

ÁLGEBRA MODERNA

I. N. Herstein

Esta obra presenta el sistema básico algebraico desde un punto de vista abstracto. Ofrece abundante material práctico al alumno.

En la mayoría de los capítulos se hace el intento de ayudar al estudiante a comprender el significado de los resultados generales obtenidos mediante su aplicación a problemas particulares.

Estudia el concepto algebraico básico, teoría de grupos, que sirve como uno de los bloques de construcción fundamentales de la gran estructura que hoy se llama álgebra abstracta. Trata conceptos algebraicos igualmente básicos como anillos, campos, espacios vectoriales y álgebra lineal.

Se incluyen también los elementos de la teoría de Galois. Como material complementario se dan varios resultados relativos a campos finitos, demostrándose el teorema de Wedderburn sobre anillos finitos con división; un teorema de Frobenius y un teorema de los cuatro cuadrados.

Presenta un gran número de problemas para complementar pruebas que aparecen a lo largo del texto, para ilustrar y



Luis Arroyo Gómez S.

McGraw Hill
1977

Algebra moderna

Traducción:

Federico Velasco Coba
Coordinador académico del
Instituto de Geofísica
Facultad de Ciencias
Universidad Veracruzana

Revisión técnica:

Emilio Lluis Riera
Instituto de Matemáticas
Facultad de Ciencias
Universidad Nacional Autónoma de México

BIBLIOTECA NACIONAL MATEMÁTICA SUPERIOR

Mayo 29

Deben Fotocopias : ~~Javier Lomus~~
Atildefonso Mora
Nelsi
Alexander

I. N. Herstein

Algebra moderna

- Grupos
- Anillos
- Campos
- Teoría de Galois



**Editorial Trillas
México**

Título de esta obra en inglés:

Topics in Algebra
versión autorizada en español de la
primera edición publicada en inglés por
© 1964, Blaisdell Publishing Company
Division of Ginn and Company
Waltham, Massachusetts, E. U. A.

Primera edición en español, 1970
Reimpresiones, 1973, 1974, 1976 y 1979

Quinta reimpresión, noviembre 1980

*La presentación y disposición en conjunto de
ÁLGEBRA MODERNA,
son propiedad del editor. Prohibida la reproducción
parcial o total de esta obra, por cualquier medio o método,
sin autorización por escrito del editor*

*Derechos reservados en lengua española conforme a la ley
© 1970, Editorial Trillas, S. A.,
Av. Río Churubusco 385 Pte., México 13, D. F.*

*Miembro de la Cámara Nacional de la
Industria Editorial. Reg. núm. 158*

Impreso en México

ISBN 968-24-0137-2

Indice general

Capítulo 1 NOCIONES PRELIMINARES 11

1. Teorías de conjuntos 12
2. Aplicaciones 21
3. Los enteros 28

Capítulo 2 TEORÍA DE GRUPOS 37

1. Definición de grupo 39
2. Algunos ejemplos de grupos 40
3. Algunos lemas preliminares 42
4. Subgrupos 45
5. Relación entre los números de elementos 52
6. Subgrupos normales y grupos cociente 56
7. Homomorfismos 61
8. Automorfismos 72
9. El teorema de Cayley 77
10. Grupos de permutaciones 81
11. Otro principio de conteo 87
12. El teorema de Sylow 97

Capítulo 3 TEORÍA DE ANILLOS 103

1. Definición y ejemplos de anillos 103
2. Algunas clases especiales de anillos 108
3. Homomorfismos 113
4. Ideales y anillos cociente 115
5. Más ideales y más anillos cociente 119
6. El campo de cocientes de un dominio entero 123
7. Anillos euclidianos 126
8. Un anillo euclíadiano particular 133
9. Anillos de polinomios 136
10. Polinomios sobre el campo racional 143
11. Anillos de polinomios sobre anillos conmutativos 145

**Capítulo 4
ESPACIOS VECTORIALES Y MÓDULOS 155**

1. Conceptos básicos elementales 156
2. Independencia lineal y bases 162
3. Espacios duales 171
4. Espacios con producto interior 178
5. Módulos 188

**Capítulo 5
CAMPOS 197**

1. Extensión de campos 198
2. La transcendencia de e 207
3. Raíces de polinomios 210
4. Construcciones con regla y compás 220
5. Más acerca de raíces 224
6. Elementos de la teoría de Galois 229
7. Solubilidad por radicales 243

**Capítulo 6
TRANSFORMACIONES LINEALES 251**

1. El álgebra de las transformaciones lineales 252
2. Raíces características 261
3. Matrices 265
4. Formas canónicas: forma triangular 279
5. Formas canónicas: transformaciones nilpotentes 287
6. Formas canónicas. Una descomposición de V : forma de Jordan 294
7. Formas canónicas: forma canónica racional 303
8. Traza y transpuesta 312
9. Determinantes 321
10. Transformaciones hermitianas, unitarias y normales 338
11. Formas cuadráticas reales 353

**Capítulo 7
TÓPICOS SELECTOS 359**

1. Campos finitos 361
2. Teorema de Wedderburn sobre anillos finitos con división 365
3. Teorema de Frobenius 374
4. Cuaternios enteros y el teorema de los cuatro cuadrados 377

ÍNDICE ANALITICO 385

Prólogo

LA IDEA de escribir este libro y, lo que es más importante, el deseo de hacerlo como lo hemos elaborado, surgió de un curso que dio quien esto escribe en el año académico 1959-1960 en la Universidad de Cornell. Los asistentes a este curso eran, en su mayor parte, los alumnos de segundo año más dotados para las matemáticas de Cornell. Planeé el curso con el deseo de experimentar cuáles serían los resultados de presentarles material ligeramente más elevado al que es usual enseñar en álgebra al nivel de tercero y cuarto años.

He intentado que este libro sea, tanto en contenido como en grado de dificultad, algo intermedio entre dos grandes clásicos, *A Survey of Modern Algebra* de Birkhoff y MacLane y *Modern Algebra* de Van der Waerden.

En años recientes han ocurrido cambios muy marcados en la instrucción matemática que se da en las universidades norteamericanas. Estos cambios son más notables en los últimos años antes de la graduación y el comienzo de los estudios para graduados. Temas que desde hace algunos años se consideraban materia propia en cursos para graduados semiavanzados han ido filtrándose hasta que hoy se enseñan incluso en el primer curso de álgebra abstracta. Convencido de que esta filtración continuará e incluso se intensificará en los próximos años, expongo en el libro, diseñado para usarse como una primera introducción al álgebra, material que hasta aquí se ha considerado como un poco avanzado para esta etapa educativa.

Hay siempre el gran peligro, cuando tratamos ideas abstractas, de introducirlas demasiado repentinamente y sin una base suficiente de ejemplos que las hagan verosímiles o naturales. Con el fin de mitigar esta circunstancia, he tratado de motivar los conceptos de antemano y de ilustrarlos con ejemplos concretos. Una de las pruebas más significativas del valor de un concepto abstracto es lo que tanto él como los resultados que de su uso surgen nos dicen en las situaciones familiares. En casi cada uno de los capítulos hemos intentado hacer ver el significado de los resultados generales mediante su aplicación a problemas particulares. Por ejemplo, en el capítulo sobre anillos, el teorema de los dos cuadrados de Fermat se expone como una consecuencia directa de la teoría desarrollada para los anillos euclidianos.

Los temas escogidos para estudio lo han sido no solamente porque se ha hecho habitual presentarlos a este nivel o porque son importantes en el desarrollo general; sino también teniendo en cuenta su "concreción". Por esta razón, decidimos omitir el teorema de Jordan-Hölder que, desde luego podríamos haber incluido fácilmente en los resultados que derivamos sobre grupos. Sin embargo, apreciar este resultado en su verdadero valor requiere una gran visión, y usarlo en forma adecuada, una digresión demasiado grande. Es cierto que puede desarrollarse toda la teoría de la dimensión de un espacio vectorial como uno de sus corolarios; pero, en una primera presentación, abordar así el problema parece demasiado fantasioso y poco natural para un problema tan básico y apegado a tierra. Tampoco hay mención alguna de productos tensoriales o construcciones análogas. Habiendo, como hay, tanto tiempo y oportunidad para hacernos abstractos, ¿por qué acometer esta tarea desde el principio?

Una palabra acerca de los problemas. Hay un gran número de ellos. Ciertamente, un estudiante que los resolviera todos sería realmente extraordinario. Algunos se presentan solo para completar pruebas que aparecen en el texto; otros, para ilustrar y hacer prácticas sobre los resultados obtenidos. Muchos, se incluyeron no tanto para que sean resueltos como para que se manejen ideas importantes al intentar su solución. El valor de un problema no está tanto en conseguir su solución como en las ideas y bosquejos de ideas que hace surgir en el presunto solucionador. Hay otros que se incluyen como material preliminar a lo que va a presentarse más tarde, con la esperanza e idea de que esto sirva como fundamento para la subsecuente teoría y para hacer, además, más naturales ideas, definiciones y argumentos a medida que se van presentando. Algunos problemas aparecen más de una vez. Los problemas que por una u otra razón me parecieron difíciles, tienen una estrella (y algunas veces dos). Claro que tampoco en esto habrá acuerdo entre los matemáticos; muchos pensarán que algunos problemas sin estrella deberían tenerla y viceversa.

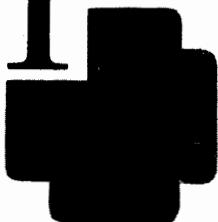
Estoy, naturalmente, en deuda con muchas personas por diversas sugerencias, comentarios y críticas. Para mencionar unas cuantas: Charles Curtis, Marshall Hall, Nathan Jacobson, Arthur Mattuck y Maxwell Rosenlicht. Debo mucho a Daniel Gorenstein e Irving Kaplansky por las numerosas conversaciones acerca del libro, su material y la forma de abordar muchos de sus temas. Debo, sobre todo, dar las gracias a George Seligman por las muchas e incisivas sugerencias y observaciones que me ha hecho sobre la presentación, tanto en lo que a estilo como en lo que a contenido se refiere. Damos también las gracias a Francis McNary de la dirección de Ginn and Company por su ayuda y cooperación. Finalmente, deseo hacer público mi agradecimiento a la John Simon Guggenheim Memorial Foundation; este libro se escribió en parte con su apoyo mientras el autor estaba en Roma como becado de la Guggenheim.



Algebra moderna

CAPITULO

1



Nociones preliminares

UNO DE los aspectos sorprendentes de la matemática del siglo veinte ha sido el reconocimiento del poder de un método abstracto. Ha hecho nacer esto un gran cuerpo de nuevos resultados y problemas, y en realidad nos ha conducido a la apertura de nuevas áreas de las matemáticas, de cuya mera existencia no se había ni sospechado.

Con estos desarrollos no solo nos han llegado nuevas matemáticas sino una visión fresca y, junto con ella, nuevas pruebas de resultados clásicos. La reducción de un problema a sus aspectos esenciales básicos nos ha revelado con frecuencia el emplazamiento adecuado de resultados considerados especiales y aislados, y nos ha mostrado interrelaciones entre áreas en las que nunca se había pensado que existiera alguna conexión.

El álgebra, que ha surgido como fruto natural de todo esto, no solamente es una materia de vida y vigor independientes —es una de las áreas más importantes de la investigación matemática habitual— sino que sirve también como hilo unificador que entrelaza a casi todas las matemáticas —geometría, teoría de los números, análisis, topología e incluso matemática aplicada.

Este libro ha sido pensado como una introducción a aquella parte de las matemáticas que hoy en día se conoce con el nombre de álgebra abstracta. El término "abstracto" es altamente subjetivo; lo que es abstracto para una persona con mucha frecuencia es concreto y poco elaborado para otra, y viceversa. En relación a las actividades de investigación corrientes en álgebra, podría describirse como "no demasiado abstracta"; desde el punto de vista de alguien educado en el cálculo y que está viendo el presente material por primera vez, puede muy bien ser descrito como "enteramente abstracto".

Sea como fuere, nos vamos a ocupar de la introducción y desarrollo de algunos de los sistemas algebraicos más importantes —grupos, anillos, espacios vectoriales, campos. Un sistema algebraico puede describirse como un conjunto de objetos, junto con algunas operaciones para combinarlos.

Antes de estudiar conjuntos restringidos en una forma cualquiera, por ejemplo, con operaciones, será necesario que consideremos conjuntos en general y algunas nociones respecto a ellos. En el otro extremo del espectro, necesitaremos alguna información acerca de un conjunto particular, el conjunto de los enteros. El propósito de este capítulo es el de discutir estos temas y derivar algunos resultados acerca de ellos a los que nos podamos referir, cuando la ocasión surja, posteriormente en el libro.

I. TEORÍA DE CONJUNTOS

No intentaremos una definición formal de un conjunto ni intentaremos sentar las bases para una teoría axiomática de la teoría de conjuntos. En lugar de ello, abordaremos el problema, desde un punto de vista operacional e intuitivo pensando en un conjunto como en una colección de objetos. En la mayoría de nuestras aplicaciones trataremos con cosas específicas y la nebulosa noción de conjunto se nos presentará en ella como algo perfectamente reconocible. Para aquellos cuyo gusto tienda más hacia el lado formal y abstracto, podemos considerar un conjunto como una noción primitiva que no se define.

Unas pocas observaciones sobre notación y terminología. Dado un conjunto S usaremos a lo largo de todo el libro la notación $a \in S$ para que se lea " a es un elemento de S ". De igual modo, $a \notin S$ se leerá " a no es un elemento de S ". El conjunto A se dirá que es un *subconjunto* del conjunto S si todo elemento en A es un elemento de S ; es decir, si $a \in A$ implica $a \in S$.

Escribiremos esto como $A \subset S$ (o, a veces, como $S \supset A$) que puede leerse como "A está contenido en S" (o S contiene a A). Esta notación no descarta la posibilidad de que $A = S$. Por cierto, ¿qué es lo que quiere decirse por igualdad de dos conjuntos? Para nosotros, lo que significará siempre es que ambos contienen los mismos elementos, es decir, que todo elemento que está en uno está en el otro y viceversa. En términos del símbolo para la relación de contención, los dos conjuntos A y B son iguales, lo que escribimos $A = B$, si ambas relaciones $A \subset B$ y $B \subset A$ se verifican simultáneamente. El procedimiento común para probar la igualdad de dos conjuntos, algo que necesitaremos hacer con frecuencia, es demostrar que las dos relaciones de contención opuestas se verifican para ellos. Un subconjunto A de S se llamará subconjunto *propio* de S si $A \subset S$, pero $A \neq S$ (A no es igual a S).

El conjunto *vacío* es el conjunto que no tiene elemento alguno; es un subconjunto de todo conjunto.

Una observación final, exclusivamente sobre notación: dado un conjunto S usaremos constantemente la notación $A = \{a \in S | P(a)\}$ que leeremos "A es el conjunto de todos los elementos en S para los que se verifica la propiedad P ". Por ejemplo, si S es el conjunto de los enteros y A el subconjunto de los enteros positivos, entonces podemos describir A como $A = \{a \in S | a > 0\}$. Otro ejemplo: si S es el conjunto que consiste en los objetos (1), (2), ..., (10), entonces el subconjunto A consistente en los objetos (1), (4), (7), (10) puede describirse por $A = \{(i) \in S | i = 3n + 1, n = 0, 1, 2, 3\}$.

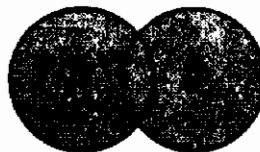
Dados dos conjuntos podemos combinarlos para formar nuevos conjuntos. No hay nada de sagrado ni de particular acerca de este número dos; podemos emplear el mismo procedimiento para cualquier número de conjuntos, finito o infinito y, ciertamente, así lo haremos. Primero, lo haremos solo con dos, porque ilustra la construcción general y no queda oscurecido con dificultades adicionales por la notación.

DEFINICIÓN. La *unión* de los dos conjuntos A y B , escrita $A \cup B$, es el conjunto $\{x | x \in A \text{ o } x \in B\}$.

Unas palabras acerca del uso de "o". En el castellano común y corriente, cuando decimos que una cosa es de esta forma o de esta otra, implicamos que no es de ambas. El "o" matemático es diferente por completo, al menos cuando estamos hablando de teoría de conjuntos. *Porque cuando decimos que x está en A o x está en B lo que queremos decir es que x está al menos en uno de los dos A o B, y puede ser que esté en ambos.*

Consideremos unos cuantos ejemplos de unión de dos conjuntos. Para cualquier conjunto A , $A \cup A = A$; en realidad, siempre que B es un subconjunto de A , $A \cup B = A$. Si A es el conjunto $\{x_1, x_2, x_3\}$ (es decir, el conjunto cuyos elementos son x_1, x_2, x_3), y B es el conjunto $\{y_1, y_2, x_1\}$, entonces $A \cup B = \{x_1, x_2, x_3, y_1, y_2\}$. Si A es el conjunto de todas las personas

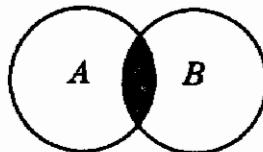
que tienen el pelo rubio y B es el conjunto de todas las personas que fuman, entonces $A \cup B$ consiste en todas las personas que, o tienen el pelo rubio o fuman, o tienen las dos características juntas. Gráficamente, podemos ilustrar la unión de dos conjuntos A y B con el diagrama siguiente.



Aquí, A es el círculo de la izquierda, B el de la derecha, y $A \cup B$ toda la parte sombreada.

DEFINICIÓN. La *intersección* de los dos conjuntos A y B , escrita $A \cap B$, es el conjunto $\{x | x \in A \text{ y } x \in B\}$.

La intersección de A y B es, pues, el conjunto de todos los elementos que están en ambos A y B . En analogía con los ejemplos usados para ilustrar la unión de dos conjuntos, veamos a qué es igual la intersección en esos mismos casos. Para cualquier conjunto A , $A \cap A = A$; en realidad, si B es un subconjunto cualquiera de A , entonces $A \cap B = B$. Si A es el conjunto $\{x_1, x_2, x_3\}$ y B el conjunto $\{y_1, y_2, x_1\}$, entonces $A \cap B = \{x_1\}$ (estamos suponiendo que ninguna y es una x). Si A es el conjunto de todas las personas que tienen el pelo rubio, y B es el conjunto de todas las personas que fuman, entonces $A \cap B$ es el conjunto de todas las personas rubias que fuman. Gráficamente, podemos ilustrar la intersección de los dos conjuntos A y B de la manera siguiente:



Aquí A es el círculo a la izquierda, B el de la derecha, mientras que la intersección es la parte sombreada.

Dos conjuntos se dice que son *ajenos* si su intersección es vacía, es decir, el conjunto vacío. Por ejemplo, si A es el conjunto de enteros positivos y B el conjunto de enteros negativos, entonces A y B son ajenos. Nótese, sin embargo, que si C es el conjunto de enteros no negativos y D es el conjunto de enteros no positivos, entonces no son ajenos, pues su intersección consiste del entero 0, luego no es vacío.

Antes de generalizar la unión e intersección de dos conjuntos a un número arbitrario de conjuntos, nos gustaría probar una pequeña proposición que correlaciona la unión con la intersección. Es, este, el primero

de una gran colección de resultados de este tipo; algunos de ellos pueden verse en los problemas que aparecen al final de esta sección.

PROPOSICIÓN. *Para tres conjuntos cualesquiera A , B , C tenemos*

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

Prueba. La prueba consistirá en demostrar, en primer lugar, que $(A \cap B) \cup (A \cap C) \subset A \cap (B \cup C)$ y después la relación inversa $A \cap (B \cup C) \subset (A \cap B) \cup (A \cap C)$.

Probemos primero que $(A \cap B) \cup (A \cap C) \subset A \cap (B \cup C)$. Como $B \subset B \cup C$, es inmediato que $A \cap B \subset A \cap (B \cup C)$. De una forma análoga, $A \cap C \subset A \cap (B \cup C)$. Por tanto,

$$(A \cap B) \cup (A \cap C) \subset (A \cap (B \cup C)) \cup (A \cap (B \cup C)) = A \cap (B \cup C).$$

Pasemos ahora a la otra dirección. Dado un elemento $x \in A \cap (B \cup C)$, es claro que debe, en primer lugar, ser un elemento de A . Además, como elemento de $B \cup C$ debe estar en B o en C . Supongamos lo primero; entonces como un elemento tanto de A como de B , x debe estar en $A \cap B$. La segunda posibilidad, es decir, $x \in C$, nos lleva a $x \in A \cap C$. Por tanto, en cualquiera de las eventualidades $x \in (A \cap B) \cup (A \cap C)$, de donde tenemos $A \cap (B \cup C) \subset (A \cap B) \cup (A \cap C)$.

Las dos relaciones de contención opuestas, al combinarse, nos dan la igualdad que afirma la proposición.

Continuamos con la discusión de conjuntos para extender la noción de unión y la de intersección a colecciones arbitrarias de conjuntos.

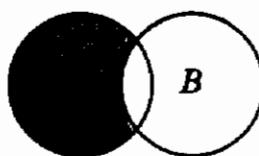
Dado un conjunto T decimos que T sirve como un *conjunto de índices* para la familia $\mathcal{F} = \{A_\alpha\}$ de conjuntos si para cada $\alpha \in T$ existe un conjunto A_α en la familia \mathcal{F} . El conjunto de índices T puede ser cualquiera, finito o infinito. A menudo usamos el conjunto de enteros positivos como conjunto de índices, pero, repetimos, T puede ser cualquier conjunto no vacío.

Por la *unión* de los conjuntos A_α , donde α está en T , entendemos el conjunto $\{x | x \in A_\alpha \text{ para, al menos, un } \alpha \in T\}$. Denotaremos tal conjunto por $\bigcup_{\alpha \in T} A_\alpha$. Por la *intersección* de los conjuntos A_α donde α está en T , entendemos el conjunto $\{x | x \in A_\alpha \text{ para todo } \alpha \in T\}$; denotaremos este conjunto por $\bigcap_{\alpha \in T} A_\alpha$. Los conjuntos A_α son *mutuamente ajenos* si para $\alpha \neq \beta$, $A_\alpha \cap A_\beta$ es el conjunto vacío.

Por ejemplo, si S es el conjunto de los números reales, y si T es el conjunto de los números racionales, sea, para $\alpha \in T$, $A_\alpha = \{x \in S | x > \alpha\}$. Es un fácil ejercicio ver que $\bigcup_{\alpha \in T} A_\alpha = S$ mientras que $\bigcap_{\alpha \in T} A_\alpha$ es el conjunto vacío. Los conjuntos A_α no son mutuamente ajenos.

DEFINICIÓN. Dados dos conjuntos A , B , entonces el *conjunto diferencia* $A - B$ es el conjunto $\{x \in A | x \notin B\}$.

Volviendo a nuestras pequeñas ilustraciones, si A es el círculo a la izquierda y B es el de la derecha, entonces $A - B$ es el área sombreada.



Obsérvese que para cualquier conjunto B , el conjunto A satisface $A = (A \cap B) \cup (A - B)$. (Pruébese.) Obsérvese, además, que $B \cap (A - B)$ es el conjunto vacío. Un caso particular de interés en la diferencia de dos conjuntos es cuando uno de ellos es un subconjunto del otro. En ese caso, si B es un subconjunto de A , a $A - B$ le llamamos el *complemento de B en A*.

Aun necesitaremos construir un tipo de conjunto más, partiendo de dos conjuntos A y B , su *producto cartesiano* $A \times B$. Este conjunto $A \times B$ se define como el conjunto de todos los pares ordenados (a, b) donde $a \in A$ y $b \in B$ y donde convenimos en que el par (a_1, b_1) es igual al par (a_2, b_2) si y sólo si $a_1 = a_2$ y $b_1 = b_2$.

Unas cuantas observaciones respecto al producto cartesiano. Dados los dos conjuntos A y B podemos construir los conjuntos $A \times B$ y $B \times A$. Como conjuntos, son distintos, aunque sintámos que deben estar estrechamente relacionados. Dados tres conjuntos A , B y C , podemos construir muchos productos cartesianos partiendo de ellos; por ejemplo, el conjunto $A \times D$, donde $D = B \times C$; el conjunto $E \times C$, donde $E = A \times B$; y también el conjunto de todas las ternas ordenadas (a, b, c) donde $a \in A$, $b \in B$ y $c \in C$. Obtenemos, así, tres conjuntos distintos, aunque también aquí sentimos que estos conjuntos deben tener una relación estrecha. Desde luego, podemos continuar este proceso con más y más conjuntos. Para ver la relación exacta entre ellos tendremos que esperar a la sección próxima, donde discutiremos las correspondencias biyectivas.

Dado un conjunto de índices T podríamos definir el producto cartesiano de los conjuntos A_x cuando x varía sobre T ; como no necesitaremos un producto tan general, no nos molestaremos en definirlo.

Finalmente, podemos considerar el producto cartesiano de un conjunto A por sí mismo, $A \times A$. Nótese que si el conjunto A es un conjunto finito con n elementos, entonces el conjunto $A \times A$ es también un conjunto finito, pero tiene n^2 elementos. El conjunto de elementos (a, a) en $A \times A$ se llama la *diagonal* de $A \times A$.

Un subconjunto R de $A \times A$ se dice que es una *relación de equivalencia* sobre A si:

- 1) $(a, a) \in R$ para todo $a \in A$;
- 2) $(a, b) \in R$ implica $(b, a) \in R$;
- 3) $(a, b) \in R$ y $(b, c) \in R$ implica que $(a, c) \in R$.

En lugar de hablar de subconjuntos de $A \times A$ podemos hablar acerca de una relación binaria (una entre dos elementos de A) sobre A mismo, conviniendo en que diremos que b está relacionado a a si $(a, b) \in R$. Las propiedades (1), (2), (3) del subconjunto R se traducen inmediatamente en las propiedades (1), (2), (3) de la definición que sigue.

DEFINICIÓN. La relación binaria, \sim , sobre A se dice que es una *relación de equivalencia* sobre A si para a, b, c cualesquiera en A :

- 1) $a \sim a$;
- 2) $a \sim b$ implica $b \sim a$;
- 3) $a \sim b$ y $b \sim c$ implica $a \sim c$.

La primera de estas propiedades se llama *reflexividad*, la segunda, *simetría* y, la tercera, *transitividad*.

El concepto de relación de equivalencia es extraordinariamente importante y juega un papel central en todas las matemáticas. Lo ilustraremos con unos cuantos ejemplos.

Ejemplo 1. Sea S un conjunto cualquiera y definamos \sim en S por $a \sim b$ para $a, b \in S$ si y sólo si $a = b$. Hemos definido claramente, así, una relación de equivalencia sobre S . En realidad, una relación de equivalencia es una generalización de la igualdad, que mide la igualdad hasta una cierta propiedad.

Ejemplo 2. Sea S el conjunto de todos los enteros. Para $a, b \in S$ conveníamos en que $a \sim b$ si $a - b$ es un entero par. Verificamos que esto define una relación de equivalencia sobre S .

- 1) Como $0 = a - a$ es par, $a \sim a$.
- 2) Si $a \sim b$, es decir, si $a - b$ es par, entonces $b - a = -(a - b)$ es también par, luego $b \sim a$.
- 3) Si $a \sim b$ y $b \sim c$, entonces tanto $a - b$ como $b - c$ son pares, luego $a - c = (a - b) + (b - c)$ es también par, lo cual prueba que $a \sim c$.

Ejemplo 3. Sea S el conjunto de todos los enteros y sea $n > 1$ un entero fijo. Para $a, b \in S$ definimos \sim por $a \sim b$ si $a - b$ es un múltiplo de n . Dejamos como ejercicio probar que esto define una relación de equivalencia en S .

Ejemplo 4. Sea S el conjunto de todos los triángulos del plano. Dos triángulos los definimos como equivalentes si son semejantes (es decir, tienen ángulos correspondientes iguales). Esto define una relación de equivalencia sobre S .

En lugar de hablar de subconjuntos de $A \times A$ podemos hablar acerca de una relación binaria (una entre dos elementos de A) sobre A mismo, conviniendo en que diremos que b está relacionado a a si $(a, b) \in R$. Las propiedades (1), (2), (3) del subconjunto R se traducen inmediatamente en las propiedades (1), (2), (3) de la definición que sigue.

DEFINICIÓN. La relación binaria, \sim , sobre A se dice que es una *relación de equivalencia* sobre A si para a, b, c cualesquiera en A :

- 1) $a \sim a$;
- 2) $a \sim b$ implica $b \sim a$;
- 3) $a \sim b$ y $b \sim c$ implica $a \sim c$.

La primera de estas propiedades se llama *reflexividad*, la segunda, *simetría* y, la tercera, *transitividad*.

El concepto de relación de equivalencia es extraordinariamente importante y juega un papel central en todas las matemáticas. Lo ilustraremos con unos cuantos ejemplos.

Ejemplo 1. Sea S un conjunto cualquiera y definamos \sim en S por $a \sim b$ para $a, b \in S$ si y sólo si $a = b$. Hemos definido claramente, así, una relación de equivalencia sobre S . En realidad, una relación de equivalencia es una generalización de la igualdad, que mide la igualdad hasta una cierta propiedad.

Ejemplo 2. Sea S el conjunto de todos los enteros. Para $a, b \in S$ conveníamos en que $a \sim b$ si $a - b$ es un entero par. Verificamos que esto define una relación de equivalencia sobre S .

- 1) Como $0 = a - a$ es par, $a \sim a$.
- 2) Si $a \sim b$, es decir, si $a - b$ es par, entonces $b - a = -(a - b)$ es también par, luego $b \sim a$.
- 3) Si $a \sim b$ y $b \sim c$, entonces tanto $a - b$ como $b - c$ son pares, luego $a - c = (a - b) + (b - c)$ es también par, lo cual prueba que $a \sim c$.

Ejemplo 3. Sea S el conjunto de todos los enteros y sea $n > 1$ un entero fijo. Para $a, b \in S$ definimos \sim por $a \sim b$ si $a - b$ es un múltiplo de n . Dejamos como ejercicio probar que esto define una relación de equivalencia en S .

Ejemplo 4. Sea S el conjunto de todos los triángulos del plano. Dos triángulos los definimos como equivalentes si son semejantes (es decir, tienen ángulos correspondientes iguales). Esto define una relación de equivalencia sobre S .

Ejemplo 5. Sea S el conjunto de todos los puntos del plano. Dos puntos a y b se definen como equivalentes si equidistan del origen. Una sencilla comprobación verifica que esto define una relación de equivalencia sobre S .

Hay muchas más relaciones de equivalencia; nos encontraremos con unas cuantas más a lo largo de este libro.

DEFINICIÓN. Si A es un conjunto y \sim es una relación de equivalencia sobre A , entonces la *clase de equivalencia* de $a \in A$ es el conjunto $\{x \in A | a \sim x\}$. Lo escribimos “ $cl(a)$ ”.

En los ejemplos que acabamos de discutir, ¿cuáles son las clases de equivalencia? En el ejemplo 1, la clase de equivalencia de a consiste tan solo en a . En el ejemplo 2 la clase de equivalencia de a consiste en todos los enteros de la forma $a + 2m$ donde $m = 0, \pm 1, \pm 2, \dots$; en este ejemplo hay solamente dos clases de equivalencia distintas, a saber, $cl(0)$ y $cl(1)$. En el ejemplo 3, la clase de equivalencia de a consiste en todos los enteros de la forma $a + kn$ donde $k = 0, \pm 1, \pm 2, \dots$; aquí hay n clases de equivalencia distintas, a saber, $cl(0), cl(1), \dots, cl(n-1)$. En el ejemplo 5 la clase de equivalencia de a consiste en todos los puntos del plano que se encuentran sobre la circunferencia que tiene su centro en el origen y pasa por a .

Aunque hemos dado algunas definiciones, introducido algunos conceptos, e incluso establecido una sencilla pequeña proposición, podríamos decir que, sinceramente, hasta el momento no hemos probado resultado alguno que tenga cierta sustancia. Vamos ahora a probar el primer resultado genuino del libro. La prueba de este teorema no es muy difícil —realmente es muy sencilla— pero no por ello el resultado que en él se presenta deja de ser de un gran uso para nosotros.

TEOREMA 1.A. *Las distintas clases de equivalencia de una relación de equivalencia sobre A nos proporcionan una descomposición de A como una unión de subconjuntos mutuamente ajenos. Recíprocamente, dada una descomposición de A como unión de subconjuntos mutuamente ajenos y no vacíos, podemos definir una relación de equivalencia sobre A para la que estos subconjuntos sean las distintas clases de equivalencia.*

Prueba. Denotemos por \sim la relación de equivalencia sobre A .

Observemos primero que como para cualquier $a \in A$, $a \sim a$, a debe estar en $cl(a)$, de donde la unión de las $cl(a)$ es todo A . Afirmamos ahora que dadas dos clases de equivalencia o son ajenas o son iguales. Supongamos, en efecto, que $cl(a)$ y $cl(b)$ no son ajenas; entonces hay un elemento $x \in cl(a) \cap cl(b)$. Como $x \in cl(a)$, $a \sim x$; como $x \in cl(b)$, $b \sim x$, de donde, por la simetría de la relación, $x \sim b$. Pero $a \sim x$ y $x \sim b$ implica por la transitividad de la relación $a \sim b$. Supongamos ahora que $y \in cl(b)$; entonces $b \sim y$. Entonces de $a \sim b$ y $b \sim y$ se deduce que $a \sim y$, es decir, que $y \in cl(a)$. Luego todo elemento en $cl(b)$ está en $cl(a)$, lo que prueba que $cl(b) \subset cl(a)$. El

argumento es claramente simétrico, de donde concluimos que $\text{cl}(a) \subset \text{cl}(b)$. Y las dos relaciones de contención opuestas nos dicen que $\text{cl}(a) = \text{cl}(b)$.

Hemos demostrado así que las distintas $\text{cl}(a)$ son mutuamente ajenas y que su unión es A . Esto prueba la primera mitad del teorema. ¡Vayamos por la otra mitad!

Supongamos que $A = \cup A_\alpha$ donde las A_α son conjuntos mutuamente ajenos y no vacíos (α está en algún conjunto de índices T). ¿Cómo los usaremos para definir una relación de equivalencia? El procedimiento es claro; dado un elemento a en A está *exactamente en un* A_α . Definimos para $a, b \in A$, $a \sim b$ si a y b están en la misma A_α . Dejamos como ejercicio probar que esto es una relación de equivalencia sobre A y que las distintas clases de equivalencia son las A_α .

Problemas

1. a) Si A es un subconjunto de B y B es un subconjunto de C , pruébese que A es un subconjunto de C .
 b) Si $B \subset A$ pruébese que $A \cup B = A$ y recíprocamente.
 c) Si $B \subset A$ pruébese que para cualquier conjunto C se tiene $B \cup C \subset A \cup C$ y $B \cap C \subset A \cap C$.
2. a) Pruébese que $A \cap B = B \cap A$ y $A \cup B = B \cup A$.
 b) Pruébese que $(A \cap B) \cap C = A \cap (B \cap C)$.
3. Pruébese que $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.
4. Para un subconjunto C de S denotemos por C' el complemento de C en S . Para cualesquier dos subconjuntos A, B de S pruébense las *leyes de De Morgan*:
 a) $(A \cap B)' = A' \cup B'$.
 b) $(A \cup B)' = A' \cap B'$.
5. Para un conjunto finito C , indiquemos por $o(C)$ el número de elementos de C . Si A y B son conjuntos finitos pruébese que $o(A \cup B) = o(A) + o(B) - o(A \cap B)$.
6. Si A es un conjunto finito de n elementos, pruébese que A tiene exactamente 2^n subconjuntos distintos.
7. Una encuesta muestra que al 63% de los norteamericanos les gusta el queso y al 76% las manzanas. ¿Qué se puede decir acerca del porcentaje de norteamericanos a los que les gusta el queso y las manzanas? (No se asegura que las estadísticas que hemos mencionado sean muy exactas.)
8. Dados dos conjuntos A y B , su *diferencia simétrica* es, por definición, $(A - B) \cup (B - A)$. Pruébese que la diferencia simétrica de A y B es igual a $(A \cup B) - (A \cap B)$.

9. Sea S un conjunto y S^* el conjunto cuyos elementos son los distintos subconjuntos de S . En S^* definimos una adición y una multiplicación como sigue: si $A, B \in S^*$ (recuérdese, esto significa que son subconjuntos de S):

- 1) $A + B = (A - B) \cup (B - A)$.
- 2) $A \cdot B = A \cap B$.

Pruébese que las siguientes leyes gobiernan estas operaciones:

- a) $(A + B) + C = A + (B + C)$.
- b) $A \cdot (B + C) = A \cdot B + A \cdot C$.
- c) $A \cdot A = A$.
- d) $A + A =$ conjunto vacío.
- e) Si $A + B = A + C$ entonces $B = C$.

(El sistema que acabamos de describir es un ejemplo de álgebra booleana.)

10. Para cada uno de los conjuntos y relaciones que abajo se dan, determiníñense cuáles definen relaciones de equivalencia.

- a) S es el conjunto de todos los humanos vivos y $a \sim b$ si a y b tienen un antecesor común.
- b) S es el conjunto de todos los humanos vivos y $a \sim b$ si a vive a menos de 100 kilómetros de distancia de donde vive b .
- c) S es el conjunto de todos los humanos vivos y $a \sim b$ si a y b tienen el mismo padre.
- d) S es el conjunto de todos los números reales, $a \sim b$ si $a = \pm b$.
- e) S es el conjunto de los enteros, $a \sim b$ si simultáneamente se verifica que $a > b$ y $b > a$.
- f) S es el conjunto de todas las rectas del plano, y $a \sim b$ si a es paralela a b .

11. a) La propiedad (2) de una equivalencia nos dice que si $a \sim b$ entonces $b \sim a$; la propiedad (3) nos dice que $a \sim b$ y $b \sim c$ implica $a \sim c$. ¿Qué hay de equivocado en la siguiente prueba de que las propiedades (2) y (3) implican la propiedad (1)? Sea $a \sim b$; entonces $b \sim a$, de donde, por la propiedad (3) (usando $a = c$), $a \sim a$.

- b) ¿Puede el lector sugerirnos una alternativa de la propiedad (1) que nos asegure que las propiedades (2) y (3) implican la propiedad (1)?

12. En el ejemplo 3 de una relación de equivalencia dado en el texto, pruébese que la relación definida es una relación de equivalencia y que hay exactamente n clases distintas de equivalencia, a saber, $\text{cl}(0)$, $\text{cl}(1)$, ..., $\text{cl}(n-1)$.

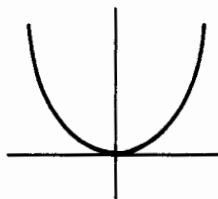
13. Complétense la prueba de la segunda mitad del teorema 1(a).

2. APPLICACIONES

Vamos ahora a introducir el concepto de aplicación o función de un conjunto en otro. Sin exageración, esta es probablemente la noción más importante y universal presente siempre a través de todas las matemáticas. Es difícil que sea algo nuevo para cualquiera de nosotros, pues hemos estado considerando aplicaciones desde los primeros días de nuestra educación en matemáticas. Cuando se nos pedía dibujar la gráfica de la relación $y = x^2$ lo que se nos pedía era simplemente estudiar la aplicación particular que lleva cada número real a su cuadrado.

Hablando en forma un poco libre, una aplicación de un conjunto S en otro conjunto T es una "regla" (sea lo que sea lo que esta palabra pueda significar) que asocia con cada elemento en S un elemento único t en T . Definiremos el término aplicación de un modo más formal y preciso, pero el propósito de la definición es el de permitirnos pensar y hablar en los términos anteriores. Pensaremos en ellos como en reglas, procedimientos o mecanismos que nos transportan de un conjunto a un otro.

Motivemos un poco la definición que vamos a dar. El punto de vista que tomaremos es el de considerar la aplicación definida por su "gráfica". Ilustraremos esto con el ejemplo familiar $y = x^2$ definida sobre los números reales S y tomando valores también sobre S . Para este conjunto S , $S \times S$, el conjunto de todos los pares (a, b) puede verse como el plano, con el par (a, b) que corresponde al punto cuyas coordenadas son a y b respectivamente. En este plano marcamos todos aquellos puntos cuyas coordenadas son de la forma (x, x^2) y denominaremos a este conjunto de puntos la gráfica de $y = x^2$. Representamos este conjunto por el dibujo siguiente.



Para encontrar el "valor" de la función o la aplicación en el punto $x = a$ miramos en el punto de la gráfica cuya primera coordenada es a y damos como valor de la función en $x = a$ el valor de la segunda coordenada.

Es este, más o menos, el método que usaremos para definir una aplicación de un conjunto en otro.

DEFINICIÓN. Si S y T son conjuntos no vacíos, entonces *una aplicación* de S en T es un subconjunto M de $S \times T$, tal que para *toda* $s \in S$ hay un solo $t \in T$ tal que el par ordenado (s, t) está en M .

Esta definición nos sirve para dar precisión al concepto de aplicación, pero casi nunca la usaremos en esa forma. En lugar de ellos preferiremos pensar en aplicación como una regla que asocia con cada elemento s de S algún elemento t de T , siendo la regla, asociar $s \in S$ con $t \in T$ (o aplicar o transformar s en t) si y sólo si $(a, t) \in M$. Diremos que t es la imagen de s bajo la aplicación.

Digamos ahora algo acerca de la notación para estas cosas. Sea σ una aplicación de S en T ; a menudo denotamos esto escribiendo $\sigma: S \rightarrow T$ o $S \xrightarrow{\sigma} T$. Si t es la imagen de s bajo σ escribiremos algunas veces esto así: $\sigma: s \rightarrow t$; más a menudo representaremos este hecho por $t = s\sigma$. Nótese que escribimos la aplicación σ a la derecha. No hay una consistencia general en este uso; muchas personas lo escribirían como $t = \sigma(s)$. Como regla general, los algebraistas lo escribirán a la derecha y otros matemáticos a la izquierda. En realidad, ni nosotros mismos seremos muy consistentes en esto; cuando queramos hacer hincapié en la naturaleza funcional de σ podemos escribir $t = \sigma(s)$.

EJEMPLOS DE APLICACIONES. En todos los ejemplos se supone que los conjuntos no son el vacío.

Ejemplo 1. Sea S un conjunto cualquiera; definamos $\iota: S \rightarrow S$ por $s = s\iota$ para todo $s \in S$. Esta aplicación ι se llama *aplicación identidad* de S .

Ejemplo 2. Sean S y T conjuntos cualesquiera, y sea t_0 un elemento de T . Definamos $\tau: S \rightarrow T$ por $\tau s \rightarrow t_0$ para todo $s \in S$.

Ejemplo 3. Sea S el conjunto de todos los números racionales positivos y $T = J \times J$ donde J es el conjunto de los enteros. Dado un número racional s podemos escribirlo como $s = \frac{m}{n}$ donde m y n no tienen ningún factor común. Definamos $\tau: S \rightarrow T$ por $\tau s = (m, n)$.

Ejemplo 4. Sea J el conjunto de los enteros y $S = \{(m, n) \in J \times J \mid n \neq 0\}$; sea T el conjunto de los números racionales; definimos $\tau: S \rightarrow T$ por $(m, n)\tau = \frac{m}{n}$ para todo (m, n) en S .

Ejemplo 5. Sea J el conjunto de los enteros y $S = J \times J$. Definimos $\tau: S \rightarrow J$ por $(m, n)\tau = m + n$.

Nótese que, en el ejemplo 5, la adición en la misma J puede representarse en términos de una aplicación de $J \times J$ en J . Dado un conjunto arbitrario S a una aplicación de $S \times S$ en S le llamaremos *operación binaria* sobre S . Dada una tal aplicación $\tau: S \times S \rightarrow S$ podríamos utilizarla para definir un “producto” * en S afirmando que $a * b = c$ si $(a, b)\tau = c$.

Ejemplo 6. Sean S y T conjuntos cualesquiera; definamos $\tau: S \times T \rightarrow S$ por $(a, b)\tau = a$ para todo $(a, b) \in S \times T$. A esta τ se le llama la proyección de $S \times T$ sobre S . En forma análoga, podríamos definir la proyección de $S \times T$ sobre T .

Ejemplo 7. Sea S el conjunto consistente en los elementos x_1, x_2, x_3 . Definamos $\tau: S \rightarrow S$ por $x_1\tau = x_2, x_2\tau = x_3, x_3\tau = x_1$.

Ejemplo 8. Sea S el conjunto de los enteros y T el conjunto consistente en los elementos E y 0 . Definamos $\tau: S \rightarrow T$ conviniendo en que $n\tau = E$ si n es par y $n\tau = 0$ si n es impar.

Si S es cualquier conjunto, sea $\{x_1, \dots, x_n\}$ su subconjunto consistente en los elementos x_1, x_2, \dots, x_n de S . En particular, $\{x\}$ es el subconjunto de S cuyo único elemento es x . Dado S lo podemos usar para construir un nuevo conjunto S^* , el conjunto cuyos elementos son los subconjuntos de S . A S^* le llamamos *conjunto de subconjuntos* de S . Así, por ejemplo, si $S = \{x_1, x_2\}$ entonces S^* tiene exactamente cuatro elementos, a saber, $a_1 = \text{conjunto vacío}, a_2 = \text{el subconjunto } S \text{ de } S, a_3 = \{x_1\}, a_4 = \{x_2\}$. La relación de S a S^* , en general, es muy interesante; examinaremos algunas de sus propiedades en los problemas.

Ejemplo 9. Sea S un conjunto, y $T = S^*$. Definamos $\tau: S \rightarrow T$ como $s\tau = \text{complemento de } \{s\} \text{ en } S = S - \{s\}$.

Ejemplo 10. Sea S un conjunto con una relación de equivalencia, y T el conjunto de clases de equivalencia en S (nótese que T es un subconjunto de S^*). Definamos $\tau: S \rightarrow T$ por $s\tau = \text{cl}(s)$.

Dejamos ahora los ejemplos para continuar con la discusión general. Dada una aplicación $\tau: S \rightarrow T$ definimos para $t \in T$ como *imagen inversa* de t con respecto a τ el conjunto $\{s \in S | t = s\tau\}$. En el ejemplo 8, la imagen inversa de E es el subconjunto de S consistente en los enteros pares. Puede suceder que para algún t en T su imagen inversa con respecto a τ sea el vacío, es decir, que t no sea imagen bajo τ de ningún elemento de S . En el ejemplo 3 que acabamos de discutir, el elemento $(4, 2)$ no es la imagen de ningún elemento de S bajo la τ usada; en el ejemplo 9, S , como un elemento de S^* , no es la imagen bajo la τ usada de ningún elemento de S .

DEFINICIÓN. La aplicación τ de S en T se dice que es suprayectiva sobre T si dado $t \in T$ cualquiera siempre existe un elemento $s \in S$ tal que $t = s\tau$.

Si llamamos al subconjunto $S\tau = \{x \in T | x = s\tau \text{ para algún } s \in S\}$ la *imagen* de S bajo τ , entonces τ es suprayectiva si la imagen de S bajo τ es todo T . Nótese que en los ejemplos 1, 4, 5, 6, 7, 8 y 10 las aplicaciones que se presentan son todas suprayectivas.

Otro tipo especial de aplicación aparece con frecuencia y es importante: las aplicaciones inyectivas (también denominadas uno a uno).

DEFINICIÓN. La aplicación τ de S en T se dice que es una aplicación inyectiva si siempre que $s_1 \neq s_2$ entonces $s_1\tau \neq s_2\tau$.

En términos de imágenes inversas, la aplicación τ es inyectiva si para toda $t \in T$ la imagen inversa de t es o vacía o un conjunto consistente en un solo elemento. En los ejemplos discutidos, las aplicaciones de los ejemplos 1, 3, 7 y 9 son todas inyectivas.

¿En qué momento diremos que dos aplicaciones de S en T son iguales? Una definición natural para esto es la de que deben tener el mismo efecto sobre cada elemento de S , es decir, la imagen de cualquier elemento en S bajo cada una de las aplicaciones debe ser la misma. De un modo un poco más formal:

DEFINICIÓN. Dos aplicaciones σ y τ de S en T se dice que son *iguales* si $s\sigma = s\tau$ para todo $s \in S$.

Consideremos la siguiente situación: tenemos una aplicación σ de S en T y otra aplicación τ de T en U . ¿Podemos componer estas aplicaciones para obtener una aplicación de S en U ? La forma más natural y obvia de hacer esto es enviar un elemento dado s de S en dos etapas a U , primero aplicando σ a s y después aplicando τ al elemento resultante $s\sigma$ de T . Esta es la base de la definición que sigue.

DEFINICIÓN. Sean $\sigma: S \rightarrow T$ y $\tau: T \rightarrow U$. Entonces la *composición* de σ y τ (llamada también *producto*) es la aplicación $\sigma \circ \tau: S \rightarrow U$ definida por $s(\sigma \circ \tau) = (s\sigma)\tau$ para todo $s \in S$.

Nótese que el orden de los eventos se lee de izquierda a derecha; $\sigma \circ \tau$ se lee, primero efectúese σ y luego sigase con τ . Aquí también este asunto de izquierda-derecha no es uniforme. Los matemáticos que escriben sus aplicaciones a la izquierda leerán $\sigma \circ \tau$ significando que primero se ha de efectuar τ y luego σ . De acuerdo con esto resulta que al leer un libro de matemáticas uno debe estar completamente seguro de cuál es la convención que el libro sigue para escribir el producto de dos aplicaciones. Repetimos, *para nosotros $\sigma \circ \tau$ siempre significará: primero aplíquese σ y luego τ* .

Ilustramos la composición de σ y τ con unos cuantos ejemplos.

Ejemplo 1. Sea $S = \{x_1, x_2, x_3\}$ y sea $T = S$. Sea $\sigma: S \rightarrow S$ definida por $x_1\sigma = x_2$, $x_2\sigma = x_3$ y $x_3\sigma = x_1$; sea $\tau: S \rightarrow S$ dada por $x_1\tau = x_1$, $x_2\tau = x_3$ y $x_3\tau = x_2$. Entonces $x_1(\sigma \circ \tau) = (x_1\sigma)\tau = x_2\tau = x_3$, $x_2(\sigma \circ \tau) = (x_2\sigma)\tau = x_3\tau = x_2$, $x_3(\sigma \circ \tau) = (x_3\sigma)\tau = x_1\tau = x_1$. Al mismo tiempo podemos calcular $\tau \circ \sigma$ porque en este caso también esto tiene sentido. Tenemos entonces $x_1(\tau \circ \sigma) = (x_1\tau)\sigma = (x_1\sigma) = x_2$, $x_2(\tau \circ \sigma) = (x_2\tau)\sigma = x_3\sigma = x_1$, $x_3(\tau \circ \sigma) = (x_3\tau)\sigma = x_2\sigma = x_3$. Nótese que $x_2 = x_1(\tau \circ \sigma)$, mientras que $x_3 = x_1(\sigma \circ \tau)$, de donde $\sigma \circ \tau \neq \tau \circ \sigma$.

Ejemplo 2. Sea S el conjunto de los enteros, T el conjunto $S \times S$, y supongamos que $\sigma: S \rightarrow T$ está definido por $m\sigma = (m-1, 1)$. Sea $U = S$ y supongamos que $\tau: T \rightarrow U (= S)$ está definida por $(m, n)\tau = m+n$. Tenemos, entonces, $\sigma \circ \tau: S \rightarrow S$ mientras que $\tau \circ \sigma: T \rightarrow T$; incluso hablar de la igualdad de $\sigma \circ \tau$ y $\tau \circ \sigma$ no tendría sentido, ya que no actúan sobre el mismo espacio. Vamos ahora a calcular $\sigma \circ \tau$ como una aplicación de S en sí mismo y luego $\tau \circ \sigma$ como uno de T en sí mismo.

Dado $m \in S$, $m\sigma = (m-1, 1)$, de donde $m(\sigma \circ \tau) = (m\sigma)\tau = (m-1, 1)\tau = (m-1)+1 = m$. Luego $\sigma \circ \tau$ es la aplicación identidad (a la que, por cierto, también a veces llamaremos aplicación idéntica) de S en sí mismo. Y, ¿qué hay acerca de $\tau \circ \sigma$? Dado $(m, n) \in T$, $(m, n)\tau = m+n$, de donde $(m, n)(\tau \circ \sigma) = ((m, n)\tau)\sigma = (m+n)\sigma = (m+n-1, 1)$. Nótese que $\tau \circ \sigma$ no es la aplicación idéntica de T en sí mismo; no es ni, incluso, una aplicación suprayectiva T .

Ejemplo 3. Sea S el conjunto de los números reales, T el conjunto de los enteros, y $U = \{E, 0\}$. Definamos $\sigma: S \rightarrow T$ por $s\sigma = \text{máximo entero menor que o igual a } s$, y $\tau: T \rightarrow U$ por $n\tau = E$ si n es par, $n\tau = 0$ si n es impar. Nótese que, en este, caso $\tau \circ \sigma$ no puede definirse. Calculamos $\sigma \circ \tau$ para dos números reales $s = \frac{8}{3}$ y $s = \pi$. Como $\frac{8}{3} = 2 + \frac{2}{3}$, $(\frac{8}{3})\sigma = 2$, de donde $(\frac{8}{3})(\sigma \circ \tau) = (\frac{8}{3}\sigma)\tau = (2)\tau = E$; $(\pi)\sigma = 3$, de donde $\pi(\sigma \circ \tau) = (\pi\sigma)\tau = (3)\tau = 0$.

Para las aplicaciones de conjuntos, siempre que los productos requeridos tengan sentido, se verifica una ley asociativa general. Este es el contenido del

LEMA 1.1 (LEY ASOCIATIVA). Si $\sigma: S \rightarrow T$, $\tau: T \rightarrow U$ y $\mu: U \rightarrow V$, entonces $(\sigma \circ \tau) \circ \mu = \sigma \circ (\tau \circ \mu)$.

Prueba. Nótese primero que $\sigma \circ \tau$ tiene sentido y lleva S en U , por tanto, $(\sigma \circ \tau) \circ \mu$ tiene también sentido y lleva S en V . Análogamente, $\sigma \circ (\tau \circ \mu)$ está definido y lleva S en V . Podemos, pues, hablar de igualdad o no igualdad de $(\sigma \circ \tau) \circ \mu$ y $\sigma \circ (\tau \circ \mu)$.

Para probar la igualdad afirmada debemos, simplemente, demostrar que para cualquier $s \in S$, $s((\sigma \circ \tau) \circ \mu) = s(\sigma \circ (\tau \circ \mu))$. Pero, de acuerdo con la definición de composición de aplicaciones, $s((\sigma \circ \tau) \circ \mu) = (s(\sigma \circ \tau))\mu = ((s\sigma)\tau)\mu$ mientras que $s(\sigma \circ (\tau \circ \mu)) = (s\sigma)(\tau \circ \mu) = ((s\sigma)\tau)\mu$. Luego, los elementos $s((\sigma \circ \tau) \circ \mu)$ y $s(\sigma \circ (\tau \circ \mu))$ son ciertamente iguales. Lo que prueba el lema.

Ahora demostraremos que si dos aplicaciones σ y τ tienen propiedades de cierto tipo y σ está definido, τ tiene también esas propiedades.

LEMA 1.2. Sean $\sigma: S \rightarrow T$ y $\tau: T \rightarrow U$, entonces:

- 1) $\sigma \circ \tau$ es suprayectiva si tanto σ como τ lo son;
- 2) $\sigma \circ \tau$ es inyectiva si tanto σ como τ lo son.

Prueba. Probaremos solo la parte (2) dejando la prueba de la parte (1) como ejercicio.

Supongamos que $s_1, s_2 \in S$ y que $s_1 \neq s_2$. Como σ es inyectiva, $s_1\sigma \neq s_2\sigma$. Como τ es inyectiva y $s_1\sigma$ y $s_2\sigma$ son elementos distintos de T , $(s_1\sigma)\tau \neq (s_2\sigma)\tau$ de donde $s_1(\sigma \circ \tau) = (s_1\sigma)\tau \neq (s_2\sigma)\tau = s_2(\sigma \circ \tau)$, probando que, ciertamente, $\sigma \circ \tau$ es inyectiva con lo que queda probado el lema.

Supongamos que σ es una aplicación inyectiva de S sobre T ; llamaremos entonces a σ correspondencia biyectiva entre S y T . Dada una $t \in T$ cualquiera, por ser σ suprayectiva existe un elemento $s \in S$ tal que $t = s\sigma$; por ser inyectiva esta s ha de ser única. Definimos la aplicación $\sigma^{-1} : T \rightarrow S$ por $s = t\sigma^{-1}$ si y sólo si $t = s\sigma$. La aplicación σ^{-1} se le llama inversa de σ . Calculemos la aplicación de S en sí mismo $\sigma \circ \sigma^{-1}$. Dado $s \in S$ sea $t = s\sigma$, de donde, por definición, $s = t\sigma^{-1}$; así pues $s(\sigma \circ \sigma^{-1}) = (s\sigma)\sigma^{-1} = t\sigma^{-1} = s$. Y hemos demostrado que $\sigma \circ \sigma^{-1}$ es la aplicación identidad de S sobre sí misma. Un cálculo análogo nos revela que $\sigma^{-1} \circ \sigma$ es la aplicación identidad de T sobre sí mismo.

Recíprocamente, si $\sigma : S \rightarrow T$ es tal que existe un $\mu : T \rightarrow S$ con la propiedad de que $\sigma \circ \mu$ y $\mu \circ \sigma$ son las aplicaciones idénticas sobre S y T respectivamente, entonces afirmamos que σ es una correspondencia biyectiva entre S y T . Observemos primero que σ es suprayectiva, pues dada $t \in T$, $t = t(\mu \circ \sigma) = (t\mu)\sigma$ (ya que $\mu \circ \sigma$ es la identidad sobre T), luego t es la imagen bajo σ del elemento $t\mu$ en S . Observemos luego que σ es inyectiva pues si $s_1\sigma = s_2\sigma$, usando $\sigma \circ \mu$ es la identidad sobre S , tenemos $s_1 = s_1(\sigma \circ \mu) = (s_1\sigma)\mu = (s_2\sigma)\mu = s_2(\sigma \circ \mu) = s_2$. Con lo que hemos probado el

LEMA 1.3. *La aplicación $\sigma : S \rightarrow T$ es una correspondencia biyectiva entre S y T si y sólo si existe una aplicación $\mu : T \rightarrow S$ tal que $\sigma \circ \mu$ y $\mu \circ \sigma$ son las aplicaciones identidad sobre S y T , respectivamente.*

DEFINICIÓN. Si S es un conjunto no vacío entonces $A(S)$ es el conjunto de todas las aplicaciones biyectivas de S sobre sí mismo.

A parte de su propio interés intrínseco $A(S)$ juega un papel central y universal cuando se considera el sistema matemático conocido con el nombre de grupo (capítulo 2). Es por esto por lo que damos a continuación el siguiente teorema sobre su naturaleza. Todas las partes constituyentes del teorema han sido probadas en los distintos lemas, de modo que enunciamos el teorema sin prueba.

TEOREMA 1.B. *Si σ, τ, μ son elementos de $A(S)$, entonces:*

- 1) $\sigma \circ \tau$ está en $A(S)$;
- 2) $(\sigma \circ \tau) \circ \mu = \sigma \circ (\tau \circ \mu)$;
- 3) existe un elemento ι (la aplicación idéntica) en $A(S)$ tal que $\sigma \circ \iota = \iota \circ \sigma = \sigma$;
- 4) existe un elemento $\sigma^{-1} \in A(S)$ tal que $\sigma \circ \sigma^{-1} = \sigma^{-1} \circ \sigma = \iota$.

Terminamos esta sección con una observación sobre $A(S)$. Supongamos que $A(S)$ tiene más de dos elementos; sean x_1, x_2, x_3 tres elementos distintos de S ; definamos la aplicación $\sigma : S \rightarrow S$ por $x_1\sigma = x_2, x_2\sigma = x_3, x_3\sigma = x_1$, $s\sigma = s$ para cualquier $s \in S$ distinta de x_1, x_2, x_3 . Definamos la

aplicación $\tau : S \rightarrow S$ por $x_2\tau = x_3$, $x_3\tau = x_2$ y $s\tau = s$ para cualquier $s \in S$ diferente de x_2 , x_3 . Es claro que tanto σ como τ están en $A(S)$. Un simple cálculo demuestra que $x_1(\sigma \circ \tau) = x_3$, pero que $x_1(\tau \circ \sigma) = x_2 \neq x_3$. Luego $\sigma \circ \tau \neq \tau \circ \sigma$. Esto es,

LEMA 1.4. Si S tiene más de dos elementos, podemos encontrar dos elementos σ , τ en $A(S)$ tales que $\sigma \circ \tau \neq \tau \circ \sigma$.

Problemas

1. Determinese en cada uno de los siguientes casos si $\sigma : S \rightarrow T$ es suprayectiva, inyectiva y determinese la imagen inversa de $t \in T$ cualquiera bajo σ .

- a) $S =$ conjunto de los números reales, $T =$ conjunto de los números reales no negativos, $s\sigma = s^2$.
- b) $S =$ conjunto de los números reales no negativos, $T =$ conjunto de los números reales no negativos, $s\sigma = s^2$.
- c) $S =$ conjunto de los enteros, $T =$ conjunto de los enteros, $s\sigma = s^2$.
- d) $S =$ conjunto de los enteros, $T =$ conjunto de los enteros, $s\sigma = 2s$.

2. Si S y T son conjuntos no vacíos, pruébese que existe una correspondencia biyectiva entre $S \times T$ y $T \times S$.

3. Si S , T y U son conjuntos no vacíos, pruébese que existe una correspondencia biyectiva entre:

- a) $(S \times T) \times U$ y $S \times (T \times U)$.
- b) Cada uno de los conjuntos de la parte (a) y el conjunto de ternas ordenadas (s, t, u) donde $s \in S$, $t \in T$, $u \in U$.
- 4. a) Si hay una correspondencia biyectiva entre S y T , pruébese que hay una correspondencia biyectiva entre T y S .
- b) Si hay una correspondencia biyectiva entre S y T y una entre T y U , pruébese que hay una correspondencia biyectiva entre S y U .

5. Si ι es la aplicación idéntica sobre S , pruébese que para cualquier σ en $A(S)$, $\sigma \circ \iota = \iota \circ \sigma = \sigma$.

*6. Si S es un conjunto cualquiera, pruébese que es *imposible* encontrar una aplicación de S sobre S^* .

7. Si el conjunto S tiene n elementos, pruébese que $A(S)$ tiene $n!$ (factorial de n) elementos.

8. Si el conjunto S tiene un número finito de elementos, pruébese que:
- a) Si σ transforma S sobre S entonces σ es inyectiva.

- b) Si σ es una aplicación inyectiva de S en sí mismo, entonces σ es suprayectiva.
 c) Pruébese, con un ejemplo, que tanto la parte (a) como la (b) son falsas si S no tiene un número finito de elementos.

9. Pruébese que los recíprocos de ambas partes del lema 1.2 son falsos, es decir:

- a) Si $\sigma \circ \tau$ es suprayectiva, no es necesario que ambas σ y τ lo sean.
 b) Si $\sigma \circ \tau$ es inyectiva, no es necesario que ambas σ y τ sean inyectivas.

10. Pruébese que hay una correspondencia biyectiva entre el conjunto de los enteros y el conjunto de los números racionales.

11. Si $\sigma : S \rightarrow T$ y si A es un subconjunto de S , la *restricción de σ a A* , σ_A , está definida por $a\sigma_A = a$ para cualquier $a \in A$. Pruébese que:

- a) σ_A define una aplicación de A en T .
 b) σ_A es inyectiva si lo es σ .
 c) σ_A puede muy bien ser inyectiva aunque no lo sea σ .

12. Si $\sigma : S \rightarrow T$ y A es un subconjunto de S tal que $A\sigma \subset A$, pruébese que $(\sigma \circ \tau)_A = \sigma_A \circ \tau_A$. $\tau \in ?$

13. Un conjunto S se dice que es *infinito* si hay una correspondencia biyectiva entre S y un subconjunto propio de S . Pruébese que:

- a) El conjunto de los enteros es infinito.
 b) El conjunto de los números reales es infinito.
 c) Si un conjunto S tiene un subconjunto A que es infinito, entonces S debe ser infinito.

(Nota: De acuerdo con el resultado del problema 8, un conjunto finito, en el sentido usual, no es infinito.)

*14. Si S es infinito y puede ponerse en correspondencia biyectiva con el conjunto de los enteros, pruébese que hay una correspondencia biyectiva entre S y $S \times S$.

*15. Dados dos conjuntos S y T decimos que $S < T$ (S es más pequeño que T) si hay una aplicación de T sobre S , pero *ninguna* de S sobre T . Pruébese que si $S < T$ y $T < U$ entonces $S < U$.

16. Si S y T son conjuntos finitos de m y n elementos respectivamente, pruébese que si $m < n$ entonces $S < T$.

3. LOS ENTEROS

Finalizamos este capítulo con una breve discusión del conjunto de los enteros. No haremos intento alguno de construirlos axiomáticamente, suponiendo, en lugar de ello, que ya tenemos el conjunto de los enteros y

que conocemos muchas de sus propiedades elementales. Entre ellas incluimos el principio de inducción matemática (que usaremos libremente a través de este libro) y el hecho de que un conjunto no vacío de enteros positivos siempre contiene un elemento mínimo. En cuanto a notación, los símbolos familiares $a > b$, $a \leq b$, $|a|$, etc., aparecerán con su significado habitual. Para evitar la repetición una y otra vez de que algo es un entero, convenimos en que *todos los símbolos de esta sección escritos en letras bastardillas minúsculas, representarán enteros.*

Dados a y b , con $b \neq 0$, podemos dividir a por b para obtener un residuo no negativo r que es menor en tamaño que b ; es decir, podemos encontrar m y r tales que $a = mb + r$ donde $0 \leq r < |b|$. Este hecho se conoce como el *algoritmo de Euclides* y suponemos que le es familiar al lector.

Decimos que $b \neq 0$ divide a a si $a = mb$ para algún m . Para indicar que b divide a a escribiremos $b|a$, y para indicar que b no divide a a , $b \nmid a$. Nótese que si $a|1$ entonces $a = \pm 1$, que cuando $a|b$ y también $b|a$ entonces $a = \pm b$, y que cualquier $b \neq 0$ divide a 0. Si $b|a$, llamamos a b un divisor de a . Nótese que si b es un divisor de g y de h , entonces es un divisor de $mg + nh$ para enteros arbitrarios m y n . Dejamos la verificación de todas estas afirmaciones como ejercicio.

DEFINICIÓN. El entero positivo c se dice que es el *máximo común divisor* de a y b si:

- 1) c es un divisor de a y de b ;
- 2) cualquier divisor de a y b es un divisor de c .

Usaremos la notación (a, b) para representar al máximo común divisor de a y b . Como se exige que el máximo común divisor sea positivo, $(a, b) = (a, -b) = (-a, b) = (-a, -b)$. Por ejemplo, $(60, 24) = (60, -24) = 12$. Otro comentario: el simple hecho de que hayamos definido lo que debe entenderse por máximo común divisor, no garantiza que éste exista. Tendremos que probar su existencia. Sin embargo, podemos decir que si existe entonces ha de ser único, pues si tuviéramos c_1 y c_2 que satisficieran ambas condiciones de la anterior definición entonces $c_1|c_2$ y $c_2|c_1$, de donde tendríamos $c_1 = \pm c_2$; pero la exigencia de positividad nos da $c_1 = c_2$. Nuestra primera tarea es, pues, la de demostrar la existencia de (a, b) . Al hacerlo, en el próximo lema, probaremos realmente un poco más, a saber, que (a, b) debe tener una forma particular.

LEMA 1.5. Si a y b son enteros, no ambos cero, entonces (a, b) existe; podemos, además, encontrar enteros m_0 y n_0 tales que $(a, b) = m_0a + n_0b$.

Pruéba. Sea \mathfrak{M} el conjunto de todos los enteros de la forma $ma + nb$ donde m y n toman valores enteros cualesquiera. Como uno de los dos a y b es distinto de cero, hay enteros distintos de cero en \mathfrak{M} . Como $x = ma + nb$ está en \mathfrak{M} , $-x = (-m)a + (-n)b$ también está en \mathfrak{M} ; por

tanto, \mathfrak{M} siempre tiene algunos números positivos. Pero entonces hay en \mathfrak{M} un entero positivo mínimo c ; por estar en \mathfrak{M} , c tiene la forma $c = m_0a + n_0b$. Afirmamos que $c = (a, b)$.

Notemos primero que si $d|a$ y $d|b$, entonces $d|(m_0a + n_0b)$, de donde $d|c$. Debemos ahora demostrar que $c|a$ y $c|b$. Dado un elemento cualquiera $x = ma + nb$ de \mathfrak{M} , por el algoritmo de Euclides tenemos $x = tc + r$ donde $0 \leq r < c$. Es decir, escribiéndolo explícitamente, $ma + nb = t(m_0a + n_0b) + r$, donde $r = (m - tm_0)a + (n - tn_0)b$ y debe, por tanto, estar en \mathfrak{M} . Como $0 \leq r < c$, por la elección de c , $r = 0$. Así pues, $x = tc$; hemos probado que $c|x$ para cualquier $x \in \mathfrak{M}$. Pero $a = 1a + 0b \in \mathfrak{M}$ y $b = 0a + 1b \in \mathfrak{M}$, de donde $c|a$ y $c|b$.

Hemos probado que c satisface las propiedades requeridas para ser (a, b) con lo que hemos probado el lema.

DEFINICIÓN. Los enteros a y b son *primos relativos* si $(a, b) = 1$.

Como una consecuencia inmediata del lema 1.5 tenemos el

COROLARIO. Si a y b son primos relativos, podemos encontrar enteros m y n tales que $ma + nb = 1$.

Introduciremos ahora otro concepto familiar. El de número primo. Entenderemos por esto un entero que no tiene una factorización distinta de las triviales. Por razones técnicas excluimos el 1 del conjunto de números primos. En la sucesión 2, 3, 5, 7, 11, todos son números primos; también en $-2, -3, -5, -7, -11$ son todos números primos. Como al factorizar, que un número sea negativo no introduce ninguna diferencia esencial, para nosotros los números primos serán siempre positivos.

DEFINICIÓN. El entero $p > 1$ es un *número primo* si sus únicos divisores son ± 1 y $\pm p$.

Otra forma de enunciar esto es decir que un entero p (mayor que 1) es un número primo si y sólo si dado otro entero n cualquiera, entonces $(p, n) = 1$ o $p|n$. Como veremos pronto, los números primos son los bloques de construcción de los enteros. Pero primero necesitamos la siguiente observación importante.

LEMA 1.6. Si a es un primo relativo a b , pero $a|bc$, entonces $a|c$.

Prueba. Como a y b son primos relativos, podemos encontrar, por el corolario al lema 1.5, enteros m y n tales que $ma + nb = 1$. Así pues $mac + nbc = c$. Pero $a|mac$ y, por hipótesis, $a|nbc$; luego $a|(mac + nbc)$. Como $mac + nbc = c$, concluimos que $a|c$, que es precisamente lo que el lema afirma.

Como consecuencia inmediata del lema y de la definición de número primo está el importante

COROLARIO. Si un número primo divide al producto de ciertos enteros, debe dividir al menos a uno de estos enteros.

Dejamos al lector la prueba del corolario.

Hemos afirmado que los números primos sirven como bloques de construcción para el conjunto de los enteros. El enunciado preciso de esto se da en el "teorema de factorización única".

TEOREMA I.C. Cualquier entero positivo $a > 1$ puede factorizarse en forma única como $a = p_1^{x_1} p_2^{x_2} \dots p_t^{x_t}$, donde $p_1 > p_2 > \dots > p_t$ son números primos y donde cada $x_i > 0$.

Prueba. El teorema en la forma en que lo hemos enunciado consiste realmente en dos subteoremas; el primero afirma la posibilidad de factorizar el entero dado como un producto de potencias de primos; el segundo nos asegura que esta descomposición es única. Probaremos el teorema mismo probando cada uno de estos dos subteoremas separadamente.

Se nos presenta de inmediato la pregunta: ¿cómo nos arreglaremos para probar el teorema? Un método natural de ataque consiste en el empleo de la inducción matemática. Unas pocas palabras acerca de éste; usaremos la siguiente versión de la inducción matemática: si la proposición $P(m_0)$ es cierta y si el que $P(r)$ sea cierta para toda r tal que $m_0 \leq r < k$ implica la verdad de $P(k)$, entonces $P(n)$ es cierta para todo $n \geq m_0$. Esta variante del principio de inducción puede mostrarse que es una consecuencia de la propiedad básica de los enteros, que afirma que cualquier conjunto no vacío de enteros positivos tiene un elemento mínimo (véase el problema 10).

Probamos primero que todo entero $a > 1$ puede factorizarse como un producto de factores potencias de primos; nuestra forma de atacar al problema es mediante la inducción matemática.

Ciertamente, $m_0 = 2$, como es un número primo, tiene una representación como un producto de potencias de primos.

Supongamos que cualquier entero r , $2 \leq r < k$ pueda ser factorizado como un producto de potencias de primos. Si el mismo k es un número primo, entonces es un producto de potencias de primos. Si no, es un número primo, entonces $k = uv$ donde $1 < u < k$ y $1 < v < k$. Por la hipótesis de inducción, como tanto u como v son menores que k , cada uno de ellos puede factorizarse como un producto de potencias de primos. Luego $k = uv$ es también un producto tal. Hemos demostrado que la verdad de la proposición para todos los enteros r , $2 \leq r < k$, implica su verdad para k . Por consiguiente, por el básico principio de inducción, la proposición es cierta para todos los enteros $n \geq m_0 = 2$; es decir, todo entero $n \geq 2$ es un producto de potencias de primos.

Pasemos ahora al problema de la unicidad. También aquí usaremos la inducción matemática, y en la forma en que la hemos acabado de usar.

Supongamos que

$$a = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r} = q_1^{\beta_1} q_2^{\beta_2} \dots q_s^{\beta_s}$$

donde $p_1 > p_2 > \dots > p_r$, $q_1 > q_2 > \dots > q_s$ son números primos, y donde cada $\alpha_i > 0$ y cada $\beta_i > 0$. Nuestro objetivo es probar que:

- 1) $r = s$;
- 2) $p_1 = q_1, p_2 = q_2, \dots, p_r = q_s$;
- 3) $\alpha_1 = \beta_1, \alpha_2 = \beta_2, \dots, \alpha_r = \beta_s$.

Para $a = 2$ esto es claramente cierto. Procediendo por inducción suponemos que es cierto para todos los enteros u , $2 \leq u < a$. Ahora bien,

$$a = p_1^{\alpha_1} \dots p_r^{\alpha_r} = q_1^{\beta_1} \dots q_s^{\beta_s}$$

y como $\alpha_1 > 0$, $p_1 | a$, luego $p_1 | q_1^{\beta_1} \dots q_s^{\beta_s}$. Pero como p_1 es un número primo, según el corolario al lema 1.6, se sigue fácilmente que $p_1 = q_i$ para algún i . Luego $q_i \geq q_1 = p_1$. Análogamente, como $q_1 | a$ llegamos a la conclusión de que $q_1 = p_j$ para algún j , de donde $p_1 \geq p_j = q_1$. En resumen, hemos demostrado que $p_1 = q_1$. Por tanto $a = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r} = p_1^{\beta_1} q_2^{\beta_2} \dots q_s^{\beta_s}$. Afirmamos que esto obliga a que $\alpha_1 = \beta_1$. (¡Pruébese!) Pero entonces $b = a/p_1^{\alpha_1} = p_2^{\alpha_2} \dots p_r^{\alpha_r} = q_2^{\beta_2} \dots q_s^{\beta_s}$. Si $b = 1$, entonces $\alpha_2 = \dots = \alpha_r = 0$ y $\beta_2 = \dots = \beta_s = 0$; es decir, $r = s = 1$, y el enunciado se verifica. Si $b > 1$, entonces como $b < a$ podemos aplicar nuestra hipótesis de inducción a b para obtener:

- 1) el número de distintos factores potencias de primos (en b) es igual en ambos lados, es decir, $r - 1 = s - 1$, luego $r = s$;
- 2) $\alpha_2 = \beta_2, \dots, \alpha_r = \beta_r$;
- 3) $p_2 = q_2, \dots, p_r = q_r$.

Junto con la información que ya hemos obtenido, a saber, $p_1 = q_1$ y $\alpha_1 = \beta_1$, es esto precisamente lo que estamos intentando probar. Vemos pues que la hipótesis de la unicidad de la factorización para los enteros menores que a implica la unicidad de la factorización para a . En consecuencia la inducción se ha completado y la afirmación de la unicidad de la factorización ha sido probada.

Cambiamos ahora un poco de dirección para estudiar la importante noción de congruencia módulo un entero dado. Como veremos más tarde la relación que ahora introducimos es un caso especial de una mucho más general que puede definirse en un contexto mucho más amplio.

DEFINICIÓN. Sea $n > 0$ un entero fijo. Definimos $a \equiv b$ mód n si $n | (a - b)$.

Nos referimos a esta relación como “congruencia módulo n ”, n se llama módulo de la relación, y a $a \equiv b$ mód n lo leemos “ a es congruente con b módulo n ”. Tenemos, por ejemplo, $73 \equiv 4$ mód 23 , $21 \equiv -9$ mód 10 , etc.

Esta relación de congruencia tiene las siguientes propiedades básicas:

LEMA 1.7

- 1) La relación “congruencia módulo n ” define una relación de equivalencia en el conjunto de los enteros.
- 2) Esta relación de equivalencia tiene n distintas clases de equivalencia.
- 3) Si $a \equiv b$ mód n y $c \equiv d$ mód n entonces $a+c \equiv b+d$ mód n y $ac \equiv bd$ mód n .
- 4) Si $ab \equiv ac$ mód n y a es primo relativo con n , entonces $b \equiv c$ mód n .

Prueba. Verificamos primero que la relación “congruencia módulo n ” es una relación de equivalencia. Como $n \mid 0$, tenemos, ciertamente, que $n \mid (a-a)$ de donde $a \equiv a$ mód n para toda a . Además, si $a \equiv b$ mód n entonces $n \mid (a-b)$, y, por tanto, $n \mid (b-a) = -(a-b)$; luego $b \equiv a$ mód n . Finalmente, si $a \equiv b$ mód n y $b \equiv c$ mód n , entonces $n \mid (a-b)$ y $n \mid (b-c)$, de donde $n \mid \{(a-b)+(b-c)\}$, es decir, $n \mid (a-c)$. Esto, desde luego, implica que $a \equiv c$ mód n .

Denotemos a la clase de equivalencia de esta relación a la que pertenece a por el símbolo $[a]$; y la llamamos *clase de congruencia* (mód n) de a . Dado un entero cualquiera a , por el algoritmo euclíadiano $a = kn+r$ donde $0 \leq r < n$. Pero entonces $a \in [r]$, luego $[a] = [r]$. Hay, por tanto, cuando más n distintas clases de congruencia; a saber, $[0], [1], \dots, [n-1]$. Pero éstas son distintas, pues si $[i] = [j]$ y, digamos, $0 \leq i < j < n$, entonces $n \mid (j-i)$ donde $j-i$ es un entero positivo menor que n , lo que es obviamente imposible. Hay, por consiguiente, exactamente n distintas clases de congruencia $[0], [1], \dots, [n-1]$. Y ya tenemos probadas las aserciones (1) y (2) del lema.

Probemos ahora la parte (3). Supongamos que $a \equiv b$ mód n y que $c \equiv d$ mód n ; luego $n \mid (a-b)$ y $n \mid (c-d)$. Tenemos entonces $n \mid \{(a-b)+(c-d)\}$ y por tanto $n \mid \{(a+c)-(b+d)\}$. Pero entonces $a+c \equiv b+d$ mód n . Además $n \mid \{(a-b)c + (c-d)b\} = ac - bd$, de donde $ac \equiv bd$ mód n .

Observemos finalmente que si $ab \equiv ac$ mód n y si a es primo con n , entonces el hecho de que $n \mid a(b-c)$, por el lema 1.6, implica que $n \mid (b-c)$ y por tanto que $b \equiv c$ mód n .

Si a no es un primo relativo con n , el resultado de la parte (4) puede ser falso; por ejemplo, $2 \cdot 3 \equiv 4 \cdot 3$ mód 6, aunque $2 \not\equiv 4$ mód 6.

El lema 1.7 nos abre ciertas posibilidades interesantes. Sea J_n el conjunto de las clases de congruencia mód n ; es decir, $J_n = \{[0], [1], \dots, [n-1]\}$. Dados dos elementos $[i]$ y $[j]$ en J_n , definamos:

$$\begin{aligned} a) \quad [i]+[j] &= [i+j]; \\ b) \quad [i][j] &= [ij]. \end{aligned}$$

Afirmamos que el lema nos asegura que esta “adición” y esta “multiplicación” están bien definidas; es decir, que si $[i] = [i']$ y $[j] = [j']$, entonces

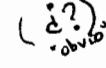
$[i]+[j] = [i'+j'] = [i'+j'] = [i']+[j']$ y que $[i][j] = [i'][j']$. (¡Verifíquese!) Estas operaciones en J_n tienen las siguientes interesantes propiedades (cuyas pruebas dejamos como ejercicios): para cualesquiera $[i], [j], [k]$ en J_n

- 1) $[i]+[j] = [j]+[i]$
- 2) $[i][j] = [j][i]$ } leyes commutativas;
- 3) $([i]+[j])+[k] = [i]+([j]+[k])$
- 4) $([i][j])[k] = [i]([j][k])$ } leyes asociativas;
- 5) $[i]([j]+[k]) = [i][j]+[i][k]$ ley distributiva;
- 6) $[0]+[i] = [i];$
- 7) $[1][i] = [i].$

* Una observación más: si $n = p$ es un número primo y si $[a] \neq [0]$ está en J_p , entonces hay un elemento $[b]$ en J_p tal que $[a][b] = [1]$.

El conjunto J_n juega un papel importante en álgebra y teoría de números. Se llama conjunto de enteros mód n ; antes de que sigamos mucho más lejos habremos tenido ocasión de conocerlo bastante bien.

Problemas

1. Si $a|b$ y $b|a$, pruébese que $a = \pm b$.
2. Si b es un divisor de g y de h , pruébese que es un divisor de $mg + nh$.
3. Si a y b son enteros, el *mínimo común múltiplo* de a y b , al que representaremos por $[a, b]$, está definido como el entero positivo d tal que: 1) $a|d$ y $b|d$.
2) Siempre que $a|x$ y $b|x$, entonces $d|x$.
4. Si $a|x$ y $b|x$ y $(a, b) = 1$ pruébese que $(a, b)|x$. 
5. Si $a = p_1^{\alpha_1} \dots p_k^{\alpha_k}$ y $b = p_1^{\beta_1} \dots p_k^{\beta_k}$ donde los p_i son números primos distintos y donde todo $\alpha_i \geq 0$, y todo $\beta_i \geq 0$, pruébese que:
1) $(a, b) = p_1^{\delta_1} \dots p_k^{\delta_k}$ donde $\delta_i = \min\{\alpha_i, \beta_i\}$ para cada i .
2) $[a, b] = p_1^{\gamma_1} \dots p_k^{\gamma_k}$ donde $\gamma_i = \max\{\alpha_i, \beta_i\}$ para cada i .
6. Dadas a, b , al aplicar el algoritmo eucliano sucesivamente tenemos:

$$a = q_0b + r_1, \quad 0 \leq r_1 < |b|$$

$$b = q_1r_1 + r_2, \quad 0 \leq r_2 < r_1$$

$$r_1 = q_2r_2 + r_3, \quad 0 \leq r_3 < r_2$$

⋮

$$r_k = q_{k+1}r_{k+1} + r_{k+2}, \quad 0 \leq r_{k+2} < r_{k+1}.$$

Como los enteros r_k son decrecientes y todos son no negativos, hay un primer entero n tal que $r_{n+1} = 0$. Pruébese que $r_n = (a, b)$. (Consideramos aquí $r_0 = |b|$.)

7. Úsese el método del problema 6 para calcular:

- a) (1128, 33).
- b) (6540, 1206).

8. Para comprobar que n es un número primo, pruébese que es suficiente demostrar que no es divisible por ningún primo p tal que $p \leq \sqrt{n}$.

9. Pruébese que $n > 1$ es un primo si y sólo si para cualquier a o $(a, n) = 1$ o $n|a$.

10. Suponiendo que todo conjunto no vacío de enteros positivos tiene un elemento mínimo, pruébese que:

- a) Si la proposición P es tal que

- 1) $P(m_0)$ es cierta
 - 2) la verdad de $P(m-1)$ implica la verdad de $P(m)$
- entonces $P(n)$ es cierta para toda $n \geq m_0$.
- b) Si la proposición P es tal que
 - 1) $P(m_0)$ es cierta
 - 2) $P(m)$ es cierta siempre que $P(a)$ es cierta para toda a tal que $m_0 \leq a < m$,
- entonces $P(n)$ es cierta para toda $n \geq m_0$.

11. Pruébese que la adición y la multiplicación usadas en J_n están bien definidas.

12. Pruébense las propiedades (1-7) para la adición y la multiplicación en J_n .

13. Si $(a, n) = 1$, pruébese que uno puede encontrar $[b] \in J_n$ tal que $[a][b] = [1]$ en J_n .

*14. Si p es un número primo, pruébese que para cualquier entero a , $a^p \equiv a \pmod p$.

15. Si $(m, n) = 1$, dadas a y b , pruébese que existe una x tal que $x \equiv a \pmod m$ y $x \equiv b \pmod n$.

16. Pruébese el corolario al lema 1.6.

17. Pruébese que n es un número primo si y sólo si en J_n $[a][b] = [0]$ implica que $[a] = [0]$ o $[b] = [0]$.

Lecturas supplementarias

Para conjuntos y números cardinales:

BIRKHOFF, G. y MACLANE, S., *A Brief Survey of Modern Algebra*. The Macmillan Company, Nueva York, 1953.

CAPÍTULO 2

Teoría de grupos

EN ESTE capítulo emprenderemos el estudio del objeto algebraico conocido como "grupo" que sirve como uno de los bloques de construcción fundamentales de la gran estructura que hoy se llama álgebra abstracta. En capítulos posteriores echaremos una mirada a algunos de los otros, tales como anillos, campos, espacios vectoriales y álgebras lineales. Aparte de que ya se ha hecho tradicional comenzar con el estudio de los grupos, hay razones naturales convincentes para esta elección. Para comenzar, los grupos, como sistemas con una sola operación, se prestan a la más simple de las descripciones formales. Sin embargo, a pesar de esta simplicidad de descripción los conceptos fundamentales del álgebra tales como homomorfismo, construcción cociente, etc., que juegan un papel tan importante

en todas las estructuras algebraicas —en realidad, en todas las matemáticas— entran aquí en una forma pura y reveladora.

Permítasenos en este punto, antes de que los detalles nos abrumen, echar una rápida ojeada al camino que vamos a recorrer. En el álgebra abstracta tenemos ciertos sistemas básicos que, en la historia y el desarrollo de las matemáticas, han alcanzado posiciones de importancia extraordinaria. Éstos son, generalmente, conjuntos con cuyos elementos podemos operar algebraicamente —por lo que entendemos que podemos combinar dos elementos del conjunto, quizá de varias maneras, para obtener un tercer elemento también del conjunto— y, además, suponemos que estas operaciones algebraicas están sujetas a ciertas reglas que se indican explícitamente en lo que se llaman axiomas o postulados definitorios del sistema. En este marco abstracto, intentaremos probar teoremas acerca de estas mismas estructuras generales, esperando siempre que cuando estos resultados se apliquen a una realización particular y concreta del sistema abstracto, afluirán hechos y conocimientos de la estructura interna del ejemplo que se discuta y que habrían quedado oscurecidos para nosotros por el volumen de información sin importancia que se nos presenta en todo caso particular.

Nos gustaría subrayar que estos sistemas algebraicos y los axiomas que los definen deben tener cierta naturalidad. Deben surgir de la experiencia que resulta de observar muchos ejemplos; deben ser ricos en resultados significativos. Sentarse, hacer una lista de unos cuantos axiomas y proceder al estudio del sistema así descrito, no resulta un procedimiento adecuado de trabajo en matemáticas. Admitimos que esto es lo que hacen algunos, pero la mayor parte de los matemáticos descartarán estos ensayos como matemáticas mediocres. Los sistemas que se estudian, son estudiados porque casos particulares de tales estructuras han aparecido una y otra vez, porque alguien, finalmente, notó que estos casos particulares eran realmente concreciones de un fenómeno general, porque alguien nota analogías entre dos objetos matemáticos aparentemente disímiles y ello le dirige hacia una investigación sobre las raíces de estas analogías. Para citar un ejemplo, hacia finales del siglo XVIII y comienzos del XIX se estaba estudiando caso tras caso de este objeto matemático que hoy conocemos como grupo; pero, sin embargo, no fue sino hasta ya bastante avanzado el siglo XIX que se introdujo la noción de grupo abstracto. Las únicas estructuras algebraicas hasta ahora encontradas que han resistido el embate del tiempo y han sobrevivido y crecido en importancia, son las basadas en un amplio y alto pilar de casos particulares. Entre matemáticos, nadie discute ni la belleza ni la importancia del primer ejemplo que, para discutir, hemos elegido los grupos.

1. DEFINICIÓN DE GRUPO

Parece aconsejable que en este punto recordemos una situación discutida en el primer capítulo. Dado un conjunto arbitrario no vacío S definimos $A(S)$ como el conjunto de todas las aplicaciones *biyectivas* del conjunto S sobre sí mismo. Para cualesquier dos elementos $\sigma, \tau \in A(S)$ introducimos un producto al que representábamos por $\sigma \circ \tau$, y una investigación posterior nos mostraba la certeza de los siguientes hechos acerca de los elementos de $A(S)$ sometidos a este producto:

- 1) Siempre que $\sigma, \tau \in A(S)$, entonces se sigue que $\sigma \circ \tau$ está también en $A(S)$. Describimos esto diciendo que $A(S)$ es *cerrado* respecto al producto (a veces decimos, "cerrado respecto a la multiplicación").
- 2) Para cualesquier tres elementos $\sigma, \tau, \mu \in A(S)$, $\sigma \circ (\tau \circ \mu) = (\sigma \circ \tau) \circ \mu$. A esta relación se le llama *ley asociativa*.
- 3) Hay un elemento muy especial $\iota \in A(S)$ que satisface $\iota \circ \sigma = \sigma \circ \iota = \sigma$ para todo $\sigma \in A(S)$. A tal elemento se le llama *elemento identidad* de $A(S)$.
- 4) Para todo $\sigma \in A(S)$ hay un elemento, al que representamos por σ^{-1} , también en $A(S)$, tal que $\sigma \circ \sigma^{-1} = \sigma^{-1} \circ \sigma = \iota$. Esta situación generalmente se describe diciendo que todo elemento de $A(S)$ tiene un *inverso* en $A(S)$.

Se verifica también otro hecho acerca de $A(S)$; a saber, que siempre que S tiene tres o más elementos podemos encontrar dos elementos $\alpha, \beta \in A(S)$ tales que $\alpha \circ \beta \neq \beta \circ \alpha$. Esta posibilidad que contradice nuestra experiencia e intuición matemática habituales, introduce una riqueza en $A(S)$ de la que, a no ser por ello, habría carecido.

Con este ejemplo como modelo y un gran conocimiento de muchas situaciones matemáticas y mucha abstracción construimos la siguiente

DEFINICIÓN. Un conjunto no vacío de elementos G se dice que forma un *grupo* si en G está definida una operación binaria, llamada producto y denotada por (\cdot) tal que:

- 1) $a, b \in G$ implica que $a \cdot b \in G$ (cierre).¹
- 2) $a, b, c \in G$ implica que $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ (ley asociativa).
- 3) Existe un elemento $e \in G$ tal que $a \cdot e = e \cdot a = a$ para todo $a \in G$ (existencia de un elemento identidad en G).
- 4) Para todo $a \in G$ existe un elemento $a^{-1} \in G$ tal que $a \cdot a^{-1} = a^{-1} \cdot a = e$ (existencia de inversos en G).

¹ Cuando los "resultados" de una operación entre elementos de un conjunto son también elementos del conjunto, también es frecuente decir que el conjunto es "estable" respecto a la operación. [N. del T.]

Considerando el origen de esta definición no es nada sorprendente que para todo conjunto no vacío S , el conjunto $A(S)$ sea un grupo. Hemos, pues, presentado ya una inagotable fuente de interesantes grupos concretos. Veremos posteriormente (en un teorema debido a Cayley) que estos $A(S)$ constituyen, en cierto sentido, una familia universal de grupos. Si S tiene tres o más elementos, recordemos que pueden encontrarse elementos $\sigma, \tau \in A(S)$ tales que $\sigma \circ \tau \neq \tau \circ \sigma$. Nos induce esto a remarcar una clase de grupos muy especiales, pero muy importantes, en la siguiente definición.

DEFINICIÓN. Un grupo G se dice que es *abeliano* (o *comutativo*) si para cualesquier $a, b \in G$ se tiene: $a \cdot b = b \cdot a$.

Un grupo que no es abeliano se llama, lo que parece bastante natural, *no abeliano*; como hemos visto, una familia de ejemplos de tales grupos sabemos muy bien que los grupos no abelianos ciertamente existen.

Otra característica natural de un grupo G es el número de elementos de que consta. Llamamos a este *orden* de G y lo denotamos por $o(G)$. Este número es, desde luego, más interesante cuando es finito. En tal caso decimos que G es un *grupo finito*.

Para ver que existen grupos finitos que no son triviales nótese que si el conjunto S contiene n elementos, entonces el grupo $A(S)$ tiene n elementos. (! Pruébese !) Este ejemplo, ciertamente muy importante, se denotará siempre que aparezca en este libro por S_n . Le llamaremos *grupo simétrico* de grado n . En la próxima sección haremos una disección más o menos completa de S_3 que es un grupo no abeliano de orden 6.

2. ALGUNOS EJEMPLOS DE GRUPOS

EJEMPLO 1. Supongamos que G está constituido por el conjunto de los enteros $0, \pm 1, \pm 2, \dots$ con $a \cdot b$, para $a, b \in G$, definida como la suma usual entre enteros, es decir, con $a \cdot b = a + b$. El lector puede verificar fácilmente que G es un grupo abeliano infinito en el que 0 juega el papel de e , y $-a$ el de a^{-1} .

EJEMPLO 2. Supongamos que G consiste en los números reales 1 y -1 con la multiplicación entre números reales como operación. G es entonces, un grupo abeliano de orden 2.

EJEMPLO 3. Sea $G = S_3$, el grupo de todas las aplicaciones biyectivas del conjunto $\{x_1, x_2, x_3\}$ sobre sí mismo, con el producto que definimos en el capítulo 1. G es un grupo de orden 6. Haremos una pequeña digresión antes de volver a S_3 .

Para tener una notación más clara, no solamente en S_3 sino en cualquier grupo G , definamos para cualquier $a \in G$, $a^0 = e$, $a^1 = a$, $a^2 = a \cdot a$,

$a^3 = a \cdot a^2, \dots, a^k = a \cdot a^{k-1}$, y $a^{-2} = (a^{-1})^2$, $a^{-3} = (a^{-1})^3$, etc. El lector puede verificar que las reglas habituales de los exponentes siguen teniendo validez, es decir, que para dos enteros cualesquiera (positivos, negativos o nulos) m, n ,

$$1) \quad a^m \cdot a^n = a^{m+n}$$

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

(Vale la pena hacer notar que, en esta notación, si G es el grupo del ejemplo 1, a^n representa al entero na .)

