \mathrm{End}_K(V)$, llamamos **polinomio característico** de h al polinomio característico de una de sus matrices coordenadas.

(13.6) **Teorema**: Sea h un endomorfismo de V, un K espacio vectorial de dimensión n. Los valores propios de h son exactamente las raíces en K del polinomio caracteristico de h.

Demostración: Sea A una matriz coordenada de h en alguna base (a_i) de V.

t es valor propio de $h \iff$ rang $(h - t.1_V) =$ rang $(A - t.I_n) =$ rang $(t.I_n - A) < n$ $\iff \det(t.I_n - A) = 0 \iff t$ es raíz del polinomio característico de A, que es el polinomio característico de h.

(¿Se utiliza algún resultado significativo para obtener la última equivalencia?)

(13.7) **Definición**: Si $h : V \longrightarrow V$ es un endomorfismo y t es un valor propio para h, llamaremos **subespacio fundamental** y lo denotaremos $S(t, h) = \{v \in V | h(v) = tv\}$.

Evidentemente es un subespacio vectorial de V ya que $S(t, h) = \text{Ker}(h - t.1_V)$.

(13.8) **Teorema**: Sean $v_1, \ldots, v_r$ son vectores propios de h, asociados a valores propios $t_1, \ldots, t_r$, donde $t_i \neq t_j$ si $i \neq j$. Entonces $v_1, \ldots, v_r$ son linealmente independientes.

Demostración: Por inducción sobre r. El resultado es obvio si $r = 1$ ya que un vector propio es distinto de cero. Supongamos que $r > 1$ y que el resultado es válido para $r - 1$ vectores. Si $s_1v_1 + s_2v_2 + \ldots + s_rv_r = \bar{0}$, entonces $\bar{0} = (h - t_1.1_V)(s_1v_1 + s_2v_2 + \ldots + s_rv_r) = s_1(h - t_1.1_V)(v_1) + s_2(h - t_1.1_V)(v_2) + \ldots + s_r(h - t_1.1_V)(v_r) = s_2(t_2 - t_1)v_2 + \ldots + s_r(t_r - t_1)v_r$. Por la hipótesis inductiva $s_i(t_i - t_1) = 0, i = 2, \ldots, r$ y como $(t_i - t_1) \neq 0$, se sigue que $s_i = 0, i = 2, \ldots, r$ y por tanto también $s_1 = 0$.

(13.9) **Corolario**: Los subespacios fundamentales de un endomorfismo son independientes.

(13.10) **Definición**: Un endomorfismo h de V se dice **diagonalizable** si tiene una matriz coordenada diagonal.

Una matriz $A \in \text{Mat}(n \times n, K)$ se dice **diagonalizable** si es matriz coordenada de un endomorfismo diagonalizable, esto es si A es semejante a una matriz D diagonal.

Notas: 1) Es claro que h es diagonalizable si y sólo si V tiene una base de vectores propios para h. Equivalentemente, h es diagonalizable si y sólo si existen $d_i \in K$ y $r \geq 1$ tales que $V = S(d_1, h) \oplus \ldots \oplus S(d_r, h)$.

2) Si h es diagonalizable y D es una matriz coordenada diagonal de h, entonces pol. car. $(h) = |xI - D| = \begin{vmatrix} x - d_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & x - d_n \end{vmatrix} = (x - d_1).\ldots.(x - d_n)$

(13.11) **Definición**: Si t es un valor propio para el endomorfismo h de V, llamaremos **multiplicidad algebraica** de t en h a la multiplicidad de t como raíz del polinomio caracteristico de h, es decir al mayor entero positivo m tal que $(x - t)^m | \det(xI - A)$, siendo

A una matriz coordenada de h. La denotaremos $m.a(t)$. Llamaremos **multiplicidad geométrica** de t a $\dim(S(t,h))$.

(13.12) **Teorema**: Sea $A \in \mathrm{Mat}(n\times n, K)$. Supongamos que $|xI-A|$ se factoriza como producto de factores de grado 1 en $K[x]$. Sea $h : V \longrightarrow V$ un endomorfismo que tiene a A por matriz coordenada. Sean $t_1, \ldots, t_r$ los distintos valores propios de h. Entonces:

i) h es diagonalizable $\iff V = S(t_1,h) \oplus \ldots \oplus S(t_r,h)$.

ii) h es diagonalizable $\iff m.a(t_i) = m.g(t_i),\ \forall i = 1, \ldots, r$.

iii) Si todos los factores de grado 1 de $|xI - A|$ son distintos, entonces A es diagonalizable.

Demostración: De la nota se sigue i).

ii) Si h es diagonalizable y $m.a(t_i) = m$, entonces t_i aparece m veces en la diagonal de D, una matriz coordenada diagonal de h, así $\dim_K(S(t_i,h)) = \dim_K(\mathrm{Ker}(h - t_i 1_V)) = n - \mathrm{rang}\ (D - t_i I) = n - (n-m) = m$.

Recíprocamente, como la suma de las multiplicidades algebraicas de los diferentes valores propios es n, se tiene: $\sum_{i=1}^{r} m.g.(t_i) = \sum_{i=1}^{r} m.a.(t_i) = n$, luego por i) y (13.9), h es diagonalizable.

iii) En este caso habrá $t_1, \ldots, t_n$ valores propios distintos entre sí, luego necesariamente es $\dim(S(t_i,h)) = 1, \forall i$ y así $m.g.(t_i) = m.a.(t_i), \forall i = 1, \ldots, r$ y se aplica ii).

BIBLIOGRAFIA.

1. T.W. Hungerford, *Algebra*, Springer, New York, 1974.

2. B. Jacob, *Linear Algebra*, Freeman and Company, San Francisco, 1989.

3. J. Sancho de San Román, *Algebra Lineal y Geometría*, Octavio y Félez, Zaragoza, 1976.

4. K. Spindler, *Abstract Algebra with Applications, vol I, Vector spaces and groups*, Dekker, New York, 1994