Con esta notación a nuestra disposición, examinemos S_3 más de cerca. Consideremos la aplicación ϕ definida sobre el conjunto x_1, x_2, x_3 por

$$\begin{aligned}\phi: \quad x_1 &\rightarrow x_2 \\ &x_2 \rightarrow x_1 \\ &x_3 \rightarrow x_3\end{aligned}$$

y la aplicación

$$\begin{aligned}\psi: \quad x_1 &\rightarrow x_2 \\ &x_2 \rightarrow x_3 \\ &x_3 \rightarrow x_1\end{aligned}$$

Podemos fácilmente comprobar que $\phi^2 = e$, $\psi^3 = e$ y que

$$\begin{aligned}\phi \cdot \psi: \quad x_1 &\rightarrow x_3 \\ &x_2 \rightarrow x_2 \\ &x_3 \rightarrow x_1\end{aligned}$$

mientras que

$$\begin{aligned}\psi \cdot \phi: \quad x_1 &\rightarrow x_1 \\ &x_2 \rightarrow x_3 \\ &x_3 \rightarrow x_2\end{aligned}$$

Es claro que $\phi \cdot \psi \neq \psi \cdot \phi$ puesto que no llevan a x_1 a la misma imagen. Como $\psi^3 = e$, se sigue que $\psi^{-1} = \psi^2$. Calculemos ahora la acción de $\psi^{-1} \cdot \phi$ sobre x_1, x_2, x_3 . Como $\psi^{-1} = \psi^2$ y

$$\begin{aligned}\psi^2: \quad x_1 &\rightarrow x_3 \\ &x_2 \rightarrow x_1 \\ &x_3 \rightarrow x_2\end{aligned}$$

tenemos que

$$\begin{aligned}\psi^{-1} \cdot \phi: \quad x_1 &\rightarrow x_3 \\ &x_2 \rightarrow x_2 \\ &x_3 \rightarrow x_1\end{aligned}$$

En otras palabras, $\phi \cdot \psi = \psi^{-1} \cdot \phi$. Consideremos los elementos $e, \phi, \psi, \psi^2, \phi \cdot \psi, \psi \cdot \phi$; todos ellos son distintos y pertenecen a G (puesto que G es cerrado), que solamente tiene seis elementos. Así pues, esta lista enumera todos los elementos de G . Se podría preguntar, por ejemplo, ¿cuál es el

elemento de la lista igual a $\psi \cdot (\phi \cdot \psi)$? Usando $\phi \cdot \psi = \psi^{-1} \cdot \phi$ vemos que $\psi \cdot (\phi \cdot \psi) = \psi \cdot (\psi^{-1} \cdot \phi) = (\psi \cdot \psi^{-1}) \cdot \phi = e \cdot \phi = \phi$. De más interés es la forma de $(\phi \cdot \psi) \cdot (\psi \cdot \phi) = \phi \cdot (\psi \cdot (\psi \cdot \phi)) = \phi \cdot (\psi^2 \cdot \phi) = \phi \cdot (\psi^{-1} \cdot \phi) = \phi \cdot (\phi \cdot \psi) = \phi^2 \cdot \psi = e \cdot \psi = \psi$. (El lector no debe asustarse de la larga y tediosa cadena de igualdades que aquí aparecen. Va a ser ésta la última vez que seamos tan tediosamente concienzudos.) Usando las mismas técnicas que hemos empleado, el lector puede calcular para alegrarse de cualquier otro u otros de los 25 productos que no implican a e . Algunos de ellos aparecerán en los ejercicios.

EJEMPLO 4. Sea n un entero cualquiera. Construimos un grupo de orden n como sigue: G consistirá en todos los símbolos a^i , $i = 0, 1, 2, \dots, n-1$ en donde insistimos en que $a^0 = a^n = e$, $a^i \cdot a^j = a^{i+j}$ si $i+j \leq n$ y $a^i \cdot a^j = a^{i+j-n}$ si $i+j > n$. El lector puede verificar que esto es un grupo. Se le conoce como *grupo cíclico* de orden n .

Una realización geométrica del grupo del ejemplo 4 puede obtenerse como sigue: sea S la circunferencia en el plano de radio 1, y sea ρ_n una rotación de un ángulo de $2\pi/n$ (radianes). Entonces $\rho_n \in A(S)$ genera un grupo de orden n ; a saber, $\{e, \rho_n, \rho_n^2, \dots, \rho_n^{n-1}\}$.

3. ALGUNOS LEMAS PRELIMINARES

Ya hemos dedicado a la teoría de los grupos varias páginas y aún no hemos probado ni un solo hecho acerca de ellos. Ya es tiempo de que remediemos esta situación. Aunque hemos de admitir que los primeros resultados que vamos a probar no son muy excitantes (la verdad es que son bastante insípidos), veremos en seguida que son extraordinariamente útiles. El aprendizaje del alfabeto no fue probablemente la parte más interesante de nuestra educación infantil, sin embargo, una vez que lo dominamos, qué panoramas más fascinantes se abrieron ante nosotros.

Comenzamos con el

LEMA 2.1. *Si G es un grupo, entonces*

- el elemento identidad de G es único;*
- todo $a \in G$ tiene un inverso único en G ;*
- para todo $a \in G$, $(a^{-1})^{-1} = a$;*
- para $a, b \in G$, $(a \cdot b)^{-1} = b^{-1} \cdot a^{-1}$.*

Prueba. Antes de que comencemos propiamente con la prueba parece aconsejable que veamos qué es lo que vamos a probar. En la parte (a) queremos probar que si dos elementos e y f de G gozan de la propiedad de que para todo $a \in G$, $a = a \cdot e = e \cdot a = a \cdot f = f \cdot a$, entonces $e = f$. En la parte (b) nuestro objetivo es demostrar que si $x \cdot a = a \cdot x = e$ y

$y \cdot a = a \cdot y = e$, donde todos los tres a , x , y están en G , entonces $x = y$.

Consideremos primero la parte (a). Como $e \cdot a = a$ para todo $a \in G$, tenemos que, en particular, $e \cdot f = f$. Pero, por otra parte, $b \cdot f = b$ para todo $b \in G$, luego debemos tener $e \cdot f = e$. Juntando estas dos fracciones de información obtenemos $f = e \cdot f = e$, y, por tanto, $e = f$.

En lugar de probar la parte (b) probaremos algo más fuerte que nos traerá inmediatamente la parte (b) como consecuencia. Supongamos que para a en G , $a \cdot x = e$ y $a \cdot y = e$; entonces, obviamente, $a \cdot x = a \cdot y$. Hagamos de esto nuestro punto de partida. Es decir, supongamos que $a \cdot x = a \cdot y$ con a , x , y en G . Hay un elemento b en G tal que $b \cdot a = e$ (por lo que hasta ahora sabemos puede que haya varios de tales b). Por tanto, $b \cdot (a \cdot x) = b \cdot (a \cdot y)$. Usando la ley asociativa esto nos lleva a que

$$x = e \cdot x = (b \cdot a) \cdot x = b \cdot (a \cdot x) = b \cdot (a \cdot y) = (b \cdot a) \cdot y = e \cdot y = y.$$

Hemos probado, en realidad, que en un grupo G , $a \cdot x = a \cdot y$ implica que $x = y$. Análogamente, podemos probar que $x \cdot a = y \cdot a$ implica que $x = y$. Esto quiere decir que en los grupos podemos cancelar, siempre que sea del mismo lado, en las ecuaciones. Pero debe tenerse presente que de que $a \cdot x = y \cdot a$ no puede concluirse que $x = y$, pues no tenemos medio alguno de saber que $a \cdot x = x \cdot a$. Queda esto ilustrado en S_3 con $a = \phi$, $x = \psi$, $y = \psi^{-1}$.

La parte (c) se sigue de esto si observamos que $a^{-1} \cdot (a^{-1})^{-1} = e = a^{-1} \cdot a$; cancelando la a^{-1} a la izquierda nos da $(a^{-1})^{-1} = a$. Esto es, para grupos en general, lo análogo del resultado familiar, digamos por ejemplo, $-(-5) = 5$, en los números reales respecto a la suma.

La parte (d) es la más trivial de todas, pues $(a \cdot b) \cdot (b^{-1} \cdot a^{-1}) = a \cdot ((b \cdot b^{-1}) \cdot a^{-1}) = a \cdot (e \cdot a^{-1}) = a \cdot a^{-1} = e$, y luego, de acuerdo con la definición de inverso, $(a \cdot b)^{-1} = b^{-1} \cdot a^{-1}$.

Ciertos resultados obtenidos en la prueba son de suficiente importancia para enunciarlos expresamente como hacemos ahora en el

LEMA 2.2. *Dados a , b en el grupo G , entonces las ecuaciones $a \cdot x = b$ y $y \cdot a = b$ tienen soluciones únicas para x y y en G . En particular, las dos leyes de cancelación,*

$$a \cdot u = a \cdot w \text{ implica } u = w$$

y

$$u \cdot a = w \cdot a \text{ implica } u = w$$

se verifican en G .

Dejamos al cuidado del lector los pocos detalles necesarios para la prueba de este lema.

Problemas

1. Determíñese, en cada uno de los siguientes casos, si el sistema descrito es o no un grupo. En caso negativo, señálense cuál o cuáles de los axiomas de grupo no se verifican.

- a) $G = \text{conjunto de todos los enteros, } a \cdot b = a - b.$
- b) $G = \text{conjunto de todos los enteros positivos, } a \cdot b = ab, \text{ el producto habitual de los enteros.}$
- c) $G = a_0, a_1, \dots, a_6 \text{ donde}$
 $a_i \cdot a_j = a_{i+j} \text{ si } i+j < 7$
 $a_i \cdot a_j = a_{i+j-7} \text{ si } i+j \geq 7$
 $\text{(por ejemplo } a_5 \cdot a_4 = a_{5+4-7} = a_2 \text{ ya que } 5+4 = 9 > 7).$
- d) $G = \text{conjunto de todos los números racionales con denominadores impares, } a \cdot b = a+b, \text{ la adición usual de los números racionales.}$

2. Pruébese que si G es un grupo abeliano, entonces, para todo $a, b \in G$ y todos los enteros n , $(a \cdot b)^n = a^n \cdot b^n$.

3. Si G es un grupo tal que $(a \cdot b)^2 = a^2 \cdot b^2$ para todo $a, b \in G$ demuéstrese que G ha de ser abeliano.

*4. Si G es un grupo en el cual $(a \cdot b)^i = a^i \cdot b^i$ para tres enteros consecutivos i y para todos los $a, b \in G$, demuéstrese que G es abeliano.

5. Pruébese que la conclusión del problema 4 no tiene validez si suponemos la relación $(a \cdot b)^i = a^i \cdot b^i$ solamente para dos enteros consecutivos.

6. Proporcionese en S_3 un ejemplo de dos elementos x, y tales que $(x \cdot y)^2 \neq x^2 \cdot y^2$.

7. Pruébese que en S_3 hay cuatro elementos que satisfacen $x^2 = e$ y tres elementos que satisfacen $y^3 = e$.

8. Si G es un grupo finito, pruébese que existe un entero positivo N tal que $a^N = e$ para todo $a \in G$.

- 9. a) Si el grupo G tiene tres elementos, pruébese que ha de ser abeliano.
 b) Hágase la parte (a) si G tiene cuatro elementos.
 c) Hágase la parte (a) si G tiene cinco elementos.

10. Demuéstrese que si todo elemento de G es su propio inverso, entonces G es abeliano.

11. Si G es un grupo de orden par, pruébese que tiene un elemento $a \neq e$ tal que $a^2 = e$.

12. Sea G un conjunto no vacío cerrado respecto a un producto asociativo que además satisface:

- a) Existe un $e \in G$ tal que $a \cdot e = a$ para toda $a \in G$.
- b) Dado $a \in G$, existe un elemento $y(a) \in G$ tal que $a \cdot y(a) = e$.

Pruébese que G debe ser un grupo bajo este producto.

13. Pruébese, mediante un ejemplo, que la conclusión del problema 12 es falsa si en lugar de las anteriores hipótesis hubiésemos supuesto que:

- a') Existe un $e \in G$ tal que $a \cdot e = a$ para toda $a \in G$.
- b') Dada $a \in G$ existe $y(a) \in G$ tal que $y(a) \cdot a = e$.

14. Supongamos que un conjunto *finito* G es cerrado respecto a un producto asociativo y que las dos leyes de cancelación se verifican en G . Pruébese que G debe ser un grupo.

15. a) Usando el resultado del problema 14, pruébese que los enteros no cero módulo p , con p un número primo, forman un grupo respecto a la multiplicación módulo p .
 b) Pruébese lo dicho en la parte (a) para los enteros no cero primos relativos con n bajo la multiplicación mód n .

16. En el problema 14 demuéstrese, mediante un ejemplo, que si solo se supone la validez de una de las leyes de cancelación; entonces la conclusión no necesariamente se sigue.

17. Pruébese que en el problema 14 existe una infinidad de ejemplos que satisfacen las condiciones y no son grupos.

18. Para un $n > 2$ cualquiera constrúyase un grupo abeliano de orden $2n$
(Sugerencia: imítense las relaciones en S_3 .)

19. Si S es un conjunto cerrado respecto a una operación asociativa pruébese que no importa cómo pongamos los paréntesis en $a_1 a_2 \dots a_n$, pero conservando el orden de los elementos, siempre se obtiene el mismo elemento en S [por ejemplo, $(a_1 \cdot a_2) \cdot (a_3 \cdot a_4) = a_1 \cdot (a_2 \cdot (a_3 \cdot a_4))$; úsese la inducción sobre n].

4. SUBGRUPOS

Antes de volver al estudio de los grupos, desearíamos cambiar nuestra notación ligeramente. Es molesto seguir usando el (\cdot) para la operación de grupo; de aquí en adelante prescindiremos de él y en lugar de escribir $a \cdot b$ para $a, b \in G$ denotaremos tal producto simplemente por ab .

En general, no estaremos interesados en subconjuntos arbitrarios de un grupo G pues no reflejan el hecho de que G tiene una estructura algebraica. Todos los subconjuntos que consideraremos serán de los que tengan propiedades algebraicas derivadas de las de G . Los subconjuntos más naturales de entre los de tales tipos se introducen en la siguiente

DEFINICIÓN. Un subconjunto H de un grupo G se dice que es *subgrupo* de G si respecto al producto en G , H mismo forma un grupo.

Hacemos notar que es claro que si H es un subgrupo de G , y K es un subgrupo de H , entonces K es un subgrupo de G .

Sería útil tener algún criterio para decidir si un subconjunto dado d ϵ un grupo es un subgrupo. Es este el propósito de los próximos dos lemas.

LEMA 2.3. *Un subconjunto no vacío H del grupo G es un subgrupo de G si y sólo si*

- 1) $a, b \in H$ implica que $ab \in H$;
- 2) $a \in H$ implica que $a^{-1} \in H$.

Prueba. Si H es un subgrupo de G , entonces es obvio que (1) y (2) deben verificarse.

Supongamos, recíprocamente, que H es un subconjunto de G para el que se verifican (1) y (2). Para confirmar que H es un subgrupo todo lo que se necesita verificar es que $e \in H$ y que la ley asociativa se verifica para los elementos de H . Como la ley asociativa es válida para G , es claro que también es válida para H que es un subconjunto de G . Si $a \in H$, según (2) $a^{-1} \in H$, luego de acuerdo con (1) $e = aa^{-1} \in H$. Y esto completa la prueba.

En el caso especial de un grupo finito, la situación se hace aún más sencilla, pues en tal caso podemos prescindir de la condición (2).

LEMA 2.4. *Si H es un subconjunto finito no vacío de un grupo G , y H es cerrado respecto a la multiplicación, entonces H es un subgrupo de G .*

Prueba. A la luz del lema 2.3 no necesitamos otra cosa que demostrar que siempre que $a \in H$, entonces $a^{-1} \in H$. Supongamos que $a \in H$; entonces $a^2 = aa \in H$, $a^3 = a^2a \in H$, ..., $a^m \in H$, ... ya que H es, por hipótesis, cerrado. Luego la colección infinita de elementos $a, a^2, \dots, a^m, \dots$ debe estar contenida en H , que es un subconjunto finito de G . Luego debe haber repeticiones en esta colección de elementos; es decir, para algunos enteros r, s con $r > s > 0$, $a^r = a^s$. Por la cancelación en G , $a^{r-s} = e$ (de donde e está en H); como $r-s-1 \geq 0$, $a^{r-s-1} \in H$ y $a^{-1} = a^{r-s-1}$ ya que $aa^{r-s-1} = a^{r-s} = e$. Por tanto, $a^{-1} \in H$, lo que completa la prueba del lema.

El lema nos dice que para comprobar si un subconjunto de un grupo finito es un subgrupo, lo único que necesitamos ver es que sea cerrado respecto a la multiplicación.

Ahora, quizá debamos ver algunos grupos y algunos de sus subgrupos. G es siempre un subgrupo de sí mismo; también el conjunto consistente tan solo en e es un subgrupo de G . Ninguno de ellos es particularmente interesante en su función de subgrupo, de modo que los describimos como subgrupos triviales. Los subgrupos entre estos dos extremos, a los que llamaremos subgrupos no triviales, son los que más nos interesarán.

EJEMPLO 1. Sea G el grupo de los enteros bajo la adición, H el subconjunto consistente en todos los múltiplos de 5. El lector debe comprobar que H es un subgrupo.

En este ejemplo no hay nada extraordinario acerca del 5; de manera semejante podríamos definir el subgrupo H_n como el subconjunto de G consistente en todos los múltiplos de n . H_n es, para cualquier n , siempre un subgrupo. ¿Qué es lo que puede decirse acerca de $H_n \cap H_m$? Quizá sea prudente probar con $H_6 \cap H_9$.

EJEMPLO 2. Sea S un conjunto cualquiera, $A(S)$ el conjunto de todas las aplicaciones biyectivas de S sobre sí mismo, que sabemos es un grupo bajo la composición de aplicaciones. Si $x_0 \in S$, sea $H(x_0) = \{\phi \in A(S) | x_0\phi = x_0\}$. $H(x_0)$ es un subgrupo de $A(S)$. Si para $x_1 \neq x_0 \in S$ definimos $H(x_1)$, ¿qué es $H(x_0) \cap H(x_1)$?

EJEMPLO 3. Sea G un grupo cualquiera y $a \in G$. Sea $(a) = \{a^i | i = 0, \pm 1, \pm 2, \dots\}$. (a) es un subgrupo de G (! verifíquese!); se llama *subgrupo cíclico generado por a*. Esto nos da una fácil fórmula para producir subgrupos de G . Si para una cierta elección de a , $G = (a)$, entonces decimos que G es un *grupo cíclico*. Tales grupos son muy particulares, pero juegan un papel muy importante en la teoría de grupos, especialmente en la parte que trata de los grupos abelianos. Desde luego, los grupos cíclicos son abelianos, pero la afirmación recíproca es falsa.

EJEMPLO 4. Sea G un grupo, y W un subconjunto de G . Sea (W) el conjunto de todos los elementos de G representables como un producto de elementos de W elevados a potencias de exponente positivo, negativo o cero. (W) es el *subgrupo de G generado por W* y es el mínimo subgrupo de G que contiene a W . En realidad, (W) es la intersección de todos los subgrupos de G que contienen W (hay al menos un tal subgrupo, pues G es un subgrupo de G que contiene a W).

DEFINICIÓN. Sea G un grupo, H un subgrupo de G ; para $a, b \in G$ decimos que a es congruente con b mód H , lo que escribimos: $a \equiv b$ mód H , si $ab^{-1} \in H$.

LEMA 2.5. La relación $a \equiv b$ mód H es una relación de equivalencia.

Prueba. De acuerdo con el capítulo I podemos ver que para probar el lema 2.5 debemos verificar las siguientes tres condiciones: para todas $a, b, c \in G$

- 1) $a \equiv a$ mód H ;
- 2) $a \equiv b$ mód H implica $b \equiv a$ mód H ;
- 3) $a \equiv b$ mód H , $b \equiv c$ mód H implica $a \equiv c$ mód H .

Estudiemos, una por una, estas condiciones.

- 1) Para probar que $a \equiv a$ mód H debemos probar, usando la definición

de congruencia mód H , que $aa^{-1} \in H$. Como H es un subgrupo de G , $e \in H$, y como $aa^{-1} = e$, $aa^{-1} \in H$, que es lo que se nos pedía demostrar.

2) Supongamos que $a \equiv b$ mód H , es decir, supongamos $ab^{-1} \in H$; queremos deducir de esto que $b \equiv a$ mód H , o, lo que es lo mismo, que $ba^{-1} \in H$. Como $ab^{-1} \in H$, que es un subgrupo de G , $(ab^{-1})^{-1} \in H$; pero según el lema 2.1, $(ab^{-1})^{-1} = ba^{-1}$, luego $ba^{-1} \in H$ y $b \equiv a$ mód H .

3) Finalmente, pedimos que $a \equiv b$ mód H y $b \equiv c$ mód H implique $a \equiv c$ mód H . La primera congruencia se traduce en $ab^{-1} \in H$, la segunda en $bc^{-1} \in H$; como H es un subgrupo de G , $(ab^{-1})(bc^{-1}) \in H$. Pero $ac^{-1} = aec^{-1} = a(b^{-1}b)c^{-1} = (ab^{-1})(bc^{-1})$, luego $ac^{-1} \in H$, de lo que se sigue que $a \equiv c$ mód H .

Esto determina que la congruencia módulo H sea una verdadera relación de equivalencia de acuerdo a lo definido en el capítulo 1, y todos los resultados acerca de relaciones de equivalencia pueden utilizarse en el examen de esta relación particular.

Unas palabras acerca de la notación usada. Si G fuera el grupo de enteros con la adición como operación, y $H = H_n$ el subgrupo consistente en todos los múltiplos de n , entonces, en G , la relación $a \equiv b$ mód H , es decir, $ab^{-1} \in H$, con la notación aditiva se leería $a - b$ es un múltiplo de n . Esta es la congruencia módulo n habitual de la teoría de los números. En otras palabras, la relación que hemos definido usando un grupo y un subgrupo arbitrarios es la generalización natural de una relación familiar en un grupo familiar.

DEFINICIÓN. Si H es un subgrupo de G , y $a \in G$, entonces $Ha = \{ha \mid h \in H\}$. A Ha se le llama *clase lateral derecha de H en G* .

LEMA 2.6. *Para todo $a \in G$,*

$$Ha = \{x \in G \mid a \equiv x \text{ mód } H\}.$$

Prueba. Sea $[a] = \{x \in G \mid a \equiv x \text{ mód } H\}$. Mostramos primero que $Ha \subset [a]$. Pues si $h \in H$, entonces $a(ha)^{-1} = a(a^{-1}h^{-1}) = h^{-1}$, luego está en H ya que H es un subgrupo de G . Según la definición de congruencia mód H esto implica que $ha \in [a]$ para todo $h \in H$, luego $Ha \subset [a]$.

Supongamos, ahora, que $x \in [a]$. Entonces $ax^{-1} \in H$, luego $(ax^{-1})^{-1} = xa^{-1}$ está también en H . Es decir, $xa^{-1} = h$ para algún $h \in H$. Multiplicando ambos lados por a a la derecha obtenemos $x = ha$, luego $x \in Ha$. Por tanto, $[a] \subset Ha$. Habiendo probado las dos inclusiones $[a] \subset Ha$ y $Ha \subset [a]$, podemos concluir que $[a] = Ha$, que es lo que afirma el lema.

En la terminología del capítulo 1, $[a]$, y por tanto Ha , es la clase de equivalencia de a en G . De acuerdo con el teorema 1.a estas clases de equivalencia constituyen una descomposición de G en subconjuntos ajenos. *Por tanto, dos clases laterales derechas de H en G o son idénticas o no tienen elemento común alguno.*

Afirmamos ahora que entre dos clases laterales derechas de H en G , Ha y Hb , existe una correspondencia biyectiva, a saber, la que a cada elemento $ha \in Ha$, donde $h \in H$, asocia el elemento $hb \in Hb$. Claramente esta aplicación es sobre Hb . Afirmamos que es una correspondencia biyectiva. En efecto, si $h_1a = h_2b$, con $h_1, h_2 \in H$, entonces por la ley de cancelación válida en G , $h_1 = h_2$ y, por tanto, $h_1a = h_2a$. Y esto prueba el

LEMA 2.7. *Hay una correspondencia biyectiva entre dos clases laterales derechas cualesquiera de H en G .*

El lema 2.7 es de máximo interés cuando H es un grupo finito, pues entonces lo que afirma es que dos clases laterales derechas de H tienen el mismo número de elementos. ¿Cuántos elementos tiene una clase lateral derecha de H ? Bien, nótese que $H = He$ es, él mismo, una clase lateral derecha de H , luego cualquier clase lateral derecha de H en G tiene $o(H)$ elementos. Supongamos ahora que G es un grupo finito y sea k el número de clases laterales derechas distintas de H en G . Según los lemas 2.6 y 2.7, dos clases laterales derechas cualesquiera distintas de H en G no tienen ningún elemento en común, y cada una de ellas tiene $o(H)$ elementos.

Como cualquier $a \in G$ está en una única clase lateral derecha, Ha , las clases laterales derechas llenan G . Luego si k representa el número de distintas clases laterales derechas de H en G , debemos tener que $ko(H) = o(G)$. Hemos probado un famoso teorema debido a Lagrange, a saber,

TEOREMA 2.A. *Si G es un grupo finito y H es un subgrupo de G , entonces $o(H)$ es un divisor de $o(G)$.*

DEFINICIÓN. Si H es un subgrupo de G , el índice de H en G es el número de distintas clases laterales derechas de H en G .

Representaremos al índice de H en G por $i_G(H)$. Si G es un grupo finito, $i_G(H) = \frac{o(G)}{o(H)}$, como vimos claramente en la prueba del teorema de Lagrange. Es muy posible que un grupo infinito G tenga un subgrupo $H \neq G$ que sea de índice finito en G .

Puede que a este nivel sea difícil para el estudiante ver la importancia extrema de este resultado. A medida que avancemos en la materia más y más, también más y más se irá dando cuenta de su carácter fundamental. A causa de su importancia, el teorema merece un escrutinio más estrecho, un poco más de análisis, así es que posteriormente damos una forma ligeramente diferente de llegar a su prueba. En realidad, el procedimiento en seguida delineado no es, en modo alguno, diferente del que ya hemos dado. La introducción del concepto de congruencia mód H facilitó la enumeración de elementos que usamos más adelante, y obvió la necesidad de comprobar que los nuevos elementos introducidos en cada etapa no aparecían antes.

Suponemos, pues, de nuevo que G es un grupo finito y que H es un subgrupo de él. Sea h_1, h_2, \dots, h_r una lista completa de los elementos de H , $r = o(H)$. Si $H = G$ no hay nada que probar. Supongamos entonces que $H \neq G$; hay, pues, entonces una $a \in G$, tal que $a \notin H$. Vamos a hacer ahora una lista de elementos en dos renglones como sigue

$$h_1, h_2, \dots, h_r$$

$$h_1 a, h_2 a, \dots, h_r a.$$

Afirmamos que todos los elementos del segundo renglón son diferentes uno de otro y también diferentes de cualesquiera de los elementos del primer renglón. Si dos cualesquiera del segundo renglón fueran iguales, entonces $h_i a = h_j a$ con $i \neq j$, pero, de acuerdo con la ley de cancelación, esto nos llevaría a que $h_i = h_j$, que es una contradicción. Si un elemento del segundo renglón fuera igual a uno del primero, entonces $h_i a = h_j$, de donde $a = h_i^{-1} h_j \in H$ ya que H es un subgrupo de G ; pero esto contradice que $a \notin H$.

Tenemos pues, hasta el momento, una lista de $2o(H)$ elementos; si estos elementos son todos los de G , ya hemos terminado la prueba. Si no, hay un $b \in G$ que no aparece en ninguno de los dos renglones. Consideremos la nueva lista

$$h_1, h_2, \dots, h_r$$

$$h_1 a, h_2 a, \dots, h_r a$$

$$h_1 b, h_2 b, \dots, h_r b.$$

Como antes, poco más o menos, podríamos mostrar que no hay en el tercer renglón dos elementos iguales y que ningún elemento del tercer renglón aparece en ninguno de los dos primeros. Tenemos, pues, en nuestra lista $3o(H)$ elementos. Continuando en esta forma, cada nuevo elemento introducido da lugar a $o(H)$ nuevos elementos. Como G es un grupo finito, terminaremos por agotar todos los elementos de G . Pero si terminamos usando k renglones para enumerar todos los elementos del grupo, habremos enumerado $ko(H)$ elementos distintos, de donde $ko(H) = o(G)$.

Es esencial señalar que el recíproco del teorema de Lagrange es falso —un grupo G no debe tener necesariamente un subgrupo de orden m solo porque m sea un divisor de $o(G)$. Por ejemplo, existe un grupo de orden 12 que no tiene subgrupo alguno de orden 6. El lector puede encontrar un ejemplo de este fenómeno; donde hay que buscarlo es en S_4 , el grupo simétrico de grado 4 que tiene un subgrupo de orden 12 que cumplirá nuestro requerimiento.

El teorema de Lagrange tiene algunos corolarios muy importantes. Pero antes de presentarlos daremos la siguiente definición.

DEFINICIÓN. Si G es un grupo y $a \in G$, el *orden* (o *periodo*) de a es el entero positivo mínimo m tal que $a^m = e$.

Si no existe tal entero decimos que a es de orden infinito. Usamos el símbolo $o(a)$ para indicar el orden de a . Recordemos otra de nuestras notaciones: para dos enteros u, v , $u|v$ se lee “ u es divisor de v ”.

COROLARIO 1. *Si G es un grupo finito y $a \in G$, entonces $o(a)|o(G)$.*

Prueba. Con el teorema de Lagrange ya a nuestra disposición, parece lo más natural probar el corolario exhibiendo un subgrupo de G cuyo orden sea $o(a)$. El elemento a mismo, nos proporciona un tal subgrupo que no es otro que el subgrupo cíclico (a) de G generado por a ; (a) consiste en e, a, a^2, \dots | ¿Cuántos elementos hay en (a) ? Afirmamos que este número es el orden de a . Claramente, como $a^{o(a)} = e$, este subgrupo tiene cuando más $o(a)$ elementos. Si tuviera realmente un número de elementos menor que este número, entonces tendríamos $a^i = a^j$ para algunos enteros $0 \leq i < j < o(a)$. De donde $a^{j-i} = e$, con $0 < j-i < o(a)$ lo que contradice el significado de $o(a)$. Así pues, el subgrupo cíclico generado por a tiene $o(a)$ elementos, de donde, según el teorema de Lagrange, $o(a)|o(G)$.

COROLARIO 2. *Si G es un grupo finito y $a \in G$, entonces $a^{o(G)} = e$.*

Prueba. De acuerdo con el corolario 1, $o(a)|o(G)$; luego $o(G) = mo(a)$. Por tanto, $a^{o(G)} = a^{mo(a)} = (a^{o(a)})^m = e^m = e$.

Un caso particular del corolario 2 es de gran interés en la teoría de números. Para todos los enteros n , se define la función de Euler ϕ mediante $\phi(1) = 1$, y para todos los $n > 1$, $\phi(n) =$ número de enteros positivos menores que n y primos respecto a n . Por ejemplo, $\phi(8) = 4$ ya que solamente 1, 3, 5 y 7 son los números menores que 8 que son primos con 8. En el problema 15(b) al final de la sección 3 se le pidió al lector que probase que los números menores que n y primos con n formaban un grupo respecto a la multiplicación mód n . Este grupo tiene orden $\phi(n)$. Si aplicamos el corolario 2 a este grupo, obtenemos el

COROLARIO 3 (DE EULER). *Si n es un entero positivo y a es primo con n , entonces $a^{\phi(n)} \equiv 1$ mód n .*

Para aplicar el corolario 2 se debe reemplazar a por su residuo al dividirlo por n . Si n fuese un número primo p , entonces $\phi(p) = p - 1$. Si a es un entero primo relativo con p , entonces, de acuerdo con el corolario 3, $a^{p-1} \equiv 1$ mód p , de donde $a^p \equiv a$ mód p . Si, por el contrario, a no es primo con p , como p es un número primo, debemos tener que $p|a$, de modo que $a \equiv 0$ mód p , de donde, $0 \equiv a^p \equiv a$ mód p también aquí. Por tanto

COROLARIO 4 (DE FERMAT). *Si p es un número primo y a es un entero cualquiera, entonces $a^p \equiv a$ mód p .*

COROLARIO 5. Si G es un grupo finito cuyo orden es un número primo p , entonces G es un grupo cíclico.

Prueba. Afirmamos primero que G sólo tiene como subgrupos los triviales, pues si H es un subgrupo de G , como $o(H)$ debe dividir a $o(G)$ solo quedan dos posibilidades, a saber, $o(H) = 1$ o $o(H) = p$. La primera de ellas implica $H = \{e\}$, mientras que la segunda implica $H = G$. Supongamos ahora que $a \neq e \in G$, y sea $H = \langle a \rangle$. H es un subgrupo de G y $H \neq \{e\}$ puesto que $a \neq e$ y $a \in H$. Luego $H = G$. Lo que nos dice que G es cíclico y que todo elemento de G es una potencia de a .

La sección 4 es de gran importancia para todo lo que sigue, no solo por sus resultados, sino porque el espíritu de las pruebas que aquí aparecen es muy característico de la teoría de grupos. El estudiante debe esperar otros argumentos muy semejantes a los que aquí ha encontrado. Sería muy sensato, por su parte, que asimilase ahora el material y el método de un modo completo, y no más adelante, cuando quizás sea demasiado tarde.

5. RELACIÓN ENTRE LOS NÚMEROS DE ELEMENTOS

Como antes hemos explicado, si H es un subgrupo de G y $a \in G$, entonces Ha consiste en todos los elementos de la forma ha donde $h \in H$. Generalicemos esta noción. Si H y K son dos subgrupos de G , sea $HK = \{x \in G \mid x = hk, h \in H, k \in K\}$. Hagamos una pausa y consideremos un ejemplo; en S_3 sea $H = \{e, \phi\}$, $K = \{e, \phi\psi\}$. Como $\phi^2 = (\phi\psi)^2 = e$, tanto H como K son subgrupos. ¿Qué puede decirse de HK ? Usando solamente la definición de HK puede verse que HK consiste en los elementos $e, \phi, \phi\psi, \phi^2\psi = \psi$. Como HK consta de cuatro elementos y 4 no es un divisor de 6, el orden de S_3 , según el teorema de Lagrange HK no puede ser un subgrupo de S_3 . (Desde luego, hubiéramos podido verificar esto directamente pero no es perjudicial seguir recordando el teorema de Lagrange.) Podíamos intentar encontrar por qué HK no es un subgrupo. Nótese que $KH = \{e, \phi, \phi\psi, \phi\psi\phi = \psi^{-1}\} \neq HK$. Es por esto precisamente por lo que HK no llega a ser un grupo, como vemos en el próximo lema.

LEMA 2.8. HK es un subgrupo de G si y sólo si $HK = KH$.

Prueba. Supongamos primero que $HK = KH$; es decir, que si $h \in H$ y $k \in K$, entonces $hk = k_1 h_1$ para algún $k_1 \in K$, $h_1 \in H$ (no es necesario que $k_1 = k$ o $h_1 = h$). Para probar que HK es un subgrupo, debemos verificar que es cerrado y que todo elemento de HK tiene su inverso en HK . Demostremos, en primer lugar, que es cerrado; supongamos, pues, $x = hk \in HK$ y $y = h'k' \in HK$. Entonces $xy = hkh'k'$, pero como $kh' \in KH = HK$, $kh' = h_2 k_2$ con $h_2 \in H$ y $k_2 \in K$. De donde $xy = h(h_2 k_2)k'$

$= (hh_2)(k_2 k') \in HK$, y HK es cerrado. Además $x^{-1} = (hk)^{-1} = k^{-1} h^{-1} \in \epsilon KH = HK$, luego $x^{-1} \in HK$. Luego HK es un subgrupo de G .

Por otra parte, si HK es un subgrupo de G , entonces para cualquier $h \in H$ y $k \in K$, $h^{-1} k^{-1} \in HK$ y por tanto $kh = (h^{-1} k^{-1})^{-1} \in HK$. Luego $HK \subset HK$. Si x es ahora un elemento cualquiera de HK , $x^{-1} = hk \in HK$, luego $x = (x^{-1})^{-1} = (hk)^{-1} = k^{-1} h^{-1} \in KH$, luego $HK \subset KH$. Luego $HK = KH$.

Un caso especial interesante es cuando G es un grupo abeliano, pues en ese caso evidentemente $HK = KH$. Luego, como consecuencia, tenemos el siguiente

COROLARIO. *Si H y K son subgrupos de un grupo abeliano G , entonces HK es un subgrupo de G .*

Si H y K son subgrupos de un grupo G , hemos visto que HK no necesariamente es un subgrupo de G . Tiene, sin embargo, sentido que nos preguntemos: ¿cuántos elementos distintos hay en el subconjunto HK ? Denotando tal número por $o(HK)$, probamos el

TEOREMA 2.B. *Si H y K son subgrupos finitos de G de órdenes $o(H)$ y $o(K)$ respectivamente, entonces*

$$o(HK) = \frac{o(H)o(K)}{o(H \cap K)}.$$

Prueba. Aunque no hay necesidad de prestar especial atención al caso particular en que $H \cap K = \{e\}$, considerar este caso, que carece de algunas de las complejidades, que se presentan en el caso general es ciertamente revelador. Lo que debemos demostrar aquí es que $o(HK) = o(H)o(K)$. Debería preguntarse, ¿cómo podría ser de otro modo? La contestación debe ser esta: si hacemos una lista de todos los elementos hk , con $h \in H$ y $k \in K$, algún elemento de la lista debería aparecer por lo menos dos veces. O, lo que es lo mismo, para algunos $h \neq h_1$ de H , debería haber k y k_1 de K tales que $hk = h_1 k_1$. Pero entonces $h_1^{-1} h = k_1 k^{-1}$; pero como $h_1 \in H$, h_1^{-1} debe también estar en H , luego $h_1^{-1} h \in H$. Análogamente, $k_1 k^{-1} \in K$. Como $h_1^{-1} h = k_1 k^{-1}$, $h_1^{-1} h \in H \cap K = \{e\}$, luego $h_1^{-1} h = e$, de donde $h = h_1$, una contradicción. Hemos, pues, probado que no puede haber ninguna coincidencia de valores, por tanto, en este caso tenemos, ciertamente, que $o(HK)$ es igual a $o(H)o(K)$.

Con esta experiencia, apoyándonos estamos listos para atacar al caso general. Como antes, debemos preguntarnos: ¿cuántas veces aparece un elemento hk dado como producto en la lista de HK ? Afirmamos que debe aparecer $o(H \cap K)$ veces. Para verlo, hacemos primero observar que si $h_1 \in H \cap K$, entonces

$$1) \quad hk = (hh_1)(h_1^{-1}k),$$

donde $hh_1 \in H$, como $h \in H$, $h_1 \in H \cap K \subset HK$ y $h^{-1}k \in K$ ya que $h_1^{-1} \in H \cap K \subset K$ y $k \in K$. Luego hk está duplicado en el producto, al menos $o(H \cap K)$ veces. Pero si $hk = h'k'$, entonces $h^{-1}h' = k(k')^{-1} = u$, y $u \in H \cap K$, y $h' = hu$ y $k' = u^{-1}k$; luego, todas las duplicaciones aparecieron en (1). En consecuencia hk aparece en la lista de HK exactamente $o(H \cap K)$ veces. Por tanto, el número de elementos distintos en HK es el número total de la lista de HK , que es $o(H)o(K)$, dividido por el número de veces que un elemento dado aparece, es decir, por $o(H \cap K)$. Lo que prueba el teorema.

Supongamos que H y K son subgrupos del grupo finito G y que $o(H) > \sqrt{o(G)}$, $o(K) > \sqrt{o(G)}$. Como $HK \subset G$, $o(HK) \leq o(G)$. Sin embargo,

$$o(G) \geq o(HK) = \frac{o(H)o(K)}{o(H \cap K)} > \frac{\sqrt{o(G)}\sqrt{o(G)}}{o(H \cap K)} = \frac{o(G)}{o(H \cap K)},$$

luego $o(H \cap K) > 1$. Luego $H \cap K \neq (e)$. Y hemos probado el siguiente

COROLARIO. Si H y K son subgrupos de G y $o(H) > \sqrt{o(G)}$ y $o(K) > \sqrt{o(G)}$, entonces $H \cap K \neq (e)$.

Aplicamos este corolario a un grupo muy especial. Supongamos que G es un grupo finito de orden pq donde p y q son números primos con $p > q$. Afirmamos que G puede tener cuando más un subgrupo de orden p . En efecto, supongamos que H y K son subgrupos de orden p . De acuerdo con el corolario $H \cap K \neq (e)$, y como subgrupo de H que por tener orden primo no tiene subgrupos que no sean triviales, debemos concluir que $H \cap K = H$. De igual modo puede verse que $H \cap K = K$. De donde $H = K$, probándose así que hay cuando más un subgrupo de orden p . Más adelante veremos que hay al menos un subgrupo de orden p , lo que, combinado con lo que acabamos de ver, nos dirá qué hay exactamente un subgrupo de orden p en G . Partiendo de este resultado podremos determinar completamente la estructura de G .

Problemas

1. Si H y K son subgrupos de G , entonces $H \cap K$ es un subgrupo de G .
2. Para un subgrupo H de G definase como clase lateral izquierda de H en G el conjunto de todos los elementos de la forma ah , $h \in H$. Pruébese que hay una correspondencia biyectiva entre el conjunto de clases laterales izquierdas de H en G y el conjunto de clases laterales derechas de H en G .
3. Si G no tiene subgrupos distintos de los triviales, pruébese que debe tener orden primo.
4. Sea G el grupo de los enteros con la adición como operación, y H_n el subgrupo consistente en todos los múltiplos de un entero fijo n . Determinense el índice de H_n en G y escríbanse todas las clases laterales de H_n en G .

5. En el problema 4, ¿qué es $H_n \cap H_m$?

*6. Si G es un grupo y H y K son dos subgrupos de G de índice finito en G , pruébese que $H \cap K$ es de índice finito en G . ¿No puede encontrar el lector una cota superior para el índice de $H \cap K$ en G ?

7. Supongamos que la aplicación τ_{ab} , para a, b números reales, transforma los reales en los números reales de acuerdo con la regla, $\tau_{ab}: x \rightarrow ax + b$. Sea $G = \{\tau_{ab} | a \neq 0\}$. Pruébese que G es un grupo bajo la composición de aplicaciones. Encuéntrese la fórmula para $\tau_{ab}\tau_{cd}$.

8. En el problema 7, sea $H = \{\tau_{ab} \in G | a \text{ es racional}\}$. Pruébese que H es un subgrupo de G . Hágase una lista de todas las clases laterales derechas de H en G .

9. a) En el problema 8 demuéstrese que toda clase lateral izquierda de H en G es una clase lateral derecha de H en G .

b) Proporciónese un ejemplo de un grupo G y un subgrupo H tales que no toda clase lateral izquierda de H en G sea una clase lateral derecha de H en G .

10. En el grupo del problema 7, sea $N = \{\tau_{1b} \in G\}$. Pruébese que:

a) N es un subgrupo de G .

b) Si $a \in G$ y $n \in N$, entonces $ana^{-1} \in N$.

11. Si un grupo abeliano tiene grupos de órdenes n y m , respectivamente, demuéstrese que tiene un subgrupo cuyo orden es el mínimo común múltiplo de n y m .

12. Si $a \in G$, definamos $N(a) = \{x \in G | xa = ax\}$. Demuéstrese que $N(a)$ es un subgrupo de G . $N(a)$ se llama generalmente *normalizador* o *centralizador* de a en G .

13. Si G es un grupo, el *centro* de G , Z , está definido por $Z = \{z \in G | zx = xz \text{ para todo } x \in G\}$. Pruébese que Z es un subgrupo de G .

14. Pruébese que cualquier subgrupo de un grupo cíclico es, él mismo, un grupo cíclico.

15. ¿Cuántos generadores tiene un grupo cíclico de orden n ? [$b \in G$ es un generador si $(b) = G$.]

16. Si $a \in G$ y $a^m = e$, pruébese que $o(a) | m$.

17. Si en el grupo G , $a^5 = e$ y $aba^{-1} = b^2$ para $a, b \in G$, encuéntrese $o(b)$.

*18. Sea G un grupo abeliano finito en el que el número de soluciones en G de la ecuación $x^n = e$ es, cuando más, n para todo entero positivo n . Pruébese que G debe ser un grupo cíclico. *fral. P100 Ex 11.18*

⑥ SUBGRUPOS NORMALES Y GRUPOS COCIENTE

Sea G el grupo S_3 y sea H el subgrupo $\{e, \phi\}$. Como el índice de H en G es 3, hay tres clases laterales derechas de H en G y tres clases laterales izquierdas de H en G . Hacemos sendas listas de ellas

	<u>Derechas</u>	<u>Izquierdas</u>
$P_0 = e$		
$P_1 = \psi$	$H = \{e, \phi\}$	$H = \{e, \phi\}$
$P_2 = \psi^2$	$H\psi = \{\psi, \phi\psi\}$	$\psi H = \{\psi, \psi\phi = \phi\psi^2\}$
$M_1 = \psi\phi$	$H\psi^2 = \{\psi^2, \phi\psi^2\}$	$\psi^2 H = \{\psi^2, \psi^2\phi = \phi\psi\}$
$M_2 = \psi^2\phi$		
$M_3 = \phi$		

Una rápida inspección nos muestra el hecho interesante de que la clase lateral derecha $H\psi$ no es una clase lateral izquierda. Luego, al menos para este subgrupo, las nociones de clase lateral derecha y clase lateral izquierda no necesariamente coinciden.

En $G = S_3$ consideremos el subgrupo $N = \{e, \psi, \psi^2\}$. Como el índice de N en G es 2, hay dos clases laterales izquierdas y dos clases laterales derechas de N en G . Hagamos sendas listas de ambas:

	<u>Derechas</u>	<u>Izquierdas</u>
	$N = \{e, \psi, \psi^2\}$	$N = \{e, \psi, \psi^2\}$
	$N\phi = \{\phi, \psi\phi, \psi^2\phi\}$	$\phi N = \{\phi, \phi\psi, \phi\psi^2\}$ $= \{\phi, \psi^2\phi, \psi\phi\}$

Una rápida inspección nos muestra aquí que toda clase lateral izquierda de N en G es una clase lateral derecha en G y reciprocamente. Vemos, así, que para algunos subgrupos la noción de clase lateral izquierda coincide con la de clase lateral derecha, mientras que para algunos otros subgrupos estos conceptos difieren.

Es un tributo al genio de Galois recordar que él fue el primero en reconocer que aquellos subgrupos para los que las clases laterales derechas e izquierdas coinciden, son subgrupos distinguidos. En matemáticas, frecuentemente el problema fundamental es el de reconocer y descubrir cuáles son los conceptos relevantes; una vez que esto se consigue, es muy posible que la tarea esté más que a medio hacer.

Vamos a definir esta clase especial de subgrupos de un modo ligeramente diferente que luego mostraremos que es equivalente al expresado en las observaciones anteriores.

DEFINICIÓN. Un subgrupo N de G se dice que es un *subgrupo normal* de G si para toda $g \in G$ y toda $n \in N$, $gng^{-1} \in N$.

Es claro que si por gNg^{-1} representamos el conjunto de todos los gng^{-1} ,

con $n \in N$, entonces N es un subgrupo normal de G si y sólo si $gNg^{-1} \subset N$ para todo $g \in G$.

LEMA 2.9. *N es un subgrupo normal de G si y sólo si $gNg^{-1} = N$ para todo $g \in G$.*

Prueba. Si $gNg^{-1} = N$ para toda $g \in G$, es claro que $gNg^{-1} \subset N$, luego N es normal en G .

Supongamos que N es normal en G . Entonces si $g \in G$, $gNg^{-1} \subset N$ y $g^{-1}Ng = g^{-1}N(g^{-1})^{-1} \subset N$. Pero como $g^{-1}Ng \subset N$, $N = g(g^{-1}Ng)g^{-1} \subset gNg^{-1} \subset N$, de donde $N = gNg^{-1}$.

Para evitar aquí una posible confusión, subrayamos que el lema 2.9 *no dice* que para todo $n \in N$ y para toda $g \in G$, $gng^{-1} = n$. No. Esto tal vez sea falso. Tomemos por ejemplo como grupo G el S_3 y como N el subgrupo $\{e, \psi, \psi^2\}$. Si calculamos $\phi N \phi^{-1}$ obtenemos $\{e, \phi\psi\phi^{-1}, \phi\psi^2\phi^{-1}\} = \{e, \psi^2, \psi\}$, pero, sin embargo, $\phi\psi\phi^{-1} \neq \psi$. Todo lo que necesitamos es que el conjunto de elementos gNg^{-1} sea el mismo que el conjunto de elementos N .

Volvemos ahora al problema de la igualdad de las clases laterales izquierdas y las clases laterales derechas.

LEMA 2.10. *El subgrupo N de G es un subgrupo normal de G si y sólo si toda clase lateral izquierda de N en G es una clase lateral derecha de N en G .*

Prueba. Si N es un subgrupo normal de G , entonces para todo $g \in G$, $gNg^{-1} = N$, de donde $(gNg^{-1})g = Ng$; esto es, $gN = Ng$, es decir, la clase lateral izquierda gN es la clase lateral derecha Ng .

Supongamos, recíprocamente, que toda clase lateral izquierda de N en G es una clase lateral derecha de N en G . Es decir, para $g \in G$, la clase lateral izquierda gN debe ser también una clase lateral derecha. ¿Pero cuál puede ser?

Como $g = ge \in gN$, cualquiera que sea la clase lateral derecha que resulte ser, gN debe contener al elemento g ; pero g está en la clase lateral derecha Ng y dos clases laterales derechas distintas no tienen ningún elemento en común. (¿Recuerda el lector la prueba del teorema de Lagrange?) Luego esta clase lateral derecha es única. De donde se sigue que $gN = Ng$. En otras palabras, $gNg^{-1} = Ngg^{-1} = N$, y N es, por tanto, un subgrupo normal de G .

Hemos ya definido lo que entendemos por HK siempre que H y K son subgrupos de G . Podemos extender fácilmente esta definición a subconjuntos arbitrarios, lo que hacemos definiendo para dos subconjuntos cualesquiera A y B de G , $AB = \{x \in G \mid x = ab, a \in A, b \in B\}$. Como un caso particular, ¿qué puede decirse cuando $A = B = H$ es un subgrupo de G ? Entonces

$HH = \{h_1 h_2 \mid h_1, h_2 \in H\} \subset H$, ya que H es cerrado respecto a la multiplicación. Pero $HH \supset He = H$ ya que $e \in H$. Luego $HH = H$.

Supongamos que N es un subgrupo normal de G y que $a, b \in G$. Consideremos $(Na)(Nb)$; como N es normal en G , $aN = Na$, y por tanto

$$NaNb = N(aN)b = N(Na)b = NNab = Nab.$$

¡Qué mundo de posibilidades nos abre esta fórmula! Pero antes de proseguir, para énfasis y futura referencia, enunciamos este resultado como el

LEMA 2.11. *Un subgrupo N de G es un subgrupo normal de G si y sólo si el producto de dos clases laterales derechas de N en G es de nuevo una clase derecha de N en G .*

Prueba. Si N es normal en G , el resultado acaba de probarse. La prueba de la otra parte es uno de los problemas que están al final de esta sección.

Supongamos que N es un subgrupo normal de G . La fórmula $NaNb = Nab$, para $a, b \in G$ es altamente sugestiva; el producto de clases laterales derechas es una clase lateral derecha. ¿Podemos utilizar este producto para dar al conjunto de las clases laterales derechas una estructura de grupo? ¡Así es, ciertamente! Este tipo de construcción, que se presenta en matemáticas a menudo y al que usualmente se le llama construir una *estructura cociente*, es de máxima importancia.

Denotemos por G/N la colección de las clases laterales derechas de N en G (es decir, los elementos de G/N son ciertos subconjuntos de G) y usemos el producto de subconjuntos de G para que nos suministre un producto en G/N .

Afirmamos respecto a este producto:

1) $X, Y \in G/N$ implica $XY \in G/N$; pues $X = Na$, $Y = Nb$ para algunos $a, b \in G$, y $XY = NaNb = Nab \in G/N$.

2) $X, Y, Z \in G/N$ implica $X = Na$, $Y = Nb$, $Z = Nc$ con $a, b, c \in G$, y, por tanto, $(XY)Z = (NaNb)Nc = N(ab)Nc = N(ab)c = Na(bc)$ (pues G es asociativo) $= Na(Nbc) = Na(NbNc) = X(YZ)$. Por tanto, el producto en G/N satisface la ley asociativa.

3) Considérese el elemento $N = Ne \in G/N$. Si $X \in G/N$, $X = Na$, $a \in G$, entonces $XN = NaNe = Nae = Na = X$, y análogamente también $NX = X$. Por tanto Ne es un elemento identidad para G/N .

4) Supongamos $X = Na \in G/N$ (donde $a \in G$); es claro que $Na^{-1} \in G/N$, y $NaNa^{-1} = Naa^{-1} = Ne$. Análogamente $Na^{-1}Na = Ne$. De donde Na^{-1} es el inverso de Na en G/N .

Pero un sistema que satisface (1), (2), (3), (4) es exactamente lo que llamamos un grupo; es decir

TEOREMA 2.C. *Si G es un grupo y N un subgrupo normal de G , entonces G/N es también un grupo. Se le llama grupo cociente o grupo factor de G por N .*

Si, además, G es un grupo finito, ¿cuál es el orden de G/N ? Como G/N tiene como sus elementos las clases laterales derechas de N en G , y como el número de estas es precisamente $i_G(N) = \frac{o(G)}{o(N)}$, podemos decir

LEMÁ 2.12. Si G es un grupo finito y N es un subgrupo normal de G , entonces $o(G/N) = \frac{o(G)}{o(N)}$.

Cerramos esta sección con un ejemplo.

Sea G el grupo aditivo de los enteros (el grupo que con la adición como operación forman los enteros) y sea N el conjunto de todos los múltiplos de 3. Como la operación en G es la adición, escribiremos las clases laterales de N en G como $N+a$ en lugar de como Na . Consideremos las tres clases laterales $N, N+1, N+2$. Afirmamos que estas son todas las clases laterales de N en G . En efecto, dada $a \in G$, $a = 3b+c$ donde $b \in G$ y $c = 0, 1$ o 2 (a es el resto en la división de a por 3). Por tanto, $N+a = N+3b+c = (N+3b)+c = N+c$ ya que $3b \in N$. Así pues, toda clase lateral es, como afirmábamos, o N , o $N+1$ o $N+2$, y $G/N = \{N, N+1, N+2\}$. ¿Cómo sumamos elementos en G/N ? Nuestra fórmula $NaNb = Nab$ se traduce en: $(N+1)+(N+2) = N+3 = N$ ya que $3 \in N$; $(N+2)+(N+2) = N+4 = N+1$, y así sucesivamente. Sin ser específico, se siente que G/N está íntimamente relacionado con los enteros módulo 3 bajo la adición. Es claro que lo que hicimos con 3 lo podemos hacer con cualquier entero n , en cuyo caso el grupo factor nos sugeriría una relación con los enteros módulo n bajo la adición. Aclararemos este tipo de relación en la sección próxima.

Problemas

*1. Si H es un subgrupo de G tal que el producto de dos clases laterales derechas de H en G es de nuevo una clase lateral derecha de H en G , pruébese que H es normal en G .

Cor 2.1 Marshall 2. Si G es un grupo y H es un subgrupo de índice 2 en G , pruébese que H es un subgrupo normal de G .

3. Si N es un subgrupo normal de G y H es un subgrupo cualquiera de G , pruébese que NH es un subgrupo de G .

4. Pruébese que la intersección de dos subgrupos normales de G es un subgrupo normal de G .

5. Si H es un subgrupo de G y N es un subgrupo normal de G , pruébese que $H \cap N$ es un subgrupo normal de H .

6. Pruébese que todo subgrupo de un grupo abeliano es normal.

Ilustrar con el ^{e)} *7. ¿Es cierto lo recíproco de lo afirmado en el problema 6? En caso grupo de los afirmativo, pruébese; en caso contrario, proporcionese un ejemplo de un cuaternio grupo no abeliano cuyos subgrupos sean todos normales.

8. Sea G un grupo y H un subgrupo de G . Sea, para $g \in G$, $gHg^{-1} = \{ghg^{-1} | h \in H\}$. Pruébese que gHg^{-1} es un subgrupo de G .

9. Supongamos que H es el único subgrupo de orden $o(H)$ en el grupo finito G . Pruébese que H es un subgrupo normal de G .

10. Si H es un subgrupo de G , sea $N(H) = \{g \in G | gHg^{-1} = H\}$. Pruébese:

- 1) $N(H)$ es un subgrupo de G .
- 2) H es normal en $N(H)$.
- 3) Si H es un subgrupo normal del subgrupo K en G , entonces $K \subset N(H)$ (es decir, $N(H)$ es el máximo subgrupo de G en el que H es normal).
- 4) H es normal en G si y sólo si $N(H) = G$.

11. Si N y M son subgrupos normales de G , pruébese que NM es también un subgrupo normal de G .

*12. Supongamos que N y M son dos subgrupos normales de G y que $N \cap M = \{e\}$. Demuéstrese que, entonces, para cualesquiera $n \in N$, $m \in M$, se tiene $nm = mn$.

13. Si un subgrupo cíclico T de G es normal en G , pruébese que todo subgrupo de T es normal en G .

*14. Pruébese, por medio de un ejemplo, que pueden encontrarse tres grupos $E \subset F \subset G$, donde E sea normal en F , F sea normal en G , pero E no sea normal en G .

15. Si N es normal en G y $a \in G$ es de orden $o(a)$, pruébese que el orden m de Na en G/N es un divisor de $o(a)$.

16. Si N es un subgrupo normal en el grupo finito G tal que $i_G(N)$ y $o(N)$ son primos relativos, demuéstrese que cualquier elemento $x \in G$ que satisfaga $x^{o(N)} = e$ debe ser de N .

17. Definamos G como el conjunto de todos los símbolos formales $x^i y^j$, $i = 0, 1, 2, \dots, n-1$, conviniendo en que

$$x^i y^j = x^{i'} y^{j'} \text{ si y sólo si } i = i' \text{ y } j = j'$$

$$x^2 = y^n = e, n > 2$$

$$xy = y^{-1}x.$$

a) Encuéntrese la forma del producto $(x^i y^j)(x^k y^m)$ del tipo $x^a y^b$.

- b) Usando lo anterior, pruébese que G es un grupo no abeliano de orden $2n$.
- c) Si n es impar, pruébese que el centro de G es $\{e\}$, mientras que si n es par el centro de G es mayor que $\{e\}$.

A este grupo se le conoce como grupo *diédrico*. Se obtiene una realización geométrica de él, como sigue: sea y una rotación del plano euclíadiano alrededor del origen con ángulo de giro de $2\pi/n$ radianes, y sea x la reflexión respecto al eje vertical. G es el grupo de movimientos del plano generado por y y x .

7 HOMOMORFISMOS

Las ideas y resultados de esta sección están íntimamente relacionadas con las de la sección precedente. Si hay una idea central común a todos los aspectos del álgebra moderna, tal es la noción de homomorfismo. Indicamos con ello una aplicación de un sistema algebraico a un sistema algebraico análogo que preserva la estructura. Precisamos la idea, en lo que a grupos se refiere, en la definición que sigue.

DEFINICIÓN. Una aplicación ϕ de un grupo G en un grupo \bar{G} se dice que es un *homomorfismo* si para $a, b \in G$ cualesquiera siempre se tiene $\phi(ab) = \phi(a)\phi(b)$.

Nótese que en el primer miembro de esta relación, es decir, en el término $\phi(ab)$, el producto ab se calcula en G usando el producto de elementos de G , mientras que en el segundo miembro de la relación, es decir, en el término $\phi(a)\phi(b)$, el producto es el de elementos en \bar{G} .

Ejemplo 0. $\phi(x) = e$ para todo $x \in G$ Es este trivialmente un homomorfismo. Análogamente $\phi(x) = x$ para todo $x \in G$ es un homomorfismo de G en G .

Ejemplo 1. Sea G el grupo aditivo de los números reales (es decir, ab para $a, b \in G$ es realmente el número real $a+b$) y sea \bar{G} el grupo de los números reales distintos de cero con la multiplicación ordinaria entre números reales como operación de grupo. Definamos $\phi: G \rightarrow \bar{G}$ por $\phi(a) = 2^a$. Para verificar que esta aplicación es un homomorfismo, debemos comprobar si es cierto que $\phi(ab) = \phi(a)\phi(b)$, recordando que, según el producto del primer miembro, entendemos la operación en G (es decir, la adición). Esto es, debemos comprobar, lo que sabemos que es cierto, que $2^{a+b} = 2^a 2^b$. Como 2^a es siempre positivo, la imagen de ϕ no es todo \bar{G} , luego ϕ es un homomorfismo de G en \bar{G} , pero no sobre \bar{G} .

Ejemplo 2. Sea $G = S_3 = \{e, \phi, \psi, \psi^2, \phi\psi, \phi\psi^2\}$ y $\bar{G} = \{e, \phi\}$. Definamos la aplicación $f: G \rightarrow \bar{G}$ por $f(\phi^i\psi^j) = \phi^i$. Tenemos, pues, $f(e) = e$, $f(\phi) = \phi$, $f(\psi) = e$, $f(\psi^2) = e$, $f(\phi\psi) = \phi$, $f(\phi\psi^2) = \phi$. El lector debe verificar que la f así definida es un homomorfismo.

Ejemplo 3. Sea G el grupo aditivo de los enteros y $\bar{G} = G$. Para el entero $x \in G$ definamos ϕ por $\phi(x) = 2x$. Que ϕ es un homomorfismo se sigue de que $\phi(x+y) = 2(x+y) = 2x+2y = \phi(x)+\phi(y)$.

Ejemplo 4. Sea G el grupo de los números reales distintos de cero bajo la multiplicación, $\bar{G} = \{1, -1\}$, donde $1 \cdot 1 = 1$, $(-1)(-1) = 1$, $1(-1) = (-1)1 = -1$. Definamos $\phi: G \rightarrow \bar{G}$ por $\phi(x) = 1$ si x es positivo y $\phi(x) = -1$ si x es negativo. El hecho de que ϕ sea un homomorfismo es equivalente a la afirmación: positivo por positivo es positivo, positivo por negativo es negativo, negativo por negativo es positivo.

Ejemplo 5. Sea G el grupo aditivo de los enteros y \bar{G}_n el grupo aditivo de los enteros módulo n . Definamos ϕ por $\phi(x) =$ residuo de la división de x por n . Es fácil verificar que esto es un homomorfismo.

El resultado del siguiente lema nos proporciona una clase infinita de ejemplos de homomorfismos. Cuando probemos el teorema 2.D resultará que en cierto sentido este lema nos proporciona el ejemplo más general de homomorfismo.

LEMA 2.13. *Supongamos que G es un grupo y que N es un subgrupo normal de G ; definamos la aplicación ϕ de G en G/N por $\phi(x) = Nx$ para todo $x \in G$. Entonces, ϕ es un homomorfismo de G sobre G/N .*

Prueba. En realidad, no hay nada que probar, pues este hecho lo hemos probado ya varias veces. Pero en beneficio de la comprensión repetimos la prueba una vez más.

Que ϕ es suprayectiva es trivial pues todo elemento $X \in G/N$ es de la forma $X = Ny$, $y \in G$, luego $X = \phi(y)$. Para probar el cumplimiento de la propiedad multiplicativa requerida para que ϕ sea un homomorfismo solamente notamos que si $x, y \in G$, $\phi(xy) = NxNy = NxNy = \phi(x)\phi(y)$.

En el lema 2.13 y en los ejemplos que lo preceden, un hecho que se nos muestra es que un homomorfismo no necesita ser inyectivo; pero hay una cierta uniformidad en este proceso de desviación debe ser inyectiva. Pondremos esto en claro dentro de pocas líneas.

DEFINICIÓN. Si ϕ es un homomorfismo de G en \bar{G} , el *núcleo de ϕ* , K_ϕ , se define por $K_\phi = \{x \in G \mid \phi(x) = \bar{e}, \bar{e} = \text{elemento identidad de } \bar{G}\}$.

Antes de investigar las propiedades de K_ϕ es aconsejable establecer que, como conjunto, K_ϕ no es vacío. Esto es lo que dice la primera parte del siguiente lema.

LEMA 2.14. Si ϕ es un homomorfismo de G en \bar{G} , entonces :

- 1) $\phi(e) = \bar{e}$, el elemento unidad de \bar{G} ;
- 2) $\phi(x^{-1}) = \phi(x)^{-1}$ para todo $x \in G$.

Prueba. Para probar (1) simplemente calculamos $\phi(x)\bar{e} = \phi(x) = \phi(xe) = \phi(x)\phi(e)$, de donde, de acuerdo con la propiedad de cancelación en \bar{G} , tenemos que $\phi(e) = \bar{e}$.

Para establecer (2) observemos que $\bar{e} = \phi(e) = \phi(xx^{-1}) = \phi(x)\phi(x^{-1})$, de donde, según la definición de $\phi(x)^{-1}$ en \bar{G} , obtenemos el resultado $\phi(x^{-1}) = \phi(x)^{-1}$.

El argumento usado en la prueba del lema debe recordar a cualquier lector que haya conocido un desarrollo de los logaritmos, el usado en probar los resultados familiares de que $\log 1 = 0$ y $\log(1/x) = -\log x$; no es esto una coincidencia, pues la aplicación $\phi : x \rightarrow \log x$ es un homomorfismo del grupo multiplicativo de los números reales positivos en el grupo aditivo de los números reales.

El lema 2.14 nos muestra que e está en el núcleo de cualquier homomorfismo, de modo que tal núcleo no es vacío. Pero podemos decir más aún.

LEMA 2.15. Si ϕ es un homomorfismo de G en \bar{G} de núcleo K , entonces K es un subgrupo normal de G .

Prueba. Primero debemos comprobar que K es un subgrupo de G . Para ver esto se debe mostrar que K es cerrado respecto a la multiplicación y que todo elemento de K tiene su inverso también en K .

Si $x, y \in K$, entonces $\phi(x) = \bar{e}, \phi(y) = \bar{e}$, donde \bar{e} es el elemento identidad de \bar{G} , y por tanto $\phi(xy) = \phi(x)\phi(y) = \bar{e}\bar{e} = \bar{e}$, de donde $xy \in K$. Además, si $x \in K$, $\phi(x) = \bar{e}$, de donde, según el lema 2.14, $\phi(x^{-1}) = \phi(x)^{-1} = \bar{e}^{-1} = \bar{e}$; luego $x^{-1} \in K$. K es, de acuerdo con esto, un subgrupo de G .

Para probar la normalidad de K debe establecerse que para cualquier $g \in G$ y para cualquier $k \in K$, $gkg^{-1} \in K$; en otras palabras, debe probarse que $\phi(gkg^{-1}) = \bar{e}$ siempre que $\phi(k) = \bar{e}$. Pero $\phi(gkg^{-1}) = \phi(g)\phi(k)\phi(g^{-1}) = \phi(g)\bar{e}\phi(g)^{-1} = \phi(g)\phi(g)^{-1} = \bar{e}$. Y esto completa la prueba del lema 2.15.

Sea ahora ϕ un homomorfismo del grupo G sobre el grupo \bar{G} , y supongamos que K es el núcleo de ϕ . Si $\bar{g} \in \bar{G}$, decimos que un elemento $x \in G$ es una *imagen inversa* de \bar{g} bajo ϕ si $\phi(x) = \bar{g}$. ¿Cuáles son todas las imágenes inversas de \bar{g} ? Para $\bar{g} = \bar{e}$ tenemos la contestación, a saber (por su definición), K . ¿Pero qué hay con elementos $\bar{g} \neq \bar{e}$? Bien, supongamos que $x \in G$ es una imagen inversa de \bar{g} ; ¿podemos escribir otras? Es claro que sí, pues si $k \in K$ y $y = kx$, entonces $\phi(y) = \phi(kx) = \phi(k)\phi(x) = \bar{e}\bar{g} = \bar{g}$. Así pues, todos los elementos de Kx están en la imagen inversa de \bar{g} siempre que x lo esté. ¿Pueden haber otros? Supongamos que $\phi(z) = \bar{g} = \phi(x)$. Ignorando el término intermedio nos quedamos con $\phi(z) = \phi(x)$, y, por tanto, $\phi(z)\phi(x)^{-1} = \bar{e}$. Pero $\phi(x)^{-1} = \phi(x^{-1})$, de donde $\bar{e} = \phi(z)\phi(x)^{-1}$

$= \phi(z)\phi(x^{-1}) = \phi(zx^{-1})$, de donde, en consecuencia, $zx^{-1} \in K$; luego $z \in Kx$. En otras palabras, hemos demostrado que Kx consta exactamente de todas las imágenes inversas de \bar{g} , siempre que x sea una sola de esas imágenes. Recordaremos esto mediante el siguiente

LEMA 2.16. *Si ϕ es un homomorfismo de G sobre \bar{G} de núcleo K , entonces el conjunto de todas las imágenes inversas de $\bar{g} \in \bar{G}$ bajo ϕ en G está dado por Kx donde x es una imagen inversa particular cualquiera de \bar{g} en G .*

Inmediatamente se presenta un caso particular, a saber, la situación cuando $K = \{e\}$. Pero en tal caso el lema 2.16 lo que nos dice es que cualquier $\bar{g} \in \bar{G}$ tiene exactamente una imagen inversa. Es decir, ϕ es una aplicación inyectiva. La recíproca es trivialmente cierta, es decir, si ϕ es un homomorfismo inyectivo de G en (incluso no sobre) \bar{G} , su núcleo debe consistir exactamente en e .

DEFINICIÓN. Un homomorfismo ϕ de G en \bar{G} se dice que es un *isomorfismo* si ϕ es inyectivo.[†]

DEFINICIÓN. Dos grupos G , G^* se dice que son *isomorfos* si hay un isomorfismo de G sobre G^* . En este caso escribimos $G \approx G^*$.

Dejamos al lector verificar los siguientes tres hechos:

- 1) $G \approx G$.
- 2) $G \approx G^*$ implica $G^* \approx G$.
- 3) $G \approx G^*$, $G^* \approx G^{**}$ implica $G \approx G^{**}$.

Cuando dos grupos son isomorfos, entonces, en cierto sentido, son iguales. Difieren en que sus elementos se denominan en forma distinta. El isomorfismo nos da la clase de esta diferencia de denominación, y con ella, conociendo un determinado cálculo en un grupo, podemos efectuar el cálculo análogo en el otro. El isomorfismo es como un diccionario que nos permite traducir una frase de un idioma a una frase, de igual significación, en otro idioma. (Desgraciadamente, no existen diccionarios tan perfectos, porque en los idiomas las palabras no tienen significado único y no aparecen estos cambios de significación en las traducciones literales.) Pero, decir solamente que una oración dada en un lenguaje puede expresarse en otro no nos lleva muy lejos; lo que se necesita es el diccionario para efectuar la traducción. Análogamente, puede ser de poco interés saber que dos grupos son isomorfos, y ser lo realmente interesante el propio isomorfismo. Así, siempre que probemos que dos grupos son isomorfos, intentaremos exhibir una aplicación precisa que produzca este isomorfismo.

[†] En la actualidad parece más frecuente llamar a tal homomorfismo un *monomorfismo*, reservando el término *isomorfismo* para los homomorfismos que son, a la vez, inyectivo y suprayectivo (N. del T.).

Volviendo por un momento al lema 2.16, vemos en él un medio de caracterizar, en términos del núcleo cuando un homomorfismo es realmente un isomorfismo.

COROLARIO. *Un homomorfismo ϕ de G en \bar{G} con núcleo K_ϕ es un isomorfismo de G en \bar{G} si y sólo si $K_\phi = \{e\}$.*

Este corolario nos proporciona una técnica estándar para probar que dos grupos son isomorfos. Primero encontramos un homomorfismo de uno sobre el otro, y luego probamos que el núcleo de este homomorfismo consiste solamente en el elemento identidad. Una ilustración de este método aparece en la prueba del muy importante

TEOREMA 2.D. *Sea ϕ un homomorfismo de G sobre \bar{G} con núcleo K . Entonces $G/K \approx \bar{G}$.*

Prueba. Consideremos el diagrama

$$\begin{array}{ccc} G & \xrightarrow{\phi} & \bar{G} \\ \sigma \downarrow & & \\ G & \xrightarrow{\quad} & K \end{array}$$

donde $\sigma(g) = Kg$.

Nos gustaría completar esto hasta

$$\begin{array}{ccc} G & \xrightarrow{\phi} & \bar{G} \\ \sigma \downarrow \psi & \nearrow & \\ G & \xrightarrow{\quad} & K \end{array}$$

Parece claro que para construir la aplicación ψ de G/K en \bar{G} , debemos usar G como un intermediario, y también que esta construcción debe ser relativamente poco complicada. ¿Qué más natural que completar el diagrama de acuerdo con el que sigue?

$$\begin{array}{ccc} g & \xrightarrow{\quad} & \phi(g) \\ \downarrow \psi & \nearrow & \\ Kg & \xrightarrow{\quad} & \end{array}$$

Con este preámbulo definimos formalmente la aplicación ψ de G/K a \bar{G} por: si $X \in G/K$, $X = Kg$, entonces $\psi(X) = \phi(g)$. Inmediatamente surge un

problema: ¿está esta aplicación bien definida? Si $X \in G/K$, puede escribirse como Kg de varias formas (por ejemplo, $Kg = Kkg$, $k \in K$); pero si $X = Kg = Kg'$, $g, g' \in G$, entonces, por una parte, $\psi(X) = \phi(g)$, y por la otra, $\psi(X) = \phi(g')$. Para que la aplicación ψ tenga sentido debemos tener que $\phi(g) = \phi(g')$. Supongamos pues que $Kg = Kg'$; entonces $g = kg'$ donde $k \in K$, y de aquí $\phi(g) = \phi(kg') = \phi(k)\phi(g') = e\phi(g') = \phi(g')$ ya que $k \in K$, el núcleo de ϕ .

Ahora comprobaremos que ψ es *suprayectiva*. Esto es claro, ya que si $\bar{x} \in \bar{G}$, $\bar{x} = \phi(g)$, $g \in G$ (ya que ϕ es suprayectiva) de modo que $\bar{x} = \phi(g) = \psi(Kg)$.

Si $X, Y \in G/K$, $X = Kg$, $Y = Kf$, $g, f \in G$, entonces $XY = KgKf = Kgf$, de modo que $\psi(XY) = \psi(Kgf) = \phi(gf) = \phi(g)\phi(f)$ ya que ϕ es un homomorfismo de G sobre \bar{G} . Pero $\psi(X) = \psi(Kg) = \phi(g)$, y $\psi(Y) = \psi(Kf) = \phi(f)$, de donde vemos que $\psi(XY) = \psi(X)\psi(Y)$, y ψ es un homomorfismo de G/K sobre \bar{G} .

Para probar que ψ es un isomorfismo de G/K sobre \bar{G} , lo único que nos queda por demostrar es que el núcleo de ψ es el elemento unidad de G/K . Como el elemento unidad de G/K es $K = Ke$, debemos demostrar que si $\psi(Kg) = e$, entonces $Kg = Ke = K$. Esto es ahora fácil, pues $e = \psi(Kg) = \phi(g)$, de modo que $\phi(g) = e$, de donde g está en el núcleo de ϕ , es decir, en K . Pero entonces $Kg = K$ ya que K es un subgrupo de G . Y ya hemos juntado todas las piezas. Hemos mostrado un homomorfismo inyectivo de G/K sobre \bar{G} . Luego $G/K \approx \bar{G}$, y el teorema 2.D queda probado.

El teorema 2.D es importante, porque nos dice, en forma precisa, qué grupos podemos esperar se aparezcan como imágenes homomórficas de un grupo dado. Éstos deben ser expresables en la forma G/K donde K es normal en G . Pero, según el lema 2.13, para cualquier subgrupo normal N de G , G/N es una imagen homomórfica de G . Hay, pues, una correspondencia biyectiva entre las imágenes homomórficas de G y los subgrupos normales de G . Si se tuvieran que encontrar todas las imágenes homomórficas de G podríamos hacerlo, sin dejar nunca G , como sigue: encontraremos todos los subgrupos normales N de G y construyamos todos los grupos G/N . El conjunto de los grupos así construidos nos da todas las imágenes homomórficas de G (salvo isomorfismo).

Un grupo se dice que es *simple* si no tiene imágenes homomórficas distintas de las triviales, es decir, si no tiene subgrupos normales no triviales. Una conjetura famosa y desde hace tiempo presentada, es que un grupo simple no abeliano de orden finito tiene un número par de elementos. Aún aguarda su prueba. †

Hemos afirmado que el concepto de homomorfismo es muy importante. Para reforzar esta afirmación demostraremos ahora cómo es que los métodos

† Este importante resultado acaban probarlo dos jóvenes matemáticos americanos: Walter Feit y John Thompson.

y resultados de esta sección pueden usarse para probar hechos no triviales acerca de los grupos. Cuando construimos el grupo G/N , donde N es normal en G , si resulta que conocemos la estructura de G/N conoceremos la de G "hasta N ". Es verdadero que cierta información acerca de G aparece confusa, pero a menudo queda la bastante para que de ciertos hechos sobre G/N podamos asegurar otros hechos acerca de G . Cuando fotografiamos cierta escena transferimos un objeto tridimensional a una representación bidimensional de él. Sin embargo, mirando la fotografía podemos sacar una gran cantidad de información acerca de la escena fotografiada.

En las dos aplicaciones de las ideas desarrolladas hasta ahora, que a continuación vamos a ver, las pruebas que damos no son las mejores posibles. En realidad, un poco más adelante en este mismo capítulo, se probarán estos resultados en una situación más general y de un modo más fácil. Usamos aquí esta presentación porque ilustra de modo efectivo muchos conceptos de la teoría de grupos.

APLICACIÓN 1 (TEOREMA DE CAUCHY PARA LOS GRUPOS ABELIANOS).
Supongamos que G es un grupo abeliano finito y que $p \mid o(G)$, donde p es un número primo. Entonces hay un elemento $a \neq e \in G$ tal que $a^p = e$.

Prueba. Procedemos por inducción sobre $o(G)$. En otras palabras, suponemos que el teorema es cierto para todos los grupos abelianos que tengan menos elementos que G . Basándonos en ello, queremos probar que el resultado se verifica también para G . Para comenzar con la inducción, notemos que el teorema es cierto por vacuidad para grupos que tienen un solo elemento.

Si G no tiene ningún subgrupo $H \neq (e)$, G , por el resultado de un problema anterior de este capítulo G debe ser cíclico de orden primo. Este primo debe ser p y G , ciertamente, tiene $p-1$ elementos $a \neq e$ que satisfacen $a^p = a^{o(G)} = e$.

Supongamos pues que G tiene un subgrupo $N \neq (e)$, G . Si $p \nmid o(N)$, de acuerdo con nuestra hipótesis de inducción, como $o(N) < o(G)$ y N es abeliano, hay un elemento $b \in N$, $b \neq e$, que satisface $b^p = e$; como $b \in N \subset G$ habríamos con ello exhibido un elemento del tipo requerido. Podemos, por tanto, suponer que $p \nmid o(N)$. Como G es abeliano, N es un subgrupo normal de G , de modo que G/N es un grupo. Además, $o(G/N) = \frac{o(G)}{o(N)}$, y como $p \nmid o(N)$, $p \nmid \frac{o(G)}{o(N)} < o(G)$. Además, como G es abeliano, G/N es abeliano. Luego, por nuestra hipótesis de inducción, hay un elemento $X \in G/N$ que satisface $X^p = e_1$, el elemento unidad de G/N , $X \neq e_1$. Por la forma de los elementos de G/N , $X = Nb$, $b \in G$, de forma que $X^p = (Nb)^p = Nb^p$. Como $e_1 = Ne$, $X^p = e_1$, $X \neq e_1$ se traduce por $Nb^p = N$, $Nb \neq N$. Luego $b^p \in N$, $b \notin N$. Usando uno de los corolarios del teorema de Lagrange, $(b^p)^{o(N)} = e$. Es decir, $b^{o(N)p} = e$. Sea $c = b^{o(N)}$.

Ciertamente $c^p = e$. Para demostrar que c es un elemento que satisface la conclusión del teorema debemos finalmente mostrar que $c \neq e$. Pero si $c = e$, $b^{o(N)} = e$, y, por tanto, $(Nb)^{o(N)} = N$. Combinando esto con $(Nb)^p = N$, y $p \nmid o(N)$, y siendo p un número primo, encontramos que forzosamente habría de tenerse $Nb = N$, luego $b \in N$, una contradicción. Luego $c \neq e$ y $c^p = e$, y hemos completado la inducción. Esto prueba el resultado.

APLICACIÓN 2 (TEOREMA DE SYLOW PARA GRUPOS ABELIANOS). Si G es un grupo abeliano de orden $o(G)$, y si p es un número primo tal que $p^\alpha \mid o(G)$, $p^{\alpha+1} \nmid o(G)$, entonces G tiene un subgrupo de orden p^α .

Prueba. Si $\alpha = 0$, el subgrupo (e) satisface la conclusión del resultado. Supongamos pues que $\alpha \neq 0$. Entonces $p \mid o(G)$. De acuerdo con la aplicación 1, hay un elemento $a \neq e \in G$ que satisface $a^p = e$. Sea $S = \{x \in G \mid x^{p^n} = e\}$ para algún entero n . Como $a \in S$, $a \neq e$, se sigue que $S \neq (e)$. Afirmamos ahora que S es un subgrupo de G . Como G es finito sólo debemos comprobar que S es cerrado. Si $x, y \in S$, $x^{p^n} = e$, $y^{p^n} = e$, de modo que $(xy)^{p^{n+m}} = x^{p^{n+m}}y^{p^{n+m}} = e$ (hemos usado el hecho de ser G abeliano), lo que prueba que $xy \in S$.

Afirmamos ahora que $o(S) = p^\beta$, con β entero y $0 < \beta \leq \alpha$. Pues si un primo $q \mid o(S)$, $q \neq p$, de acuerdo el resultado de la aplicación 1 habría un elemento $c \in S$, $c \neq e$ que satisfaría $c^q = e$. Sin embargo, $c^{p^n} = e$ para algún n , ya que $c \in S$, pero como p^n y q son primos entre sí, podemos encontrar enteros λ y μ tales que $\lambda q + \mu p^n = 1$, de donde $c = c^1 = c^{\lambda q + \mu p^n} = (c^q)^\lambda (c^{p^n})^\mu = e$, lo que contradiría la hipótesis $c \neq e$. Según el teorema de Lagrange $o(S) \mid o(G)$, de modo que $\beta \leq \alpha$. Supongamos $\beta < \alpha$; consideremos el grupo abeliano G/S . Como $\beta < \alpha$ y $o(G/S) = \frac{o(G)}{o(S)}$, $p \mid o(G/S)$, hay un elemento Sx , $(x \in G)$ en G/S que satisface $Sx \neq S$, $(Sx)^{p^n} = S$ para algún entero $n > 0$. Pero $S = (Sx)^{p^n} = Sx^{p^n}$, de donde $x^{p^n} \in S$; de donde $e = (x^{p^n})^{o(S)} = (x^{p^n})^{p^\beta} = x^{p^{n+\beta}}$. Por tanto, x satisface exactamente los requisitos necesarios para ponerlo en S , en otras palabras, $x \in S$, y, por consiguiente, $Sx = S$, lo que contradice $Sx \neq S$. Luego $\beta < \alpha$ es imposible, lo que nos deja como única alternativa que $\beta = \alpha$. De donde S es el subgrupo de orden p^α requerido.

Vamos a confirmar el enunciado. Supongamos que T es otro subgrupo de G de orden p^α , $T \neq S$. Como G es abeliano $ST = TS$, de modo que ST es un subgrupo de G . Según el teorema 2.b

$$o(ST) = \frac{o(S)o(T)}{o(S \cap T)} = \frac{p^\alpha p^\alpha}{o(S \cap T)}$$

y como $S \neq T$, $o(S \cap T) < p^\alpha$, lo que nos da $o(ST) = p^\gamma$, $\gamma > \alpha$. Como ST

es un subgrupo de G , $\alpha(ST)|\alpha(G)$; luego $p^{\alpha}|\alpha(G)$, lo que contradiría el hecho de que α es la máxima potencia de p que divide a $\alpha(G)$. Luego ningún subgrupo T existe, y S es el único subgrupo de orden p^{α} . Con lo que hemos probado el

COROLARIO. Si G es abeliano de orden $\alpha(G)$ y $p^{\alpha}|\alpha(G)$, y $p^{\alpha+1}\nmid\alpha(G)$, hay un subgrupo único de G de orden p^{α} .

Si consideramos $G = S_3$, que es no abeliano, $\alpha(G) = 2 \cdot 3$, vemos que G tiene tres subgrupos distintos de orden 2, a saber, $\{e, \phi\}, \{e, \phi\psi\}, \{e, \phi\psi^2\}$, de modo que vemos que el corolario que afirma la unicidad no siempre vale para los grupos no abelianos. Pero el teorema de Sylow se verifica para todos los grupos finitos.

Dejemos las aplicaciones y volvamos al desarrollo general. Supongamos que ϕ es un homomorfismo de G sobre \bar{G} de núcleo K , y supongamos que \bar{H} es un subgrupo de \bar{G} . Sea $H = \{x \in G \mid \phi(x) \in \bar{H}\}$. Afirmamos que H es un subgrupo de G , y que $H \supset K$. Que $H \supset K$ es trivial, pues si $x \in K$, $\phi(x) = \bar{e}$ está en \bar{H} , de lo que se sigue que $K \subset H$. Supongamos ahora que $x, y \in H$; se tiene de ello que $\phi(x) \in \bar{H}$ y $\phi(y) \in \bar{H}$, y de ello se deduce que $\phi(xy) = \phi(x)\phi(y) \in \bar{H}$. Por tanto, $xy \in H$, es decir, vemos que H es cerrado respecto al producto de G . Además, si $x \in H$, $\phi(x) \in \bar{H}$ y, por tanto, $\phi(x^{-1}) = \phi(x)^{-1} \in \bar{H}$, de lo que se sigue que $x^{-1} \in H$. Hemos, pues, demostrado nuestra aseveración por completo. ¿Qué más es lo que se puede decir cuando \bar{H} es normal en \bar{G} ? Sean $g \in G$, $h \in H$; entonces $\phi(h) \in \bar{H}$, de donde $\phi(ghg^{-1}) = \phi(g)\phi(h)\phi(g)^{-1} \in \bar{H}$, ya que \bar{H} es normal en \bar{G} . Dicho en otra forma, $ghg^{-1} \in H$, de lo que se sigue que H es normal en G . Debe observarse un otro punto, a saber, que el homomorfismo ϕ de G sobre \bar{G} , cuando solo se considera sobre elementos de H , induce un homomorfismo de H sobre \bar{H} , de núcleo exactamente K , ya que $K \subset H$; de acuerdo con el teorema 2.d tenemos que $\bar{H} \approx H/K$.

Supongamos recíprocamente que L es un subgrupo de G y que $K \subset L$. Sea $\bar{L} = \{\bar{x} \in \bar{G} \mid \bar{x} = \phi(l), l \in L\}$. El lector debe verificar que \bar{L} es un subgrupo de \bar{G} . ¿Podemos describir explícitamente el subgrupo $T = \{y \in G \mid \phi(y) \in \bar{L}\}$? Claramente $L \subset T$. ¿Hay algún elemento $t \in T$ que no esté en L ? Supongamos $t \in T$; entonces $\phi(t) \in \bar{L}$, de modo que de acuerdo con la definición de \bar{L} , $\phi(t) = \phi(l)$ para algún $l \in L$. Por tanto, $\phi(tl^{-1}) = \phi(t)\phi(l)^{-1} = \bar{e}$, de donde $tl^{-1} \in K \subset L$, luego t está en $Ll = L$. De modo análogo se prueba que $T \subset L$, lo que combinado con $L \subset T$ nos da $L = T$.

Hemos así establecido una correspondencia biyectiva entre el conjunto de todos los subgrupos de \bar{G} y el conjunto de todos los subgrupos de G que contienen a K . Por otra parte, en esta correspondencia un subgrupo normal de G se corresponde con un subgrupo normal de \bar{G} .

Resumiremos estos últimos párrafos en el siguiente lema.

LEMA 2.17. *Sea ϕ un homomorfismo de G sobre \bar{G} de núcleo K . Para un subgrupo \bar{H} de \bar{G} sea H el subconjunto de G definido por $H = \{x \in G \mid \phi(x) \in \bar{H}\}$. Entonces H es un subgrupo de G y $H \supset K$; si \bar{H} es normal en \bar{G} , entonces H es normal en G . Por otra parte, esta asociación establece una aplicación biyectiva del conjunto de todos los subgrupos de \bar{G} sobre el conjunto de todos los subgrupos de G que contienen K .*

Deseamos probar un teorema más general acerca de la relación entre dos grupos que son homomorfos.

TEOREMA 2.E. *Sea ϕ un homomorfismo de G sobre \bar{G} de núcleo K , y sea \bar{N} un subgrupo normal de \bar{G} y $N = \{x \in G \mid \phi(x) \in \bar{N}\}$. Entonces $G/N \approx \bar{G}/\bar{N}$. O lo que es equivalente, $G/N \approx (G/K)/(N/K)$.*

Prueba. Como ya sabemos, hay un homomorfismo θ de \bar{G} sobre \bar{G}/\bar{N} definido por $\theta(\bar{g}) = \bar{N}\bar{g}$. Definimos la aplicación $\psi : G \rightarrow \bar{G}/\bar{N}$ por $\psi(g) = \bar{N}\phi(g)$ para todo $g \in G$. En primer lugar, vemos que ψ es sobre, pues si $\bar{g} \in \bar{G}$, $\bar{g} = \phi(g)$ para algún $g \in G$, ya que ϕ es sobre, de modo que el elemento tipo $\bar{N}\bar{g}$ de \bar{G}/\bar{N} puede representarse como $\bar{N}\phi(g) = \psi(g)$.

Si $a, b \in G$, $\psi(ab) = \bar{N}\phi(ab)$ según la definición de la aplicación ψ . Pero ahora, como ϕ es un homomorfismo, $\phi(ab) = \phi(a)\phi(b)$. Luego $\psi(ab) = \bar{N}\phi(a)\phi(b) = \bar{N}\phi(a)\bar{N}\phi(b) = \psi(a)\psi(b)$. Hasta el momento, hemos demostrado que ψ es un homomorfismo de G sobre \bar{G}/\bar{N} . ¿Cuál es el núcleo T de ψ ? En primer lugar, si $n \in N$, $\phi(n) \in \bar{N}$, de modo que $\psi(n) = \bar{N}\phi(n) = \bar{N}$, el elemento identidad de \bar{G}/\bar{N} , probando que $N \subset T$. Por otra parte, si $t \in T$, $\psi(t) =$ elemento identidad de $\bar{G}/\bar{N} = \bar{N}$; pero $\psi(t) = \bar{N}\phi(t)$. Comparando estas dos evaluaciones de $\psi(t)$, llegamos a que $\bar{N} = \bar{N}\phi(t)$, lo que obliga a que $\phi(t) \in \bar{N}$; pero esto coloca a t en N de acuerdo con la definición de N . Es decir, $T \subset N$. Ya se probó que el núcleo de ψ es igual a N . Pero entonces ψ es un homomorfismo de G sobre \bar{G}/\bar{N} de núcleo N . Por el teorema 2.d, $G/N \approx \bar{G}/\bar{N}$, que es la primera parte del teorema. La última afirmación del teorema se sigue inmediatamente de la observación (que sigue como consecuencia del teorema 2.d) de que $\bar{G} \approx G/K$, $\bar{N} \approx N/K$, y $\bar{G}/\bar{N} \approx (G/K)/(N/K)$.

Problemas

1. Verifíquese en cada uno de los siguientes casos si las aplicaciones que se definen son homomorfismos y cuando lo sean determine el núcleo.

- G es el grupo de los números reales distintos de cero bajo la multiplicación, $\bar{G} = G$, $\phi(x) = x^2$ para todo $x \in G$.
- G y \bar{G} como en (a), $\phi(x) = 2^x$.
- G es el grupo aditivo de los números reales, $\bar{G} = G$, $\phi(x) = x + 1$ para todo $x \in G$.

- d) G y \bar{G} como en (c), $\phi(x) = 13x$ para todo $x \in G$.
e) G es un grupo abeliano cualquiera, $\bar{G} = G$, $\phi(x) = x^5$ para toda $x \in G$.

2. Sea G un grupo cualquiera y g un elemento fijo de G . Definamos $\phi : G \rightarrow \bar{G}$ por $\phi(x) = gxg^{-1}$. Pruébese que ϕ es un isomorfismo de G sobre \bar{G} .

3. Sea G un grupo abeliano finito de orden $o(G)$ y supongamos que el entero n es primo con $o(G)$. Pruébese que todo $g \in G$ puede escribirse como $g = x^n$ con $x \in G$. (Sugerencia: Considerese la aplicación $\phi : G \rightarrow G$ definida por $\phi(y) = y^n$, y pruébese que esta aplicación es un isomorfismo de G sobre G .)

4. a) Dado un grupo cualquiera G y un subconjunto U de G , sea \bar{U} el subgrupo mínimo de G que contiene a U . Pruébese que hay un tal subgrupo \bar{U} en G . (A \bar{U} se le llama *subgrupo generado por U*).
b) Si $gug^{-1} \in U$ para todo $g \in G$ y $u \in U$, pruébese que \bar{U} es un subgrupo normal de G .

5. Sea $U = \{xyx^{-1}y^{-1} | x, y \in G\}$. En este caso \bar{U} se escribe comúnmente como G' y se llama *subgrupo conmutador* de G .

- a) Pruébese que G' es normal en G .
b) Pruébese que G/G' es abeliano.
c) Si G/N es abeliano, pruébese que $N \supseteq G'$.
d) Pruébese que si H es un subgrupo de G y $H \supseteq G'$, entonces H es normal en G .

6. Si N y M son subgrupos normales de G , pruébese que $NM/M \approx N/N \cap M$.

7. Sea V el conjunto de todos los números reales, y para a, b reales, con $a \neq 0$, definamos $\tau_{ab} : V \rightarrow V$ por $\tau_{ab}(x) = ax + b$. Sea $G = \{\tau_{ab} | a, b$ sean reales y $a \neq 0\}$ y sea $N = \{\tau_{1b} | b \in G\}$. Pruébese que N es un subgrupo normal de G y que $G/N \approx$ grupo multiplicativo de los números reales distintos de cero.

8. Sea G el grupo diédrico definido como el conjunto de todos los símbolos formales $x^i y^j$, $i = 0, 1, j = 0, 1, \dots, n-1$ donde $x^2 = e$, $y^n = e$, $xy = y^{-1}x$. Pruébese que:

- a) El subgrupo $N = \{e, y, y^2, \dots, y^{n-1}\}$ es normal en G .
b) Que $G/N \approx W$, donde $W = \{1, -1\}$ es el grupo con la multiplicación entre números reales como operación.

9. Pruébese que el centro de un grupo es siempre un subgrupo normal.

10. Pruébese que un grupo de orden 9 es abeliano.

Caycedo, 29/17/3

*Caycedo, 6/4
Perry*

11. Si G es un grupo no abeliano de orden 6, pruébese que $G \approx S_3$.
12. Si G es abeliano y si N es un subgrupo cualquiera de G , pruébese que G/N es abeliano.

8. AUTOMORFISMOS

En la sección precedente se definió y examinó el concepto de isomorfismo de un grupo en otro. El caso especial en que el isomorfismo transforme un grupo dado en sí mismo, debe, obviamente, ser de alguna importancia. Usamos la palabra "en" con toda intención, pues existen grupos G que tienen isomorfismos que transforman G en, y no sobre, sí mismo. El más sencillo ejemplo de un caso es el siguiente: Sea G el grupo aditivo de los enteros y definamos $\phi : G \rightarrow G$ por $\phi : x \mapsto 2x$ para todo $x \in G$. Como $\phi : (x+y) = 2(x+y) = 2x+2y$, ϕ es un homomorfismo. Además, si las imágenes de x y y bajo ϕ son iguales, entonces $2x = 2y$, de donde $x = y$. ϕ es, pues, un isomorfismo. Sin embargo ϕ no es sobre, pues la imagen de cualquier entero bajo ϕ es un entero par; así, por ejemplo, 1 no aparece como imagen bajo ϕ de ningún elemento de G . Los automorfismos de un grupo sobre sí mismo serán para nosotros de un interés máximo.

DEFINICIÓN. Por *automorfismo* de un grupo G entenderemos un isomorfismo de G sobre sí mismo.

Como mencionamos en el capítulo 1, siempre que hablemos de aplicaciones de un conjunto en sí mismo escribiríremos las aplicaciones al lado derecho; así si $T : S \rightarrow S$, y $x \in S$, entonces xT es la imagen de x bajo T .

Sea I la aplicación de G que envía cada elemento sobre él mismo; es decir, la dada por $xI = x$ para toda $x \in G$. I es trivialmente un automorfismo de G . Sea $\mathcal{A}(G)$ el conjunto de todos los automorfismos de G ; siendo un subconjunto de $A(G)$, el conjunto de todas las aplicaciones inyectivas de G sobre sí mismo, para los elementos de $\mathcal{A}(G)$ podemos usar el producto definido en $A(G)$, es decir, el producto de aplicaciones. Este producto satisface entonces la propiedad asociativa en $\mathcal{A}(G)$ y, por tanto *a fortiori*, en $\mathcal{A}(G)$. Además, I , el elemento unidad de $A(G)$, está en $\mathcal{A}(G)$, luego $\mathcal{A}(G)$ no es vacío.

Un hecho importante que intentaremos establecer es que $\mathcal{A}(G)$ es un subgrupo de $A(G)$ y, por tanto, en sí mismo, un grupo. Si T_1 y T_2 están en $\mathcal{A}(G)$ sabemos ya que $T_1 T_2 \in A(G)$. Pero lo que necesitamos es que esté en el conjunto, más pequeño, $\mathcal{A}(G)$. Lo probaremos inmediatamente. Para todo $x, y \in G$, $(xy)T_1 = (xT_1)(yT_1)$, y $(xy)T_2 = (xT_2)(yT_2)$ y, por tanto,

$$\begin{aligned} (xy)T_1 T_2 &= ((xy)T_1)T_2 = ((xT_1)(yT_1))T_2 \\ &= ((xT_1)T_2)((yT_1)T_2) = (xT_1 T_2)(yT_1 T_2). \end{aligned}$$

Es decir, $T_1 T_2 \in \mathcal{A}(G)$. Solo es necesario un hecho más verificar para poder asegurar que $\mathcal{A}(G)$ es un subgrupo de $A(G)$, a saber, que si $T \in \mathcal{A}(G)$ entonces $T^{-1} \in \mathcal{A}(G)$. Si $x, y \in G$, entonces $((xT^{-1})(yT^{-1}))T = ((xT^{-1})T)((yT^{-1})T) = (xI)(yI) = xy$, luego $(xT^{-1})(yT^{-1}) = (xy)T^{-1}$, lo que nos dice que T^{-1} está en $\mathcal{A}(G)$. En resumen, hemos probado el

LEMA 2.18. *Si G es un grupo, entonces $\mathcal{A}(G)$, el conjunto de los automorfismos de G , es un grupo.*

Desde luego, por ahora, no tenemos modo alguno de saber que $\mathcal{A}(G)$ tiene elementos distintos del I . Si G es un grupo que tiene solamente dos elementos no resulta difícil convencerse que $\mathcal{A}(G)$ consiste solamente en I . Para grupos G con más de dos elementos, $\mathcal{A}(G)$ siempre tiene más de un elemento.

Lo que nos gustaría es tener una muestra más rica de automorfismos de la que ahora tenemos, la constituida solo por I . Si el grupo G es abeliano y hay algún elemento $x_0 \in G$ que satisface $x_0 \neq x_0^{-1}$, podemos escribir explícitamente un automorfismo, el que constituye la aplicación T definida por $xT = x^{-1}$ para todo $x \in G$. Para cualquier grupo G , T es suprayectiva; para cualquier grupo abeliano G , $(xy)T = (xy)^{-1} = y^{-1}x^{-1} = x^{-1}y^{-1} = (xT)(yT)$. Además $x_0T = x_0^{-1} \neq x_0$, luego $T \neq I$.

Pero la clase de los grupos abelianos es algo limitada y nos gustaría tener algunos automorfismos de los grupos no abelianos. Aunque parezca extraño, la tarea de encontrar automorfismos para tales grupos es más fácil que la de encontrarlos para los grupos abelianos.

Sea G un grupo; para $g \in G$ definamos $T_g : G \rightarrow G$ por $xT_g = g^{-1}xg$ para todo $x \in G$. Afirmamos que T_g es un automorfismo de G . En primer lugar, T_g es suprayectiva, pues dado $y \in G$, si $x = gyg^{-1}$ se tiene $xT_g = g^{-1}(x)g = g^{-1}(gyg^{-1})g = y$. Sean ahora $x, y \in G$, entonces $(xy)T_g = g^{-1}(xy)g = g^{-1}(xgg^{-1}y)g = (g^{-1}xg)(g^{-1}yg) = (xT_g)(yT_g)$. T_g es, por consiguiente, un homomorfismo de G sobre sí mismo. Afirmamos, además, que T_g es inyectiva, pues si $xT_g = yT_g$, entonces $g^{-1}xg = g^{-1}yg$, de donde, aplicando las leyes de cancelación en G , se tiene: $x = y$. T_g se llama *automorfismo interior* correspondiente a g . Si G es no abeliano, hay un par $a, b \in G$ tal que $ab \neq ba$; pero entonces $bT_a = a^{-1}ba \neq b$, luego $T_a \neq I$. Puesto que para un grupo no abeliano G siempre existirán automorfismos no triviales.

Sea $\mathfrak{J}(G) = \{T_g \in \mathcal{A}(G) | g \in G\}$. El cálculo de T_{gh} para $g, h \in G$ puede ser de algún interés. Supongamos $x \in G$; por definición tenemos: $xT_{gh} = (gh)^{-1}x(gh) = h^{-1}g^{-1}xgh = (g^{-1}xg)T_h = (xT_g)T_h = xT_gT_h$. Mirando el comienzo y el final de esta cadena puede verse que $T_{gh} = T_gT_h$. Esta pequeña observación es interesante y sugestiva. Es de interés porque inmediatamente nos dice que $\mathfrak{J}(G)$ es un subgrupo de $\mathcal{A}(G)$ (! verifíquese!). A $\mathfrak{J}(G)$ se le suele llamar *grupo de los automorfismos interiores de G* . Es sugestiva porque si consideramos la aplicación $\psi : G \rightarrow \mathcal{A}(G)$ definida por

$\psi(g) = T_g$ para toda $g \in G$, entonces $\psi(gh) = T_{gh} = T_g T_h = \psi(g)\psi(h)$. Es decir, ψ es un homomorfismo de G en $\mathcal{A}(G)$ cuya imagen es $\mathcal{J}(G)$. ¿Cuál es el núcleo de ψ ? Llamémosle K y supongamos que $g_0 \in K$. Entonces $\psi(g_0) = I$, lo que es equivalente a decir que $T_{g_0} = I$. Pero esto nos dice que para cualquier $x \in G$, $xT_{g_0} = x$; ahora bien, $xT_{g_0} = g_0^{-1}xg_0$, luego tenemos: $x = g_0^{-1}xg_0$ para todo $x \in G$. Luego $g_0x = g_0g_0^{-1}xg_0 = xg_0$; es decir, g_0 debe conmutar con todos los elementos de G . Pero el centro de G , Z , se definió precisamente como el conjunto de todos los elementos de G que conmutan con todos los elementos de G . (Véase el problema 13, sección 5.) Luego $K \subset Z$. Ahora bien, si $z \in Z$, entonces $xT_z = z^{-1}xz = z^{-1}(zx)$ (puesto que $zx = xz$) = x , de donde $T_z = I$ y, por tanto, $z \in K$. De manera que, $Z \subset K$. Como hemos probado que $K \subset Z$ y, también, que $Z \subset K$, tenemos que $Z = K$. Resumiendo, ψ es un homomorfismo de G en $\mathcal{A}(G)$ con imagen $\mathcal{J}(G)$ y núcleo Z . De acuerdo con el teorema 2.d, $\mathcal{J}(G) \approx G/Z$. Para subrayar la importancia de este resultado general lo enunciamos como lema.

LEMA 2.19. $\mathcal{J}(G) \approx G/Z$, donde $\mathcal{J}(G)$ es el grupo de los automorfismos interiores de G , y Z es el centro de G .

Supongamos que ϕ es un automorfismo de un grupo G y que $a \in G$ tiene orden n (es decir, $a^n = e$, pero para todo m tal que $0 < m < n$, $a^m \neq e$). Entonces $\phi(a)^n = \phi(a^n) = \phi(e) = e$, es decir, $\phi(a)^n = e$. Si $\phi(a)^m = e$ para algún m tal que $0 < m < n$, entonces $\phi(a^m) = \phi(a)^m = e$, lo que implica, como ϕ es inyectiva, que $a^m = e$, en contradicción con lo supuesto. Luego

LEMA 2.20. Sea G un grupo y ϕ un automorfismo de G . Si $a \in G$ es de orden $o(a) > 0$, entonces $o(\phi(a)) = o(a)$.

Los automorfismos de grupo pueden usarse como procedimiento para la construcción de nuevos grupos partiendo del grupo original. Antes de dar una explicación abstracta de esto consideraremos un caso particular.

Sea G un grupo cíclico de orden 7, es decir, G consiste en todas las potencias de a , a^i , donde suponemos que $a^7 = e$. La aplicación $\phi : a^i \rightarrow a^{2i}$, como puede comprobarse trivialmente, es un automorfismo de G de orden 3, es decir, $\phi^3 = I$. Sea x un símbolo al que formalmente sometemos a las siguientes condiciones: $x^3 = e$, $x^{-1}a^i x = \phi(a^i) = a^{3i}$, y consideremos todos los símbolos formales $x^i a^j$ con $i = 0, 1, 2, j = 0, 1, 2, \dots, 6$ y donde convenimos que $x^i a^j = x^k a^l$ si y sólo si $i \equiv k \pmod{3}$, y $j \equiv l \pmod{7}$. Multiplicamos estos símbolos según las reglas: $x^3 = a^7 = e$, $x^{-1}ax = a^3$. Por ejemplo, $xaxa^2 = x(ax)a^2 = x(xa^3)a^2 = x^2a^5$. El lector puede verificar que así se obtiene un grupo no abeliano de orden 21.

Generalmente, si G es un grupo y T un automorfismo de orden r de G que no es un automorfismo interior, escojamos un símbolo x y consideremos todos los elementos $x^i g$, $i = 0, \pm 1, \pm 2, \dots$, $g \in G$ sujetos a las condiciones $x^i g = x^{i'} g'$ si y sólo si $i \equiv i' \pmod{r}$, $g = g'$ y $x^{-1}g'x = gT^i$ para todo i .

De esta forma obtenemos un grupo mayor $\{G, T\}$; G es normal en $\{G, T\}$ y $\{G, T\}/G \approx$ grupo generado por T = grupo cíclico de orden r .

Cerramos esta sección con la determinación de $\mathcal{A}(G)$ para todos los grupos cíclicos.

Ejemplo 1. Sea G un grupo cíclico finito de orden r , $G = \langle a \rangle$, $a^r = e$. Supongamos que T es un automorfismo de G . Si conocemos aT , como $a^t T = (aT)^t$, $a^t T$ está determinado, luego gT está determinado para toda $g \in G = \langle a \rangle$. Necesitamos, pues, solo considerar imágenes posibles de a bajo T . Como $aT \in G$, y como todo elemento de G es una potencia de a , $aT = a^t$ para algún entero t tal que $0 < t < r$. Ahora bien, como T es un automorfismo, aT debe tener el mismo orden que a (lema 2.20), y afirmamos que esta condición fuerza a t a ser primo con r . Pues si $d|t$, $d|r$, entonces $(aT)^{r/d} = a^{t(r/d)} = a^{t(t/d)} = (a^t)^{t/d} = e$; luego aT tendría como orden un divisor de r/d , lo que combinado con el hecho de que aT tiene orden r , implica $d = 1$. Recíprocamente, para cualquier entero s tal que $0 < s < r$ que sea primo relativo con r , la aplicación $S : a^i \rightarrow a^{si}$ es un automorfismo de G . Por tanto, $\mathcal{A}(G)$ está en una correspondencia biyectiva con el grupo U_r de enteros menores que r y primos relativos con r con la multiplicación módulo r como operación. Afirmamos no solo que existe una tal correspondencia biyectiva, sino que existe una que, además, es un isomorfismo. Para cada entero i tal que $0 < i < r$ y r e i sean primos relativos, sea T_i el elemento de $\mathcal{A}(G)$ dado por $T_i : a \rightarrow a^i$; entonces $T_i T_j : a \rightarrow a^l \rightarrow a^{lj}$, luego $T_i T_j = T_{ij}$. La aplicación $i \rightarrow T_i$ exhibe el isomorfismo de U_r sobre $\mathcal{A}(G)$. De donde $\mathcal{A}(G) \approx U_r$.

Ejemplo 2. G es un grupo cíclico infinito. Es decir, G consiste de todos los a^i , $i = 0, \pm 1, \pm 2, \dots$ donde se supone que $a^i = e$ si y sólo si $i = 0$. Supongamos que T es un automorfismo de G . Como en el ejemplo 1, $aT = a^t$. El problema que ahora se nos presenta es, ¿qué valores de t son posibles? Como T es un automorfismo de G , transforma G sobre sí mismo, de modo que $a = gT$ para alguna $g \in G$. Luego $a = a^t T = (aT)^t$ para algún entero t . Como $aT = a^t$, debemos tener que $a = a^{ti}$, luego que $a^{ti-1} = e$. De donde $ti-1 = 0$, es decir, $ti = 1$. Claramente, como t e i son enteros esto implica que forzosamente $t = \pm 1$, y cada uno de estos dos valores da lugar a un automorfismo, $t = 1$ nos da el automorfismo identidad I , $t = -1$ genera el automorfismo $T : g \rightarrow g^{-1}$ para toda g en el grupo cíclico G . Luego, aquí $\mathcal{A}(G) \approx$ grupo cíclico de orden 2.

Problemas

1. ¿Las siguientes aplicaciones son automorfismos de sus grupos respectivos?

- a) Grupo de los enteros bajo la adición y $T : x \rightarrow -x$.
- b) Grupo multiplicativo de los reales positivos y $T : x \rightarrow x^2$.

- c) Grupo cíclico de orden 12 y $T: x \rightarrow x^3$.
d) Grupo S_3 y $T: x \rightarrow x^{-1}$.

*J.F. Cayaolo
Roma 60'*

2. Sea G un grupo, H un subgrupo de G , T un automorfismo de G . Sea $(H)T = \{hT \mid h \in H\}$. Pruébese que $(H)T$ es un subgrupo de G .

3. Sea G un grupo, T un automorfismo de G , N un subgrupo normal de G . Pruébese que $(N)T$ es un subgrupo normal de G .

4. Para $G = S_3$ pruébese que $G \approx J(G)$.

5. Para cualquier grupo G pruébese que $J(G)$ es un subgrupo normal de $A(G)$ (el grupo $A(G)/J(G)$ se llama *grupo de los automorfismos exteriores de G*).

6. Sea G un grupo de orden 4, $G = \{e, a, b, ab\}$, $a^2 = b^2 = e$, $ab = ba$. Determínese $A(G)$.

7. a) Un subgrupo C de G se dice que es un *subgrupo característico* de G si $(C)T \subset C$ para todo automorfismo T de G . Pruébese que un subgrupo característico de G debe ser un subgrupo normal de G .

b) Pruébese que el recíproco de (a) es falso.

8. Para cualquier grupo G , pruébese que el subgrupo comutador G' es un subgrupo característico de G . (Véase el problema 5 de la sección 7.)

9. Si G es un grupo, N un subgrupo normal de G y M un subgrupo característico de N , pruébese que M es un subgrupo normal de G .

10. Sea G un grupo finito, T un automorfismo de G con la propiedad de que $xT = x$ para $x \in G$ si y sólo si $x = e$. Pruébese que todo $g \in G$ puede representarse como $g = x^{-1}(xT)$ para algún $x \in G$.

11. Sea G un grupo finito y T un automorfismo de G con la propiedad de que $xT = x$ si y sólo si $x = e$. Supongamos, por otra parte, que $T^2 = I$. Pruébese que G debe ser abeliano.

*12. Sea G un grupo finito y supongamos que el automorfismo T envía más de las tres cuartas partes de los elementos de G sobre sus inversos. Pruébese que $xT = x^{-1}$ para todo $x \in G$ y que G es abeliano.

13. En el problema 12, ¿puede el lector encontrar algún ejemplo de un grupo finito que sea no abeliano y que tenga un automorfismo que transforme exactamente tres cuartas partes de los elementos de G sobre sus inversos?

*14. Pruébese que todo grupo finito con más de dos elementos tiene un automorfismo no trivial.

*15. Sea G un grupo de orden $2n$. Supongamos que la mitad de los elementos de G son de orden 2 y que la otra mitad forman un subgrupo H de orden n . Pruébese que H es de orden impar y es un subgrupo abeliano de G .

*16. Sea $\phi(n)$ la función ϕ de Euler. Si $a > 1$ es un entero, pruébese que $n \mid \phi(a^n - 1)$.

9. EL TEOREMA DE CAYLEY

Cuando los grupos aparecieron al principio en las matemáticas, generalmente procedían de una fuente específica y se presentaban en forma muy concreta. Muy a menudo era en la forma de un conjunto de transformaciones de algún objeto matemático particular. En realidad, la mayor parte de los grupos aparecieron como grupos de permutaciones, es decir, como subgrupos de S_n . ($S_n = A(S)$ cuando S es un conjunto finito de n elementos.) El matemático inglés Cayley fue el primero en notar que todo grupo podía realizarse como un subgrupo de $A(S)$ para algún S . Nuestra tarea en esta sección, será la de dar una presentación del teorema de Cayley y algunos otros resultados con él relacionados.

TEOREMA 2.F (CAYLEY). *Todo grupo es isomorfo a un subgrupo de $A(S)$ para algún S apropiado.*

Prueba. Sea G un grupo. El conjunto S será el que esté constituido por los elementos de G ; es decir, ponemos $S = G$. Si $g \in G$ definamos $\tau_g : S (= G) \rightarrow S (= G)$ por $x\tau_g = xg$ para todo $x \in G$. Si $y \in G$, entonces $y = (yg^{-1})g = (yg^{-1})\tau_g$, de modo que τ_g transforma S sobre sí mismo. Por otra parte, τ_g es inyectiva, pues si $x, y \in S$ y $x\tau_g = y\tau_g$, entonces $xg = yg$, lo que, según la ley de cancelación en los grupos, implica que $x = y$. Hemos probado que para toda $g \in G$, $\tau_g \in A(S)$.

Si $g, h \in G$, consideremos τ_{gh} . Para cualquier $x \in S = G$, $x\tau_{gh} = x(gh) = (xg)h = (x\tau_g)\tau_h = x\tau_g\tau_h$. Nótese que hemos usado la ley asociativa aquí en pasos esenciales. De $x\tau_{gh} = x\tau_g\tau_h$ deducimos que $\tau_{gh} = \tau_g\tau_h$. Por tanto, si definimos $\psi : G \rightarrow A(S)$ por $\psi(g) = \tau_g$, la relación $\tau_{gh} = \tau_g\tau_h$ nos dice que ψ es un homomorfismo. ¿Cuál es el núcleo K de ψ ? Si $g_0 \in K$, entonces $\psi(g_0) = \tau_{g_0}$, es la aplicación idéntica sobre S , de modo que para $x \in G$ y, en particular, para $e \in G$, $e\tau_{g_0} = e$. Pero $e\tau_{g_0} = eg_0 = g_0$. Luego, comparando estas dos expresiones para $e\tau_{g_0}$, concluimos que $g_0 = e$, de donde $K = \{e\}$. Luego, de acuerdo con el corolario al lema 2.16 ψ es un isomorfismo de G en $A(S)$, lo que prueba el teorema.

El teorema nos permite exhibir cualquier grupo abstracto como un objeto más concreto, a saber, como un grupo de transformaciones. Pero

tiene sus limitaciones; pues si G es un grupo finito de orden $o(G)$, entonces, usando $S = G$, como en nuestra prueba, $A(S)$ tiene $o(G)!$ elementos. Nuestro grupo G de orden $o(G)$ es algo perdido en el grupo $A(S)$ que, con sus $o(G)!$ elementos es gigantesco en comparación con G . Y nos preguntamos, ¿no podría encontrarse una S más económica, para la que $A(S)$ fuera menor? Eso es lo que vamos a intentar conseguir.

Sea G un grupo y H un subgrupo de G . Sea S el conjunto cuyos elementos son las clases laterales derechas de H en G . Es decir, $S = \{Hg | g \in G\}$. S no necesariamente ha de ser un grupo, en realidad, sería un grupo solamente si H fuera un subgrupo normal de G . Sin embargo, podemos hacer que nuestro grupo G actúe sobre S del siguiente modo natural: para $g \in G$ sea $t_g : S \rightarrow S$ el dado por $(Hx)t_g = Hxg$. Siguiendo los pasos de la prueba del teorema 2.f podemos fácilmente probar:

$$1) t_g \in A(S) \text{ para todo } g \in G.$$

$$2) t_{gh} = t_g t_h.$$

Luego la aplicación $\theta : G \rightarrow A(S)$ definida por $\theta(g) = t_g$ es un homomorfismo de G en $A(S)$. ¿Puede siempre afirmarse que θ sea un isomorfismo? Supongamos que K es el núcleo de θ . Si $g_0 \in K$, entonces $\theta(g_0) = t_{g_0}$ es la aplicación idéntica sobre S , de modo que para toda $X \in S$, $Xt_{g_0} = X$. Como todo elemento de S es una clase lateral derecha de H en G , debemos tener que $Hat_{g_0} = Ha$ para toda $a \in G$, y usando la definición de t_{g_0} , a saber, $Hat_{g_0} = Hag_0$, llegamos a la identidad $Hag_0 = Ha$ para toda $a \in G$. Por otra parte, si $b \in G$ es tal que $Hxb = Hx$ para toda $x \in G$, dando vuelta a nuestro argumento podríamos mostrar que $b \in K$. Luego $K = \{b \in G | Hxb = Hx \text{ para todo } x \in G\}$. Afirmamos que, por esta caracterización de K , K debe ser el máximo subgrupo normal de G que está contenido en H . Explicamos primero el uso de la palabra máximo; queremos decir con ella que si N es un subgrupo normal del G que está contenido en H , entonces N debe estar contenido en K . Queremos demostrar que este es el caso. Que K es un subgrupo normal de G se sigue del hecho de que es el núcleo de un homomorfismo de G . Pasemos ahora a probar que $K \subset H$. En efecto, si $b \in K$, $Hab = Ha$ para todo $a \in G$, luego en particular $Hb = Heb = He = H$, de donde $b \in H$. Finalmente, si N es un subgrupo normal de G que está contenido en H , si $n \in N$ y $a \in G$, entonces $ana^{-1} \in N \subset H$, luego $Hana^{-1} = H$; por tanto $Han = Ha$ para toda $a \in G$. De manera que, $n \in K$ por nuestra caracterización de K .

Y hemos probado el

TEOREMA 2.G. *Si G es un grupo, H un subgrupo de G y S el conjunto de todas las clases laterales derechas de H en G , entonces hay un homomorfismo θ de G en $A(S)$ y el núcleo de θ es el máximo subgrupo normal de G que está contenido en H .*

El caso $H = (e)$ nos da exactamente el teorema de Cayley (teorema 2.f). Si resulta que H no tiene ningún subgrupo normal de G distinto del (e) en él, entonces θ debe ser un isomorfismo de G en $A(S)$. En este caso, habríamos disminuido el tamaño del S usado en la prueba del teorema 2.f. Esto es interesante principalmente para grupos finitos, pues usaremos esta observación como medio para probar que ciertos grupos finitos tienen subgrupos normales no triviales, al igual que para representar ciertos grupos finitos como grupos de permutaciones sobre pequeños conjuntos.

Examinemos estas observaciones con mayor cuidado. Supongamos que G tiene un subgrupo H cuyo índice $i(H)$ (es decir, el número de clases laterales derechas de H en G) satisface $i(H)! < o(G)$. Sea S el conjunto de todas las clases laterales derechas de H en G . La aplicación, θ , del teorema 2.g no puede ser un isomorfismo, pues si lo fuera, $\theta(G)$ tendría $o(G)$ elementos y, sin embargo, sería un subgrupo de $A(S)$ que tiene $i(H)! < o(G)$ elementos. Por tanto, el núcleo de θ debe ser mayor que (e) ; como este núcleo es el máximo subgrupo normal de G que está contenido en H , podemos concluir que H contiene un subgrupo normal no trivial de G .

Pero el argumento arriba usado tiene implicaciones incluso cuando $i(H)!$ no es menor que $o(G)$. Si $o(G)$ no divide a $i(H)!$ entonces, por el teorema de Lagrange, sabemos que $A(S)$ no puede tener ningún subgrupo de orden $o(G)$, de donde ningún subgrupo isomorfo a G . Pero $A(S)$ contiene a $\theta(G)$, de donde $\theta(G)$ no puede ser isomorfo a G , es decir, θ no puede ser un isomorfismo. Pero, entonces, como antes, H debe contener un subgrupo normal no trivial de G .

Resumimos todo esto en el siguiente

LEMA 2.21. *Si G es un grupo finito, y $H \neq G$ es un subgrupo de G tal que $o(G) \nmid i(H)!$ entonces H debe contener a un subgrupo normal no trivial de G . En particular, G no puede ser simple.*

Aplicaciones

1) Sea G un grupo de orden 36. Supongamos que G tiene un subgrupo H de orden 9 (más adelante veremos que este es siempre el caso). Entonces $i(H) = 4, 4! = 24 < 36 = o(G)$ de modo que en H debe haber un subgrupo normal $N \neq (e)$, de G , de orden un divisor de 9, es decir, de orden 3 o 9.

2) Sea G un grupo de orden 99 y supongamos que H es un subgrupo de G de orden 11 (veremos más adelante que tal cosa debe suceder forzosamente). Entonces $i(H) = 9$, y como $99 \nmid 9!$ hay un subgrupo normal no trivial $N \neq (e)$ de G en H . Como H es de orden 11, que es un primo, su solo subgrupo distinto del (e) es él mismo, lo que implica que $N = H$. Es decir, H mismo es un subgrupo normal de G .

3) Sea G un grupo no abeliano de orden 6. De acuerdo con el problema 11, sección 3, hay un $a \neq e \in G$ que satisface $a^2 = e$. Luego, el subgrupo $H = \{e, a\}$ es de orden 2, y $i(H) = 3$. Supongamos, por el momento, que

sabemos que H no es normal en G . Como H tiene como únicos subgrupos él mismo y $\{e\}$, H no tiene ningún subgrupo normal no trivial de G en él. Luego G es isomorfo a un subgrupo T de orden 6 en $A(S)$, donde S es el conjunto de las clases laterales derechas de H en G . Como $o(A(S)) = i(H)! = 3! = 6$, $T = S$. En otras palabras, $G \approx A(S) = S_3$. Habríamos probado que todo grupo no abeliano de orden 6 es isomorfo a S_3 . Todo lo que queda por demostrar es que H no es normal en G . Como puede ser de algún interés, pasamos a una prueba detallada de esto. Si $H = \{e, a\}$ fuese normal en G , entonces para cada $g \in G$, como $gag^{-1} \in H$ y $gag^{-1} \neq e$, tendríamos que $gag^{-1} = a$, o lo que es lo mismo, que $ga = ag$ para todo $g \in G$. Sea $b \in G$, $b \notin H$, y consideremos $N(b) = \{x \in G \mid xb = bx\}$. Sabemos, por un problema anterior, que $N(b)$ es un subgrupo de G , y $N(b) \supset H$; $N(b) \neq H$ ya que $b \in N(b)$ y $b \notin H$. Como H es un subgrupo de $N(b)$, $o(H) | o(N(b)) | 6$. El único número par n , $2 < n \leq 6$ que divide a 6 es 6. Luego $o(N(b)) = 6$; de donde b comutaría con todos los elementos de G . Así que todo elemento de G comuta con todo elemento de G , haciendo de G un grupo abeliano, en contra de lo supuesto. Luego H no puede ser normal en G . Esta prueba es algo larga, pero ilustra algunas de las ideas que hemos presentado.

Problemas

1. Sea G un grupo; consideremos las aplicaciones de G en sí mismo, λ_g , definidas para $g \in G$ por $x\lambda_g = gx$ para toda $x \in G$. Pruébese que toda λ_g es inyectiva y suprayectiva y que $\lambda_{gh} = \lambda_g \lambda_h$.
2. Definamos λ_g como en el problema 1 y τ_g como en la prueba del teorema 2.f. Pruébese que para cualesquiera $g, h \in G$, las aplicaciones λ_g y τ_h satisfacen $\lambda_g \tau_h = \tau_h \lambda_g$. (Sugerencia: para $x \in G$ considérense $x(\lambda_g \tau_h)$ y $x(\tau_h \lambda_g)$.)
3. Si θ es una aplicación inyectiva de G sobre sí mismo tal que $\lambda_g \theta = \theta \lambda_g$ para todo $g \in G$, pruébese que $\theta = \tau_h$ para algún $h \in G$.
4. a) Si H es un subgrupo de G , pruébese que para todo $g \in G$, gHg^{-1} es un subgrupo de G .
 b) Pruébese que $W = \text{intersección de todos los } gHg^{-1}$ es un subgrupo normal de G .
5. Usando el lema 2.21 pruébese que un grupo de orden p^2 , donde p es un número primo, debe tener un subgrupo normal de orden p .
- *6. Demuéstrese que en un grupo G de orden p^2 , cualquier subgrupo normal de orden p debe encontrarse en el centro de G . (Sugerencia: si m es un entero, $m^p \equiv m \pmod{p}$.)

*7. Usando el resultado del problema 6, pruébese que cualquier grupo de orden p^2 es abeliano.

8. Si p es un número primo, pruébese que cualquier grupo G de orden $2p$ debe tener un subgrupo de orden p , y que este subgrupo es normal en G .

9. Si $\sigma(G)$ es pq con p y q números primos distintos y si G tiene un subgrupo normal de orden p y un subgrupo normal de orden q , pruébese que G es cíclico.

*10. Sea $\sigma(G) = pq$, $p > q$ primos. Pruébese:

- G tiene un subgrupo de orden p y un subgrupo de orden q .
- Si $q \nmid p-1$, entonces G es cíclico.
- Dados dos primos p, q , $q \nmid p-1$, existe un grupo no abeliano de orden pq .
- Cualesquiera dos grupos no abelianos de orden pq son isomorfos.

10. GRUPOS DE PERMUTACIONES

Hemos visto que todo grupo puede representarse isomórficamente como un subgrupo de $A(S)$ para algún conjunto S y, en particular, que un grupo finito G puede representarse como un subgrupo de S_n , para algún n , donde S_n es el grupo simétrico de grado n . Esto demuestra claramente que los grupos S_n mismos merecen un estrecho examen.

Supongamos que S es un conjunto finito de n elementos x_1, x_2, \dots, x_n . Si $\phi \in A(S) = S_n$, entonces ϕ es una aplicación inyectiva de S sobre sí mismo, y podemos describir ϕ mostrando lo que hace a cada elemento, por ejemplo, $\phi : x_1 \rightarrow x_2, x_2 \rightarrow x_4, x_4 \rightarrow x_3, x_3 \rightarrow x_1$. Pero esto es muy laborioso. Abreviamos el procedimiento si escribimos ϕ como

$$\begin{pmatrix} x_1 & x_2 & x_3 & \dots & x_n \\ x_{i_1} & x_{i_2} & x_{i_3} & \dots & x_{i_n} \end{pmatrix}$$

donde x_{i_k} es la imagen de x_i bajo ϕ . Volviendo al ejemplo que acabamos de poner, ϕ estaría representado por

$$\begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_2 & x_4 & x_1 & x_3 \end{pmatrix}.$$

Mientras que esta notación es ya más manejable hay también en ella cierto derroche, pues no parece que sirva a propósito alguno el símbolo x . Podríamos exactamente lo mismo representar la permutación por

$$\begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix}.$$

Para nuestro ejemplo tendríamos

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix}.$$

Dadas dos permutaciones θ, ψ en S_n , usando esta representación simbólica de θ y ψ , ¿cuál sería la representación de $\theta\psi$? Para calcularla podemos comenzar viendo qué es lo que $\theta\psi$ hace a x_1 (que aquí aparecerá escrito como 1). θ lleva 1 a i_1 , mientras que ψ lleva i_1 a k , digamos, luego $\theta\psi$ lleva 1 a k . Repetimos a continuación este proceso para 2, 3, ..., n . Por ejemplo, si θ es la permutación representada por

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

y ψ es la representada por

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix},$$

entonces $i_1 = 3$ y ψ lleva 3 en 2, luego $k = 2$ y $\theta\psi$ lleva 1 a 2. Análogamente $\theta\psi : 2 \rightarrow 1, 3 \rightarrow 3, 4 \rightarrow 4$. Es decir, la representación de $\theta\psi$ es

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix}.$$

Si escribimos

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

y

$$\psi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix},$$

entonces

$$\theta\psi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix}.$$

Este es el modo en que multiplicaremos los símbolos de la forma

$$\begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & \dots & n \\ k_1 & k_2 & \dots & k_n \end{pmatrix}.$$

Sea S un conjunto y $\theta \in A(S)$. Dados dos elementos $a, b \in S$ definimos $a \equiv_{\theta} b$ si y sólo si $b = a\theta^i$ para algún entero i (i puede ser positivo, negativo o cero). Afirmamos que esto define una relación de equivalencia sobre S . En efecto:

- 1) $a \equiv_{\theta} a$ ya que $a = a\theta^0 = ae$.
- 2) Si $a \equiv_{\theta} b$, entonces $b = a\theta^i$, luego $a = b\theta^{-i}$, luego $b \equiv_{\theta} a$.
- 3) Si $a \equiv_{\theta} b$ y $b \equiv_{\theta} c$, entonces $b = a\theta^i$ y $c = b\theta^j = (a\theta^i)\theta^j = a\theta^{i+j}$, lo que implica que $a \equiv_{\theta} c$.

Esta relación de equivalencia induce, de acuerdo con el teorema I.a, una descomposición de S en subconjuntos ajenos, a saber, las llamadas clases de equivalencia. A la clase de equivalencia de un elemento $s \in S$ le llamamos órbita de s bajo θ ; es decir, la órbita de s bajo θ consiste en todos los elementos $s\theta^i$, $i = 0, \pm 1, \pm 2, \dots$.

En particular, si S es un conjunto finito y $s \in S$, hay un entero positivo mínimo $l = l(s)$ dependiente de s tal que $s\theta^l = s$. La órbita de s bajo θ consta, entonces, de los elementos $s, s\theta, s\theta^2, \dots, s\theta^{l-1}$. Por ciclo de θ entendemos el conjunto ordenado $(s, s\theta, s\theta^2, \dots, s\theta^{l-1})$. Si conocemos todos los ciclos de θ es claro que conocemos θ , ya que conoceríamos la imagen de cualquier elemento bajo θ . Antes de proseguir, ilustramos estas ideas con un ejemplo. Sea

$$\theta = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 3 & 5 & 6 & 4 \end{pmatrix}$$

donde S consiste en los elementos $1, 2, \dots, 6$ (recuérdese que 1 aparece en lugar de x_1 , 2 en lugar de x_2 , etc.). Comenzando con 1, la órbita de 1 se compone de $1 = 1\theta^0, 1\theta^1 = 2, 1\theta^2 = 2\theta = 1$, es decir, la órbita de 1 es el conjunto de los elementos 1 y 2. Esto nos dice que la órbita de 2 es el mismo conjunto. La órbita de 3 consta solamente del 3; la del 4 tiene como elementos 4, $4\theta = 5, 4\theta^2 = 5\theta = 6, 4\theta^3 = 6\theta = 4$. Los ciclos de θ son pues $(1, 2), (3), (4, 5, 6)$.

Ahora procederemos con una pequeña digresión, dejando por un momento nuestra θ particular. Supongamos que por el ciclo (i_1, i_2, \dots, i_r) entendemos la permutación ψ que envía i_1 en i_2, i_2 en i_3, \dots, i_{r+1} en i_r e i_r en i_1 , y deja sin modificación todos los otros elementos de S . Así, por ejemplo, si S consiste en los elementos $1, 2, \dots, 9$, entonces el símbolo $(1, 3, 4, 2, 6)$ representa a la permutación

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

Multiplicamos los ciclos multiplicando las permutaciones que representan.

Así, si de nuevo S tiene 9 elementos,

$$(1 \ 2 \ 3) (5 \ 6 \ 4 \ 1 \ 8)$$

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

Volvamos a las ideas del párrafo penúltimo, y preguntémonos, dada la permutación

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

¿cuáles son los ciclos de θ ? Encontremos primero la órbita de 1; está compuesta de 1, $1\theta = 2$, $1\theta^2 = 2\theta = 3$, $1\theta^3 = 3\theta = 8$, $1\theta^4 = 8\theta = 5$, $1\theta^5 = 5\theta = 6$, $1\theta^6 = 6\theta = 4$, $1\theta^7 = 4\theta = 1$. Es decir, la órbita de 1 es el conjunto $\{1, 2, 3, 8, 5, 6, 4\}$. Las órbitas de 7 y 9 que se encuentra son $\{7\}$ y $\{9\}$, respectivamente. Los ciclos de θ son, pues, $(7), (9), (1, 1\theta, 1\theta^2, \dots, 1\theta^6) = (1, 2, 3, 8, 5, 6, 4)$. El lector debe verificar ahora que si toma el producto (de acuerdo a como se definió en el último párrafo) de $(1, 2, 3, 8, 5, 6, 4)$, (7) y (9) , obtendrá θ . Es decir, al menos en este caso, θ es el producto de sus ciclos.

Pero no es esto ninguna casualidad, pues resulta ahora trivial probar que el

LEMA 2.22. *Toda permutación es el producto de sus ciclos.*

Prueba. Sea θ la permutación. Entonces sus ciclos son de la forma $(s, s\theta, \dots, s\theta^{l-1})$. De acuerdo con la multiplicación de ciclos, según lo definimos anteriormente, y como los ciclos de θ son ajenos, la imagen de $s' \in S$ bajo θ , que es $s'\theta$, es la misma que la imagen de s' bajo el producto, ψ , de todos los ciclos distintos de θ . Luego θ y ψ tienen el mismo efecto sobre cada uno de los elementos de S , de donde $\theta = \psi$, que es lo que queríamos probar.

Si las observaciones anteriores no nos resultan aún claras a estas alturas, el lector debe tomar una permutación determinada, encontrar sus ciclos, tomar su producto, y verificar el lema. Al hacer esto el propio lema resultará obvio.

El lema 2.22 se plantea usualmente en la forma: *toda permutación puede expresarse en forma única como un producto de ciclos ajenos.*

Consideremos el m -ciclo $(1, 2, \dots, m)$. Un simple cálculo muestra que $(1, 2, \dots, m) = (1, 2)(1, 3) \dots (1, m)$. Más generalmente, el m -ciclo $(a_1, a_2, \dots, a_m) = (a_1, a_2)(a_1, a_3) \dots (a_1, a_m)$. Esta descomposición no es única; queremos decir con esto que un m -ciclo puede escribirse como un producto de ciclos de orden 2 en más de una forma. Por ejemplo, $(1, 2, 3) = (1, 2)(1, 3) = (3, 1)(3, 2)$. Ahora bien, como toda permutación es un producto de ciclos ajenos y todo ciclo es un producto de ciclos de orden dos, hemos probado que

LEMA 2.23. *Toda permutación es un producto de ciclos de orden 2.*

A los ciclos de orden dos les llamaremos *transposiciones*.

DEFINICIÓN. Una permutación $\theta \in S_n$ se dice que es una *permutación par* si puede representarse como un producto de un número par de transposiciones.

La definición que acabamos de dar insiste en que θ tenga una representación como un producto de un número par de transposiciones. Quizá tenga otras representaciones como un producto de un número impar de transposiciones. Lo primero que queremos ver es que tal cosa no es posible. Francamente, no nos sentimos muy satisfechos con la prueba que aquí damos de este hecho porque en ella se introduce un polinomio que parece un objeto extraño a nuestro sujeto actual de estudio.

Consideremos el polinomio en n -variables $p(x_1, \dots, x_n) = \prod_{i < j} (x_i - x_j)$.

Si $\theta \in S_n$ hagamos actuar a θ sobre el polinomio $p(x_1, \dots, x_n)$ por medio de la regla $\theta : p(x_1, \dots, x_n) = \prod_{i < j} (x_i - x_j) \rightarrow \prod_{i < j} (x_{\theta(i)} - x_{\theta(j)})$. Es claro que $\theta : p(x_1, \dots, x_n) \rightarrow \pm p(x_1, \dots, x_n)$. Por ejemplo, en S_5 , $\theta = (134)(25)$ lleva a

$$p(x_1, \dots, x_5) = (x_1 - x_2)(x_1 - x_3)(x_1 - x_4)(x_1 - x_5)(x_2 - x_3) \\ \times (x_2 - x_4)(x_2 - x_5)(x_3 - x_4)(x_3 - x_5)(x_4 - x_5)$$

en

$$(x_3 - x_5)(x_3 - x_4)(x_3 - x_1)(x_3 - x_2)(x_5 - x_4)(x_5 - x_1) \\ \times (x_5 - x_2)(x_4 - x_1)(x_4 - x_2)(x_1 - x_2)$$

que puede verificarse que es $-p(x_1, \dots, x_5)$.

Si, en particular, θ es un transposición, $\theta : p(x_1, \dots, x_n) \rightarrow -p(x_1, \dots, x_n)$. (¡Verifíquese!) Así pues, si una permutación Π puede representarse como un producto de un número par de transposiciones en una representación, Π debe dejar $p(x_1, \dots, x_n)$ fijo, de modo que cualquier representación de Π como un producto de transposiciones debe ser tal que deje $p(x_1, \dots, x_n)$ fijo; es decir, en cualquier representación es un producto de un número par de transposiciones. Esto demuestra que la definición dada de una permuta-

ción par es significativa. Llamamos a una permutación *impar* si no es una permutación par.

Los siguientes hechos resultan ahora claros.

- 1) El producto de dos permutaciones pares es una permutación par.
- 2) El producto de una permutación par y una impar es una impar (y, lo mismo, el producto de una impar por una par).
- 3) El producto de dos permutaciones impares es una permutación par.

La regla para combinar permutaciones pares e impares es como la de combinar números pares e impares en la adición. Y esto no es una coincidencia ya que esta última regla se usa para establecer 1, 2 y 3.

Sea A_n el subconjunto de S_n consistente en todas las permutaciones pares. Como el producto de dos permutaciones pares es una permutación par, A_n debe ser un subgrupo de S_n . Quizá la mejor forma de ver esto es la que sigue: sea W el grupo de los números reales 1 y -1 bajo la multiplicación. Definamos $\psi : S_n \rightarrow W$ por $\psi(s) = 1$ si s es una permutación par, $\psi(s) = -1$ si s es una permutación impar. Por las reglas 1, 2 y 3 anteriores ψ es un homomorfismo sobre W . El núcleo de ψ es precisamente A_n ; como núcleo de un homomorfismo A_n es un subgrupo normal de S_n . Según el teorema 2.d, $S_n/A_n \approx W$, luego como

$$2 = o(W) = o\left(\frac{S_n}{A_n}\right) = \frac{o(S_n)}{o(A_n)},$$

vemos que $o(A_n) = \frac{1}{2}n!$ A A_n se le llama *grupo alternante* de grado n . Resumimos nuestras observaciones en el siguiente

LEMA 2.24. *S_n tiene como subgrupo normal de índice 2 al grupo alternante, A_n , consistente en todas las permutaciones pares.*

Al final de la próxima sección volveremos nuevamente a tratar S_n .

Problemas

1. Encuéntrense las órbitas y ciclos de las siguientes permutaciones:

a) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 3 & 4 & 5 & 1 & 6 & 7 & 9 & 8 \end{pmatrix}.$

b) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 5 & 4 & 3 & 1 & 2 \end{pmatrix}.$

2. Escribanse las permutaciones del problema 1 como si fueran el producto de ciclos ajenos.

3. Exprésense como el producto de ciclos ajenos:
 - a) $(1, 2, 3)(4, 5)(1, 6, 7, 8, 9)(1, 5)$.
 - b) $(1, 2)(1, 2, 3)(1, 2)$.
4. Pruébese que $(1, 2, \dots, n)^{-1} = (n, n-1, n-2, \dots, 2, 1)$.
5. Encuéntrese la estructura cíclica de todas las potencias de $(1, 2, \dots, 8)$.
6. a) ¿Cuál es el orden de un n -ciclo?
 b) ¿Cuál es el orden del producto de los ciclos ajenos de longitudes m_1, m_2, \dots, m_k ?
 c) ¿Cómo se encuentra el orden de una permutación dada?
7. Calcúlese $a^{-1}ba$ cuando
 - 1) $a = (1, 3, 5)(1, 2)$ $b = (1, 5, 7, 9)$.
 - 2) $a = (5, 7, 9)$ $b = (1, 2, 3)$.
8. a) Dada la permutación $x = (1, 2)(3, 4)$ y la $y = (5, 6)(1, 3)$, encuéntrese una permutación a tal que $a^{-1}xa = y$.
 b) Pruébese que no existe a alguno tal que $a^{-1}(1, 2, 3)a = (1, 3)(5, 7, 8)$.
 c) Pruébese que no hay permutación a alguna tal que $a^{-1}(1, 2)a = (3, 4)(1, 5)$.
9. Determinése para qué m un m -ciclo es una permutación par.
10. Determinése cuáles de las siguientes son permutaciones pares:
 - a) $(1, 2, 3)(1, 2)$.
 - b) $(1, 2, 3, 4, 5)(1, 2, 3)(4, 5)$.
 - c) $(1, 2)(1, 3)(1, 4)(2, 5)$.
11. Pruébese que el subgrupo mínimo de S_n que contiene $(1, 2)$ y $(1, 2, \dots, n)$ es S_n . (En otras palabras, que éstos generan S_n .)
- *12. Pruébese que para $n \geq 3$ el subgrupo generado por los 3-ciclos es A_n .
- *13. Pruébese que si un subgrupo normal de A_n contiene aunque solo sea un 3-ciclo entonces debe ser forzosamente todo A_n .
- *14. Pruébese que A_5 no tiene ningún subgrupo normal $N \neq (e), A_5$.
15. Suponiendo cierto lo que se afirma en el problema 14, pruébese que cualquier subgrupo de A_5 es, cuando más, de orden 12.

11. OTRO PRINCIPIO DE CONTEO

La matemática es rica en técnica y argumentos. En esta gran variedad de situaciones una de las más esenciales herramientas de trabajo es el conteo. Y, sin embargo, aunque parezca extraño, es una de las cosas más

difíciles. Al hablar de conteo, a lo que queremos referirnos es al proceso de saber en forma precisa cuántas son todas las posibilidades en situaciones altamente complejas. Algunas veces esto puede conseguirse por un proceso exhaustivo primitivo de enumeración de caso tras caso, pero tal rutina es invariablemente pesada y viola el sentido matemático de la estética. Es siempre preferible el toque ligero, diestro, delicado, al golpe del mazo. Pero la objeción más seria al estudio caso por caso, es la de que es eficaz solo en muy pocos casos. Es por eso que en varias fases de la matemática encontramos procedimientos de conteo exactos que nos dicen con toda precisión cuántos elementos, dentro de un contexto muy amplio, satisfacen ciertas condiciones. Un gran favorito entre los matemáticos es el proceso de contar en una situación dada de dos formas diferentes; la comparación de los dos conteos se usa luego como un método de llegar a ciertas conclusiones. Generalmente hablando, se introduce una relación de equivalencia sobre un conjunto finito, se mide el tamaño de las clases de equivalencia de esta relación, y luego se iguala el número de elementos del conjunto a la suma de los órdenes de estas clases de equivalencia. Ilustraremos este tipo de enfoque en la presente sección. Introduciremos una relación, probaremos que es una relación de equivalencia, y hallaremos luego una clara descripción algebraica del tamaño de cada clase de equivalencia. De esta simple descripción surgirá una multitud de bellos y poderosos resultados acerca de los grupos finitos.

DEFINICIÓN. Si $a, b \in G$, entonces b se dice que es un *conjugado* de a en G si existe un elemento $c \in G$ tal que $b = c^{-1}ac$.

Escribiremos para esto $a \sim b$ y nos referiremos a esta relación como *conjugación*.

LEMA 2.25. *La conjugación es una relación de equivalencia sobre G .*

Prueba. Como de costumbre, para demostrar lo afirmado debemos probar que:

- 1) $a \sim a$;
- 2) $a \sim b$ implica $b \sim a$;
- 3) $a \sim b, b \sim c$ implica $a \sim c$

para todo a, b y c de G .

Probaremos cada una de tales afirmaciones.

1) Como $a = e^{-1}ae, a \sim a$, con $c = e$ sirviendo como la c de la definición de conjugación.

2) Si $a \sim b$, entonces $b = x^{-1}ax$ para algún $x \in G$, de donde, $a = (x^{-1})^{-1}b(x^{-1})$ y como $y = x^{-1} \in G$ y $a = y^{-1}by$, se sigue que $b \sim a$.

3) Supongamos que $a \sim b$ y $b \sim c$ con $a, b, c \in G$. Entonces $b = x^{-1}ax$ y $c = y^{-1}by$ para ciertos $x, y \in G$. Sustituyendo b por su expresión en la

expresión para c , obtenemos: $c = y^{-1}(x^{-1}ax)y = (xy)^{-1}a(xy)$; como $xy \in G$, se tiene, en consecuencia, que $a \sim c$.

Para $a \in G$ sea $C(a) = \{x \in G \mid a \sim x\}$. $C(a)$, la clase de equivalencia de a en G respecto a la relación que estudiamos, se llama usualmente *clase de conjugados* de a en G ; consiste en el conjunto de todos los distintos elementos de la forma $y^{-1}ay$ cuando y toma valores en G .

Nuestra atención va a circunscribirse ahora al caso en que G es un grupo finito. Supongamos que $C(a)$ tiene c_a elementos. Busquemos una descripción alternativa de c_a . Antes de hacerlo, observemos que $\sigma(G) = \sum c_a$ donde la suma corre sobre un conjunto de $a \in G$ usando una a para cada una de las clases de conjugados. Esta observación es, desde luego, simplemente una reformulación del hecho de que nuestra relación de equivalencia —conjugación— induce una descomposición de G en clases de equivalencia ajenas —las clases de conjugados. De máximo interés ahora es la evaluación de c_a .

Para lograrlo, recordamos un concepto que introdujimos en el problema 12, sección 5. Como este concepto es importante —demasiado importante para dejar su conocimiento al azar de que el lector haya o no resuelto ese problema— volvemos a lo que muy bien puede ya ser un terreno familiar para muchos de los lectores.

DEFINICIÓN. Si $a \in G$, entonces $N(a)$, el *normalizador de a en G* , es el conjunto $N(a) = \{x \in G \mid xa = ax\}$.

$N(a)$ consiste precisamente de aquellos elementos de G que comutan con a .

LEMA 2.26. $N(a)$ es un subgrupo de G .

Prueba. En este resultado el orden de G , sea este finito o infinito, carece de importancia, por lo que no ponemos restricción alguna sobre cuál sea ese orden.

Supongamos que $x, y \in N(a)$. Tenemos pues, $xa = ax$ y $ya = ay$. Por tanto $(xy)a = x(ya) = x(ay) = (xa)y = (ax)y = a(xy)$, luego $xy \in N(a)$. De $ax = xa$ se sigue que $x^{-1}a = x^{-1}(ax)x^{-1} = x^{-1}(xa)x^{-1} = ax^{-1}$, luego que x^{-1} está también en $N(a)$. Pero con ello hemos demostrado que $N(a)$ es un subgrupo de G .

Estamos ahora en buena posición para enunciar nuestro principio de conteo.

TEOREMA 2.H. Si G es un grupo finito, entonces $c_a = \frac{\sigma(G)}{\sigma(N(a))}$; en otras palabras, el número de elementos conjugados a a en G es el índice del normalizador de a en G .

Prueba. Para comenzar, la clase de conjugados de a en G , $C(a)$, consiste exactamente en todos los elementos $x^{-1}ax$ cuando x recorre G . c_a mide el número de los distintos $x^{-1}ax$. Nuestro método de prueba será mostrar que dos elementos en la misma clase lateral derecha de $N(a)$ en G dan lugar a un mismo conjugado de a , mientras que dos elementos en diferentes clases laterales derechas de $N(a)$ en G dan lugar a diferentes conjugados de a . De esta forma tendremos una correspondencia biyectiva entre conjugados de a y clases laterales derechas de $N(a)$ en G .

Supongamos que $x, y \in G$ están en una misma clase lateral derecha de $N(a)$ en G . Entonces $y = nx$ donde $n \in N(a)$ y entonces $na = an$. Por lo tanto, como $y^{-1} = (nx)^{-1} = x^{-1}n^{-1}$, $y^{-1}ay = x^{-1}n^{-1}anx = x^{-1}n^{-1}nax = x^{-1}ax$, es decir, x y y dan lugar a un mismo conjugado de a .

Si, por otra parte, x y y están en clases laterales derechas distintas de $N(a)$ en G , afirmamos que $x^{-1}ax \neq y^{-1}ay$. Si no fuera este el caso, de $x^{-1}ax = y^{-1}ay$ deduciríamos que $yx^{-1}a = ayx^{-1}$; esto, a su vez, implicaría que $yx^{-1} \in N(a)$. Pero esto nos dice que x y y están en la misma clase lateral derecha de $N(a)$ en G , lo que contradice el hecho de que están en diferentes clases laterales. Y la prueba es completa.

COROLARIO.

$$o(G) = \sum \frac{o(G)}{o(N(a))}$$

donde esta suma se efectúa tomando un elemento a en cada clase de conjugados.

Prueba. Como $o(G) = \sum c_a$, usando el teorema, el corolario se deduce de inmediato.

A la ecuación en este corolario se suele llamar *ecuación de clase* de G .

Antes de proseguir con las aplicaciones de estos resultados examinemos estos conceptos en algún grupo específico. Nada se saca al considerar grupos abelianos, porque dos elementos son conjugados si y sólo si son iguales (es decir, $c_a = 1$ para todo a). Nos fijamos, pues, de nuevo en nuestro antiguo amigo, el grupo S_3 . Sus elementos son $e, (1, 2), (1, 3), (2, 3), (1, 2, 3), (1, 3, 2)$. Enumeramos las clases conjugadas:

$$C(e) = \{e\}$$

$$C(1, 2) = \{(1, 2), (1, 3)^{-1}(1, 2)(1, 3), (2, 3)^{-1}(1, 2)(2, 3), (1, 2, 3)^{-1}(1, 2)(1, 2, 3), (1, 3, 2)^{-1}(1, 2)(1, 3, 2)\} = \{(1, 2), (1, 3), (2, 3)\} \text{ (Verifíquese)}$$

$$C(1, 2, 3) = \{(1, 2, 3), (1, 3, 2)\} \text{ (después de otra verificación).}$$

El lector debe verificar que $N((1, 2)) = \{e, (1, 2)\}$ y que $N((1, 2, 3)) = \{e, (1, 2, 3), (1, 3, 2)\}$, de modo que $c_{(1, 2)} = \frac{6}{2} = 3$ y $c_{(1, 2, 3)} = \frac{6}{3} = 2$.

APLICACIONES DEL TEOREMA 2.H. El teorema 2.h se presta inmediatamente a una poderosa aplicación. No necesitamos construcción artificial alguna para ilustrar su uso, pues los resultados que siguen que revelan la fuerza del teorema son ellos mismos teoremas de estatura e importancia.

Recordemos que el centro $Z(G)$ de un grupo G es el conjunto de todos los $a \in G$ tales que $ax = xa$ para todo $x \in G$. Obsérvese que

SUBLEMA. $a \in Z$ si y sólo si $N(a) = G$. Si G es finito, $a \in Z$ si y sólo si $o(N(a)) = o(G)$.

Prueba. Si $a \in Z$, $xa = ax$ para todo $x \in G$, de donde $N(a) = G$. Si, recíprocamente, $N(a) = G$, $xa = ax$ para toda $x \in G$, de modo que $a \in Z$. Si G es finito, $o(N(a)) = o(G)$ es equivalente a $N(a) = G$.

APLICACIÓN I

TEOREMA 2.I. Si $o(G) = p^n$ donde p es un número primo, entonces $Z(G) \neq (e)$.

Prueba. Si $a \in G$, como $N(a)$ es un subgrupo de G , $o(N(a))$, por ser un divisor de $o(G) = p^n$ debe ser de la forma $o(N(a)) = p^{n_a}$; $a \in Z(G)$ si y sólo si $n_a = n$. Escribamos la ecuación de clase para esta G , haciendo $z = o(Z(G))$. Tenemos $p^n = o(G) = \Sigma(p^n/p^{n_a})$; pero como hay exactamente z elementos tales que $n_a = n$, nos encontramos con que

$$p^n = z + \sum_{n_a < n} \frac{p^n}{p^{n_a}}.$$

Pero fijémonos en esto: p es un divisor del primer miembro; como $n_a < n$ para cada término en la Σ del segundo miembro,

$$p \left| \frac{p^n}{p^{n_a}} = p^{n-n_a} \right.$$

de modo que p es un divisor de cada uno de los términos de esta suma, de donde un divisor de la suma. Por tanto,

$$p \left| \left(p^n - \sum_{n_a < n} \frac{p^n}{p^{n_a}} \right) = z \right.$$

Como $e \in Z(G)$, $z \neq 0$; luego z es un entero positivo divisible por el primo p . Por tanto, $z > 1$. ¡Pero entonces debe haber un elemento además del e en $Z(G)$! Y esto es lo que afirma el teorema.

Dicho de otra manera, el teorema afirma que un grupo que tiene como orden la potencia de un primo debe siempre tener un centro no trivial.

Podemos ahora probar, como corolario de esto, un resultado que presentábamos en un problema anterior.

COROLARIO. *Si $o(G) = p^2$ donde p es un número primo, entonces G es abeliano.*

Prueba. Nuestra meta es demostrar que $Z(G) = G$. Por el momento sabemos ya que $Z(G) \neq (e)$ es un subgrupo de G de modo que $o(Z(G)) = p$ o p^2 . Si $o(Z(G)) = p^2$, entonces $Z(G) = G$ de acuerdo con lo afirmado. Supongamos por un momento que $o(Z) = p$; sea $a \in G$, $a \notin Z(G)$. $N(a)$ es, pues, un subgrupo de G , $Z(G) \subset N(a)$, $a \in N(a)$, de modo que $o(N(a)) > p$, y según el teorema de Lagrange $o(N(a))|o(G) = p^2$. Pero, entonces, forzosamente ha de tenerse $o(N(a)) = p^2$, de donde $a \in Z(G)$ en contra de lo supuesto. Luego $o(Z(G)) = p$ no es, realmente, posible.

APLICACIÓN 2. Usamos ahora el teorema 2.h para probar un importante teorema debido a Cauchy. Puede ser que el lector recuerde que este teorema ha sido ya probado para grupos abelianos como una aplicación de los resultados que encontramos en la sección sobre homomorfismos. En realidad, haremos uso de este caso especial en la prueba que sigue. Pero, para ser frances, en la sección que sigue inmediatamente a esta probaremos un resultado mucho más fuerte, debido a Sylow, que tiene el teorema de Cauchy como corolario inmediato en una forma que prescinde completamente del teorema 2.h. Realmente, si el teorema de Cauchy fuese nuestra última y única meta, podríamos probarlo usando tan solo los resultados más esenciales de la teoría de grupos en unas pocas líneas. (El lector puede ver la encantadora y brevíssima prueba del teorema de Cauchy encontrada por McKay y publicada en el *American Mathematical Monthly*, vol. 66 (1959), pág. 119.) Sin embargo, a pesar de todos estos argumentos en contra, presentamos aquí la prueba como una sorprendente ilustración del teorema 2.h.

TEOREMA 2.J (CAUCHY). *Si p es un número primo y $p|o(G)$, entonces G tiene un elemento de orden p .*

Prueba. Buscamos un elemento $a \neq e \in G$ que satisfaga $a^p = e$. Para probar su existencia procedemos por inducción sobre $o(G)$; es decir, suponemos que el teorema es cierto para todos los grupos T tales que $o(T) < o(G)$. No necesitamos molestarnos por el comienzo de la inducción pues el resultado es cierto por vacuidad para grupos de orden 1.

Si, para cualquier subgrupo W de G , $W \neq G$, se tuviera que $p \nmid o(W)$, entonces, de nuestra hipótesis de inducción existiría un elemento de orden p en W y habría, por tanto, un tal elemento en G . Podemos, pues, suponer que p no es un divisor del orden de ningún subgrupo propio de G . En

particular, si $a \notin Z(G)$, como $N(a) \neq G$, $p \nmid o(N(a))$. Escribamos la ecuación de clase:

$$o(G) = o(Z(G)) + \sum_{N(a) \neq G} \frac{o(G)}{o(N(a))}.$$

Como $p \mid o(G)$, $p \nmid o(N(a))$, tenemos que

$$p \mid \frac{o(G)}{o(N(a))},$$

y por tanto,

$$p \mid \sum_{N(a) \neq G} \frac{o(G)}{o(N(a))},$$

como tenemos también que $p \mid o(G)$, concluimos que

$$p \mid \left(o(G) - \sum_{N(a) \neq G} \frac{o(G)}{o(N(a))} \right) = o(Z(G)).$$

$Z(G)$ es, pues, un subgrupo de G cuyo orden es divisible por p . Pero hemos comenzado por suponer que p no es un divisor del orden de ningún subgrupo propio de G , luego $Z(G)$ no puede ser un subgrupo propio de G . Luego vemos que la única posibilidad que nos queda es aceptar que $Z(G) = G$. Pero, entonces, G es abeliano; y solo nos queda invocar el resultado ya establecido para los grupos abelianos para completar la inducción. Y esto prueba el teorema.

Concluimos esta sección considerando la relación de conjugación en una clase específica de grupos, a saber, los grupos simétricos S_n .

Dado el entero n , decimos que la sucesión de enteros positivos n_1, n_2, \dots, n_r , $n_1 \leq n_2 \leq \dots \leq n_r$, constituye una partición de n , si $n = n_1 + n_2 + \dots + n_r$. Denotemos por $p(n)$ el número de particiones de n . Determinemos $p(n)$ para pequeños valores de n :

$$p(1) = 1 \text{ ya que } 1 = 1 \text{ es la sola partición de } 1$$

$$p(2) = 2 \text{ ya que } 2 = 2 \text{ y } 2 = 1+1$$

$$p(3) = 3 \text{ ya que } 3 = 3, 3 = 1+2, 3 = 1+1+1$$

$$\begin{aligned} p(4) &= 4 \text{ ya que } 4 = 4, 4 = 1+3, 4 = 1+1+2, \\ &\quad 4 = 1+1+1+1, 4 = 2+2. \end{aligned}$$

Algunos otros son $p(5) = 7$, $p(6) = 11$, $p(61) = 1\,121\,505$. Hay mucha literatura matemática sobre $p(n)$.

Cada vez que descomponemos una permutación dada en S_n en un producto de ciclos ajenos obtenemos una partición de n ; pues si los ciclos que aparecen tienen longitudes n_1, n_2, \dots, n_r , respectivamente, $n_1 \leq n_2 \leq \dots \leq n_r$.

$\leq \dots \leq n_r$, entonces $n = n_1 + n_2 + \dots + n_r$. Diremos que una permutación $\sigma \in S_n$ tiene el ciclo de descomposición $\{n_1, n_2, \dots, n_r\}$ si puede escribirse como el producto de ciclos ajenos de longitudes n_1, n_2, \dots, n_r , $n_1 \leq n_2 \leq \dots \leq n_r$. Así pues, en S_9

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

tiene ciclo de descomposición $\{1, 1, 2, 2, 3\}$; obsérvese que $1+1+2+2+3=9$. Vamos ahora a intentar probar que dos permutaciones en S_n son conjugadas si y sólo si tienen el mismo ciclo de descomposición. Una vez probado esto es evidente que S_n tendrá exactamente $p(n)$ clases conjugadas.

Para alcanzar nuestro objetivo mostramos una regla muy simple para calcular el conjugado de una permutación dada. Supongamos que $\sigma \in S_n$ y que σ envía $i \rightarrow j$. ¿Cómo encontramos $\theta^{-1}\sigma\theta$ donde $\theta \in S_n$? Supongamos que θ manda $i \rightarrow s$ y $j \rightarrow t$; entonces $\theta^{-1}\sigma\theta$ manda $s \rightarrow t$. En otras palabras, para calcular $\theta^{-1}\sigma\theta$ reemplácese todo símbolo en σ por su imagen bajo θ . Por ejemplo, para determinar $\theta^{-1}\sigma\theta$ donde $\theta = (1, 2, 3)(4, 7)$ y $\sigma = (5, 6, 7)(3, 4, 2)$, entonces, como $\theta : 5 \rightarrow 5, 6 \rightarrow 6, 7 \rightarrow 4, 3 \rightarrow 1, 4 \rightarrow 7, 2 \rightarrow 3$, $\theta^{-1}\sigma\theta$ se obtiene de σ reemplazando en σ al 5 por el 5, al 6 por el 6, al 7 por el 4, al 3 por el 1, al 4 por el 7 y al 2 por el 3, de forma que $\theta^{-1}\sigma\theta = (5, 6, 4)(1, 7, 3)$.

Con este algoritmo para el cálculo de los conjugados se comprende que dos permutaciones que tienen la misma descomposición en ciclos son conjugadas. En efecto, si $\sigma = (a_1, a_2, \dots, a_{n_1})(b_1, b_2, \dots, b_{n_2}) \dots (x_1, x_2, \dots, x_{n_r})$ y $\tau = (\alpha_1, \alpha_2, \dots, \alpha_{n_1})(\beta_1, \beta_2, \dots, \beta_{n_2}) \dots (\chi_1, \chi_2, \dots, \chi_{n_r})$, entonces $\tau = \theta^{-1}\sigma\theta$ donde puede usarse como θ la permutación

$$\begin{pmatrix} a_1 & a_2 & \dots & a_{n_1} & b_1 & \dots & b_{n_2} & \dots & x_1 & \dots & x_{n_r} \\ \alpha_1 & \alpha_2 & \dots & \alpha_{n_1} & \beta_1 & \dots & \beta_{n_2} & \dots & \chi_1 & \dots & \chi_{n_r} \end{pmatrix}.$$

Así, por ejemplo, $(1, 2)(3, 4, 5)(6, 7, 8)$ y $(7, 5)(1, 3, 6)(2, 4, 8)$ pueden mostrarse son conjugadas usando como permutación conjugante la

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 7 & 5 & 1 & 3 & 6 & 2 & 4 & 8 \end{pmatrix}.$$

Que dos conjugadas tienen la misma descomposición en ciclos es ahora trivial pues, por nuestra regla, para calcular una conjugada reemplazamos cada elemento en un ciclo dado por su imagen bajo la permutación conjugante.

Reformulamos el resultado obtenido en la discusión previa en el siguiente lema.

LEMA 2.27. *El número de clases conjugadas en S_n es $p(n)$, el número de particiones de n .*

Como tenemos una descripción tan explícita de las clases conjugadas en S_n , podemos encontrar todos los elementos que conmutan con una permutación dada. Ilustramos esto con un caso particular muy sencillo.

Dada la permutación $(1, 2)$ en S_n , ¿qué elementos conmutan con ella? Es claro que cualquier permutación que deje fijos a 1 y a 2 permuta. Hay $(n-2)!$ de tales. También $(1, 2)$ conmuta consigo misma. De esta forma obtenemos $2(n-2)!$ elementos en el grupo generado por $(1, 2)$ y las $(n-2)!$ permutaciones que dejan 1 y 2 fijos. ¿Hay otras? Hay $n(n-1)/2$ transposiciones y éstas son precisamente todas las conjugadas de $(1, 2)$. Luego la clase conjugada de $(1, 2)$ tiene $n(n-1)/2$ elementos. Si el orden del normalizador de $(1, 2)$ es r , entonces, por nuestro principio de conteo,

$$\frac{n(n-1)}{2} = \frac{o(S_n)}{r} = \frac{n!}{r}.$$

Así pues, $r = 2(n-2)!$. Es decir, el orden del normalizador de $(1, 2)$ es $2(n-2)!$. Pero hemos exhibido $2(n-2)!$ elementos que conmutan con $(1, 2)$; luego el elemento general σ que conmuta con $(1, 2)$ es $\sigma = (1, 2)^i\tau$ donde $i = 0$ o 1 , y τ es una permutación que deja tanto a 1 como a 2 fijos.

Consideremos, como otra aplicación, la permutación $(1, 2, 3, \dots, n) \in S_n$. Afirmamos que este elemento conmuta solamente con sus potencias. Ciertamente conmuta con todas sus potencias, y esto da lugar a n elementos. Ahora bien, cualquier n -ciclo es conjugado de $(1, 2, \dots, n)$ y hay $(n-1)!$ n -ciclos distintos en S_n . Luego si u denota el orden del normalizador de $(1, 2, \dots, n)$ en S_n , como $o(S_n)/u =$ número de conjugados de $(1, 2, \dots, n)$ en $S_n = (n-1)!$,

$$u = \frac{n!}{(n-1)!} = n.$$

Luego el orden del normalizador de $(1, 2, \dots, n)$ en S_n es n . Las potencias de $(1, 2, \dots, n)$ nos han proporcionado tales n elementos, no hay lugar para otras, y hemos probado lo que habíamos afirmado.

Problemas

1. a) Pruébese que en S_n hay $\frac{1}{r} \frac{n!}{(n-r)!}$ r ciclos distintos.

b) Usando lo anterior, encuéntrese el número de conjugados que el r -ciclo $(1, 2, \dots, r)$ tiene en S_n .

- c) Pruébese que cualquier elemento σ en S_n que conmuta con $(1, 2, \dots, r)$ es de la forma $\sigma = (1, 2, \dots, r)^i\tau$, donde $i = 0, 1, 2, \dots, r$, y τ es una permutación que deja todos los elementos $1, 2, \dots, r$ fijos.
2. a) Encuéntrese el número de conjugados de $(1, 2)(3, 4)$ en S_n , para $n \geq 4$.
- b) Encuéntrese la forma de todos los elementos que conmutan con $(1, 2)(3, 4)$ en S_n .
3. Si p es un número primo, pruébese que en S_p hay $(p-1)!+1$ elementos x que satisfacen $x^p = e$.
4. Si en un grupo finito G un elemento a tiene exactamente dos conjugados, pruébese que G tiene un subgrupo normal $N \neq (e), G$.
5. a) Encuéntrense dos elementos en A_5 , el grupo alternante de grado 5, que son conjugados en S_5 , pero no en A_5 .
- b) Encuéntrense todas las clases conjugadas en A_5 y el número de elementos en cada clase conjugada.
6. a) Si N es un subgrupo normal de G y $a \in N$, muéstrese que todo conjugado de a en G está en N .
- b) Pruébese que $o(N) = \Sigma c_a$ para ciertas elecciones de a en N .
- c) Usando lo anterior y el resultado del problema 5(b), pruébese que en A_5 no hay ningún subgrupo normal N distinto del (e) y el A_5 .
7. Usando el teorema 2.i como argumento, pruébese que si $o(G) = p^n$, p un número primo, entonces G tiene un subgrupo de orden p^α para toda α tal que $0 \leq \alpha \leq n$.
8. Si $o(G) = p^n$, p un número primo, pruébese que existen subgrupos N_i , $i = 0, 1, \dots, r$ (para algún r) tales que $G = N_0 \supset N_1 \supset N_2 \supset \dots \supset N_r = (e)$ donde N_i es un subgrupo normal de N_{i-1} y donde N_{i-1}/N_i es abeliano.
9. Si $o(G) = p^n$, p un número primo, y $H \neq G$ es un subgrupo de G , muéstrese que existe un $x \in G$, $x \notin H$ tal que $x^{-1}Hx = H$.
10. Pruébese que cualquier subgrupo de orden p^{n-1} de un grupo G de orden p^n , p número primo, es normal en G .
- *11. Si $o(G) = p^n$, p un número primo, y si $N \neq (e)$ es un subgrupo normal de G , pruébese que $N \cap Z \neq (e)$ donde Z es el centro de G .
12. Si G es un grupo, Z su centro, y G/Z es cíclico, pruébese que G debe ser abeliano.
13. Pruébese que cualquier grupo de orden 15 es cíclico.

14. Pruébese que un grupo de orden 28 tiene un subgrupo normal de orden 7.

15. Pruébese que si un grupo G de orden 28 tiene un subgrupo normal de orden 4, entonces G es abeliano.

12. EL TEOREMA DE SYLOW

El teorema de Lagrange nos dice que el orden de un subgrupo de un grupo finito es un divisor del orden de ese grupo. El recíproco, sin embargo, es falso. Hay muy pocos teoremas que afirmen la existencia de subgrupos de un orden prescrito en grupos finitos arbitrarios. El más básico, y ampliamente usado, es un teorema clásico debido al matemático noruego Sylow. Presentamos aquí una prueba muy elegante y elemental del teorema de Sylow; la prueba, en la forma en que aquí aparece, se debe a Wielandt y apareció en la revista *Archiv der Matematik*, vol. 10 (1959), págs. 401-402.

TEOREMA 2.K (SYLOW). Si p es un número primo y $p^a \mid o(G)$, entonces G tiene un subgrupo de orden p^a .

Antes de entrar en la prueba del teorema haremos una pequeña digresión para hablar un poco de teoría de los números y combinatoria.

El número de formas en que podemos escoger un subconjunto de k elementos de un conjunto de n elementos puede fácilmente mostrarse que es

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Si $n = p^a m$ donde p es un número primo, y si $p^r \mid m$, pero $p^{r+1} \nmid m$, consideremos

$$\begin{aligned} \binom{p^a m}{p^a} &= \frac{(p^a m)!}{(p^a)!(p^a m - p^a)!} \\ &= \frac{p^a m(p^a m - 1) \dots (p^a m - i) \dots (p^a m - p^a + 1)}{p^a(p^a - 1) \dots (p^a - i) \dots (p^a - p^a + 1)} \end{aligned}$$

El problema es: ¿qué potencia de p divide a $\binom{p^a m}{p^a}$? Mirando este número escrito en la forma que lo acabamos de escribir, se puede ver que, excepto por el término m en el numerador, la potencia de p que divide a $(p^a m - i)$ es la misma que la que divide a $p^a - i$, de modo que todas las potencias

de p se cancelan, excepto la potencia que divide a m . Luego

$$p^r \left| \binom{p^\alpha m}{p^\alpha}, p^{r+1} \nmid \binom{p^\alpha m}{p^\alpha} \right.$$

Prueba del teorema. Sea \mathfrak{M} el conjunto de todos los subconjuntos de G que tienen p^α elementos. \mathfrak{M} tiene pues $\binom{p^\alpha m}{p^\alpha}$ elementos. Dados $M_1, M_2 \in \mathfrak{M}$

(M_1 es un subconjunto de G que tiene p^α elementos y lo mismo sucede con M_2) definamos $M_1 \sim M_2$ si existe un elemento $g \in G$ tal que $M_1 = M_2g$. Es inmediato verificar que con ello definimos una relación de equivalencia en \mathfrak{M} . Afirmamos que hay, al menos, una clase de equivalencia de elementos de \mathfrak{M} tal que el número de elementos en esta clase no es un múltiplo de p^{r+1} , pues si p^{r+1} fuese un divisor del tamaño de cada clase de equivalencia, entonces p^{r+1} sería un divisor del número de elementos en \mathfrak{M} .

Como \mathfrak{M} tiene $\binom{p^\alpha m}{p^\alpha}$ elementos y $p^{r+1} \nmid \binom{p^\alpha m}{p^\alpha}$, tal no puede ser el caso.

Sea $\{M_1, \dots, M_n\}$ una clase de equivalencia en \mathfrak{M} donde $p^{r+1} \nmid n$. Por la misma definición de equivalencia en \mathfrak{M} , si $g \in G$, para cada $i = 1, \dots, n$, $M_i g = M_j$ para algún j , $1 \leq j \leq n$. Sea $H = \{g \in G \mid M_1 g = M_1\}$. Es claro que H es un subgrupo de G , pues si $a, b \in H$, entonces $M_1 a = M_1$ y $M_1 b = M_1$, de donde $M_1 ab = (M_1 a)b = M_1 b = M_1$. Estaremos vitalmente interesados en $o(H)$. Afirmamos que $no(H) = o(G)$; dejamos que el lector efectúe la prueba, pero le sugerimos el argumento empleado al principio de la sección 11. Ahora, $no(H) = o(G) = p^\alpha m$; como $p^{r+1} \nmid n$ y $p^{r+1} \mid p^\alpha m = no(H)$, debe seguirse que $p^\alpha \mid o(H)$ y, por tanto, que $o(H) \geq p^\alpha$. Pero si $m_1 \in M_1$, entonces para todo $h \in H$, $m_1 h \in M_1$. Luego M_1 tiene al menos $o(H)$ distintos elementos. Pero M_1 era un subconjunto de G que contenía p^α elementos. Luego $p^\alpha \geq o(H)$. Lo que, combinado con $o(H) \geq p^\alpha$, nos dice que $o(H) = p^\alpha$. *Pero entonces hemos exhibido un subgrupo de G que tiene exactamente p^α elementos, a saber, H .* Esto prueba el teorema; realmente hemos hecho aun más que eso —hemos construido el subgrupo requerido.

Lo que usualmente se conoce como teorema de Sylow es un caso particular del teorema 2.k, a saber, el

COROLARIO. Si $p^m \mid o(G)$ y $p^{m+1} \nmid o(G)$, entonces G tiene un subgrupo de orden p^m .

Un subgrupo de G de orden p^m , donde $p^m \mid o(G)$, $p^{m+1} \nmid o(G)$, se llama p -subgrupo de Sylow de G . El teorema común de Sylow tiene otras dos partes más: una de ellas nos dice que cualesquiera dos p -subgrupos de Sylow de G son conjugados en G y la otra afirma que el número de

p -subgrupos de Sylow de G es de la forma $1 + kp$. No nos ocuparemos aquí de su prueba. En los problemas suplementarios, una combinación de los problemas 21-24 nos da éstas.

Problemas

1. En el grupo simétrico de grado 4, S_4 , encuéntrese un 2-subgrupo de Sylow y un 3-subgrupo de Sylow.

2. Pruébese que un grupo de orden 108 debe tener un subgrupo normal de orden 9 o 27.

Problemas suplementarios

No hay ninguna relación, en este capítulo, entre el orden en que los problemas aparecen y el orden de aparición de las secciones, que puedan ser relevantes para su solución. Nada se indica tampoco acerca de la dificultad de ninguno de los problemas.

1. a) Si G es un grupo abeliano finito con elementos a_1, a_2, \dots, a_n , pruébese que $a_1 a_2 \dots a_n$ es un elemento cuyo cuadrado es la identidad.
- b) Si el G de la parte (a) tiene más de un elemento de orden 2 o no tiene ningún elemento de orden 2, pruébese que $a_1 a_2 \dots a_n = e$.
- c) Si G tiene un elemento y de orden 2, entonces $a_1 a_2 \dots a_n = y$.
- d) (*Teorema de Wilson*). Si p es un número primo, entonces

$$(p-1)! \equiv -1(p).$$

2. Si p es un primo impar y si $1 + \frac{1}{2} + \frac{1}{3} + \dots + 1/(p-1) = a/b$, donde a y b son enteros, pruébese que $p|a$. Si $p > 3$, pruébese que $p^2|a$.

DEFINICIÓN. Un grupo G se dice que es *soluble* si existen subgrupos $G = N_0 \supset N_1 \supset N_2 \supset \dots \supset N_r = (e)$ tales que N_i es normal en N_{i-1} y N_{i-1}/N_i es abeliano.

3. Pruébese que un grupo soluble siempre tiene un subgrupo abeliano normal M tal que $M \neq (e)$.
4. Pruébese que la imagen homomorfa de un grupo soluble es soluble.
5. Si G es un grupo y N es un subgrupo normal de G tal que tanto N como G/N son solubles, pruébese que G es soluble.
6. Pruébese que un subgrupo de un grupo soluble es soluble.
7. Encuéntrense todos los automorfismos de S_3 , el grupo simétrico de grado 3.

8. Pruébese que la ecuación $x^2ax = a^{-1}$ es soluble para x en un grupo G si y sólo si a es el cubo de algún elemento en G .

9. Pruébese que $xax = b$ es soluble para x en G si y sólo si ab es el cuadrado de algún elemento en G .

10. Pruébese que $(1, 2, 3)$ no es el cubo de ningún elemento en S_n .

11. Supongamos que G es un grupo tal que $\phi(x) = x^n$ define un automorfismo de G . Pruébese que para todo $a \in G$, $a^{n^{-1}}$ está en el centro de G .

12. Si p es un número primo impar, $a \neq 0$ se dice que es un *residuo cuadrático de p* si existe un entero x tal que $x^2 \equiv a \pmod{p}$. Pruébese que:

a) Los residuos cuadráticos módulo p forman un subgrupo Q del grupo de los enteros distintos de cero mód p bajo la multiplicación.

b) $0(Q) = \frac{p-1}{2}$.

c) Si $q \in Q$, $n \notin Q$ (a n se le llama entonces un *no residuo*), entonces nq es un no residuo.

d) Si n_1 y n_2 son no residuos, entonces n_1n_2 es un residuo.

e) Si a es un residuo cuadrático de p entonces $a^{(p-1)/2} \equiv 1 \pmod{p}$.

13. Pruébese que en los enteros módulo p , p un número primo, hay cuando más n soluciones de $x^n \equiv 1 \pmod{p}$ para todo entero n .

14. Proporcionese un ejemplo de un grupo no abeliano en el que $(xy)^3 = x^3y^3$ para todo x y y .

15. Si G es un grupo, A un subgrupo de G y N un subgrupo normal de G , pruébese que si tanto A como N son solubles, entonces también lo es AN .

16. Si G es un grupo abeliano y a y b de G tienen órdenes m y n respectivamente, pruébese que hay un elemento c en G cuyo orden es el mínimo común múltiplo de m y n .

17. En el grupo abeliano finito G pruébese que el número de soluciones de $x^n = e$, donde $n|o(G)$, es un múltiplo de n .

18. Igual al problema 17, salvo que sin suponer que el grupo sea abeliano.

19. Si A y B son subgrupos de G , diremos que A es conjugado con B si $B = x^{-1}Ax$ para algún $x \in G$. Pruébese que:

a) La relación de conjugación es una relación de equivalencia sobre el conjunto de subgrupos de G .

b) Si $N(A) = \{x \in G | x^{-1}Ax = A\}$ entonces hay una correspondencia biyectiva entre las clases laterales derechas de $N(A)$ en G y los distintos conjugados de A .

20. Sea G un grupo finito y H un subgrupo de G . Para A y B subgrupos de G diremos que A es un conjugado de B relativo a H si $B = x^{-1}Ax$ para algún $x \in H$. Pruébese que:

- a) Esto define una relación de equivalencia sobre el conjunto de subgrupos de G .
- b) El número de subgrupos de G conjugados con A relativos a H es igual al índice de $N(A) \cap H$ en H .

21. a) Si G es un grupo finito y S_p un p -subgrupo de Sylow de G de orden p^m , pruébese que S_p es el único subgrupo de G de orden p^m que se encuentra en $N(S_p)$.

b) Si S_p es un p -subgrupo de Sylow de G y si a , de orden p^k , está en $N(S_p)$, entonces $a \in S_p$.

c) Pruébese que $N(N(S_p)) = N(S_p)$.

22. a) Si G es un grupo finito y si S_p es un p -subgrupo de Sylow de G pruébese que el número de conjugados de S_p en G no es un múltiplo de p .

b) Descomponiendo la clase de conjugados de S_p por el uso de la conjugación relativa a S_p , pruébese que la clase conjugada de S_p tiene $1 + kp$ subgrupos distintos (*Sugerencia: Úsese la parte (b) del problema 20 y el problema 21.*)

23. a) Si S_p es un p -subgrupo de Sylow y B un subgrupo de orden p^k en G , pruébese que si B no está contenido en algún conjugado de S_p el número de conjugados en G de S_p es un múltiplo de p .

b) Usando la parte (a) y el problema 22, pruébese que B debe estar contenido en algún conjugado de S_p .

c) Pruébese que cualesquiera dos p -subgrupos de Sylow de G (para el mismo primo p) son conjugados. (*Esta es la segunda parte del teorema de Sylow.*)

24. Usando los problemas 22 y 23, pruébese que n_p , el número de p -subgrupos de Sylow en G , es de la forma $n_p = 1 + kp$. (*Esta es la tercera parte del teorema de Sylow.*)

25. a) Usando el problema 24 pruébese que si $o(G) = 36$ entonces G tiene 1 o 4 3-subgrupos de Sylow.

b) Si $o(G) = 56$, pruébese que G tiene 1 u 8 7-subgrupos de Sylow. En el último caso, pruébese que los 2-subgrupos de Sylow de G deben ser normales en G .

26. Si $o(G) = 42$, pruébese que su 7-subgrupo de Sylow es normal.

27. Pruébese que un grupo de orden 48 debe tener un subgrupo normal de orden 8 o 16.

28. Mediante una discusión caso por caso, usando los resultados obtenidos en el capítulo, pruébese que un grupo de orden menor que 60 o es de orden primo o tiene un subgrupo no trivial normal.

29. a) Si $p > q$ son primos distintos, pruébese que un grupo de orden pq es soluble.

b) Si $q \nmid (p - 1)$, pruébese que un grupo de orden pq es cíclico.

c) Pruébese que cualesquiera dos grupos no abelianos de orden pq son isomorfos.

30. Pruébese que un grupo no puede ser escrito como la unión de los conjuntos de dos subgrupos propios.

31. Si un grupo G tiene un subgrupo propio de índice finito, pruébese que tiene un subgrupo normal de índice finito.

32. Si G es un grupo y $a \in G$ tiene orden finito y solamente un número finito de conjugados en G , pruébese que estos conjugados generan un subgrupo normal finito de G .

Lecturas supplementarias

BURNSIDE, W., *Theory of Groups of Finite Order*, segunda edición, Cambridge University Press, Cambridge, Inglaterra, 1911; Dover Publications, Inc., Nueva York, 1955.

HALL, MARSHALL, *Theory of Groups*. The Macmillan Company, Nueva York, 1959.

Topics para discusión en clase

MCKAY, JAMES H., "Another proof of Cauchy's group theorem", *American Mathematical Monthly*, vol. 66 (1959), pág. 119.

SEGAL, I. E., "The automorphisms of the symmetric group", *Bulletin of the American Mathematical Society*, vol. 46 (1940), pág. 565.

CAPÍTULO

3
+

Teoría de anillos

1. DEFINICIÓN Y EJEMPLOS DE ANILLOS

COMO indicamos en el capítulo 2, hay ciertos sistemas algebraicos que nos sirven como los bloques de construcción de las estructuras que componen la materia que actualmente se llama álgebra moderna. A estas alturas de nuestro estudio, hemos aprendido algo de uno de ellos, de los grupos. Ahora nuestro propósito es introducir y estudiar un segundo de los tales bloques, el constituido por los llamados anillos. El concepto abstracto de grupo tiene su origen en el conjunto de aplicaciones o permutaciones de un conjunto sobre sí mismo. En contraste, los anillos surgen de otra fuente bastante más familiar, el conjunto de los enteros. Veremos que están

caracterizados de acuerdo con los aspectos algebraicos de los enteros ordinarios de los que pueden considerarse una generalización.

En el próximo párrafo se aclarará que un anillo es completamente diferente de un grupo, ya que es un sistema bioperacional, en el que hay definidas dos operaciones; estas operaciones comúnmente se llaman adición y multiplicación. Sin embargo, a pesar de las diferencias, el análisis de los anillos seguirá el esquema que ya establecimos para los grupos. Tendremos los análogos de los homomorfismos, de los subgrupos normales, de los grupos factores, etc. Con la experiencia que hemos ganado en el estudio de los grupos, podremos dar las definiciones necesarias, entretejerlas con teoremas significativos, y terminar probando resultados que son interesantes e importantes acerca de objetos matemáticos que nos han sido familiares desde hace tiempo. Para citar solo un ejemplo, más adelante veremos en este libro, usando los argumentos que aquí desarrollamos, que es imposible trisecar un ángulo de 60° usando solamente regla y compás.

DEFINICIÓN. Un conjunto no vacío R se dice que es un *anillo asociativo* si en R están definidas dos operaciones, denotadas por “ $+$ ” y “ \cdot ” respectivamente tales que para cualesquiera a, b, c de R :

- 1) $a+b$ está en R .
- 2) $a+b = b+a$.
- 3) $(a+b)+c = a+(b+c)$.
- 4) Hay un elemento 0 en R tal que $a+0 = a$ (para todo a en R).
- 5) Existe un elemento $-a$ en R tal que $a+(-a) = 0$.
- 6) $a \cdot b$ está en R .
- 7) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.
- 8) $a \cdot (b+c) = a \cdot b + a \cdot c$ y $(b+c) \cdot a = b \cdot a + c \cdot a$ (las dos leyes distributivas).

Los axiomas (1) a (5) simplemente afirman que R es un grupo abeliano bajo la operación $+$ a la que llamamos adición. Los axiomas (6) y (7) nos dicen que R es cerrado bajo una operación asociativa, \cdot , a la que llamamos multiplicación. El axioma (8) sirve para correlacionar las dos operaciones de R .

Siempre que hablamos de un anillo se entenderá que estamos hablando de un anillo asociativo. Los anillos no asociativos, es decir, aquellos en que no se verifica el axioma 7, se presentan en matemáticas y son objeto de estudio, pero aquí no tendremos ocasión de considerarlos.

Puede y no suceder que exista un elemento 1 en R tal que $a \cdot 1 = 1 \cdot a = a$ para toda a en R ; si tal elemento existe diremos que R es un *anillo con elemento unitario*.

Si la multiplicación de R es tal que $a \cdot b = b \cdot a$ para todo a, b en R entonces llamamos a R *anillo conmutativo*.

Antes de comenzar a estudiar algunas propiedades de los anillos, haremos una pausa para examinar algunos ejemplos. Motivándonos en ellos definiremos varios casos especiales de anillos que son de importancia.

EJEMPLO 1. R es el conjunto de los enteros, positivos, negativos y el 0; $+$ es la adición usual y \cdot la multiplicación usual de los enteros. R es un anillo conmutativo con elemento unitario.

EJEMPLO 2. R es el conjunto de todos los enteros pares bajo las operaciones habituales de adición y multiplicación. R es un anillo conmutativo, pero no tiene elemento unitario.

EJEMPLO 3. R es el conjunto de los números racionales bajo la adición y multiplicación habituales de los números racionales. R es un anillo conmutativo con elemento unitario. Pero aún es más que eso, pues podemos ver que los elementos de R distintos del 0 forman un grupo abeliano bajo la multiplicación. Un anillo con esta última propiedad se llama *campo*.

EJEMPLO 4. R es el conjunto de los enteros módulo 7 bajo la adición y multiplicación módulo 7. Es decir, los elementos de R son los siete símbolos $\bar{0}, \bar{1}, \bar{2}, \bar{3}, \bar{4}, \bar{5}, \bar{6}$ donde:

1) $i+j = \bar{k}$ donde k es el residuo de la división de $i+j$ por 7 (así, por ejemplo, $\bar{4}+\bar{5} = \bar{2}$ ya que $4+5 = 9$, que, dividido por 7 da 2 como residuo).

2) $i \cdot j = \bar{m}$ donde m es el residuo de la división de ij por 7 (así, $\bar{3} \cdot \bar{3} = \bar{1}$, ya que $5 \cdot 3 = 15$ que tiene 1 como residuo de su división por 7). El estudiante debe verificar que R es un anillo conmutativo con elemento unidad. Pero puede decirse mucho más, a saber, puesto que:

$$\bar{1} \cdot \bar{1} = \bar{1} = \bar{6} \cdot \bar{6}$$

$$\bar{2} \cdot \bar{4} = \bar{1} = \bar{4} \cdot \bar{2}$$

$$\bar{3} \cdot \bar{5} = \bar{1} = \bar{5} \cdot \bar{3}$$

los elementos de R distintos de cero forman un grupo abeliano bajo la multiplicación. R es, pues, un campo. Como solamente tiene un número finito de elementos se llama *campo finito*.

EJEMPLO 5. R es el conjunto de los enteros módulo 6 bajo la adición y la multiplicación módulo 6. Si denotamos los elementos de R por $\bar{0}, \bar{1}, \bar{2}, \dots, \bar{5}$, vemos que $\bar{2} \cdot \bar{3} = \bar{0}$, aunque $\bar{2} \neq \bar{0}$ y $\bar{3} \neq \bar{0}$. Así pues es posible en un anillo R que $a \cdot b = 0$ sin que $a = 0$ ni $b = 0$. Esto no puede suceder en un campo (véase el problema 10 al final de la sección 2), luego el anillo R del ejemplo 5 no es, ciertamente, un campo.

En todos los ejemplos que hemos presentado, el anillo que apareció era un anillo conmutativo. Ahora presentaremos un anillo no conmutativo.

EJEMPLO 6. R será el conjunto de todos los símbolos $\alpha_{11}e_{11} + \alpha_{12}e_{12} + \alpha_{21}e_{21} + \alpha_{22}e_{22} = \sum_{i,j=1}^2 \alpha_{ij}e_{ij}$ donde todos los α_{ij} son números racionales y donde convenimos en que:

$$(1) \quad \sum_{i,j=1}^2 \alpha_{ij}e_{ij} = \sum_{i,j=1}^2 \beta_{ij}e_{ij}$$

si y sólo si para toda $i, j = 1, 2$, $\alpha_{ij} = \beta_{ij}$.

$$(2) \quad \sum_{i,j=1}^2 \alpha_{ij}e_{ij} + \sum_{i,j=1}^2 \beta_{ij}e_{ij} = \sum_{i,j=1}^2 (\alpha_{ij} + \beta_{ij})e_{ij}.$$

$$(3) \quad \left(\sum_{i,j=1}^2 \alpha_{ij}e_{ij} \right) \left(\sum_{i,j=1}^2 \beta_{ij}e_{ij} \right) = \sum_{i,j=1}^2 \gamma_{ij}e_{ij}$$

donde $\gamma_{ij} = \sum_{v=1}^2 \alpha_{iv}\beta_{vj} = \alpha_{i1}\beta_{1j} + \alpha_{i2}\beta_{2j}$

Esta multiplicación, a primera vista, parece algo complicada, pero está fundada en reglas relativamente sencillas, a saber, multiplicar $\sum \alpha_{ij}e_{ij}$ por $\sum \beta_{ij}e_{ij}$ formalmente multiplicando término a término y agrupar términos usando para ello las relaciones $e_{ij} \cdot e_{kl} = 0$ si $j \neq k$, $e_{ij} \cdot e_{ji} = e_{ii}$.

Para ilustrar la multiplicación, si $a = e_{11} - e_{21} + e_{22}$ y $b = e_{22} + 3e_{12}$, entonces

$$\begin{aligned} a \cdot b &= (e_{11} - e_{21} + e_{22}) \cdot (e_{22} + 3e_{12}) \\ &= e_{11} \cdot e_{22} + 3e_{11} \cdot e_{12} - e_{21} \cdot e_{22} - 3e_{21} \cdot e_{12} + e_{22} \cdot e_{22} + 3e_{22} \cdot e_{12} \\ &= 0 + 3e_{12} - 0 - 3e_{22} + e_{22} + 0 \\ &= 3e_{12} - 3e_{22} + e_{22} = 3e_{12} - 2e_{22}. \end{aligned}$$

Nótese que $e_{11} \cdot e_{12} = e_{12}$ mientras que $e_{12} \cdot e_{11} = 0$. Vemos, pues, que la multiplicación en R no es conmutativa. También vemos que es posible que $u \cdot v = 0$ con $u \neq 0$ y $v \neq 0$.

El lector debe verificar que R es realmente un anillo. Se le llama el anillo de las matrices racionales 2×2 . Tanto él como sus análogos ocuparán posteriormente una gran parte de nuestro estudio.

EJEMPLO 7. Sea C el conjunto de todos los símbolos (α, β) donde α y β son números reales. Definimos:

$$(1) \quad (\alpha, \beta) = (\gamma, \delta) \text{ si y sólo si } \alpha = \gamma \text{ y } \beta = \delta.$$

Introducimos en C una adición definiendo para $x = (\alpha, \beta)$ y $y = (\gamma, \delta)$

$$(2) \quad x + y = (\alpha, \beta) + (\gamma, \delta) = (\alpha + \gamma, \beta + \delta).$$

Nótese que $x + y$ sigue siendo un elemento de C . Afirmamos que C es un grupo abeliano bajo esta operación con $(0, 0)$ sirviendo como elemento

identidad para la adición, y $(-\alpha, -\beta)$ como el inverso, en la adición, de (α, β) .

Ahora que C ya está provisto de una adición, para que C sea un anillo necesitamos una multiplicación. Logramos esto al definir:

$$\text{para } X = (\alpha, \beta), Y = (\gamma, \delta) \text{ en } C$$

$$(3) \quad X \cdot Y = (\alpha, \beta) \cdot (\gamma, \delta) = (\alpha\gamma - \beta\delta, \alpha\delta + \beta\gamma).$$

Nótese que $X \cdot Y = Y \cdot X$. Además, $X \cdot (1, 0) = (1, 0) \cdot X = X$, de modo que $(1, 0)$ es un elemento unidad para C .

Observamos de nuevo que $X \cdot Y \in C$. Además, si $X = (\alpha, \beta) \neq (0, 0)$, entonces, como α, β son reales y no ambos 0, $\alpha^2 + \beta^2 \neq 0$, luego

$$Y = \left(\frac{\alpha}{\alpha^2 + \beta^2}, \frac{-\beta}{\alpha^2 + \beta^2} \right)$$

está en C . Finalmente vemos que

$$(\alpha, \beta) \cdot \left(\frac{\alpha}{\alpha^2 + \beta^2}, \frac{-\beta}{\alpha^2 + \beta^2} \right) = (1, 0).$$

Hemos demostrado totalmente que C es un campo. Si escribimos (α, β) como $\alpha + \beta i$ el lector puede verificar que C es simplemente una forma disfrazada de los familiares números complejos.

EJEMPLO 8. Este último ejemplo se llama con frecuencia el de los *cuaternios reales*. Este anillo fue descrito por primera vez por el matemático irlandés Hamilton. Inicialmente se usó mucho en el estudio de la mecánica; actualmente su interés primario es el de ser un ejemplo importante, pues aún juega un papel fundamental en la geometría y en la teoría de los números.

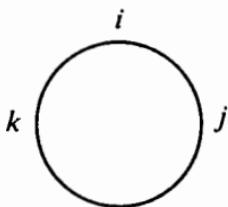
Sea Q el conjunto de todos los símbolos $\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$ donde todos los números $\alpha_0, \alpha_1, \alpha_2$ y α_3 son números reales. Convenimos en que dos de tales símbolos $\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$ y $\beta_0 + \beta_1 i + \beta_2 j + \beta_3 k$ son iguales si y sólo si $\alpha_t = \beta_t$ para $t = 0, 1, 2, 3$. Para que Q sea un anillo debemos definir $+$ y para sus elementos. Para ello definimos:

- 1) para cualquier $X = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$, y $Y = \beta_0 + \beta_1 i + \beta_2 j + \beta_3 k$ en Q , $X + Y = (\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k) + (\beta_0 + \beta_1 i + \beta_2 j + \beta_3 k) = (\alpha_0 + \beta_0) + (\alpha_1 + \beta_1)i + (\alpha_2 + \beta_2)j + (\alpha_3 + \beta_3)k$

y

- 2) $X \cdot Y = (\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k) \cdot (\beta_0 + \beta_1 i + \beta_2 j + \beta_3 k) = (\alpha_0\beta_0 - \alpha_1\beta_1 - \alpha_2\beta_2 - \alpha_3\beta_3) + (\alpha_0\beta_1 + \alpha_1\beta_0 + \alpha_2\beta_3 - \alpha_3\beta_2)i + (\alpha_0\beta_2 + \alpha_2\beta_0 + \alpha_3\beta_1 - \alpha_1\beta_3)j + (\alpha_0\beta_3 + \alpha_3\beta_0 + \alpha_1\beta_2 - \alpha_2\beta_1)k.$

Admitimos que esta fórmula para el producto parece formidable; sin embargo, su aspecto es más complicado de lo que realmente es. Resulta de multiplicar dos símbolos tales formalmente y agrupar términos usando las relaciones: $i^2 = j^2 = k^2 = ijk = -1$, $ij = -ji = k$, $jk = -kj = i$, $ki = -ik = j$. La parte última de estas relaciones, llamadas la tabla de multiplicación de las unidades cuaternionas, puede ser recordada por el pequeño diagrama que mostramos; cuando se recorre en el sentido de las manecillas del reloj se puede leer el producto, por ejemplo, $ij = k$, $jk = i$, $ki = j$, mientras que yendo en sentido contrario han de leerse los negativos.



Nótese que los elementos ± 1 , $\pm i$, $\pm j$, $\pm k$, forman un grupo no abeliano de orden 8 bajo el producto.

El lector puede probar que Q es un anillo no conmutativo en el que $0 = 0+0i+0j+0k$ y $1 = 1+0i+0j+0k$ sirven como el cero y el elemento unidad respectivamente. Ahora bien, si $X = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$ no es cero, entonces no todos los $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ son 0; como todos son reales de ello se sigue que $\beta = \alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2 \neq 0$. Luego

$$Y = \frac{\alpha_0}{\beta} - \frac{\alpha_1}{\beta}i - \frac{\alpha_2}{\beta}j - \frac{\alpha_3}{\beta}k \in Q.$$

Un simple cálculo nos muestra ahora que $X \cdot Y = 1$. Luego los elementos distintos del cero de Q forman un grupo no abeliano respecto a la multiplicación. Un anillo en que los elementos distintos del cero forman un grupo se llama *anillo con división* o *semicampo* (skew field). Desde luego, un anillo con división conmutativo es un campo. Q nos da un ejemplo de un anillo con división que no es un campo. Muchos otros ejemplos de anillos con división no conmutativos existen, pero tendríamos que extenderlos demasiado para presentar aquí otro. La investigación de la naturaleza de los anillos con división y los intentos para clasificarlos forman una parte importante del álgebra.

2. ALGUNAS CLASES ESPECIALES DE ANILLOS

Los ejemplos que acabamos de dar en la sección 1 indican claramente que aunque los anillos son una generalización directa de los enteros, ciertos hechos aritméticos a los que estamos acostumbrados en el anillo de los

enteros no tienen forzosamente que tener validez en los anillos en general. Por ejemplo, hemos visto la posibilidad de que $a \cdot b = 0$ sin que ni a ni b sean cero. Existen también ejemplos muy naturales en que $a \cdot b \neq b \cdot a$. Todas estas cosas van en contra de nuestra experiencia previa.

Por simplicidad en la expresión, prescindiremos de aquí en adelante del punto en $a \cdot b$ y escribiremos simplemente este producto como ab .

DEFINICIÓN. Si R es un anillo conmutativo entonces $a \neq 0 \in R$ se dice que es un *divisor de cero* si existe un $b \in R$, $b \neq 0$, tal que $ab = 0$.

DEFINICIÓN. Un anillo conmutativo es un *dominio entero* si no tiene divisores de cero.

El anillo de los enteros es un ejemplo de dominio entero.

DEFINICIÓN. Un anillo se dice que es un *anillo con división* si sus elementos distintos de cero forman un grupo bajo la multiplicación.

El elemento unidad bajo la multiplicación se escribirá como 1, y el inverso de un elemento a bajo la multiplicación se denominará como a^{-1} .

Damos finalmente la definición del muy importante objeto matemático conocido como campo.

DEFINICIÓN. Un *campo* es un anillo conmutativo con división.

En nuestros ejemplos de la sección 1, mostramos el anillo con división no conmutativo de los cuaternios reales y los siguientes campos; los números racionales, los números complejos, y los enteros módulo 7. El capítulo 5 se ocupará de los campos y sus propiedades.

Nosotros queremos estar en la posibilidad de calcular en anillos casi exactamente como con los números reales, recordando siempre que hay diferencias —puede suceder que $ab \neq ba$ o que no podamos dividir. Es en vista de estas posibilidades que queremos probar el próximo lema que confirma que ciertas cosas que nos gustaría que fuesen ciertas en los anillos lo son realmente.

LEMA 3.1. Si R es un anillo, entonces para todo $a, b \in R$

- 1) $a0 = 0a = 0$.
- 2) $a(-b) = (-a)b = -(ab)$.
- 3) $(-a)(-b) = ab$.

Si, además, R tiene un elemento unitario, 1, entonces

- 4) $(-1)a = -a$.
- 5) $(-1)(-1) = 1$.

Prueba

1) Si $a \in R$, entonces $a0 = a(0+0) = a0+a0$ (según la ley distributiva derecha), y como R es un grupo respecto a la adición, esta ecuación implica que $a0 = 0$.

Análogamente, $0a = (0+0)a = 0a+0a$, usando la ley distributiva izquierda, de donde, también aquí, se sigue que $0a = 0$.

2) Para probar que $a(-b) = -(ab)$ debemos mostrar que $ab + a(-b) = 0$. Pero $ab + a(-b) = a(b + (-b)) = a0 = 0$ según el uso de la ley distributiva y el resultado de la parte (1) de este lema. Análogamente $(-a)b = -(ab)$.

3) Que $(-a)(-b) = ab$ es en realidad un caso particular de la parte (2); lo remarcaremos porque su análogo en el caso de los números reales ha sido también subrayado en nuestra educación en la escuela. Tenemos con él:

$$\begin{aligned} (-a)(-b) &= -(a(-b)) && \text{(por parte la (2))} \\ &= -(-(ab)) && \text{(por parte la (2))} \\ &= ab \end{aligned}$$

pues $-(-x) = x$ es una consecuencia del hecho de que en cualquier grupo $(u^{-1})^{-1} = u$.

4) Supongamos que R tiene un elemento unitario 1; entonces $a + (-1)a = 1a + (-1)a = (1 + (-1))a = 0a = 0$, de donde $(-1)a = -a$. En particular, si $a = -1$, $(-1)(-1) = -(-1) = 1$, lo que deja establecido en la parte (5).

Con este lema ya establecido nos sentiremos libres de calcular con negativos y con 0 como lo hemos hecho anteriormente. El resultado del lema 3.1 es nuestro permiso para hacerlo. Por conveniencia, $a + (-b)$ se escribirá como $a - b$.

El lema que acabamos de probar, aunque muy útil e importante, no es ciertamente muy excitante. Prosigamos con resultados de un mayor interés. Antes de hacerlo enunciamos un principio que, aunque es completamente trivial nos provee de un arma poderosa, cuando es manejada con propiedad. Este principio no nos dice ni más ni menos que lo siguiente: si un cartero distribuye 101 cartas a 100 buzones, entonces un buzón debe tener al menos dos cartas. No suena muy prometedor como argumento, ¿no es así? Pues nos quedaremos sorprendidos. Las ideas matemáticas pueden a menudo ser muy difíciles y oscuras, pero ningún argumento de tal tipo puede esgrimirse contra este sencillísimo principio mencionado. Lo formalizamos e incluso le daremos un nombre.

EL PRINCIPIO DE LAS CASILLAS. Si n objetos se distribuyen sobre m lugares y si $n > m$ entonces algún lugar recibe al menos dos objetos.

Una formulación equivalente y que a menudo usaremos es: Si n objetos se distribuyen sobre n lugares en tal forma que ningún lugar reciba más de un objeto, entonces cada lugar recibe *exactamente* un objeto.

Inmediatamente hacemos uso de esta idea para probar que

LEMA 3.2. *Un dominio entero finito es un campo.*

Prueba. Recordemos que un dominio entero es un anillo conmutativo tal que $ab = 0$ si y sólo si al menos uno de los dos factores, a o b , es el mismo 0. Un campo, por otra parte, es un anillo conmutativo con elemento unidad en el que todo término distinto de cero tiene un multiplicativo inverso en el anillo.

Sea D un dominio entero finito. Para probar que D es un campo debemos

- 1) demostrar la existencia de un elemento $1 \in D$ tal que $a1 = a$ para todo $a \in D$;
- 2) probar que para todo elemento $a \neq 0 \in D$ existe un elemento $b \in D$ tal que $ab = 1$.

Sean x_1, x_2, \dots, x_n todos los elementos de D , y supongamos que $a \neq 0 \in D$. Consideremos los elementos x_1a, x_2a, \dots, x_na , todos ellos de D . ¡Afirmamos que todos son distintos! Pues si $x_i a = x_j a$ para $i \neq j$, entonces $(x_i - x_j)a = 0$. Como D es un dominio entero y $a \neq 0$, tendría que tenerse $x_i - x_j = 0$, de donde $x_i = x_j$ en contradicción con ser $i \neq j$. Tenemos pues que x_1a, x_2a, \dots, x_na son n distintos elementos de D que tiene exactamente n elementos. Según el principio de las casillas en tal conjunto deben aparecer todos los elementos de D ; dicho de otra manera, todo elemento y de D puede escribirse como $x_i a$ para algún x_i . En particular, como $a \in D$, $a = x_{i_0}a$ para algún $x_{i_0} \in D$. Como D es conmutativo, $a = x_{i_0}a = ax_{i_0}$. Nos proponemos demostrar que x_{i_0} actúa como un elemento unidad para todos los elementos de D . En efecto, si $y \in D$, como hemos visto, $y = x_i a$ para algún $x_i \in D$, y, por tanto, $yx_{i_0} = (x_i a)x_{i_0} = x_i(ax_{i_0}) = x_i a = y$. Luego x_{i_0} es un elemento unidad para D y lo que escribimos como 1. Ahora bien, $1 \in D$ es también realizable, según se desprende de nuestro argumento previo, como un múltiplo de a ; es decir, existe un $b \in D$ tal que $1 = ba$. Con lo que el lema ha quedado completamente probado.

COROLARIO. *Si p es un número primo entonces J_p , el anillo de los enteros módulo p , es un campo.*

Prueba. De acuerdo con el lema, es suficiente probar que J_p es un dominio entero, ya que solamente tiene un número finito de elementos. Si $a, b \in J_p$ y $ab \equiv 0$ entonces p debe dividir al entero ordinario ab , y por tanto p , por ser un primo, debe dividir a a o a b . Pero entonces o $a \equiv 0$ mód p o $b \equiv 0$ mód p , de donde en J_p una de ellas es cero.

Problemas

R es un anillo en todos los problemas.

1. Si $a, b, c, d \in R$ evalúese $(a+b)(c+d)$.
2. Pruébese que si $a, b \in R$ entonces $(a+b)^2 = a^2 + ab + ba + b^2$, donde por x^2 se debe entender xx .
3. Encuéntrese la forma del teorema del binomio en un anillo general; en otras palabras, encuéntrese una expresión para $(a+b)^n$ cuando n es un entero positivo.
4. Si todo $x \in R$ satisface $x^2 = x$ pruébese que R debe ser conmutativo. (Un anillo en el que $x^2 = x$ para todos los elementos se llama anillo booleano.)
5. Si R es un anillo, considerándolo solamente como un grupo abeliano respecto su adición en el capítulo 2 definimos lo que debe entenderse por na donde $a \in R$ y n es un entero. Pruébese que si $a, b \in R$ y n, m son enteros entonces $(na)(mb) = (nm)(ab)$.
6. Se dice que un dominio entero D es de característica 0 si la relación $ma = 0$ donde $0 \neq a \in D$ y m es un entero puede solamente verificarse si $m = 0$. Se dice que D es de característica finita si para algún $a \neq 0$ en D y algún entero $m \neq 0$, $ma = 0$. La *característica* de D se define entonces como el mínimo entero positivo p tal que $pa = 0$ para algún $a \neq 0$ en D . Pruébese que:
 - a) Si D es de característica p entonces $px = 0$ para todo $x \in D$.
 - b) La característica de un dominio entero o es 0 o un número primo.
7. Si R es un sistema que satisface todas las condiciones de los anillos con elemento unidad con la posible excepción de $a+b = b+a$, pruébese que el axioma $a+b = b+a$ debe verificarse en R , y que R es, por tanto, un anillo. (*Sugerencia:* desarrollese en dos formas $(a+b)(1+1)$.)
8. Demuéstrese que el anillo conmutativo D es un dominio entero si y sólo si para $a, b, c \in D$ con $a \neq 0$ la relación $ab = ac$ implica que $b = c$.
9. Pruébese que el lema 3.2 es falso si prescindimos de suponer que el dominio entero es finito.
10. Pruébese que todo campo es un dominio entero.
11. Usando el principio de la casilla pruébese que si m y n son enteros primos relativos y a y b enteros cualesquiera, entonces existe un entero x tal que $x \equiv a$ mód m y $x \equiv b$ mód n . (*Sugerencia:* considérense los residuos de $a, a+m, a+2m, \dots, a+(n-1)m$ en la división por n .)
12. Usando el principio de la casilla pruébese que la expansión decimal de un número racional debe, después de cierto punto, hacerse periódica.

3) HOMOMORFISMOS

Al estudiar los grupos vimos que el concepto de homomorfismo resultaba ciertamente fructífero. Esto parece sugerir que apropiadamente un análogo para anillos nos llevaría también hasta importantes ideas. Recuérdese que para los grupos un homomorfismo se definió como una aplicación tal que $\phi(ab) = \phi(a)\phi(b)$. Como un anillo tiene dos operaciones, ¿qué podría ser una extensión más natural de este tipo de fórmula que la que se presenta en la siguiente definición?

DEFINICIÓN. Una aplicación ϕ del anillo R en el anillo R' se dice que es un *homomorfismo* si

- 1) $\phi(a+b) = \phi(a)+\phi(b)$,
- 2) $\phi(ab) = \phi(a)\phi(b)$,

para $a, b \in R$ cualesquiera.

Como en el caso de los grupos, hagamos también aquí hincapié en que el $+$ y el \cdot que aparecen en los primeros miembros de las relaciones en (1) y (2) son los de R mientras que el $+$ y el \cdot que aparecen en los segundos miembros son los de R' .

Una útil observación es la de que un homomorfismo de un anillo R en un anillo R' es, si ignoramos totalmente la multiplicación en ambos anillos, al menos un homomorfismo de R en R' cuando los consideramos como grupos abelianos bajo las respectivas adiciones. Por tanto, en cuanto a la adición concierne, todas las propiedades acerca de los homomorfismos de grupos probadas en el capítulo 2 se verifican también aquí. En particular, la mera reformulación del lema 2.14 para el caso del grupo aditivo de un anillo nos da

LEMA 3.3. *Si ϕ es un homomorfismo de R en R' entonces*

- 1) $\phi(0) = 0$.
- 2) $\phi(-a) = -\phi(a)$ para todo $a \in R$.

Unas palabras de advertencia: si tanto R como R' tienen elementos unidad 1 y $1'$ respectivamente para sus multiplicaciones, no se sigue de ello necesariamente que $\phi(1) = 1'$. Sin embargo, si R' es un dominio entero, o si R' es arbitrario, pero ϕ es suprayectivo, entonces $\phi(1) = 1'$ es necesariamente cierto.

En el caso de los grupos, dado un homomorfismo asociamos con este homomorfismo cierto subconjunto del grupo al que llamamos núcleo del homomorfismo. ¿Cuál habrá de ser la definición apropiada del núcleo de un homomorfismo entre anillos? Después de todo, los anillos tienen dos operaciones, adición y multiplicación, y podría ser natural preguntar cuál de estas debe singularizarse como base para la definición. Pero la elección es clara. Dentro de la definición de cualquier grupo arbitrario está la

condición de que el anillo forme un grupo abeliano bajo la adición. La multiplicación del anillo se dejó con muchas menos restricciones, y por ello, en cierto sentido, mucho menos bajo nuestro control que la adición. Es por esto por lo que es a la adición a la que damos énfasis especial en el anillo, y damos la siguiente definición.

DEFINICIÓN. Si ϕ es un homomorfismo de R en R' entonces el *núcleo* de ϕ , $I(\phi)$, es el conjunto de todos los elementos $a \in R$ tales que $\phi(a) = 0$, el elemento cero de R' .

LEMA 3.4. Si ϕ es un homomorfismo de R en R' con núcleo $I(\phi)$, entonces:

- 1) $I(\phi)$ es un subgrupo de R bajo la adición.
- 2) Si $a \in I(\phi)$ y $r \in R$ entonces tanto ar como ra están en $I(\phi)$.

Prueba. Como ϕ es, en particular, un homomorfismo de R , como grupo aditivo, en R' , como grupo aditivo, (1) sigue de inmediato de nuestros resultados en teoría de grupos.

Para ver (2), supongamos que $a \in I(\phi)$, $r \in R$. Entonces $\phi(a) = 0$, de modo que $\phi(ar) = \phi(a)\phi(r) = 0\phi(r) = 0$, de acuerdo con el lema 3.1. Análogamente $\phi(ra) = 0$. Luego, por la propiedad definitoria de $I(\phi)$, tanto ar como ra están en $I(\phi)$.

Antes de proseguir, examinemos estos conceptos en ciertos ejemplos.

Ejemplo 1. Sean R y R' dos anillos arbitrarios y definamos $\phi(a) = 0$ para todo $a \in R$. ϕ es trivialmente un homomorfismo y $I(\phi) = R$. A ϕ se le llama el homomorfismo cero.

Ejemplo 2. Sea R un anillo, $R' = R$ y definamos ϕ por $\phi(x) = x$ para todo $x \in R$. Claramente ϕ es un homomorfismo y $I(\phi)$ consiste solamente en el 0.

Ejemplo 3. Sea $J(\sqrt{2})$ el conjunto de todos los números reales de la forma $m+n\sqrt{2}$ donde m y n son enteros; $J(\sqrt{2})$ forma un anillo bajo la adición y la multiplicación habituales de los números reales. (Verifíquese!) Definamos $\phi: J(\sqrt{2}) \rightarrow J(\sqrt{2})$ por $\phi(m+n\sqrt{2}) = m-n\sqrt{2}$. ϕ es un homomorfismo de $J(\sqrt{2})$ sobre $J(\sqrt{2})$ y su núcleo $I(\phi)$ consiste solamente en el 0. (Verifíquese.)

Ejemplo 4. Sea J el anillo de los enteros y J_n el anillo de los enteros módulo n . Definamos $\phi: J \rightarrow J_n$ por $\phi(a) =$ residuo de la división de a por n . El lector debe verificar que ϕ es un homomorfismo de J sobre J_n y que el núcleo, $I(\phi)$, de ϕ consiste en todos los múltiplos de n .

**Ejemplo 5.* Sea R el conjunto de todas las funciones reales continuas definidas sobre el intervalo unitario cerrado. R se hace un anillo bajo las habituales adición y multiplicación de funciones; que sea un anillo es una consecuencia del hecho de que la suma y el producto de dos funciones continuas es una función continua. Sea F el anillo de los números reales y definamos $\phi: R \rightarrow F$ por $\phi(f(x)) = f(\frac{1}{2})$. ϕ , es entonces un homomorfismo de R sobre F y su núcleo consiste en todas las funciones en R que se anulan en $x = \frac{1}{2}$.

En todos los ejemplos aquí dados hemos usado anillos conmutativos. Existen muchos bellos ejemplos en que los anillos no son conmutativos, pero sería prematuro discutirlos aquí.

DEFINICIÓN. Un homomorfismo de R en R' se dice que es un *isomorfismo* si es una aplicación inyectiva.

DEFINICIÓN. Dos anillos se dice que son *isomorfos* si existe un isomorfismo de uno sobre el otro.

Las notas que presentamos en el capítulo 2 acerca del significado de un isomorfismo o de la afirmación que dos grupos son isomorfos pueden aplicarse palabra por palabra al caso de los grupos. Análogamente, el criterio dado en el lema 2.16 para que un homomorfismo sea un isomorfismo se traduce directamente de grupos a anillos en la forma

LEMA 3.5. *El homomorfismo ϕ de R en R' es un isomorfismo si y sólo si $I(\phi) = (0)$.*

4. IDEALES Y ANILLOS COCIENTE

Una vez que se han establecido las ideas de homomorfismo y su núcleo para anillos, basadas ambas en nuestra experiencia con los grupos, parece que ha de ser fructuoso establecer también para anillos algo análogo al concepto de subgrupo normal. Una vez logrado esto puede esperarse que este análogo conduzca a una construcción sobre los anillos semejante a la del grupo cociente de un grupo por un subgrupo normal. Finalmente, si alguien fuera un optimista, esperaría que los teoremas sobre homomorfismo para grupos se pudieran aplicar íntegramente a los anillos.

Afortunadamente, todo esto puede hacerse proveyéndonos con ello de una técnica incisiva para el análisis de los anillos.

La primera tarea con que nos encontramos parece que es la de definir un concepto adecuado de “subgrupo normal” para anillos. Con un poco de intuición no es esto difícil. Recordemos que los subgrupos normales resultaban no ser otra cosa en último término que núcleos de homomorfismos,

aunque en sus primeras condiciones definitorias no aparecieran los homomorfismos para nada. ¿Por qué no usar esta observación como la clave de nuestra definición para anillos? El lema 3.4 nos ha proporcionado ya algunas condiciones de las que un subconjunto de un anillo debe cumplir para que pueda ser el núcleo de un homomorfismo. Tomamos ahora el punto de vista de que ya que al menos al presente no tenemos ninguna otra información de que disponer, haremos de las conclusiones del lema 3.4 nuestro punto de partida para nuestra tarea, por lo que definiremos:

DEFINICIÓN. Un subconjunto no vacío U de R se dice que es un *ideal* (bilateral) de R si:

- 1) U es un subgrupo de R bajo la adición.
- 2) Para todo $u \in U$ y $r \in R$, tanto ur como ru están en U .

La condición (2) afirma que U “absorbe” la multiplicación a la derecha y a la izquierda por elementos arbitrarios del anillo. Por esta razón U comúnmente se llama ideal bilateral. Como no tendremos ocasión alguna, aparte de la que se nos presente en alguno de los problemas de usar ningún otro concepto de ideal, solo usaremos la palabra ideal, en lugar de ideal bilateral, en todo lo que sigue.

Dado un ideal U de un anillo R , sea R/U el conjunto de todas las distintas clases laterales de U en R que se obtienen al considerar U como un subgrupo de R respecto a la adición. Adviértase que simplemente decimos clases laterales, en lugar de clases laterales derechas o clases laterales izquierdas; esto está justificado por ser R un grupo abeliano bajo la adición. Repitamos lo que hemos dicho: R/U consiste en el conjunto de las clases laterales $a + U$ donde $a \in R$. De acuerdo con los resultados del capítulo 2, R/U es automáticamente un grupo bajo la adición, lo que se consigue por la ley de composición $(a + U) + (b + U) = (a + b) + U$. Para que R/U tenga una estructura de anillo debemos definir en él una multiplicación. ¿Qué más natural que definir $(a + U)(b + U) = ab + U$? Debemos, sin embargo, estar seguros de que tiene sentido. Dicho de otra manera, estamos obligados a probar que si $a + U = a' + U$ y $b + U = b' + U$, entonces, bajo nuestra definición de multiplicación, $(a + U)(b + U) = (a' + U)(b' + U)$, o lo que es equivalente, que $ab + U = a'b' + U$. Notemos primero a este fin que como $a + U = a' + U$, $a = a' + u_1$ donde $u_1 \in U$; análogamente, $b = b' + u_2$ con $u_2 \in U$. Luego $ab = (a' + u_1)(b' + u_2) = a'b' + u_1 b' + a' u_2 + u_1 u_2$; pero como U es un ideal de R , $u_1, b' \in U$, $a' u_2 \in U$ y $u_1 u_2 \in U$. De donde $u_1 b' + a' u_2 + u_1 u_2 = u_3 \in U$. Pero entonces $ab = a'b' + u_3$, de donde deducimos que $ab + U = a'b' + u_3 + U$, y como $u_3 \in U$, $u_3 + U = U$. La consecuencia final de todo esto es que $ab + U = a'b' + U$. Hemos al menos alcanzado el principal objetivo en nuestro camino a la meta; a saber, el de introducir una multiplicación bien definida. El resto es rutina. Para establecer que R/U es un anillo, tenemos que ir comprobando que cada uno de los axiomas que definen

un anillo se verifica en R/U . Todas estas verificaciones son bastante parecidas, así que lo que vamos a hacer es escoger uno de los axiomas, la ley distributiva derecha, y probar que se verifica en R/U . El resto lo dejamos al lector como ejercicio informal. Si $X = a + U$, $Y = b + U$ y $Z = c + U$ son tres elementos de R/U , donde $a, b, c \in R$, entonces $(X+Y)Z = ((a+U)+(b+U))(c+U) = ((a+b)+U)(c+U) = (a+b)c+U = ac+bc+U = (ac+U)+(bc+U) = (a+U)(c+U)+(b+U)(c+U) = XZ+YZ$.

R/U se ha convertido ahora en anillo. Es claro que si R es conmutativo, entonces también lo es R/U , pues $(a+U)(b+U) = ab+U = ba+U = (b+U)(a+U)$. (El recíproco es falso.) Si R tiene un elemento unidad 1, entonces R/U tiene un elemento unidad $1+U$. Podemos preguntarnos: ¿en qué relación está R/U con R ? Con la experiencia que ahora tenemos nos es fácil contestar. Hay un homomorfismo ϕ de R sobre R/U dado por $\phi(a) = a+U$ para todo $a \in R$, cuyo núcleo es exactamente U . El lector debe verificar que el ϕ así definido es un homomorfismo de R sobre R/U de núcleo U .

Resumimos estas observaciones en el siguiente lema.

LEMA 3.6. *Si U es un ideal del anillo R , entonces R/U es un anillo y es una imagen homomorfa de R .*

Con esta construcción del *anillo cociente* de un anillo por un ideal, satisfactoriamente lograda, estamos listos para traer a los anillos los teoremas de homomorfismo de los grupos. Como la prueba es una aplicación palabra por palabra de la que se dio para los grupos al lenguaje de los anillos, nos limitamos a enunciar el teorema sin prueba, refiriendo al lector al capítulo 2 para la prueba.

TEOREMA 3.A. *Sean R y R' anillos y ϕ un homomorfismo de R sobre R' de núcleo U . Entonces R' es isomorfo a R/U . Además hay una correspondencia biyectiva entre el conjunto de ideales de R' y el conjunto de ideales de R que contienen a U . Esta correspondencia puede obtenerse asociando a cada ideal W' en R' el ideal W de R definido por $W' = \{x \in R \mid \phi(x) \in W'\}$. Con W así definido R/W es isomorfo a R'/W' .*

Problemas

1. Si U es un ideal de R y $1 \in U$ pruébese que $U = R$.
2. Si F es un campo pruébese que sus únicos ideales son (0) y el mismo F .
3. Pruébese que cualquier homomorfismo de un campo es o un isomorfismo o lleva todos los elementos en el 0 .

4. Si R es un anillo comutativo y $a \in R$ pruébese que:

a) $aR = \{ar \mid r \in R\}$ es un ideal bilateral de R .

b) Pruébese, mediante un ejemplo que esto puede ser falso si R no es comutativo.

5. Si U, V son ideales de R sea $U+V = \{u+v \mid u \in U, v \in V\}$. Pruébese que $U+V$ es también un ideal.

6. Si U y V son ideales de R sea UV el conjunto de todos los elementos que pueden escribirse como sumas finitas de elementos de la forma uv donde $u \in U$ y $v \in V$. Pruébese que UV es un ideal de R .

7. En el problema 6 pruébese que $UV \subset U \cap V$.

8. Si R es el anillo de los enteros, sea U el ideal consistente en todos los múltiplos de 17. Pruébese que si V es un ideal de R y $R \supset V \supset U$, entonces $o \ V = R$ o $V = U$. Generalícese el resultado.

9. Si U es un ideal de R sea $r(U) = \{x \in R \mid xu = 0 \text{ para todo } u \in U\}$. Pruébese que $r(U)$ es un ideal de R .

10. Si U es un ideal de R sea $[R:U] = \{x \in R \mid rx \in U \text{ para todo } r \in R\}$. Pruébese que $[R:U]$ es un ideal de R y que contiene a U .

11. Sea R un anillo con elemento unidad. Usando sus elementos definimos un anillo \tilde{R} conviniendo en que $a \oplus b = a+b+1$, y $a \cdot b = ab+a+b$ donde $a, b \in R$ y la adición y la multiplicación en los segundos miembros de estas relaciones son las de R .

a) Pruébese que \tilde{R} es un anillo bajo las operaciones \oplus y \cdot .

b) ¿Qué elemento es el que actúa como cero de \tilde{R} ?

c) ¿Qué elemento es el que actúa como el uno de \tilde{R} ?

d) Pruébese que R es isomorfo a \tilde{R} .

*12. En el ejemplo 6 de la sección 1 de este capítulo, discutimos el anillo de las matrices racionales 2×2 . Pruébese que este anillo no tiene ideales distintos al (0) y a sí mismo.

*13. En el ejemplo 8 de la sección 1 de este capítulo, discutimos los cuaternios reales. Usando esto como modelo definimos los cuaternios sobre los enteros módulo p , p un número primo impar, exactamente del mismo modo, sin embargo, ahora consideramos todos los símbolos de la forma $\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$ donde $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ son enteros módulo p .

a) Pruébese que este es un anillo con p^4 elementos cuyos únicos ideales son (0) y el mismo anillo.

**b) Pruébese que este anillo *no es* un anillo con división.

Si R es un anillo cualquiera, un subconjunto L de R se llama *ideal izquierdo* de R si

1) L es un subgrupo de R respecto a la adición.

2) $r \in R$ y $a \in L$ implica $ra \in L$.

(Análogamente se puede definir un *ideal derecho*.)

Un ideal es, simultáneamente, un ideal izquierdo y un ideal derecho de R .

14. Para $a \in R$ sea $Ra = \{xa \mid x \in R\}$. Pruébese que Ra es un ideal izquierdo de R .

15. Pruébese que la intersección de dos ideales izquierdos de R es un ideal izquierdo de R .

16. ¿Qué puede decir el lector acerca de la intersección de un ideal izquierdo con un ideal derecho de R ?

17. Si R es un anillo y $a \in R$ sea $r(a) = \{x \in R \mid ax = 0\}$. Pruébese que $r(a)$ es un ideal derecho de R .

18. Si R es un anillo y L un ideal izquierdo de R sea $\lambda(L) = \{x \in R \mid xa = 0 \text{ para todo } a \in L\}$. Pruébese que $\lambda(L)$ es un ideal bilateral de R .

***19.** Sea R un anillo en el que $x^3 = x$ para todo $x \in R$. Pruébese que R es un anillo comunitativo.

20. Si R es un anillo con elemento unidad 1 y ϕ es un homomorfismo de R sobre R' , pruébese que $\phi(1)$ es el elemento unidad de R' .

21. Si R es un anillo con elemento unidad 1 y ϕ es un homomorfismo de R en un dominio entero R' tal que $I(\phi) \neq R$ pruébese que $\phi(1)$ es el elemento unidad de R' . NSHOP

5. MÁS IDEALES Y MÁS ANILLOS COCIENTE

Continuamos la discusión de los ideales y de los anillos cociente.

Supongamos, al menos por el momento, que un campo es la clase más conveniente de anillo. ¿Por qué? Si no por otra razón, al menos por la de que podemos dividir en él, de manera que tanto las operaciones como los resultados se aproximan en un campo más estrechamente a lo que es nuestra experiencia con los números reales y complejos. Además, como quedó ilustrado en el problema 2 del precedente conjunto de problemas, un campo no tiene más imágenes homomórficas que él mismo o el anillo trivial consistente en el 0. No podemos, por tanto, simplificar un campo aplicándole un homomorfismo. Teniendo en cuenta todas estas observaciones es natural que intentemos ligar de alguna forma un anillo general con campos. ¿Qué es lo que aparecerá implicado en esta liga? Tenemos una maquinaria cuyas partes componentes son los homomorfismos, los ideales y los anillos cociente. Con estas forjaremos la unión.

Pero antes debemos precisar las más muy vagas observaciones del párrafo precedente. Planteamos explícitamente nuestra pregunta: ¿Bajo qué condiciones es un campo la imagen homomorfa de un anillo? Para anillos conmutativos damos una contestación completa en esta sección.

Esencial para el tratamiento del problema planteado es el recíproco del resultado del problema 2 de la lista de problemas al final de la sección 4.

LEMÁ 3.7 *Sea R un anillo conmutativo con elemento unitario cuyos únicos ideales son (0) y el mismo R . Entonces R es un campo.*

Prueba. Para tener una prueba de este lema debemos presentar, para cada $a \neq 0 \in R$, un elemento $b \neq 0 \in R$ tal que $ab = 1$.

Supongamos pues que $a \neq 0$ está en R . Consideraremos el conjunto $Ra = \{xa \mid x \in R\}$. Afirmamos que Ra es un ideal de R . Para establecer tal cosa debemos demostrar que es un subgrupo de R bajo la adición y que si $u \in Ra$ y $r \in R$ entonces ru está también en Ra . Solo necesitamos probar que ru está en Ra porque entonces ur también está, ya que $ru = ur$.

Ahora bien, si $u, v \in Ra$, entonces $u = r_1 a, v = r_2 a$ para ciertos $r_1, r_2 \in R$. Luego $u + v = r_1 a + r_2 a = (r_1 + r_2)a \in Ra$; análogamente $-u = -r_1 a = (-r_1)a \in Ra$. De donde Ra es un subgrupo aditivo de R . Además, si $r \in R$, $ru = r(r_1 a) = (rr_1)a \in Ra$. Ra satisface, por tanto, todas las condiciones definitorias de un ideal de R , de donde es un ideal de R . Nótese que tanto la ley asociativa como la ley distributiva de la multiplicación se utilizaron en la prueba de este hecho.

Según nuestras hipótesis sobre R , $Ra = (0)$ o $Ra = R$. Como $0 \neq a = 1a \in Ra$, $Ra \neq (0)$; luego nos quedamos con la otra sola posibilidad, a saber, que $Ra = R$. Esta última ecuación nos dice que todo elemento en R es el producto de a por algún elemento de R . En particular, $1 \in R$ y, por tanto, puede realizarse como un múltiplo de a ; es decir, existe un elemento $b \in R$ tal que $ba = 1$. Y esto completa la prueba del lema.

DEFINICIÓN. Un ideal $M \neq R$ ^{de} es un anillo R se dice que es un *ideal máximo* de R si siempre que U es un ideal de R tal que $M \subset U \subset R$ se tiene que $R = U$ o $M = U$.

En otras palabras, un ideal de R es un ideal máximo si es imposible intercalar un ideal entre él y el anillo total. Dado un anillo R no hay garantía alguna de que tenga algún ideal máximo. Si el anillo tiene elemento unidad, esto puede probarse admitiendo un axioma básico de las matemáticas, el llamado axioma de selección. Puede también que haya muchos ideales maximales distintos en un anillo R ; esto lo ilustraremos más adelante en el anillo de los enteros.

Hasta el momento hemos conocido muy pocos anillos. Solamente por la consideración de un concepto dado en muchos casos particulares se puede

apreciar plenamente el concepto y sus motivaciones. Por ello, antes de proseguir, examinaremos algunos ideales máximos en dos anillos específicos. Cuando lleguemos a la discusión de los anillos de polinomios exhibiremos allí todos los ideales máximos.

Ejemplo 1. Sea R el anillo de los enteros y U un ideal de R . Como U es un subgrupo de R bajo la adición, por nuestros resultados en la teoría de grupos sabemos que U consiste en todos los múltiplos de un entero fijo n_0 ; escribimos esto como $U = (n_0)$. ¿Qué valores de n_0 dan lugar a ideales máximos?

Afirmamos primero que si p es un número primo, entonces $P = (p)$ es un ideal máximo de R . En efecto, si U es un ideal de R y $U \supset P$ entonces $U = (n_0)$ para algún entero n_0 . Como $p \in P \subset U$, $p = mn_0$ para algún entero m ; como p es un número primo, esto implica que $n_0 = 1$ o que $n_0 = p$. Si $n_0 = 1$, entonces $1 \in U$, de donde $r = 1r \in U$ para todo $r \in R$, de donde se sigue que $U = R$. Si $n_0 = p$ entonces $P \subset U = (n_0) \subset P$, de donde $U = P$. De donde ningún ideal, aparte de R y P , puede ponerse entre P y R , de donde deducimos que P es máximo.

Supongamos, por otra parte, que $M = (n_0)$ es un ideal máximo de R . Afirmamos que n_0 debe ser un número primo, pues si $n_0 = ab$, donde a y b son enteros positivos, entonces $U = (a) \supset M$, de donde $U = R$ o $U = M$. Si $U = R$ entonces $a = 1$, según puede verse fácilmente; si $U = M$, entonces $a \in M$ y, por tanto, $a = rn_0$ para algún entero r , ya que todo elemento de M es un múltiplo de n_0 . Pero entonces $n_0 = ab = rn_0 b$, de donde se tiene que $rb = 1$, luego $b = 1$ y $n_0 = a$. Luego n_0 es un número primo.

En este ejemplo particular, la noción de ideal máximo se hace viva —corresponde exactamente a la noción de número primo. No se debe, sin embargo, saltar apresuradamente a ninguna generalización; esta clase de correspondencia no se verifica habitualmente para anillos más generales.

Ejemplo 2. Sea R el anillo de todas las funciones reales continuas sobre el intervalo unitario cerrado. (Véase ejemplo 5, sección 3.) Sea

$$M = \{f(x) \in R \mid f\left(\frac{1}{2}\right) = 0\}$$

M es ciertamente un ideal de R . Además, es un ideal máximo de R , pues si un ideal U contiene a M y $M \neq R$, entonces hay una función $g(x) \in U$, $g(x) \notin M$. Como $g(x) \notin M$, $g\left(\frac{1}{2}\right) = \alpha \neq 0$. Entonces $h(x) = g(x) - \alpha$ es tal que $h\left(\frac{1}{2}\right) = g\left(\frac{1}{2}\right) - \alpha = 0$, luego $h(x) \in M \subset U$. Pero $g(x)$ está también en U ; luego $\alpha = g(x) - h(x) \in U$ y por tanto $1 = \alpha\alpha^{-1} \in U$. Luego para cualquier función $t(x) \in R$, $t(x) = 1t(x) \in U$, por tanto, $U = R$. M es de esta manera un ideal máximo de R . Análogamente si y es un número real tal que $0 \leq y \leq 1$, entonces $M_y = \{f(x) \in R \mid f(y) = 0\}$ es un ideal máximo de R . Se puede mostrar (véase el problema 4 al final de esta sección) que todo ideal máximo es de esta forma. Así pues, aquí los ideales máximos se corresponden con los puntos del intervalo unitario.

Después de haber visto algunos ideales máximos en algunos anillos concretos estamos listos para continuar el desarrollo general con el siguiente.

TEOREMA 3.B. *Si R es un anillo conmutativo con elemento unidad y M es un ideal de R , entonces M es un ideal máximo de R si y sólo si R/M es un campo.*

Prueba. Supongamos primero que M es un ideal de R tal que R/M es un campo. Como R/M es un campo sus solos ideales son (0) y R/M mismo. Pero por el teorema 3.a hay una correspondencia inyectiva entre el conjunto de ideales de R/M y el conjunto de ideales de R que contienen a M . El ideal M de R se corresponde con el ideal (0) de R/M mientras que el ideal R de R se corresponde con el ideal R/M de R/M en esta aplicación inyectiva. Luego no hay ningún ideal, entre M y R aparte de estos dos. De donde M es un ideal máximo.

Por otra parte, si M es un ideal máximo de R , por la correspondencia arriba mencionada R/M tiene solamente a (0) y a sí mismo como ideales. Por otra parte R/M es conmutativo y tiene un elemento unidad ya que R tiene estas dos propiedades. Todas las condiciones del lema 3.7 se cumplen para R/M de forma que podemos concluir, de acuerdo con lo afirmado en ese lema, que R/M es un campo.

Tendremos muchas ocasiones de hacer referencia a este resultado en nuestro estudio de los anillos de polinomios y en la teoría de extensiones de campos.

Problemas

1. Sea R un anillo con elemento unitario, no necesariamente conmutativo, tal que los únicos ideales derechos de R son (0) y R . Pruébese que R es un anillo con división.

*2. Sea R un anillo tal que los únicos ideales derechos de R sean (0) y R . Pruébese que o R es un anillo con división o R es un anillo con un número primo de elementos en el que $ab = 0$ para cualesquiera $a, b \in R$.

3. Sea J el anillo de los enteros, p un número primo, y (p) el ideal de J consistente en todos los múltiplos de p . Pruébese que:

- a) $J/(p)$ es isomorfo a J_p , el anillo de los enteros módulo p :
- b) Usando el teorema 3.b y la parte (a) de este problema, demuéstrese que J_p es un campo.

**4. Sea R el anillo de todas las funciones continuas reales sobre el intervalo unitario cerrado. Si M es un ideal máximo de R pruébese que existe un número real γ , $0 \leq \gamma \leq 1$, tal que $M = M_\gamma = \{f(x) \in R \mid f(\gamma) = 0\}$.

6. EL CAMPO DE COCIENTES DE UN DOMINIO ENTERO

Recordemos que un dominio entero es un anillo commutativo D con la propiedad adicional de que no tiene divisores de cero, es decir, que si $ab = 0$ para algunos $a, b \in D$, entonces, al menos uno de los dos a o b debe ser 0. El anillo de los enteros es, desde luego, un ejemplo de dominio entero.

El anillo de los enteros tiene la atractiva propiedad de que podemos extenderlo al conjunto de los números racionales, que es un campo. ¿Podemos efectuar una construcción análoga en cualquier dominio entero? Vamos a ver, en lo que sigue, que realmente así es.

DEFINICIÓN. Un anillo R *puede sumergirse* en un anillo R' si hay un isomorfismo de R en R' . (Si R y R' tienen elementos unidad 1 y $1'$ exigimos, además, que este isomorfismo lleve 1 en $1'$.)

A R' le llamaremos sobre anillo o *extensión* de R si R puede sumergirse en R' .

Con esta definición de inmersión, probamos el

TEOREMA 3.C. *Todo dominio entero puede sumergirse en un campo.*

Prueba. Antes de hacer explícitos los detalles de la prueba, enfoquemos informalmente el problema. Sea D nuestro dominio entero; en términos vagos, el campo que buscamos debe ser el de todos los cocientes $\frac{a}{b}$ donde $a, b \in D$ y $b \neq 0$. Desde luego, en D es bien posible que $\frac{a}{b}$ carezca de sentido.

¿Qué es lo que requeriremos de estos símbolos $\frac{a}{b}$? Claramente debemos tener una respuesta a las siguientes tres preguntas:

1) ¿Cuándo es $\frac{a}{b} = \frac{c}{d}$?

2) ¿Qué es $\frac{a}{b} + \frac{c}{d}$?

3) ¿Qué es $\frac{a}{b} \frac{c}{d}$?

En la contestación de (1), ¿qué más natural que decir que $\frac{a}{b} = \frac{c}{d}$ si y sólo si $ad = bc$? En cuanto a (2) y (3), ¿por qué no ensayar lo obvio que es definir

$$\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd} \quad y \quad \frac{a}{b} \frac{c}{d} = \frac{ac}{bd}$$

En realidad, en lo que sigue, hacemos de estas consideraciones nuestra guía. Dejemos ahora la heurística y entremos en el dominio de las matemáticas, con definiciones precisas y deducciones rigurosas.

Sea \mathfrak{M} el conjunto de todos los pares ordenados (a, b) con $a, b \in D$ y $b \neq 0$. (Piénsese en (a, b) como a/b .) Definimos ahora en \mathfrak{M} la siguiente relación:

$$(a, b) \sim (c, d) \text{ si y sólo si } ad = bc.$$

Afirmamos que hemos definido con ello una relación de equivalencia sobre \mathfrak{M} . Para comprobarlo, probamos el cumplimiento de las tres condiciones definitorias de una relación de equivalencia en esta relación particular.

- 1) Si $(a, b) \in \mathfrak{M}$ entonces $(a, b) \sim (a, b)$, ya que $ab = ba$.
- 2) Si $(a, b), (c, d) \in \mathfrak{M}$ y $(a, b) \sim (c, d)$ entonces $ad = bc$, de donde $cb = da$ y, por tanto, $(c, d) \sim (a, b)$.
- 3) Si $(a, b), (c, d)$ y (e, f) están todos en \mathfrak{M} y $(a, b) \sim (c, d)$ y $(c, d) \sim (e, f)$, entonces $ad = bc$ y $cf = de$. Luego $bcf = bde$, y como $bc = ad$, de ello se sigue que $adf = bde$. Como D es conmutativo, esta relación se convierte en la $adf = bed$; como, además, D es un dominio entero y $d \neq 0$ esta relación implica, a su vez, que $af = be$. Pero entonces $(a, b) \sim (e, f)$ y nuestra relación es transitiva.

Sea $[a, b]$ la clase de equivalencia en \mathfrak{M} de (a, b) , y sea F el conjunto de todas las clases de equivalencia $[a, b]$ donde $a, b \in D$ y $b \neq 0$. F es el candidato para el campo que estamos buscando. Para hacer de F un campo debemos introducir una adición y una multiplicación para sus elementos y probar después que bajo estas operaciones F es un campo.

Tratemos primero la adición. Guiados por nuestra discusión heurística al comienzo de la prueba, definimos:

$$[a, b] + [c, d] = [ad + bc, bd].$$

Como D es un dominio entero y ambos b y d son distintos de cero, $b \neq 0$ y $d \neq 0$, tenemos también $bd \neq 0$; esto, al menos, nos dice que $[ad + bc, bd] \in F$. Afirmamos ahora que esta adición está bien definida, es decir, que si $[a, b] = [a', b']$ y $[c, d] = [c', d']$ entonces $[a, b] + [c, d] = [a', b'] + [c', d']$. Comprobémoslo. De $[a, b] = [a', b']$ se tiene que $ab' = ba'$; de $[c, d] = [c', d']$ se tiene que $cd' = dc'$. Lo que necesitamos es que estas relaciones tengan como consecuencia obligada la igualdad de $[a, b] + [c, d]$ y $[a', b'] + [c', d']$. Según la definición de adición, esto se reduce a probar que $[ad + bc, bd] = [a'd' + b'c', b'd']$, o lo que es equivalente a probar que $(ad + bc)b'd' = bd(a'd' + b'c')$. Usando $ab' = ba'$, $cd' = dc'$, esto toma la forma: $(ad + bc)b'd' = abb'd' + bcb'd' = ab'dd' + bb'cd' = ba'dd' + bb'dc' = bd(a'd' + b'c')$, que es la igualdad deseada.

Claramente $[0, b]$ actúa como un elemento cero para esta adición y $[-a, b]$ como el negativo de $[a, b]$. Es sencillo mostrar que F es un grupo abeliano bajo esta adición.

Vayamos ahora con la multiplicación en F . Guiados de nuevo por nuestra discusión heurística preliminar definimos $[a, b][c, d] = [ac, bd]$. Como en el caso de la adición, como $b \neq 0$ y $d \neq 0$, $bd \neq 0$ y, por tanto, $[ac, bd] \in F$. Un cálculo de tipo bastante parecido al que acabamos de efectuar prueba que si $[a, b] = [a', b']$ y $[c, d] = [c', d']$, entonces $[a, b][c, d] = [a', b'][c', d']$. Ahora se puede probar que los elementos distintos de cero de F (es decir, todos los elementos $[a, b]$ donde $a \neq 0$) forman un grupo abeliano bajo la multiplicación en el que $[d, d]$ actúa como el elemento unidad y donde

$$[c, d]^{-1} = [d, \bar{d}] \quad (\text{ya que, como } c \neq 0, [d, c] \text{ está en } F).$$

Con un cálculo rutinario se puede comprobar que la ley distributiva se cumple en F . F es, pues, un campo.

Todo lo que queda es mostrar que D puede sumergirse en F . Exhibiremos explícitamente un isomorfismo de D en F . Observemos primero, antes de hacer esto, que para $x \neq 0$ y $y \neq 0$ en D , $[ax, x] = [ay, y]$ puesto que $(ax)y = x(ay)$; denotemos a $[ax, x]$ por $[a, 1]$. Definamos $\phi: D \rightarrow F$ por $\phi(a) = [a, 1]$ para todo $a \in D$. Dejamos al lector la verificación de que ϕ es un isomorfismo de D en F , y de que si D tiene un elemento 1, entonces $\phi(1)$ es el elemento unidad de F . El teorema queda ahora probado en su totalidad.

A F se le suele llamar el *campo de cocientes* de D . En el caso especial en que D es el anillo de los enteros, la F así construida es, desde luego, el campo de los números racionales.

Problemas

1. Pruébese que si $[a, b] = [a', b']$ y $[c, d] = [c', d']$ entonces $[a, b][c, d] = [a', b'][c', d']$.

2. Pruébese la ley distributiva en F .

3. Pruébese que la aplicación $\phi: D \rightarrow F$ definida por $\phi(a) = [a, 1]$ es un isomorfismo de D en F .

4. Pruébese que si K es un campo cualquiera que contiene a D , entonces K contiene un subcampo isomorfo a F . En este sentido F es el campo mínimo que contiene a D .

*5. Sea R un anillo commutativo con elemento unidad. Un subconjunto no vacío S de R se llama sistema multiplicativo si:

1) $0 \notin S$.

2) $s_1, s_2 \in S$ implica $s_1 s_2 \in S$.

Sea \mathfrak{M} el conjunto de todos los pares ordenados (r, s) tales que $r \in R$ y $s \in S$. Definamos en \mathfrak{M} $(r, s) \sim (r', s')$ si existe un elemento $s'' \in S$ tal que

$$s''(rs' - sr') = 0.$$

a) Pruébese que así se ha definido una relación de equivalencia sobre \mathfrak{M} .

Denotemos la clase de equivalencia de (r, s) por $[r, s]$ y por R_s el conjunto de todas las clases de equivalencia. Definamos en R_s $[r_1, s_1] + [r_2, s_2] = [r_1 s_2 + r_2 s_1, s_1 s_2]$ y $[r_1, s_1] \cdot [r_2, s_2] = [r_1 r_2, s_1 s_2]$.

b) Pruébese que la adición y la multiplicación que acabamos de describir están bien definidas y que R_s forma un anillo bajo estas operaciones.

c) ¿Puede sumergirse R en R_s ?

d) Pruébese que la aplicación $\phi : R \rightarrow R_s$, definida por $\phi(a) = [as, s]$ es un homomorfismo de R en R_s y encuéntrese el núcleo de ϕ .

e) Pruébese que este núcleo no tiene ningún elemento de S en él.

f) Pruébese que todo elemento de la forma $[s_1, s_2]$ (donde $s_1, s_2 \in S$) en R_s tiene un inverso en R_s .

6. Sea D un dominio entero y $a, b \in D$. Supongamos que $a^n = b^n$ y que $a^m = b^m$ para m y n enteros primos relativos. Pruébese que $a = b$.

7. ANILLOS EUCLIDIANOS

La clase de anillos que ahora nos proponemos estudiar está sugerida por varios ejemplos —el anillo de los enteros, los enteros gaussianos (sección 8) y los anillos de polinomios (sección 9). La definición de esta clase está diseñada para incorporar en ella ciertas características sobresalientes de los tres ejemplos concretos que acabamos de enumerar.

DEFINICIÓN. Un dominio entero R se dice que es un *anillo euclíadiano* si para todo $a \neq 0$ en R está definido un entero no negativo $d(a)$ tal que:

- 1) Para cualesquiera $a, b \in R$, ambos distintos de cero, $d(a) \leq d(ab)$.
- 2) Para cualesquiera $a, b \in R$, ambos distintos de cero, existen $t, r \in R$ tales que $a = tb + r$, donde $r = 0$ o $d(r) < d(b)$.

No asignamos valor alguno a $d(0)$. Los enteros sirven como un ejemplo de anillo euclíadiano, donde $d(a) =$ valor absoluto de a actúa como la función que la definición requiere. En la próxima sección veremos que los enteros gaussianos también forman un anillo euclíadiano. Aparte de tal observación y los resultados que desarrollamos en esta parte, probaremos un teorema clásico en la teoría de los números debido a Fermat, a saber, que todo número primo de la forma $4n+1$ puede escribirse como la suma de dos cuadrados.

Comenzamos con el

TEOREMA 3.D. *Sea R un anillo euclíadiano y A un ideal de R . Entonces existe un elemento $a_0 \in A$ tal que A consiste exactamente en todos los a_0x para $x \in R$ cualquiera.*

Prueba. Si A consiste solamente en el elemento 0, para $a_0 = 0$ la afirmación del teorema se verifica.

Podemos suponer que $A \neq (0)$, luego que hay un $a \neq 0$ en A . Escójase un $a_0 \in A$ tal que $d(a_0)$ sea mínimo. (Como d toma solo valores enteros no negativos, esto es siempre posible.)

Supongamos que $a \in A$. Por las propiedades de los anillos euclidianos existen $t, r \in R$ tales que $a = ta_0 + r$ donde $r = 0$ o $d(r) < d(a_0)$. Como $a_0 \in A$ y A es un ideal de R , ta_0 está en A . Si combinamos esto con que $a \in A$ tenemos que $a - ta_0 \in A$; pero $r = a - ta_0$, de donde $r \in A$. Si $r \neq 0$, entonces $d(r) < d(a_0)$ dándonos un elemento r en A cuyo d -valor es menor que el de a_0 , en contradicción con nuestra elección de a_0 como elemento de A de d -valor mínimo. Por consiguiente $r = 0$ y $a = ta_0$, lo que prueba el teorema.

Introducimos la notación $(a) = \{xa | x \in R\}$ para representar al ideal de todos los múltiplos de a .

DEFINICIÓN. Un dominio entero R con elemento unidad es un *anillo de ideales principales* si todo ideal A en R es de la forma $A = (a)$ para algún $a \in R$.

Una vez que establezcamos que un anillo euclíadiano tiene un elemento unitario, en virtud del teorema 3.d, sabremos que un anillo euclíadiano es un anillo de ideales principales. Lo recíproco, sin embargo, es falso; hay anillos de ideales principales que no son anillos euclidianos. (Véase el artículo de T. Motzkin, *Bulletin of the American Mathematical Society*, vol. 55 [1949], páginas 1142-1146, titulado "El Algoritmo Euclíadiano".)

COROLARIO AL TEOREMA 3.D. *Un anillo euclíadiano posee un elemento unitario.*

Prueba. Sea R un anillo euclíadiano; entonces, como R es ciertamente un ideal de R , debe concluirse, de acuerdo con el teorema 3.d, que $R = (u_0)$ para algún $u_0 \in R$. Luego todo elemento de R es un múltiplo de u_0 . Tendremos, por tanto, en particular que $u_0 = u_0c$ para algún $c \in R$. Si $a \in R$, entonces $a = xu_0$ para algún $x \in R$, de donde $ac = (xu_0)c = x(u_0c) = xu_0 = a$. Luego se ve que c es el elemento unitario buscado.

DEFINICIÓN. Si $a \neq 0$ y b están en un anillo comutativo R , entonces a se dice que *divide a b* si existe un $c \in R$ tal que $b = ac$. Usaremos el símbolo $a|b$ para representar el hecho de que a divide a b y $a \nmid b$ para indicar que a no divide a b .

La prueba de la siguiente observación es tan simple y directa que la omitimos.

- OBSERVACIÓN. 1) Si $a|b$ y $b|c$, entonces $a|c$.
2) Si $a|b$ y $a|c$, entonces $a|(b \pm c)$.
3) Si $a|b$ entonces $a|bx$ para todo $x \in R$.

DEFINICIÓN. Si $a, b \in R$ entonces $d \in R$ se dice que es un *máximo común divisor* de a y b si:

- 1) $d|a$ y $d|b$.
2) Siempre que $c|a$ y $c|b$, entonces $c|d$.

Usaremos la notación $d = (a, b)$ para indicar que d es un máximo común divisor de a y b .

LEMA 3.8. Sea R un anillo euclíadiano. Entonces cualesquiera dos elementos a y b en R tienen un máximo común divisor d . Además $d = \lambda a + \mu b$ para algunos $\lambda, \mu \in R$.

Prueba. Sea A el conjunto de todos los elementos $ra + sb$ donde r y s son elementos cualesquiera de R . Afirmamos que A es un ideal de R . Supongamos en efecto que $x, y \in A$; entonces $x = r_1 a + s_1 b$, $y = r_2 a + s_2 b$ y, por tanto, $x \pm y = (r_1 \pm r_2)a + (s_1 \pm s_2)b \in A$. Análogamente, para $u \in R$, $ux = u(r_1 a + s_1 b) = (ur_1)a + (us_1)b \in A$.

Como A es un ideal de R , según el teorema 3.d existe un elemento $d \in A$ tal que todo elemento de A es un múltiplo de d . Por el hecho de ser d un elemento de A y de ser todos los elementos de A de la forma $ra + sb$, $d = \lambda a + \mu b$ para ciertos $\lambda, \mu \in R$. Ahora bien, de acuerdo con el corolario del teorema 3.d, R tiene un elemento unidad 1, luego $a = 1a + 0b \in A$ y $b = 0a + 1b \in A$. Luego por estar en A , ambos son múltiplos de d de donde $d|a$ y $d|b$.

Supongamos, finalmente, que $c|a$ y $c|b$; entonces $c|\lambda a$ y $c|\mu b$ de modo que c ciertamente divide a $\lambda a + \mu b = d$. Por lo tanto, d tiene todas las condiciones requeridas a un máximo común divisor y el lema queda probado.

DEFINICIÓN. Sea R un anillo comutativo con elemento unidad. Un elemento $a \in R$ es una *unidad* en R si existe un elemento $b \in R$ tal que $ab = 1$.

¡No se confunda una unidad con un elemento unitario! Una unidad en un anillo es un elemento cuyo inverso está también en el anillo.

LEMA 3.9. Sea R un dominio entero con elemento unitario y supongamos que para $a, b \in R$ se tiene, simultáneamente, que $a|b$ y que $b|a$. Entonces $a = ub$ donde u es una unidad en R .

Prueba. Como $a|b$, $b = xa$ para algún $x \in R$; como $b|a$, $a = yb$ para algún $y \in R$. Luego $b = x(yb) = (xy)b$; pero todos estos son elementos de un dominio entero, luego podemos cancelar la b y obtenemos $xy = 1$; y es, pues, una unidad en R y $a = yb$, lo que prueba el lema.

DEFINICIÓN. Sea R un anillo conmutativo con elemento unitario. Dos elementos a y b de R se dice que son *asociados* si $b = ua$ para alguna unidad u de R .

La relación de ser asociados es una relación de equivalencia. (Problema 1 al final de esta sección.) Nótese que en un anillo euclíadiano cualesquiera dos máximos comunes divisores de dos elementos dados son asociados (problema 2).

Hasta el momento no hemos hecho uso de la condición (1) en la definición de un anillo euclíadiano, a saber, que $d(a) \leq d(ab)$ para $b \neq 0$. En la siguiente prueba haremos uso de ella.

LEMA 3.10. *Sea R un anillo euclíadiano y $a, b \in R$. Si b no es una unidad en R entonces $d(a) < d(ab)$.**

Prueba. Consideremos el ideal $A = (a) = \{xa \mid x \in R\}$ de R . Por la condición (1) para anillos euclidianos, $d(a) \leq d(xa)$ para $x \neq 0$ en R . Luego el d -valor de a es el mínimo de los d -valores de elementos de A . Ahora bien, $ab \in A$; si $d(ab) = d(a)$, de acuerdo con la prueba que usamos para demostrar el teorema 3.d, como el d -valor de ab es mínimo en A , todo elemento de A es un múltiplo de ab . En particular, como $a \in A$, a debe ser un múltiplo de ab ; de donde $a = abx$ para algún $x \in R$. Como todo esto está teniendo lugar en un dominio entero, de ello se deduce que $bx = 1$. Luego b es una unidad en R , en contradicción al hecho de que no es una unidad. El resultado neto de todo esto es que $d(a) < d(ab)$.

DEFINICIÓN. En el anillo euclíadiano R un elemento π que no sea una unidad se dice que es un *elemento primo* de R si siempre que $\pi = ab$ donde a y b están en R , se tiene que uno de los dos a o b es una unidad en R .

Un elemento primo es un elemento en R que no puede ser factorizado en R en forma que no sea trivial.

LEMA 3.11. *Sea R un anillo euclíadiano. Entonces todo elemento en R o es una unidad en R o puede escribirse como el producto de un número finito de elementos primos de R .*

*Se está suponiendo, además, que tanto a como b no son cero. (N. del T.)

Prueba. La prueba es por inducción sobre $d(a)$.

Si $d(a) = d(1)$ entonces a es una unidad en R (problema 3), y por tanto, en este caso, la aserción del lema es correcta.

Suponemos que el lema es cierto para todos los elementos x en R tales que $d(x) < d(a)$. Con base en esta hipótesis intentamos probarlo para a . Esto completaría la inducción y probaría el lema.

Si a es un elemento primo de R no hay nada que probar. Supongamos, pues, que $a = bc$ donde ni b ni c son unidades de R . Según el lema 3.10, $d(b) < d(bc) = d(a)$ y $d(c) < d(bc) = d(a)$. Luego, por nuestra hipótesis de inducción, b y c pueden escribirse como un producto de un número finito de elementos primos de R ; $b = \pi_1 \pi_2 \dots \pi_n$, $c = \pi'_1 \pi'_2 \dots \pi'_m$ donde los π y los π' son elementos primos de R . Por consiguiente, $a = bc = \pi_1 \pi_2 \dots \pi_n \pi'_1 \pi'_2 \dots \pi'_m$ y de esta forma hemos factorizado a a como un producto de un número finito de elementos primos. Y esto completa la prueba.

DEFINICIÓN. En el anillo euclíadiano R , a y b en R se dice que son *primos relativos* si su máximo común divisor es una unidad de R .

Como cualquier asociado de un máximo común divisor es un máximo común divisor, y como 1 es un asociado de cualquier unidad, si a y b son primos relativos podemos suponer que $(a, b) = 1$.

LEMA 3.12. *Sea R un anillo euclíadiano. Supongamos que para $a, b, c \in R$, $a|bc$ pero $(a, b) = 1$. Entonces $a|c$.*

Prueba. Como hemos visto en el lema 3.8, el máximo común divisor de a y b puede realizarse en la forma $\lambda a + \mu b$. Luego, por nuestra hipótesis, $\lambda a + \mu b = 1$. Multiplicando esta relación por c obtenemos $\lambda ac + \mu bc = c$. Pero $a|\lambda ac$ siempre y $a|\mu bc$ ya que $a|bc$ por hipótesis; luego $a|(\lambda ac + \mu bc) = c$. Que es lo que afirma el lema.

Deseamos demostrar que los elementos primos en un anillo euclíadiano juegan el mismo papel que los números primos en los enteros. Si π en R es un elemento primo de R y $a \in R$ entonces $\pi|a$ o $(\pi, a) = 1$, pues, en particular, (π, a) es un divisor de π de donde debe ser π o 1 (o cualquier unidad). Si $(\pi, a) = 1$, la mitad de nuestra afirmación está probada; si $(\pi, a) = \pi$, como $(\pi, a)|a$ tenemos que $\pi|a$, y la otra mitad de nuestra afirmación es cierta.

LEMA 3.13. *Si π es un elemento primo en el anillo euclíadiano R y $\pi|ab$ donde $a, b \in R$, entonces π divide al menos a uno de los dos elementos a o b .*

Prueba. Supongamos que π no divide a a ; entonces $(\pi, a) = 1$. Aplicando ahora el lema 3.12 tenemos que $\pi|b$.

COROLARIO. Si π es un elemento primo en el anillo euclíadiano R y $\pi | a_1 a_2 \dots a_n$, entonces π divide al menos a uno de los factores a_1, a_2, \dots, a_n .

Llevemos un poco más adelante la analogía entre elementos primos y números primos y probemos que

TEOREMA 3.E. (TEOREMA DE LA FACTORIZACIÓN ÚNICA). Sea R un anillo euclíadiano y $a \neq 0$ un elemento de R que no es una unidad. Supongamos que $a = \pi_1 \pi_2 \dots \pi_n = \pi'_1 \pi'_2 \dots \pi'_m$ donde los π_i y los π'_j son elementos primos de R . Entonces $n = m$ y cada π_i , $1 \leq i \leq n$ es un asociado de algún π'_j , $1 \leq j \leq m$ y reciprocamente, cada π'_k es un asociado de algún π_q .

Prueba. Fijémonos en la relación $a = \pi_1 \pi_2 \dots \pi_n = \pi'_1 \pi'_2 \dots \pi'_m$, de donde $\pi_1 | \pi'_1 \pi'_2 \dots \pi'_m$. Por el lema 3.13 π_1 debe dividir a algún π'_j ; como π_1 y π'_j son ambos elementos primos de R y $\pi_1 | \pi'_j$, ambos elementos deben ser asociados y $\pi'_j = u_1 \pi_1$ donde u_1 es una unidad de R . Tenemos, pues, $\pi_1 \pi_2 \dots \pi_n = \pi'_1 \pi'_2 \dots \pi'_m = u_1 \pi'_2 \dots \pi'_{i-1} \pi'_{i+1} \dots \pi'_m$; cancelamos π_1 y queda $\pi_2 \dots \pi_n = u_1 \pi'_2 \dots \pi'_{i-1} \pi'_{i+1} \dots \pi'_m$. Repitiendo el razonamiento sobre esta relación con π_2 , y así sucesivamente, después de n pasos el primer miembro se hace 1, y el segundo un producto de un cierto número de π' (el exceso de m sobre n). Esto obligaría a que $n \leq m$ ya que las π' no son unidades. Análogamente, $m \leq n$, de modo que $n = m$. Y a lo largo del proceso demostrativo hemos probado también que cada uno de los π_i tiene algún π'_j como asociado y reciprocamente.

Al combinar el lema 3.11 y el teorema 3.e tenemos que todo elemento distinto de cero en un anillo euclíadiano R puede ser escrito en forma única (salvo asociados) como un producto de elementos primos o es una unidad en R .

Terminamos la sección determinando todos los ideales máximos en un anillo euclíadiano.

En el teorema 3.d probamos que cualquier ideal A en el anillo euclíadiano R es de la forma $A = (a_0)$ donde $(a_0) = \{xa_0 | x \in R\}$. Preguntamos ahora: ¿cuáles son las condiciones que impuestas sobre a_0 aseguran que A es un ideal máximo de R ? Para esta pregunta tenemos una contestación sencilla y precisa, a saber:

LEMA 3.14. El ideal $A = (a_0)$ es un ideal máximo del anillo euclíadiano R si y sólo si a_0 es un elemento primo de R .

Prueba. Probamos primero que si a_0 no es un elemento primo, entonces $A = (a_0)$ no es un ideal máximo. En efecto, supongamos que $a_0 = bc$ donde $b, c \in R$ y ni b ni c son unidades. Sea $B = (b)$; entonces ciertamente $a_0 \in B$ de modo que $A \subset B$. Afirmamos además que $A \neq B$ y que $B \neq R$.

Si $B = R$ entonces $1 \in B$ de modo que $1 = xb$ para algún $x \in R$, luego b sería una unidad en R , lo que va contra lo supuesto. Por otra parte, si

$A = B$ entonces $b \in B = A$, de donde $b = xa_0$ para algún $x \in R$. Combinado esto con $a_0 = bc$ tendríamos $a_0 = xca_0$, de donde $xc = 1$. Pero entonces c sería una unidad en contra también de lo que hemos supuesto. Por tanto ni A ni R son iguales a B , y como $A \subset B$, A no puede ser un ideal máximo de R .

Recíprocamente, supongamos que a_0 es un elemento primo de R y que U es un ideal de R tal que $A = (a_0) \subset U \subset R$. Por el teorema 3.d, $U = (u_0)$. Como $a_0 \in A \subset U = (u_0)$, $a_0 = xu_0$ para algún $x \in R$. Pero a_0 es un elemento primo de R , de lo que se sigue que o x o u_0 es una unidad de R . Si u_0 es una unidad de R entonces $U = R$ (véase el problema 5). Si, por otra parte, x es una unidad de R , entonces $x^{-1} \in R$ y la relación $a_0 = xu_0$ nos da $u_0 = x^{-1}a_0 \in A$, ya que A es un ideal de R . Esto implica que $U \subset A$; de donde, junto con $A \subset U$, se concluye que $U = A$. Por tanto, no hay ningún ideal de R que se encuentre propiamente entre A y R . Lo que nos dice que A es un ideal máximo de R .

Problemas

1. Pruébese que en un anillo comutativo con elemento unitario la relación “ a es un asociado de b ” es una relación de equivalencia.
2. Pruébese que en cualquier anillo euclíadiano dos máximos comunes divisores cualesquiera de a y b son asociados.
3. Pruébese que una condición necesaria y suficiente para que el elemento a de un anillo euclíadiano sea una unidad es que $d(a) = d(1)$.
4. Pruébese que en un anillo euclíadiano (a, b) puede encontrarse como sigue:

$$b = q_0a + r_1 \quad \text{donde } d(r_1) < d(a)$$

$$a = q_1r_1 + r_2 \quad \text{donde } d(r_2) < d(r_1)$$

$$r_1 = q_2r_2 + r_3 \quad \text{donde } d(r_3) < d(r_2)$$

•

•

•

$$r_{n-1} = q_nr_n$$

$$\text{y } r_n = (a, b).$$

5. Pruébese que si un ideal U de un anillo R contiene una unidad de R entonces $U = R$.
6. Pruébese que las unidades en un anillo comutativo con elemento unitario forman un grupo abeliano.
7. Dados dos elementos a y b en el anillo euclíadiano R , su *mínimo común múltiplo* $c \in R$ es un elemento de R tal que $a|c$, $b|c$ y, además, siempre que $a|x$ y $b|x$ con $x \in R$, entonces $c|x$. Pruébese que cualesquiera dos elementos del anillo euclíadiano R tienen al menos un mínimo común múltiplo en R .

8. En el problema 7, si el mínimo común múltiplo de a y b se denota por $[a, b]$, pruébese que $[a, b] = \frac{ab}{(a, b)}$.

8. UN ANILLO EUCLIDIANO PARTICULAR

Una abstracción en matemáticas gana importancia cuando, particularizada a un ejemplo específico, arroja nueva luz sobre este ejemplo. Vamos nosotros ahora a particularizar la noción de anillo euclíadiano a un anillo concreto, el anillo de los enteros gaussianos. Al aplicar los resultados generales obtenidos acerca de los anillos euclidianos a los enteros gaussianos obtendremos un teorema absolutamente nada trivial acerca de los números primos, teorema debido a Fermat.

Denotemos por $J[i]$ el conjunto de todos los números complejos de la forma $a+bi$ donde a y b son enteros. Bajo la adición y multiplicación habituales de los números complejos, $J[i]$ forma un dominio entero llamado el dominio de los *enteros gaussianos*.

Nuestro primer objetivo es exhibir $J[i]$ como un anillo euclíadiano. Para conseguirlo, introducimos una función $d(x)$ definida para todo elemento distinto de cero de $J[i]$ que satisface:

- 1) $d(x)$ es un entero no negativo para todo $x \neq 0 \in J[i]$.
- 2) $d(x) \leq d(xy)$ para todo $y \neq 0$ en $J[i]$.
- 3) Dados $u, v \in J[i]$ existen $t, r \in J[i]$ tales que $v = tu + r$ donde $r = 0$ o $d(r) < d(u)$.

Nuestro candidato para esta función d es el siguiente: si $x = a+bi \in J[i]$, entonces $d(x) = a^2 + b^2$. La $d(x)$ así definida satisface, sin duda, la propiedad (1); en efecto, si $x \neq 0 \in J[i]$, entonces $d(x) \geq 1$. Como es bien sabido, para cualesquiera dos números complejos (no necesariamente en $J[i]$) x, y , $d(xy) = d(x)d(y)$, luego si x y y están, además, en $J[i]$ y $y \neq 0$, entonces como $d(y) \geq 1$, $d(x) = d(x)1 \leq d(x)d(y) = d(xy)$, lo que muestra que la condición (2) se satisface. Todos nuestros esfuerzos deben ahora dirigirse a demostrar que la condición (3) también se verifica para esta función d en $J[i]$. Hacemos esto en la prueba del

TEOREMA 3.F. $J[i]$ es un anillo euclíadiano.

Prueba. Como hicimos notar en la anterior discusión, para probar el teorema 3.F solamente necesitamos probar que para cualesquiera $x, y \in J[i]$ existen $t, r \in J[i]$ tales que $y = tx + r$ donde $r = 0$ o $d(r) < d(x)$.

Establecemos primero esto para un caso muy especial, a saber, aquél en que y es arbitraria en $J[i]$, pero en que x es un entero positivo (ordinario) n .

Supongamos que $y = a + bi$; por el algoritmo de la división para el anillo de los enteros podemos encontrar enteros u, v tales que $a = un + u_1$ y $b = vn + v_1$ donde u_1 y v_1 son enteros que satisfacen $|u_1| \leq \frac{1}{2}n$ y $|v_1| \leq \frac{1}{2}n$. Sea $t = u + vi$ y $r = u_1 + v_1 i$; entonces $y = a + bi = un + u_1 + (vn + v_1)i = (u + vi)n + u_1 + v_1 i = tn + r$. Como $d(r) = d(u_1 + v_1 i) = u_1^2 + v_1^2 \leq n^2/4 + n^2/4 < n^2 = d(n)$, vemos que en este caso particular hemos mostrado que $y = tn + r$ con $r = 0$ o $d(r) < d(n)$.

Pasamos ahora al caso general; sean $x \neq 0$ y y elementos arbitrarios de $J[i]$. Tenemos así que $x\bar{x}$, donde \bar{x} es el conjugado complejo de x , es un entero positivo n . Aplicando el resultado del anterior párrafo a los elementos $y\bar{x}$ y n vemos que hay elementos $t, r \in J[i]$ tales que $y\bar{x} = tn + r$ con $r = 0$ o $d(r) < d(n)$. Poniendo en esta relación $n = x\bar{x}$ obtenemos $d(y\bar{x} - tx\bar{x}) < d(n) = d(x\bar{x})$; aplicando a esto el hecho de que $d(y\bar{x} - tx\bar{x}) = d(y - tx)d(\bar{x})$ y $d(x\bar{x}) = d(x)d(\bar{x})$ obtenemos que $d(y - tx)d(\bar{x}) < d(x)d(\bar{x})$. Como $x \neq 0$, $d(\bar{x})$ es un entero positivo y esta desigualdad se simplifica y nos da $d(y - tx) < d(x)$. Tenemos pues la representación $y = tx - r_0$ con $r_0 = y - tx$, con t y r_0 en $J[i]$ y como acabamos de ver, $r_0 = 0$ o $d(r_0) = d(y - tx) < d(x)$. Y esto prueba el teorema.

Como se ha probado que $J[i]$ es un anillo euclíadiano, podemos usar los resultados que hemos establecido acerca de esta clase de anillos en la sección previa al anillo euclíadiano que ahora estudiamos, el $J[i]$.

LEMA 3.15. *Sea p un entero primo y supongamos que para algún entero c primo relativo con p , podemos encontrar enteros x y y tales que $x^2 + y^2 = cp$. Entonces p puede escribirse como la suma de los cuadrados de dos enteros, es decir, existen enteros a y b tales que $p = a^2 + b^2$*

Prueba. El anillo de los enteros es un subanillo de $J[i]$. Supongamos que el entero p es también un elemento primo en $J[i]$. Como $cp = x^2 + y^2 = (x + yi)(x - yi)$, por el lema 3.13, $p|(x + yi)$ o $p|(x - yi)$ en $J[i]$. Pero si $p|(x + yi)$ entonces $x + yi = p(u + vi)$ lo que nos diría que $x = pu$ y $y = pv$ de forma que p dividiría también a $x - yi$. Pero entonces $p^2|(x + yi)(x - yi) = cp$, de lo que se seguiría que $p|c$ en contra de lo supuesto. Análogamente, si $p|(x - yi)$. Luego p no es un elemento primo en $J[i]$! Por tanto $p = (a + bi)(g + di)$ donde $a + bi$ y $g + di$ están en $J[i]$ y ninguno de ellos es una unidad en $J[i]$. Pero esto significa que ni $a^2 + b^2 = 1$ ni $g^2 + d^2 = 1$ (véase el problema 2). De $p = (a + bi)(g + di)$, se sigue fácilmente que $p = (a - bi)(g - di)$. Luego $p^2 = (a + bi)(g + di)(a - bi)(g - di) = (a^2 + b^2)(g^2 + d^2)$. Luego $(a^2 + b^2)|p^2$, luego $a^2 + b^2 = 1, p$ o p^2 ; $a^2 + b^2 \neq 1$, puesto que $a + bi$ no es una unidad en $J[i]$; $a^2 + b^2 \neq p^2$, pues de otra forma $g^2 + d^2 = 1$, en contra del hecho de que $g + di$ no es una unidad en $J[i]$. Luego la sola posibilidad que nos queda es que $a^2 + b^2 = p$ y el lema queda entonces establecido.

Los números primos impares se dividen en dos clases, los que tienen 1 como residuo al dividirse por 4 y aquellos otros que tienen residuo 3. Queremos probar que todo número primo de la primera clase puede escribirse como la suma de dos cuadrados, mientras que ningún primo de la segunda puede representarse de tal modo.

LEMA 3.16. *Si p es un número primo de la forma $4n+1$ entonces puede resolverse la congruencia $x^2 \equiv -1 \pmod{p}$.*

Prueba. Sea $x = 1 \cdot 2 \cdot 3 \cdots \frac{p-1}{2}$. como $p-1 = 4n$, en este producto

para la obtención de x hay un número par de términos, en consecuencia de lo cual

$x = (-1)(-2)(-3) \cdots \left(-\left(\frac{p-1}{2}\right)\right)$. Pero $p-k \equiv -k \pmod{p}$, de modo que

$$x^2 \equiv \left(1 \cdot 2 \cdots \frac{p-1}{2}\right) (-1)(-2) \cdots \left(-\left(\frac{p-1}{2}\right)\right)$$

$$\equiv 1 \cdot 2 \cdots \frac{p-1}{2} \frac{p+1}{2} \cdots (p-1)$$

$$\equiv (p-1)! = -1 \pmod{p}.$$

Estamos usando aquí el teorema de Wilson, que anteriormente probamos, a saber, que si p es un número primo $(p-1)! \equiv -1 \pmod{p}$.

Ilustremos este resultado. Si $p = 13$, $x = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 = 720 = 5 \pmod{13}$, y $5^2 = -1 \pmod{13}$.

TEOREMA 3.G (FERMAT). *Si p es un número primo de la forma $4n+1$ entonces $p = a^2 + b^2$ para algunos enteros a y b .*

Prueba. De acuerdo con el lema 3.16 existe un x tal que $x^2 \equiv -1 \pmod{p}$. Puede escogerse x de forma que $0 \leq x \leq p-1$ ya que solamente necesitamos usar el residuo de x en la división por p . Podemos restringir el tamaño de x aún más de modo que satisfaga $|x| \leq p/2$. Pues si $x > p/2$ entonces $y = p-x$ satisface $y^2 \equiv -1 \pmod{p}$ y $|y| \leq p/2$. Podemos, pues, suponer que tenemos un entero x tal que $|x| \leq p/2$ y $x^2 + 1$ es un múltiplo de p , digamos cp . Ahora bien, $cp = x^2 + 1 \leq p^2/4 + 1 < p^2$, de donde $c < p$, y, por tanto, $p \nmid c$. De donde, según el lema 3.15 obtenemos que $p = a^2 + b^2$ para algunos enteros a y b , probando así el teorema.

Problemas

1. Encuéntrense todas las unidades en $J[i]$.
2. Si $a+bi$ no es una unidad en $J[i]$ pruébese que $a^2+b^2 > 1$.
3. Encuéntrese el máximo común divisor en $J[i]$ de:
 - a) $3+4i$ y $4-3i$.
 - b) $11+7i$ y $18-i$.
4. Pruébese que si p es un número primo de la forma $4n+3$ entonces no hay x alguno tal que $x^2 \equiv -1 \pmod{p}$.
5. Pruébese que ningún primo de la forma $4n+3$ puede escribirse como a^2+b^2 donde a y b sean números enteros.
6. Pruébese que hay un número infinito de primos de la forma $4n+3$.
- *7. Pruébese que existe un número infinito de primos de la forma $4n+1$.
- *8. Determínense todos los elementos primos en $J[i]$.
- *9. Determínense todos los enteros positivos que pueden escribirse como una suma de dos cuadrados (de enteros).

9. ANILLOS DE POLINOMIOS

En nuestra educación matemática se nos introdujo muy pronto —generalmente en los primeros años de secundaria— al estudio de los polinomios. Durante una temporada que parecía que no iba a tener fin se nos obligaba hasta el punto de aburrimiento insopportable a factorizarlos, multiplicarlos, dividirlos y simplificarlos. La facilidad en factorizar un cuadrático se interpretaba como una muestra de genuino talento matemático.

Posteriormente, en los primeros años de universidad, los polinomios hacen de nuevo aparición en un marco algo distinto. Ahora son funciones, con sus valores, y nos preocupan su continuidad, sus derivadas, sus integrales y sus máximos y sus mínimos.

También aquí nos interesaremos en los polinomios, pero desde un punto de vista distinto de cualquiera de los dos que hemos enumerado. Para nosotros, los polinomios serán simplemente elementos de un cierto anillo y lo que nos interesará serán las propiedades algebraicas de ese anillo. Nuestro interés primario en ellos estriba en que son anillos euclidianos con propiedades que serán decisivas en la discusión que posteriormente tendremos de campos y extensiones de campos.

Sea F un campo. Llamaremos *anillo de polinomios* sobre F en la indeterminada x , y representaremos por $F[x]$, al conjunto de todos los símbolos $a_0+a_1x+\dots+a_nx^n$, donde n puede ser cualquier entero no negativo y

donde los coeficientes $a_0, a_1, a_2, \dots, a_n$ están todos en F . Para que $F[x]$ sea un anillo debemos ser capaces de reconocer cuando dos de sus elementos son iguales y saber cómo sumarlos y multiplicarlos de forma que los axiomas definitorios de un anillo se verifiquen en $F[x]$. Este será nuestro objetivo inicial.

Podríamos evitar la frase "el conjunto de todos los símbolos" que acabamos de usar por la introducción de un argumento apropiado de sucesiones, pero parece más conveniente seguir el camino más familiar a la mayor parte de nuestros lectores.

DEFINICIÓN. Si $p(x) = a_0 + a_1x + \dots + a_nx^n$ y $q(x) = b_0 + b_1x + \dots + b_nx^n$ están, ambos, en $F[x]$ entonces $p(x) = q(x)$ si y sólo si para todo entero $i \geq 0$, $a_i = b_i$.

Así pues, dos polinomios se dice son iguales si y sólo si sus correspondientes coeficientes son iguales.

DEFINICIÓN. Si $p(x) = a_0 + a_1x + \dots + a_nx^n$ y $q(x) = b_0 + b_1x + \dots + b_nx^n$ están ambos en $F[x]$ entonces $p(x) + q(x) = c_0 + c_1x + \dots + c_nx^n$ donde para cada i , $c_i = a_i + b_i$.

En otras palabras, para sumar dos polinomios se suman los coeficientes de sus términos semejantes. Para sumar $1+x$ y $3-2x+x^2$ consideramos $1+x$ como $1+x+0x^2$ y sumamos luego de acuerdo con la fórmula dada para obtener como su suma $4-x+x^2$.

La operación más complicada es la única que nos queda por definir en $F[x]$, la multiplicación.

DEFINICIÓN. Si $p(x) = a_0 + a_1x + \dots + a_nx^n$ y $q(x) = b_0 + b_1x + \dots + b_nx^n$ entonces $p(x)q(x) = c_0 + c_1x + \dots + c_kx^k$ donde $c_i = a_i b_0 + a_{i-1} b_1 + a_{i-2} b_2 + \dots + a_0 b_i$.

Esta definición no dice más que para multiplicar dos polinomios se multiplican los símbolos formalmente, se usa la relación $x^\alpha x^\beta = x^{\alpha+\beta}$ y se reducen los términos semejantes. Ilustremos la definición con un ejemplo:

$$p(x) = 1 + x - x^2 \quad q(x) = 2 + x^2 + x^3.$$

Aquí $a_0 = 1$, $a_1 = 1$, $a_2 = -1$, $a_3 = a_4 = \dots = 0$, y $b_0 = 2$, $b_1 = 0$, $b_2 = 1$, $b_3 = 1$, $b_4 = b_5 = \dots = 0$. De donde

$$c_0 = a_0 b_0 = 1 \cdot 2 = 2,$$

$$c_1 = a_1 b_0 + a_0 b_1 = 1 \cdot 2 + 1 \cdot 0 = 2,$$

$$c_2 = a_2 b_0 + a_1 b_1 + a_0 b_2 = (-1)(2) + 1 \cdot 0 + 1 \cdot 1 = -1,$$

$$c_3 = a_3 b_0 + a_2 b_1 + a_1 b_2 + a_0 b_3 = (0)(2) + (-1)(0) + 1 \cdot 1 + 1 \cdot 1 = 2,$$

$$c_4 = a_4 b_0 + a_3 b_1 + a_2 b_2 + a_1 b_3 + a_0 b_4 = (0)(2) + (0)(0) + (-1)(1) \\ + (1)(1) + 1(0) = 0,$$

$$c_5 = a_5 b_0 + a_4 b_1 + a_3 b_2 + a_2 b_3 + a_1 b_4 + a_0 b_5 = (0)(2) + (0)(0) + (0)(1) \\ + (-1)(1) + (1)(0) + (0)(0) = -1,$$

$$c_6 = a_6 b_0 + a_5 b_1 + a_4 b_2 + a_3 b_3 + a_2 b_4 + a_1 b_5 + a_0 b_6 = (0)(2) + (0)(0) \\ + (0)(1) + (0)(1) + (-1)(0) + (1)(0) + (1)(0) = 0,$$

$$c_7 = c_8 = \dots = 0.$$

Por tanto, de acuerdo con nuestra definición,

$$(1+x-x^2)(2+x^2+x^3) = c_0 + c_1 x + \dots = 2 + 2x - x^2 + 2x^3 - x^5.$$

Si se efectúa esta multiplicación de acuerdo con lo que el lector aprendió en sus estudios de secundaria se verá que llegamos a la misma contestación. Nuestra definición de producto es la que el lector ya conocía.

Sin más examen afirmamos que $F[x]$ es un anillo bajo estas operaciones, que su multiplicación es conmutativa y que tiene un elemento unitario. La verificación de todas estas afirmaciones se la dejamos al lector.

DEFINICIÓN. Si $f(x) = a_0 + a_1 x + \dots + a_n x^n \neq 0$ y $a_n \neq 0$, entonces el grado de $f(x)$, escrito: $\deg f(x)$, es n .

Es decir, el grado de $f(x)$ es el mayor entero i para el que el i -ésimo coeficiente de $f(x)$ no es 0. No definimos el grado del polinomio cero. Decimos que un polinomio es una *constante* si es de grado 0 o es el polinomio cero. La función de grado definida sobre los elementos de $F[x]$ distintos del cero nos provee de la función $d(x)$ necesaria para que $F[x]$ sea un anillo euclíadiano.

LEMA 3.17. Si $f(x), g(x)$ son dos elementos distintos del cero de $F[x]$ entonces $\deg(f(x)g(x)) = \deg f(x) + \deg g(x)$.

Prueba. Supongamos que $f(x) = a_0 + a_1 x + \dots + a_m x^m$ y $g(x) = b_0 + b_1 x + \dots + b_n x^n$ y que $a_m \neq 0$ y $b_n \neq 0$. Tenemos pues, $\deg f(x) = m$ y $\deg g(x) = n$. Por definición $f(x)g(x) = c_0 + c_1 x + \dots + c_k x^k$ donde $c_i = a_i b_0 + a_{i-1} b_1 + \dots + a_1 b_{i-1} + a_0 b_i$. Nosotros afirmamos que $c_{m+n} = a_m b_n \neq 0$ y que $c_i = 0$ para todo $i > m+n$. Que $c_{m+n} = a_m b_n$ se puede ver por la definición de inmediato. Pero qué hay sobre c_i para $i > m+n$? c_i es la suma de los términos de la forma $a_j b_{i-j}$; como $i = j + (i-j) > m+n$, entonces o $j > m$ o $(i-j) > n$. Pero entonces uno de los dos, a_j o b_{i-j} es 0, luego $a_j b_{i-j} = 0$; como c_i es la suma de una colección de ceros él mismo es 0, y nuestra afirmación ha quedado probada. Luego el coeficiente más

alto distinto de cero de $f(x)g(x)$ es c_{m+n} , de donde $\deg f(x)g(x) = m+n = \deg f(x) + \deg g(x)$.

COROLARIO. Si $f(x)$ y $g(x)$ son elementos de $F[x]$ distintos de cero, entonces $\deg f(x) \leq \deg f(x)g(x)$.

Prueba. Como $\deg f(x)g(x) = \deg f(x) + \deg g(x)$, y como $\deg g(x) \geq 0$ este resultado se sigue inmediatamente del lema.

COROLARIO. $F[x]$ es un dominio entero.

Dejamos al lector la prueba de este corolario.

Como $F[x]$ es un dominio entero, a la luz del teorema 3.c podemos construir por ello su campo de cocientes. Este campo se compone simplemente de todos los cocientes de polinomios y se llama el campo de las *funciones racionales en x sobre F*.

La función $\deg f(x)$ definida para todo $f(x) \neq 0$ en $F[x]$ tiene las siguientes propiedades:

- 1) $\deg f(x)$ es un entero no negativo.
- 2) $\deg f(x) \leq \deg f(x)g(x)$ para todo $g(x) \neq 0$ en $F[x]$.

Para que $F[x]$ sea un anillo euclíadiano con la función de grado actuando como la d -función de un anillo euclíadiano, necesitamos aun que dadas $f(x), g(x) \in F[x]$ existan $t(x), r(x)$ en $F[x]$ tales que $f(x) = t(x)g(x) + r(x)$ en donde $r(x) = 0$ o $\deg r(x) < \deg g(x)$. Probamos esto en el siguiente lema.

LEMA 3.18. (EL ALGORITMO DE LA DIVISIÓN). Dados dos polinomios $f(x)$ y $g(x)$ de $F[x]$ con $g(x) \neq 0$, existen entonces dos polinomios $t(x)$ y $r(x)$ en $F[x]$ tales que $f(x) = t(x)g(x) + r(x)$ donde $r(x) = 0$ o $\deg r(x) < \deg g(x)$.

Prueba. La prueba no es en realidad nada más que el proceso de la "división larga" que todos usamos para dividir un polinomio entre otro.

Si el grado de $f(x)$ es menor que el de $g(x)$ nada hay que probar, pues nada más tenemos que poner $t(x) = 0$, $r(x) = f(x)$ y, ciertamente, tenemos $f(x) = 0g(x) + f(x)$ donde $\deg f(x) < \deg g(x)$ o $f(x) = 0$.

Podemos, pues, suponer que $f(x) = a_0 + a_1x + \dots + a_mx^m$ y $g(x) = b_0 + b_1x + \dots + b_nx^n$ con $a_m \neq 0$ y $b_n \neq 0$ y $m \geq n$.

Sea $f_1(x) = f(x) - (a_m/b_n)x^{m-n}g(x)$; entonces $\deg f_1(x) \leq m-1$, de donde por inducción sobre el grado de $f(x)$ podemos suponernos que $f_1(x) = t_1(x)g(x) + r(x)$ donde $r(x) = 0$ o $\deg r(x) < \deg g(x)$. Pero entonces $f(x) - (a_m/b_n)x^{m-n}g(x) = t_1(x)g(x) + r(x)$, lo cual, por transposición, nos da $f(x) = ((a_m/b_n)x^{m-n} + t_1(x))g(x) + r(x)$. Si ponemos $t(x) = (a_m/b_n)x^{m-n} + t_1(x)$ tendremos que $f(x) = t(x)g(x) + r(x)$, donde $t(x), r(x) \in F[x]$ y donde $r(x) = 0$ o $\deg r(x) < \deg g(x)$. Lo que prueba el lema.

Con este último lema hemos llenado todos los requisitos necesarios para exhibir a $F[x]$ como un anillo euclíadiano y tenemos ya el derecho de decir:

TEOREMA 3.H. *$F[x]$ es un anillo euclíadiano.*

Todos los resultados de la sección 7 son ahora aplicables y los enumeramos, para este caso particular nuestro, en los siguientes lemas. Podría ser muy instructivo para el lector intentar probarlos directamente, adaptando los argumentos usados en la sección 7 a nuestro anillo particular $F[x]$ y su función euclíadiana, el grado.

LEMA 3.19. *$F[x]$ es un anillo de ideales principales.*

LEMA 3.20. *Dados dos polinomios $f(x)$ y $g(x)$ en $F[x]$ tienen siempre un máximo común divisor $d(x)$ que puede escribirse en la forma $d(x) = \lambda(x)f(x) + \mu(x)g(x)$.*

¿Qué es lo que corresponde a un elemento primo?

DEFINICIÓN. Un polinomio $p(x)$ en $F[x]$ se dice que es *irreducible* sobre F si siempre que $p(x) = a(x)b(x) \in F[x]$, entonces uno de los dos, $a(x)$ o $b(x)$, tiene grado cero (es decir, es una constante).

La irreducibilidad depende del campo; por ejemplo, el polinomio $x^2 + 1$ es irreducible sobre el campo real, pero no sobre el campo complejo, pues en este último $x^2 + 1 = (x+i)(x-i)$, donde $i^2 = -1$.

LEMA 3.21. *Cualquier polinomio en $[F[x]]$ puede escribirse en forma única como un producto de polinomios irreducibles en $F[x]$.**

LEMA 3.22. *El ideal $A = (p(x))$ en $F[x]$ es un ideal máximo si y sólo si $p(x)$ es irreducible sobre F .*

En el capítulo 5 volveremos a considerar de forma mucho más detenida a este campo $F[x]/(p(x))$, pero por el momento vamos a calcular un ejemplo.

Sea F el campo de los números racionales y consideremos el polinomio $p(x) = x^3 - 2$ en $F[x]$. Puede verificarse fácilmente que es irreducible sobre F , de donde $F[x]/(x^3 - 2)$ es un campo. ¿Qué aspecto tienen sus elementos? Comencemos por convenir en representar por $A = (x^3 - 2)$ al ideal de $F[x]$ generado por $x^3 - 2$.

Cualquier elemento en $F[x]/(x^3 - 2)$ es una clase lateral de la forma $f(x) + A$ del ideal A con $f(x) \in F[x]$. Ahora bien, dado un polinomio cualquiera $f(x) \in F[x]$, por el algoritmo de la división,

$f(x) = t(x)(x^3 - 2) + r(x)$ donde $r(x) = 0$ o $\deg r(x) < \deg(x^3 - 2) = 3$. Luego $r(x) = a_0 + a_1 x + a_2 x^2$, donde a_0, a_1, a_2 están en F ; por consiguiente $f(x) + A = a_0 + a_1 x + a_2 x^2 + t(x)(x^3 - 2) + A = a_0 + a_1 x + a_2 x^2 + A$ ya que $t(x)(x^3 - 2)$ está en A , de donde por la adición y la multiplicación en $F[x]/(x^3 - 2)$, $f(x) + A = (a_0 + A) + a_1(x + A) + a_2(x + A)^2$. Si ponemos $t = x + A$, entonces todo elemento en $F[x]/(x^3 - 2)$ es de la forma $a_0 + a_1 t + a_2 t^2$ con a_0, a_1, a_2 en F . ¿Y qué puede decirse acerca de t ? Como $t^3 - 2 = (x + A)^3 - 2 = x^3 - 2 + A = A = 0$ (ya que A es el elemento cero de $F[x]/(x^3 - 2)$) vemos que $t^3 = 2$.

Además, si $a_0 + a_1 t + a_2 t^2 = b_0 + b_1 t + b_2 t^2$, entonces $(a_0 - b_0) + (a_1 - b_1)t + (a_2 - b_2)t^2 = 0$, de donde $(a_0 - b_0) + (a_1 - b_1)x + (a_2 - b_2)x^2$ está en $A = (x^3 - 2)$. ¿Cómo puede ser esto si cada uno de los elementos de A tiene como grado al menos 3? Solamente si $a_0 - b_0 + (a_1 - b_1)x + (a_2 - b_2)x^2 = 0$, es decir, solamente si $a_0 = b_0, a_1 = b_1$ y $a_2 = b_2$. Luego cada elemento en $F[x]/(x^3 - 2)$ tiene una representación única de la forma $a_0 + a_1 t + a_2 t^2$ donde $a_0, a_1, a_2 \in F$. Según el lema 3.22 $F[x]/(x^3 - 2)$ es un campo. Será interesante ver esto directamente; todo lo que se requiere es probar que si $a_0 + a_1 t + a_2 t^2 \neq 0$, entonces tiene un inverso de la forma $\alpha + \beta t + \gamma t^2$. De aquí que lo que tenemos que hacer es hallar la solución para α, β , y γ de la relación $(a_0 + a_1 t + a_2 t^2)(\alpha + \beta t + \gamma t^2) = 1$, en donde no todos los tres a_0, a_1 y a_2 son iguales a cero. Efectuando la multiplicación y usando $t^3 = 2$ obtenemos $(a_0\alpha + 2a_2\beta + 2a_1\gamma) + (a_1\alpha + a_0\beta + 2a_2\gamma)t + (a_2\alpha + a_1\beta + a_0\gamma)t^2 = 1$; luego:

$$a_0\alpha + 2a_2\beta + 2a_1\gamma = 1$$

$$a_1\alpha + a_0\beta + 2a_2\gamma = 0$$

$$a_2\alpha + a_1\beta + a_0\gamma = 0.$$

Podemos intentar resolver estas tres ecuaciones en las tres incógnitas α, β, γ . Cuando lo hacemos, se ve que sólo existe una solución si y sólo si:

$$a_0^3 + 2a_1^3 + 4a_2^3 - 6a_0a_1a_2 \neq 0.$$

Por tanto, el problema de probar directamente que $F[x]/(x^3 - 2)$ es un campo se reduce a probar que la sola solución en números racionales de

$$(1) \quad a_0^3 + 2a_1^3 + 4a_2^3 = 6a_0a_1a_2$$

es la solución $a_0 = a_1 = a_2 = 0$. Probémoslo. Si existe una solución en números racionales, quitando denominadores podemos demostrar que existe una solución donde a_0, a_1 y a_2 son enteros. Luego podemos suponer que a_0, a_1 y a_2 son enteros que satisfacen (1). Afirmamos ahora que podemos suponer que a_0, a_1, a_2 no tienen más común divisor que 1, pues si $a_0 = b_0d, a_1 = b_1d$ y $a_2 = b_2d$ donde d es su máximo común divisor, entonces, sustituyendo en (1), obtenemos $d^3(b_0^3 + 2b_1^3 + 4b_2^3) = d^3(6b_0b_1b_2)$ y, por tanto, $b_0^3 + 2b_1^3 + 4b_2^3 = 6b_0b_1b_2$. El problema se ha reducido, por

tanto a probar, que (1) no tiene ninguna solución en enteros que sean primos relativos. Pero entonces (1) implica que a_0^3 es par; puesto que a_0 es par; sustituyendo a_0 por $2\alpha_0$ en (1) tenemos $4a_0^3 + a_1^3 + 2a_2^3 = 6a_0a_1a_2$. Luego a_1^3 , y por tanto a_1 , es par; $a_1 = 2\alpha_1$. Sustituyendo en (1) obtenemos $2a_0^3 + 4a_1^3 + a_2^3 = 6a_0a_1a_2$. Luego a_2^3 , y por tanto a_2 , es par! Pero entonces a_0, a_1 y a_2 tienen a 2 como factor común! Esto contradice el hecho de que son primos relativos, y hemos así probado que la ecuación $a_0^3 + 2a_1^3 + 4a_2^3 = 6a_0a_1a_2$ no tiene ninguna otra solución racional que $a_0 = a_1 = a_2 = 0$. Por tanto, podemos hallar las soluciones para α, β , y γ y $\frac{F[x]}{(x^3 - 2)}$ se ha visto, directamente, que es un campo.

Problemas

1. Encuéntrese el máximo común divisor de los siguientes polinomios sobre F , el campo de los números racionales:

- a) $x^3 - 6x^2 + x + 4$ y $x^5 - 6x + 1$.
- b) $x^2 + 1$ y $x^6 + x^3 + x + 1$.

2. Pruébese que:

- a) $x^2 + x + 1$ es irreducible sobre F , el campo de los enteros módulo 2.
- b) $x^2 + 1$ es irreducible sobre los enteros módulo 7.
- c) $x^3 - 9$ es irreducible sobre los enteros mód 31.
- d) $x^3 - 9$ es irreducible sobre los enteros mód 11.

3. Sean F y K dos campos con $F \subset K$, y supongamos $f(x), g(x) \in F[x]$ son primos relativos en $F[x]$. Pruébese que son primos relativos en $K[x]$.

4. a) Pruébese que $x^2 + 1$ es irreducible sobre el campo F de los enteros mód 11 y pruébese directamente que $F[x]/(x^2 + 1)$ es un campo que tiene 121 elementos.

b) Pruébese que $x^2 + x + 4$ es irreducible sobre F , el campo de los enteros mód 11 y pruébese directamente que $F[x]/(x^2 + x + 4)$ es un campo con 121 elementos.

*c) Pruébese que los campos de la parte (a) y la parte (b) son isomorfos.

5. Sea F el campo de los números reales. Pruébese que $F[x]/(x^2 + 1)$ es un campo isomorfo al campo de los números complejos.

*6. Definamos la derivada $f'(x)$ del polinomio

$$f(x) = a_0 + a_1x + \dots + a_nx^n$$

como

$$f'(x) = a_1 + 2a_2x + 3a_3x^2 + \dots + na_nx^{n-1}$$

Pruébese que si $f(x) \in F[x]$, donde F es el campo de los números racionales, entonces $f(x)$ es divisible por el cuadrado de un polinomio si y sólo si $f(x)$ y $f'(x)$ tienen un máximo común divisor $d(x)$ de grado positivo.

7. Si $f(x)$ está en $F[x]$, donde F es el campo de los enteros módulo p , p un primo, y $f(x)$ es irreducible sobre F de grado n , pruébese que $F[x]/(f(x))$ es un campo con p^n elementos.

10. POLINOMIOS SOBRE EL CAMPO RACIONAL

Especializamos la discusión general a la de los polinomios cuyos coeficientes son números racionales. La mayor parte del tiempo, los coeficientes serán en realidad enteros. Nos ocuparemos fundamentalmente de la irreducibilidad de tales polinomios.

DEFINICIÓN. Se dice que el polinomio $f(x) = a_0 + a_1 x + \dots + a_n x^n$, donde $a_0, a_1, a_2, \dots, a_n$ son enteros, es *primitivo* si el máximo común divisor de a_0, a_1, \dots, a_n es 1.

LEMA 3.23. Si $f(x)$ y $g(x)$ son polinomios primitivos entonces $f(x)g(x)$ es un polinomio primitivo.

Prueba. Sea $f(x) = a_0 + a_1 x + \dots + a_n x^n$ y $g(x) = b_0 + b_1 x + \dots + b_m x^m$. Supongamos que el lema fuera falso; entonces todos los coeficientes de $f(x)g(x)$ serían divisibles por algún entero mayor que 1, de donde, por algún número primo p . Como $f(x)$ es primitivo, p no divide a alguno de los coeficientes a_i . Sea a_j el primer coeficiente de $f(x)$ al que p no divide. Análogamente, sea b_k el primer coeficiente de $g(x)$ al que p no divide. En $f(x)g(x)$ el coeficiente de x^{j+k} , c_{j+k} , es

$$(1) \quad c_{j+k} = a_j b_k + (a_{j+1} b_{k-1} + a_{j+2} b_{k-2} + \dots + a_{j+k} b_0) \\ + (a_{j-1} b_{k+1} + a_{j-2} b_{k+2} + \dots + a_0 b_{j+k}).$$

Ahora bien, por nuestra elección de b_k , $p \nmid b_{k-1}, b_{k-2}, \dots$, de modo que $p \nmid (a_{j+1} b_{k-1} + a_{j+2} b_{k-2} + \dots + a_{j+k} b_0)$. Análogamente, por nuestra elección de a_j , $p \nmid a_{j-1}, a_{j-2}, \dots$, de modo que $p \nmid (a_{j-1} b_{k+1} + a_{j-2} b_{k+2} + \dots + a_0 b_{j+k})$. Por hipótesis, $p \mid c_{j+k}$. Luego por (1), $p \mid a_j b_k$, lo que es imposible ya que $p \nmid a_j$ y $p \nmid b_k$. Con lo que el lema queda probado.

DEFINICIÓN. El *contenido* del polinomio $f(x) = a_0 + a_1 x + \dots + a_n x^n$, donde todos los a son enteros, es el máximo común divisor de los enteros a_0, a_1, \dots, a_n .

Es claro que todo polinomio $p(x)$ con coeficientes enteros puede escribirse como $p(x) = dq(x)$, donde d es el contenido de $p(x)$ y $q(x)$ es un polinomio primitivo.

TEOREMA 3.1 (LEMA DE GAUSS). Si el polinomio primitivo $f(x)$ puede factorizarse como el producto de dos polinomios de coeficientes racionales,

entonces puede factorizarse como el producto de dos polinomios de coeficientes enteros.

Prueba. Supongamos que $f(x) = u(x)v(x)$ donde $u(x)$ y $v(x)$ tienen coeficientes racionales. Quitando denominadores y sacando los factores comunes podemos escribir entonces $f(x) = (a/b)\lambda(x)\mu(x)$, donde a y b son enteros y donde tanto $\lambda(x)$ como $\mu(x)$ tienen coeficientes enteros y son primativos. Luego $bf(x) = a\lambda(x)\mu(x)$. El contenido del primer miembro es b , ya que $f(x)$ es primitivo; como $\lambda(x)$ y $\mu(x)$ son primativos, según el lema 3.23 $\lambda(x)\mu(x)$ es primitivo, luego el contenido del segundo miembro es a . Por lo tanto, $a = b$, $(a/b) = 1$ y $f(x) = \lambda(x)\mu(x)$ donde $\lambda(x)$ y $\mu(x)$ tienen coeficientes enteros. Y esto es lo que el teorema afirma.

DEFINICIÓN. Un polinomio se dice que es *entero mónico* si todos sus coeficientes son enteros y su coeficiente más alto es 1.

Luego un polinomio entero mónico es simplemente uno de la forma $x^n + a_1x^{n-1} + \dots + a_n$, donde todos los a son enteros. Es claro que todo polinomio entero mónico es primitivo.

COROLARIO. Si un polinomio entero mónico se factoriza como el producto de dos polinomios no constantes de coeficientes racionales, entonces se factoriza como el producto de dos polinomios enteros monicos.

Dejamos la prueba de este corolario como un ejercicio para el lector.

El problema de decidir si un polinomio dado es o no irreducible puede ser difícil y laborioso. Pocos criterios existen que declaren que un polinomio dado es o no irreducible. Uno de estos pocos es el siguiente.

TEOREMA 3.J (EL CRITERIO DE EISENSTEIN). Sea $f(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$ un polinomio con coeficientes enteros. Supongamos que para algún número primo p , $p \nmid a_n$, $p \mid a_1$, $p \mid a_2, \dots, p \mid a_0$ y $p^2 \nmid a_0$. Entonces $f(x)$ es irreducible sobre los racionales.

Prueba. Sin pérdida de generalidad podemos suponer que $f(x)$ es primitivo, pues el sacar el máximo común divisor de sus coeficientes no modifica la hipótesis, ya que $p \nmid a_n$. Si $f(x)$ se factoriza como un producto de dos polinomios racionales, por el lema de Gauss, se factoriza también como el producto de dos polinomios de coeficientes enteros. Luego si suponemos que $f(x)$ es reducible, entonces

$$f(x) = (b_0 + b_1x + \dots + b_rx^r)(c_0 + c_1x + \dots + c_sx^s),$$

donde las b y las c son enteros y donde $r > 0$ y $s > 0$. Comparando los coeficientes de ambos miembros tenemos primero $a_0 = b_0c_0$. Como $p \mid a_0$,

p debe dividir a uno de los dos, b_0 o c_0 . Como $p^2 \nmid a_0$, p no puede dividir a la vez a ambos b_0 y c_0 . Supongamos que $p \mid b_0$ y $p \nmid c_0$. No todos los coeficientes b_0, \dots, b_r pueden ser divisibles por p ; de otro modo todos los coeficientes de $f(x)$ serían divisibles por p , lo que es manifiestamente falso, ya que $p \nmid a_n$. Sea b_k el primer b no divisible por p , $k \leq r < n$. Tenemos entonces que $p \mid b_{k-1}$ y a las b anteriores. Pero $a_k = b_k c_0 + b_{k-1} c_1 + b_{k-2} c_2 + \dots + b_0 c_k$, y $p \mid a_k$ y $p \mid b_{k-1}, b_{k-2}, \dots, b_0$, de modo que $p \mid b_k c_0$. Pero $p \nmid c_0$ y $p \nmid b_k$, lo que entra en conflicto con que $p \mid b_k c_0$. Esta contradicción prueba que nosotros no pudimos haber factorizado $f(x)$. Luego $f(x)$ es irreducible.

Problemas

1. Sea D un anillo euclíadiano y F su campo de cocientes. Pruébese el lema de Gauss para polinomios con coeficientes en D factorizados como productos de polinomios en F .
2. Si p es un número primo, pruébese que el polinomio $x^n - p$ es irreducible sobre los racionales. Por el criterio de Eisenstein.
3. Pruébese que el polinomio $1 + x + \dots + x^{p-1}$, donde p es un número primo, es irreducible sobre el campo de los números racionales. (Sugerencia: Considérese el polinomio $1 + (x+1) + (x+1)^2 + \dots + (x+1)^{p-1}$, y úsese el criterio de Eisenstein.)
4. Si m y n son enteros primos relativos y si $\left(x - \frac{m}{n}\right) \mid (a_0 + a_1 x + \dots + a_r x^r)$, donde las a son enteros, pruébese que $m \mid a_0$ y $n \mid a_r$.
5. Si a es racional y $x-a$ divide a un polinomio entero mónico, pruébese que a debe ser un entero. Concurrencia de 4.1.

11. ANILLOS DE POLINOMIOS SOBRE ANILLOS COMMUTATIVOS

Al definir el anillo de polinomios en una variable sobre un campo F , no se hizo ningún uso esencial del hecho de que F fuera un campo; todo lo que se usó fue que F era un anillo conmutativo. La naturaleza de campo de F solamente se hizo sentir en la prueba de que $F[x]$ era un anillo euclíadiano.

Podemos, pues, imitar lo que hicimos con campos con anillos más generales. Mientras que algunas cualidades se perderán como, por ejemplo, la del “euclidismo”, veremos que quedan las suficientes para llevarnos a resultados interesantes. Podíamos haber desarrollado el sujeto con esta generalidad desde el principio, para luego haber obtenido resultados par-

ticulares para $F[x]$, cuando F era un campo. Sin embargo, sentimos que es más conveniente ir de lo concreto a lo abstracto que de lo abstracto a lo concreto. El precio que pagamos por ello es repetición, pero incluso esto sirve a un propósito, a saber, el de consolidar las ideas. Por la experiencia que ganamos al tratar de polinomios sobre campos, nos podemos permitir ser un poco más esquemáticos en las pruebas que aquí demos.

Sea R un anillo comutativo con elemento unitario. Por *el anillo de polinomios en x sobre R* , $R[x]$, entendemos el conjunto de todos los símbolos formales $a_0 + a_1 x + \dots + a_m x^m$, donde a_0, a_1, \dots, a_m están en R , y donde la igualdad, la adición y la multiplicación se definen exactamente igual que como se definieron en la sección 9. Como en tal sección, $R[x]$ es un anillo comutativo con elemento unitario.

Definimos ahora el *anillo de polinomios en las n variables x_1, \dots, x_n sobre R* , $R[x_1, \dots, x_n]$, como sigue: sea $R_1 = R[x_1]$, $R_2 = R_1[x_2]$, el anillo de polinomios en x_2 sobre $R_1, \dots, R_n = R_{n-1}[x_n]$. A R_n se le llama el anillo de polinomios en x_1, \dots, x_n sobre R . Sus elementos son de la forma $\sum a_{i_1, i_2, \dots, i_n} x_1^{i_1} x_2^{i_2} \dots x_n^{i_n}$, con la igualdad y la adición definidas por los coeficientes y la multiplicación por el uso de la ley distributiva y la regla de exponentes $(x_1^{i_1} x_2^{i_2} \dots x_n^{i_n})(x_1^{j_1} x_2^{j_2} \dots x_n^{j_n}) = x_1^{i_1+j_1} x_2^{i_2+j_2} \dots x_n^{i_n+j_n}$. De particular importancia es el caso en que $R = F$ es un campo; obtenemos entonces el anillo de polinomios en n variables sobre un campo.

Va a interesarnos en particular la influencia de la estructura de R sobre la de $R[x_1, \dots, x_n]$. El primer resultado en tal dirección es el que sigue.

LEMA 3.24. *Si R es un dominio entero, entonces también lo es $R[x]$.*

Prueba. Para $0 \neq f(x) = a_0 + a_1 x + \dots + a_m x^m$, donde $a_m \neq 0$, en $R[x]$, definimos el *grado* de $f(x)$ como m ; es decir, el grado de $f(x)$ es el índice del coeficiente distinto de cero más alto de $f(x)$. Si R es un dominio entero dejamos como ejercicio probar que $\deg(f(x)g(x)) = \deg f(x) + \deg g(x)$. Pero entonces para $f(x) \neq 0, g(x) \neq 0$, es imposible tener $f(x)g(x) = 0$. Es decir, $R[x]$ es un dominio entero.

El uso reiterado del lema nos da como resultado inmediato el siguiente

COROLARIO. *Si R es un dominio entero, entonces también lo es $R[x_1, \dots, x_n]$.*

En particular, cuando F es un campo, $F[x_1, \dots, x_n]$ debe ser un dominio entero. Como tal, podemos construir su campo de cocientes; a tal campo le llamamos *campo de funciones racionales en x_1, \dots, x_n sobre F* y le denotamos por $F(x_1, \dots, x_n)$. Este campo desempeña un papel vital en geometría algebraica. Para nosotros será de máxima importancia en nuestra discusión, en el capítulo 5, de la teoría de Galois.

Pero queremos interrelaciones más profundas entre las estructuras de R y de $R[x_1, \dots, x_n]$ que las que aparecen en el lema 3.24, y es en tal dirección en la que ahora vamos a avanzar.

Exactamente en la misma forma en que lo hicimos para los anillos euclidianos, podemos hablar ahora de divisibilidad, unidades, etc., en dominios enteros arbitrarios, R , con elemento unitario. Dos elementos a y b en R se dice que están *asociados* si $a = ub$ donde u es una unidad en R . Un elemento a que no es una unidad en R se dirá que es *irreducible* (o un *elemento primo*) si siempre que $a = bc$ con b y c , ambos, de R se tiene que uno de ambos b o c es una unidad en R . Un elemento irreducible es, pues, un elemento que no puede ser factorizado de modo “no trivial”.

DEFINICIÓN. Un dominio entero R con elemento unidad es un *dominio de factorización única* si

- cualquier elemento distinto de cero de R es una unidad o puede escribirse como el producto de un número finito de elementos irreducibles de R ;
- la descomposición de la parte (a) es única salvo el orden y asociación de los elementos irreducibles.

El teorema 3.e afirma que un anillo eucliano es un dominio de factorización única. Lo recíproco, sin embargo, es falso; por ejemplo, el anillo $F[x_1, x_2]$ donde F es un campo no es ni siquiera un anillo de ideales principales (de donde se deduce que no puede ser un anillo eucliano), pero como veremos pronto es un dominio de factorización única.

En los anillos conmutativos generales podemos hablar de los máximos comunes divisores de elementos; la principal dificultad es que éstos, en general, tal vez no existan. Pero en los dominios de factorización única su existencia es segura. Este hecho no es difícil de probar y lo dejamos como un ejercicio; igualmente fáciles de probar son las otras partes del siguiente

LEMA 3.25. Si R es un dominio de factorización única y si a y b están en R , entonces a y b tienen un máximo común divisor (a, b) en R . Además, si a y b son primos relativos (es decir, si $(a, b) = 1$), entonces siempre que $a|bc$ se tiene $a|c$.

COROLARIO. Si $a \in R$ es un elemento irreducible y $a|bc$, entonces $a|b$ o $a|c$.

Queremos ahora presentar la versión apropiada del lema de Gauss (teorema 3.i) que probamos para polinomios con coeficientes enteros, para el anillo $R[x]$, donde R es un dominio con factorización única.

Dado el polinomio $f(x) = a_0 + a_1 x + \dots + a_m x^m$ en $R[x]$, entonces el *contenido* de $f(x)$ es, por definición, el máximo común divisor de a_0, a_1, \dots, a_m . Es único, salvo unidades de R . Denotaremos al contenido

de $f(x)$ por $c(f)$. Un polinomio en $R[x]$ se dice que es *primitivo* si su contenido es 1 (es decir, es una unidad de R). Dado un polinomio cualquiera $f(x) \in R[x]$ podemos escribir $f(x) = af_1(x)$ donde $a = c(f)$ y donde $f_1(x) \in R[x]$ es primitivo. (¡Pruébese!) Excepto por multiplicación por unidades de R esta descomposición de $f(x)$, como un elemento de R por un polinomio primitivo en $R[x]$, es única. (¡Pruébese!)

La prueba del lema 3.23 puede aplicarse íntegramente a nuestra presente situación; el único cambio que debe hacerse en la prueba es reemplazar el número primo p por un elemento irreducible de R . Así pues, tenemos

LEMA 3.26. *Si R es un dominio de factorización única, entonces el producto de dos polinomios primitivos en $R[x]$ es también un polinomio primitivo en $R[x]$.*

Dados $f(x)$ y $g(x)$ en $R[x]$ podemos escribir $f(x) = af_1(x)$, $g(x) = bg_1(x)$, donde $a = c(f)$ y $b = c(g)$ y donde $f_1(x)$ y $g_1(x)$ son primitivos. Tenemos pues: $f(x)g(x) = abf_1(x)g_1(x)$. Según el lema 3.26, $f_1(x)g_1(x)$ es primitivo. Luego el contenido del producto $f(x)g(x)$ es ab , es decir, $c(f)c(g)$. Y hemos probado el siguiente

COROLARIO. *Si R es un dominio de factorización única y si $f(x)$ y $g(x)$ están en $R[x]$, entonces $c(fg) = c(f)c(g)$ (salvo unidades).*

Por una sencilla inducción el corolario se extiende al producto de un número finito de polinomios. Es decir, se tiene: $c(f_1f_2\dots f_k) = c(f_1)c(f_2)\dots c(f_k)$.

Sea R un dominio de factorización única. Por ser un dominio entero, de acuerdo con el teorema 3.c, tiene un campo de cocientes F . Podemos considerar $R[x]$ como un subanillo de $F[x]$. Dado un polinomio cualquiera $f(x) \in F[x]$, entonces $f(x) = (f_0(x)/a)$, donde $f_0(x) \in R[x]$ y $a \in R$. (¡Pruébese!) Es natural preguntar por la relación, en términos de reducibilidad e irreducibilidad, de un polinomio en $R[x]$ considerado como un polinomio en el anillo más grande $F[x]$.

LEMA 3.27. *Si $f(x)$ en $R[x]$ es a la vez prima e irreducible como un elemento de $R[x]$, entonces es irreducible como un elemento de $F[x]$. Recíprocamente, si el elemento primitivo $f(x)$ en $R[x]$ es irreducible como un elemento de $F[x]$, es también irreducible como un elemento de $R[x]$.*

Prueba. Supongamos que el elemento primitivo $f(x)$ en $R[x]$ es irreducible en $R[x]$, pero es reducible en $F[x]$. Así pues, $f(x) = g(x)h(x)$, donde $g(x)$ y $h(x)$ están en $F[x]$ y son de grado positivo. Pero ahora $g(x) = (g_0(x)/a)$ y $h(x) = (h_0(x)/b)$, donde a y b pertenecen a R y donde $g_0(x), h_0(x) \in R[x]$. Además, $g_0(x) = \alpha g_1(x)$, $h_0(x) = \beta h_1(x)$, donde

$\alpha = c(g_0)$, $\beta = c(h_0)$, y $g_1(x)$, $h_1(x)$ son primas en $R[x]$. Tenemos pues que $f(x) = (\alpha\beta/ab)g_1(x)h_1(x)$, de donde $abf(x) = \alpha\beta g_1(x)h_1(x)$. Por el lema 3.26, $g_1(x)h_1(x)$ es primo, de donde el contenido del segundo miembro es $\alpha\beta$. Como $f(x)$ es primo, el contenido del primer miembro es ab ; pero entonces $ab = \alpha\beta$; la implicación de esto es que $f(x) = g_1(x)h_1(x)$, con lo que hemos obtenido una factorización no trivial de $f(x)$ en $R[x]$, en contrario con lo que supusimos en la hipótesis. (Nota: esta factorización es no trivial ya que $g_1(x)$ y $g(x)$, y $h_1(x)$ y $h(x)$, son del mismo grado, luego no pueden ser unidades en $R[x]$ (véase problema 4).) Dejamos la prueba de la parte recíproca del teorema como un ejercicio.

LEMA 3.28. *Si R es un dominio de factorización única y si $p(x)$ es un polinomio primo en $R[x]$, entonces puede ser factorizado en una forma única como el producto de elementos irreducibles en $R[x]$.*

Prueba. Cuando consideramos $p(x)$ como un elemento en $F[x]$, por el lema 3.21, podemos factorizarlo como $p(x) = p_1(x)\dots p_k(x)$ donde $p_1(x), p_2(x), \dots, p_k(x)$ son polinomios irreducibles en $F[x]$. Cada $p_i(x) = (f_i(x)/a_i)$, donde $f_i(x) \in R[x]$ y $a_i \in R$; además, $f_i(x) = c_i q_i(x)$, donde $c_i = c(f_i)$ y donde $q_i(x)$ es primo en $R[x]$. Así p es, para toda i , $p_i(x) = (c_i q_i(x)/a_i)$, donde $a_i, c_i \in R$ y donde $q_i(x) \in R[x]$ es primo. Como $p_i(x)$ es irreducible en $F[x]$, $q_i(x)$ debe también ser irreducible en $F[x]$, de donde, por el lema 3.27, es irreducible en $R[x]$.

Tenemos ahora

$$p(x) = p_1(x) \cdots p_k(x) = \frac{c_1 c_2 \cdots c_k}{a_1 a_2 \cdots a_k} q_1(x) \cdots q_k(x),$$

de donde $a_1 a_2 \cdots a_k p(x) = c_1 c_2 \cdots c_k q_1(x) \cdots q_k(x)$. Usando la primalidad de $p(x)$ y de $q_1(x) \cdots q_k(x)$, podemos considerar como contenido del primer miembro a $a_1 a_2 \cdots a_k$, y como contenido del segundo miembro a $c_1 c_2 \cdots c_k$, luego $a_1 a_2 \cdots a_k = c_1 c_2 \cdots c_k$, de donde $p(x) = q_1(x) \cdots q_k(x)$. Hemos factorizado $p(x)$, en $R[x]$, como un producto de elementos irreducibles.*

¿Podemos factorizarlo de otro modo? Si $p(x) = r_1(x) \cdots r_k(x)$, donde los $r_i(x)$ son irreducibles en $R[x]$, por la primalidad de $p(x)$, cada $r_i(x)$ debe ser primo, de donde irreducible en $F[x]$ por el lema 3.27. Pero por el lema 3.21 sabemos que la factorización en $F[x]$ es única; el resultado neto de esto es que las $r_i(x)$ y las $q_i(x)$ son iguales (salvo asociación) en algún orden, de donde $p(x)$ tiene una factorización única como producto de irreducibles en $R[x]$.

Tenemos ahora toda la información necesaria para probar el principal teorema de esta sección.

*Realmente, lo que tenemos es: $c_1 c_2 \cdots c_k = u a_1 a_2 \cdots a_k$, con u una unidad. Luego $p(x) = (u q_1(x)) \cdots q_k(x)$. Hay que probar que u es de R . (N. del T.)

TEOREMA 3.K. Si R es un dominio de factorización única, entonces también lo es $R[x]$.

Prueba. Sea $f(x)$ un elemento arbitrario en $R[x]$. Podemos escribir $f(x)$ en forma única como $f(x) = cf_1(x)$ donde $c = c(f)$ está en R y $f_1(x)$ en $R[x]$, es primitivo. Por el lema 3.28, podemos descomponer $f_1(x)$ en forma única como producto de elementos irreducibles de $R[x]$. ¿Y qué hay acerca de c ? Supongamos que $c = a_1(x)a_2(x)\dots a_m(x)$ en $R[x]$; entonces $0 = \deg c = \deg(a_1(x)) + \deg(a_2(x)) + \dots + \deg(a_m(x))$. Por tanto, cada $a_i(x)$ debe ser de grado 0, es decir, debe ser un elemento de R . En otras palabras, las solas factorizaciones de c como un elemento de $R[x]$ son las que tienen como elemento de R . En particular, un elemento irreducible en R lo es también en $R[x]$. Como R es un dominio de factorización única, c tiene una factorización única como producto de elementos irreducibles de R , de donde $R[x]$.

Si juntamos la factorización única de $f(x)$ en la forma $cf_1(x)$, donde $f_1(x)$ es primitivo y $c \in R$, con la factorización única de c y de $f_1(x)$, hemos probado el teorema.

Dado R como un dominio de factorización única, entonces $R_1 = R[x_1]$ es también un dominio de factorización única. Luego $R_2 = R_1[x_2] = R[x_1, x_2]$ es también un dominio de factorización única. Continuando con este esquema obtenemos

COROLARIO 1. Si R es un dominio de factorización única entonces también lo es $R[x_1, \dots, x_n]$.

Un caso especial del corolario 1, pero de interés e importancia independientes es el

COROLARIO 2. Si F es un campo, entonces $F[x_1, \dots, x_n]$ es un dominio de factorización única.

Problemas

1. Pruébese que $R[x]$ es un anillo comutativo con elemento unidad siempre que R lo sea.

2. Pruébese que $R[x_1, \dots, x_n] = R[x_{i_1}, \dots, x_{i_n}]$ donde (i_1, \dots, i_n) es una permutación de $(1, 2, \dots, n)$.

3. Si R es un dominio entero, pruébese que para $f(x)$ y $g(x)$ en $R[x]$, $\deg(f(x), g(x)) = \deg(f(x)) + \deg(g(x))$.

4. Si R es un dominio entero con elemento unitario, pruébese que cualquier unidad en $R[x]$ debe ser también una unidad en R .

5. Sea R un anillo conmutativo sin elementos *nilpotentes* no nulos, es decir, tal que $a^n = 0$ implique $a = 0$. Si $f(x) = a_0 + a_1x + \dots + a_nx^n$ en $R[x]$ es un divisor de cero, pruébese que hay un elemento $b \neq 0$ en R tal que $ba_0 = ba_1 = \dots = ba_n = 0$.

*6. Resuélvase el problema 5 prescindiendo de la hipótesis de que R no tiene elementos nilpotentes distintos del cero.

*7. Si R es un anillo conmutativo con elemento unitario, pruébese que $a_0 + a_1x + \dots + a_nx^n$ en $R[x]$ tiene un inverso en $R[x]$ (es decir, es una unidad en $R[x]$) si y sólo si a_0 es una unidad en R y a_1, \dots, a_n son elementos nilpotentes en R .

8. Pruébese que cuando F es un campo, $F[x_1, x_2]$ no es un anillo de ideales principales.

9. Pruébese en forma completa el lema 3.25 y su corolario.

10. a) Si R es un dominio de factorización única, pruébese que todo $f(x) \in R[x]$ puede escribirse como $f(x) = af_1(x)$, donde $a \in R$ y $f_1(x)$ es un primitivo.

b) Pruébese que la descomposición de la parte (a) es única (hasta asociados).

11. Si R es un dominio entero, y si F es su campo de cocientes, pruébese que cualquier elemento $f(x)$ en $F[x]$ puede escribirse como $f(x) = (f_0(x)/a)$, donde $f_0(x) \in R[x]$ y $a \in R$.

12. Pruébese la "parte recíproca" del lema 3.27.

13. Pruébese el corolario 2 al teorema 3.k.

14. Pruébese que un anillo de ideales principales es un dominio de factorización única.

15. Si J es el anillo de los enteros, pruébese que $J[x_1, \dots, x_n]$ es un dominio de factorización única.

Problemas supplementarios

1. Sea R un anillo conmutativo; un ideal P de R se dice que es un *ideal primo* de R si $ab \in P$, $a, b \in R$ implica que $a \in P$ o $b \in P$. Pruébese que P es un ideal primo de R si y sólo si P/R es un dominio entero.

2. Sea R un anillo conmutativo con elemento unidad; pruébese que todo ideal máximo de R es un ideal primo.

3. Proporcionese un ejemplo de un anillo en el que algún ideal primo no sea un ideal máximo.

4. Si R es un anillo conmutativo finito (es decir, que tiene solamente un número finito de elementos) con elemento unidad, pruébese que todo ideal primo de R es un ideal máximo de R .

5. Si F es un campo, pruébese que $F[x]$ es isomorfo a $F[t]$.

6. Encuéntrense todos los automorfismos σ de $F[x]$ con la propiedad de que $\sigma(f) = f$ para todo $f \in F$.

7. Si R es un anillo conmutativo, sea $N = \{x \in R \mid x^n = 0 \text{ para algún entero } n\}$. Pruébese que:

a) N es un ideal de R .

b) En $\bar{R} = R/N$, si $\bar{x}^m = 0$ para algún m , entonces $\bar{x} = 0$.

8. Sea R un anillo conmutativo y supongamos que A es un ideal de R . Sea $N(A) = \{x \in R \mid x^n \in A \text{ para algún } n\}$. Pruébese que:

a) $N(A)$ es un ideal de R que contiene a A .

b) $N(N(A)) = N(A)$.

9. Si n es un entero, sea J_n el anillo de los enteros módulo n . Describábase N (véase el problema 7) para J_n en términos de n .

10. Si A y B son ideales en un anillo R tales que $A \cap B = (0)$, pruébese que para cualesquiera $a \in A$ y $b \in B$, $ab = 0$.

11. Si R es un anillo, sea $Z(R) = \{x \in R \mid xy = yx \text{ para todo } y \in R\}$. Pruébese que $Z(R)$ es un subanillo de R .

12. Si R es un anillo con división, pruébese que $Z(R)$ es un campo.

13. Encuéntrese un polinomio de grado 3 irreducible sobre el anillo de los enteros módulo 3, J_3 . Úsese para construir un campo con 27 elementos.

14. Constrúyase un campo con 625 elementos.

15. Si F es un campo y $p(x) \in F[x]$, pruébese que en el anillo

$$R = \frac{F[x]}{(p(x))},$$

N (véase el problema 7) es (0) si y sólo si $p(x)$ no es divisible por el cuadrado de ningún polinomio.

16. Pruébese que el polinomio $f(x) = 1 + x + x^3 + x^4$ no es irreducible sobre ningún campo F .

17. Pruébese que el polinomio $f(x) = x^4 + 2x + 2$ es irreducible sobre el campo de los números racionales.

18. Pruébese que si F es un campo finito, su característica debe ser un número primo p y F contiene p^n elementos para algún entero. Pruébese, además, que si $a \in F$ entonces $a^{p^n} = a$.

19. Pruébese que cualquier ideal distinto de cero en los enteros gaussianos $J[i]$ debe contener algún entero positivo.

20. Pruébese que si R es un anillo en el que $a^4 = a$ para todo $a \in R$ entonces R debe ser conmutativo.

21. Sean R y R' y ϕ una aplicación de R en R' que satisface:

- a) $\phi(x+y) = \phi(x)+\phi(y)$ para todo $x, y \in R$.
- b) $\phi(xy) = \phi(x)\phi(y)$ o $\phi(y)\phi(x)$.

Pruébese que para todo $a, b \in R$, $\phi(ab) = \phi(a)\phi(b)$ o que para todo $a, b \in R$, $\phi(a) = \phi(b)\phi(a)$. (*Sugerencia:* Si $a \in R$, hágase $W_a = \{x \in R \mid \phi(ax) = \phi(a)\phi(x)\}$ y $U_a = \{x \in R \mid \phi(ax) = \phi(x)\phi(a)\}$.)

Lecturas suplementarias

ZARISKI, OSCAR, y SAMUEL, PIERRE, *Commutative Algebra*, vol. 1. D. Van Nostrand Company, Inc., Princeton, Nueva Jersey, 1958.

McCoy, N. H., *Rings and Ideals*, Carus Monograph Series, No. 8. Open Court Publishing Company, La Salle, Illinois, 1948.

Tópicos para discusión en clase

MOTZKIN, T., "The Euclidean algorithm", *Bulletin of the American Mathematical Society*, vol. 55 (1949), págs. 1142-1146.

CAPITULO

4
+

Espacios vectoriales y módulos

HASTA EL momento hemos introducido y estudiado los grupos y los anillos; los primeros tienen su motivación en el conjunto de las aplicaciones biyectivas de un conjunto sobre sí mismo, los últimos, en el conjunto de los enteros. El tercer modelo algebraico que vamos ahora a considerar —el de espacio vectorial— puede, en gran parte, considerarse que tiene sus orígenes en tópicos de la geometría y de la física.

Su descripción será reminiscente de las de los grupos y los anillos —en realidad, parte de su estructura es la de un grupo abeliano— pero un espacio vectorial difiere de estas dos estructuras previas en que uno de los productos en él definidos, usa elementos que son ajenos a él mismo. Estas observaciones se entenderán plenamente cuando demos la definición de espacio vectorial.

Los espacios vectoriales deben su importancia al hecho de que muchos modelos entre los que surgen como soluciones de problemas específicos, resultan ser espacios vectoriales. Es por esa razón por la que los conceptos básicos que introducimos en ellos tienen una cierta universalidad y los hemos encontrado, y los seguiremos encontrando, en contextos muy diversos. Entre estas nociones fundamentales están las de dependencia lineal, base y dimensión, las que desarrollaremos en este capítulo. Son, éstas, argumentos poderosos y efectivos en todas las ramas de las matemáticas; nosotros haremos un uso inmediato y libre de ellas en muchas partes esenciales del capítulo 5 que trata de teoría de campos.

Intimamente entrelazados con los espacios vectoriales están los homomorfismos de un espacio vectorial en otro (o en sí mismo). Estos homomorfismos constituirán la mayor parte del material que estudiaremos en el capítulo 6.

En la última parte del presente capítulo estudiaremos los módulos, una generalización de los espacios vectoriales; hablando a *grossó modo*, un módulo es un espacio vectorial sobre un anillo en vez de sobre un campo. Para módulos finitamente generados sobre anillos euclidianos probaremos el teorema fundamental para bases. Este resultado nos permite dar una descripción completa y la construcción de todos los grupos abelianos finitos.

1. CONCEPTOS BÁSICOS ELEMENTALES

DEFINICIÓN. Un conjunto no vacío V se dice que es un *espacio vectorial* sobre un campo F si V es un grupo abeliano respecto a una operación que denotamos por $+$, y si para todo $\alpha \in F$, $v \in V$ está definido un elemento, escrito como αv , de V , con las siguientes propiedades:

- 1) $\alpha(v + w) = \alpha v + \alpha w$
- 2) $(\alpha + \beta)v = \alpha v + \beta v$
- 3) $\alpha(\beta v) = (\alpha\beta)v$
- 4) $1v = v$

para cualesquiera $\alpha, \beta \in F$ y $v, w \in V$ (donde el 1 representa el elemento unitario de F en la multiplicación).

Nótese que en el axioma (1) anterior, el $+$ es el de V , mientras que en el axioma (2), en el primer miembro el $+$ es el de F y en el segundo miembro, el de V .

Usaremos sistemáticamente las siguientes notaciones:

- a) F representará un campo.
- b) Las letras griegas minúsculas serán elementos de F ; nos referiremos con frecuencia a los elementos de F como *escalares*.

- c) Las letras latinas mayúsculas representarán espacios vectoriales sobre F .
- d) Las letras latinas minúsculas denotarán elementos de espacios vectoriales. A los elementos de un espacio vectorial les llamaremos con frecuencia *vectores*.

Si ignoramos el hecho de que V tiene dos operaciones definidas sobre él y lo vemos por un momento simplemente como un grupo abeliano bajo $+$, el axioma (1) no afirma más que el hecho de que la multiplicación de los elementos de V por un escalar fijo α define un homomorfismo del grupo abeliano V en sí mismo. De acuerdo con el lema 4.1, que en seguida presentaremos, puede verse que si $\alpha \neq 0$ este homomorfismo es un isomorfismo de V sobre V .

Esto sugiere que muchos aspectos de la teoría de espacios vectoriales (y también de la de anillos) podrían haber sido desarrollados como una parte de la teoría de grupos, si hubiésemos generalizado la noción de grupo a la de *grupo con operadores*. Para los estudiantes que están familiarizados con el álgebra abstracta, éste es el punto de vista preferido; como no suponemos familiaridad alguna por parte del lector con el álgebra abstracta, creemos que tal enfoque podía dar lugar a una introducción excesivamente rápida a las ideas del tema con ninguna experiencia que actúe como guía.

EJEMPLO 1. Sea F un campo y K un campo que contiene a F como subcampo. Consideramos K como un espacio vectorial sobre F usando como el $+$ del espacio vectorial la adición de elementos de K , y definiendo para $\alpha \in F, v \in K, \alpha v$ como el producto de α y v como elementos en el campo K . Los axiomas (1), (2), (3) para un espacio vectorial son entonces consecuencias de la ley distributiva derecha, de la ley distributiva izquierda y de la ley asociativa, respectivamente, que valen para K por ser un anillo.

EJEMPLO 2. Sea F un campo y sea V la totalidad de todos los n -tuples, $(\alpha_1, \dots, \alpha_n)$ donde todos los $\alpha_i \in F$. Dos elementos $(\alpha_1, \dots, \alpha_n)$ y $(\beta_1, \dots, \beta_n)$ de V se definen como iguales si y sólo si $\alpha_i = \beta_i$ para todo $i = 1, 2, \dots, n$. Introducimos ahora las operaciones requeridas en V para hacer de él un espacio vectorial definiendo:

- 1) $(\alpha_1, \dots, \alpha_n) + (\beta_1, \dots, \beta_n) = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \dots, \alpha_n + \beta_n)$
- 2) $y(\alpha_1, \dots, \alpha_n) = (y\alpha_1, \dots, y\alpha_n)$ para $y \in F$.

Es fácil verificar que con estas operaciones V es un espacio vectorial sobre F . Como tal espacio vectorial aparecerá con frecuencia le asignaremos un símbolo, a saber, $F^{(n)}$.

EJEMPLO 3. Sea F un campo cualquiera y sea $V = F[x]$ el conjunto de todos los polinomios en x sobre F . Pretendemos ignorar, por el momento,

el hecho de que en $F[x]$ podemos multiplicar dos cualesquiera de sus elementos y nos concentraremos solamente en el hecho de que dos polinomios pueden sumarse y de que un polinomio puede siempre multiplicarse por un elemento de F . Con estas operaciones naturales, $F[x]$ es un espacio vectorial sobre F .

EJEMPLO 4. Sea V_n el conjunto de todos los polinomios de $F[x]$ de grado menor que n . Usando las operaciones habituales para polinomios de suma y multiplicación, V_n es un espacio vectorial sobre F .

¿Cuál es la relación entre los ejemplos 4 y 2? Cualquier elemento de V_n es de la forma $\alpha_0 + \alpha_1 x + \dots + \alpha_{n-1} x^{n-1}$, donde $\alpha_i \in F$; si transformamos este elemento en el elemento $(\alpha_0, \alpha_1, \dots, \alpha_{n-1})$ en $F^{(n)}$ podríamos razonablemente esperar, una vez que el homomorfismo y el isomorfismo hayan sido definidos, que encontremos que V_n y $F^{(n)}$ sean isomorfos como espacios vectoriales.

DEFINICIÓN. Si V es un espacio vectorial sobre F y si $W \subset V$, entonces W es un *subespacio* de V si bajo las operaciones de V , W mismo forma un espacio vectorial sobre F . Es decir, W es un subespacio de V siempre que $w_1, w_2 \in W, \alpha, \beta \in F$ implique que $\alpha w_1 + \beta w_2 \in W$.

Nótese que el espacio vectorial definido en el ejemplo 4 es un subespacio del definido en el ejemplo 3. Ejemplos adicionales de espacios vectoriales y de subespacios pueden encontrarse en los problemas al final de esta sección.

DEFINICIÓN. Si U y V son espacios vectoriales sobre F , entonces la aplicación T de U en V se dice que es un *homomorfismo* si

- 1) $(u_1 + u_2)T = u_1 T + u_2 T$
- 2) $(\alpha u_1)T = \alpha(u_1 T)$

para cualesquiera $u_1, u_2 \in U$, y $\alpha \in F$.

Como en nuestros previos modelos, un homomorfismo es una aplicación que preserva toda la estructura algebraica de nuestro sistema.

Si T es, además, inyectiva le llamamos *isomorfismo*. El núcleo de T es, por definición, $\{u \in U | uT = 0\}$ donde 0 es el elemento identidad de la adición en V . Es un ejercicio probar que el núcleo de T es un subespacio de U y que T es un isomorfismo si y sólo si su núcleo es $\{0\}$. Dos espacios vectoriales se dice que son *isomorfos* si hay un isomorfismo de uno sobre el otro.

El conjunto de todos los homomorfismos de U en V se representará por $\text{Hom}(U, V)$. Son de particular interés para nosotros dos casos especiales, $\text{Hom}(U, F)$ y $\text{Hom}(U, U)$. Estudiaremos muy pronto el primero de éstos; el segundo, que puede mostrarse que es un anillo, se llama *anillo de las transformaciones lineales de U* . Más adelante, en este mismo libro, dedicaremos mucho de nuestro tiempo a un estudio detallado de $\text{Hom}(U, U)$.

Propiamente comenzamos el estudio del tema con un lema operacional que, como en el caso de los anillos, nos permite efectuar ciertos cálculos naturales y sencillos en los espacios vectoriales. En el enunciado del lema, 0 representa el cero de la adición en V , o el de la adición en F , y $-v$ el inverso aditivo del elemento v de V .

LEMÁ 4.1. *Si V es un espacio vectorial sobre F , entonces*

- 1) $\alpha 0 = 0$ para $\alpha \in F$.
- 2) $ov = 0$ para $v \in V$.
- 3) $(-\alpha)v = -(\alpha v)$ para $\alpha \in F, v \in V$.
- 4) Si $v \neq 0$ entonces $\alpha v = 0$ implica que $\alpha = 0$.

Prueba. La prueba es muy fácil y sigue los mismos lineamientos que la prueba del resultado análogo para anillos; la damos, por tal razón, brevemente y con pocas explicaciones.

- 1) Como $\alpha 0 = \alpha(0+0) = \alpha 0 + \alpha 0$, entonces $\alpha 0 = 0$.
- 2) Como $ov = (o+o)v = ov + ov$, entonces $ov = 0$.
- 3) Como $0 = (\alpha + (-\alpha))v = \alpha v + (-\alpha)v$, $(-\alpha)v = -(\alpha v)$.
- 4) Si $\alpha v = 0$ y $\alpha \neq 0$, entonces

$$0 = \alpha^{-1}0 = \alpha^{-1}(\alpha v) = (\alpha^{-1}\alpha)v = 1v = v.$$

El lema que acabamos de probar muestra que la multiplicación por el cero de V o por el cero de F siempre da como resultado el cero de V . No habrá, pues, ningún peligro de confusión porque usemos el mismo símbolo para ambos, y eso es lo que haremos de aquí en adelante. Usaremos el símbolo 0 tanto para el cero de V como para el cero de F .

Sea V un espacio vectorial sobre F y sea W un subespacio de V . Considerando V y W simplemente como grupos abelianos, construimos el grupo cociente V/W ; sus elementos son las clases laterales $v+W$ donde $v \in V$. La commutatividad de la adición nos asegura, de acuerdo a lo que hemos visto en el capítulo 2 sobre teoría de grupos, que V/W es un grupo abeliano. Vamos a intentar hacer de él un espacio vectorial. Si $\alpha \in F$ y $v+W \in V/W$, definimos $\alpha(v+W) = \alpha v + W$. Como es usual, probamos primero que este producto está bien definido, es decir, que si $v+W = v'+W$, entonces $\alpha(v+W) = \alpha(v'+W)$. Ahora bien, como $v+W = v'+W$, $v-v' \in W$; como W es un subespacio, $\alpha(v-v')$ debe estar también en W . Usando la parte (3) del lema 4.1 (véase el problema 1) esto nos dice que $\alpha v - \alpha v' \in W$ y por tanto que $\alpha v + W = \alpha v' + W$. Por tanto, $\alpha(v+W) = \alpha v + W = \alpha v' + W = \alpha(v'+W)$; con lo que el producto se ha probado que está bien definido. La verificación de los axiomas característicos de los espacios vectoriales para V/W es rutinaria y la dejamos como ejercicio. Y hemos probado el

LEMÁ 4.2. Si V es un espacio vectorial sobre F y si W es un subespacio de V , entonces V/W es un espacio vectorial sobre F , donde para $v_1 + W$, $v_2 + W \in V/W$ y $\alpha \in F$

- 1) $(v_1 + W) + (v_2 + W) = (v_1 + v_2) + W$.
- 2) $\alpha(v_1 + W) = \alpha v_1 + W$.

A V/W se le suele llamar el *espacio cociente* de V por W .

Sin más preliminares formulamos ahora el primer teorema de homomorfismo para espacios vectoriales; no damos prueba alguna sino que referimos al lector a la prueba del teorema 2.d.

TEOREMA 4.A. Si T es un homomorfismo de U sobre V de núcleo W , entonces V es isomorfo a U/W . Recíprocamente, si U es un espacio vectorial y W un subespacio de U entonces hay un homomorfismo de U sobre U/W .

Los otros teoremas sobre homomorfismo aparecerán como ejercicios al final de esta sección.

DEFINICIÓN. Sea V un espacio vectorial sobre F y sean U_1, \dots, U_n subespacios de V . Se dice entonces que V es la *suma directa interna* de U_1, \dots, U_n si todo elemento $v \in V$ puede escribirse de una forma y sólo de una forma como $v = u_1 + u_2 + \dots + u_n$ donde $u_i \in U_i$.

Dado un número finito cualquiera de espacios vectoriales sobre F , V_1, \dots, V_n , consideremos el conjunto V de todas las n -adas ordenadas (v_1, \dots, v_n) donde $v_i \in V_i$. Convenimos en que dos elementos (v_1, \dots, v_n) y (v'_1, \dots, v'_n) de V son iguales si y sólo si para cada i , $v_i = v'_i$. Sumamos dos elementos de tal tipo mediante la regla $(v_1, \dots, v_n) + (w_1, \dots, w_n) = (v_1 + w_1, \dots, v_n + w_n)$. Finalmente, si $\alpha \in F$ y $(v_1, \dots, v_n) \in V$ definimos $\alpha(v_1, \dots, v_n)$ como $(\alpha v_1, \dots, \alpha v_n)$. Comprobar que los axiomas característicos de los espacios vectoriales se cumplen para V con operaciones como las que para él acabamos de definir, no ofrece dificultad alguna. Así pues, el mismo V es un espacio vectorial sobre F . Llamamos a V la *suma directa externa* de V_1, \dots, V_n y lo denotamos escribiendo $V = V_1 \oplus \dots \oplus V_n$.

TEOREMA 4.B. Si V es la suma directa interna de U_1, \dots, U_n , entonces V es isomorfo a la suma directa externa de U_1, \dots, U_n .

Prueba. Dada $v \in V$, v puede escribirse, por hipótesis, de un modo y sólo de un modo en la forma $v = u_1 + u_2 + \dots + u_n$ donde $u_i \in U_i$; definamos la aplicación T de V en $U_1 \oplus \dots \oplus U_n$ por $vT = (u_1, \dots, u_n)$. Como v tiene una única representación de esta forma, T está bien definida. Es claramente suprayectiva, pues para un elemento arbitrario $(w_1, \dots, w_n) \in U_1 \oplus \dots \oplus U_n$ se tiene wT igual a él, con $w = w_1 + \dots + w_n$. Dejamos para el lector la prueba de que T es inyectiva y un homomorfismo.

En razón del isomorfismo probado en el teorema 4.b, nos referiremos de aquí en adelante tan sólo a sumas directas, sin especificar si tales sumas son interiores o exteriores.

Problemas

1. Demuéstrese que en todo espacio vectorial $\alpha(v-w) = \alpha v - \alpha w$.
2. Pruébese que los espacios vectoriales de los ejemplos 2 y 4 son isomorfos.
3. Pruébese que el núcleo de un homomorfismo es un subespacio.
4. a) Si F es el campo de los números reales pruébese que el conjunto de las funciones reales continuas sobre el intervalo cerrado $[0, 1]$ forma un espacio vectorial sobre F .
b) Pruébese que aquellas funciones de la parte (a) para las que existe la n -ésima derivada ($n = 1, 2, \dots$), forman un subespacio.
5. a) Sea F el campo de los números reales y sea V el conjunto de todas las sucesiones $(a_1, a_2, \dots, a_n, \dots)$, $a_i \in F$, donde la igualdad, la adición, y la multiplicación por los escalares se definen por componentes. Pruébese que V es un espacio vectorial sobre F .
b) Sea $W = \{(a_1, \dots, a_n, \dots) \in V \mid \lim_{n \rightarrow \infty} a_n = 0\}$. Pruébese que W es un subespacio de V .
*c) Sea $U = \{(a_1, \dots, a_n, \dots) \in V \mid \sum_{i=1}^{\infty} a_i^2 \text{ es finito}\}$. Pruébese que U es un subespacio de V y está contenido en W .
6. Si U y V son espacios vectoriales sobre F , definase una adición y una multiplicación por escalares en $\text{Hom}(U, V)$ a modo de hacer de $\text{Hom}(U, V)$ un espacio vectorial sobre F .
- *7. Usando el resultado del problema 6, pruébese que $\text{Hom}(F^{(n)}, F^{(m)})$ es isomorfo a $F^{(nm)}$ como espacio vectorial.
8. Si $n > m$, pruébese que hay un homomorfismo de $F^{(n)}$ sobre $F^{(m)}$ con un núcleo W que es isomorfo a $F^{(n-m)}$.
9. Si $v \neq 0 \in F^{(n)}$, pruébese que hay un elemento $T \in \text{Hom}(F^{(n)}, F)$ tal que $vT \neq 0$.
10. Pruébese que existe un isomorfismo de $F^{(n)}$ en $\text{Hom}(\text{Hom}(F^{(n)}, F), F)$.
11. Si U y W son subespacios de V , pruébese que $U + W = \{v \in V \mid v = u + w, u \in U, w \in W\}$ es un subespacio de V .

12. Pruébese que la intersección de dos subespacios de V es un subespacio de V .

13. Si A y B son subespacios de V , pruébese que $(A+B)/B$ es isomorfo a $A/(A \cap B)$.

14. Si T es un homomorfismo de U sobre V con núcleo W , pruébese que hay una correspondencia biyectiva entre los subespacios de V y los subespacios de U que contienen a W .

15. Sea V un espacio vectorial sobre F y sean V_1, \dots, V_n subespacios de V . Supongamos que $V = V_1 + V_2 + \dots + V_n$ (véase problema 11), y que $V_i \cap (V_1 + \dots + V_{i-1} + V_{i+1} + \dots + V_n) = (0)$ para todo $i = 1, 2, \dots, n$. Pruébese que V es la suma directa interna de V_1, \dots, V_n .

16. Sea $V = V_1 \oplus \dots \oplus V_n$; pruébese que en V hay subespacios V_i isomorfos a los V_i tales que V es la suma directa interna de los V_i .

17. Sea T definida sobre $F^{(2)}$ por $(x_1, x_2)T = (\alpha x_1 + \beta x_2, \gamma x_1 + \delta x_2)$ donde $\alpha, \beta, \gamma, \delta$ son ciertos elementos fijos de F .

a) Pruébese que T es un homomorfismo de $F^{(2)}$ en sí mismo.

b) Encuéntrense cuáles son las condiciones necesarias y suficientes que $\alpha, \beta, \gamma, \delta$ para que T sea un isomorfismo.

18. Supongamos a T definido sobre $F^{(3)}$ por $(x_1, x_2, x_3)T = (\alpha_{11}x_1 + \alpha_{12}x_2 + \alpha_{13}x_3, \alpha_{21}x_1 + \alpha_{22}x_2 + \alpha_{23}x_3, \alpha_{31}x_1 + \alpha_{32}x_2 + \alpha_{33}x_3)$. Muéstrese que T es un homomorfismo de $F^{(3)}$ en él mismo y determinense las condiciones necesarias y suficientes que han de satisfacer las α_{ij} para que T sea un isomorfismo.

19. Sea T un homomorfismo de V en W . Usando T , defínase un homomorfismo T^* de $\text{Hom}(W, F)$ en $\text{Hom}(V, F)$.

20. a) Pruébese que $F^{(1)}$ no es isomorfo a $F^{(n)}$ para $n > 1$.

b) Pruébese que $F^{(2)}$ no es isomorfo a $F^{(3)}$.

21. Si V es un espacio vectorial sobre un campo infinito F , pruébese que V no puede escribirse como unión de los conjuntos de un número finito de subespacios propios.

2. INDEPENDENCIA LINEAL Y BASES

Si observamos un poco más detenidamente dos de los ejemplos descritos en la sección previa, el ejemplo 4 y el ejemplo 3, notamos que aunque tienen muchas propiedades comunes, hay entre ellos una diferencia esencial. Consiste tal diferencia en el hecho de que en el primero podemos encontrar un número finito de elementos $1, x, x^2, \dots, x^{n-1}$ tales que todo elemento

puede escribirse como una combinación de ellos con coeficientes de F , mientras que en el último ningún conjunto finito de elementos de tal tipo existe.

Intentaremos ahora examinar con algún detalle, los espacios vectoriales que pueden generarse, como en el caso del ejemplo 4, por un conjunto finito de elementos.

DEFINICIÓN. Si V es un espacio vectorial sobre F y si $v_1, \dots, v_n \in V$, entonces cualquier elemento de la forma $\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n$, donde los $\alpha_i \in F$, se llama *combinación lineal* sobre F de v_1, \dots, v_n .

Como, en general, estamos trabajando con un cierto campo fijo F , diremos frecuentemente combinación lineal en lugar de combinación lineal sobre F . Análogamente, deberá entenderse que cuando hablamos de un espacio vectorial lo que queremos decir es un espacio vectorial sobre F .

DEFINICIÓN. Si S es un subconjunto no vacío del espacio vectorial V , entonces $L(S)$, el *espacio generado* por S , es el conjunto de todas las combinaciones lineales de conjuntos finitos de elementos de S .

Pusimos, después de todo, en $L(S)$ los elementos requeridos por los axiomas definitorios de un espacio vectorial, por lo que no es sorprendente que nos encontremos con que

LEMA 4.3. *$L(S)$ es un subespacio de V .*

Prueba. Si v y w están en $L(S)$, entonces $v = \lambda_1 s_1 + \dots + \lambda_n s_n$ y $w = \mu_1 t_1 + \dots + \mu_m t_m$, donde los λ y los μ están en F y los s_i y los t_i están todos en S . Por tanto, para $\alpha, \beta \in F$, $\alpha v + \beta w = \alpha(\lambda_1 s_1 + \dots + \lambda_n s_n) + \beta(\mu_1 t_1 + \dots + \mu_m t_m) = (\alpha\lambda_1)s_1 + \dots + (\alpha\lambda_n)s_n + (\beta\mu_1)t_1 + \dots + (\beta\mu_m)t_m$, y, por tanto, es también un elemento de $L(S)$. Y con ello hemos demostrado que $L(S)$ es un subespacio de V .

La prueba de cada una de las partes del siguiente lema es realmente sencilla y la dejamos como un ejercicio al lector.

LEMA 4.4. *Si S, T son subconjuntos de V , entonces*

- 1) $S \subset T$ implica $L(S) \subset L(T)$.
- 2) $L(S \cup T) = L(S) + L(T)$.
- 3) $L(L(S)) = L(S)$.

DEFINICIÓN. El espacio vectorial V se dice que es *de dimensión finita* (sobre F) si existe un subconjunto finito S en V tal que $V = L(S)$.

Adviértase que $F^{(n)}$ es de dimensión finita sobre F , pues para S consistente en el conjunto de los vectores $(1, 0, \dots, 0)$, $(0, 1, 0, \dots, 0)$, ..., $(0, 0, \dots, 0, 1)$ se tiene: $V = L(S)$.

Aunque ya hemos definido lo que entendemos por un espacio de dimensión finita, aún no hemos definido lo que se entiende por la dimensión de un espacio. Tal definición aparecerá dentro de poco.

DEFINICIÓN. Si V es un espacio vectorial y si v_1, \dots, v_n están en V , decimos que tales vectores son *linealmente dependientes* sobre F si existen elementos $\lambda_1, \dots, \lambda_n$ en F , no todos cero, tales que $\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n = 0$.

Si los vectores v_1, \dots, v_n no son linealmente dependientes sobre F , se dice que son *linealmente independientes* sobre F . También aquí acortaremos a menudo la expresión “linealmente dependientes sobre F ” reemplazándola por “linealmente dependientes”. Nótese que si v_1, \dots, v_n son linealmente independientes, entonces ninguno de ellos puede ser 0, pues si, por ejemplo, $v_1 = 0$, entonces $\alpha v_1 + 0v_2 + \dots + 0v_n = 0$ para cualquier $\alpha \neq 0$ en F .

En $F^{(3)}$ es fácil verificar que $(1, 0, 0)$, $(0, 1, 0)$, y $(0, 0, 1)$ son linealmente independientes, mientras que $(1, 1, 0)$, $(3, 1, 3)$ y $(5, 3, 3)$ son linealmente dependientes.

Señalemos que la dependencia lineal no depende solamente de los vectores, sino también del campo. Por ejemplo, el campo de los números complejos es un espacio vectorial sobre el campo de los números reales y también un espacio vectorial sobre el propio campo de los números complejos. Los elementos $v_1 = 1$, $v_2 = i$ son linealmente independientes sobre el campo de los números reales, pero linealmente dependientes sobre el de los complejos, ya que $iv_1 + (-1)v_2 = 0$.

El concepto de dependencia lineal es absolutamente básico y muy importante. Pasemos a ver algunas de sus propiedades.

LEMA 4.5. Si $v_1, \dots, v_n \in V$ son linealmente independientes, entonces todo elemento del subespacio por ellos generado tiene una representación única de la forma $\lambda_1 v_1 + \dots + \lambda_n v_n$, con los $\lambda_i \in F$.

Prueba. Por definición, todo elemento del espacio generado por ellos es de la forma $\lambda_1 v_1 + \dots + \lambda_n v_n$. Para mostrar la unicidad, debemos probar que si $\lambda_1 v_1 + \dots + \lambda_n v_n = \mu_1 v_1 + \dots + \mu_n v_n$, entonces $\lambda_1 = \mu_1$, $\lambda_2 = \mu_2$, ..., $\lambda_n = \mu_n$. Pero si $\lambda_1 v_1 + \dots + \lambda_n v_n = \mu_1 v_1 + \dots + \mu_n v_n$, entonces ciertamente tenemos $(\lambda_1 - \mu_1)v_1 + (\lambda_2 - \mu_2)v_2 + \dots + (\lambda_n - \mu_n)v_n = 0$, lo que por la independencia lineal de v_1, \dots, v_n trae como consecuencia obligada $\lambda_1 - \mu_1 = 0$, $\lambda_2 - \mu_2 = 0$, ..., $\lambda_n - \mu_n = 0$.

El próximo teorema, aunque muy fácil y a primera vista, de naturaleza en cierta forma técnica, tiene como consecuencia resultados que constituyen los verdaderos fundamentos del tema. Enumeraremos algunos de ellos

como corolarios; los otros aparecerán en la sucesión de lemas y teoremas que a continuación presentaremos.

TEOREMA 4.c. *Si v_1, \dots, v_n están en V , entonces o son linealmente independientes o algún v_k es una combinación lineal de los que le preceden, v_1, \dots, v_{k-1} .*

Prueba. Si v_1, \dots, v_n son linealmente independientes, no hay, naturalmente, nada que probar. Supongamos entonces que $\alpha_1 v_1 + \dots + \alpha_n v_n = 0$ expresión en la que no todos los α son 0. Sea k el mayor entero para el que $\alpha_k \neq 0$. Como $\alpha_i = 0$ para $i > k$, $\alpha_1 v_1 + \dots + \alpha_k v_k = 0$, lo que, como $\alpha_k \neq 0$, implica que $v_k = \alpha_k^{-1}(-\alpha_1 v_1 - \alpha_2 v_2 - \dots - \alpha_{k-1} v_{k-1}) = (-\alpha_k^{-1} \alpha_1) v_1 + \dots + (-\alpha_k^{-1} \alpha_{k-1}) v_{k-1}$. Luego v_k es una combinación lineal de sus predecesores.

COROLARIO 1. *Si v_1, \dots, v_n de V generan a W y si v_1, \dots, v_k son linealmente independientes, entonces podemos encontrar un subconjunto de v_1, \dots, v_n de la forma $v_1, \dots, v_k, v_{i_1}, \dots, v_{i_r}$ consistente de elementos linealmente independientes que generen también a W .*

Prueba. Si v_1, \dots, v_n son linealmente independientes, la afirmación del teorema está probada. En caso contrario, saquemos de este conjunto la primera v_j que sea una combinación lineal de los elementos que la preceden. Como v_1, \dots, v_k son linealmente independientes, $j > k$. El subconjunto así construido, $v_1, \dots, v_k, \dots, v_{j-1}, v_{j+1}, \dots, v_n$, tiene $n-1$ elementos. Es claro que el subespacio que generan está contenido en W . Pero afirmamos que es igual a W , pues para toda $w \in W$, w puede escribirse como una combinación lineal de v_1, \dots, v_n , pero en esta combinación lineal podemos reemplazar v_j por una combinación lineal de v_1, \dots, v_{j-1} . Por lo tanto, w es una combinación lineal de $v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_n$.

Continuando una y otra vez con este proceso, llegamos a un subconjunto $v_1, \dots, v_k, v_{i_1}, \dots, v_{i_r}$ que generan aún W , pero en el que no hay elemento alguno que sea una combinación lineal de los que le preceden. Por el teorema 4.c, los elementos $v_1, \dots, v_k, v_{i_1}, \dots, v_{i_r}$ deben ser linealmente independientes.

COROLARIO 2. *Si V es un espacio de dimensión finita, entonces contiene un conjunto finito v_1, \dots, v_n de elementos linealmente independientes que generan V .*

Prueba. Como V es de dimensión finita está generada por un número finito de elementos u_1, \dots, u_m . Según el corolario 1 podemos encontrar un número finito de éstos, al que denotamos por v_1, \dots, v_n , consistente en elementos linealmente independientes y que generan a V .

DEFINICIÓN. Un subconjunto S de un espacio vectorial V se llama *base* de V si S consiste en elementos linealmente independientes (es decir, S es tal que cualquier número finito de elementos de S es un conjunto de vectores linealmente independiente) y $V = L(S)$.

Usando esta terminología, podemos reformular el Corolario 2 en la siguiente forma.

COROLARIO 3. *Si V es un espacio vectorial de dimensión finita y si u_1, \dots, u_m generan a V , entonces hay un cierto subconjunto de u_1, \dots, u_m que forma una base de V .*

El corolario 3 afirma que un espacio de dimensión finita posee una base que contiene a un número finito de elementos v_1, \dots, v_n . Junto con el lema 4.5 nos dice que todo elemento de V tiene una representación única de la forma $\alpha_1 v_1 + \dots + \alpha_n v_n$ con $\alpha_1, \dots, \alpha_n$ en F .

Veamos algunas de las implicaciones heurísticas de estas observaciones. Supongamos que V es un espacio de dimensión finita sobre F ; como acabamos de ver, V tiene entonces una base v_1, \dots, v_n . Por tanto, todo elemento v de V tiene una expresión única de la forma $v = \alpha_1 v_1 + \dots + \alpha_n v_n$. Transformemos V en $F^{(n)}$ definiendo como imagen de $\alpha_1 v_1 + \dots + \alpha_n v_n$ a $(\alpha_1, \dots, \alpha_n)$. Por la unicidad de la expresión en esta forma, la aplicación está bien definida, es biyectiva y suprayectiva; puede demostrarse que tiene todas las propiedades requeridas para ser un isomorfismo. Por tanto, V es isomorfo a $F^{(n)}$ para un cierto n , donde en realidad n es el número de elementos en alguna base de V sobre F . Si alguna otra base de V tuviera m elementos, por las mismas razones V sería isomorfo con $F^{(m)}$. Luego como tanto $F^{(n)}$ como $F^{(m)}$ serían isomorfos con V , habrían de ser también isomorfos uno con otro.

Se presenta entonces de forma natural un problema. ¿Bajo qué condiciones sobre n y m son $F^{(n)}$ y $F^{(m)}$ isomorfos? Nuestra intuición nos sugiere que esto solo puede suceder cuando $n = m$. ¿Por qué? Por una parte, si F fuera un campo con un número finito de elementos —por ejemplo, si $F = J_p$ (los enteros módulo el número primo p)— entonces $F^{(n)}$ tendría p^n elementos, mientras que $F^{(m)}$ tendría p^m . El isomorfismo implicaría que tienen el mismo número de elementos, luego habríamos de tener $n = m$. Por otra parte, si F fuera el campo de los números reales, entonces $F^{(n)}$ (en lo que al lector puede parecer una vaga forma geométrica de expresión) representa un n -espacio real, y nuestro sentido geométrico nos dice que un n -espacio es diferente de un m -espacio cuando $n \neq m$. Podemos pues esperar que si F es un campo cualquiera, entonces $F^{(m)}$ y $F^{(n)}$ son isomorfos solamente si $m = n$. Esto es equivalente, de acuerdo a nuestra discusión anterior, a afirmar que dos bases cualesquiera de V deben tener el mismo número de elementos. Es pensando en la demostración de tal afirmación como objetivo que probamos el siguiente lema.

LEMA 4.6. Si v_1, \dots, v_n es una base de V sobre F y w_1, \dots, w_m son elementos linealmente independientes de V , entonces $m \leq n$.

Prueba. Todo vector en V y, en particular, w_m , es una combinación lineal de v_1, \dots, v_n . Por tanto, los vectores w_m, v_1, \dots, v_n son linealmente dependientes. Además, generan a V ya que v_1, \dots, v_n lo generan. Luego existe algún subconjunto propio de tal conjunto, $w_m, v_{i_1}, \dots, v_{i_k}$ con $k \leq n-1$ que forma una base de V . Para formar esta nueva base hemos cambiado un w por al menos un v_i . Repitamos el procedimiento con el conjunto $w_{m-1}, w_m, v_{i_1}, \dots, v_{i_k}$. De este conjunto linealmente dependiente podemos extraer, de acuerdo con lo dicho por el corolario 1 del teorema 4.c, una base de la forma $w_{m-1}, w_m, v_{j_1}, \dots, v_{j_s}$, $s \leq n-2$. Repitiendo este proceso llegaremos a obtener una base de V de la forma $w_2, \dots, w_{m-1}, w_m, v_\alpha, v_\beta, \dots$; como w_1 no es una combinación lineal de w_2, \dots, w_{m-1} , la anterior base debe incluir sin duda algún v . Para llegar a esta base hemos introducido $m-1$ elementos w , y cada una de estas introducciones nos ha costado al menos una v y, sin embargo, todavía nos queda al menos una v . Luego $m-1 \leq n-1$ y, por tanto, $m \leq n$.

Este lema tiene como consecuencia (que nosotros enunciamos como corolarios) los resultados básicos que muestran la naturaleza de la dimensión de un espacio vectorial. Estos corolarios son de importancia máxima en todo lo que sigue, no solamente en este capítulo, sino en el resto del libro y, en realidad, en todas las matemáticas. Los corolarios son, todos, teoremas por derecho propio.

COROLARIO 1. Si V es un espacio de dimensión finita sobre F , entonces cualesquiera dos bases de V tienen el mismo número de elementos.

Prueba. Sea v_1, \dots, v_n una base de V sobre F y w_1, \dots, w_m una otra. Como, en particular, w_1, \dots, w_m es un conjunto de vectores linealmente independientes sobre F , hemos de tener, de acuerdo con el lema 4.6, que $m \leq n$. Intercambiando ahora los papeles de las v y las w , obtenemos: $n \leq m$. Lo que, con la desigualdad anterior, nos dice: $n = m$.

COROLARIO 2. $F^{(n)}$ es isomorfo a $F^{(m)}$ si y sólo si $n = m$.

Prueba. $F^{(n)}$ tiene como una de sus bases el conjunto de n vectores, $(1, 0, \dots, 0), (0, 1, 0, \dots, 0), \dots, (0, 0, \dots, 0, 1)$. Análogamente, $F^{(m)}$ tiene una base de m vectores. Un isomorfismo transforma una base suprayectiva una base (problema 4 del final de esta sección), de donde, según el corolario 1, $m = n$.

El corolario 2 pone sobre una base firme las observaciones heurísticas que anteriormente hicimos sobre los posibles isomorfismos de $F^{(n)}$ sobre

$F^{(n)}$. Como dijimos en aquellas observaciones, V es isomorfo a $F^{(n)}$ para algún n . Por el corolario 2, esta n es única, luego

COROLARIO 3. *Si V es de dimensión finita sobre F , entonces V es isomorfo a $F^{(n)}$ para un entero único n ; tal n es el número de elementos de cualquier base de V sobre F .*

DEFINICIÓN. Al entero n del corolario 3 se le llama *dimensión de V sobre F* .

La dimensión de V sobre F es, pues, el número de elementos de cualquier base de V sobre F .

Escribiremos $\dim V$ por dimensión de V sobre F , salvo cuando ocasionalmente deseemos subrayar el papel que juega el campo F , en cuyo caso escribiremos $\dim_F V$.

COROLARIO 4. *Dos espacios de dimensión finita sobre F cualesquiera de igual dimensión son isomorfos.*

Prueba. Si esta dimensión es n , entonces cada uno de ellos es isomorfo a $F^{(n)}$, luego isomorfo uno de otro.

¿Cuánta libertad tenemos para la construcción de bases de un espacio V ? El próximo lema asegura que, comenzando con cualquier conjunto de vectores independientes, podemos extenderlo a una base de V .

LEMA 4.7. *Si V es de dimensión finita sobre F y $u_1, \dots, u_m \in V$ es un conjunto de vectores linealmente independiente, entonces podemos encontrar vectores u_{m+1}, \dots, u_{m+r} en V tales que $u_1, \dots, u_m, u_{m+1}, \dots, u_{m+r}$ sea una base de V .*

Prueba. Como V es de dimensión finita, tiene una base; sea v_1, \dots, v_n una base de V . Como estos vectores generan a V , los vectores $u_1, \dots, u_m, v_1, \dots, v_n$ también generan a V . De acuerdo con el corolario 1 al teorema 4.c, hay un subconjunto de ellos de la forma $u_1, \dots, u_m, v_{i_1}, \dots, v_{i_r}$ de elementos linealmente independientes que generan V . Para probar el lema, simplemente hacemos $u_{m+1} = v_{i_1}, \dots, u_{m+r} = v_{i_r}$.

¿Cuál es la relación entre la dimensión de una imagen homomorfa de V y la dimensión de V ? La contestación nos la proporciona el siguiente lema.

LEMA 4.8. *Si V es de dimensión finita y W es un subespacio de V , entonces W es de dimensión finita, $\dim W \leq \dim V$ y $\dim V/W = \dim V - \dim W$.*

Prueba. Según el lema 4.6, si $n = \dim V$, entonces cualesquiera $n+1$ elementos de V son linealmente dependientes; en particular, cualesquiera $n+1$ elementos en W son linealmente dependientes. Podemos, pues, encontrar un conjunto máximo de elementos linealmente independientes de W , w_1, \dots, w_m y $m \leq n$. Si $w \in W$, entonces w_1, \dots, w_m, w es un conjunto linealmente dependiente, de donde $\alpha w + \alpha_1 w_1 + \dots + \alpha_m w_m = 0$ para ciertas α_i no todas las cuales son cero. Si $\alpha = 0$, por la independencia lineal de las w_i , tendríamos que concluir que todas las α_i son cero, una contradicción. Luego $\alpha \neq 0$, y por lo tanto $w = -\alpha^{-1}(\alpha_1 w_1 + \dots + \alpha_m w_m)$. Por consiguiente w_1, \dots, w_m generan a W ; por tanto, W es de dimensión finita sobre F y, además, tiene una base de m elementos, con $m \leq n$. De la definición de dimensión se sigue entonces que $\dim W \leq \dim V$.

Sea ahora w_1, \dots, w_m una base de W . De acuerdo con el lema 4.7 podemos extender a una base, $w_1, \dots, w_m, v_1, \dots, v_r$ de V , donde $m+r = \dim V$ y $r = \dim W$.

Sean $\bar{v}_1, \dots, \bar{v}_r$ las imágenes en V/W de v_1, \dots, v_r . Como cualquier vector $v \in V$ es de la forma $v = \alpha_1 w_1 + \dots + \alpha_m w_m + \beta_1 v_1 + \dots + \beta_r v_r$, entonces \bar{v} , la imagen de v , es de la forma $\bar{v} = \beta_1 \bar{v}_1 + \dots + \beta_r \bar{v}_r$ (pues $\bar{w}_1 = \bar{w}_2 = \dots = \bar{w}_m = 0$). Por tanto, $\bar{v}_1, \dots, \bar{v}_r$ generan V/W . Afirmamos que son linealmente independientes, pues si $\gamma_1 v_1 + \dots + \gamma_r v_r = 0$ entonces $y_1 v_1 + \dots + y_r v_r \in W$ y, por tanto, $y_1 v_1 + \dots + y_r v_r = \lambda_1 w_1 + \dots + \lambda_m w_m$, lo que por la independencia lineal del conjunto $w_1, \dots, w_m, v_1, \dots, v_r$ implicaría que $y_1 = \dots = y_r = \lambda_1 = \dots = \lambda_m = 0$. Hemos mostrado, pues, que V/W tiene una base de r elementos y, por tanto, $\dim V/W = r = \dim V - m = \dim V - \dim W$.

COROLARIO. Si A y B son subespacios de dimensión finita de un espacio vectorial V , entonces $A+B$ es de dimensión finita y $\dim(A+B) = \dim(A) + \dim(B) - \dim(A \cap B)$.

Prueba. Según el resultado del problema 13 al final de la sección 1, $\frac{A+B}{B} \approx \frac{A}{A \cap B}$, y como A y B son de dimensión finita, de ello se deduce

$$\dim(A+B) - \dim B = \dim\left(\frac{A+B}{B}\right) = \dim\left(\frac{A}{A \cap B}\right) = \dim A - \dim(A \cap B).$$

De donde, por transposición, obtenemos el resultado del lema.

Problemas

1. Pruébese el lema 4.4.
2. a) Si F es el campo de los números reales, pruébese que los vectores $(1, i, 0, 0)$, $(0, 1, -1, 0)$, y $(0, 0, 0, 3)$ en $F^{(4)}$ son linealmente independientes sobre F .

- b) ¿Qué condiciones sobre la característica de F harían a los tres vectores de la parte (a) linealmente dependientes?
3. Si V tiene una base de n elementos, pruébese una prueba detallada de que V es isomorfo a $F^{(n)}$.
4. Si T es un isomorfismo de V sobre W , pruébese que T transforma una base de V sobre una base de W .
5. Si V es de dimensión finita y T es un isomorfismo de V en V , pruébese que T debe transformar V sobre V .
- *6. Si V es de dimensión finita y T es un homomorfismo de V sobre V , pruébese que T debe ser inyectivo y, por tanto, un isomorfismo.
7. Si V es de dimensión n pruébese que cualquier conjunto de n vectores linealmente independientes en V forma una base de V .
8. Si V es de dimensión finita y W es un subespacio de V tal que $\dim V = \dim W$, pruébese que $V = W$.
9. Si V es de dimensión finita y T es un homomorfismo de V en sí mismo que no es suprayectivo, pruébese que hay algún $v \neq 0$ en V tal que $vT = 0$.
10. Sea F un campo y sea $F[x]$ el conjunto de todos los polinomios en x sobre F . Pruébese que $F[x]$ no es de dimensión finita sobre F .
11. Sea $V_n = \{p(x) \in F[x]: \deg p(x) < n\}$. Defínase T por
 $(\alpha_0 + \alpha_1 x + \dots + \alpha_{n-1} x^{n-1}) T = \alpha_0 + \alpha_1(x+1) + \alpha_2(x+1)^2 + \dots + \alpha_{n-1}(x+1)^{n-1}$.
Pruébese que T es un isomorfismo de V_n sobre sí mismo.
12. Sea $W = \{\alpha_0 + \alpha_1 x + \dots + \alpha_{n-1} x^{n-1} \in F[x]: \alpha_0 + \alpha_1 + \dots + \alpha_{n-1} = 0\}$. Demuéstrese que W es un subespacio de V_n y encuéntrese una base de W sobre F .
13. Sea v_1, \dots, v_n una base de V y sean w_1, \dots, w_n n elementos cualesquiera de V . Defínase T sobre V por $(\lambda_1 v_1 + \dots + \lambda_n v_n) T = \lambda_1 w_1 + \dots + \lambda_n w_n$.
- a) Pruébese que T es un homomorfismo de V en sí mismo.
b) ¿Cuándo es T un isomorfismo?
14. Demuéstrese que cualquier homomorfismo de V en sí mismo, cuando V es de dimensión finita, puede realizarse como en el problema 13 por elección apropiada de elementos w_1, \dots, w_n .
15. Volvamos al problema 13. Como v_1, \dots, v_n es una base de V , todo $w_i = \alpha_{i1} v_1 + \dots + \alpha_{in} v_n$, $\alpha_{ij} \in F$. Pruébese que los n^2 elementos α_{ij} de F determinan el homomorfismo T .

- *16. Si $\dim_F V = n$, pruébese que $\dim_F (\text{Hom}(V, V)) = n^2$.
17. Si V es de dimensión finita y W es un subespacio de V , pruébese que hay un subespacio W_1 de V tal que $V = W \oplus W_1$.

3. ESPACIOS DUALES

Dados dos espacios vectoriales cualesquiera, V y W , sobre un campo F , hemos definido $\text{Hom}(V, W)$ como el conjunto de todos los homomorfismos del espacio vectorial de V en W . Por el momento, $\text{Hom}(V, W)$ es simplemente un conjunto sin ninguna estructura en él. Procederemos ahora a introducir operaciones en este conjunto que lo convertirán en un espacio vectorial sobre F . Realmente, ya hemos indicado cómo conseguir esto en las descripciones de algunos de los problemas en las secciones anteriores. No obstante, nos proponemos tratar del asunto aquí de un modo más formal.

Sean S y T dos elementos cualesquiera de $\text{Hom}(V, W)$; esto quiere decir que ambos son homomorfismos de espacio vectorial de V en W . Recordando la definición de un tal homomorfismo, debemos tener $(v_1 + v_2)S = v_1 S + v_2 S$ y $(\alpha v_1)S = \alpha(v_1 S)$ para todo $v_1, v_2 \in V$ y todo $\alpha \in F$. Todo esto se repite también para T .

Necesitamos primero introducir una suma para estos elementos S y T en $\text{Hom}(V, W)$. ¿Qué más natural que definir $S+T$ suponiendo que $r(S+T) = rS+rT$ para todo $r \in V$? Debemos, desde luego, verificar que $S+T$ está en $\text{Hom}(V, W)$. Por la misma definición de $S+T$, si $v_1, v_2 \in V$, entonces $(v_1 + v_2)(S+T) = (v_1 + v_2)S + (v_1 + v_2)T$; como $(v_1 + v_2)S = v_1 S + v_2 S$ y $(v_1 + v_2)T = v_1 T + v_2 T$, y como la adición en W es conmutativa, tenemos: $(v_1 + v_2)(S+T) = v_1 S + v_1 T + v_2 S + v_2 T$. Invocando de nuevo la definición de $S+T$, el segundo miembro de esta relación toma la forma $v_1(S+T) + v_2(S+T)$. Un cálculo análogo demuestra que $(\alpha v)(S+T) = \alpha(v(S+T))$. Por consiguiente, $S+T$ está en $\text{Hom}(V, W)$. Sea 0 el homomorfismo de V en W que envía todo elemento de V sobre el elemento cero de W ; para $S \in \text{Hom}(V, W)$, definamos $-S$ por $v(-S) = -(vS)$. Es inmediato que $\text{Hom}(V, W)$ es un grupo abeliano bajo la adición que acabamos de definir.

Una vez que hemos conseguido introducir la estructura de un grupo abeliano sobre $\text{Hom}(V, W)$, volvemos ahora nuestra atención al problema de definir λS para $\lambda \in F$ y $S \in \text{Hom}(V, W)$, pues nuestra última meta es la de hacer de $\text{Hom}(V, W)$ un espacio vectorial sobre F . Para $\lambda \in F$ y $S \in \text{Hom}(V, W)$, definimos λS por $r(\lambda S) = \lambda(rS)$ para todo $r \in V$. Dejamos al lector la demostración de que λS está en $\text{Hom}(V, W)$ y de que bajo la operación que así hemos definido $\text{Hom}(V, W)$ es un espacio vectorial sobre F . Sin embargo, no tenemos ninguna seguridad de que $\text{Hom}(V, W)$ tenga elementos distintos del homomorfismo 0 . Sea como fuere, hemos probado el

LEMA 4.9. *Hom (V, W) es un espacio vectorial sobre F bajo las operaciones que acabamos de describir.*

Resultados como el del lema 4.9 nos dan realmente muy poca información; sin embargo, nos confirman que las definiciones que hemos dado son razonables. Preferiríamos algunos resultados sobre Hom (V, W) que nos dieran más información. Un resultado de tal tipo nos lo proporciona el

TEOREMA 4.D. *Si V y W son, respectivamente, de dimensiones m y n sobre F, entonces Hom (V, W) es de dimensión mn sobre F.*

Prueba. Probaremos el teorema exhibiendo explícitamente una base de Hom (V, W) sobre F que tenga mn elementos.

Sea v_1, \dots, v_m una base de V sobre F y w_1, \dots, w_n una para W sobre F. Si $v \in V$, entonces $v = \lambda_1 v_1 + \dots + \lambda_m v_m$ donde $\lambda_1, \dots, \lambda_m$ son elementos únicamente definidos de F; definamos $T_{ij} : V \rightarrow W$ por $v T_{ij} = \lambda_i w_j$. Desde el punto de vista de las bases mencionadas, lo que estamos haciendo es simplemente definir $v_k T_{ij} = 0$ para $k \neq i$ y $v_i T_{ij} = w_j$. Es un ejercicio sencillo establecer que T_{ij} está en Hom (V, W). Como i puede ser uno cualquiera de los números 1, 2, ..., m y j uno cualquiera de los 1, 2, ..., n , hay mn de tales T_{ij} .

Nuestra tesis es que estos mn elementos constituyen una base de Hom (V, W) sobre F. Sea, en efecto, $S \in \text{Hom}(V, W)$; como $v_1 S \in W$, y como cualquier elemento de W es una combinación lineal sobre F de w_1, \dots, w_n , $v_1 S = \alpha_{11} w_1 + \dots + \alpha_{1n} w_n$, para ciertos $\alpha_{11}, \alpha_{12}, \dots, \alpha_{1n}$ en F. En realidad $v_i S = \alpha_{i1} w_1 + \dots + \alpha_{in} w_n$ para $i = 1, 2, \dots, m$. Consideremos $S_0 = \alpha_{11} T_{11} + \alpha_{12} T_{12} + \dots + \alpha_{1n} T_{1n} + \alpha_{21} T_{21} + \dots + \alpha_{2n} T_{2n} + \dots + \alpha_{m1} T_{m1} + \dots + \alpha_{mn} T_{mn}$. Calculemos $v_k S_0$ para el vector de la base v_k . Tenemos $v_k S_0 = v_k (\alpha_{11} T_{11} + \dots + \alpha_{m1} T_{m1} + \dots + \alpha_{mn} T_{mn}) = \alpha_{11} (v_k T_{11}) + \alpha_{12} (v_k T_{12}) + \dots + \alpha_{m1} (v_k T_{m1}) + \dots + \alpha_{mn} (v_k T_{mn})$. Como $v_k T_{ij} = 0$ para $i \neq k$ y $v_k T_{kj} = w_j$, esta suma se reduce a $v_k S_0 = \alpha_{k1} w_1 + \dots + \alpha_{kn} w_n$ que, como vemos, no es otra cosa que $v_k S$. Así pues, los homomorfismos S_0 y S coinciden sobre una base de V. Afirmamos que tal cosa implica $S_0 = S$ (véase el problema 3 al final de esta sección). Pero S_0 es una combinación lineal de los T_{ij} , de donde S debe ser la misma combinación lineal. En resumen, hemos demostrado que los mn elementos $T_{11}, T_{12}, \dots, T_{1n}, \dots, T_{mn}$, generan Hom (V, W) sobre F.

Para probar que forman una base de Hom (V, W) sobre F solo queda probar su independencia lineal sobre F. Supongamos que $\beta_{11} T_{11} + \beta_{12} T_{12} + \dots + \beta_{1n} T_{1n} + \dots + \beta_{i1} T_{i1} + \dots + \beta_{in} T_{in} + \dots + \beta_{m1} T_{m1} + \dots + \beta_{mn} T_{mn} = 0$ con todos los β_{ij} de F. Aplicando esto a v_k tenemos: $0 = v_k (\beta_{11} T_{11} + \dots + \beta_{ij} T_{ij} + \dots + \beta_{mn} T_{mn}) = \beta_{k1} w_1 + \beta_{k2} w_2 + \dots + \beta_{kn} w_n$ ya que $v_k T_{ij} = 0$ para $i \neq k$ y $v_k T_{kj} = w_j$. Pero w_1, \dots, w_n son linealmente independientes sobre F, luego $\beta_{kj} = 0$ para cualesquiera k y j . Luego los T_{ij} son linealmente

independientes sobre F , de donde puede afirmarse que forman una base de $\text{Hom}(V, W)$ sobre F .

Una consecuencia inmediata del teorema 4.d es que siempre que $V \neq (0)$ y $W \neq (0)$ son espacios de dimensión finita, entonces $\text{Hom}(V, W)$ no consiste solamente del elemento 0, pues su dimensión sobre F es $mn \geq 1$.

Algunos casos especiales del teorema 4.d son de gran importancia y los enumeramos como corolarios.

COROLARIO 1. Si $\dim_F V = m$, entonces $\dim_F \text{Hom}(V, W) = m^2$.

Prueba. En el teorema basta poner $V = W$ y, por tanto, $m = n$. Entonces $mn = m^2$.

COROLARIO 2. Si $\dim_F V = m$, entonces $\dim_F \text{Hom}(V, F) = m$.

Prueba. Como un espacio vectorial, F es de dimensión 1 sobre F . Aplicando el teorema obtenemos $\dim_F \text{Hom}(V, F) = m$.

El corolario 2 tiene la interesante consecuencia de que si V es de dimensión finita sobre F entonces es isomorfo a $\text{Hom}(V, F)$, pues, según el corolario tienen la misma dimensión sobre F , de donde según el corolario 4 del lema 4.6 deben ser isomorfos. Este isomorfismo tiene muchas limitaciones. Expliquémonos. Depende, en gran medida, de que V sea de dimensión finita, pues si V no lo fuera no existiría ningún isomorfismo de tal tipo. No hay ninguna construcción formal y elegante de este isomorfismo que pueda aplicarse a todos los espacios vectoriales. Depende mucho de las particularidades de la situación que se estudie. Dentro de pocas páginas veremos, sin embargo, que para los espacios $\text{Hom}(\text{Hom}(V, F), F)$, sí existe un homomorfismo elegante, cualquiera que sea el espacio V .

DEFINICIÓN. Si V es un espacio vectorial, entonces su *espacio dual* es $\text{Hom}(V, F)$.

Usaremos la notación \hat{V} como notación para el espacio dual del V . Un elemento de \hat{V} se llamará *funcional lineal* sobre V en F .

Si V no es de dimensión finita entonces \hat{V} es usualmente demasiado grande y rara para que sea de interés. Para tales espacios vectoriales tenemos, a menudo, otras estructuras adicionales, tales como una topología, en ellos, y entonces, como espacio dual, no se toma generalmente a todo el \hat{V} sino un subespacio de él adecuadamente restringido. Si V es de dimensión finita su espacio dual está siempre definido, como hemos dicho, como todo el espacio $\text{Hom}(V, F)$.

En la prueba del teorema 4.d construimos una base de $\text{Hom}(V, W)$ usando una base particular de V y una de W . La construcción dependía

fundamentalmente de las bases particulares de V y W que hubiésemos escogido. Si hubiéramos escogido otras bases habríamos terminado con una base diferente para $\text{Hom}(V, W)$. Como principio general, es preferible dar pruebas, siempre que sea posible, que no dependan de una elección de bases. A tales pruebas se les suele llamar invariantes. Una prueba o construcción invariante tiene la ventaja, aparte de la puramente estética, sobre una prueba o construcción en que se usen bases de las cuales no necesitamos preocuparnos de en qué medida depende todo de la particular selección de las bases que se haya hecho.

Los elementos de \hat{V} son funciones definidas sobre V y con sus valores en F . Conservando la notación funcional, escribiremos usualmente los elementos de \hat{V} usando las letras f, g, \dots , y denotaremos a los valores en F de un elemento $v \in V$ como $f(v)$ (en lugar de por vf).

Sea V un espacio vectorial de dimensión finita sobre F y sea v_1, \dots, v_n una base de V ; sea ϑ_i el elemento de \hat{V} definido por $\vartheta_i(v_j) = 0$ para $i \neq j$, y $\vartheta_i(v_i) = 1$, y $\vartheta_i(\alpha_1 v_1 + \dots + \alpha_i v_i + \dots + \alpha_n v_n) = \alpha_i$. En realidad las ϑ_i no son nada más que las T_{ij} introducidas en la prueba del teorema 4.d, para las que, en este caso, hay que tener en cuenta que $W = F$ es unidimensional sobre F . Conocemos, pues, que $\vartheta_1, \dots, \vartheta_n$ forman una base de \hat{V} . Llamamos a esta base la *base dual* de la v_1, \dots, v_n . Si $v \neq 0 \in V$, por el lema 4.7 podemos encontrar una base de la forma $v_1 = v, v_2, \dots, v_n$ y, por tanto, hay un elemento en \hat{V} , llamémosle ϑ_1 , tal que $\vartheta_1(v_1) = \vartheta_1(v) = 1 \neq 0$. Con lo que hemos probado

LEMA 4.10. *Si V es de dimensión finita y $v \neq 0 \in V$, entonces hay un elemento $f \in \hat{V}$ tal que $f(v) \neq 0$.*

En realidad, el lema 4.10 es también cierto si V no es de dimensión finita, pero como no tenemos necesidad de este resultado y su prueba envolvería cuestiones de lógica que no son adecuadas en esta etapa, omitimos la prueba.

Sea $v_0 \in V$, donde V es un espacio vectorial cualquiera sobre F . Cuando f varía sobre \hat{V} y v_0 se conserva fijo, $f(v_0)$ define una funcional sobre \hat{V} en F ; nótese que lo único que estamos haciendo es intercambiando los papeles de función y variable. Denotemos esta función por T_{v_0} ; en otras palabras $T_{v_0}(f) = f(v_0)$ para cualquier $f \in \hat{V}$. ¿Qué puede decirse acerca de T_{v_0} ? Tenemos para comenzar: $T_{v_0}(f+g) = (f+g)(v_0) = f(v_0) + g(v_0) = T_{v_0}(f) + T_{v_0}(g)$; además, $T_{v_0}(\lambda f) = (\lambda f)(v_0) = \lambda f(v_0) = \lambda T_{v_0}(f)$. ¡Luego T_{v_0} está en el espacio dual del \hat{V} ! Representamos este espacio por $\hat{\mathcal{F}}$ y nos referimos a él como el *segundo dual* del V .

Dado cualquier elemento $v \in V$, podemos asociar con él un elemento T_v en $\hat{\mathcal{F}}$. Definimos la transformación $\psi : V \rightarrow \hat{\mathcal{F}}$ por $v\psi = T_v$ para todo $v \in V$. ¿Es ψ un homomorfismo de V en $\hat{\mathcal{F}}$? Claro que sí. Tenemos, en efecto: $T_{v+w}(f) = f(v+w) = f(v) + f(w) = T_v(f) + T_w(f) = (T_v + T_w)(f)$, luego $T_{v+w} = T_v + T_w$, es decir, $(v+w)\psi = v\psi + w\psi$. Análogamente, para $\lambda \in F$,

$(\lambda v)\psi = \lambda(v\psi)$. Luego ψ define un homomorfismo de V en \hat{V} . La construcción de ψ no usa ninguna base particular ni ninguna propiedad particular de V ; es un ejemplo de construcción invariantes.

¿Cuándo es ψ un isomorfismo? Para contestar a esto debemos saber cuándo $v\psi = 0$ o, lo que es equivalente, cuándo $T_v = 0$. Pero si $T_v = 0$ entonces $0 = T_v(f) = f(v)$ para toda $f \in \hat{V}$. Pero como ya señalamos, sin prueba, para un espacio vectorial general, dado $v \neq 0$ hay una $f \in \hat{V}$ con $f(v) \neq 0$. Realmente tal afirmación la demostramos para V de dimensión finita. Luego para V de dimensión finita (y en realidad, para V cualquiera) ψ es un isomorfismo. Pero cuando V es de dimensión finita, ψ es un isomorfismo sobre \hat{V} y, en cambio, cuando V no lo es, ψ no es suprayectiva.

Si V es de dimensión finita, de acuerdo con el segundo corolario al teorema 4.4, V y \hat{V} son de la misma dimensión; análogamente, \hat{V} y $\hat{\hat{V}}$ son de la misma dimensión; como ψ es un isomorfismo de V en \hat{V} la igualdad de dimensiones obliga a que ψ sea suprayectiva. Y hemos probado el

LEMA 4.11. *Si V es de dimensión finita, entonces ψ es un isomorfismo de V sobre \hat{V} .*

De aquí en adelante, identificaremos V con \hat{V} , y recordaremos que esta identificación se efectúa a través del isomorfismo ψ .

DEFINICIÓN. Si W es un subespacio de V , entonces el *aniquilador* de W , $A(W) = \{f \in \hat{V} \mid f(w) = 0 \text{ para todo } w \in W\}$.

Dejamos como ejercicio para el lector la verificación del hecho de que $A(W)$ es un subespacio de \hat{V} . Claramente si $U \subset W$, entonces $A(U) \supset A(W)$.

Sea W un subespacio de V , y V de dimensión finita. Si $f \in \hat{V}$, sea \tilde{f} la restricción de f a W ; \tilde{f} está, pues, definida sobre W por $\tilde{f}(w) = f(w)$ para toda $w \in W$. Como $f \in \hat{V}$, es claro que $\tilde{f} \in \hat{W}$. Consideremos la aplicación $T: \hat{V} \rightarrow \hat{W}$ definida por $fT = \tilde{f}$ para toda $f \in \hat{V}$. T es, así, un homomorfismo de \hat{V} en \hat{W} . ¿Cuál es el núcleo de T ? Si f está en el núcleo de T , entonces la restricción de f a W debe ser 0; es decir, $f(w) = 0$ para todo $w \in W$. Además, recíprocamente, si $f(w) = 0$ para toda $w \in W$, entonces f está en el núcleo de T . Por tanto, el núcleo de T es exactamente $A(W)$.

Afirmamos ahora que la aplicación T es sobre \hat{W} . Lo que debemos demostrar es que dado cualquier elemento $h \in \hat{W}$, entonces h es la restricción de algún $f \in \hat{V}$, es decir, que $h = \tilde{f}$. Según el lema 4.7, si w_1, \dots, w_m es una base de W , entonces podemos expandirla a una base de V de la forma $w_1, \dots, w_m, v_1, \dots, v_r$, donde $r+m = \dim V$. Sea W_1 el subespacio de V generado por v_1, \dots, v_r . Entonces $V = W \oplus W_1$. Si $h \in \hat{W}$, definamos $f \in \hat{V}$ del siguiente modo: escribamos $v \in V$ como $v = w+w_1$, $w \in W$, $w_1 \in W_1$, entonces $f(v) = h(w)$. Es fácil ver que f está en \hat{V} y que $\tilde{f} = h$. Luego $h = fT$ y, por tanto, T transforma \hat{V} sobre \hat{W} . Como el núcleo de T es

$A(W)$, por el teorema 4.a \hat{W} es isomorfo a $\hat{V}/A(W)$. En particular, ambos espacios tienen la misma dimensión. Sea $m = \dim W$, $n = \dim V$, y $r = \dim A(W)$. De conformidad con el corolario 2 al teorema 4.d, $m = \dim \hat{W}$ y $n = \dim \hat{V}$. Pero según el lema 4.8, $\dim \hat{V}/A(W) = \dim \hat{V} - \dim A(W) = n - r$, luego $m = n - r$ y, por tanto, de donde transponiendo, $r = n - m$. Con lo que hemos probado el

TEOREMA 4.E. Si V es de dimensión finita y W es un subespacio de V , entonces \hat{W} es isomorfo a $\hat{V}/A(W)$ y $\dim A(W) = \dim V - \dim W$.

COROLARIO. $A(A(W)) = W$.

Prueba. Recordemos que para que el corolario tenga aunque solo sea sentido, como $W \subset V$ y $A(A(W)) \subset \hat{V}$, hemos tenido que identificar V con \hat{V} . Ahora bien, $W \subset A(A(W))$, pues si $w \in W$, entonces $w\Psi = T_w$ actúa sobre V por $T_w(f) = f(w)$ y, por tanto, es 0 para toda $f \in A(W)$. Pero $\dim A(A(W)) = \dim \hat{V} - \dim A(W)$ (aplicando el teorema al espacio vectorial \hat{V} y su subespacio $A(W)$) de modo que $\dim A(A(W)) = \dim \hat{V} - \dim A(W) = \dim V - (\dim V - \dim W) = \dim W$. Como $W \subset A(A(W))$ y son de la misma dimensión, de ello se sigue que $W = A(A(W))$.

El teorema 4.e tiene aplicación al estudio de los sistemas de *ecuaciones lineales homogéneas*. Consideremos el sistema de m ecuaciones con n incógnitas

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = 0$$

$$a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = 0$$

.

.

$$a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = 0$$

donde las a_{ij} están en F . Queremos saber el número de soluciones linealmente independientes (x_1, \dots, x_n) que hay en $F^{(n)}$ para este sistema.

Sea U el subespacio de $F^{(n)}$ generado por los m vectores $(a_{11}, a_{12}, \dots, a_{1n})$, $(a_{21}, a_{22}, \dots, a_{2n})$, ..., $(a_{m1}, a_{m2}, \dots, a_{mn})$, y supongamos que U es de dimensión r . En tal caso, diremos que el sistema de ecuaciones es de *rango r*.

Usemos $v_1 = (1, 0, \dots, 0)$, $v_2 = (0, 1, 0, \dots, 0)$, ..., $v_n = (0, 0, \dots, 0, 1)$ como una base de $F^{(n)}$ y sea $\vartheta_1, \vartheta_2, \dots, \vartheta_n$ su base dual en $\hat{F}^{(n)}$. Cualquier $f \in \hat{F}^{(n)}$ es de la forma $f = x_1\vartheta_1 + x_2\vartheta_2 + \dots + x_n\vartheta_n$, donde $x_i \in F$ para toda i . ¿Cuándo $f \in A(U)$? En tal caso, como

$$(a_{11}, \dots, a_{1n}) \in U, 0 = f(a_{11}, a_{12}, \dots, a_{1n}) =$$

$$f(a_{11}v_1 + \dots + a_{1n}v_n) = (x_1\vartheta_1 + x_2\vartheta_2 + \dots + x_n\vartheta_n)(a_{11}v_1 + \dots + a_{1n}v_n) =$$

$$x_1a_{11} + x_2a_{12} + \dots + x_na_{1n} \text{ ya que } \vartheta_i(v_j) = 0 \text{ para } i \neq j \text{ y } \vartheta_i(v_i) = 1.$$

Asimismo, las otras ecuaciones del sistema son satisfechas. Recíprocamente, toda solución (x_1, \dots, x_n) del sistema de ecuaciones homogéneas da lugar a un elemento $x_1\hat{v}_1 + \dots + x_n\hat{v}_n$ en $A(U)$. Vemos, por ello, que el número de soluciones linealmente independientes del sistema de ecuaciones es la dimensión de $A(U)$, la que, por el teorema 4.e es $n-r$. Y hemos probado el siguiente

TEOREMA 4.F. *Si el sistema de ecuaciones lineales homogéneas:*

$$a_{11}x_1 + \dots + a_{1n}x_n = 0$$

$$a_{21}x_1 + \dots + a_{2n}x_n = 0$$

.

.

$$a_{m1}x_1 + \dots + a_{mn}x_n = 0,$$

donde $a_{ij} \in F$ es de rango r entonces hay $n-r$ soluciones linealmente independientes en $F^{(n)}$.

COROLARIO. Si $n > m$, es decir, si el número de incógnitas es superior al número de ecuaciones, entonces hay una solución (x_1, \dots, x_n) donde no todos los x_1, \dots, x_n son cero.

Prueba. Como U está generado por m vectores, y $m < n$, $r = \dim U \leq m < n$; la aplicación del teorema 4.f da lugar al corolario.

Problemas

1. Pruébese que $A(W)$ es un subespacio de \hat{V} .
2. Si S es un subconjunto de V , sea $A(S) = \{f \in \hat{V} \mid f(s) = 0 \text{ para todo } s \in S\}$. Pruébese que $A(S) = A(L(S))$ donde $L(S)$ es el espacio generado por S .
3. Si $S, T \in \text{Hom}(V, W)$ y $v_i S = v_i T$ para todos los elementos v_i de una base de V , pruébese que $S = T$.
4. Complétese la prueba, dándose todos los detalles, de que $\text{Hom}(V, W)$ es un espacio vectorial sobre F .
5. Si ψ denota la aplicación usada en el texto de V en \hat{V} , suministrese una prueba completa de que ψ es un homomorfismo entre espacios vectoriales de V en \hat{V} .
6. Si V es de dimensión finita y $v_1 \neq v_2$ están en V , pruébese que hay una $f \in \hat{V}$ tal que $f(v_1) \neq f(v_2)$.

7. Si W_1 y W_2 son subespacios de V , que es de dimensión finita, describáse $A(W_1 + W_2)$ en términos de $A(W_1)$ y $A(W_2)$.

8. Si V es de dimensión finita y W_1 y W_2 son subespacios de V , describáse $A(W_1 \cap W_2)$ en términos de $A(W_1)$ y $A(W_2)$.

9. Si F es el campo de los números reales, encuéntrese $A(W)$ donde:

a) W es generado por $(1, 2, 3)$ y $(0, 4, -1)$.

b) W es generado por $(0, 0, 1, -1)$, $(2, 1, 1, 0)$ y $(2, 1, 1, -1)$.

10. Encuéntrense los rangos de los siguientes sistemas de ecuaciones lineales homogéneas sobre F , el campo de los números reales, y encuéntrense todas las soluciones.

a) $x_1 + 2x_2 - 3x_3 + 4x_4 = 0$

$$x_1 + 3x_2 - x_3 = 0$$

$$6x_1 + x_3 + 2x_4 = 0.$$

b) $x_1 + 3x_2 + x_3 = 0$

$$x_1 + 4x_2 + x_3 = 0.$$

c) $x_1 + x_2 + x_3 + x_4 + x_5 = 0$

$$x_1 + 2x_2 = 0$$

$$4x_1 + 7x_2 + x_3 + x_4 + x_5 = 0$$

$$x_2 - x_3 - x_4 - x_5 = 0.$$

11. Si f y g están en \hat{V} y son tales que $f(x) = 0$ implica $g(x) = 0$, pruébese que $g = \lambda f$ para algún $\lambda \in F$.

4. ESPACIOS CON PRODUCTO INTERIOR

En nuestra discusión de espacios vectoriales la naturaleza específica de F como campo, aparte del hecho de que es un campo, no ha jugado realmente papel alguno. En esta sección no vamos a seguir considerando espacios vectoriales sobre campos arbitrarios F , sino que vamos a restringir F al campo de los números reales o al campo de los números complejos. En el primer caso, vamos a llamar a V *espacio vectorial real*, en el segundo, *espacio vectorial complejo*.

Todos nosotros hemos tenido alguna experiencia con espacios vectoriales reales —en realidad, tanto la geometría analítica como el tema del análisis vectorial tratan de ellos. ¿Qué conceptos de los allí usados pueden trasladarse a un marco más abstracto? Para comenzar, en estos ejemplos concretos teníamos la idea de longitud; en segundo lugar, teníamos también la idea de perpendicularidad o, más generalmente, la de ángulo. Todos éstos se hacen casos especiales del concepto de un producto punto (a menudo llamado también producto escalar o interior).

Recordemos algunas propiedades del producto punto tal como se presenta en el caso especial de los vectores reales tridimensionales. Dados los vectores $v = (x_1, x_2, x_3)$ y $w = (y_1, y_2, y_3)$, donde las x y las y son números reales, el producto punto de v y w , representado por $v \cdot w$, se definía por $v \cdot w = x_1 y_1 + x_2 y_2 + x_3 y_3$. Nótese que la longitud de v venía dada por $\sqrt{v \cdot v}$ y el ángulo θ entre v y w estaba determinado por

$$\cos \theta = \frac{v \cdot w}{\sqrt{v \cdot v} \sqrt{w \cdot w}}.$$

¿Qué propiedades formales tenía este producto punto? Enumeramos algunas:

- 1) $v \cdot v \geq 0$ y $v \cdot v = 0$ si y sólo si $v = 0$,
- 2) $v \cdot w = w \cdot v$,
- 3) $u \cdot (av + bw) = a(u \cdot v) + b(u \cdot w)$,

para vectores u, v, w y números reales a, b , cualesquiera.

Todo lo que se ha dicho puede llevarse a los espacios vectoriales complejos. Pero para tener definiciones geométricamente razonables debemos hacer algunas modificaciones. Si definimos simplemente $v \cdot w = x_1 y_1 + x_2 y_2 + x_3 y_3$ para $v = (x_1, x_2, x_3)$ y $w = (y_1, y_2, y_3)$, donde las x y las y son números complejos, entonces es muy posible que $v \cdot v = 0$ con $v \neq 0$; queda esto ilustrado con el vector $v = (1, i, 0)$. En realidad, $v \cdot v$ no necesita ser ni siquiera real. Si, como en el caso real, queremos que $v \cdot v$ represente de algún modo la longitud de v , queremos que esta longitud sea real y que un vector que no sea nulo no pueda tener una longitud cero.

Podemos conseguir todo ello con solo alterar la definición de producto escalar ligeramente. Si $\bar{\alpha}$ denota el conjugado complejo del número complejo α , volviendo a los v y w del párrafo anterior, definamos $v \cdot w = x_1 \bar{y}_1 + x_2 \bar{y}_2 + x_3 \bar{y}_3$. Para vectores reales, esta nueva definición coincide con la antigua; por otra parte, para vectores complejos arbitrarios $v \neq 0$, no solamente es $v \cdot v$ real, sino que también positivo. Tenemos, pues, la posibilidad de introducir de una forma natural una longitud no negativa. Pero perdemos algo; por ejemplo, no sigue siendo cierto que $v \cdot w = w \cdot v$. En realidad, la relación exacta entre estos dos miembros es $v \cdot w = \overline{w \cdot v}$. Enumeremos unas pocas propiedades de este producto punto:

- 1) $v \cdot w = \overline{w \cdot v}$
- 2) $v \cdot v \geq 0$, y $v \cdot v = 0$ si y sólo si $v = 0$.
- 3) $(\alpha v + \beta w) \cdot w = \alpha(v \cdot w) + \beta(w \cdot w)$
- 4) $u \cdot (av + bw) = \bar{\alpha}(u \cdot v) + \bar{\beta}(u \cdot w)$,

para cualesquiera números complejos α, β y cualesquiera vectores u, v, w .

Repetimos que en lo que sigue F será el campo de los números complejos o el campo de los números reales.

DEFINICIÓN. El espacio vectorial V sobre F se dice que es un *espacio con producto interior* si para dos vectores cualesquiera $u, v \in V$ está definido un elemento (u, v) de F tal que:

- 1) $(u, v) = \overline{(v, u)}$
- 2) $(u, u) \geq 0$ y $(u, u) = 0$ si y sólo si $u = 0$
- 3) $(\alpha u + \beta v, w) = \alpha(u, w) + \beta(v, w)$

para cualesquiera $u, v, w \in V$ y $\alpha, \beta \in F$.

Son ahora pertinentes unas pocas observaciones acerca de las propiedades (1), (2) y (3). Una función que las satisface se llama *producto interior*. Si F es el campo de los números complejos, la propiedad (1) implica que (u, u) es real, y por tanto la propiedad (2) tiene sentido. Usando (1) y (3) vemos que $(u, \alpha v + \beta w) = \overline{(\alpha v + \beta w, u)} = \alpha \overline{(v, u)} + \beta \overline{(w, u)} = \bar{\alpha} \overline{(v, u)} + \bar{\beta} \overline{(w, u)} = \bar{\alpha}(u, v) + \bar{\beta}(u, w)$.

Hacemos ahora una pausa para ver algunos ejemplos de espacios con producto interior.

EJEMPLO 1. Definamos en $F^{(n)}$, para $u = (\alpha_1, \dots, \alpha_n)$ y $v = (\beta_1, \dots, \beta_n)$, $(u, v) = \alpha_1 \beta_1 + \alpha_2 \beta_2 + \dots + \alpha_n \beta_n$. Definimos así un producto escalar sobre $F^{(n)}$.

EJEMPLO 2. Definamos en $F^{(2)}$ para $u = (\alpha_1, \alpha_2)$ y $v = (\beta_1, \beta_2)$, $(u, v) = 2\alpha_1 \beta_1 + \alpha_1 \beta_2 + \alpha_2 \beta_1 + \alpha_2 \beta_2$. Es fácil verificar que esto define un producto escalar sobre $F^{(2)}$.

EJEMPLO 3. Sea V el conjunto de todas las funciones de valores complejos definidas sobre el intervalo cerrado unitario $[0, 1]$. Si $f(t), g(t) \in V$ definamos $(f(t), g(t)) = \int_0^1 f(t)g(t) dt$. Dejamos para el lector la verificación de que esto define un producto interior sobre V .

Para el resto de esta sección V representará a un espacio con producto interior.

DEFINICIÓN. Si $v \in V$, entonces la *longitud* de v (o *norma* de v), que representaremos por $\|v\|$, está definida como $\|v\| = \sqrt{(v, v)}$.

LEMA 4.12. Si $u, v \in V$ y $\alpha, \beta \in F$, entonces $(\alpha u + \beta v, \alpha u + \beta v) = \alpha \bar{\alpha}(u, u) + \alpha \beta(u, v) + \bar{\alpha} \beta(v, u) + \beta \bar{\beta}(v, v)$.

Prueba. Por la propiedad (3) de las definitorias de un producto interior, $(\alpha u + \beta v, \alpha u + \beta v) = \alpha(u, \alpha u + \beta v) + \beta(v, \alpha u + \beta v)$; pero $(u, \alpha u + \beta v) = \bar{\alpha}(u, u) + \beta(u, v)$ y $(v, \alpha u + \beta v) = \bar{\alpha}(v, u) + \bar{\beta}(v, v)$. Haciendo las correspondientes

sustituciones en la expresión para $(\alpha u + \beta v, \alpha u + \beta v)$, obtenemos el resultado deseado.

COROLARIO. $\|u\| = |\alpha| \|u\|$.

Prueba. $\|\alpha u\|^2 = (\alpha u, \alpha u) = \alpha \bar{\alpha} (u, u)$ por lema 4.12 (con $v = 0$). Como $\alpha \bar{\alpha} = |\alpha|^2$ y $(u, u) = \|u\|^2$, tomando las raíces cuadradas obtenemos $\|\alpha u\| = |\alpha| \|u\|$.

Haremos ahora una pequeña digresión para probar un resultado muy elemental y familiar acerca de las ecuaciones cuadráticas reales.

LEMA 4.13. Si a, b y c son números reales tales que $a > 0$ y $a\lambda^2 + 2b\lambda + c \geq 0$ para todo número real λ , entonces $b^2 \leq ac$.

Prueba. Completando los cuadrados,

$$a\lambda^2 + 2b\lambda + c = \frac{1}{a}(a\lambda + b)^2 + \left(c - \frac{b^2}{a}\right).$$

Como esto es mayor que o igual a 0 para todo λ , en particular debe ser cierto para $\lambda = -\frac{b}{a}$. Luego $c - \frac{b^2}{a} \geq 0$, y como $a > 0$, obtenemos $b^2 \leq ac$.

Pasamos ahora a probar una desigualdad muy importante, usualmente conocida como la *desigualdad de Schwarz*.

TEOREMA 4.G. Si $u, v \in V$, entonces $|(u, v)| \leq \|u\| \|v\|$.

Prueba. Si $u = 0$ entonces tanto (u, v) como $\|u\| \|v\|$ son iguales a cero, luego el resultado en este caso es cierto.

Supongamos por el momento que (u, v) es real y $u \neq 0$. Segundo el lema 4.12 para cualquier número real λ , $0 \leq (\lambda u + v, \lambda u + v) = \lambda^2(u, u) + 2(u, v)\lambda + (v, v)$. Sea $a = (u, u)$, $b = (u, v)$ y $c = (v, v)$; para tales a , b y c se satisface la hipótesis del lema 4.13, luego $b^2 \leq ac$. Es decir, $(u, v)^2 \leq (u, v)(v, v)$ de donde es inmediato que $|(u, v)| \leq \|u\| \|v\|$.

Si $\alpha = (u, v)$ no es real, entonces es claro que no es cero, de modo que u/α es significativo. Ahora bien,

$$\left(\frac{u}{\alpha}, v\right) = \frac{1}{\alpha}(u, v) = \frac{1}{(u, v)}(u, v) = 1,$$

y, por tanto, es ciertamente real. Por el caso de la desigualdad de Schwarz que hemos discutido en el anterior párrafo,

$$1 = \left| \left(\frac{u}{\alpha}, v \right) \right| \leq \left\| \frac{u}{\alpha} \right\| \|v\|;$$

ya que

$$\left\| \frac{u}{\alpha} \right\| = \frac{1}{|\alpha|} \|u\|,$$

obtenemos

$$1 \leq \frac{\|u\| \|v\|}{|\alpha|},$$

de donde $|\alpha| \leq \|u\| \|v\|$. Haciendo $\alpha = (u, v)$, obtenemos $|(u, v)| \leq \|u\| \|v\|$, es decir, el resultado deseado.

Algunos de los casos particulares de la desigualdad de Schwarz son en sí de gran interés. Señalamos dos de ellos.

1) Si $V = F^{(n)}$ con $(u, v) = \alpha_1 \beta_1 + \dots + \alpha_n \beta_n$, donde $u = (\alpha_1, \dots, \alpha_n)$ y $v = (\beta_1, \dots, \beta_n)$, entonces el teorema 4.g implica que

$$|\alpha_1 \beta_1 + \dots + \alpha_n \beta_n|^2 \leq (|\alpha_1|^2 + \dots + |\alpha_n|^2)(|\beta_1|^2 + \dots + |\beta_n|^2).$$

2) Si V es el conjunto de todas las funciones complejas continuas con $(0, 1)$ como dominio, y con producto interior definido por $(f(t), g(t)) =$

$= \int_0^1 f(t) \overline{g(t)} dt$, entonces el teorema 4.g implica que

$$\left| \int_0^1 f(t) \overline{g(t)} dt \right|^2 \leq \int_0^1 |f(t)|^2 dt \int_0^1 |g(t)|^2 dt.$$

El concepto de perpendicularidad es en extremo útil e importante en geometría. Introducimos su análogo en los espacios vectoriales con producto interior.

DEFINICIÓN: Si $u, v \in V$, entonces se dice que u es ortogonal a v si $(u, v) = 0$.

Nótese que si u es ortogonal a v , entonces v es ortogonal a u , pues $(v, u) = (\overline{u}, \overline{v}) = \overline{0} = 0$.

DEFINICIÓN. Si W es un subespacio de V , el *complemento ortogonal* de W , W^\perp , se define por $W^\perp = \{x \in V \mid (x, w) = 0 \text{ para toda } w \in W\}$.

LEMA 4.14. W^\perp es un subespacio de V .

Prueba. Si $a, b \in W^\perp$, entonces para cualesquiera $\alpha, \beta \in F$ y toda $w \in W$, $(\alpha a + \beta b, w) = \alpha(a, w) + \beta(b, w) = 0$, ya que $a, b \in W^\perp$.

Nótese que $w \in W \cap W^\perp = \{0\}$, pues si $w \in W \cap W^\perp$ debe ser ortogonal a sí misma, es decir, debemos tener $(w, w) = 0$. Las propiedades definitorias

de un espacio con producto interior descartan la posibilidad de tal resultado a menos que $w = 0$.

Una de nuestras metas es mostrar que $V = W + W^\perp$. Una vez que hayamos hecho esto, la nota acabada de hacer cobrará algún interés, pues que implicará que V es la suma directa de W y W^\perp .

DEFINICIÓN. El conjunto de vectores $\{v_i\}$ en V se dice que es *ortonormal* si y sólo si:

- 1) todo v_i es de longitud 1 (es decir, $(v_i, v_i) = 1$)
- 2) para $i \neq j$, $(v_i, v_j) = 0$.

LEMA 4.15. Si $\{v_i\}$ es un conjunto ortonormal, entonces los vectores en $\{v_i\}$ son linealmente independientes. Si $w = \alpha_1 v_1 + \dots + \alpha_n v_n$, entonces $\alpha_i = (w, v_i)$ para $i = 1, 2, \dots, n$.

Prueba. Supongamos que $\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n = 0$. Entonces $0 = (\alpha_1 v_1 + \dots + \alpha_n v_n, v_i) = \alpha_1 (v_1, v_i) + \dots + \alpha_n (v_n, v_i)$. Como $(v_j, v_i) = 0$ para $j \neq i$, mientras que $(v_i, v_i) = 1$, esta ecuación se reduce a $\alpha_i = 0$. Luego las v_j son linealmente independientes.

Si $w = \alpha_1 v_1 + \dots + \alpha_n v_n$, entonces un cálculo como el anterior nos da $(w, v_i) = \alpha_i$.

Análogo en espíritu y en prueba al lema 4.15 es el

LEMA 4.16. Si $\{v_1, \dots, v_n\}$ es un conjunto ortonormal en V y si $w \in V$, entonces $u = w - (w, v_1)v_1 - (w, v_2)v_2 - \dots - (w, v_i)v_i - \dots - (w, v_n)v_n$ es ortogonal a cada uno de los v_1, v_2, \dots, v_n .

Prueba. Calculando (u, v_i) para cualquier $i \leq n$, el uso de la ortonormalidad de v_1, \dots, v_n nos da el resultado.

La construcción que vamos a efectuar en la prueba del siguiente teorema es una que aparece y reaparece en muchas partes de la matemática. Es un procedimiento básico y se conoce como *proceso de ortogonalización de Gram-Schmidt*. Aunque aquí estaremos trabajando en un espacio de dimensión finita con producto interior, el proceso de Gram-Schmidt trabaja igualmente bien en situaciones de dimensión finita.

TEOREMA 4.H. Sea V un espacio de dimensión finita con producto interior; entonces V tiene un conjunto ortonormal como base.

Prueba. Sea V de dimensión n sobre F y sea v_1, \dots, v_n una base de V . Partiendo de esta base construiremos un conjunto ortonormal de n vectores;

según el lema 4.15 este conjunto es linealmente independiente de modo que formará una base de V .

Comenzamos con la construcción. Buscamos n vectores w_1, \dots, w_n cada uno de los cuales queremos que tenga longitud 1 y tales que para $i \neq j$, $(w_i, w_j) = 0$. En realidad, terminaremos por construirlos en la siguiente forma: w_1 será un múltiplo de v_1 , w_2 estará en el subespacio generado por w_1 y v_2 , w_3 en el generado por w_1, w_2 y v_3 , y más generalmente, w_i en el generado por w_1, w_2, \dots, w_{i-1} y v_i .

Sea

$$w_1 = \frac{v_1}{\|v_1\|};$$

entonces

$$(w_1, w_1) = \left(\frac{v_1}{\|v_1\|}, \frac{v_1}{\|v_1\|} \right) = \frac{1}{\|v_1\|^2} (v_1, v_1) = 1,$$

de donde $\|w_1\| = 1$. Nos preguntamos ahora: ¿para qué valores de α es $\alpha w_1 + v_2$ ortogonal a w_1 ? Todo lo que necesitamos es que $(\alpha w_1 + v_2, w_1) = 0$, es decir, $\alpha(w_1, w_1) + (v_2, w_1) = 0$. Como $(w_1, w_1) = 1$, $\alpha = -(v_2, w_1)$ nos resolverá tal ecuación. Sea $u_2 = -(v_2, w_1)w_1 + v_2$; u_2 es ortogonal a w_1 ; como v_1 y v_2 son linealmente independientes, w_1 y v_2 deben ser linealmente independientes, y por tanto $u_2 \neq 0$. Sea $w_2 = (u_2/\|u_2\|)$; entonces $\{w_1, w_2\}$ es un conjunto ortonormal. Continuemos. Sea $u_3 = -(v_3, w_1)w_1 - (v_3, w_2)w_2 + v_3$; una simple comprobación verifica que $(u_3, w_1) = (u_3, w_2) = 0$. Como w_1, w_2 y v_3 son linealmente independientes (pues w_1, w_2 están en el espacio generado por v_1 y v_2), $u_3 \neq 0$. Sea $w_3 = (u_3/\|u_3\|)$; entonces $\{w_1, w_2, w_3\}$ es un conjunto ortonormal. El camino que nos queda es ahora claro. Supongamos que hemos construido w_1, w_2, \dots, w_i , todos en subespacio generado por v_1, \dots, v_i , formando un conjunto ortonormal. ¿Cómo construimos el siguiente, w_{i+1} ? Simplemente construyendo $u_{i+1} = -(v_{i+1}, w_1)w_1 - (v_{i+1}, w_2)w_2 - \dots - (v_{i+1}, w_i)w_i + v_{i+1}$. Que $u_{i+1} \neq 0$ y que es ortogonal a cada uno de los w_1, \dots, w_i es de comprobación que dejamos como ejercicio para el lector. Finalmente, hacemos $w_{i+1} = (u_{i+1}/\|u_{i+1}\|)$.

De esta forma, dados r elementos linealmente independientes en V podemos construir un conjunto ortonormal que tenga r elementos. En particular, cuando $\dim V = n$, partiendo de cualquier base de V podemos construir un conjunto ortonormal que tenga n elementos. Y esto nos provee de la base requerida de V .

Ilustramos la construcción usada en la última prueba en un caso concreto. Sea F el campo de los números reales y sea V el conjunto de los polinomios en una variable x sobre F que tienen grado 2 o menor. Definimos en V un producto interior por: si $p(x), q(x) \in V$, entonces $(p(x), q(x)) =$

$\int_{-1}^1 p(x)q(x)dx$. Comencemos con la base $v_1 = 1, v_2 = x, v_3 = x^2$. Siguiendo la construcción que hemos visto

$$w_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{\int_{-1}^1 1 dx}} = \frac{1}{\sqrt{2}};$$

$$u_2 = -(v_2, w_1)w_1 + v_2,$$

que después de los cálculos se reduce a $u_2 = x$, y por tanto

$$w_2 = \frac{u_2}{\|u_2\|} = \frac{x}{\sqrt{\int_{-1}^1 x^2 dx}} = \frac{\sqrt{3}}{\sqrt{2}}x;$$

finalmente, $u_3 = -(v_3, w_1)w_1 - (v_3, w_2)w_2 + v_3 = \frac{-1}{3} + x^2$, y por tanto

$$w_3 = \frac{u_3}{\|u_3\|} = \frac{\frac{-1}{3} + x^2}{\sqrt{\int_{-1}^1 \left(\frac{-1}{3} + x^2\right)^2 dx}} = \frac{\sqrt{10}}{4}(-1 + 3x^2).$$

Mencionamos anteriormente el próximo teorema como uno de nuestros objetivos. Estamos ahora en posibilidad de probarlo.

TEOREMA 4.1. *Si V es un espacio de dimensión finita con producto interior y W un subespacio de V , entonces $V = W + W^\perp$. Aún más, V es la suma directa de W y W^\perp .*

Prueba. A causa de la naturaleza altamente geométrica del resultado, y por ser realmente básico, damos varias pruebas. En la primera, haremos uso del teorema 4.4 y de algunos de los anteriores lemas. A la segunda, la motivaremos geométricamente.

Primera prueba. Como subconjunto del espacio con producto interior V , W mismo es un espacio con producto interior (su producto interior no siendo otro que la restricción del de V a W). Así pues, podemos encontrar un conjunto ortonormal w_1, \dots, w_r en W que sea una base de W . Si $v \in V$, según el lema 4.16, $v_0 = v - (v, w_1)w_1 - (v, w_2)w_2 - \dots - (v, w_r)w_r$ es ortogonal a cada uno de los w_1, \dots, w_r y por tanto ortogonal a W . Así pues, $v_0 \in W^\perp$, y como $v = v_0 + ((v, w_1)w_1 + \dots + (v, w_r)w_r)$, $v \in W + W^\perp$. Por tanto, $V = W + W^\perp$. Como además $W \cap W^\perp = \{0\}$, esta suma es directa.

Segunda prueba. En esta prueba supondremos que F es el campo de los números reales. La prueba puede aplicarse, en casi la misma forma, para los números complejos; sin embargo, en este último caso se tiene que hacer frente a algunos detalles más que pueden tender a oscurecer las ideas centrales que se están utilizando.

Sea $v \in V$; supongamos que pudieramos encontrar un vector $w_0 \in W$ tal que $\|v - w_0\| \leq \|v - w\|$ para todo $w \in W$. Afirmamos que entonces $(v - w_0, w) = 0$ para todo $w \in W$, es decir, que $v - w_0 \in W^\perp$.

Si $w \in W$, entonces $w_0 + w \in W$, en consecuencia de lo cual

$$(v - w_0, v - w_0) \leq (v - (w_0 + w), v - (w_0 + w)).$$

Pero el primer miembro es $(w, w) + (v - w_0, v - w_0) - 2(v - w_0, w)$, lo que nos lleva a $2(v - w_0, w) \leq (w, w)$ para todo $w \in W$. Si m es un número positivo cualquiera, como $\frac{w}{m} \in W$, tenemos que

$$\frac{2}{m}(v - w_0, w) = 2\left(v - w_0, \frac{w}{m}\right) \leq \left(\frac{w}{m}, \frac{w}{m}\right) = \frac{1}{m^2}(w, w),$$

y por tanto $2(v - w_0, w) \leq (1/m)(w, w)$ para cualquier entero positivo m . Pero $(1/m)(w, w) \rightarrow 0$ cuando $m \rightarrow \infty$, de donde $2(v - w_0, w) \leq 0$. Análogamente, $-w \in W$, y por tanto $0 \leq -2(v - w_0, w) = 2(v - w_0, -w) \leq 0$, lo que nos dice que $(v - w_0, w) = 0$ para todo $w \in W$. Luego $v - w_0 \in W^\perp$, de donde $v \in w_0 + W^\perp \subset W + W^\perp$.

Para terminar con la segunda prueba debemos demostrar la existencia de un $w_0 \in W$ tal que $\|v - w_0\| \leq \|v - w\|$ para todo $w \in W$. Indicamos esquemáticamente dos formas de probar la existencia de un tal w_0 .

Sea u_1, \dots, u_k una base de W ; por tanto, cualquier $w \in W$ es de la forma $w = \lambda_1 u_1 + \dots + \lambda_k u_k$. Sea $\beta_{ij} = (u_i, u_j)$ y $\gamma_i = (v, u_i)$ para $v \in V$. Tenemos pues, $(v - w, v - w) = (v - \lambda_1 u_1 - \dots - \lambda_k u_k, v - \lambda_1 u_1 - \dots - \lambda_k u_k) = (v, v) - \sum \lambda_i \lambda_j \beta_{ij} - 2 \sum \lambda_i \gamma_i$. Esta función cuadrática en las λ es *no negativa* y por lo tanto, por los resultados del cálculo, tiene un mínimo. Las λ para este mínimo, $\lambda_1^{(0)}, \lambda_2^{(0)}, \dots, \lambda_k^{(0)}$, nos dan el vector buscado, $w_0 = \lambda_1^{(0)} u_1 + \dots + \lambda_k^{(0)} u_k$ en W .

Una segunda forma de exhibir un tal minimizante w , es como sigue. Definamos en V una métrica ζ por $\zeta(x, y) = \|x - y\|$; es fácil ver que ζ es una métrica propia sobre V , y V es ahora un espacio métrico. Sea $S = \{w \in W \mid \|v - w\| \leq \|v\|\}$; en esta métrica S es un conjunto compacto (pruébese) y por tanto la función continua $f(w) = \|v - w\|$ definida para $w \in S$ toma un mínimo para algún punto $w_0 \in S$. Dejamos al lector la verificación de que w_0 es el vector deseado, el que satisface $\|v - w_0\| \leq \|v - w\|$ para todo $w \in W$.

COROLARIO. Si V es un espacio de dimensión finita con producto interior y W es un subespacio de V , entonces $(W^\perp)^\perp = W$.

Prueba. Si $w \in W$, entonces para cualquier $u \in W^\perp$, $(w, u) = 0$, de donde $W \subset (W^\perp)^\perp$. Pero $V = W + W^\perp$ y $V = W^\perp + (W^\perp)$; de donde deducimos, como las sumas son sumas directas que $\dim(W) = \dim((W^\perp)^\perp)$. Como $W \subset (W^\perp)^\perp$ y es de la misma dimensión que $(W^\perp)^\perp$, de ello se sigue que $W = (W^\perp)^\perp$.

Problemas

En todos los problemas que siguen V es un espacio con producto interior sobre F .

1. Si F es el campo real y V es $F^{(3)}$, demuéstrese que la desigualdad de Schwarz implica que el coseno de un ángulo es de valor absoluto menor o igual que 1.

2. Si F es el campo real, encuéntrense todas las tétradas de números reales (a, b, c, d) tales que para $u = (\alpha_1, \alpha_2), v = (\beta_1, \beta_2) \in F^{(2)}$, $(u, v) = a\alpha_1\beta_1 + b\alpha_2\beta_2 + c\alpha_1\beta_2 + d\alpha_2\beta_1$, define un producto interior sobre $F^{(2)}$.

3. En V , definase la *distancia* $\zeta(u, v)$ de u a v por $\zeta(u, v) = \|u - v\|$. Pruébese que:

- 1) $\zeta(u, v) \geq 0$ y $\zeta(u, v) = 0$ si y sólo si $u = v$.
- 2) $\zeta(u, v) = \zeta(v, u)$.
- 3) $\zeta(u, v) \leq \zeta(u, w) + \zeta(w, v)$ (desigualdad del triángulo).

4. Si $\{w_1, \dots, w_m\}$ es un conjunto ortonormal en V , pruébese que

$$\sum_{i=1}^m |(w_i, v)|^2 \leq \|v\|^2 \text{ para cualquier } v \in V.$$

(desigualdad de Bessel)

5. Si V es de dimensión finita y si $\{w_1, \dots, w_m\}$ es un conjunto ortonormal en V tal que $\sum_{i=1}^m |(w_i, v)|^2 = \|v\|^2$ para todo $v \in V$, pruébese que $\{w_1, \dots, w_m\}$ debe ser una base de V .

6. Si $\dim V = n$ y si $\{w_1, \dots, w_m\}$ es un conjunto ortonormal en V , pruébese que existen vectores w_{m+1}, \dots, w_n , tales que $\{w_1, \dots, w_m, w_{m+1}, \dots, w_n\}$ es un conjunto ortonormal (y una base de V).

7. Úsese el resultado del problema 6 para dar otra prueba del teorema 4.i.

8. Pruébese en V la ley del paralelogramo:

$$\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2 + \|v\|^2).$$

Explíquese qué significa geométricamente en el caso especial $V = F^{(3)}$, donde F es el campo real, y donde el producto interior es el producto punto habitual.

9. Sea V el conjunto de las funciones reales $y = f(x)$ que satisfacen $d^2y/dx^2 + 9y = 0$.

a) Pruébese que V es un espacio vectorial real bidimensional.

b) Defínase en V , $(y, z) = \int_0^\pi yz \, dx$. Encuéntrese una base ortonormal en V .

10. Sea V el conjunto de las funciones reales $y = f(x)$ que satisfacen

$$\frac{d^3y}{dx^3} - 6 \frac{d^2y}{dx^2} + 11 \frac{dy}{dx} - 6y = 0.$$

a) Pruébese que V es un espacio vectorial real tridimensional.

b) Defínase en V , $(u, v) = \int_{-\infty}^0 uv \, dx$. Demuéstrese que esto define un producto interior sobre V y encuéntrese una base ortonormal para V .

11. Si W es un subespacio de V y $v \in V$ satisface $(v, w) + (w, v) \leq (w, w)$ para todo $w \in W$, pruébese que $(v, w) = 0$ para todo $w \in W$.

12. Si V es un espacio finito dimensional con producto interior y si f es una funcional lineal sobre V (es decir, si $f \in V'$), pruébese que hay un $u_0 \in V$ tal que $f(v) = (v, u_0)$ para todo $v \in V$.

5. MÓDULOS

El concepto de módulo es una generalización del de espacio vectorial; en lugar de restringir los escalares a que constituyan un campo permitimos en el caso de los módulos que constituyan tan solo un anillo.

Esta sección tiene muchas definiciones, pero solo un teorema importante. Pero las definiciones son tan cercanas en espíritu a las que ya hicimos para los vectores que las principales ideas que aquí debemos desarrollar no deberían estar enterradas en un mar de definiciones.

DEFINICIÓN. Sea R un anillo cualquiera; un conjunto no vacío M se dice que es un R -módulo (o un módulo sobre R) si M es un grupo abeliano

bajo una operación $+$, tal que para cada $r \in R$ y $m \in M$ existe un elemento rm en M de tal modo que se verifica:

- 1) $r(a+b) = ra+rb$
- 2) $r(sa) = (rs)a$
- 3) $(r+s)a = ra+sa$

para cualesquiera $a, b \in M$ y $r, s \in R$.

Si R tiene un elemento unitario, 1, y si $1m = m$ para todo elemento $m \in M$, entonces a M lo llamaremos un R -módulo **unitario**. Nótese que si R es un campo, un R -módulo unitario no es otra cosa que un espacio vectorial sobre R . *Todos nuestros módulos serán unitarios.*

Hablando propiamente, deberíamos haber llamado al objeto que acabamos de definir un R -módulo *izquierdo* porque permitimos la multiplicación por elementos de R tan solo por la izquierda. Análogamente, podríamos definir lo que debe entenderse por un R -módulo *derecho*. Nosotros no faremos tal distinción derecha-izquierda y convendremos en que por R -módulo entenderemos un R -módulo izquierdo.

EJEMPLO 1. Todo grupo abeliano G es un módulo sobre el anillo de los enteros.

Para verlo, nada más tenemos que escribir la operación de G como $+$ y convenir en que na , para $a \in G$ y n un entero, tenga el significado que le asignamos en el capítulo 2. Las reglas usuales de los exponentes en los grupos abelianos se traducen en las propiedades requeridas para hacer de G un módulo sobre los enteros. Nótese que es un módulo unitario.

EJEMPLO 2. Sea R un anillo cualquiera y sea M un ideal izquierdo de R . Para $r \in R$ y $m \in M$ sea rm el producto de estos elementos como elementos de R . La definición de ideal izquierdo implica que $rm \in M$, mientras que los axiomas que definen un anillo nos aseguran que M es un R -módulo. (En este ejemplo por un anillo entendemos un anillo asociativo para poder asegurar que $r(sm) = (rs)m$.)

EJEMPLO 3. El caso especial en que $R = M$; cualquier anillo R es un R -módulo sobre sí mismo.

EJEMPLO 4. Sea R un anillo cualquiera y sea λ un ideal izquierdo de R . Sea M el conjunto de todas las clases laterales $a + \lambda$, donde $a \in R$, de λ en R .

Definamos en $M(a + \lambda) + (b + \lambda) = (a + b) + \lambda$ y $r(a + \lambda) = ra + \lambda$. Puede probarse que M es un R -módulo. (Véase el problema 2 al final de esta sección.)

M se escribe usualmente como $R - \lambda$ (o, algunas veces, como R/λ) y se llama **módulo diferencia** (**cociente**) de R por λ .

Un subgrupo aditivo A del R -módulo M se llama **submódulo** de M si siempre que $r \in R$ y $a \in A$, entonces $ra \in A$.

Dado un R -módulo M y un submódulo A podríamos construir el módulo cociente M/A de una forma análoga a la forma en que construimos grupos cociente, anillos cociente y espacios cociente. Podríamos también hablar acerca de los homomorfismos de un R -módulo en otro, y probar los teoremas apropiados sobre homomorfismo. Hacemos esto en los problemas al final de esta sección.

Nuestro interés en los módulos marcha en una dirección un poco diferente; intentaremos encontrar una descomposición interesante para módulos sobre ciertos anillos.

DEFINICIÓN. Si M es un R módulo y M_1, \dots, M_s son submódulos de M , entonces se dice que M es la **suma directa** de M_1, \dots, M_s si todo elemento $m \in M$ puede escribirse de modo único como $m = m_1 + m_2 + \dots + m_s$ donde $m_1 \in M_1, m_2 \in M_2, \dots, m_s \in M_s$.

Como en el caso de espacios vectoriales, si M es la suma directa de M_1, \dots, M_s , entonces M debe ser isomorfo, como módulo, al conjunto de todos los s -tuples, (m_1, \dots, m_s) , donde el i -ésimo componente m_i es un elemento cualquiera de M_i , donde la adición es por componentes, y donde $r(m_1, \dots, m_s) = (rm_1, rm_2, \dots, rm_s)$ para $r \in R$. Así pues, conocer la estructura de cada M_i , nos capacitaría para conocer la estructura de M .

De particular interés y simplicidad son los módulos generados por un solo elemento; a tales módulos les llamamos **cíclicos**. Para ser precisos:

DEFINICIÓN. Un R -módulo M se dice que es **cíclico** si hay un elemento $m_0 \in M$ tal que todo $m \in M$ es de la forma $m = rm_0$, donde $r \in R$.

Para R , el anillo de los enteros, un módulo cíclico sobre R no es otra cosa que un grupo cíclico.

Necesitamos aún una definición más, a saber,

DEFINICIÓN. Un R -módulo M se dice que es **finitamente generado** si existen elementos $a_1, \dots, a_n \in M$ tales que todo m en M es de la forma $m = r_1 a_1 + r_2 a_2 + \dots + r_n a_n$.

Una vez hechas todas las definiciones que necesitábamos, vamos ahora al teorema que es la razón primaria por la cual esta sección existe. Se llama a menudo el **teorema fundamental sobre módulos finitamente generados** sobre anillos euclidianos. Exigiremos en él que R sea un anillo eucladiano (véase el capítulo 3, sección 7); pero el teorema se verifica en el contexto más general en que solo exigimos de R que sea un dominio de ideales principales.

TEOREMA 4.3. *Sea R un anillo eucladiano; entonces cualquier R -módulo finitamente generado, M , es la suma directa de un número finito de submódulos cíclicos.*

Prueba. Antes de dejarnos envolver con la operación de la prueba, veamos lo que el teorema expresa. La hipótesis de que M es finitamente generado nos dice que hay un conjunto de elementos $a_1, \dots, a_n \in M$ tales que todo elemento en M puede expresarse en la forma $r_1 a_1 + r_2 a_2 + \dots + r_n a_n$ donde las $r_i \in R$. La conclusión del teorema expresa que cuando R está propiamente condicionada podemos encontrar algún otro conjunto de elementos b_1, \dots, b_q en M tales que todo elemento $m \in M$ pueda expresarse en forma única como $m = s_1 b_1 + \dots + s_q b_q$ con $s_i \in R$. Una observación respecto a esa unicidad; no significa que los s_i son únicos, en realidad tal afirmación puede ser falsa; simplemente afirma que los elementos $s_i b_i$ sí son únicos. Es decir, si $m = s_1 b_1 + \dots + s_q b_q$ y $m = s'_1 b_1 + \dots + s'_q b_q$, de ello no podemos concluir que $s_1 = s'_1, s_2 = s'_2, \dots, s_q = s'_q$, sino que lo que podemos inferir es que $s_1 b_1 = s'_1 b_1, \dots, s_q b_q = s'_q b_q$.

Otra observación antes de que comencemos con el argumento técnico. Aunque el teorema se formula para un anillo eucladiano general, solo daremos la prueba con todos sus detalles para el caso especial del anillo de los enteros. Al final indicaremos las ligeras modificaciones necesarias para hacer que la prueba sirva para el caso más general. Hemos escogido este camino para impedir que las ideas generales, que son las mismas en todos los casos, queden oscurecidas por detalles técnicos que no son de importancia alguna.

Estamos, pues, simplemente suponiendo que M es un grupo abeliano que tiene un conjunto generador finito. Llaremos a aquellos conjuntos generadores que tienen un mínimo de elementos *conjuntos generadores mínimos* y al número de elementos en un conjunto generador mínimo el *rango de M* .

Nuestra prueba procede ahora por inducción sobre el rango de M .

Si el rango de M es 1, entonces M está generado por un solo elemento, de donde resulta que es cíclico; en este caso el teorema es cierto. Supongamos que el teorema es cierto para todos los grupos abelianos de rango $q-1$, y que M es de rango q .

Dado un conjunto generador mínimo a_1, \dots, a_q de M , si cualquier relación de la forma $n_1 a_1 + n_2 a_2 + \dots + n_q a_q = 0$ (n_1, \dots, n_q enteros) implica que $n_1 a_1 = n_2 a_2 = \dots = n_q a_q = 0$, entonces M es la suma directa de M_1, M_2, \dots, M_q , donde cada M_i es el módulo cíclico (es decir, el subgrupo) generado por a_i , con lo que ya tendríamos comprobado en tal caso el enunciado. Por consiguiente, dado un conjunto generador mínimo cualquiera b_1, \dots, b_q de M , debe haber enteros r_1, \dots, r_q tales que $r_1 b_1 + \dots + r_q b_q = 0$ y en el que no todos los $r_1 b_1, r_2 b_2, \dots, r_q b_q$ sean 0. De entre tales posibles relaciones para todos los conjuntos generadores mínimos, hay un entero

positivo posible mínimo que aparece como coeficiente. Sea este entero s_1 y sea el conjunto generador en el que ocurre a_1, \dots, a_q . Así pues

$$(1) \quad s_1 a_1 + s_2 a_2 + \dots + s_q a_q = 0.$$

Afirmamos que si $r_1 a_1 + \dots + r_q a_q = 0$ entonces $s_1 | r_1$; pues si $r_1 = ms_1 + t$, $0 \leq t < s_1$, multiplicando la ecuación (1) por m y restando de $r_1 a_1 + \dots + r_q a_q = 0$, esto nos lleva a $ta_1 + (r_2 - ms_2)a_2 + \dots + (r_q - ms_q)a_q = 0$; como $t < s_1$ y s_1 es el entero positivo posible mínimo en una tal relación, debemos tener $t = 0$.

Afirmamos ahora que $s_1 | s_i$ para $i = 2, \dots, q$. Supongamos que no fuera así; entonces $s_1 \nmid s_2$, por ejemplo, luego $s_2 = m_2 s_1 + t$, $0 < t < s_1$. Tenemos ahora que $a'_1 = a_1 + m_2 a_2, a_2, \dots, a_q$ también generan M y, además, $s_1 a'_1 + ta_2 + s_3 a_3 + \dots + s_q a_q = 0$; luego t aparece como coeficiente en alguna relación entre elementos de un conjunto generador mínimo. Pero esto implica, por la misma elección de s_1 , que o $t = 0$ o $t \geq s_1$. Nos tenemos que quedar con $t = 0$, es decir, con que $s_1 | s_2$. Y, análogamente, para las otras s_i . Escribamos $s_i = m_i s_1$.

Consideremos los elementos $a_1^* = a_1 + m_2 a_2 + m_3 a_3 + \dots + m_q a_q, a_2, \dots, a_q$. Es claro que generan M ; además, $s_1 a_1^* = s_1 a_1 + m_2 s_1 a_2 + \dots + m_q s_1 a_q = s_1 a_1 + s_2 a_2 + \dots + s_q a_q = 0$. Si $r_1 a_1^* + r_2 a_2 + \dots + r_q a_q = 0$, sustituyendo a a_1^* , obtenemos una relación entre a_1, \dots, a_q en que el coeficiente de a_1 es r_1 ; luego $s_1 | r_1$ y por lo tanto $r_1 a_1^* = 0$. Si M_1 es el módulo cíclico generado por a_1^* y M_2 es el submódulo de M generado por a_2, \dots, a_q , lo que acabamos de demostrar es que $M_1 \cap M_2 = (0)$. Pero $M_1 + M_2 = M$, ya que a_1^*, a_2, \dots, a_q generan M . Luego M es la suma directa de M_1 y M_2 . Como M_2 está generado por a_2, \dots, a_q , su rango es, cuando más, $q-1$ (en realidad, es $q-1$), de forma que, según la hipótesis de inducción, M_2 es la suma directa de módulos cíclicos. Con lo que, en resumen, hemos descompuesto a M en suma directa de módulos cíclicos.

COROLARIO. *Cualquier grupo abeliano finito es el producto (suma) directo de grupos cíclicos.*

Prueba. El grupo abeliano finito G es claro que está finitamente generado, puesto que lo genera el conjunto finito de todos sus elementos. La aplicación del teorema 4.j nos da el corolario.

Supongamos que R es un anillo euclidiano con la función de Euclides d . Modificamos la prueba dada para los enteros para una válida para R como sigue.

- 1) En lugar de escoger s_1 como el entero positivo mínimo posible que aparece en cualquier relación entre los elementos de un conjunto generador, lo escogemos como aquel elemento de R entre los de valor d mínimo que aparece en cualquier relación.

- 2) En la prueba de que $s_1 | r_1$ en cualquier relación $r_1 a_1 + \dots + r_q a_q = 0$, el único cambio necesario es que $r_1 = ms_1 + t$ donde

$$t = 0 \text{ o } d(t) < d(s_1);$$

el resto es igual. Lo mismo ha de decirse para la prueba de que $s_1 | s_i$.

Así pues, con estos cambios menores la prueba es válida para anillos euclidianos en general, quedando, pues, el teorema 4.j completamente probado.

Volvemos ahora a los grupos abelianos finitos. Sea G un grupo abeliano finito de orden $p_1^{n_1} p_2^{n_2} \dots p_k^{n_k}$, donde las p_i son primos distintos. Entonces G es el producto directo de sus subgrupos de Sylow S_{p_1}, \dots, S_{p_k} . Para lo que vamos a decir, será suficiente que de lo anterior consideremos tan solo el caso en que G es de orden p^n , con p número primo.

G es el producto directo de grupos cíclicos G_1, \dots, G_k de órdenes p^{n_1}, \dots, p^{n_k} respectivamente, donde $n_1 \geq n_2 \geq \dots \geq n_k$. Llámemos a n_1, n_2, \dots, n , los *invariantes* de G . (Muy a menudo es a $p^{n_1}, p^{n_2}, \dots, p^{n_k}$ a los que se llaman los invariantes.) ¿Qué puede decirse acerca de éstos?

El orden de $G_1 G_2$ es

$$\frac{o(G_1)o(G_2)}{o(G_1 \cap G_2)},$$

y como el producto de G_1 y G_2 es directo, $G_1 \cap G_2 = (e)$; es decir, $o(G_1 G_2) = o(G_1)o(G_2) = p^{n_1}p^{n_2} = p^{n_1+n_2}$. Continuando de este modo obtenemos $p^n = o(G) = o(G_1 G_2 \dots G_k) = p^{n_1+\dots+n_k}$. Luego $n_1 + n_2 + \dots + n_k = n$. En los términos previamente usados (capítulo 2, sección 11) n_1, \dots, n_k forman una partición de n .

Dado un grupo abeliano de orden p^n , llegamos a una partición de n . Por otra parte, dada una partición de n , $n = n_1 + n_2 + \dots + n_k$, podemos construir un grupo abeliano de orden p^n de la siguiente forma: sea G_1 un grupo cíclico de orden p^{n_1} , G_2 uno de orden p^{n_2} , ..., G_k uno de orden p^{n_k} ; sea G el producto directo exterior de G_1, \dots, G_k . G es un grupo abeliano de orden p^n .

Por tanto, para cada partición hay un grupo abeliano y para cada grupo abeliano una partición. Si mostramos que los invariantes de G caracterizan a G hasta isomorfismo, tendremos una correspondencia biyectiva entre las particiones de n y los grupos abelianos no isomorfos de orden p^n . Puede demostrarse que este es el caso, es decir, que G es isomorfo a G_1 , donde ambos son grupos abelianos de orden p^n , si y sólo si tienen los mismos invariantes (véase el problema 15 al final de esta sección).

Sea $p(n)$ el número de particiones de n . Entonces

TEOREMA 4.K. *El número de grupos no abelianos no isomorfos de orden p^n es $p(n)$.*

COROLARIO. El número de grupos abelianos no isomorfos de orden $p_1^{n_1} p_2^{n_2} \dots p_k^{n_k}$, todos los p_i primos distintos, es $p(n_1)p(n_2) \dots p(n_k)$.

Problemas

1. Verifíquese que la afirmación hecha en el ejemplo 1 de que todo grupo abeliano es un módulo sobre el anillo de los enteros, es cierta.

2. Verifíquese que el conjunto del ejemplo 4 es un R -módulo.

3. Supongamos que R es un anillo con elemento unidad y que M es un módulo sobre R , pero que no es unitario. Pruébese que existe una $m \neq 0$ en M tal que $rm = 0$ para todo $r \in R$.

Dados dos R módulos M y N , entonces la aplicación T de M en N se llama *homomorfismo* (o *R -homomorfismo*, o *homomorfismo de módulos*) si

$$a) (m_1 + m_2)T = m_1 T + m_2 T$$

$$b) (rm_1)T = r(m_1 T)$$

para $m_1, m_2 \in M$ cualesquiera, y todo $r \in R$.

4. Si T es un homomorfismo de M en N , sea $K(T) = \{x \in M \mid xT = 0\}$. Pruébese que $K(T)$ es un submódulo de M y que $I(T) = \{xT \mid x \in M\}$ es un submódulo de N .

5. El homomorfismo T se dice que es un *isomorfismo* si es inyectivo. Pruébese que T es un isomorfismo si y sólo si $K(T) = 0$.

6. Sean M, N, Q tres R -módulos, y sean T un homomorfismo de M en N y S un homomorfismo de N en Q . Defínase $TS : M \rightarrow Q$ por $m(TS) = (mT)S$ para cualquier $m \in M$. Pruébese que TS es un R -homomorfismo de M en Q y determíñese su núcleo.

7. Si M es un R -módulo y A es un submódulo de M , defínase el módulo cociente M/A (úsele lo análogo en grupos, anillos y espacios vectoriales como guía) de forma que sea un R -módulo y pruébese que hay un R -homomorfismo de M sobre M/A .

8. Si T es un homomorfismo de M sobre N con $K(T) = A$, pruébese que N es isomorfo (como módulo) a M/A .

9. Si A y B son submódulos de M pruébese que:

a) $A \cap B$ es un submódulo de M .

b) $A+B = \{a+b \mid a \in A, b \in B\}$ es un submódulo de M .

c) $(A+B)/B$ es isomorfo a $A/(A \cap B)$.

10. Un R -módulo M se dice que es *irreducible* si sus únicos submódulos son (0) y M . Pruébese que todo R -módulo unitario irreducible es cíclico.

11. Si M es un R -módulo irreducible pruébese que M es cíclico o que para todo $m \in M$ y $r \in R$, $rm = 0$.

*12. Si M es un R -módulo irreducible tal que $rm \neq 0$ para algún $r \in R$ y algún $m \in M$, pruébese que cualquier R -homomorfismo T de M en M es un isomorfismo de M sobre M o que $mT = 0$ para todo $m \in M$.

13. Sea M un R -módulo y sea $E(M)$ el conjunto de todos los R -homomorfismos de M en M . Háganse las definiciones apropiadas de adición y multiplicación de elementos de $E(M)$ de forma que $E(M)$ se haga un anillo. (Sugerencia: imítense lo que se ha hecho para $\text{Hom}(V, V)$, para V un espacio vectorial.)

*14. Si M es un R -módulo irreducible tal que $rm \neq 0$ para algún $r \in R$ y algún $m \in M$, pruébese que $E(M)$ es un anillo con división. (A este resultado se le conoce como *lema de Schur*.)

*15. Demuéstrese que cualesquiera dos grupos abelianos finitos que tienen los mismos invariantes son isomorfos.

16. ¿Cuántos grupos abelianos no isomorfos hay de orden 2^5 ? Dígase como se construirían todos.

17. Describábase cómo se construirían todos los grupos abelianos de orden 16 200. ¿Cuántos grupos no isomorfos hay de tal orden?

18. Proporcionérese una prueba completa del teorema 4.j para módulos finitamente generados sobre anillos euclidianos.

19. Sea M un R -módulo; si $m \in M$, sea $\lambda(m) = \{x \in R \mid xm = 0\}$. Demuéstrese que $\lambda(m)$ es un ideal izquierdo de R . Se le llama el *orden* de m .

20. Si λ es un ideal izquierdo de R y si M es un R -módulo, pruébese que para $m \in M$, $\lambda m = \{xm \mid x \in \lambda\}$ es un submódulo de M .

*21. Sea M un R -módulo irreducible en el que $rm \neq 0$ para algún $r \in R$ y $m \in M$. Sea $m_0 \neq 0 \in M$ y sea $\lambda(m_0) = \{x \in R \mid xm_0 = 0\}$.

a) Pruébese que $\lambda(m_0)$ es un ideal izquierdo máximo de R (es decir, que si λ es un ideal izquierdo de R tal que $R \supset \lambda \supset \lambda(m_0)$, entonces $\lambda = R$ o $\lambda = \lambda(m_0)$).

b) Pruébese que, como R -módulos, M es isomorfo a $R - \lambda(m_0)$ (véase el ejemplo 4).

Lecturas supplementarias

HALMOS, P. R., *Finite-Dimensional Vector Spaces*, segunda edición. D. Van Nostrand Company, Inc., Princeton, 1958.

CAPITULO

5

Campos

EN NUESTRA discusión de anillos *remarcamos* una clase especial a cuyos elementos llamamos campos. Un campo, recordemos, es un anillo conmutativo con elemento unidad en el que todo elemento distinto de cero tiene un inverso multiplicativo. Dicho de otra manera, un campo es un anillo conmutativo en el que podemos dividir por cualquier elemento distinto de cero.

Los campos juegan un papel central en álgebra. Por una parte, los resultados respecto a ellos encuentran importantes aplicaciones en la teoría de números. Por otra parte, su teoría enmarca el tema de la teoría de ecuaciones que trata cuestiones acerca de las raíces de polinomios.

En nuestro desarrollo tocaremos solo ligeramente el campo de los

números algebraicos. En lugar de ello, nuestro mayor énfasis recaerá sobre los aspectos de la teoría de campos que tienen importancia para la teoría de ecuaciones. Aunque no trataremos el tema en su forma más amplia ni más general, llegaremos en él suficientemente lejos para poder introducir algunas de las bellas ideas, debidas al brillante matemático francés Evaristo Galois, que han servido como inspiración y guía para el álgebra como es en la actualidad.

1. EXTENSIÓN DE CAMPOS

En esta sección nos ocuparemos de la relación de un campo con otro. Sea F un campo; un campo K se dice que es una *extensión* de F si K contiene a F . Es decir, K es una extensión de F si F es un subcampo de K . A lo largo de todo este capítulo, F denotará un campo dado y K una extensión de F .

Como antes hemos señalado en el capítulo sobre espacios vectoriales, si K es una extensión de F , entonces, bajo las operaciones ordinarias del campo K , K es un espacio vectorial sobre F . Como espacio vectorial podemos hablar de dependencia lineal, dimensión, bases, etc., en K respecto a F .

DEFINICIÓN. El *grado* de K sobre F es la dimensión de K como espacio vectorial sobre F .

Denotaremos siempre el grado de K sobre F por $[K : F]$. Para nosotros es de particular interés el caso en que $[K : F]$ es finito, es decir, el caso en que K es de dimensión finita como espacio vectorial sobre F . Esta situación se describe diciendo que K es una *extensión finita* de F .

Comenzamos con un resultado acerca de las extensiones finitas, relativamente sencillo, pero, al mismo tiempo, altamente efectivo.

TEOREMA 5.A. Si L es una extensión finita de K y K una extensión finita de F , entonces L es una extensión finita de F . Además, $[L : F] = [L : K][K : F]$.

Prueba. La estrategia que vamos a emplear en la prueba es la de escribir explícitamente una base de L sobre F . En esta forma no solo demostramos que L es una extensión finita de F , sino que, ciertamente, probamos el resultado más preciso y que constituye realmente el corazón del teorema, que $[L : F] = [L : K][K : F]$.

Supongamos, entonces, que $[L : K] = m$ y que $[K : F] = n$. Sea v_1, \dots, v_m una base de L sobre K y w_1, \dots, w_n una base de K sobre F . ¿Qué podría ser más elegante y natural que el que los elementos $v_i w_j$, donde $i = 1, \dots, m$ y $j = 1, \dots, n$, sirvieran como base de L sobre F ? Al menos tal colección tiene el número exacto de elementos. Procedemos ahora a demostrar que es cierto que forman una base de L sobre F . ¿Qué necesitamos para que

esto quede establecido? Debemos primero demostrar que todo elemento de L es una combinación lineal de estos mn elementos con coeficientes en F , y a continuación debemos demostrar que estos mn elementos son linealmente independientes sobre F .

Sea t un elemento cualquiera de L . Como todo elemento de L es una combinación lineal de v_1, \dots, v_m con coeficientes en K , el elemento t debe ser en particular de esa forma. Luego $t = k_1 v_1 + \dots + k_m v_m$, donde los elementos k_1, \dots, k_m están todos en K . Pero todo elemento de K es una combinación lineal de w_1, \dots, w_n con coeficientes en F . Luego $k_1 = f_{11} w_1 + \dots + f_{1n} w_n, \dots, k_i = f_{i1} w_1 + \dots + f_{in} w_n, \dots, k_m = f_{m1} w_1 + \dots + f_{mn} w_n$, donde todos los f_{ij} están en F .

Sustituyendo por estas expresiones a k_1, \dots, k_m en $t = k_1 v_1 + \dots + k_m v_m$ obtenemos $t = (f_{11} w_1 + \dots + f_{1n} w_n) v_1 + \dots + (f_{m1} w_1 + \dots + f_{mn} w_n) v_m$. Efectuando las multiplicaciones indicadas y usando las leyes distributiva y asociativa, llegamos finalmente a $t = f_{11} v_1 w_1 + \dots + f_{1n} v_1 w_n + \dots + f_{i1} v_i w_1 + \dots + f_{mn} v_m w_n$. Como los f_{ij} están en F , hemos expresado t como combinación lineal sobre F de los elementos $v_i w_j$. Por tanto, los elementos $v_i w_j$ generan ciertamente a L sobre F y, por tanto, satisfacen la primera propiedad que se requiere de una base.

Debemos demostrar que los elementos $v_i w_j$ son linealmente independientes sobre F . Supongamos que $f_{11} v_1 w_1 + \dots + f_{1n} v_1 w_n + \dots + f_{i1} v_i w_1 + \dots + f_{mn} v_m w_n = 0$, donde los f_{ij} están, todos, en F . Nuestro objetivo es probar que $f_{ij} = 0$ para i, j cualesquiera. Reagrupando la anterior expresión se tiene $(f_{11} w_1 + \dots + f_{1n} w_n) v_1 + \dots + (f_{i1} w_1 + \dots + f_{in} w_n) v_i + \dots + (f_{m1} w_1 + \dots + f_{mn} w_n) v_m = 0$.

Como las w_i están en K y como $K \supset F$, todos los elementos $k_i = f_{i1} w_1 + \dots + f_{in} w_n$ están en K . Y tenemos $k_1 v_1 + \dots + k_m v_m = 0$ con $k_1, \dots, k_m \in K$. Pero, por hipótesis, v_1, \dots, v_m forman una base de L sobre K , luego, en particular, deben ser linealmente independientes sobre K . El resultado final de esto es que $k_1 = k_2 = \dots = k_m = 0$. Usando los valores explícitos de los k_i tenemos

$$f_{i1} w_1 + \dots + f_{in} w_n = 0 \quad \text{para } i = 1, 2, \dots, m.$$

Pero si invocamos ahora el hecho de que los w_i son linealmente independientes sobre F , llegamos a la conclusión de todo los f_{ij} han de ser nulos. En otras palabras, hemos probado que los $v_i w_j$ son linealmente independientes sobre F . Y de esa forma satisfacen la otra propiedad requerida por una base.

Hemos logrado, por fin, probar que los mn elementos $v_i w_j$ forman una base de L sobre F . Luego $[L : F] = mn$; como $m = [L : K]$ y $n = [\bar{K} : F]$, hemos obtenido el resultado buscado, que $[L : F] = [L : K][K : F]$.

Supongamos que L, K, F son tres campos en la relación $L \supset K \supset F$, y supongamos, además, que $[L : F]$ es finito. Es claro que cualesquier elementos sobre L linealmente independientes sobre K son también lineal-

mente independientes sobre F . Luego la hipótesis de que $[L : F]$ es finito fuerza la conclusión de que $[L : K]$ es también finito. Además, como K es un subespacio de L , $[K : F]$ es finito. Por el teorema, $[L : F] = [L : K][K : F]$, de donde $[K : F][L : F]$. Y hemos probado el siguiente

COROLARIO. *Si L es una extensión finita de F y si K es un subcampo de L que contiene a F , entonces $[K : F] \mid [L : F]$.*

Así, por ejemplo, si $[L : F]$ es un número primo, entonces no puede haber ningún campo propiamente entre F y L . Dentro de poco, en la sección 4, cuando discutamos la construcción de ciertas figuras geométricas con regla y compás, el corolario será muy significativo.

DEFINICIÓN. Un elemento $a \in K$ se dice que es *algebraico sobre F* si existen elementos $\alpha_0, \alpha_1, \dots, \alpha_n$ en F , no todos 0, tales que $\alpha_0 a^n + \alpha_1 a^{n-1} + \dots + \alpha_n = 0$.

Si el polinomio $q(x) \in F[x]$, el anillo de los polinomios en x sobre F , y si $q(x) = \beta_0 x^m + \beta_1 x^{m-1} + \dots + \beta_m$, entonces, para cualquier elemento $b \in K$, por $q(b)$ entenderemos el elemento $\beta_0 b^m + \beta_1 b^{m-1} + \dots + \beta_m$ de K . En la expresión comúnmente usada, $q(b)$ es el *valor* del polinomio $q(x)$ obtenido al sustituir la x por la b . El elemento b se dice que *satisface* $q(x)$ si $q(b) = 0$. En estos términos, $a \in K$ es algebraico sobre F si existe un polinomio distinto de cero $p(x) \in F[x]$ que a satisface, es decir, para el cual $p(a) = 0$.

Sea K una extensión de F y sea a de K . Sea \mathfrak{M} la colección de todos los subcampos de K que contienen tanto a F como a a . \mathfrak{M} no es vacío, pues el propio K es un elemento de \mathfrak{M} . Pero, como es fácil probar, la intersección de cualquier número de subcampos de K es ella misma un subcampo de K . Luego la intersección de todos aquellos subcampos de K que son miembros de \mathfrak{M} es un subcampo de K . Denotamos a este subcampo por $F(a)$. ¿Cuáles son sus propiedades? Ciertamente, contiene tanto a F como a a , pues ello es cierto para todos los subcampos de K que son miembros de \mathfrak{M} . Por otra parte, por la misma definición de intersección, todo subcampo de K en \mathfrak{M} contiene a $F(a)$, y $F(a)$ mismo está en \mathfrak{M} . Luego $F(a)$ es el *mínimo subcampo de K que contiene tanto a F como a a* . Llamamos a $F(a)$ el subcampo obtenido por la *adición de a a F* .

Nuestra descripción de $F(a)$, por el momento, ha sido puramente externa. Damos ahora una descripción alternativa más constructiva de $F(a)$. Consideremos todos los elementos de K que pueden expresarse en la forma $\beta_0 + \beta_1 a + \dots + \beta_s a^s$; aquí las β pueden tomar valores cualesquiera sobre F , y s puede ser cualquier entero no negativo. Como elementos en K , tal elemento puede dividirse por otro con tal de que el último sea distinto de cero. Sea U el conjunto de todos los cocientes. Dejamos como ejercicio probar que U es un subcampo de K .

Por una parte, U contiene ciertamente a F y a a , de donde $U \supset F(a)$. Por otra, cualquier subcampo de K que contiene tanto a F como a a , en virtud de que como subcampo ha de ser cerrado respecto a la adición y a la multiplicación, debe contener todos los elementos $\beta_0 + \beta_1 a + \dots + \beta_s a^s$ para $\beta_i \in F$ cualesquiera. Así pues, $F(a)$ debe contener todos estos elementos y, por ser un subcampo de K , también debe contener a los cocientes de tales elementos. Por tanto, $F(a) \supset U$. Las dos relaciones, $U \subset F(a)$ y $U \supset F(a)$, implican, desde luego, que $U = F(a)$. De esta forma hemos obtenido una construcción interna de $F(a)$, a saber, la U .

Ligamos ahora la propiedad de que $a \in K$ sea algebraica sobre F con propiedades macroscópicas del campo $F(a)$ mismo.

TEOREMA 5.B. *El elemento $a \in K$ es algebraico sobre F si y sólo si $F(a)$ es una extensión finita de F .*

Prueba. Como suele suceder con muchas de las proposiciones del tipo “si y sólo si”, una mitad de la prueba será directa y simple, mientras que la otra mitad será más profunda y complicada.

Supongamos que $F(a)$ es una extensión finita de F y que $[F(a) : F] = m$. Consideremos los elementos $1, a, a^2, \dots, a^m$; todos están en $F(a)$ y son $m+1$. De acuerdo con el lema 4.6, estos elementos son linealmente dependientes sobre F . Por tanto, hay elementos $\alpha_0, \alpha_1, \dots, \alpha_m$ en F , no todos 0, tales que $\alpha_0 1 + \alpha_1 a + \alpha_2 a^2 + \dots + \alpha_m a^m = 0$. Luego a es algebraico sobre F y satisface el polinomio distinto de cero $p(x) = \alpha_0 + \alpha_1 x + \dots + \alpha_m x^m$ en $F[x]$ de grado cuando más $m = [F(a) : F]$. Y esto prueba la parte “si” del teorema.

Pasamos ahora a la parte “sólo si”. Supongamos que a en K es algebraica sobre F . Por hipótesis, a satisface algún polinomio distinto de cero en $F[x]$; sea $p(x)$ uno de los polinomios en $F[x]$ de grado positivo mínimo para los que $p(a) = 0$. Afirmamos que $p(x)$ es irreducible sobre F . Pues supongamos que $p(x) = f(x)g(x)$ donde $f(x), g(x) \in F[x]$; entonces, $0 = p(a) = f(a)g(a)$ (véase el problema 1) y, como $f(a)$ y $g(a)$ son elementos del campo K , el hecho de que su producto sea cero obliga a que $f(a) = 0$ o $g(a) = 0$. Como $p(x)$ es de grado positivo mínimo entre los que toman en a el valor cero, debemos concluir que debe verificarse o que $\deg f(x) \geq \deg p(x)$ o que $\deg g(x) \geq \deg p(x)$. Pero esto prueba la irreductibilidad de $p(x)$.

Definimos la aplicación ψ de $F[x]$ en $F(a)$ como sigue. Para cualquier $h(x) \in F[x]$, $h(x)\psi = h(a)$. Dejamos al lector verificar que ψ es un homomorfismo entre anillos del anillo $F[x]$ en el campo $F(a)$ (véase el problema 1). ¿Qué es V el núcleo de ψ ? Por la definición de ψ , $V = \{h(x) \in F[x] \mid h(a) = 0\}$. Además, $p(x)$ es un elemento de grado mínimo en el ideal V de $F[x]$. Por los resultados de la sección 9, capítulo 3, todo elemento en V es un múltiplo de $p(x)$, y como $p(x)$ es irreducible, según el lema 3.22, V es un ideal máximo de $F[x]$. De acuerdo con el teorema 3.b, $F[x]/V$ es un campo.

Pero, según el teorema general de homomorfismos para anillos (teorema 3.a), $F[x]/V$ es isomorfo a la imagen de $F[x]$ bajo ψ . Resumiendo, hemos mostrado que la imagen de $F[x]$ bajo ψ es un subcampo de $F(a)$ que contiene tanto a F como a K ; según la propia definición de $F(a)$ nos vemos obligados a concluir que la imagen de $F[x]$ bajo ψ es todo $F(a)$. En pocas palabras, $F[x]/V$ es isomorfo a $F(a)$.

Pero $V = (p(x))$, el ideal generado por $p(x)$; y por esto afirmamos que la dimensión de $F[x]/V$, como espacio vectorial sobre F , es precisamente igual a $\deg p(x)$ (véase el problema 2). En vista del isomorfismo entre $F[x]/V$ y $F(a)$ obtenemos el hecho de que $[F(a) : F] = \deg p(x)$. Por tanto, $[F(a) : F]$ es seguramente finito; y es este el contenido de la parte “sólo si” del teorema. Nótese que, realmente, hemos probado aún más, a saber, que $[F(a) : F]$ es igual al grado del polinomio de grado mínimo satisfecho por a sobre F .

La prueba que acabamos de dar ha sido de desarrollo un poco lento, pero esto ha sido deliberado. La ruta seguida contiene importantes ideas y liga conceptos y resultados desarrollados anteriormente en nuestra exposición. Ninguna parte de las matemáticas es una isla.

Vamos ahora a rehacer la parte “sólo si” trabajando más en el interior de $F(a)$. Este “rehacer” es, en realidad, idéntico a la prueba que acabamos de dar; la única diferencia es que las distintas piezas están unidas en forma un poco diferente.

Sea $p(x)$ de nuevo un polinomio sobre F de los de grado positivo mínimo satisfechos por a . Un tal polinomio se llama *polinomio mínimo* para a sobre F . Podemos suponer que el coeficiente de la potencia máxima de x es 1, es decir, que $p(x)$ es mónico; en tal caso podemos hablar de “el polinomio mínimo” para a sobre F , pues cualesquiera dos polinomios mínimos monicos para a sobre F son iguales. (Pruébese.) Supongamos que $p(x)$ es de grado n ; tenemos, pues, $p(x) = x^n + \alpha_1 x^{n-1} + \dots + \alpha_n$ donde las α_i están en F . Por hipótesis, $a^n + \alpha_1 a^{n-1} + \dots + \alpha_n = 0$, de donde $a^n = -\alpha_1 a^{n-1} - \alpha_2 a^{n-2} - \dots - \alpha_n$. ¿Qué hay acerca de a^{n+1} ? Por lo anterior, $a^{n+1} = -\alpha_1 a^n - \alpha_2 a^{n-1} - \dots - \alpha_n a$; si sustituimos en el segundo miembro de esta relación a a^n por la expresión anteriormente hallada, realizamos a a^{n+1} como una combinación lineal de los elementos 1, a , ..., a^{n-1} sobre F . Continuando con este proceso, obtenemos que a^{n+k} para $k \geq 0$ es una combinación lineal sobre F de 1, a , a^2 , ..., a^{n-1} .

Consideremos ahora $T = \{\beta_0 + \beta_1 a + \dots + \beta_{n-1} a^{n-1} \mid \beta_0, \beta_1, \dots, \beta_{n-1} \in F\}$. Es inmediato que T es cerrado respecto a la adición y en vista de las observaciones hechas en el párrafo anterior también es claro que es cerrado respecto a la multiplicación. Cualquier otra cosa que, además, pueda ser, cuando menos hemos demostrado que T es un anillo. Además, T contiene a F y a a . Deseamos ahora demostrar que T es más que solamente un anillo, que T es, en realidad, un campo.

Sea $u \neq 0$, $u = \beta_0 + \beta_1 a + \dots + \beta_{n-1} a^{n-1}$ un elemento de T , y $h(x) = \beta_0 + \beta_1 x + \dots + \beta_{n-1} x^{n-1} \in F[x]$. Como $u \neq 0$, y $u = h(a)$, tenemos que $h(a) \neq 0$, de donde $p(x) \nmid h(x)$. Por la irreducibilidad de $p(x)$, $p(x)$ y $h(x)$ deben, por tanto, ser primos relativos. Luego podemos encontrar polinomios $s(x)$ y $t(x)$ en $F[x]$ tales que $p(x)s(x) + h(x)t(x) = 1$. Pero, entonces, $1 = p(a)s(a) + h(a)t(a) = h(a)t(a)$, ya que $p(a) = 0$; teniendo en cuenta que $u = h(a)$ obtenemos $ut(a) = 1$. La inversa de u es, pues, $t(a)$; en $t(a)$ todas las potencias de a mayores que $n-1$ pueden reemplazarse por combinaciones lineales de $1, a, \dots, a^{n-1}$ sobre F , de donde $t(a) \in T$. Hemos demostrado que todo elemento distinto de cero de T tiene su inverso en T ; por consiguiente, T es un campo. Pero $T \subset F(a)$ y F y a están, ambos, contenidos en T , de donde resulta que $T = F(a)$. Hemos identificado $F(a)$ como el conjunto de todas las expresiones $\beta_0 + \beta_1 a + \dots + \beta_{n-1} a^{n-1}$.

Tenemos, pues, que T es generado sobre F por los elementos $1, a, \dots, a^{n-1}$, luego, por consecuencia, $[T : F] \leq n$. Pero los elementos $1, a, a^2, \dots, a^{n-1}$ son linealmente independientes sobre F , pues cualquier relación de la forma $\gamma_0 + \gamma_1 a + \dots + \gamma_{n-1} a^{n-1}$, con los elementos $\gamma_i \in F$, lleva a la conclusión de que a satisface el polinomio $\gamma_0 + \gamma_1 x + \dots + \gamma_{n-1} x^{n-1}$ sobre F de grado menor que n . Esta contradicción prueba la independencia lineal de $1, a, \dots, a^{n-1}$ y, por tanto, estos elementos forman realmente una base de T sobre F , de donde, en realidad, sabemos ahora que $[T : F] = n$. Como $T = F(a)$, de ellos se sigue el resultado $[F(a) : F] = n$.

DEFINICIÓN. El elemento $a \in K$ se dice que es *algebraico de grado n* sobre F si satisface un polinomio distinto de cero sobre F de grado n , pero no satisface a ningún polinomio de grado menor.

En el transcurso de la prueba del teorema 5.b (en cada una de las pruebas que dimos), probamos un resultado un poco más sólido que el que se enuncia en el teorema, a saber,

TEOREMA 5.c. Si $a \in K$ es algebraico de grado n sobre F , entonces $[F(a) : F] = n$.

Este resultado se adapta a muchos usos. Damos a continuación como una consecuencia inmediata de él, el siguiente muy interesante

TEOREMA 5.d. Si a, b en K son algebraicos sobre F , entonces $a \pm b$, ab y a/b (si $b \neq 0$) son todos algebraicos sobre F . En otras palabras, los elementos en K que son algebraicos sobre F forman un subcampo de K .

Prueba. Supongamos que a es algebraico de grado m sobre F mientras que b es algebraico de grado n sobre F . Por el teorema 5.c el subcampo $T = F(a)$ de K es de grado m sobre F . Como b es algebraico de grado

n sobre F , a fortiori es algebraico de grado cuando más n sobre T . Pero $[W : F] = [W : T][T : F]$ por el teorema 5.a; por tanto, $[W : F] \leq mn$ y W es, pues, una extensión finita de F . Pero a y b están, ambos, en W , de donde $a \pm b$, ab y a/b están en W . Por el teorema 5.b; como $[W : F]$ es finito, estos elementos deben ser algebraicos sobre F , probando por ello el teorema.

También aquí hemos probado algo más. Como $[W : F] \leq mn$ todo elemento en W satisface un polinomio de grado cuando más mn sobre F , de donde el

COROLARIO. Si a y b de K son algebraicos sobre F de grados m y n respectivamente, entonces $a \pm b$, ab y a/b (si $b \neq 0$) son algebraicos sobre F y de grado cuando más mn .

En la prueba del último teorema hicimos dos extensiones del campo F . A la primera, le llamamos T ; era simplemente el campo $F(a)$. A la segunda, le llamábamos W y era $T(b)$. Así pues, $W = (F(a))(b)$ y es costumbre representarlo por $F(a, b)$. Análogamente, podríamos hablar de $F(b, a)$; no es difícil probar que $F(a, b) = F(b, a)$. Siguiendo este patrón podemos definir $F(a_1, a_2, \dots, a_n)$ para elementos a_1, \dots, a_n de K .

DEFINICIÓN. La extensión K de F se llama *extensión algebraica* de F si todo elemento de K es algebraico sobre F .

Probamos un resultado más a lo largo de los caminos que los teoremas hasta ahora probados nos han marcado.

TEOREMA 5.E. Si L es una extensión algebraica de K y si K es una extensión algebraica de F , entonces L es una extensión algebraica de F .

Prueba. Sea u un elemento arbitrario cualquiera de L ; nuestro objetivo es mostrar que u satisface algún polinomio no trivial con coeficientes en F . ¿Qué información tenemos hasta el presente? Ciertamente, sabemos que u satisface algún polinomio $x^n + \sigma_1 x^{n-1} + \dots + \sigma_n$ donde $\sigma_1, \dots, \sigma_n$ están en K . Pero K es algebraico sobre F ; por tanto, por varios usos del teorema 5.c, $M = F(\sigma_1, \dots, \sigma_n)$ es una extensión finita de F . Como u satisface el polinomio $x^n + \sigma_1 x^{n-1} + \dots + \sigma_n$ cuyos coeficientes están en M , u es algebraico sobre M . En virtud de lo dicho en el teorema 5.b de ello resulta que $M(u)$ es una extensión finita de M . Sin embargo, por el teorema 5.a, $[M(u) : F] = [M(u) : M][M : F]$, de donde $M(u)$ es una extensión finita de F . Pero esto implica que u es algebraico sobre F , lo que completa la prueba del teorema.

Una rápida descripción del teorema 5.e: algebraica sobre algebraica es algebraica.

Los resultados precedentes son de interés especial en el caso particular en que F es el campo de los números racionales y K el campo de los números complejos.

DEFINICIÓN. Un número complejo se dice que es un *número algebraico* si es algebraico sobre el campo de los números racionales.

Un número complejo que no es algebraico se llama *transcendente*. En la actual etapa no tenemos motivo alguno para suponernos que hay números transcendentales. En la próxima sección probaremos que el familiar número real e es transcendente. Esto, desde luego, establecerá la existencia de los números transcendentales. En realidad, existen en gran abundancia; en un sentido muy bien definido hay más números transcendentales que algebraicos.

El teorema 5.d aplicado a los números algebraicos prueba el hecho interesante de que *los números algebraicos forman un campo*; es decir, que la suma, producto y cociente de números algebraicos son también números algebraicos.

El teorema 5.e cuando se usa junto con el llamado “teorema fundamental del álgebra” tiene como implicación que las raíces de un polinomio cuyos coeficientes son números algebraicos son ellas misma números algebraicos.

Problemas

1. Pruébese que la aplicación $\psi : F[x] \rightarrow F(a)$ definida por $h(x)\psi = h(a)$ es un homomorfismo.
2. Sea F un campo y sea $F[x]$ el anillo de los polinomios en x sobre F . Sea $g(x)$ de grado n en $F[x]$ y sea $V = (g(x))$ el ideal generado por $g(x)$ en $F[x]$. Pruébese que $F[x]/V$ es un espacio vectorial de dimensión n sobre F .
3. a) Si V es un espacio vectorial de dimensión finita sobre el campo K , y si F es un subcampo de K tal que $[K : F]$ es finito, demuéstrese que V es un espacio vectorial de dimensión finita sobre F y que, además, $\dim_F(V) = (\dim_K(V))([K : F])$.
 b) Pruébese que el teorema 5.a es un caso particular del resultado de la parte (a).
4. a) Sea R el campo de los números reales y Q el campo de los números racionales. En R , $\sqrt{2}$ y $\sqrt{3}$ son ambos números algebraicos sobre Q . Exhíbase un polinomio de grado 4 sobre Q satisfecho por $\sqrt{2} + \sqrt{3}$.
 b) ¿Cuál es el grado de $\sqrt{2} + \sqrt{3}$ sobre Q ? Pruébese la contestación.
 c) ¿Cuál es el grado de $\sqrt{2}\sqrt{3}$ sobre Q ?
5. Con la misma notación que en el problema 4, demuéstrese que $\sqrt{2} + \sqrt[3]{5}$ es algebraica sobre Q de grado 6.

- *6. a) Encuéntrese un elemento $u \in R$ tal que $Q(\sqrt{2}, \sqrt[3]{5}) = Q(u)$.
 b) En $Q(\sqrt{2}, \sqrt[3]{5})$ caratterízense todos los elementos w tales que $Q(w) \neq Q(\sqrt{2}, \sqrt[3]{5})$.
7. a) Pruébese que $F(a, b) = F(b, a)$.
 b) Si (i_1, i_2, \dots, i_n) es una permutación cualquiera de $(1, 2, \dots, n)$, pruébese que
- $$F(a_1, \dots, a_n) = F(a_{i_1}, a_{i_2}, \dots, a_{i_n}).$$

8. Si $a, b \in K$ son algebraicos sobre F de grados m y n , respectivamente, y si m y n son primos relativos, pruébese que $F(a, b)$ es de grado mn sobre F .

9. Supongamos que F es un campo que tiene un número finito de elementos, q .

- a) Pruébese que hay un número primo p tal que $\underbrace{a+a+\dots+a}_{p \text{ veces}} = 0$ para todo $a \in F$.
 b) Pruébese que $q = p^n$ para algún entero n .
 c) Si $a \in F$ pruébese que $a^q = a$.
 d) Si $b \in K$ es algebraico sobre F pruébese que $b^{q^m} = b$ para algún $m > 0$.

Un número algebraico a se dice que es un *entero algebraico* si satisface una ecuación de la forma $a^m + \alpha_1 a^{m-1} + \dots + \alpha_m = 0$, donde $\alpha_1, \dots, \alpha_m$ son enteros.

10. Si a es número algebraico cualquiera, pruébese que hay un entero positivo n tal que na es un entero algebraico.

11. Si el número racional r es también un entero algebraico, pruébese que r debe ser un entero ordinario.

12. Si a es un entero algebraico y m es un entero ordinario, pruébese que:
 a) $a+m$ es un entero algebraico.
 b) ma es un entero algebraico.

13. Si α es un entero algebraico que satisface $\alpha^3 + \alpha + 1 = 0$ y β es un entero algebraico que satisface $\beta^2 + \beta - 3 = 0$, pruébese que $\alpha + \beta$ y $\alpha\beta$ son enteros algebraicos.

- **14. a) Pruébese que la suma de dos enteros algebraicos es un entero algebraico.
 b) Pruébese que el producto de dos enteros algebraicos es un entero algebraico.
15. a) Pruébese que $\sin 1^\circ$ es un entero algebraico.
 b) Basándose en la parte (a), pruébese que $\sin m^\circ$ es un entero algebraico para cualquier entero m .

2. LA TRANSCENDENCIA DE e

En la definición de números algebraicos y transcendentales señalamos que podía demostrarse que existen números transcendentales. Una forma de conseguir tal prueba de existencia sería la de demostrar que algún determinado número es transcendente.

En 1851, Liouville dio un criterio para la algebraicidad de un número complejo; usándolo, fue capaz de enumerar una larga colección de números transcendentales. Por ejemplo, de su trabajo se desprende que el número $0.10100100000100\dots 10\dots$ es transcendente; aquí el número de ceros entre dos unos sucesivos va siendo $1!, 2!, \dots, n!, \dots$.

Esto, ciertamente, resuelve el problema de existencia. Sin embargo, el problema de si algunos números familiares dados eran transcendentales aún persistía. El primer éxito en esta dirección fue de Hermite, quien en 1873 dio una prueba de que e es transcendente. Hilbert simplificó mucho la prueba de Hermite. La prueba que aquí daremos es una variación, debida a Hurwitz, de la prueba de Hilbert.

El número π ofrecía mayores dificultades. Fueron finalmente vencidas por Lindemann, quien, en 1882, presentó una prueba de que π es transcendental. Una consecuencia inmediata de esto es el hecho de que es imposible cuadrar el círculo con regla y compás, pues tal construcción nos llevaría a un número algebraico θ tal que $\theta^2 = \pi$. Pero si θ es algebraico, también lo es θ^2 , en virtud de lo cual π sería algebraico en contradicción con el resultado de Lindemann.

En 1934, trabajando independientemente, Gelfond y Schneider probaron que si a y b son números algebraicos y b es irracional, entonces a^b es transcendente. Esto contestaba afirmativamente el problema planteado por Hilbert sobre si $2^{\sqrt{2}}$ era transcendente.

Para aquellos interesados en proseguir con este tema de los números transcendentales, recomendamos elogiosamente los encantadores libros, uno C. L. Siegel, titulado *Transcendental Numbers* (Princeton University Press), y otro, de I. Niven, titulado *Irrational Numbers* (Carus Monographs).

Probar que e es irracional es fácil; probar que π es irracional es mucho más difícil. Para una prueba sutil y elegante de esto último véase el artículo de Niven titulado "Una prueba sencilla de que π es irracional", *Bulletin of the American Mathematical Society*, vol. 53 (1947), pág. 509.

Discutamos ahora la transcendencia de e . Aparte de su interés intrínseco, su prueba nos ofrece un cambio de paso. Hasta este momento todos nuestros argumentos han sido de naturaleza algebraica; ahora, durante un momento, volvemos a los terrenos, más familiares, del cálculo. La prueba misma usará solamente cálculo elemental; el resultado necesario más profundo será el teorema del valor medio.

TEOREMA 5.F. *El número e es transcendente.*

Prueba. En la prueba usaremos la notación estándar $f^{(i)}(x)$ para denotar la i -ésima derivada de $f(x)$ respecto a x .

Supongamos que $f(x)$ es un polinomio de grado r con coeficientes reales. Sea $F(x) = f(x) + f^{(1)}(x) + f^{(2)}(x) + \dots + f^{(r)}(x)$. Calculamos $(d/dx)(e^{-x}F(x))$; usando el hecho de que $f^{(r+1)}(x) = 0$ (pues $f(x)$ es de grado r) y la propiedad básica de e , es decir, que $(d/dx)e^x = e^x$, obtenemos $(d/dx)(e^{-x}F(x)) = -e^{-x}f(x)$.

El teorema del valor medio afirma que si $g(x)$ es continuamente diferenciable sobre el intervalo cerrado $[x_1, x_2]$ entonces

$$\frac{g(x_1) - g(x_2)}{x_1 - x_2} = g^{(1)}(x_1 + \theta(x_2 - x_1)), \quad \text{donde } 0 < \theta < 1.$$

Aplicamos esto a nuestra función $e^{-x}F(x)$ que, ciertamente, satisface todas las condiciones requeridas para que se cumpla el teorema del valor medio sobre el intervalo cerrado $[x_1, x_2]$ donde $x_1 = 0$ y $x_2 = k$, donde k es un entero positivo cualquiera. Obtenemos, entonces, que $e^{-k}F(k) - F(0) = -e^{-\theta_k k}f(\theta_k k)k$, donde θ_k depende de k y es un número real entre 0 y 1. Multiplicando esta relación por e^k obtenemos $F(k) - F(0)e^k = -e^{(1-\theta_k)k}f(\theta_k k)$. Lo escribimos explícitamente para distintos valores de h :

$$(1) \quad \begin{aligned} F(1) - eF(0) &= -e^{(1-\theta_1)}f(\theta_1) = \epsilon_1 \\ F(2) - e^2F(0) &= -2e^{2(1-\theta_2)}f(2\theta_2) = \epsilon_2 \\ &\vdots \\ F(n) - e^nF(0) &= -ne^{n(1-\theta_n)}f(n\theta_n) = \epsilon_n. \end{aligned}$$

Supongamos ahora que e fuera un número algebraico; entonces satisfaría alguna relación de la forma

$$(2) \quad c_n e^n + c_{n-1} e^{n-1} + \dots + c_1 e + c_0 = 0,$$

donde c_0, c_1, \dots, c_n son enteros y $c_0 > 0$.

En las relaciones (1), multipliquemos la primera ecuación por c_1 , la segunda por c_2 y así sucesivamente; sumando estas obtenemos $c_1 F(1) + c_2 F(2) + \dots + c_n F(n) - F(0)(c_1 e + c_2 e^2 + \dots + c_n e^n) = c_1 \epsilon_1 + c_2 \epsilon_2 + \dots + c_n \epsilon_n$.

En vista de la relación (2), $c_1 e + c_2 e^2 + \dots + c_n e^n = -c_0$, de donde la anterior ecuación se simplifica hasta tomar la forma

$$(3) \quad c_0 F(0) + c_1 F(1) + \dots + c_n F(n) = c_1 \epsilon_1 + \dots + c_n \epsilon_n.$$

Toda esta discusión tiene validez para la $F(x)$ construida a partir de un polinomio arbitrario $f(x)$. Vemos ahora lo que todo esto implica para un polinomio muy particular, un polinomio que fue Hermite el primero que usó, a saber,

$$f(x) = \frac{1}{(p-1)!} x^{p-1} (1-x)^p (2-x)^p \cdots (n-x)^p.$$

Aquí, p puede ser cualquier número primo escogido de forma que $p > n$ y $p > c_0$. Partiendo de este polinomio examinaremos atentamente $F(0)$, $F(1)$, ..., $F(n)$ y haremos una estimación de la magnitud de $\epsilon_1, \epsilon_2, \dots, \epsilon_n$.

Cuando se desarrolla, $f(x)$ es un polinomio de la forma

$$\frac{(n!)^p}{(p-1)!}x^{p-1} + \frac{a_0 x^p}{(p-1)!} + \frac{a_1 x^{p+1}}{(p-1)!} + \dots,$$

donde a_0, a_1, \dots , son enteros.

Cuando $i \geq p$ afirmamos que $f^{(i)}(x)$ es un polinomio con coeficientes que son enteros, todos los cuales son múltiplos de p . (Pruébese. Véase el problema 2.) Así pues, para cualquier entero j , $f^{(i)}(j)$, para $i \geq p$ es un entero y es un múltiplo de p .

Ahora bien, según su misma definición, $f(x)$ tiene una raíz de multiplicidad p en $x = 1, 2, \dots, n$. Luego para $j = 1, 2, \dots, n$, $f(j) = 0, \dots, f^{(p-1)}(j) = 0$. Pero $F(j) = f(j) + f^{(1)}(j) + \dots + f^{(p-1)}(j) + f^{(p)}(j) + \dots + f^{(r)}(j)$; por la anterior discusión para $j = 1, 2, \dots, n$, $F(j)$ es un entero y es un múltiplo de p .

¿Y qué hay acerca de $F(0)$? Como $f(x)$ tiene una raíz de multiplicidad $p-1$ en $x = 0$, $f(0) = f^{(1)}(0) = \dots = f^{(p-2)}(0) = 0$. Para $i \geq p$, $f^{(i)}(0)$ es un entero que a la vez es un múltiplo de p . Pero $f^{(p-1)}(0) = (n!)^p$ y como $p > n$ y es un número primo, $p \nmid (n!)^p$ de forma que $f^{(p-1)}(0)$ es un entero no divisible por p . $F(0) = f(0) + f^{(1)}(0) + \dots + f^{(p-2)}(0) + f^{(p-1)}(0) + f^{(p)}(0) + \dots + f^{(r)}(0)$, concluimos que $F(0)$ es un entero no divisible por p . Como $c_0 > 0$ y $p > c_0$, y como $p \nmid F(0)$ mientras que $p \mid F(1), p \mid F(2), \dots, p \mid F(n)$, podemos asegurar que $c_0 F(0) + c_1 F(1) + \dots + c_n F(n)$ es un entero y no es divisible por p .

Pero de acuerdo con (3), $c_0 F(0) + c_1 F(1) + \dots + c_n F(n) = c_1 \epsilon_1 + \dots + c_n \epsilon_n$. ¿Qué puede decirse acerca de ϵ_i ? Recordemos que

$$\epsilon_i = \frac{-e^{i(1-\theta_i)}(1-i\theta_i)^p \cdots (n-i\theta_i)^p (i\theta_i)^{p-1} i}{(p-1)!},$$

donde $0 < \theta_i < 1$. Así pues

$$|\epsilon_i| \leq e^n \frac{n^p (n!)^p}{(p-1)!}.$$

Cuando $p \rightarrow \infty$

$$\frac{e^n n^p (n!)^p}{(p-1)!} \rightarrow 0,$$

(Pruébese.) De donde resulta que podemos encontrar un número primo mayor que ambos c_0 y n y suficientemente grande para que $|c_1 \epsilon_1 + \dots + c_n \epsilon_n| < 1$. Pero $c_1 \epsilon_1 + \dots + c_n \epsilon_n = c_0 F(0) + \dots + c_n F(n)$, luego debe ser un entero; como es más pequeño que 1 nuestra sola posible conclusión es que

$c_1\epsilon_1 + \dots + c_n\epsilon_n = 0$. En consecuencia, $c_0F(0) + \dots + c_nF(n) = 0$; pero esto carece en absoluto de sentido, ya que sabemos que $p \nmid (c_0F(0) + \dots + c_nF(n))$, mientras que $p \mid 0$. Esta contradicción, que nace de la hipótesis de que e es algebraico, prueba que e debe ser trascendente.

Problemas

- Usando para e la serie infinita, $e = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{1}{m!} + \dots$,

pruébese que e es irracional.

- Si $g(x)$ es un polinomio con coeficientes enteros, pruébese que si p es un número primo entonces para $i \geq p$,

$$\frac{d^i}{dx^i} \left(\frac{g(x)}{(p-1)!} \right)$$

es un polinomio con coeficientes enteros cada uno de los cuales es divisible por p .

- Si a es un número real cualquiera, pruébese que $(a^m/m!) \rightarrow 0$ cuando $m \rightarrow \infty$.

- Si $m > 0$ y n son enteros, pruébese que $e^{m/n}$ es trascendente.

3. RAÍCES DE POLINOMIOS

En la sección 1 discutimos elementos de una extensión dada K de F que eran algebraicos sobre F , es decir, elementos que satisfacían polinomios en $F[x]$. Volvemos ahora el problema al revés; dado un polinomio $p(x)$ en $F[x]$, queremos encontrar un campo K que sea una extensión de F en el que $p(x)$ tenga una raíz. Ya no tenemos el campo K a nuestra disposición; en realidad, nuestro primer objetivo es construirlo. Una vez que lo hayamos construido lo examinaremos detenidamente y veremos qué consecuencias podemos derivar de tal examen.

DEFINICIÓN. Si $p(x) \in F[x]$ entonces un elemento a que se encuentra en algún campo extensión del F se llama *raíz* de $p(x)$ si $p(a) = 0$.

Comenzamos con el familiar resultado conocido como el *teorema del residuo*.

LEMA 5.1. Si $p(x) \in F[x]$ y si K es una extensión de F , entonces para cualquier elemento $b \in K$, $p(x) = (x-b)q(x) + p(b)$, donde $q(x) \in K[x]$ y donde $\deg q(x) = \deg p(x) - 1$.

Prueba. Como $F \subset K$, $F[x]$ está contenido en $K[x]$, de donde podemos considerar $p(x)$ como se encontrara en $K[x]$. Por el algoritmo de la división para polinomios en $K[x]$, $p(x) = (x - b)q(x) + r$ donde $q(x) \in K[x]$ y donde $r = 0$ o $\deg r < \deg(x - b) = 1$. Así pues, o $r = 0$ o $\deg r = 0$; en cualquiera de los casos r debe ser un elemento de K . ¿Pero qué elemento de K es exactamente? Como $p(x) = (x - b)q(x) + r$, $p(b) = (b - b)q(b) + r = r$. Por tanto, $p(x) = (x - b)q(x) + p(b)$. Que el grado de $q(x)$ es menor en una unidad que el de $p(x)$ es fácil de verificar y se deja para el lector.

COROLARIO. Si $a \in K$ es una raíz de $p(x) \in F[x]$, donde $F \subset K$, entonces en $K[x]$, $(x - a) | p(x)$.

Prueba. De acuerdo con el lema 5.1, en $K[x]$, $p(x) = (x - a)q(x) + p(a) = (x - a)q(x)$, ya que $p(a) = 0$. Por tanto, $(x - a) | p(x)$ en $K[x]$.

DEFINICIÓN. El elemento $a \in K$ es una raíz de $p(x) \in F[x]$ de *multiplicidad m* si $(x - a)^m | p(x)$, mientras que $(x - a)^{m+1} \nmid p(x)$.

Una pregunta razonable es la siguiente: ¿cuántas raíces puede tener un polinomio en un campo dado? Antes de contestar debemos ponernos de acuerdo en cómo debemos contar una raíz de multiplicidad m . *Siempre la contaremos como m raíces*. Aceptada esta convención, podemos probar el

LEMA 5.2. Un polinomio de grado n sobre un campo tiene cuando más n raíces en cualquier campo extensión.

Prueba. Procedemos por inducción sobre n , el grado del polinomio $p(x)$. Si $p(x)$ es de grado 1, entonces debe ser de la forma $\alpha x + \beta$, donde α, β están en un campo F y donde $\alpha \neq 0$. Cualquier a tal que $p(a) = 0$ debe entonces implicar que $\alpha a + \beta = 0$ de donde se concluye que $a = -\beta/\alpha$. Es decir, $p(x)$ tiene la raíz única $-\beta/\alpha$, de donde la conclusión del lema es, ciertamente, válida en este caso.

Suponiendo que el resultado sea cierto en cualquier campo para todos los polinomios de grado menor que n , supongamos que $p(x)$ es de grado n sobre F . Sea K una extensión cualquiera de F . Si $p(x)$ no tiene ninguna raíz en K entonces lo afirmado es cierto, pues el número de raíces en K , a saber, cero, es definitivamente cuando más n . Supongamos, pues, que $p(x)$ tiene al menos una raíz $a \in K$ y que a es una raíz de multiplicidad m . Como $(x - a)^m | p(x)$, de ello se sigue que $m \leq n$. Ahora $p(x) = (x - a)^m q(x)$ donde $q(x) \in K[x]$ es de grado $n - m$. Del hecho de que $(x - a)^{m+1} \nmid p(x)$ deducimos que $(x - a) \nmid q(x)$, de donde, según el corolario al lema 5.1, a no es una raíz de $q(x)$. Si $b \neq a$ es una raíz, en K , de $p(x)$ entonces $0 = p(b) = (b - a)^m q(b)$; pero, como $b - a \neq 0$ y como estamos en un campo, concluimos que $q(b) = 0$. Es decir, cualquier raíz de $p(x)$ en K distinta de la a debe

ser una raíz de $q(x)$. Como $q(x)$ es de grado $n-m < n$, $q(x)$ tiene, de acuerdo con nuestra hipótesis de inducción, cuando más $n-m$ raíces en K , que, junto con la otra raíz a , contada m veces, nos dice que $p(x)$ tiene cuando más $m+(n-m) = n$ raíces en K . Esto completa la inducción y prueba el lema.

Se debe subrayar que la commutatividad es esencial en el lema 5.2. Si consideramos el anillo de cuaternios reales, al que solo le falta ser commutativo para ser un campo, entonces el polinomio $x^2 + 1$ tiene al menos tres raíces, i , j y k (en realidad tiene un número infinito de raíces). En una dirección un poco diferente necesitamos, aunque el anillo sea commutativo, que sea un dominio entero además, pues si $ab = 0$ con $a \neq 0$ y $b \neq 0$ en el anillo commutativo R , entonces el polinomio ax de grado 1 sobre R tiene al menos dos distintas raíces $x = 0$ y $x = b$ en R .

Los dos lemas anteriores, aunque interesantes, son de interés secundario. Seguimos ahora con nuestra primitiva tarea, la de hacernos de extensiones adecuadas de F en las que un polinomio dado tenga raíces. Una vez que hayamos hecho esto, seremos capaces de analizar tales extensiones con el suficiente grado de precisión para obtener resultados. El paso más importante en la construcción lo damos en el próximo teorema. El argumento usado nos recordará algunos de los usados en la sección 1.

TEOREMA 5.G. *Si $p(x)$ es un polinomio en $F[x]$ de grado $n \geq 1$ y es irreducible sobre F , entonces hay una extensión E de F tal que $[E:F] = n$ en que $p(x)$ tiene una raíz.*

Prueba. Sea $F[x]$ el anillo de polinomios en x sobre F y sea $V = (p(x))$ el ideal de $F[x]$ generado por $p(x)$. Según el lema 3.22 V es un ideal máximo de $F[x]$, de donde, de acuerdo con el teorema 3.b, $E = F[x]/V$ es un campo. Mostraremos que esta E satisface las conclusiones del teorema.

Necesitamos demostrar primero que E es una extensión de F ; ¡aunque en realidad no lo es! Pero sea F la imagen de F en E , es decir, sea $F = \{\alpha + V | \alpha \in F\}$. Afirmamos que F es un campo isomorfo a F ; en realidad, si ψ es la aplicación de $F[x]$ en $\frac{F[x]}{V} = E$ definida por $f(x)\psi = f(x) + V$, entonces la restricción a F induce un isomorfismo de F sobre F (*¡pruébese!*). Usando este isomorfismo, identificamos F y F ; *de este modo podemos considerar a E como una extensión de F* .

Afirmamos que E es una extensión finita de F de grado $n = \deg p(x)$, pues los elementos $1 + V, x + V, (x + V)^2 = x^2 + V, \dots, (x + V)^n = x^n + V, \dots, (x + V)^{n-1} = x^{n-1} + V$ forman una base de E sobre F (*¡pruébese!*). Por conveniencia de notación denotemos al elemento $x\psi = x + V$ en el campo E por a . Dado $f(x) \in F[x]$, ¿qué es $f(x)\psi$? Afirmamos que es simplemente $f(a)$, pues como ψ es un homomorfismo, si $f(x) = \beta_0 + \beta_1 x + \dots + \beta_k x^k$, entonces

$f(x)\psi = \beta_0\psi + (\beta_1\psi)(x\psi) + \dots + (\beta_k\psi)(x\psi)^k$, y usando la identificación antes indicada de $\beta\psi$ con β , vemos que $f(x)\psi = f(a)$. En particular, como $p(x) \in V$, $p(x)\psi = 0$; pero $p(x)\psi = p(a)$. Luego el elemento $a = x\psi$ en E es una raíz de $p(x)$. El campo E se ha demostrado que satisface todas las propiedades requeridas en la conclusión del teorema 5.g y, por tanto, este teorema ha quedado probado.

Una consecuencia inmediata de este teorema es el

COROLARIO. Si $f(x) \in F[x]$, entonces hay una extensión finita E de F en que $f(x)$ tiene una raíz. Además, $[E : F] \leq \deg f(x)$.

Prueba. Sea $p(x)$ un factor irreducible de $f(x)$; cualquier raíz de $p(x)$ es una raíz de $f(x)$. De acuerdo con el teorema, hay una extensión E de F con $[E : F] = \deg p(x) \leq \deg f(x)$ en que $p(x)$ y, por tanto, $f(x)$, tiene una raíz.

Aunque, en realidad, se trata de un corolario al anterior corolario, el siguiente teorema es de tanta importancia que lo singularizamos como tal, como teorema.

TEOREMA 5.h. Sea $f(x) \in F[x]$ de grado $n \geq 1$. Entonces hay una extensión E de F de grado más $n!$ en que $f(x)$ tiene n raíces (y, por tanto, un juego completo de raíces).

Prueba. En el enunciado del teorema una raíz de multiplicidad m se cuenta, desde luego, como m raíces.

Según el anterior corolario hay una extensión E_0 de F con $[E_0 : F] \leq n$ en que $f(x)$ tiene una raíz α . Así pues, en $E_0[x]$, $f(x)$ se factoriza como $f(x) = (x - \alpha)q(x)$ donde $q(x)$ es de grado $n - 1$. Usando inducción (o continuando el anterior proceso) hay una extensión E de E_0 de grado cuando más $(n - 1)!$ en que $q(x)$ tiene $n - 1$ raíces. Como cualquier raíz de $f(x)$ es o α o una raíz de $q(x)$, obtenemos en E todas las n raíces de $f(x)$. Y tenemos, además, $[E : F] = [E : E_0][E_0 : F] \leq (n - 1)!n = n!$. Con lo que todas las partes del teorema han quedado probadas.

El teorema 5.h afirma la existencia de una extensión finita E en la que el polinomio dado $f(x)$, de grado n sobre F , tiene n raíces. Si $f(x) = a_0x^n + a_1x^{n-1} + \dots + a_n$, $a_0 \neq 0$ y si las n raíces en E son $\alpha_1, \dots, \alpha_n$, haciendo uso del corolario al lema 5.1, $f(x)$ puede factorizarse sobre E como $f(x) = a_0(x - \alpha_1)(x - \alpha_2) \dots (x - \alpha_n)$. Así pues, $f(x)$ se descompone completamente sobre E como un producto de factores *lineales* (de primer grado). Como una extensión finita de F existe con esta propiedad, existe también una extensión finita de F de grado mínimo que también disfruta de esta propiedad de descomposición de $f(x)$ como un producto de factores lineales.

Para una tal extensión mínima, ningún subcampo propio tiene la propiedad de que $f(x)$ se factorice sobre él en un producto de factores lineales. Esto nos lleva a la siguiente

DEFINICIÓN. Si $f(x) \in F[x]$, una extensión finita E de F se dice que es un *campo de descomposición* de $f(x)$ sobre F si $f(x)$ puede ser descompuesto en un producto de factores lineales sobre E (es decir, en $E[x]$), pero no en ningún subcampo propio de E .

Repetimos: el teorema 5.h nos garantiza la existencia de campos de descomposición. En realidad nos dice aún más, pues nos asegura que dado un polinomio de grado n sobre F hay un campo de descomposición de este polinomio que es una extensión de F de grado cuando más $n!$ sobre F . Veremos más adelante que esta cota superior $n!$ es realmente alcanzada; es decir, dado n , podemos encontrar un campo F y un polinomio de grado n en $F[x]$ tal que el campo de descomposición de $f(x)$ sobre F tenga grado $n!$.

Equivalente a la definición que dimos de campo de descomposición de $f(x)$ sobre F es la proposición: E es un campo de descomposición de $f(x)$ sobre F si E es una extensión mínima de F en la que $f(x)$ tiene n raíces, donde $n = \deg f(x)$.

Surge de inmediato una pregunta: dados dos campos de descomposición E_1 y E_2 del mismo polinomio $f(x)$ en $F[x]$, ¿cuál es la relación de uno con otro? A primera vista no tenemos derecho alguno a suponer que exista entre ellos relación alguna. Nuestro siguiente objetivo es demostrar que ciertamente están íntimamente relacionados; en realidad, que son isomorfos con un isomorfismo que deja fijos todos los elementos de F . Es en tal dirección en la que ahora nos moveremos.

Sean F y F' dos campos y sea τ un isomorfismo de F sobre F' . Por conveniencia, denotemos la imagen de cualquier $\alpha \in F$ bajo τ por α' ; es decir, $\alpha\tau = \alpha'$. Mantendremos esta notación durante unas páginas que siguen.

¿Podemos usar τ para definir un isomorfismo entre $F[x]$ y $F'[t]$, los respectivos anillos de polinomios sobre F y F' ? ¿Por qué no probar lo obvio? Para un polinomio arbitrario $f(x) = \alpha_0 x^n + \alpha_1 x^{n-1} + \dots + \alpha_n \in F[x]$ definimos τ^* por $f(x)\tau^* = (\alpha_0 x^n + \alpha_1 x^{n-1} + \dots + \alpha_n)\tau^* = \alpha'_0 t^n + \alpha'_1 t^{n-1} + \dots + \alpha'_n$.

Es un problema sin complicaciones que dejamos al lector, el verificar el siguiente

LEMA 5.3. τ^* define un isomorfismo de $F[x]$ sobre $F'[t]$ con la propiedad de que $\alpha\tau^* = \alpha'$ para todo $\alpha \in F$.

Si $f(x)$ está en $F[x]$ escribiremos $f(x)\tau^*$ como $f'(t)$. El lema 5.3 implica inmediatamente que las factorizaciones de $f(x)$ en $F[x]$ van a dar a factorizaciones análogas de $f'(t)$ en $F'[t]$, y viceversa. En particular, $f(x)$ es irreducible en $F[x]$ si y solo si $f'(t)$ es irreducible en $F'[t]$.

Pero por el momento no estamos particularmente interesados en los anillos de polinomios, sino más bien en las extensiones de F . Recordemos que en la prueba del teorema 5.b empleamos anillos cociente de anillos de polinomios para obtener extensiones adecuadas de F . En consecuencia, debería ser natural para nosotros estudiar la relación entre $F[x]/(f(x))$ y $F'[t]/(f'(t))$, donde $(f(x))$ denota el ideal generado por $f(x)$ en $F[x]$ y $(f'(t))$ el generado por $f'(t)$ en $F'[t]$. El próximo lema, que es importante para esta cuestión, es realmente parte de un resultado más general sobre la teoría de anillos, pero nos contentaremos con él como resulta al aplicarse en nuestro muy particular caso.

LEMA 5.4. *Hay un isomorfismo τ^{**} de $F[x]/(f(x))$ sobre $F'[t]/(f'(t))$ con la propiedad de que para todo $\alpha \in F$, $\alpha\tau^{**} = \alpha'$.*

Prueba. Antes de comenzar con la prueba misma, debemos aclarar cómo debe entenderse la última parte del enunciado del lema. Como hemos hecho ya antes varias veces, podemos considerar a F como inmerso en $F[x]/(f(x))$ mediante la identificación de cada elemento $\alpha \in F$ con la clase lateral $\alpha + (f(x))$ de $F[x]/(f(x))$. Análogamente, podemos considerar a F' contenido en $F'[t]/(f'(t))$. El isomorfismo τ^{**} se supone entonces que satisface $[\alpha + (f(x))] \tau^{**} = \alpha' + (f'(t))$.

Buscamos un isomorfismo τ^{**} de $F[x]/(f(x))$ sobre $F'[t]/(f'(t))$. ¿Qué puede ser más simple o más natural que probar la τ^{**} definida por $[g(x) + (f(x))] \tau^{**} = g'(t) + (f'(t))$ para todo $g(x) \in F[x]$? Dejamos como un ejercicio llenar los diferentes detalles para probar que la τ^{**} así definida está bien definida y que es un isomorfismo de $F[x]/(f(x))$ sobre $F'[t]/(f'(t))$ con las propiedades necesarias para verificar el enunciado del lema 5.4.

Para nuestro propósito —que es el de probar la unicidad de los campos de descomposición — el lema 5.4 nos da la línea de ataque, pues podemos probar ahora que

TEOREMA 5.1. *Si $p(x)$ es irreducible en $F[x]$ y si v es una raíz de $p(x)$, entonces $F(v)$ es isomorfo a $F'(w)$ donde w es una raíz de $p'(t)$; además, este isomorfismo σ puede escogerse de modo que*

$$1) \quad v\sigma = w$$

$$2) \quad \alpha\sigma = \alpha' \text{ para todo } \alpha \in F.$$

Prueba. Sea v una raíz del polinomio irreducible $p(x)$ perteneciente (la raíz) a alguna extensión K de F . Sea $M = \{f(x) \in F[x] \mid f(v) = 0\}$. M es trivialmente un ideal de $F[x]$, y $M \neq F[x]$. Como $p(x) \in M$ y es un polinomio irreducible, tenemos que $M = (p(x))$. Como en la prueba del teorema 5.b transfórmese $F[x]$ en $F(v) \subset K$ con la aplicación ψ definida por $q(x)\psi = q(v)$ para todo $q(x) \in F[x]$. Vimos antes (en la prueba del teorema 5.b) que ψ

transforma $F[x]$ sobre $F(v)$. El núcleo de ψ es precisamente M , luego debe ser $(p(x))$. Según el teorema fundamental del homomorfismo de anillos hay un isomorfismo ψ^* de $F[x]/(p(x))$ sobre $F(v)$. Nótese, además, que $\alpha\psi^* = \alpha$ para todo $\alpha \in F$. Resumiendo: ψ^* es un isomorfismo de $F[x]/(p(x))$ sobre $F(v)$ que deja todos los elementos de F fijos y con la propiedad de que $v = [x + (p(x))] \psi^*$.

Como $p(x)$ es irreducible en $F[x]$, $p'(t)$ es irreducible en $F'[t]$ según el lema 5.3) y, por tanto, hay un isomorfismo θ^* de $F'[t]/(p'(t))$ sobre $F'(w)$ donde w es una raíz de $p'(t)$ tal que θ^* deja fijos todos los elementos de F' y tal que $[t + (p'(t))] \theta^* = w$.

Unimos ahora todas las piezas para probar el teorema 5.i. De acuerdo con el lema 5.4 hay un isomorfismo τ^{**} de $F[x]/(p(x))$ sobre $F'[t]/(p'(t))$ que coincide con τ sobre F y que lleva $x + (p(x))$ sobre $t + (p'(t))$. Consideremos la aplicación $\sigma = (\psi^*)^{-1} \tau^{**} \theta^*$ (motivada por

$$F(v) \xrightarrow{(\psi^*)^{-1}} \frac{F[x]}{(p(x))} \xrightarrow{\tau^{**}} \frac{F'[t]}{(p'(t))} \xrightarrow{\theta^*} F'(w)$$

de $F(v)$ sobre $F'(w)$). Es un isomorfismo de $F(v)$ sobre $F'(w)$ ya que todas las aplicaciones ψ^* , τ^{**} , θ^* son isomorfismos suprayectivos. Además, como $v = [x + (p(x))] \psi^*$, $v\sigma = (v(\psi^*)^{-1}) \tau^{**} \theta^* = ([x + (p(x))] \tau^{**}) \theta^* = [t + (p'(t))] \theta^* = w$. Además, para $\alpha \in F$, $\alpha\sigma = (\alpha(\psi^*)^{-1}) \tau^{**} \theta^* = (\alpha\tau^{**}) \theta^* = \alpha'\theta^* = \alpha'$. Hemos mostrado que σ es un isomorfismo que satisface todos los requerimientos del isomorfismo en el enunciado del teorema. Luego el teorema 5.i ha sido probado.

Un caso particular, pero en sí de interés, es el que aparece en el siguiente corolario.

COROLARIO. *Si $p(x) \in F[x]$ es irreducible y si a, b son dos raíces de $p(x)$, entonces $F(a)$ es isomorfo a $F(b)$ con un isomorfismo que lleva a en b y que deja fijos todos los elementos de F .*

Llegamos ahora al teorema que es, como antes indicamos, la piedra angular sobre la que descansa toda la teoría de Galois. Para nosotros es el punto focal de toda esta sección.

TEOREMA 5.J. *Cualesquiera campos de descomposición E y E' de los polinomios $f(x) \in F[x]$ y $f'(t) \in F'[t]$, respectivamente, son isomorfos con un isomorfismo ϕ con la propiedad de que $\alpha\phi = \alpha'$ para todo $\alpha \in F$. (En particular, cualesquiera dos campos de descomposición del mismo polinomio sobre un campo dado F son isomorfos con un isomorfismo que deja fijos todos los elementos de F .)*

Prueba. Nos gustaría usar un argumento inductivo; para hacerlo necesitamos un indicador de tamaño valuado en los enteros que podamos hacer decrecer por medio de una u otra técnica. Usaremos como indicador el grado de algún campo de descomposición sobre el campo inicial. Tal vez parezca artificioso (en realidad, incluso puede ser artificioso), pero lo usamos, como pronto veremos, porque el teorema 5.i nos proporciona el mecanismo para disminuirlo.

Si $[E:F] = 1$, entonces $E = F$, de donde $f(x)$ se descompone en un producto de factores lineales sobre el mismo F . Segundo el lema 5.3 $f'(t)$ se descompone sobre F' en un producto de factores lineales, de donde $E' = F'$. Pero entonces $\phi = \tau$ nos proporciona el isomorfismo de E sobre E' que coincide con τ sobre F .

Supongamos que el resultado es cierto para cualquier campo F_0 y cualquier polinomio $f(x) \in F_0[x]$ con tal de que el grado de algún campo de descomposición E_0 de $f(x)$ sea menor que n sobre F_0 , es decir, con tal de que $[E_0 : F_0] < n$.

Supongamos que $[E:F] = n > 1$, donde E es un campo de descomposición de $f(x)$ sobre F . Como $n > 1$, $f(x)$ tiene un factor irreducible $p(x)$ de grado $r > 1$. Sea $p'(t)$ el correspondiente factor irreducible de $f'(t)$. Como E descompone a $f(x)$, un juego completo de raíces de $f(x)$ y, por tanto, a priori, de raíces de $p(x)$, están en E . De esta manera, hay una $v \in E$ tal que $p(v) = 0$; de acuerdo con el teorema 5.c, $[F(v):F] = r$. Análogamente, hay una $w \in E'$ tal que $p'(w) = 0$. Segundo el teorema 5.i hay un isomorfismo ϕ de $F(v)$ sobre $F'(w)$ con la propiedad de que $\alpha\sigma = \alpha'$ para toda $\alpha \in F$.

Como $[F(v):F] = r > 1$,

$$[E : F(v)] = \frac{[E : F]}{[F(v) : F]} = \frac{n}{r} < n.$$

Afirmamos que E es un campo de descomposición para $f(x)$ considerado como un polinomio sobre $F_0 = F(v)$, pues ningún subcampo de E , conteniendo F_0 y, por tanto, F , puede descomponer $f(x)$, ya que E se supone es un campo de descomposición de $f(x)$ sobre F . Análogamente, E' es un campo de descomposición para $f'(t)$ sobre $F'_0 = F'(w)$. Por nuestra hipótesis de inducción hay un isomorfismo ϕ de E sobre E' tal que $a\phi = a\sigma$ para todo $a \in F_0$. Pero para cada $\alpha \in F$, $\alpha\sigma = \alpha'$ de donde para $\alpha \in F \subset F_0$, $\alpha\phi = \alpha\sigma = \alpha'$. Esto completa la inducción y prueba el teorema.

Para ver la verdad de la parte “en particular...”, sea $F = F'$ y sea τ la aplicación idéntica $\alpha\tau = \alpha$ para todo $\alpha \in F$. Supongamos que E_1 y E_2 son dos campos de descomposición de $f(x) \in F[x]$. Considerando $E_1 = E \supset F$ y $E_2 = E' \supset F' = F$ y aplicando el teorema que acabamos de probar, tenemos que E_1 y E_2 son isomorfos con un isomorfismo que deja fijos todos los elementos de F .

En vista del hecho de que dos campos de descomposición del mismo polinomio sobre F son isomorfos y con un isomorfismo que deja fijos todos los elementos de F , podemos hablar justificadamente *del* campo de descomposición, y no de *un* campo de descomposición, pues que este es esencialmente único.

EJEMPLOS

1. Sea F un campo cualquiera y sea $p(x) = x^2 + \alpha x + \beta$, $\alpha, \beta \in F$, perteneciente a $F[x]$. Si K es una extensión cualquiera de F en que $p(x)$ tiene una raíz, a , entonces el elemento $b = -\alpha - a$ igual que en K es también una raíz de $p(x)$. Si $b = a$ es fácil comprobar que $p(x)$ debe, entonces, ser $p(x) = (x-a)^2$ y, por tanto, ambas raíces de $p(x)$ están en K . Si $b \neq a$, entonces, de nuevo, ambas raíces de $p(x)$ están en K . En consecuencia $p(x)$ puede ser descompuesto por cualquier extensión de grado 2 de F . Podíamos también haber obtenido este resultado directamente aplicando el teorema 5.h.
2. Sea F el campo de los números racionales, y sea $f(x) = x^3 - 2$. En el campo de los números complejos las tres raíces de $f(x)$ son $\sqrt[3]{2}$, $\omega \sqrt[3]{2}$, $\omega^2 \sqrt[3]{2}$, donde $\omega = (-1 + \sqrt{3}i)/2$ y donde $\sqrt[3]{2}$ es una raíz cúbica real de 2. Pero aquí $F(\sqrt[3]{2})$ no puede descomponer $x^3 - 2$, pues, como un subcampo del campo real, no puede contener el número complejo y no real $\sqrt[3]{2}$. Sin determinarlo explícitamente, ¿qué podemos decir acerca de E , el campo de descomposición de $x^3 - 2$ sobre F ? Segundo el teorema 5.h, $[E:F] \leqslant 3! = 6$; por la anterior observación, como $x^3 - 2$ es irreducible sobre F y como $[F(\sqrt[3]{2}):F] = 3$, de acuerdo con el corolario al teorema 5.a, $3 = [F(\sqrt[3]{2}):F][E:F]$. Finalmente, $[E:F] > [F(\sqrt[3]{2}):F] = 3$. La única posibilidad es que $[E:F] = 6$. Podíamos, desde luego, obtener este resultado haciendo dos extensiones $F_1 = F(\sqrt[3]{2})$ y $E = F_1(\omega)$ y mostrando que ω satisface una ecuación cuadrática irreducible sobre F_1 .
3. Sea F el campo de los números racionales y sea $f(x) = x^4 + x^2 + 1 \in F[x]$. Afirmamos que $E = F(\omega)$, donde $\omega = (-1 + \sqrt{3}i)/2$, es un campo de descomposición de $f(x)$. Luego $[E:F] = 2$, muy lejos del máximo posible $4! = 24$.

Problemas

1. En la prueba del lema 5.1, demostrar que el grado de $q(x)$ es menor en una unidad que el de $p(x)$.
2. En la prueba del teorema 5.g, pruébese con todo detalle que los elementos $1 + V, x + V, \dots, x^{n-1} + V$, forman una base de E sobre F .
3. Pruébese el lema 5.3 con todos los detalles.

4. Demuéstrese que τ^{**} en el lema 5.4 está bien definido y es un isomorfismo de $F[x]/(f(x))$ sobre $F'[t]/(f'(t))$.

5. En el ejemplo 3 al final de esta sección pruébese que $F(\omega)$ es el campo de descomposición de $x^4 + x^2 + 1$.

6. Sea F el campo de los números racionales. Determinense los grados de los campos de descomposición de los siguientes polinomios sobre F .

- a) $x^4 + 1$.
- b) $x^6 + 1$.
- c) $x^4 - 2$.
- d) $x^5 - 1$.
- e) $x^9 + x^3 + 1$.

7. Si p es un número primo, pruébese que el campo de descomposición sobre F , el campo de los números racionales, del polinomio $x^p - 1$ es de grado $p - 1$.

**8. Si $n > 1$, pruébese que el campo de descomposición de $x^n - 1$ sobre el campo de los números racionales es de grado $\Phi(n)$ donde Φ es la función Φ de Euler.

*9. Si F es el campo de los números racionales, encontrar las condiciones necesarias y suficientes sobre a y b para que el campo de descomposición de $x^3 + ax + b$ tenga exactamente 3 como grado sobre F .

10. Sea p un número primo y sea $F = J_p$, el campo de enteros módulo p .

- a) Pruébese que hay un polinomio irreducible de grado 2 sobre F .
- b) Úsese este polinomio para construir un campo con p^2 elementos.
- *c) Pruébese que dos polinomios irreducibles cualesquiera de grado 2 sobre F nos llevan a campos isomorfos con p^2 elementos.

11. Si E es una extensión de F y si $f(x) \in F[x]$ y ϕ es un automorfismo de E que deja fijos todos los elementos de F , pruébese que ϕ debe llevar una raíz de $f(x)$ que pertenezca a E en una raíz de $f(x)$ en E .

12. Pruébese que $F(\sqrt[3]{2})$, donde F es el campo de los números racionales, no tiene ningún automorfismo aparte del idéntico.

13. Usando el resultado del problema 11, pruébese que si el número complejo α es una raíz del polinomio $p(x)$ de coeficientes reales, entonces $\bar{\alpha}$, el conjugado complejo de α , es también una raíz de $p(x)$.

14. Usando el resultado del problema 11, pruébese que si m es un entero que no es un cuadrado perfecto y si $\alpha + \beta\sqrt{m}$ (α, β racionales) es la raíz de un polinomio $p(x)$ que tiene *coeficientes racionales*, entonces $\alpha - \beta\sqrt{m}$ es también una raíz de $p(x)$.

*15. Si F es el campo de los números reales, pruébese que si ϕ es un automorfismo de F , entonces ϕ deja fijos todos los elementos de F .

16. a) Encuéntrense *todas* las tétradas reales $t = a_0 + a_1i + a_2j + a_3k$ que satisfacen $t^2 = -1$.

*b) Para t , como en la parte (a), pruébese que podemos encontrar una tétrada real s tal que $sts^{-1} = i$.

4. CONSTRUCCIONES CON REGLA Y COMPÁS

Hacemos una pausa en nuestro desarrollo general para examinar algunas implicaciones de los resultados obtenidos hasta ahora en algunas situaciones geométricas familiares.

Un número real α se dice que es un *número constructible* si únicamente usando la regla y compás podemos construir un segmento rectilíneo de longitud α . Suponemos que nos han dado alguna unidad de longitud fundamental. Recuérdese que según la geometría de secundaria podemos construir, con regla y compás, una recta perpendicular y una recta paralela a una recta dada que pase por un punto dado. Basándonos en esto es un fácil ejercicio (véase el problema 1) probar que si α y β son números constructibles entonces también lo son $\alpha \pm \beta$, $\alpha\beta$, y, cuando $\beta \neq 0$, α/β . Por tanto, el conjunto de números constructibles forma un subcampo, W , del campo de los números reales.

En particular, como $1 \in W$, W debe contener a F_0 , el campo de los números racionales. Deseamos estudiar la relación de W con el campo racional.

Como tendremos muchas ocasiones de usar la frase “construir con regla y compás” (y variantes de ella) *las palabras construir, constructible, construcción, siempre significarán con regla y compás*.

Si $w \in W$ podemos llegar a w partiendo del campo racional por un número *finito* de construcciones.

Sea F un subcampo cualquiera del campo de los números reales. Consideremos todos los puntos (x, y) en el plano real euclíadiano cuyas dos coordenadas, x y y , están en F ; llamamos al conjunto de estos puntos el *plano de F* . Cualquier recta que una dos puntos en el plano de F tiene una ecuación de la forma $ax + by + c = 0$ donde a , b y c están todos en F (véase el problema 1). Además, cualquier circunferencia que tenga como centro un punto en el plano de F y como radio un elemento de F tiene una ecuación de la forma $x^2 + y^2 + ax + by + c = 0$, donde todos los a , b , c están en F (véase el problema 3). Llamamos a tales rectas y circunferencias *rectas y circunferencias en F* .

Dadas dos rectas en F que se intersecan en el plano real, entonces su punto de intersección es un punto en el plano de F (véase el problema 4). Por otra parte, la intersección de una recta en F y una circunferencia en F

no necesariamente nos va a dar un punto en el plano de F . Pero usando el hecho de que la ecuación de una recta en F es de la forma $ax+by+c=0$ y la de una circunferencia en F es de la forma $x^2+y^2+dx+ey+f=0$, donde a, b, c, d, e, f están, todas, en F , podemos demostrar que cuando una recta y una circunferencia de F se intersecan en el plano real, entonces se intersecan o en un punto en el plano de F o en el plano de $F(\sqrt{\gamma})$ para algún positivo γ de F (véase el problema 5). Finalmente, la intersección de dos circunferencias en F puede realizarse como la de una recta en F y una circunferencia en F , pues si estas dos circunferencias son $x^2+y^2+a_1x+b_1y+c_1=0$ y $x^2+y^2+a_2x+b_2y+c_2=0$, entonces su intersección es la intersección de cualquiera de ellas con la recta $(a_1-a_2)x+(b_1-b_2)y+(c_1-c_2)=0$, luego también nos da un punto o en el plano de F o en el plano de $F(\sqrt{\gamma})$ para algún positivo γ en F .

Así pues, las rectas y las circunferencias de F nos llevan a puntos o en F o en extensiones cuadráticas de F . Si ahora estamos en $F(\sqrt{\gamma_1})$ para alguna extensión cuadrática de F , entonces las rectas y las circunferencias en $F(\sqrt{\gamma_1})$ se intersecan en puntos en el plano de $F(\sqrt{\gamma_1}, \sqrt{\gamma_2})$ donde γ_2 es un número positivo en $F(\sqrt{\gamma_1})$. Un punto es constructible partiendo de F si podemos encontrar números reales $\lambda_1, \dots, \lambda_n$, tales que $\lambda_1^2 \in F, \lambda_2^2 \in F(\lambda_1), \lambda_3^2 \in F(\lambda_1, \lambda_2), \dots, \lambda_n^2 \in F(\lambda_1, \dots, \lambda_{n-1})$, tales que el punto está en el plano $F(\lambda_1, \dots, \lambda_n)$. Recíprocamente, si $\gamma \in F$ es tal que $\sqrt{\gamma}$ es real, entonces podemos realizar γ como una intersección de rectas y circunferencias en F (véase el problema 6). Así pues, un punto es constructible partiendo de F si y sólo si podemos encontrar un número finito de números reales $\lambda_1, \dots, \lambda_n$, tales que

$$1) [F(\lambda_1):F] = 1 \text{ o } 2$$

$$2) [F(\lambda_1, \dots, \lambda_i):F(\lambda_1, \dots, \lambda_{i-1})] = 1 \text{ o } 2 \text{ para } i = 1, 2, \dots, n,$$

y, además, que nuestro punto se encuentre en el plano de $F(\lambda_1, \dots, \lambda_n)$.

Hemos dicho que un número real α es constructible si empleando la regla y el compás podemos construir un segmento rectilíneo de longitud α . Pero esto, en términos de la anterior construcción, se traduce en : α es constructible si, comenzando en el plano de los racionales, F_0 , podemos incluir α en un campo obtenido partiendo del F_0 por un número finito de extensiones cuadráticas. Y esto es el

TEOREMA 5.K. *El número real α es constructible si y sólo si podemos encontrar un número finito de números reales $\lambda_1, \dots, \lambda_n$ tales que :*

$$1) \lambda_1^2 \in F_0,$$

$$2) \lambda_i^2 \in F_0(\lambda_1, \dots, \lambda_{i-1}) \text{ para } i = 1, 2, \dots, n,$$

tales que $\alpha \in F_0(\lambda_1, \dots, \lambda_n)$.

Podemos, además, calcular el grado de $F_0(\lambda_1, \dots, \lambda_n)$ sobre F_0 , pues según el teorema 5.a, $[F_0(\lambda_1, \dots, \lambda_n):F_0] = [F_0(\lambda_1, \dots, \lambda_n):F_0(\lambda_1, \dots, \lambda_{n-1})] \cdots [F_0(\lambda_1, \dots, \lambda_i):F_0(\lambda_1, \dots, \lambda_{i-1})] \cdots [F_0(\lambda_1):F_0]$. Como cada término del producto es 1 o 2, tenemos que $[F_0(\lambda_1, \dots, \lambda_n):F_0] = 2^r$, y tenemos, por tanto, el

COROLARIO 1. *Si α es constructible entonces α se encuentra en alguna extensión de los racionales de grado una potencia de 2.*

Si α es constructible, según el anterior corolario 1, hay un subcampo K de los reales tal que $\alpha \in K$ y tal que $[K:F_0] = 2^r$. Pero $F_0(\alpha) \subset K$, de donde, de acuerdo con el corolario al teorema 5.a $[F_0(\alpha):F_0][K:F_0] = 2^r$; por tanto $[F_0(\alpha):F_0]$ es también una potencia de 2. Pero si α satisface un polinomio irreducible de grado k sobre F_0 , hemos probado en el teorema 5.c que $[F_0(\alpha):F_0] = k$. De donde tenemos el importante criterio para no constructibilidad.

COROLARIO 2. *Si el número real α satisface un polinomio irreducible sobre el campo de los números racionales de grado k , y si k no es una potencia de 2, entonces α no es constructible.*

Este último corolario nos permite llegar a una conclusión en el viejo problema de la trisección de un ángulo por regla y compás, pues probamos que

TEOREMA 5.L. *Es imposible, usando solo regla y compás, trisecar el ángulo de 60° .*

Prueba. Si pudiéramos trisecar 60° con regla y compás, entonces la longitud $\alpha = \cos 20^\circ$ podría construirse. Recordemos en este momento la identidad $\cos 3\theta = 4\cos^3\theta - 3\cos\theta$. Haciendo $\theta = 20^\circ$ y recordando que $\cos 60^\circ = \frac{1}{2}$, obtenemos $4\alpha^3 - 3\alpha = \frac{1}{2}$, de donde $8\alpha^3 - 6\alpha - 1 = 0$. Luego α es una raíz del polinomio $8x^3 - 6x - 1$ sobre el campo racional. Pero este polinomio es irreducible sobre el campo racional (problema 7(a)), y como su grado es 3, que ciertamente no es una potencia de 2, según el corolario 2 al teorema 6.k, α no es constructible. Así pues, 60° no puede trisecarse con solo el empleo de la regla y el compás.

Otro viejo problema es el de la duplicación del cubo, es decir, el de la construcción de un cubo cuyo volumen sea dos veces el de un cubo dado. Si el cubo original es el cubo unitario, la resolución del problema lleva consigo la construcción de una longitud α tal que $\alpha^3 = 2$. Como el polinomio $x^3 - 2$ es irreducible sobre los racionales (problema 7(b)), según el corolario 2 al teorema 5.k, α no es constructible. Así pues,

TEOREMA 5.M. *Con el uso exclusivo de la regla y el compás es imposible duplicar el cubo.*

Queremos exhibir otra figura geométrica que no puede construirse con regla y compás, a saber, el heptágono regular. Para llevar a efecto tal construcción se requeriría la constructibilidad de $\alpha = 2 \cos(2\pi/7)$. Pero afirmamos que α satisface a $x^3 + x^2 - 2x - 1$ (problema 8) y que este polinomio es irreducible sobre el campo de los números racionales (problema 7(c)). De donde usando de nuevo el corolario 2 al teorema 5.k, obtenemos el

TEOREMA 5.N. *Es imposible construir un heptágono regular con regla y compás.*

Problemas

1. Pruébese que si α, β son constructibles entonces lo son también $\alpha \pm \beta, \alpha\beta$, y α/β (cuando $\beta \neq 0$).

2. Pruébese que una recta en F tiene una ecuación de la forma $ax + by + c = 0$ con a, b y c en F .

3. Pruébese que una circunferencia en F tiene una ecuación de la forma $x^2 + y^2 + ax + by + c = 0$, con a, b y c en F .

4. Pruébese que dos rectas en F que se intersecan en el plano real, se intersecan en un punto en el plano de F .

5. Pruébese que una recta en F que se interseca en el plano real con una circunferencia en F lo hace en un punto que está o en el plano de F o en el plano de $F(\sqrt{\gamma})$ donde γ es un número positivo en F .

6. Si $\gamma \in F$ es positivo, pruébese que $\sqrt{\gamma}$ es realizable como una intersección de rectas y circunferencias en F .

7. Pruébese que los siguientes polinomios son irreducibles sobre el campo de los números racionales.

a) $8x^3 - 6x - 1$.

b) $x^3 - 2$.

c) $x^3 + x^2 - 2x - 1$.

8. Pruébese que $2 \cos \frac{2\pi}{7}$ satisface a $x^3 + x^2 - 2x - 1$. (*Sugerencia:*

Úsese la igualdad $2 \cos \frac{2\pi}{7} = e^{2\pi i/7} + e^{-2\pi i/7}$.)

9. Pruébese que el pentágono regular es constructible.

10. Pruébese que el hexágono regular es constructible.

11. Pruébese que el pentadecágono regular es constructible.
12. Pruébese que es posible trisecar el ángulo de 72° .
13. Pruébese que un eneágono regular no es constructible.
- *14. Pruébese que el polígono regular de 17 lados es constructible.

5. MÁS ACERCA DE RAÍCES

Volvemos a la exposición general. Sea F un campo y, como usualmente, $F[x]$ el anillo de los polinomios en x sobre F .

DEFINICIÓN. Si $f(x) = \alpha_0 x^n + \alpha_1 x^{n-1} + \dots + \alpha_i x^{n-i} + \dots + \alpha_{n-1} x + \alpha_n$ es un polinomio en $F[x]$, entonces la *derivada* de $f(x)$, representada por $f'(x)$, es el polinomio $f'(x) = n\alpha_0 x^{n-1} + (n-1)\alpha_1 x^{n-2} + \dots + (n-i)\alpha_i x^{n-i-1} + \dots + \alpha_{n-1}$ de $F[x]$.

Dar esta definición o probar las propiedades básicas formales de la derivada en cuanto a polinomios se refiere, no requiere el concepto de límite. Pero, como el campo F es arbitrario, podemos esperar que pasen algunas cosas extrañas. Por ejemplo, si F es de característica $p \neq 0$ la derivada del polinomio x^p es $px^{p-1} = 0$. Así pues, el resultado común del cálculo de que un polinomio cuya derivada es cero debe ser una constante, no sigue siendo válido. Pero si la característica de F es 0 y si $f'(x) = 0$ para $f(x) \in F[x]$ es cierto que $f(x) = a \in F$ (véase el problema 1). Incluso cuando la característica de F es $p \neq 0$ podemos aún describir los polinomios con derivada cero; si $f'(x) = 0$ entonces $f(x)$ es un polinomio en x^p (véase el problema 2).

Probamos ahora las análogas de las reglas formales de diferenciación que tan bien conocemos.

LEMA 5.5. *Para cualesquiera $f(x), g(x) \in F[x]$ y cualquier $\alpha \in F$*

- 1) $(f(x) + g(x))' = f'(x) + g'(x);$
- 2) $(\alpha f(x))' = \alpha f'(x);$
- 3) $(f(x)g(x))' = f'(x)g(x) + f(x)g'(x).$

Prueba. Las pruebas de las partes (1) y (2) son extraordinariamente fáciles y se dejan como ejercicio. Para probar la parte (3) nótese que, de acuerdo con las partes (1) y (2), es suficiente probarla en el caso muy especial $f(x) = x^i$ y $g(x) = x^j$ donde tanto i como j son positivos. Pero entonces $f(x)g(x) = x^{i+j}$, de donde $(f(x)g(x))' = (i+j)x^{i+j-1}$; pero $f'(x)g(x) = ix^{i-1}x^j = ix^{i+j-1}$ y $f(x)g'(x) = jx^i x^{j-1} = jx^{i+j-1}$; de donde, en consecuencia, $f'(x)g(x) + f(x)g'(x) = (i+j)x^{i+j-1} = (f(x)g(x))'$.

Recuérdese que en el cálculo elemental se muestra la equivalencia entre la existencia de una raíz múltiple de una función y la anulación simultánea de la función y su derivada en un punto dado. Incluso dentro de nuestro actual marco, en el que F es un campo arbitrario, existe una tal interrelación.

LEMA 5.6. *El polinomio $f(x) \in F[x]$ tiene una raíz múltiple si y sólo si $f(x)$ y $f'(x)$ tienen un factor común no trivial (es decir, de grado positivo).*

Prueba. Antes de probar el lema, parece adecuado que hagamos observar que si $f(x)$ y $g(x)$ en $F[x]$ tienen un factor común no trivial en $K[x]$, para una K extensión de F , entonces tienen un factor común no trivial en $F[x]$. En efecto, si fueran primos relativos como elementos en $F[x]$, entonces podrían encontrarse dos polinomios $a(x)$ y $b(x)$ en $F[x]$ tales que $a(x)f(x) + b(x)g(x) = 1$. Como esta relación también se verifica para estos elementos vistos como elementos de $K[x]$, deberían ser también primos relativos en $K[x]$.

Vamos ahora con el lema. De la observación que acabamos de hacer podemos suponer, sin pérdida de generalidad, que las raíces de $f(x)$ se encuentran todas en F (de otra manera extendemos F hasta K , el campo de descomposición de $f(x)$). Si $f(x)$ tiene una raíz múltiple α entonces $f(x) = (x - \alpha)^m q(x)$ donde $m > 1$. Pero, como puede calcularse de inmediato, $((x - \alpha)^m)' = m(x - \alpha)^{m-1}$, de donde, según el lema 5.5, $f'(x) = (x - \alpha)^m q'(x) + m(x - \alpha)^{m-1} q(x) = (x - \alpha)r(x)$, ya que $m > 1$. Pero esto nos dice que $f(x)$ y $f'(x)$ tienen $x - \alpha$ como factor común, con lo que el lema queda probado en una dirección.

Por otra parte, si $f(x)$ no tiene ninguna raíz múltiple, entonces $f(x) = (x - \alpha_1)(x - \alpha_2) \dots (x - \alpha_n)$, donde las α_i son todas distintas (estamos suponiendo que $f(x)$ es mónico). Pero entonces $f'(x) = \sum_{i=1}^n (x - \alpha_1) \dots (\widehat{x - \alpha_i}) \dots (x - \alpha_n)$ donde la $\widehat{}$ determina el término que se ha suprimido. Afirmamos que ninguna raíz de $f(x)$ es una raíz de $f'(x)$, pues si $f'(\alpha_i) = \prod_{j \neq i} (\alpha_i - \alpha_j) \neq 0$, ya que las raíces son todas distintas. Pero si $f(x)$ y $f'(x)$ tienen un factor común no trivial, tienen una raíz común, a saber, cualquier raíz de este factor común. El resultado neto es que $f(x)$ y $f'(x)$ no tienen ningún factor común no trivial, con lo que el lema ha sido probado en la otra dirección.

COROLARIO 1. *Si $f(x) \in F[x]$ es irreducible, entonces :*

- 1) *Si la característica de F es 0, $f(x)$ no tiene raíces múltiples.*
- 2) *Si la característica de F es $p \neq 0$, $f(x)$ tiene una raíz múltiple sólo si es de la forma $f(x) = g(x^p)$.*

Prueba. Como $f(x)$ es irreducible, sus únicos factores en $F[x]$ son 1 y $f(x)$. Si $f(x)$ tiene una raíz múltiple, entonces $f(x)$ y $f'(x)$ tienen un factor

común no trivial de acuerdo con el lema, de donde $f(x) | f'(x)$. Pero como el grado de $f'(x)$ es menor que el de $f(x)$, la única forma posible de que esto suceda es que $f'(x)$ sea 0. En característica 0 esto implica que $f(x)$ es una constante, que no tiene ninguna raíz; cuando la característica es $p \neq 0$, esto obliga a que $f(x) = g(x^p)$.

Volveremos dentro de un momento a discutir las implicaciones del corolario 1 más completamente. Pero antes, para su posterior uso en el capítulo 7 en nuestro tratamiento de campos finitos, probaremos un caso más bien particular

COROLARIO 2. *Si F es un campo de característica $p \neq 0$, entonces el polinomio $x^{p^n} - x \in F[x]$, tiene, para $n \geq 1$, raíces distintas.*

Prueba. La derivada de $x^{p^n} - x$ es $p^n x^{p^n-1} - 1 = -1$, ya que F es de característica p . Por tanto, $x^{p^n} - x$ y su derivada son ciertamente primos relativos, lo que, según el lema, implica que $x^{p^n} - x$ no tiene raíces múltiples.

El corolario 1 no descarta la posibilidad de que en característica $p \neq 0$ un polinomio irreducible pueda tener raíces múltiples. Para fijar ideas, exhibimos un ejemplo en donde lo dicho es lo que realmente sucede. Sea F_0 un campo de característica 2 y sea $F = F_0(x)$ el campo de las funciones racionales en x sobre F_0 . Afirmamos que el polinomio $t^2 - x$ en $F[t]$ es irreducible sobre F y que sus raíces son iguales. Para probar la irreducibilidad debemos demostrar que no hay ninguna función racional en $F_0(x)$ cuyo cuadrado sea x ; este es el contenido del problema 4. Para ver que $t^2 - x$ tiene una raíz múltiple, nótese que su derivada (la derivada es con respecto a t , pues x estando en F , se considera como una constante) es $2t = 0$. Desde luego, el ejemplo análogo funciona para cualquier característica prima.

Ahora que hemos visto que la posibilidad es una realidad, se señala una aguda diferencia entre los casos de característica 0 y los de característica p . La presencia de polinomios irreducibles con raíces múltiples en el último caso, nos lleva hasta muchas sutilezas tan interesantes como complicadas. Su estudio requiere un tratamiento más elaborado y sofisticado que preferimos evitar en este nivel. *Por tanto, para el resto de este capítulo convenimos en que todos los campos que aparecen en el texto propiamente dicho, son campos de característica 0.*

DEFINICIÓN. La extensión K de F es una extensión simple de F si $K = F(\alpha)$, para algún α en K .

En característica 0 (o en extensiones propiamente condicionadas en característica $p \neq 0$; véase problema 14) todas las extensiones finitas son realizables como extensiones simples. Este resultado es el

TEOREMA 5.P. Si F es de característica 0 y si a y b son algebraicos sobre F , entonces existe un elemento $c \in F(a, b)$ tal que $F(a, b) = F(c)$.

Prueba. Sean $f(x)$ y $g(x)$, de grados m y n , los polinomios irreducibles sobre F satisfechos por a y b respectivamente. Sea K una extensión de F en que tanto $f(x)$ como $g(x)$ se descomponen completamente. Como la característica de F es 0 todas las raíces de $f(x)$ son distintas, y lo mismo ocurre con las de $g(x)$. Sean las raíces de $f(x)$, $a = a_1, a_2, \dots, a_m$ y las de $g(x)$, $b = b_1, b_2, \dots, b_n$.

Si $j \neq 1$, entonces $b_j \neq b_1 = b$, de donde la ecuación $a_i + \lambda b_j = a_1 + \lambda b_1 = a + \lambda b$ tiene solamente una solución λ en K , a saber,

$$\lambda = \frac{a_i - a}{b - b_1}.$$

Como F es de característica 0 tiene un número infinito de elementos, de donde resulta que podemos encontrar un elemento $\gamma \in F$ tal que $a_i + \gamma b_j \neq a + \gamma b$ para todo i y para toda $j \neq 1$. Sea $c = a + \gamma b$; nuestra tesis es que $F(c) = F(a, b)$. Como $c \in F(a, b)$ no hay duda de que $F(c) \subset F(a, b)$. Demostremos que tanto a como b están en $F(c)$ de lo que se sigue que $F(a, b) \subset F(c)$.

Como b satisface al polinomio $g(x)$ sobre F , lo satisface también cuando lo consideramos un polinomio sobre $K = F(c)$. Además, si $h(x) = f(c - \gamma x)$, entonces $h(x) \in K[x]$ y $h(b) = f(c - \gamma b) = f(a) = 0$, ya que $a = c - \gamma b$. Luego en una extensión de K , $h(x)$ y $g(x)$ tienen $x - b$ como factor común. Aseguramos que $x - b$ es, en realidad, su máximo común divisor. Pues si $b_j \neq b$ es otra raíz de $g(x)$, entonces $h(b_j) = f(c - \gamma b_j) = 0$, ya que, por nuestra elección de γ , $c - \gamma b_j$ para $j \neq 1$ esquiva todas las raíces a_i de $f(x)$. Además, como $(x - b)^2 \nmid g(x)$, $(x - b)^2$ no puede dividir al máximo común divisor de $h(x)$ y $g(x)$. Así pues, $x - b$ es el máximo común divisor de $h(x)$ y $g(x)$ sobre alguna extensión de K . Pero entonces tienen un máximo común divisor no trivial sobre K , que debe ser un divisor de $x - b$. Como el grado de $x - b$ es 1, vemos que el máximo común divisor de $g(x)$ y $h(x)$ en $K[x]$ es exactamente $x - b$. Luego $x - b \in K[x]$, de donde $b \in K$; recordando que $K = F(c)$, obtenemos que $b \in F(c)$. Como $a = c - \gamma b$, y como $b, c \in F(c)$, $\gamma \in F \subset F(c)$, tenemos que $a \in F(c)$, de donde $F(a, b) \subset F(c)$. Las dos relaciones de contención opuestas nos dicen que $F(a, b) = F(c)$.

Un simple argumento de inducción extiende el resultado de dos elementos a cualquier número finito, es decir, si $\alpha_1, \dots, \alpha_n$ son algebraicos sobre F , entonces hay un elemento $c \in F(\alpha_1, \dots, \alpha_n)$ tales que $F(c) = F(\alpha_1, \dots, \alpha_n)$. Luego el

COROLARIO. Cualquier extensión finita de un campo de característica 0 es una extensión simple.

Problemas

1. Si F es de característica 0 y $f(x) \in F[x]$ es tal que $f'(x) = 0$, pruébese que $f(x) = \alpha \in F$.
2. Si F es de característica $p \neq 0$ y si $f(x) \in F[x]$ es tal que $f'(x) = 0$, pruébese que $f(x) = g(x^n)$ para algún polinomio $g(x) \in F[x]$.
3. Pruébese que $(f(x) + g(x))' = f'(x) + g'(x)$ y que $(\alpha f(x))' = \alpha f'(x)$ para $f(x), g(x) \in F[x]$ y $\alpha \in F$.
4. Pruébese que no hay ninguna función racional en $F(x)$ tal que su cuadrado sea x .
5. Complétense la inducción necesaria para establecer el corolario al teorema 5.p.

Un elemento a en una extensión K de F se llama *separable sobre F* si satisface un polinomio sobre F que no tiene raíces múltiples. Una extensión K de F se llama *separable sobre F* si todos sus elementos son separables sobre F . Un campo F se llama *perfecto* si todas las extensiones finitas de F son separables.

6. Pruébese que cualquier campo de característica 0 es perfecto.
7. a) Si F es de característica $p \neq 0$ muéstrese que para $a, b \in F$, $(a+b)^{p^m} = a^{p^m} + b^{p^m}$.
b) Si F es de característica $p \neq 0$ y si K es una extensión de F , sea $T = \{a \in K \mid a^{p^n} \in F \text{ para algún } n\}$. Pruébese que T es un subcampo de K .
8. Si K, T, F son como en el problema 7(b), pruébese que cualquier automorfismo de K que deja fijos todos los elementos de F deja también fijos todos los elementos de T .
- *9. Demuéstrese que un campo F de característica $p \neq 0$ es perfecto si y sólo si para cualquier $a \in F$ podemos encontrar un $b \in F$ tal que $b^p = a$.
10. Usando el resultado del problema 9, pruébese que cualquier campo finito es perfecto.
- **11. Si K es una extensión de F pruébese que el conjunto de elementos en K que son separables sobre F forma un subcampo de K .
12. Si F es de característica $p \neq 0$ y si K es una extensión finita de F , pruébese que dado $a \in K$ o $a^{p^n} \in F$ para algún n o podemos encontrar un entero m tal que $a^{p^m} \notin F$ y es separable sobre F .
13. Si K y F son como en el problema 12, y si ningún elemento que está en K , pero no en F , es separable sobre F , pruébese que dado $a \in K$ podemos encontrar un entero n , dependiente de a , tal que $a^{p^n} \in F$.

14. Si K es una extensión finita y separable de F , pruébese que K es una extensión simple de F .

15. Si uno de los elementos a o b es separable sobre F , pruébese que $F(a, b)$ es una extensión simple de F .

6. ELEMENTOS DE LA TEORÍA DE GALOIS

Dado un polinomio $p(x)$ en $F[x]$, el anillo de polinomios en x sobre F , asociaremos con $p(x)$ un grupo al que llamaremos el *grupo de Galois* de $p(x)$. Hay una relación muy estrecha entre las raíces de un polinomio y su grupo de Galois; en realidad, el grupo de Galois resultará ser un cierto grupo de permutaciones de las raíces del polinomio. Haremos un estudio de estas ideas en esta y las próximas secciones.

Introduciremos este grupo por medio del campo de descomposición de $p(x)$ sobre F , quedando definido el grupo de Galois de $p(x)$ como un cierto grupo de automorfismos de este campo de descomposición. Es esta la razón de que en tantos de los teoremas que vamos ahora a ver nos ocupemos de los automorfismos de un campo. Entre los subgrupos del grupo de Galois y los subcampos del campo de descomposición, existe una hermosa dualidad que expresa el teorema fundamental de la teoría de Galois (teorema 5.v). De esto derivaremos una condición para la solubilidad por medio de radicales de las raíces de un polinomio en términos de la estructura algebraica de su grupo de Galois. De esta condición derivaremos, a su vez, el clásico resultado de Abel sobre la no solubilidad por radicales del polinomio general de grado 5. Durante el proceso derivaremos, también, como resultados colaterales, teoremas que, de por sí, son de gran interés. Uno de ellos será el teorema fundamental sobre funciones simétricas. Nuestro enfoque del tema se basa en el tratamiento dado por Artin.

Recuérdese que estamos suponiendo que todos nuestros campos son de característica 0, de donde resulta que podemos hacer (y haremos) libre uso del teorema 5.p y su corolario.

Por un *automorfismo* del campo K entenderemos, como es común, una aplicación σ de K sobre sí mismo tal que $\sigma(a+b) = \sigma(a)+\sigma(b)$ y $\sigma(ab) = \sigma(a)\sigma(b)$ para $a, b \in K$ cualesquiera. Dos automorfismos σ y τ de K se dice que son distintos si $\sigma(a) \neq \tau(a)$ para al menos un elemento $a \in K$.

Comenzamos con el siguiente

TEOREMA 5.Q. *Si K es un campo y si $\sigma_1, \dots, \sigma_n$ son distintos automorfismos de K , entonces es imposible encontrar elementos a_1, \dots, a_n , no todos 0, en K , tales que $a_1\sigma_1(u) + a_2\sigma_2(u) + \dots + a_n\sigma_n(u) = 0$ para todo $u \in K$.*

Prueba. Supongamos que pudiéramos encontrar un conjunto de elementos a_1, \dots, a_n en K , no todos cero, tales que $a_1\sigma_1(u) + \dots + a_n\sigma_n(u) = 0$ para todo $u \in K$.

para todo $u \in K$. Entonces podríamos encontrar una relación tal que tuviera tan pocos términos como fuera posible; renumerando, si fuera preciso, podemos suponer que esta relación mínima es

$$(1) \quad a_1\sigma_1(u) + \dots + a_m\sigma_m(u) = 0$$

donde a_1, \dots, a_m son todos diferentes de 0.

Si m fuera igual a 1 entonces $a_1\sigma_1(u) = 0$ para todo $u \in K$, lo que nos llevaría a $a_1 = 0$, en contra de lo supuesto. Podemos, pues, suponer que $m > 1$. Como los automorfismos son distintos hay un elemento $c \in K$ tal que $\sigma_1(c) \neq \sigma_m(c)$. Como $cu \in K$ para todo $u \in K$, la relación (1) debe también verificarse para cu , es decir, $a_1\sigma_1(cu) + a_2\sigma_2(cu) + \dots + a_m\sigma_m(cu) = 0$ para todo $u \in K$. Usando la hipótesis de que las σ son automorfismos de K , esta relación toma la forma

$$(2) \quad a_1\sigma_1(c)\sigma_1(u) + a_2\sigma_2(c)\sigma_2(u) + \dots + a_m\sigma_m(c)\sigma_m(u) = 0.$$

Multiplicando la relación (1) por $\sigma_1(c)$ y restando el resultado de (2) obtenemos

$$(3) \quad a_2(\sigma_2(c) - \sigma_1(c))\sigma_2(u) + \dots + a_m(\sigma_m(c) - \sigma_1(c))\sigma_m(u) = 0.$$

Si hacemos $b_i = a_i(\sigma_i(c) - \sigma_1(c))$ para $i = 2, \dots, m$, entonces los b_i están en K , $b_m = a_m(\sigma_m(c) - \sigma_1(c)) \neq 0$, ya que $a_m \neq 0$ y $\sigma_m(c) - \sigma_1(c) \neq 0$, aunque $b_2\sigma_2(u) + \dots + b_m\sigma_m(u) = 0$ para todo $u \in K$. Esto produce una relación más corta, en contra de la elección que hicimos; luego el teorema está probado.

DEFINICIÓN. Si G es un grupo de automorfismos de K , entonces el *campo fijo* de G es el conjunto de todos los elementos $a \in K$ tales que $\sigma(a) = a$ para todo $\sigma \in G$.

Nótese que esta definición tiene sentido, incluso si G no es un grupo, sino simplemente un conjunto de automorfismos de K . Pero el campo fijo de un conjunto de automorfismos y el del grupo de automorfismos generado por este conjunto (en el grupo de todos los automorfismos de K) son iguales (problema 1), de donde nada perdemos por definir el concepto solo para grupos de automorfismos. Además, únicamente estaremos interesados en los campos fijos de grupos de automorfismos.

Habiendo llamado en la anterior definición *campo fijo* de G al conjunto que allí se define, sería agradable comprobar que la terminología empleada en este caso es en verdad exacta. Es lo que nos dice el

LEMA 5.7. *El campo fijo de G es un subcampo de K .*

Prueba. Sean a, b elementos del campo fijo de G . Para todo $\sigma \in G$, tenemos, pues, $\sigma(a) = a$ y $\sigma(b) = b$. Pero entonces, $\sigma(a \pm b) = \sigma(a) \pm$

$\sigma(b) = a \pm b$, y de la misma forma, $\sigma(ab) = \sigma(a)\sigma(b) = ab$; de donde $a \pm b$ y ab están también en el campo fijo de G . Si $b \neq 0$, entonces $\sigma(b^{-1}) = \sigma(b)^{-1} = b^{-1}$, de donde b^{-1} también se encuentra en el campo fijo de G . Luego hemos verificado que el campo fijo de G es, ciertamente, un subcampo de K .

Nos ocuparemos de los automorfismos de un campo que se comportan de una forma determinada sobre un subcampo dado.

DEFINICIÓN. Sea K un campo y sea F un subcampo de K . Entonces el grupo de automorfismos de K relativos a F , que representaremos por $G(K, F)$, es el conjunto de todos los automorfismos de K que dejan fijos todos los elementos de F ; es decir, el automorfismo σ de K está en $G(K, F)$ si y sólo si $\sigma(\alpha) = \alpha$ para todo $\alpha \in F$.

No es sorprendente, y es muy fácil de probar, el siguiente

LEMA 5.8 $G(K, F)$ es un subgrupo del grupo de todos los automorfismos de K .

Dejamos la prueba de este lema al lector. Una observación : K contiene el campo de los números racionales F_0 , ya que K es de característica 0 y es fácil ver que el campo fijo de cualquier grupo de automorfismos de K , siendo un campo, debe contener a F_0 . De aquí que todo número racional permanece fijo en todo automorfismo de K .

Hacemos una pausa para examinar unos cuantos ejemplos de los conceptos que acabamos de presentar.

EJEMPLO 1. Sea K el campo de los números complejos y sea F el campo de los números reales. Calculamos $G(K, F)$. Si σ es un automorfismo cualquiera de K , como $i^2 = -1$, $\sigma(i)^2 = \sigma(i^2) = \sigma(-1) = -1$, de donde $\sigma(i) = \pm i$. Si, además, σ deja fijos a todos los reales, entonces para cualquier $a + bi$ donde a y b son reales, $\sigma(a + bi) = \sigma(a) + \sigma(b)i = a \pm bi$. Cada una de estas posibilidades, es decir, la aplicación $\sigma_1(a + bi) = a + bi$ y $\sigma_2(a + bi) = a - bi$ define un automorfismo de K ; σ_1 es el automorfismo identidad y σ_2 la conjugación compleja. Así pues, $G(K, F)$ es un grupo de orden 2.

¿Cuál es el campo fijo de $G(K, F)$? Debe, ciertamente, contener a F , ¿pero contiene algo más? Si $a + bi$ está en el campo fijo de $G(K, F)$ entonces $a + bi = \sigma_2(a + bi) = a - bi$ de donde $b = 0$ y $a = a + bi \in F$. En este caso vemos que el campo fijo de $G(K, F)$ es precisamente el mismo F .

EJEMPLO 2. Sea F_0 el campo de los números racionales y sea $K = F_0(\sqrt[3]{2})$ donde $\sqrt[3]{2}$ es la raíz cúbica real de 2. Todo elemento en K es de la forma $a_0 + a_1\sqrt[3]{2} + a_2(\sqrt[3]{2})^2$ donde a_0, a_1 y a_2 son números racionales. Si σ es un

automorfismo de K , entonces $\sigma(\sqrt[3]{2})^3 = \sigma((\sqrt[3]{2})^3) = \sigma(2) = 2$, de donde $\sigma(\sqrt[3]{2})$ debe también ser una raíz cúbica de 2 perteneciente a K . Pero hay solamente una raíz cúbica real de 2, y como K es un subcampo del campo real, debemos tener que $\sigma(\sqrt[3]{2}) = \sqrt[3]{2}$. Pero entonces $\sigma(\alpha_0 + \alpha_1 \sqrt[3]{2} + \alpha_2 (\sqrt[3]{2})^2) = \alpha_0 + \alpha_1 \sqrt[3]{2} + \alpha_2 (\sqrt[3]{2})^2$, es decir, σ es el automorfismo identidad de K . Vemos, pues, que $G(K, F_0)$ consta solo de la aplicación identidad, y en este caso el campo fijo de $G(K, F_0)$ no es F_0 , sino que en realidad es bastante mayor, pues es todo K .

EJEMPLO 3. Sea F_0 el campo de los números racionales y sea $\omega = e^{2\pi i/5}$; tenemos pues que $\omega^5 = 1$ y que ω satisface al polinomio $x^4 + x^3 + x^2 + x + 1$ sobre F_0 . Por el criterio de Eisenstein se puede probar que $x^4 + x^3 + x^2 + x + 1$ es irreducible sobre F_0 (véase el problema 3). Así pues, $K = F_0(\omega)$ es de grado 4 sobre F_0 y todo elemento de K es de la forma $\alpha_0 + \alpha_1 \omega + \alpha_2 \omega^2 + \alpha_3 \omega^3$ donde todos los $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ están en F_0 . Ahora bien, para cualquier automorfismo σ de K , $\sigma(\omega) \neq 1$, ya que $\sigma(1) = 1$, y $\sigma(\omega)^5 = \sigma(\omega^5) = \sigma(1) = 1$, de donde $\sigma(\omega)$ es también una raíz quinta de la unidad. En consecuencia, $\sigma(\omega)$ puede solamente ser $\omega, \omega^2, \omega^3$ o ω^4 . Afirmando que cada una de estas posibilidades ocurre realmente, pues definamos las cuatro aplicaciones $\sigma_1, \sigma_2, \sigma_3$ y σ_4 por $\sigma_i(\alpha_0 + \alpha_1 \omega + \alpha_2 \omega^2 + \alpha_3 \omega^3) = \alpha_0 + \alpha_1 (\omega^i) + \alpha_2 (\omega^i)^2 + \alpha_3 (\omega^i)^3$, para $i = 1, 2, 3, 4$. Cada uno de ellos define un automorfismo de K (problema 4). Por tanto, como $\sigma \in G(K, F_0)$ está completamente determinado por $\sigma(\omega)$, $G(K, F_0)$ es un grupo de orden 4, con σ_1 como su elemento unidad. Como $\sigma_2^2 = \sigma_4, \sigma_2^3 = \sigma_3$ y $\sigma_2^4 = \sigma_1$, $G(K, F_0)$ es un grupo cíclico de orden 4. Se puede fácilmente probar que el campo fijo de $G(K, F_0)$ es F_0 (problema 5). El subgrupo $A = \{\sigma_1, \sigma_4\}$ de $G(K, F_0)$ tiene como su campo fijo el conjunto de todos los elementos $\alpha_0 + \alpha_2 (\omega^2 + \omega^3)$, que es una extensión de F_0 de grado 2.

Los ejemplos, aunque ilustrativos, son aún demasiado especiales, pues puede observarse que en cualquiera de ellos $G(K, F)$ resulta ser un grupo cíclico. Esto es extraordinariamente atípico, pues, en general, $G(K, F)$ no necesita ser ni siquiera abeliano (véase el teorema 5.a). Pero, a pesar de su carácter especial, traen a luz ciertos hechos importantes. Por una parte, muestran que debemos estudiar el efecto de los automorfismos sobre las raíces de los polinomios y, por otra, subrayan que F no necesariamente ha de ser igual a todo el campo fijo de $G(K, F)$. Los casos en que esto sucede son muy convenientes y son situaciones a las que dentro de poco dedicaremos mucho tiempo y esfuerzo.

Calculamos ahora una importante cota de la magnitud de $G(K, F)$.

TEOREMA 5.R. Si K es una extensión finita de F , entonces $G(K, F)$ es un grupo finito y su orden, $o(G(K, F))$, satisface $o(G(K, F)) \leq [K : F]$.

Prueba. Sea $[K:F] = n$ y supongamos que u_1, \dots, u_n es una base de K sobre F . Supongamos que podemos encontrar $n+1$ automorfismos distintos $\sigma_1, \sigma_2, \dots, \sigma_{n+1}$ en $G(K, F)$. De acuerdo con el corolario al teorema 4.f el sistema de n ecuaciones lineales homogéneas en las $n+1$ incógnitas x_1, \dots, x_{n+1} :

$$\begin{aligned} \sigma_1(u_1)x_1 + \sigma_2(u_1)x_2 + \cdots + \sigma_{n+1}(u_1)x_{n+1} &= 0 \\ \vdots \\ \sigma_1(u_i)x_1 + \sigma_2(u_i)x_2 + \cdots + \sigma_{n+1}(u_i)x_{n+1} &= 0 \\ \vdots \\ \sigma_1(u_n)x_1 + \sigma_2(u_n)x_2 + \cdots + \sigma_{n+1}(u_n)x_{n+1} &= 0 \end{aligned}$$

tiene una solución no trivial (no toda 0) $x_1 = a_1, \dots, x_{n+1} = a_{n+1}$ en K . Luego

$$1) \quad a_1\sigma_1(u_i) + a_2\sigma_2(u_i) + \cdots + a_{n+1}\sigma_{n+1}(u_i) = 0$$

para $i = 1, 2, \dots, n$.

Como cada uno de los σ_i deja fijo a todo elemento de F y como un elemento arbitrario t de K es de la forma $t = x_1u_1 + \dots + x_nu_n$ con x_1, \dots, x_n en F , entonces, por el sistema de ecuaciones (1), tenemos $a_1\sigma_1(t) + \dots + a_{n+1}\sigma_{n+1}(t) = 0$ para toda $t \in K$. Pero esto contradice el resultado del teorema 5.q. Luego el teorema 5.r ha sido probado.

El teorema 5.r es de importancia central en la teoría de Galois. Pero aparte del papel que allí juega nos sirve también para probar un resultado clásico concerniente a las funciones racionales simétricas. Este resultado sobre funciones simétricas, a su vez juega un papel importante en la teoría de Galois.

Hagamos primero algunas observaciones sobre el campo de las funciones racionales en n variables sobre un campo F . Recordemos que en la sección 11 del capítulo 3 definimos el anillo de los polinomios en las n variables x_1, \dots, x_n sobre F y de esto pasamos a definir el campo de las funciones racionales en x_1, \dots, x_n , $F(x_1, \dots, x_n)$, sobre F como el anillo de todos los cocientes de tales polinomios.

Sea S_n el grupo simétrico de grado n considerado como si actuara sobre el conjunto $[1, 2, \dots, n]$; para $\sigma \in S_n$ e i un entero con $1 \leq i \leq n$, sea $\sigma(i)$ la imagen de i bajo σ . Podemos hacer actuar a S_n sobre $F(x_1, \dots, x_n)$ en la siguiente forma: para $\sigma \in S_n$ y $r(x_1, \dots, x_n) \in F(x_1, \dots, x_n)$, definimos la aplicación que lleva $r(x_1, \dots, x_n)$ sobre $r(x_{\sigma(1)}, \dots, x_{\sigma(n)})$. Representaremos a esta aplicación de $F(x_1, \dots, x_n)$ sobre sí mismo también por σ . Es obvio que estas aplicaciones definen automorfismos de $F(x_1, \dots, x_n)$. ¿Cuál es el campo fijo de $F(x_1, \dots, x_n)$ respecto a S_n ? Consiste simplemente en todas las funciones racionales $r(x_1, \dots, x_n)$ tales que $r(x_1, \dots, x_n) = r(x_{\sigma(1)}, \dots,$

$x_{\sigma(n)}$) para todo $\sigma \in S_n$. Pero estos son precisamente aquellos elementos en $F(x_1, \dots, x_n)$ que se conocen como *funciones racionales simétricas*. Como son el campo fijo de S_n forman un subcampo de $F(x_1, \dots, x_n)$ llamado el campo de las funciones racionales simétricas al que representaremos por S . Nos ocuparemos de estos tres problemas :

- 1) ¿A qué es igual $[F(x_1, \dots, x_n):S]$?
- 2) ¿Qué es $G(F(x_1, \dots, x_n), S)$?
- 3) ¿Podemos describir S en términos de alguna extensión simple particular de F ?

Contestaremos a estas tres preguntas simultáneamente.

Podemos presentar explícitamente algunas funciones particularmente sencillas de S construidas con x_1, \dots, x_n conocidas como *funciones simétricas elementales* en x_1, \dots, x_n . Las definimos como sigue :

$$a_1 = x_1 + x_2 + \cdots + x_n = \sum_{i=1}^n x_i$$

$$a_2 = \sum_{i < j} x_i x_j$$

$$a_3 = \sum_{i < j < k} x_i x_j x_k$$

.

$$a_n = x_1 x_2 \cdots x_n.$$

Probar que estas son funciones simétricas se deja como ejercicio. Para $n = 2, 3$ y 4 las escribimos explícitamente a continuación.

$$n = 2$$

$$a_1 = x_1 + x_2.$$

$$a_2 = x_1 x_2.$$

$$n = 3$$

$$a_1 = x_1 + x_2 + x_3.$$

$$a_2 = x_1 x_2 + x_1 x_3 + x_2 x_3.$$

$$a_3 = x_1 x_2 x_3.$$

$$n = 4$$

$$a_1 = x_1 + x_2 + x_3 + x_4.$$

$$a_2 = x_1 x_2 + x_1 x_3 + x_1 x_4 + x_2 x_3 + x_2 x_4 + x_3 x_4.$$

$$a_3 = x_1 x_2 x_3 + x_1 x_2 x_4 + x_1 x_3 x_4 + x_2 x_3 x_4.$$

$$a_4 = x_1 x_2 x_3 x_4.$$

Nótese que cuando $n = 2$, x_1 y x_2 son las raíces del polinomio $t^2 - a_1 t + a_2$, cuando $n = 3$, x_1 , x_2 y x_3 son las raíces de $t^3 - a_1 t^2 + a_2 t - a_3$, y cuando $n = 4$, x_1 , x_2 , x_3 y x_4 son, todas, raíces de $t^4 - a_1 t^3 + a_2 t^2 - a_3 t + a_4$.

Como a_1, \dots, a_n están, todos, en S el campo $F(a_1, \dots, a_n)$ obtenido por la adjunción de a_1, \dots, a_n a F debe encontrarse en S . Nuestro objetivo es ahora doble, a saber, probar que

$$1) [F(x_1, \dots, x_n):S] = n!.$$

$$2) S = F(a_1, \dots, a_n).$$

Como el grupo S_n es un grupo de automorfismos de $F(x_1, \dots, x_n)$ que deja a S fijo, $S_n \subset G(F(x_1, \dots, x_n), S)$. Luego, según el teorema 5.r, $[F(x_1, \dots, x_n):S] \geq o(G(F(x_1, \dots, x_n), S)) \geq o(S_n) = n!$. Si pudiéramos demostrar que $[F(x_1, \dots, x_n):F(a_1, \dots, a_n)] \leq n!$, entonces, como $F(a_1, \dots, a_n)$, es un subcampo de S , tendríamos $n! \geq [F(x_1, \dots, x_n):F(a_1, \dots, a_n)] = [F(x_1, \dots, x_n):S] [S:F(a_1, \dots, a_n)] \geq n!$. Pero entonces tendríamos que $[F(x_1, \dots, x_n):S] = n!$, $[S:F(a_1, \dots, a_n)] = 1$ y, por tanto, $S = F(a_1, \dots, a_n)$, y, finalmente, $S_n = G(F(x_1, \dots, x_n), S)$ (esto último por lo afirmado en la segunda oración de este párrafo). Estas son precisamente las conclusiones que buscamos.

Así pues, para concluir con todo este asunto solo debemos probar que $[F(x_1, \dots, x_n):F(a_1, \dots, a_n)] \leq n!$. Para ver esto, observemos primero que el polinomio $p(t) = t^n - a_1 t^{n-1} + a_2 t^{n-2} \dots + (-1)^n a_n$, que tiene coeficientes en $F(a_1, \dots, a_n)$, se factoriza sobre $F(x_1, \dots, x_n)$ como $p(t) = (t - x_1)(t - x_2) \dots (t - x_n)$ (éste es en realidad el origen de las funciones simétricas elementales). Así pues, $p(t)$ de grado n sobre $F(a_1, \dots, a_n)$, se descompone en un producto de factores lineales sobre $F(x_1, \dots, x_n)$. No puede descomponerse sobre un subcampo propio de $F(x_1, \dots, x_n)$ que contenga a $F(a_1, \dots, a_n)$, pues este subcampo tendría entonces que contener tanto a F como a cada una de las raíces de $p(t)$, es decir, a x_1, x_2, \dots, x_n ; pero entonces este subcampo sería todo $F(x_1, \dots, x_n)$. Así pues, vemos que $F(x_1, \dots, x_n)$ es el campo de descomposición del polinomio $p(t) = t^n - a_1 t^{n-1} + \dots + (-1)^n a_n$ sobre $F(a_1, \dots, a_n)$. Como $p(t)$ es de grado n , según el teorema 5.h, tenemos $[F(x_1, \dots, x_n):F(a_1, \dots, a_n)] \leq n!$. De donde todas nuestras afirmaciones quedan probadas. Resumimos todo este estudio en el siguiente básico e importante resultado.

TEOREMA 5.s. *Sea F un campo y $F(x_1, \dots, x_n)$ el campo de las funciones racionales en x_1, \dots, x_n sobre F . Supongamos que S es el campo de las funciones racionales simétricas; entonces*

$$1) [F(x_1, \dots, x_n):S] = n!.$$

$$2) G(F(x_1, \dots, x_n), S) = S_n, \text{el grupo simétrico de grado } n.$$

$$3) \text{Si } a_1, \dots, a_n \text{ son las funciones simétricas elementales en } x_1, \dots, x_n, \text{ entonces } S = F(a_1, \dots, a_n).$$

- 4) $F(x_1, \dots, x_n)$ es el campo de descomposición sobre $F(a_1, \dots, a_n) = S$ del polinomio $t^n - a_1 t^{n-1} + a_2 t^{n-2} \dots + (-1)^n a_n$.

Mencionamos anteriormente que dado un entero cualquiera n es posible construir un campo y un polinomio de grado n sobre este campo cuyo campo de descomposición sea del máximo grado posible, $n!$, sobre este campo. El teorema 5.s nos proporciona explícitamente tal ejemplo, pues si hacemos $S = F(a_1, \dots, a_n)$, el campo de las funciones racionales en n variables a_1, \dots, a_n y consideramos el campo de descomposición del polinomio $t^n - a_1 t^{n-1} + a_2 t^{n-2} \dots + (-1)^n a_n$ sobre S , entonces vemos que es de grado $n!$ sobre S .

La parte (3) del teorema 5.s es un teorema muy clásico. *Afirma que una función racional simétrica en n variables es una función racional en las funciones simétricas elementales de estas variables.* Este resultado puede hacerse aún más sólido : un polinomio simétrico en n variables es un *polinomio* en sus funciones simétricas elementales (véase el problema 7). Este resultado se conoce como el *teorema sobre polinomios simétricos*.

En los ejemplos discutidos de grupos de automorfismos de campos y de campos fijos bajo tales grupos, vimos que podía muy bien suceder que F fuera realmente menor que el campo fijo total de $G(K, F)$. Ciertamente, F está siempre contenido en este campo, pero no necesariamente lo llena. Así pues, imponer la condición sobre una extensión K de F que F sea precisamente el campo fijo de $G(K, F)$ es una limitación genuina sobre el tipo de extensión de F que estamos considerando. Es en esta clase de extensión en la que estamos más interesados.

DEFINICIÓN. K es una extensión normal de F si K es una extensión finita de F tal que F es el campo fijo de $G(K, F)$.

Otro modo de decir lo mismo: si K es una extensión normal de F , entonces todo elemento de K que no está en F sufre alteración por algún elemento de $G(K, F)$. En los ejemplos discutidos, los ejemplos 1 y 3 eran extensiones normales, mientras que el ejemplo 2 no lo era.

Una consecuencia inmediata de la hipótesis de normalidad es que nos permite calcular con gran precisión el tamaño del campo fijo de cualquier subgrupo de $G(K, F)$ y, en particular, dar más fuerza al enunciado del teorema 5.r, cambiando la desigualdad que en él aparece en una igualdad.

TEOREMA 5.t. *Sea K una extensión normal de F y sea H un subgrupo de $G(K, F)$; sea $K_H = \{x \in K \mid \sigma(x) = x \text{ para toda } \sigma \in H\}$ el campo fijo de H . Entonces :*

- 1) $[K:K_H] = o(H)$.
- 2) $H = G(K, K_H)$

(En particular, cuando $H = G(K, F)$, $[K:F] = o(G(K, F))$.)

Prueba. Como todos los elementos de H dejan fijos a todos los elementos de K_H , es claro que $H \subset G(K, K_H)$. De acuerdo con el teorema 5.r sabemos que $[K:K_H] \geq o(G(K, K_H))$; y como $o(G(K, K_H)) \geq o(H)$ tenemos las desigualdades $[K:K_H] \geq o(G(K, K_H)) \geq o(H)$. Si pudiéramos demostrar que $[K:K_H] = o(H)$ se seguiría de inmediato que $o(H) = o(G(K, K_H))$, y como un subgrupo de $G(K, K_H)$ con el orden de $G(K, K_H)$ tendríamos $H = G(K, K_H)$. Luego solo nos queda, por demostrar que $[K:K_H] = o(H)$ para haber demostrado todo.

Según el teorema 5.p existe un $a \in K$ tal que $K = K_H(a)$; esta a debe, por tanto, satisfacer un polinomio irreducible sobre K_H de grado $m = [K:K_H]$ y ningún polinomio no trivial de grado más bajo (teorema 5.c). Sean los elementos de H los $\sigma_1, \sigma_2, \dots, \sigma_h$ donde σ_1 es la identidad de $G(K, F)$ y donde $h = o(H)$. Consideremos las funciones simétricas elementales de $a = \sigma_1(a), \sigma_2(a), \dots, \sigma_h(a)$, a saber:

$$\begin{aligned}\alpha_1 &= \sigma_1(a) + \sigma_2(a) + \cdots + \sigma_h(a) = \sum_{i=1}^h \sigma_i(a) \\ \alpha_2 &= \sum_{i < j} \sigma_i(a) \sigma_j(a) \\ &\vdots \\ \alpha_h &= \sigma_1(a) \sigma_2(a) \cdots \sigma_h(a).\end{aligned}$$

Cada α_i es invariante bajo cualquier $\alpha \in H$ (!Pruébese!). Así pues, por la definición de K_H , $\alpha_1, \alpha_2, \dots, \alpha_h$ son todos los elementos de K_H . Pero a (lo mismo que $\sigma_2(a), \dots, \sigma_h(a)$) es una raíz del polinomio $p(x) = (x - \sigma_1)(x - \sigma_2(a)) \cdots (x - \sigma_h(a)) = x^h - \alpha_1 x^{h-1} + \alpha_2 x^{h-2} + \cdots + (-1)^h \alpha_h$ que tiene todos sus coeficientes en K_H . Por la naturaleza de a esto obliga a que $h \geq m = [K:K_H]$, de donde $o(H) \geq [K:K_H]$. Como ya sabemos que $o(H) \leq [K:K_H]$ sabemos que $o(H) = [K:K_H]$, la conclusión deseada.

Cuando $H = G(K, F)$, por la normalidad de K sobre F , $K_H = F$; por consiguiente, para este caso particular tenemos el resultado $[K:F] = o(G(K, F))$.

Estamos acercándonos rápidamente al teorema central de la teoría de Galois. Lo que aún falta es la relación entre los campos de descomposición y las extensiones normales. Llenamos esta falla con el

TEOREMA 5.u. *K es una extensión normal de F si y sólo si K es el campo de descomposición de algún polinomio sobre F .*

Prueba. En una dirección la prueba nos recordará mucho la del teorema 5.t.

Supongamos que K es una extensión normal de F ; según el teorema 5.p, $K = F(a)$. Consideremos el polinomio $p(x) = (x - \sigma_1(a))(x - \sigma_2(a)) \dots (x - \sigma_n(a))$ sobre K , donde $\sigma_1, \sigma_2, \dots, \sigma_n$ son todos los elementos de $G(K, F)$. Desarrollando $p(x)$ vemos que $p(x) = x^n - \alpha_1 x^{n-1} + \alpha_2 x^{n-2} + \dots + (-1)^n \alpha_n$ donde $\alpha_1, \dots, \alpha_n$ son las funciones simétricas elementales en $a = \sigma_1(a), \sigma_2(a), \dots, \sigma_n(a)$. Pero entonces $\alpha_1, \dots, \alpha_n$ son, cada una, invariantes con respecto a toda $\sigma \in G(K, F)$, de donde, por la normalidad de K sobre F , todas deben estar en F . Por tanto, K descompone al polinomio $p(x) \in F[x]$ en un producto de factores lineales. Como a es una raíz de $p(x)$ y como a genera K sobre F , a no puede estar en ningún subcampo propio de K que contenga a F . Luego K es el campo de descomposición de $p(x)$ sobre F .

Ahora en la otra dirección; esto es un poco más complicado. Apartamos una pieza de la prueba en el

LEMA 5.9. *Sea K el campo de descomposición de $f(x)$ en $F[x]$ y sea $p(x)$ un factor irreducible de $f(x)$ en $F[x]$. Si las raíces de $p(x)$ son $\alpha_1, \dots, \alpha_r$, entonces para cada i existe un automorfismo σ_i en $G(K, F)$ tal que $\sigma_i(\alpha_1) = \alpha_i$.*

Prueba. Como cualquier raíz de $p(x)$ es una raíz de $f(x)$, tal raíz debe encontrarse en K . Sean α_1, α_i dos raíces cualesquiera de $p(x)$. De acuerdo con el teorema 5.i hay un isomorfismo τ de $F_1 = F(\alpha_1)$ sobre $F'_1 = F(\alpha_i)$ que lleva α_1 sobre α_i y deja todos los elementos de F fijos. Ahora bien, K es el campo de descomposición de $f(x)$ considerado como un polinomio sobre F_1 ; análogamente, K es el campo de descomposición de $f(x)$ considerado como un polinomio sobre F'_1 . Según el teorema 5.j hay un isomorfismo σ_i de K sobre K (luego un automorfismo de K) que coincide con τ sobre F_1 . Pero entonces $\sigma_i(\alpha_1) = \tau(\alpha_1) = \alpha_i$ y σ_i deja a todos los elementos de F fijos. Esto es, desde luego, exactamente lo que afirma el lema 5.9.

Volvemos ahora a nuestra tarea de completar la prueba del teorema 5.u. Supongamos que K es el campo de descomposición del polinomio $f(x)$ en $F[x]$. Queremos demostrar que K es normal sobre F . Procedemos por inducción sobre $[K:F]$, suponiendo que para cualquier par de campos K_1, F_1 con $[K_1:F_1]$ menor que $[K:F]$, siempre que K_1 es el campo de descomposición sobre F_1 de un polinomio en $F_1[x]$, entonces K_1 es normal sobre F_1 .

Si $f(x) \in F[x]$ se descompone en factores lineales sobre F , entonces $K = F$, que ciertamente es una extensión normal de F . Así pues, supongamos que $f(x)$ tiene un factor irreducible $p(x) \in F[x]$ de grado $r > 1$. Las r raíces distintas $\alpha_1, \alpha_2, \dots, \alpha_r$ de $p(x)$ todas se encuentran en K y K es el campo de descomposición de $f(x)$ considerado como un polinomio sobre $F(\alpha_1)$.

Como

$$[K : F(\alpha_1)] = \frac{[K : F]}{[F(\alpha_1) : F]} = \frac{n}{r} < n,$$

de acuerdo con nuestra hipótesis de inducción, K es una extensión normal de $F(\alpha_1)$.

Sea $\theta \in K$ fija para cualquier automorfismo $\sigma \in G(K, F)$; queremos demostrar que θ está en F . Ahora bien, cualquier automorfismo en $G(K, F(\alpha_1))$ deja, ciertamente, fija a F , de donde deja a θ fija; por la normalidad de K sobre $F(\alpha_1)$, esto implica que θ está en $F(\alpha_1)$. Así pues

$$1) \quad \theta = \lambda_0 + \lambda_1 \alpha_1 + \lambda_2 \alpha_1^2 + \cdots + \lambda_{r-1} \alpha_1^{r-1} \text{ donde } \lambda_0, \dots, \lambda_{r-1} \in F.$$

De conformidad con el lema 5.9 hay un automorfismo σ_i de K , $\sigma_i \in G(K, F)$, tal que $\sigma_i(\alpha_1) = \alpha_i$; como éste σ_i deja θ y toda λ_j fijas, aplicándolo a (1) obtenemos

$$2) \quad \theta = \lambda_0 + \lambda_1 \alpha_i + \lambda_2 \alpha_i^2 + \cdots + \lambda_{r-1} \alpha_i^{r-1} \text{ para } i = 1, 2, \dots, r.$$

Así pues, el polinomio $q(x) = \lambda_{r-1} x^{r-1} + \lambda_{r-2} x^{r-2} + \cdots + \lambda_1 x + (\lambda_0 - \theta)$ en $K[x]$, de grado cuando más $r-1$, tiene las r distintas raíces $\alpha_1, \alpha_2, \dots, \alpha_r$. Esto puede suceder solamente si todos los coeficientes son cero; en particular $\lambda_0 - \theta = 0$, de donde $\theta = \lambda_0$, luego está en F . Esto completa la inducción y prueba que K es una extensión normal de F . El teorema 5.u está completamente probado.

DEFINICIÓN. Sea $f(x)$ un polinomio en $F[x]$ y sea K su campo de descomposición sobre F . El *grupo de Galois* de $f(x)$ es el grupo $G(K, F)$ de todos los automorfismos de K que dejan fijos todos los elementos de F .

Nótese que el grupo de Galois de $f(x)$ puede considerarse como un grupo de permutaciones de sus raíces, pues si α es una raíz de $f(x)$ y si $\sigma \in G(K, F)$ entonces $\sigma(\alpha)$ es también una raíz de $f(x)$.

Llegamos ahora al resultado conocido como el *teorema fundamental de la teoría de Galois*. Establece una correspondencia biyectiva entre los subcampos del campo de descomposición de $f(x)$ y los subgrupos de su grupo de Galois. Además da un criterio para que un subcampo de una extensión normal sea él mismo una extensión normal de F . Este teorema fundamental se usará en la próxima sección para derivar condiciones para la solubilidad por radicales de las raíces de un polinomio.

TEOREMA 5.v. *Sea $f(x)$ un polinomio en $F[x]$, K su campo de descomposición sobre F y $G(K, F)$ su grupo de Galois. Para cualquier subcampo T de K que contiene a F sea $G(K, T) = \{\sigma \in G(K, F) \mid \sigma(t) = t \text{ para todo } t \in T\}$ y para cualquier subgrupo H de $G(K, F)$ sea $K_H = \{x \in K \mid \sigma(x) = x \text{ para todo } \sigma \in H\}$*

todo $\sigma \in H\}$. Entonces la asociación de T con $G(K, T)$ establece una correspondencia biyectiva del conjunto de subcampos de K que contienen a F sobre el conjunto de subgrupos de $G(K, F)$ tal que :

- 1) $T = K_{G(K, T)}$.
- 2) $H = G(K, K_H)$.
- 3) $[K:T] = o(G(K, T))$, $[T:F] = \text{índice de } G(K, T) \text{ en } G(K, F)$.
- 4) T es una extensión normal de F si y sólo si $G(K, T)$ es un subgrupo normal de $G(K, F)$.
- 5) Cuando T es una extensión normal de F , entonces $G(T, F)$ es isomorfo a $G(K, F)/G(K, T)$.

Prueba. Como K es el campo de descomposición de $f(x)$ sobre F es también el campo de descomposición de $f(x)$ sobre cualquier subcampo T que contenga a F ; por tanto, según el teorema 5.u, K es una extensión normal de T . Así pues, por la definición de normalidad, T es el campo fijo de $G(K, T)$, es decir, $T = K_{G(K, T)}$, probando así (1).

Como K es una extensión normal de F , de acuerdo con el teorema 5.t, dado un subgrupo H de $G(K, F)$, entonces $H = G(K, K_H)$ que es lo que se afirma en la parte (2). Además, esto demuestra que cualquier subgrupo de $G(K, F)$ se presenta en la forma $G(K, T)$, de donde la asociación de T con $G(K, T)$ transforma el conjunto de todos los subcampos de K que contienen a F sobre el conjunto de todos los subgrupos de $G(K, F)$. Que es inyectiva es claro, pues, si $G(K, T_1) = G(K, T_2)$, entonces, por la parte (1), $T_1 = K_{G(K, T_1)} = K_{G(K, T_2)} = T_2$.

Como K es normal sobre T , tenemos, al aplicar de nuevo el teorema 5.t, $[K:T] = o(G(K, T))$; pero entonces, $o(G(K, F)) = [K:F] = [K:T][T:F] = o(G(K, T))[T:F]$, de donde

$$[T:F] = \frac{o(G(K, F))}{o(G(K, T))} = \text{índice de } G(K, T)$$

en $G(K, F)$. Y ésta es la parte (3).

Las únicas partes que quedan por probar son las que conciernen a la normalidad. Haremos primero la siguiente observación. T es una extensión normal de F si y sólo si para cada $\sigma \in G(K, F)$, $\sigma(T) \subset T$. ¿Por qué? Sabemos por el teorema 5.p que $T = F(a)$; así pues, si $\sigma(T) \subset T$ entonces $\sigma(a) \in T$ para todo $\sigma \in G(K, F)$. Pero como vimos en la prueba del teorema 5.u esto implica que T es el campo de descomposición de $p(x) = \prod_{\sigma \in G(K, F)} (x - \sigma(a))$

que tiene coeficientes en F . Como campo de descomposición T , por el teorema 5.u, es una extensión normal de F . Recíprocamente, si T es una extensión normal de F , entonces $T = F(a)$, donde el polinomio mínimo de a , $p(x)$, sobre F tiene todas sus raíces en T (teorema 5.u). Pero para cualquier $\sigma \in G(K, F)$, $\sigma(a)$ es también una raíz de $p(x)$, de donde $\sigma(a)$ debe estar en T .

Como T está generado por σ sobre F tenemos que $\sigma(T) \subset T$, para todo $\sigma \in G(K, F)$.

Así pues, T es una extensión normal de F si y sólo si para todo $\sigma \in G(K, F)$, $\tau \in G(K, T)$ y $t \in T$, $\sigma(t) \in T$ y, por tanto, $\tau(\sigma(t)) = \sigma(t)$; es decir, si y sólo si $\sigma^{-1}\tau\sigma(t) = t$. Pero esto dice que T es normal sobre F si y sólo si $\sigma^{-1}G(K, T) \subset G(K, T)$ para todo $\sigma \in G(K, F)$. Siendo esta última condición precisamente la que define $G(K, T)$ como un subgrupo normal de $G(K, F)$, vemos que la parte (4) queda probada.

Finalmente, si T es normal sobre F , dado $\sigma \in G(K, F)$, como $\sigma(t) \in T$, σ induce un automorfismo σ_* de T definido por $\sigma_*(t) = \sigma(t)$ para todo $t \in T$. Como σ_* deja a todo elemento de F fijo, σ_* debe estar en $G(T, F)$. Además, como es evidente, para cualquier $\sigma, \psi \in G(K, F)$, $(\sigma\psi)_* = \sigma_*\psi_*$ de donde la aplicación de $G(K, F)$ en $G(T, F)$ definida por $\sigma \rightarrow \sigma_*$ es un homomorfismo de $G(K, F)$ en $G(T, F)$. ¿Qué es el núcleo de este homomorfismo? Consiste en todos los elementos σ en $G(K, F)$ tal que σ_* es la aplicación identidad sobre T . Es decir, el núcleo es el conjunto de todos los $\sigma_* \in G(K, F)$ tales que $t = \sigma_*(t) = \sigma(t)$; por la misma definición, tenemos que el núcleo es exactamente $G(K, T)$. La imagen de $G(K, F)$ en $G(T, F)$, según el teorema 2.d, es isomorfa a $G(K, F)/G(K, T)$, cuyo orden es $o(G(K, F))/o(G(K, T)) = [T:F]$ (por parte 3) = $o(G(T, F))$ (como establece el teorema 5.t.). Así pues, la imagen de $G(K, F)$ en $G(T, F)$ es todo $G(T, F)$ y, por tanto, $G(T, F)$ es isomorfo a $G(K, F)/G(K, T)$. Esto termina la prueba de la parte (5) y con ello completamos la prueba del teorema 5.v.

Problemas

1. Si K es un campo y S un conjunto de homomorfismos de K , demuestre que el campo fijo de S y el de \bar{S} (el subgrupo del grupo de todos los automorfismos de K generados por S) son idénticos.

2. Pruébese el lema 5.8.

3. Usando el criterio de Eisenstein, pruébese que $x^4 + x^3 + x^2 + x + 1$ es irreducible sobre el campo de los números racionales.

4. En el ejemplo 3 del texto, pruébese que cada una de las aplicaciones σ_i que allí se definieron es un automorfismo de $F_0(\omega)$.

5. En el ejemplo 3, pruébese que el campo fijo de $F_0(\omega)$ bajo $\sigma_1, \sigma_2, \sigma_3$ y σ_4 es precisamente F_0 .

6. Pruébese directamente que cualquier automorfismo de K debe dejar fijos todos los racionales.

*7. Pruébese que un polinomio simétrico en x_1, \dots, x_n es un polinomio en las funciones simétricas elementales en x_1, \dots, x_n .

8. Exprésense los siguientes como polinomios en las funciones simétricas elementales en x_1, x_2 y x_3 .

- a) $x_1^2 + x_2^2 + x_3^2$.
- b) $x_1^3 + x_2^3 + x_3^3$.
- c) $(x_1 - x_2)^2(x_1 - x_3)^2(x_2 - x_3)^2$.

9. Si $\alpha_1, \alpha_2, \alpha_3$ son las raíces del polinomio cúbico $x^3 + 7x^2 - 8x + 3$, encuéntrese el polinomio cúbico cuyas raíces son :

- a) $\alpha_1^2, \alpha_2^2, \alpha_3^2$.
- b) $\frac{1}{\alpha_1}, \frac{1}{\alpha_2}, \frac{1}{\alpha_3}$.
- c) $\alpha_1^3, \alpha_2^3, \alpha_3^3$.

***10.** Pruébense las *identidades de Newton*, es decir, si $\alpha_1, \alpha_2, \dots, \alpha_n$ son las raíces de $f(x) = x^n + a_1x^{n-1} + a_2x^{n-2} + \dots + a_n$ y si $s_k = \alpha_1^k + \alpha_2^k + \dots + \alpha_n^k$, entonces

- a) $s_k + a_1s_{k-1} + a_2s_{k-2} + \dots + a_{k-1}s_1 + ka_k = 0$ si $k = 1, 2, \dots, n$.
- b) $s_k + a_1s_{k-1} + \dots + a_ns_{k-n} = 0$ para $k > n$.
- c) Para $n = 5$, aplíquese la parte (a) para determinar s_2, s_3, s_4 y s_5 .

11. Pruébese que las funciones simétricas elementales en x_1, \dots, x_n son, ciertamente, funciones simétricas en x_1, \dots, x_n .

12. Si $p(x) = x^n - 1$, pruébese que el grupo de Galois de $p(x)$ sobre el campo de los números racionales es abeliano.

El número complejo ω es una *raíz n -ésima primitiva de la unidad* si $\omega^n = 1$ pero $\omega^m \neq 1$ para $0 < m < n$. F_0 denotará el campo de los números racionales.

- 13.** a) Pruébese que hay $\phi(n)$ raíces n -ésimas primitivas de la unidad donde $\phi(n)$ es la función ϕ de Euler.
- b) Si ω es una raíz n -ésima primitiva de la unidad, pruébese que $F_0(\omega)$ es el campo de descomposición de $x^n - 1$ sobre F_0 (y por tanto es una extensión normal de F_0).
- c) Si $\omega_1, \dots, \omega_{\phi(n)}$ son las $\phi(n)$ raíces n -ésimas primitivas de la unidad, pruébese que cualquier automorfismo de $F_0(\omega_1)$ lleva ω_1 en algún ω_i .
- d) Pruébese que $[F_0(\omega_1):F_0] \leq \phi(n)$.

- 14.** La notación es como la del problema 13.

- *a) Pruébese que hay un automorfismo σ_i de $F_0(\omega_1)$ que lleva ω_1 en ω_i .
- b) Pruébese que el polinomio $p_n(x) = (x - \omega_1)(x - \omega_2) \dots (x - \omega_{\phi(n)})$

tiene coeficientes racionales. El polinomio $p_n(x)$ se llama el n -ésimo *polinomio ciclotímico*.

*c) Pruébese que en realidad los coeficientes de $p_n(x)$ son enteros.

15. Úsense los resultados de los problemas 13 y 14 para probar que $p_n(x)$ es irreducible sobre F_0 para todo $n \geq 1$.

16. Para $n = 3, 4, 6$ y 8 , calcúlese $p_n(x)$ explícitamente, demuéstrese que tiene coeficientes enteros y pruébese directamente que es irreducible sobre F_0 .

17. a) Pruébese que el grupo de Galois de $x^3 - 2$ sobre F_0 es isomorfo a S_3 , el grupo simétrico de grado 3.

b) Encuéntrese el campo de descomposición K de $x^3 - 2$ sobre F_0 .

c) Para cada subgrupo H de S_3 encuéntrese K_H y compruébese que la correspondencia da en el teorema 5.v.

d) Encuéntrese una extensión normal en K de grado 2 sobre F_0 .

18. Si el campo F contiene una raíz n -ésima primitiva de la unidad pruébese que el grupo de Galois de $x^n - a$, para $a \in F$, es abeliano.

7. SOLUBILIDAD POR RADICALES

Dado el polinomio específico $x^2 + 3x + 4$ sobre el campo de los números racionales F_0 , de acuerdo con la fórmula cuadrática para sus raíces, sabemos que estas son $(-3 \pm \sqrt{-7})/2$; así pues, el campo $F_0(\sqrt{7}i)$ es el campo de descomposición de $x^2 + 3x + 4$ sobre F_0 . Hay, por consiguiente, un elemento $\gamma = -7$ en F_0 tal que el campo extensión $F_0(\omega)$ donde $\omega^2 = \gamma$ es tal que contiene todas las raíces de $x^2 + 3x + 4$.

Desde un punto de vista ligeramente diferente, dado el polinomio cuadrático *general* $p(x) = x^2 + a_1x + a_2$ sobre F , podemos considerarlo como un polinomio *particular* sobre el campo $F(a_1, a_2)$ de las funciones racionales en las dos variables a_1 y a_2 sobre F ; en la extensión obtenida por la adjunción de ω a $F(a_1, a_2)$ donde $\omega^2 = a_1^2 - 4a_2 \in F(a_1, a_2)$ encontramos todas las raíces de $p(x)$. Hay una fórmula que expresa todas las raíces de $p(x)$ en términos de a_1 , a_2 y raíces cuadradas de funciones racionales de a_1 y a_2 .

Para una ecuación cúbica la situación es muy semejante; dada la ecuación general cúbica $p(x) = x^3 + a_1x^2 + a_2x + a_3$ puede darse una fórmula explícita, incluyendo combinaciones de raíces cuadradas y raíces cúbicas de funciones racionales en a_1 , a_2 y a_3 . Aunque en forma algo complicada las fórmulas de Cardano nos las dan explícitamente: Sean $p = a_2 - (a_1^2/3)$ y

$$q = \frac{2a_1^3}{27} - \frac{a_1a_2}{3} + a_3$$

y sea

$$P = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{p^3}{27} + \frac{q^2}{4}}}$$

y

$$Q = \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{p^3}{27} + \frac{q^2}{4}}}$$

(con raíces cúbicas propiamente escogidas); entonces las raíces de $p(x)$ son $P+Q-(\alpha_1/3)$, $\omega P+\omega^2 Q-(\alpha_1/3)$ y $\omega^2 P+\omega Q-(\alpha_1/3)$ donde $\omega \neq 1$ es una raíz cúbica de 1. Estas fórmulas solo nos sirven para ilustrar que, según la adjunción de una cierta raíz cuadrada y luego una raíz cúbica a $F(\alpha_1, \alpha_2, \alpha_3)$ llegamos a un campo en el que $p(x)$ tiene sus raíces.

Para polinomios de cuarto grado, que no daremos explícitamente, mediante el uso de operaciones racionales y raíces cuadradas podemos reducir el problema al de resolver cierta raíz cúbica, de modo que también aquí puede darse una fórmula que exprese las raíces en términos de combinaciones de radicales de funciones racionales de los coeficientes.

Para polinomios de grado quinto o más alto, no puede darse tal fórmula universal radical, pues demostraremos que es imposible expresar sus raíces, en general, de este modo.

Dado un campo F y un polinomio $p(x) \in F[x]$ decimos que $p(x)$ es soluble por radicales sobre F si podemos encontrar una sucesión finita de campos $F_1 = F(\omega_1)$, $F_2 = F_1(\omega_2)$, ..., $F_k = F_{k-1}(\omega_k)$ tal que $\omega_1^{r_1} \in F$, $\omega_2^{r_2} \in F_1$, ..., $\omega_k^{r_k} \in F_{k-1}$ tal que las raíces de $p(x)$ se encuentren todas en F_k .

Si K es el campo de descomposición de $p(x)$ sobre F , entonces $p(x)$ es soluble por radicales sobre F si podemos encontrar una sucesión de campos como anteriormente tales que $K \subset F_k$. Una observación importante y que usaremos posteriormente en la prueba del teorema 5.x, es que si puede encontrarse un tal F_k , podemos, sin pérdida de generalidad, suponer que sea una extensión normal de F ; dejamos la prueba de esta afirmación como problema (problema 1).

Por polinomio general de grado n sobre F , $p(x) = x^n + a_1 x^{n-1} + \dots + a_n$ entendemos lo siguiente: Sea $F(a_1, \dots, a_n)$ el campo de funciones racionales en las n variables a_1, \dots, a_n sobre F , y considérese el polinomio particular $p(x) = x^n + a_1 x^{n-1} + \dots + a_n$ sobre el campo $F(a_1, \dots, a_n)$. Decimos que es soluble por radicales si es soluble por radicales sobre $F(a_1, \dots, a_n)$. Esto expresa realmente la idea intuitiva de "encontrar una fórmula" para las raíces de $p(x)$ que implique combinaciones de raíces m -ésimas para varias m , de funciones racionales en a_1, a_2, \dots, a_n . Para $n = 2, 3$ y 4 señalamos que esto puede hacerse siempre. Para $n \geq 5$; Abel probó que no puede hacerse. Pero esto no excluye la posibilidad de que un polinomio dado sobre F pueda resolverse por radicales. En realidad, daremos un criterio

para esto en términos del grupo de Galois del polinomio. Pero primero debemos desarrollar unos pocos resultados de teoría pura de grupos. Algunos de estos aparecieron como problemas al final del capítulo 2; pero, sin embargo, los haremos aquí oficialmente.

DEFINICIÓN. Un grupo G se dice que es *soluble* si podemos encontrar una cadena finita de subgrupos $G = N_0 \supset N_1 \supset N_2 \supset \dots \supset N_k = (e)$ donde cada N_i sea un subgrupo normal de N_{i-1} y tal que cada grupo factor N_{i-1}/N_i sea abeliano.

Todo grupo abeliano es soluble, pues simplemente se toma $N_0 = G$ y $N_1 = (e)$ para satisfacer la anterior definición. El grupo simétrico de grado 3, S_3 , es soluble. En efecto, si tomamos $N_1 = \{e, (1, 2, 3), (1, 3, 2)\}$, N_1 es un subgrupo normal de S_3 y S_3/N_1 y $N_1/(e)$ son, ambos, abelianos por ser de órdenes 2 y 3, respectivamente. Se puede demostrar que S_4 es soluble (problema 3). Para $n \geq 5$ demostraremos en el teorema 5.w que S_n no es soluble.

Busquemos una descripción alternativa para la solubilidad. Dado el grupo G y los elementos a y b de G , entonces el *comutador* de a y b es el elemento $a^{-1}b^{-1}ab$. El *subgrupo comutador*, G' , de G es el subgrupo de G generado por todos los comutadores de G . (No es necesariamente cierto que el conjunto de los comutadores mismo forme un subgrupo de G .) Vimos en un ejercicio anterior que G' es un subgrupo normal de G . Además, el grupo G/G' es abeliano, pues dados dos elementos cualesquiera en él, aG' , bG' , con $a, b \in G$, entonces

$$(aG')(bG') = abG' = ba(b^{-1}a^{-1}ab)G' =$$

(como $a^{-1}b^{-1}ab \in G' = baG' = (bG')(aG')$). Por otra parte, si M es un subgrupo normal de G tal que G/M es abeliano, entonces $M \supset G'$, pues dados $a, b \in G$, entonces $(aM)(bM) = (bM)(aM)$ de donde deducimos $abM = baM$, luego $a^{-1}b^{-1}abM = M$ y, por tanto, $a^{-1}b^{-1}ab \in M$. Como M contiene todos los comutadores, contiene al grupo que estos generan, es decir, a G' .

G' es un grupo por derecho propio, así que podemos hablar de su grupo comutador $G^{(2)} = (G')$. Este es el subgrupo de G generado por todos los elementos $(a')^{-1}(b')^{-1}a'b'$ donde $a', b' \in G'$. Es fácil probar que no solo es $G^{(2)}$ un subgrupo normal de G' , sino también un subgrupo normal de G (problema 4). Continuando de esta forma definimos los subgrupos comutadores más altos $G^{(m)}$ por $G^{(m)} = (G^{(m-1)})'$. Todo $G^{(m)}$ es un subgrupo normal de G (problema 4) y $G^{(m-1)}/G^{(m)}$ es un grupo abeliano.

En términos de estos subgrupos comutadores más altos de G , tenemos un criterio sucinto de solubilidad, a saber,

LEMA 5.10. G es soluble si y sólo si $G^{(k)} = (e)$ para algún entero k .

Prueba. Si $G^{(k)} = (e)$ sea $N_0 = G$, $N_1 = G'$, $N_2 = G^{(2)}$, ..., $N_k = G^{(k)} = (e)$. Tenemos $G = N_0 \supset N_1 \supset N_2 \supset \dots \supset N_k = (e)$; con cada N_i por normal en G , ciertamente, también normal en N_{i-1} . Finalmente,

$$\frac{N_{i-1}}{N_i} = \frac{G^{(i-1)}}{G^{(i)}} = \frac{G^{(i-1)}}{(G^{(i-1)})'}$$

Luego es abeliano. Así pues, según la definición de solubilidad de un grupo, G es un grupo soluble.

Recíprocamente, si G es un grupo soluble, hay una cadena $G = N_0 \supset N_1 \supset N_2 \supset \dots \supset N_k = (e)$ donde cada N_i es normal en N_{i-1} y donde N_{i-1}/N_i es abeliano. Pero, entonces, el subgrupo comutador N'_{i-1} de N_{i-1} debe estar contenido en N_i . Así pues, $N_1 \supset N'_1 = G'$, $N_2 \supset N'_2 \supset (G')' = G^{(2)}$, $N_3 \supset N'_3 \supset (G^{(2)})' = G^{(3)}$, ..., $N_i \supset G^{(i)}$, $(e) = N_k \supset G^{(k)}$. De donde resulta que $G^{(k)} = (e)$.

COROLARIO. Si G es un grupo soluble y si \bar{G} es una imagen homomórfica de G , entonces \bar{G} es soluble.

Prueba. Como \bar{G} es una imagen homomórfica de G , es inmediato que $(\bar{G})^{(k)}$ es la imagen de $G^{(k)}$. Como $G^{(k)} = (e)$ para alguna k , $(\bar{G})^{(k)} = (e)$ para la misma k , de donde, de acuerdo con el lema, \bar{G} es soluble.

El siguiente lema es clave en la prueba de la familia infinita de grupos S_n , con $n \geq 5$, no es soluble; aquí S_n es el grupo simétrico de grado n .

LEMA 5.11. Sea $G = S_n$ donde $n \geq 5$; entonces $G^{(k)}$ para $k = 1, 2, \dots$, contiene todo ciclo de orden 3 de S_n .

Prueba. Observemos primero que para un grupo arbitrario G , si N es un subgrupo normal de G entonces N' debe también ser un subgrupo normal de G (problema 5).

Afirmamos que si N es un subgrupo normal de $G = S_n$ donde $n \geq 5$, que contiene todo ciclo de orden 3 en S_n , entonces N' debe también contener todo ciclo de orden 3. Pues supongamos $a = (1, 2, 3)$, $b = (1, 4, 5)$ de N (estamos aquí usando que $n \geq 5$); entonces $a^{-1}b^{-1}ab = (3, 2, 1)(5, 4, 1)(1, 2, 3)(1, 4, 5) = (1, 4, 2)$, como comutador de elementos de N debe estar en N' . Como N' es un subgrupo normal de G , para cualquier $\pi \in S_n$, $\pi^{-1}(1, 4, 2)\pi$ debe estar también en N' . Escojamos π en S_n tal que $\pi(1) = i_1$, $\pi(4) = i_2$ y $\pi(2) = i_3$, donde $i_1, i_2 \in i_3$ son cualesquiera tres enteros distintos en el rango de 1 a n ; entonces $\pi^{-1}(1, 4, 2)\pi = (i_1, i_2, i_3)$ está en N' . Luego N' contiene todos los ciclos de orden 3.

Haciendo $N = G$, que es ciertamente normal en G y contiene todos los ciclos de orden tres, tenemos que G' contiene todos los ciclos de orden 3;

como G' es normal en G , $G^{(2)}$ contiene todos los ciclos de orden 3; como $G^{(2)}$ es normal en G , $G^{(3)}$ contiene todos los ciclos de orden 3. Continuando de esta forma llegamos a la conclusión de que $G^{(k)}$ contiene todos los ciclos de orden 3 para cualquier k .

Una consecuencia directa de este lema es el resultado interesante para la teoría de grupos de que

TEOREMA 5.w. S_n no es soluble para $n \geq 5$.

Prueba. Si $G = S_n$, según el lema 5.11, $G^{(k)}$ contiene todos los ciclos de orden 3 de S_n para todo k . Por tanto, $G^{(k)} \neq (e)$ para toda k , de donde de acuerdo con el lema 5.10 G no puede ser soluble.

Interrelacionamos ahora la solubilidad por radicales de $p(x)$ con la solubilidad como grupo del grupo de Galois de $p(x)$. La misma terminología es altamente sugestiva de que una tal relación existe. Pero primero necesitamos un resultado acerca del grupo de Galois de un cierto tipo de polinomio.

LEMA 5.12. Supongamos que el campo F tenga todas las raíces n -ésimas de la unidad (para un cierto determinado n) y supongamos que $a \neq 0$ está en F . Sea $x^n - a \in F[x]$ y sea K su campo de descomposición sobre F . Entonces:

- 1) $K = F(u)$, donde u es cualquier raíz de $x^n - a$.
- 2) El grupo de Galois de $x^n - a$ sobre F es abeliano.

Prueba. Como F contiene a todas las raíces n -ésimas de la unidad, contiene $\xi = e^{2\pi i/n}$; nótese que $\xi^n = 1$ pero $\xi^m \neq 1$ para $0 < m < n$.

Si $u \in K$ es cualquier raíz de $x^n - a$, entonces $u, \xi u, \xi^2 u, \dots, \xi^{n-1} u$ son todas las raíces de $x^n - a$. Que son raíces, es evidente; que son distintas se sigue de que si $\xi^i u = \xi^j u$ con $0 \leq i < j < n$, entonces como $u \neq 0$ y $(\xi^i - \xi^j)u = 0$, debemos tener $\xi^i = \xi^j$, lo que es imposible ya que $\xi^{j-i} = 1$ con $0 < j-i < n$. Como $\xi \in F$, todos los $u, \xi u, \dots, \xi^{n-1} u$ estén en $F(u)$, luego $F(u)$ descompone $x^n - a$; como ningún subcampo propio de $F(u)$ que contenga a F contiene también a u , ningún subcampo propio de $F(u)$ puede descomponer a $x^n - a$. Así pues, $F(u)$ es el campo de descomposición de $x^n - a$, y hemos probado que $K = F(u)$.

Si σ, τ son dos elementos cualesquiera de $x^n - a$, es decir, si σ, τ son automorfismos de $K = F(u)$ que dejan todos los elementos de F fijos, entonces como tanto $\sigma(u)$ como $\tau(u)$ son raíces de $x^n - a$, $\sigma(u) = \xi^i u$ y $\tau(u) = \xi^j u$ para algunas i y j . Así pues, $\sigma\tau(u) = \sigma(\xi^j u) = \xi^j \sigma(u)$ (ya que $\xi^j \in F$) = $\xi^i \xi^j u = \xi^{i+j} u$; análogamente, $\tau\sigma(u) = \xi^{i+j} u$. Por tanto, $\sigma\tau$ y $\tau\sigma$ coinciden sobre u y sobre F , de donde, en todo $K = F(u)$. Pero entonces $\sigma\tau = \tau\sigma$, de donde el grupo de Galois es abeliano.

Nótese que el lema dice que cuando F tiene todas las raíces n -ésimas de la unidad, entonces, adjuntando una raíz de $x^n - a$ a F , donde $a \in F$, tenemos todo el campo de descomposición de $x^n - a$, luego éste debe ser una extensión normal de F .

Suponemos para el resto de la sección que F es un campo que contiene todas las raíces n -ésimas de la unidad para todo entero n . Tenemos

TEOREMA 5.x. *Si $p(x) \in F[x]$ es soluble por radicales sobre F , entonces el grupo de Galois sobre F de $p(x)$ es un grupo soluble.*

Prueba. Sea K el campo de descomposición de $p(x)$ sobre F ; el grupo de Galois de $p(x)$ sobre F es $G(K, F)$. Como $p(x)$ es soluble por radicales existe una sucesión de campos

$$F \subset F_1 = F(\omega_1) \subset F_2 = F_1(\omega_2) \subset \dots \subset F_k = F_{k-1}(\omega_k),$$

donde $\omega_1 \in F$, $\omega_2 \in F_1$, ..., $\omega_k \in F_{k-1}$ y donde $K \subset F_k$. Como dijimos podemos suponer, sin pérdida de generalidad, que F_k es una extensión normal de F . Como extensión normal de F , F_k es también una extensión normal de cualquier campo intermedio, de donde F_k es una extensión normal de cada una de las F_i .

Según el lema 5.12 toda F_i es una extensión normal de F_{i-1} y como F_k es normal sobre F_{i-1} , de acuerdo con el teorema 5.v, $G(F_k, F_i)$ es un subgrupo normal en $G(F_k, F_{i-1})$. Consideremos la cadena:

$$1) \quad G(F_k, F) \supset G(F_k, F_1) \supset G(F_k, F_2) \supset \dots \supset G(F_k, F_{k-1}) \supset (e).$$

Como acabamos de hacer notar, cada grupo en la cadena es un subgrupo normal en el que le precede. Como F_i es una extensión normal de F_{i-1} , de acuerdo con el teorema fundamental de la teoría de Galois (teorema 5.v) el grupo de F_i sobre F_{i-1} , $G(F_i, F_{i-1})$ es isomorfo a $G(F_k, F_{i-1})/G(F_k, F_i)$. Pero según el lema 5.12, $G(F_i, F_{i-1})$ es un grupo abeliano. Luego todos los grupos cociente $G(F_k, F_{i-1})/G(F_k, F_i)$ de la cadena (1) es abeliano.

¡Luego el grupo $G(F_k, F)$ es soluble! Como $K \subset F_k$ es una extensión normal de F (por ser un campo de descomposición), según el teorema 5.v, $G(F_k, K)$ es un subgrupo normal de $G(F_k, F)$ y $G(K, F)$ es isomorfo a $G(F_k, F)/G(F_k, K)$. Así pues, $G(K, F)$ es una imagen homomórfica de $G(F_k, F)$ que es un grupo soluble; por el corolario del lema 5.10, el mismo $G(K, F)$ debe entonces ser un grupo soluble. Como $G(K, F)$ es el grupo de Galois de $p(x)$ sobre F , el teorema ha sido probado.

Hacemos dos observaciones sin prueba.

- 1) El recíproco del teorema 5.x es también cierto, es decir, si el grupo de Galois de $p(x)$ sobre F es soluble, entonces $p(x)$ es soluble por radicales sobre F .
- 2) El teorema 5.x y su recíproco son ciertos incluso si F no contiene raíces de la unidad.

Recordando lo que se entiende por polinomio general de grado n sobre F , $p(x) = x^n + a_1 x^{n-1} + \dots + a_n$, y lo que se entiende por soluble por radicales, cerramos el capítulo con el gran teorema clásico de Abel

TEOREMA 5.Y. *El polinomio general de grado $n \geq 5$ no es soluble por radicales.*

Prueba. En el teorema 5.s demostramos que si $F(a_1, \dots, a_n)$ es el campo de las funciones racionales en las n variables a_1, \dots, a_n , entonces el grupo de Galois del polinomio $p(t) = t^n + a_1 t^{n-1} + \dots + a_n$ sobre $F(a_1, \dots, a_n)$ era S_n , el grupo simétrico de grado n . De acuerdo con el teorema 5.w S_n no es un grupo soluble cuando $n \geq 5$, así pues, según el teorema 5.x $p(t)$ no es soluble por radicales sobre $F(a_1, \dots, a_n)$ cuando $n \geq 5$.

Problemas

*1. Si $p(x)$ es soluble por radicales sobre F , pruébese que puede encontrarse una sucesión de campos

$$F \subset F_1 = F(\omega_1) \subset F_2 = F_1(\omega_2) \subset \dots \subset F_k = F_{k-1}(\omega_k)$$

donde $\omega_1 \in F$, $\omega_2 \in F_1, \dots, \omega_k \in F_{k-1}$, con F_k conteniendo todas las raíces de $p(x)$ tal que F_k es *normal* sobre F .

2. Pruébese que un subgrupo de un grupo soluble es soluble.
3. Pruébese que S_4 es un grupo soluble.
4. Si G es un grupo, pruébese que todos los $G^{(k)}$ son subgrupos normales de G .
5. Si N es un subgrupo normal de G , pruébese que N' debe también ser un subgrupo normal de G .
6. Pruébese que el grupo alternante (el grupo de las permutaciones pares en S_n) A_n , tiene subgrupos normales no triviales para $n \geq 5$.

Lecturas supplementarias

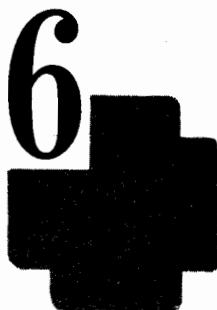
- ARTIN, E., *Galois Theory*, segunda edición. Notre Dame Mathematical Lectures, número 2.
- POLLARD, H., *Theory of Algebraic Numbers*, Carus Monographs, número 9. John Wiley and Sons, Inc., Nueva York, 1950.
- VAN DER WAERDEN, B. L., *Modern Algebra*, vol. I. Ungar Publishing Company, Nueva York, 1949.
- WEISNER, L., *Theory of Equations*. The Macmillan Company, Nueva York, 1938.

- SIEGEL, C. L., *Transcendental Numbers*, Annals of Mathematical Studies, número 16. Princeton University Press, Princeton, Nueva Jersey, 1949.
- NIVEN, I., *Irrational Numbers*, Carus Monographs, número 11. John Wiley & Sons, Inc., Nueva York, 1956.

Tópicos para discusión en clase

- NIVEN, I., "A simple proof of the irrationality of π ", *Bulletin of the American Mathematical Society*, vol. 53 (1947), pág. 509.

CAPITULO



Transformaciones lineales

EN EL capítulo 4 definimos: para dos espacios vectoriales cualesquiera V y W sobre el mismo campo F , el conjunto $\text{Hom}(V, W)$ de todos los homomorfismos de espacio vectorial de V en W . En realidad, introdujimos en $\text{Hom}(V, W)$ las operaciones de adición y multiplicación por escalares (elementos de F) en tal forma que $\text{Hom}(V, W)$ mismo se nos hizo un espacio vectorial.

De mucho mayor interés es el caso especial $V = W$, pues aquí, aparte de las operaciones de espacio vectorial, podemos, además, introducir una multiplicación con la cual $\text{Hom}(V, W)$ se hace un anillo. Provisto de esta doble naturaleza —la de espacio vectorial y la de anillo— $\text{Hom}(V, W)$ adquiere una estructura sumamente rica. Son, esta estructura y sus con-

secuencias, las que imparten tanta vida y brillo al tema y las que justifican más plenamente la creación del concepto abstracto de espacio vectorial.

Nuestro mayor interés se concentrará en $\text{Hom}(V, V)$ con un V que no será un espacio vectorial arbitrario, sino un espacio vectorial de dimensión finita sobre un campo F . La dimensionalidad finita de V impone sobre $\text{Hom}(V, V)$ que cada uno de sus elementos satisfaga un polinomio sobre F . Este hecho, quizás más que cualquier otro, nos facilita el camino para entrar en el conocimiento de $\text{Hom}(V, V)$ y nos permite explorar profunda y efectivamente su estructura.

La materia que aquí vamos a considerar se denomina con frecuencia *álgebra lineal*. Contiene a la isomorfa *teoría de matrices*. La afirmación de que sus resultados son de uso cotidiano en todos los aspectos de las matemáticas (y de lo que no son matemáticas) no es de ningún modo exagerada.

Un mito popular afirma que los matemáticos gozan con la falta de aplicaciones de su disciplina y se molestan cuando alguno de sus resultados es “manchado,” por el uso en el mundo exterior. ¡Es una rematada tontería! Es cierto que un matemático no basa sus juicios de valor en la aplicabilidad de un resultado dado fuera de las propias matemáticas sino que más bien los basa en un criterio matemático algo intrínseco y a veces intangible. Pero también es cierto que lo recíproco es falso —la utilidad de un resultado nunca ha disminuido su valor matemático. Un caso perfecto para ilustrar el punto es el del álgebra lineal; tenemos en ella matemáticas reales, interesantes y excitantes por sí y, sin embargo, es probablemente la parte de las matemáticas que encuentra las aplicaciones más amplias —en la física, la química, la economía, en realidad en casi toda ciencia o pseudociencia.

1. EL ÁLGEBRA DE LAS TRANSFORMACIONES LINEALES

Sea V un espacio vectorial sobre un campo F y $\text{Hom}(V, V)$, como antes, el conjunto de todos los homomorfismos de espacio vectorial de V en sí mismo. En la sección 3 del capítulo 4 vimos que $\text{Hom}(V, V)$ forma un espacio vectorial sobre F donde para $T_1, T_2 \in \text{Hom}(V, V)$, $T_1 + T_2$ queda definido por $v(T_1 + T_2) = vT_1 + vT_2$ para todo $v \in V$, y donde para $\alpha \in F$, αT_1 está definido por $v(\alpha T_1) = \alpha(vT_1)$.

Para $T_1, T_2 \in \text{Hom}(V, V)$, como $vT_1 \in V$ para cualquier $v \in V$, $(vT_1)T_2$ tiene sentido. Como hemos hecho para aplicaciones de un conjunto en sí mismo, definimos $T_1 T_2$ por $v(T_1 T_2) = (vT_1)T_2$ para cualquier $v \in V$. Afirmando ahora que $T_1 T_2 \in \text{Hom}(V, V)$. Para probarlo, debemos demostrar que para cualesquiera $\alpha, \beta \in F$ y cualesquiera $u, v \in V$, $(\alpha u + \beta v)(T_1 T_2) = \alpha(u(T_1 T_2)) + \beta(v(T_1 T_2))$. Calculamos:

$$\begin{aligned} (\alpha u + \beta v)(T_1 T_2) &= ((\alpha u + \beta v)T_1)T_2 \\ &= (\alpha(uT_1) + \beta(vT_1))T_2 \\ &= \alpha(uT_1)T_2 + \beta(vT_1)T_2 \\ &= \alpha(u(T_1 T_2)) + \beta(v(T_1 T_2)). \end{aligned}$$

Dejamos como ejercicio la prueba de las siguientes propiedades de este producto en $\text{Hom}(V, V)$:

- 1) $(T_1 + T_2)T_3 = T_1T_3 + T_2T_3$
- 2) $T_3(T_1 + T_2) = T_3T_1 + T_3T_2$
- 3) $T_1(T_2T_3) = (T_1T_2)T_3$
- 4) $\alpha(T_1T_2) = (\alpha T_1)T_2 = T_1(\alpha T_2)$

para cualesquiera $T_1, T_2, T_3 \in \text{Hom}(V, V)$ y toda $\alpha \in F$.

Nótese que las propiedades (1), (2) y (3) anteriores son exactamente las que se necesitan para hacer de $\text{Hom}(V, V)$ un anillo asociativo. La propiedad (4) entremezcla el carácter de $\text{Hom}(V, V)$ como espacio vectorial con su carácter como anillo.

Nótese, además, que hay un elemento I en $\text{Hom}(V, V)$, definido por $vI = v$ para todo $v \in V$, con la propiedad de que $TI = IT = T$ para todo $T \in \text{Hom}(V, V)$. $\text{Hom}(V, V)$ es, pues, un anillo con elemento unitario. Además, si en la anterior propiedad (4) hacemos $T_2 = I$, obtenemos $\alpha T_1 = T_1(\alpha I)$. Como $(\alpha I)T_1 = \alpha(IT_1) = \alpha T_1$, vemos que $(\alpha I)T_1 = T_1(\alpha I)$ para todo $T_1 \in \text{Hom}(V, V)$, luego αI commuta con todo elemento de $\text{Hom}(V, V)$. De aquí en adelante escribiremos siempre αI simplemente como α .

DEFINICIÓN. Un anillo asociativo A se llama *álgebra sobre F* si A es un espacio vectorial sobre F tal que para cualesquiera $a, b \in A$ y $\alpha \in F$, $\alpha(ab) = (\alpha a)b = a(\alpha b)$.

Homomorfismos, isomorfismos, ideales, etc., se definen para las álgebras como se definieron para los anillos, con la exigencia suplementaria de que deben, para el caso de las álgebras, preservar también la estructura de espacio vectorial.

Nuestras anteriores observaciones indican que $\text{Hom}(V, V)$ es un álgebra sobre F . Para hacer la notación más manejable escribiremos de aquí en adelante $A(V)$ en lugar de $\text{Hom}(V, V)$; siempre que queramos enfatizar el papel del campo F emplearemos como notación $A_F(V)$.

DEFINICIÓN. Una transformación lineal sobre V , como espacio vectorial sobre F , es un elemento de $A_F(V)$.

A veces, nos referiremos a $A(V)$ como el *anillo, o álgebra de las transformaciones lineales sobre V* .

Para álgebras arbitrarias A , con elemento unitario, sobre un campo F , podemos probar el análogo del teorema de Cayley para grupos; a saber:

LEMA 6.1. Si A es un álgebra con elemento unitario sobre F , entonces A es isomorfa a una subálgebra de $A(V)$ para algún espacio vectorial V sobre F .

Prueba. Como A es un álgebra sobre F , debe ser un espacio vectorial sobre F . Usaremos $V = A$ para probar el teorema.

Si $a \in A$, sea $T_a : A \rightarrow A$ definida por $vT_a = va$ para todo $v \in A$. Afirmamos que T_a es una transformación lineal sobre $V (= A)$. Por la ley distributiva derecha $(v_1 + v_2)T_a = (v_1 + v_2)a = v_1a + v_2a = v_1T_a + v_2T_a$. Como A es un álgebra, $(\alpha v)T_a = (\alpha v)a = \alpha(va) = \alpha(vT_a)$ para $v \in A$, $\alpha \in F$. Luego T_a es ciertamente una transformación lineal sobre A .

Consideremos la aplicación $\psi : A \rightarrow A(V)$ definido por $a\psi = T_a$ para todo $a \in A$. Afirmamos que ψ es un isomorfismo de A en $A(V)$. Para comenzar, si $a, b \in A$ y $\alpha, \beta \in F$, entonces para todo $v \in A$, $vT_{\alpha a + \beta b} = v(\alpha a + \beta b) = \alpha(va) + \beta(vb)$ (por la ley distributiva izquierda y el hecho de que A es un álgebra sobre F) $= \alpha(vT_a) + \beta(vT_b) = v(\alpha T_a + \beta T_b)$, ya que tanto T_a como T_b son transformaciones lineales. En consecuencia $T_{\alpha a + \beta b} = \alpha T_a + \beta T_b$, de donde ψ es un homomorfismo de espacios vectoriales de A en $A(V)$. A continuación calculamos, para $a, b \in A$, $vT_{ab} = v(ab) = (va)b = (vT_a)b = v(T_a b)$ (hemos usado la ley asociativa de A en este cálculo), lo que implica que $T_{ab} = T_a T_b$. Hemos, pues, visto que ψ es también un homomorfismo de anillos de A . Hasta el momento hemos probado que ψ es un homomorfismo de A , como álgebra, en $A(V)$. Todo lo que queda por hacer es determinar el núcleo de ψ . Sea $a \in A$ un elemento del núcleo de ψ ; entonces $a\psi = 0$, de donde $T_a = 0$ y por tanto $vT_a = 0$ para todo $v \in V$. Ahora bien, $V = A$, y A tiene un elemento unidad e , de donde $eT_a = 0$. Sin embargo, $0 = eT_a = ea = a$, con lo que probamos que $a = 0$. El núcleo de ψ debe, por tanto, consistir solamente en 0, lo que implica que ψ es un isomorfismo de A en $A(V)$. Y esto completa la prueba del lema.

El lema señala el papel universal que juegan las álgebras particulares $A(V)$, puesto que en estas podemos encontrar copias isomorfas de cualquier álgebra.

Sea A un álgebra con elemento unidad e sobre F , y sea $p(x) = \alpha_0 + \alpha_1x + \dots + \alpha_nx^n$ un polinomio en $F[x]$. Para $a \in A$, representaremos por $p(a)$ al elemento $\alpha_0e + \alpha_1a + \dots + \alpha_na^n$ en A . Si $p(a) = 0$ diremos que a satisface a $p(x)$.

LEMÁ 6.2. *Sea A un álgebra con elemento unitario sobre F , y supongamos que A es de dimensión m sobre F . Entonces, todo elemento de A satisface algún polinomio no trivial en $F[x]$ de grado cuando más m .*

Prueba. Sea e el elemento unitario de A ; si $a \in A$, consideremos los $n+1$ elementos e, a, a^2, \dots, a^n de A . Como A es m -dimensional sobre F , por el lema 4.6, e, a, a^2, \dots, a^n , por ser $m+1$, deben ser linealmente dependientes sobre F . En otras palabras, hay elementos $\alpha_0, \alpha_1, \dots, \alpha_m$ en F , no todos cero, tales que $\alpha_0e + \alpha_1a + \dots + \alpha_ma^n = 0$. Pero entonces a satisface al polinomio no trivial $q(x) = \alpha_0 + \alpha_1x + \dots + \alpha_mx^m$, de grado, cuando más m , en $F[x]$.

Si V es un espacio vectorial de dimensión finita sobre F , de dimensión n , por el corolario 1 del teorema 4.d, $A(V)$ es de dimensión n^2 sobre F . Como $A(V)$ es un álgebra sobre F , podemos aplicar el lema 6.2 para obtener que todo elemento de $A(V)$ satisface un polinomio sobre F de grado, cuando más n^2 . Este hecho será de significación capital en todo lo que sigue, de modo que lo destacamos enunciándolo como un

TEOREMA 6.A. *Si V es un espacio vectorial de dimensión- n sobre F , entonces, dado un elemento cualquiera T de $A(V)$, existe un polinomio no trivial $q(x) \in F[x]$, de grado, cuando más n^2 , tal que $q(T) = 0$.*

Veremos después que podemos afirmar mucho más acerca del grado de $q(x)$; en realidad, veremos más adelante que puede escogerse un tal $q(x)$ con grado, cuando más n . Este hecho constituye un famoso teorema sobre nuestro actual tema de estudio conocido como el teorema de Cayley-Hamilton. Por el momento podemos seguir adelante sin una estimación más precisa del grado de $q(x)$; todo lo que necesitamos es que exista un $q(x)$ adecuado.

Como para V de dimensión finita, dado $T \in A(V)$ existe algún polinomio $q(x)$ para el que $q(T) = 0$, existe en $F[x]$ un polinomio $p(x)$ de grado mínimo no trivial con tal propiedad. Llamamos a $p(x)$ *polinomio mínimo* para T sobre F . Si T satisface un polinomio $h(x)$, entonces $p(x) | h(x)$.

DEFINICIÓN. Un elemento $T \in A(V)$ se llama *invertible a la derecha* si existe un $S \in A(V)$ tal que $TS = I$. (Aquí I denota al elemento unidad de $A(V)$.)

Análogamente, podemos definir que T es *invertible a la izquierda* si existe un $U \in A(V)$ tal que $UT = I$. Si T es invertible tanto a la derecha como a la izquierda y si $TS = UT = I$, es entonces un fácil probar, como ejercicio que $S = U$ y que S es único.

DEFINICIÓN. Un elemento T de $A(V)$ es *invertible o regular* si es invertible tanto a la derecha como a la izquierda; es decir, si hay un elemento $S \in A(V)$ tal que $ST = TS = I$. Escribimos entonces S como T^{-1} .

Un elemento en $A(V)$ que no es regular se llama *singular*.

Es completamente posible que un elemento en $A(V)$ sea invertible a la derecha, pero que no sea invertible. He aquí un ejemplo: sea F el campo de los números reales y sea $V = F[x]$, el conjunto de todos los polinomios en x sobre F . Definamos S en V por $q(x)S = \frac{d}{dx}q(x)$ y T por $q(x)T = \int_1^x q(x)dx$. Entonces $ST \neq I$, mientras que $TS = I$. Como veremos en un

momento, si V es finito dimensional sobre F entonces un elemento en $A(V)$ que es invertible a la derecha es invertible.

TEOREMA 6.B. *Si V es de dimensión finita sobre F entonces $T \in A(V)$ es invertible si y sólo si el término constante del polinomio mínimo para T no es 0.*

Prueba. Sea $p(x) = \alpha_0 + \alpha_1 x + \dots + \alpha_k x^k$, $\alpha_k \neq 0$ el polinomio mínimo para T sobre F .

Si $\alpha_0 \neq 0$, como $0 = p(T) = \alpha_k T^k + \alpha_{k-1} T^{k-1} + \dots + \alpha_1 T + \alpha_0$, tenemos

$$\begin{aligned} 1 &= T \left(-\frac{1}{\alpha_0} (\alpha_k T^{k-1} + \alpha_{k-1} T^{k-2} + \dots + \alpha_1) \right) \\ &= \left(-\frac{1}{\alpha_0} (\alpha_k T^{k-1} + \dots + \alpha_1) \right) T. \end{aligned}$$

Por lo tanto, $S = -\frac{1}{\alpha_0} (\alpha_k T^{k-1} + \dots + \alpha_1)$ actúa como un inverso para T , de donde T es invertible.

Supongamos, por otra parte, que T es invertible, pero $\alpha_0 = 0$. Así pues, $0 = \alpha_1 T + \alpha_2 T^2 + \dots + \alpha_k T^k = (\alpha_1 + \alpha_2 T + \dots + \alpha_k T^{k-1}) T$. Multiplicando esta relación a la derecha por T^{-1} obtenemos $\alpha_1 + \alpha_2 T + \dots + \alpha_k T^{k-1} = 0$, de donde T satisface el polinomio $q(x) = \alpha_1 + \alpha_2 x + \dots + \alpha_k x^{k-1}$ en $F[x]$. Como el grado de $q(x)$ es menor que el de $p(x)$, esto es imposible. Por tanto, $\alpha_0 \neq 0$, y la otra mitad del teorema queda probada.

COROLARIO 1. *Si V es de dimensión finita sobre F y si $T \in A(V)$ es invertible, entonces T^{-1} es una expresión polinomial en T sobre F .*

Prueba. Como T es invertible, por el teorema, $\alpha_0 + \alpha_1 T + \dots + \alpha_k T^k = 0$ con $\alpha_0 \neq 0$. Pero entonces $T^{-1} = -\frac{1}{\alpha_0} (\alpha_1 + \alpha_2 T + \dots + \alpha_k T^{k-1})$.

COROLARIO 2. *Si V es de dimensión finita sobre F y si $T \in A(V)$ es singular, entonces existe un $S \neq 0$ en $A(V)$ tal que $ST = TS = 0$.*

Prueba. Como T no es regular, el término constante de su polinomio mínimo debe ser 0. Es decir, $p(x) = \alpha_1 x + \dots + \alpha_k x^k$, donde $0 = \alpha_1 T + \dots + \alpha_k T^k$. Si $S = \alpha_1 + \dots + \alpha_k T^{k-1}$, entonces $S \neq 0$ (ya que $\alpha_1 + \dots + \alpha_k x^{k-1}$ es de grado menor que $p(x)$) y $ST = TS = 0$.

COROLARIO 3. *Si V es de dimensión finita sobre F y si $T \in A(V)$ es invertible a la derecha, entonces T es invertible.*

Prueba. Sea $TU = 1$. Si T fuera singular, habría un $S \neq 0$ tal que $ST = 0$. Pero, $0 = (ST)U = S(TU) = S1 = S \neq 0$. Una contradicción. Luego T es regular.

Deseamos aplicar la información contenida en el teorema 6.b y sus corolarios desde $A(V)$ a la acción de T sobre V . Un resultado de la mayor importancia a este respecto es

TEOREMA 6.C. *Si V es finito dimensional sobre F entonces $T \in A(V)$ es singular si y sólo si existe un $v \neq 0$ en V tal que $vT = 0$.*

Prueba. De acuerdo al corolario 2 del teorema 6.b, T es singular si y sólo si hay un $S \neq 0$ en $A(V)$ tal que $ST = TS = 0$. Como $S \neq 0$, hay un elemento $w \in V$ tal que $wS \neq 0$.

Sea $v = wS$; entonces $vT = (wS)T = w(ST) = w0 = 0$. Hemos presentado un vector no cero v , en V , al que T suprime. Recíprocamente, si $vT = 0$ con $v \neq 0$, dejamos como ejercicio probar que T no es invertible.

Buscamos aún otra caracterización de la singularidad o regularidad de una transformación lineal en términos de su acción global sobre V .

DEFINICIÓN. Si $T \in A(V)$, entonces la *imagen* de T , VT , está definida por $VT = \{vT \mid v \in V\}$.

Es fácil mostrar que la imagen de T es un subespacio vectorial de V . Consiste simplemente en todas las imágenes según T de los elementos de V . Notése que el rango de T es todo V si y sólo si T es suprayectiva.

TEOREMA 6.D. *Si V es de dimensión finita sobre F , entonces $T \in A(V)$ es regular si y sólo si T transforma V sobre V .*

Prueba. Como sucede con frecuencia, la mitad de lo afirmado es trivial; a saber, si T es regular, entonces, dado $v \in V$, $v = (vT^{-1})T$, de donde $VT = V$ y T es suprayectiva.

Por otra parte, supongamos que T no es regular. Debemos mostrar que T no es suprayectiva. Como T es singular, según el teorema 6.c, existe un vector $v_1 \neq 0$ en V tal que $v_1 T = 0$. De acuerdo con el lema 4.7 podemos completar partiendo de v_1 una base v_1, v_2, \dots, v_n de V . Entonces todo elemento de VT es una combinación lineal de los elementos $w_1 = v_1 T$, $w_2 = v_2 T, \dots, w_n = v_n T$. Como $w_1 = 0$, VT está generado por los $n-1$ elementos w_2, \dots, w_n , luego $\dim VT \leq n-1 < n = \dim V$. Pero entonces VT debe ser diferente de V ; es decir, T no es suprayectiva.

El teorema 6.d establece que podemos distinguir los elementos regulares de los singulares, en el caso de dimensión finita, de acuerdo a que sus

imágenes sean o no todo V . Si $T \in A(V)$, esto puede reformularse como sigue: T es regular si y sólo si $\dim(VT) = \dim V$. Esto sugiere que podríamos usar $\dim(VT)$ no solamente como una prueba para la regularidad sino incluso como una medida del grado de la singularidad (o falta de regularidad) para un $T \in A(V)$ dado.

DEFINICIÓN. Si V es de dimensión finita sobre F entonces el rango de T es la dimensión de VT , la imagen de T , sobre F .

Denotamos al rango de T por $r(T)$. A un extremo del espectro, si $r(T) = \dim V$, T es regular (y, por tanto, absolutamente no singular). En el otro extremo, si $r(T) = 0$, entonces $T = 0$ y entonces T es todo lo singular que se puede ser. El rango, como una función sobre $A(V)$, es una función importante y ahora investigaremos algunas de sus propiedades.

LEMA 6.3. Si V es de dimensión finita sobre F entonces para $S, T \in A(V)$

- 1) $r(ST) \leq r(T)$
- 2) $r(TS) \leq r(T)$
(y, por tanto, $r(ST) \leq \min\{r(T), r(S)\}$)
- 3) $r(ST) = r(TS) = r(T)$ para S regular en $A(V)$.

Prueba. Demostraremos en su orden (1), (2) y (3).

1) Como $VS \subset V$, $V(ST) = (VS)T \subset VT$, de donde, de acuerdo con el lema 4.8, $\dim(V(ST)) \leq \dim VT$; es decir, $r(ST) \leq r(T)$.

2) Supongamos que $r(T) = m$. Entonces VT tiene una base de m elementos, w_1, w_2, \dots, w_m . Pero entonces $(VT)S$ está generado por w_1S, w_2S, \dots, w_mS , de donde tiene una dimensión igual o menor que m . Como $r(TS) = \dim(V(TS)) = \dim((VT)S) \leq m = \dim VT = r(T)$, hemos probado (2).

3) Si S es invertible, entonces $VS = V$, de donde $V(ST) = (VS)T = VT$. Por tanto, $r(ST) = \dim(V(ST)) = \dim(VT) = r(T)$. Por otra parte, si VT tiene w_1, \dots, w_m como base, la regularidad de S implica que w_1S, \dots, w_mS son linealmente independientes. (¡Pruébese!) Como estos vectores generan $V(TS)$, forman una base de $V(TS)$. Pero entonces $r(TS) = \dim(V(TS)) = \dim(VT) = r(T)$.

COROLARIO. Si $T \in A(V)$ y si $S \in A(V)$ es regular, entonces $r(T) = r(STS^{-1})$.

Prueba. Según la parte (3) del lema, $r(STS^{-1}) = r(S(TS^{-1})) = r(TS^{-1})S = r(T)$.

Problemas

En todos los problemas, *salvo aviso expreso en contra*, V representará un espacio vectorial de dimensión finita sobre un campo F .

1. Pruébese que $S \in A(V)$ es regular si y sólo si siempre que $v_1, \dots, v_n \in V$ son linealmente independientes, entonces $v_1 S, \dots, v_n S$ son también linealmente independientes.
2. Pruébese que $T \in A(V)$ queda completamente determinado por sus valores sobre una base de V .
3. Pruébese el lema 6.1 incluso cuando A no tiene elemento unidad.
4. Si A es el campo de los números complejos y F es el campo de los números reales, entonces A es un álgebra sobre F de dimensión 2. Para $a = \alpha + \beta i$ en A , calcúlese la acción de T_a (véase el lema 6.1) sobre una base de A sobre F .
5. Si V es bidimensional sobre F y $A = A(V)$, obténgase una base de A sobre F y calcúlese T_a para cada a en tal base.
6. Si $\dim_F V > 1$, pruébese que $A(V)$ no es comutativa.
7. En $A(V)$, sea $Z = \{T \in A(V) \mid ST = TS \text{ para toda } S \in A(V)\}$. Pruébese que Z consiste simplemente en los múltiplos del elemento unidad de $A(V)$ por los elementos de F .
- *8. Si $\dim_F V > 1$, pruébese que $A(V)$ no tiene ideales bilaterales aparte de (0) y $A(V)$.
- **9. Pruébese que la conclusión del problema 8 es falsa si V no es de dimensión finita sobre F .
10. Si V es un espacio vectorial arbitrario sobre F y si $T \in A(V)$ es invertible tanto a la derecha como a la izquierda, pruébese que el inverso derecho de T y su inverso izquierdo son iguales. Partiendo de ello, pruébese que el inverso es único.
11. Si V es un espacio vectorial arbitrario sobre F y si $T \in A(V)$ es invertible a la derecha con un inverso derecho *único*, pruébese que T es invertible.
12. Pruébese que los elementos regulares de $A(V)$ forman un grupo.
13. Si F es el campo de los enteros módulo 2 y V es bidimensional sobre F , calcúlese el grupo de elementos regulares de $A(V)$ y pruébese que este grupo es isomorfo a S_3 , el grupo simétrico de grado 3.
- *14. Si F es un campo finito con q elementos, calcúlese el orden del grupo de los elementos regulares de $A(V)$ cuando V es bidimensional sobre F .

*15. Hágase el problema 14 cuando V es de dimensión n sobre F .

*16. Si V es de dimensión finita, pruébese que todo elemento en $A(V)$ puede escribirse como una suma de elementos regulares.

17. Un elemento $E \in A(V)$ se llama *idempotente* si $E^2 = E$. Si $E \in A(V)$ es un idempotente, pruébese que $V = V_0 \oplus V_1$ donde $v_0 E = 0$ para todo $v_0 \in V_0$ y $v_1 E = v_1$ para todo $v_1 \in V_1$.

18. Si $T \in A_F(V)$, con F de característica distinta de 2, satisface $T^3 = T$, pruébese que $V = V_0 \oplus V_1 \oplus V_2$ donde:

- 1) $v_0 \in V_0$ implica $v_0 T = 0$
- 2) $v_1 \in V_1$ implica $v_1 T = v_1$
- 3) $v_2 \in V_2$ implica $v_2 T = -v_2$.

*19. Si V es de dimensión finita y $T \neq 0 \in A(V)$, pruébese que hay un $S \in A(V)$ tal que $E = TS \neq 0$ es un idempotente.

20. El elemento $T \in A(V)$ se llama *nilpotente* si $T^m = 0$ para algún m . Si T es nilpotente y si $vT = \alpha v$ para algún $v \neq 0$ en V , con $\alpha \in F$, pruébese que $\alpha = 0$.

21. Si $T \in A(V)$ es nilpotente, pruébese que $\alpha_0 + \alpha_1 T + \alpha_2 T^2 + \dots + \alpha_k T^k$ es regular, con tal de que $\alpha_0 \neq 0$.

22. Si A es un álgebra de dimensión finita sobre F y si $a \in A$, pruébese que para algún entero $k > 0$ y algún polinomio $p(x) \in F[x]$, $a^k = a^{k+1}p(a)$.

*23. Usando el resultado del problema 22, pruébese que para $a \in A$ hay un polinomio $q(x) \in F[x]$ tal que $a^k = a^{2k}q(a)$.

24. Usando el resultado del problema 23, pruébese que dada $a \in A$ o a es nilpotente o hay un elemento $b \neq 0$ en A de la forma $b = ah(a)$, donde $h(x) \in F[x]$, tal que $b^2 = b$.

25. Si A es un álgebra sobre F (no necesariamente de dimensión finita) y si para $a \in A$, $a^2 - a$ es nilpotente, pruébese que o a es nilpotente o hay un elemento de la forma $b = ah(a) \neq 0$, donde $h(x) \in F[x]$, tal que $b^2 = b$.

*26. Si $T \neq 0 \in A(V)$ es singular, pruébese que hay un elemento $S \in A(V)$ tal que $TS = 0$ pero $ST \neq 0$.

27. Sea V bidimensional sobre F con base v_1, v_2 . Supongamos que $T \in A(V)$ es tal que $v_1 T = \alpha v_1 + \beta v_2$ y $v_2 T = \gamma v_1 + \delta v_2$ donde $\alpha, \beta, \gamma, \delta \in F$. Encuéntrese un polinomio distinto de cero en $F[x]$ de grado 2 satisfecho por T .

28. Si V es tridimensional sobre F con base v_1, v_2, v_3 y si $T \in A(V)$ es tal que $v_i T = \alpha_{i1} v_1 + \alpha_{i2} v_2 + \alpha_{i3} v_3$ para $i = 1, 2, 3$, con todos los $\alpha_{ij} \in F$, encuéntrese un polinomio de grado 3 en $F[x]$ satisfecho por T .

29. Sea V n -dimensional sobre F con una base v_1, \dots, v_n . Supongamos que $T \in A(V)$ es tal que

$$v_1 T = v_2, v_2 T = v_3, \dots, v_{n-1} T = v_n,$$

$$v_n T = -\alpha_n v_1 - \alpha_{n-1} v_2 - \dots - \alpha_1 v_n,$$

donde $\alpha_1, \dots, \alpha_n \in F$. Pruébese que T satisface el polinomio

$$p(x) = x^n + \alpha_1 x^{n-1} + \alpha_2 x^{n-2} + \dots + \alpha_n \text{ sobre } F.$$

30. Si $T \in A(V)$ satisface un polinomio $q(x) \in F[x]$, pruébese que para $S \in A(V)$, S regular, STS^{-1} también satisface $q(x)$.

31. a) Si F es el campo de los números racionales y si V es tridimensional sobre F con una base v_1, v_2, v_3 , calcúlese la clase de $T \in A(V)$ definido por

$$v_1 T = v_1 - v_2$$

$$v_2 T = v_1 + v_3$$

$$v_3 T = v_2 + v_3.$$

b) Encuéntrese un vector $v \in V$, $v \neq 0$, tal que $vT = 0$.

32. Pruébese que la imagen de T y $U = \{v \in V \mid vT = 0\}$ son subespacios de V .

33. Si $T \in A(V)$, sea $V_0 = \{v \in V \mid vT^k = 0 \text{ para algún } k\}$. Pruébese que V_0 es un subespacio y que si $vT^m V_0$, entonces $v \in V_0$.

34. Pruébese que el polinomio mínimo de T sobre F divide a todos los polinomios satisfechos por T sobre F .

*35. Si $n(T)$ es la dimensión de la U del problema 32, pruébese que $r(T) + n(T) = \dim V$.

2. RAÍCES CARACTERÍSTICAS

En lo que falta de este capítulo nuestro interés se limitará a transformaciones lineales sobre espacios vectoriales de dimensión finita. Así pues, de aquí en adelante, V denotará siempre un espacio vectorial de dimensión finita sobre un campo F .

El álgebra $A(V)$ tiene un elemento unitario; para mayor comodidad, en la notación la representaremos por 1. Por el símbolo $\lambda - T$, para $\lambda \in F$, $T \in A(V)$, representaremos $\lambda 1 - T$.

DEFINICIÓN. Si $T \in A(V)$, entonces $\lambda \in F$ se llama *raíz característica* (o *valor propio*, o *eigen valor*) de T si $\lambda - T$ es singular.

Deseamos caracterizar la propiedad de ser una raíz característica en el comportamiento de T sobre V . Hacemos esto en el

TEOREMA 6.E. *El elemento $\lambda \in F$ es una raíz característica de $T \in A(V)$ si y sólo si para algún $v \neq 0$ en V , $vT = \lambda v$.*

Prueba. Si λ es una raíz característica de T , entonces $\lambda - T$ es singular, de donde, según el teorema 6.c, hay un vector $v \neq 0$ en V tal que $v(\lambda - T) = 0$. Pero entonces $\lambda v = vT$.

Por otra parte, si $vT = \lambda v$ para algún $v \neq 0$, entonces $v(\lambda - T) = 0$, de donde, de nuevo por el teorema 6.c, $\lambda - T$ debe ser singular y, por tanto, λ es una raíz característica de T .

LEMA 6.4. *Si $\lambda \in F$ es una raíz característica de $T \in A(V)$, entonces para cualquier polinomio $q(x) \in F[x]$, $q(\lambda)$ es una raíz característica de $q(T)$.*

Prueba. Supongamos que $\lambda \in F$ es una raíz característica de T . De acuerdo con el teorema 6.e, hay un vector distinto de cero, v , tal que $vT = \lambda v$. ¿Qué es lo que ocurre con vT^2 ?

Tenemos: $vT^2 = (\lambda v)T = \lambda(vT) = \lambda(\lambda v) = \lambda^2 v$. Continuando de esta forma vemos que $vT^k = \lambda^k v$ para cualquier entero positivo k . Si $q(x) = \alpha_0 x^m + \alpha_1 x^{m-1} + \dots + \alpha_m$, $\alpha_i \in F$, entonces $q(T) = \alpha_0 T^m + \alpha_1 T^{m-1} + \dots + \alpha_m$, de donde $vq(T) = v(\alpha_0 T^m + \alpha_1 T^{m-1} + \dots + \alpha_m) = \alpha_0(vT^m) + \alpha_1(vT^{m-1}) + \dots + \alpha_m v = (\alpha_0 \lambda^m + \alpha_1 \lambda^{m-1} + \dots + \alpha_m)v = q(\lambda)v$, por la observación que anteriormente hicimos. Por tanto $v(q(\lambda) - q(T)) = 0$, de donde, según el teorema 6.e, $q(\lambda)$ es una raíz característica de $q(T)$.

Como consecuencia inmediata del lema 6.4, en realidad como un simple caso especial (pero extremadamente importante), tenemos

TEOREMA 6.F. *Si $\lambda \in F$ es una raíz característica de $T \in A(V)$, entonces λ es una raíz del polinomio mínimo de T . En particular, T solamente tiene un número finito de raíces características en F .*

Prueba. Sea $p(x)$ el polinomio mínimo de T sobre F ; tenemos, pues, $p(T) = 0$. Si $\lambda \in F$ es una raíz característica de T , hay una $v \neq 0$ en V con $vT = \lambda v$. Como en la prueba del lema 6.4, $vp(T) = p(\lambda)v$; pero $p(T) = 0$, lo que, por tanto, implica que $p(\lambda)v = 0$. Como $v \neq 0$, debemos tener por las propiedades de un espacio vectorial que $p(\lambda) = 0$. Por lo tanto, λ es una raíz de $p(x)$. Como $p(x)$ tiene solamente un número finito de raíces (en realidad, como $\deg p(x) \leq n^2$ donde $n = \dim_F V$, $p(x)$ tiene, cuando más, n^2 raíces) en F , puede solamente haber un número finito de raíces características de T en F .

Si $T \in A(V)$ y si $S \in A(V)$ es regular, entonces $(STS^{-1})^2 = STS^{-1}STS^{-1} = ST^2S^{-1}$, $(STS^{-1})^3 = ST^3S^{-1}, \dots, (STS^{-1})^i = ST^iS^{-1}$. Por consiguiente, para cualquier $q(x) \in F[x]$, $q(STS^{-1}) = Sq(T)S^{-1}$. En particular,

si $q(T) = 0$, entonces $q(STS^{-1}) = 0$. Luego si $p(x)$ es el polinomio mínimo para T , entonces de lo dicho se sigue fácilmente que $p(x)$ es también el polinomio mínimo para STS^{-1} . Y hemos probado:

LEMA 6.5. Si $T, S \in A(V)$ y si S es regular, entonces T y STS^{-1} tienen el mismo polinomio mínimo.

DEFINICIÓN. El elemento $0 \neq v \in V$ se llama *vector característico* de T perteneciente a la raíz característica $\lambda \in F$ si $vT = \lambda v$.

¿Qué relación, si es que alguna, debe existir entre los vectores característicos de T pertenecientes a las distintas raíces características? A esto lo contesta el siguiente

TEOREMA 6.F'. Si $\lambda_1, \dots, \lambda_k$ de F son distintas raíces características de $T \in A(V)$ y si v_1, \dots, v_k son vectores característicos de T pertenecientes a $\lambda_1, \dots, \lambda_k$ respectivamente, entonces v_1, \dots, v_k son linealmente independientes sobre F .

Prueba. Para que el teorema requiera prueba, k debe ser mayor que 1, de modo que suponemos que $k > 1$.

Si v_1, \dots, v_k linealmente dependientes sobre F , entonces hay una relación de la forma $\alpha_1 v_1 + \dots + \alpha_k v_k = 0$, donde $\alpha_1, \dots, \alpha_k$ son todos de F y no todos ellos son 0. Entre todas las relaciones tales hay una que tiene un número mínimo de coeficientes distintos de cero. Renumerando los vectores de modo adecuado si fuera necesario, podemos suponer que esta relación más corta es

$$1) \quad \beta_1 v_1 + \dots + \beta_j v_j = 0, \quad \beta_1 \neq 0, \dots, \beta_j \neq 0.$$

Sabemos que $v_i T = \lambda_i v_i$, luego, aplicando T a la ecuación (1), obtenemos

$$2) \quad \lambda_1 \beta_1 v_1 + \dots + \lambda_j \beta_j v_j = 0.$$

Multiplicando la ecuación (1) por λ_1 y restando de la ecuación (2) obtenemos

$$(\lambda_2 - \lambda_1) \beta_2 v_2 + \dots + (\lambda_j - \lambda_1) \beta_j v_j = 0.$$

Pero $\lambda_i - \lambda_1 \neq 0$ para $i > 1$ y $\beta_i \neq 0$, de donde $(\lambda_i - \lambda_1) \beta_i \neq 0$. Pero entonces hemos construido una relación más corta que la (1) entre v_1, v_2, \dots, v_k . Esta contradicción prueba el teorema.

COROLARIO 1. Si $T \in A(V)$ y si $\dim_F V = n$, entonces T puede tener cuando más n raíces características en F .

Prueba. Cualquier conjunto de vectores linealmente independientes en V puede tener cuando más n elementos. Como cualquier conjunto de raíces características distintas de T , según el teorema 6.F', da lugar a un conjunto

correspondiente de vectores característicos linealmente independientes, de ellos se sigue el

COROLARIO 2. Si $T \in A(V)$ y si $\dim_F V = n$ y si T tiene n raíces características distintas en F , entonces hay una base de V sobre F que se compone de vectores característicos de T .

Dejamos la prueba de este corolario al lector. El corolario 2 no es sino el primero de toda una clase de teoremas en los que se nos dirá que una transformación lineal dada tiene cierta base muy conveniente del espacio vectorial sobre la cual nos es muy sencillo describir su acción

Problemas

En todos los problemas V es un espacio vectorial sobre F .

1. Si $T \in A(V)$ y si $q(x) \in F[x]$ es tal que $q(T) = 0$, ¿es cierto que toda raíz de $q(x)$ en F es una raíz característica de T ? Pruébese que esto es cierto o proporcionese un ejemplo en que se muestre que es falso.

2. Si $T \in A(V)$ y si $p(x)$ es el polinomio mínimo para T sobre F , supongamos que $p(x)$ tiene todas sus raíces en F . Pruébese que toda raíz de $p(x)$ es una raíz característica de T .

3. Sea V bidimensional sobre el campo F de los números reales con una base v_1, v_2 . Encuéntrense las raíces características y los correspondientes vectores característicos para las siguientes T :

- a) $v_1 T = v_1 + v_2, \quad v_2 T = v_1 - v_2.$
- b) $v_1 T = 5v_1 + 6v_2, \quad v_2 T = -7v_2.$
- c) $v_1 T = v_1 + 2v_2, \quad v_2 T = 3v_1 + 6v_2.$

4. Sea V como en el problema 3 y supongamos que $T \in A(V)$ es tal que $v_1 T = \alpha v_1 + \beta v_2, v_2 T = \gamma v_1 + \delta v_2$ donde $\alpha, \beta, \gamma, \delta$ están en F .

- a) Encuéntrense las condiciones necesarias y suficientes para que 0 sea una raíz característica de T en términos de $\alpha, \beta, \gamma, \delta$.
- b) Encuéntrense, en términos de $\alpha, \beta, \gamma, \delta$, condiciones necesarias y suficientes para que T tenga dos raíces características distintas en F .

5. Si V es bidimensional sobre un campo F pruébese que todo elemento en $A(V)$ satisface un polinomio de grado 2 sobre F .

*6. Si V es bidimensional sobre F y si $S, T \in A(V)$, pruébese que $(ST - TS)^2$ commuta con todos los elementos de $A(V)$.

7. Pruébese el corolario 2 al teorema 6.f.

8. Si V es n -dimensional sobre F y si $T \in A(V)$ es nilpotente (es decir, tal que $T^k = 0$ para algún k), pruébese que $T^n = 0$. (*Sugerencia:* si $v \in V$ úsese el hecho de que v, vT, vT^2, \dots, vT^n deben ser linealmente independientes sobre F .)

3. MATRICES

Aunque ya llevamos algún tiempo tratando de transformaciones, siempre lo hemos hecho en una forma impersonal y un poco lejana; para nosotros, una transformación lineal ha sido un símbolo (muy a menudo T) que actúa en una cierta forma sobre un espacio vectorial. Vemos, cuando pensamos en lo hasta aquí hecho, que fuera de los pocos ejemplos concretos con que nos hemos encontrado en los problemas, nunca nos hemos enfrentado con transformaciones lineales específicas. Al mismo tiempo, es claro que si hemos de proseguir con el tema un poco más lejos a menudo se presentará la necesidad de hacer un estudio completo y detallado de una transformación lineal dada. Para mencionar un problema preciso, si se nos presenta una transformación lineal (y suponiendo por el momento que tenemos medios para reconocerla), ¿cómo podemos arreglárnoslas para encontrar, de una forma práctica y calculable, sus raíces características?

Lo que primero buscamos es una notación sencilla o, quizás más precisamente, una representación sencilla para las transformaciones lineales. Llegaremos a ello mediante el uso de una base particular del espacio vectorial y por el uso de la acción de una transformación lineal sobre esta base. Una vez que se ha conseguido todo esto, por medio de las operaciones en $A(V)$ podemos inducir operaciones para los símbolos creados que hagan de ellos un álgebra. Este nuevo objeto, infundido de una vida algebraica propia, puede estudiarse como una entidad matemática que tiene un interés por sí misma. Este estudio es lo que comprende la llamada *teoría de matrices*.

Pero ignorar el origen de estas matrices, es decir, investigar el conjunto de símbolos independientemente de lo que representan, puede ser costoso, porque estaríamos desperdiциando una gran cantidad de información útil.

En lugar de ello, nosotros siempre usaremos las interrelaciones entre el abstracto $A(V)$ y lo concreto, el álgebra de matrices, para obtener información de una sobre la otra.

Sea V un espacio vectorial n -dimensional sobre un campo F y sea v_1, \dots, v_n una base de V sobre F . Si $T \in A(V)$ entonces T está determinado en cualquier vector tan pronto como conoczamos su acción sobre una base de V . Como T transforma V en V , $v_1 T, v_2 T, \dots, v_n T$ deben estar todos en V . Como elementos de V cada uno de estos es realizable de un único modo como

combinación lineal de v_1, \dots, v_n sobre F . Así pues:

$$v_1 T = \alpha_{11} v_1 + \alpha_{12} v_2 + \cdots + \alpha_{1n} v_n$$

$$v_2 T = \alpha_{21} v_1 + \alpha_{22} v_2 + \cdots + \alpha_{2n} v_n$$

$$v_i T = \alpha_{i1} v_1 + \alpha_{i2} v_2 + \cdots + \alpha_{in} v_n$$

.

$$v_n T = \alpha_{n1} v_1 + \alpha_{n2} v_2 + \cdots + \alpha_{nn} v_n,$$

donde cada $\alpha_{ij} \in F$. Este sistema de ecuaciones puede escribirse más compactamente como

$$v_i T = \sum_{j=1}^n \alpha_{ij} v_j, \quad \text{para } i = 1, 2, \dots, n.$$

El conjunto ordenado de n^2 números α_{ij} en F describe completamente a T . Nos servirán como medio para representar T .

DEFINICIÓN. Sea V un espacio vectorial de dimensión n sobre F y sea v_1, \dots, v_n una base para V sobre F . Si $T \in A(V)$ entonces la matriz de T en la base v_1, \dots, v_n , a la que representaremos por $m(T)$, es

$$m(T) = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{pmatrix}$$

donde $v_i T = \sum_j \alpha_{ij} v_j$.

Una matriz es entonces un arreglo ordenado en forma de cuadrado de elementos de F con, hasta el momento, ninguna otra característica, que representa el efecto de una transformación lineal sobre una base dada.

Examinemos un ejemplo. Sea F un campo y sea V el conjunto de todos los polinomios en x de grado $n-1$ o menor sobre F . Definamos D sobre V por $(\beta_0 + \beta_1 x + \dots + \beta_{n-1} x^{n-1})D = \beta_1 + 2\beta_2 x + \dots + i\beta_i x^{i-1} \dots + (n-1)\beta_{n-1} x^{n-2}$. Es trivial comprobar que D es una transformación lineal sobre V ; como el lector habrá visto, se trata simplemente del operador de diferenciación.

¿Cuál es la matriz de D ? La pregunta carece de sentido a menos que especifiquemos una base de V . En primer lugar, calculemos la matriz de D en la base $v_1 = 1, v_2 = x, v_3 = x^2, \dots, v_i = x^{i-1}, \dots, v_n = x^{n-1}$. Ahora

bien,

$$\begin{aligned}
 v_1 D &= 1 D = 0 = 0v_1 + 0v_2 + \cdots + 0v_n \\
 v_2 D &= xD = 1 = 1v_1 + 0v_2 + \cdots + 0v_n \\
 &\vdots \\
 v_i D &= x^{i-1} D = (i-1)x^{i-2} \\
 &= (C-1)v_{i-1} + 0v_n + \cdots + 0v_{i-2} + (i-1)v_{i-1} + 0v_1 \\
 &\quad + \cdots + 0v_n \\
 &\vdots \\
 v_n D &= x^{n-1} D = (n-1)x^{n-2} \\
 &= 0v_1 + 0v_2 + \cdots + 0v_{n-2} + (n-1)v_{n-1} + 0v_n.
 \end{aligned}$$

Si volvemos a la propia definición de matriz de una transformación lineal en una base dada, vemos que la matriz de D en la base v_1, \dots, v_n , $m_1(D)$, es un

$$m_1(D) = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 3 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & (n-1) & 0 \end{pmatrix}$$

Pero nada hay de especial en la base que acabamos de usar ni en como numeramos sus elementos. Supongamos que nos limitamos a reordenar los elementos de esta base; obtenemos entonces una base tan buena como la anterior $w_1 = x^{n-1}$, $w_2 = x^{n-2}$, ..., $w_i = x^{n-i}$, ..., $w_n = 1$. ¿Cuál es, con respecto a esta nueva base, la matriz de la misma transformación lineal? Tenemos ahora,

$$\begin{aligned}
 w_1 D &= x^{n-1} D = (n-1)x^{n-2} \\
 &= 0w_1 + (n-1)w_2 + 0w_3 + \cdots + 0w_n \\
 &\vdots \\
 w_i D &= x^{n-i} D = (n-i)x^{n-i-1} \\
 &= 0w_1 + \cdots + 0w_i + (n-i)w_{i+1} + 0w_{i+2} + \cdots + 0w_n \\
 &\vdots \\
 w_n D &= 1 D = 0 = 0w_1 + 0w_2 + \cdots + 0w_n,
 \end{aligned}$$

de donde $m_2(D)$, la matriz de D en esta base es

$$m_2(D) = \begin{pmatrix} 0 & (n-1) & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & (n-2) & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & (n-3) & \cdots & 0 & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \vdots & \ddots & & & & & \\ 0 & 0 & 0 & \cdots & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & \cdots & 0 & 0 \end{pmatrix}$$

Antes de terminar con este ejemplo, calculemos la matriz de D en otra base más de V sobre F . Sea $u_1 = 1, u_2 = 1+x, u_3 = 1+x^2, \dots, u_n = 1+x^{n-1}$; es fácil verificar que u_1, \dots, u_n forman una base de V sobre F . ¿Cuál es la matriz de D en esta base? Como

$$\begin{aligned} u_1 D &= 1D = 0 = 0u_1 + 0u_2 + \dots + 0u_n \\ u_2 D &= (1+x)D = 1 = 1u_1 + 0u_2 + \dots + 0u_n \\ u_3 D &= (1+x^2)D = 2x = 2(u_2 - u_1) = -2u_1 + 2u_2 + 0u_3 + \dots + 0u_n \\ &\vdots \\ u_n D &= (1+x^{n-1})D = (n-1)x^{n-2} = (n-1)(u_n - u_1) \\ &= -(n-1)u_1 + 0u_2 + \dots + 0u_{n-2} + (n-1)u_{n-1} + 0u_n. \end{aligned}$$

la matriz $m_3(D)$ de D en esta base es

$$m_3(D) = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ -2 & 2 & 0 & \cdots & 0 & 0 \\ -3 & 0 & 3 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \cdots & 0 & 0 \\ -(n-1) & 0 & 0 & \cdots & (n-1) & 0 \end{pmatrix}$$

Por el ejemplo que hemos estudiado vemos que las matrices de D , para las tres bases usadas dependían completamente de las bases. Aunque diferentes las unas de las otras representan, sin embargo, a la misma trans-

formación lineal D , y podríamos haber reconstruido D partiendo de una cualquiera de ellas si conociéramos la base usada en su determinación. Pero, aunque diferente, sería de esperar que existiera alguna relación entre $m_1(D)$, $m_2(D)$ y $m_3(D)$. Esta relación será la que determinaremos exactamente más tarde.

Como la base a usar en cualquier ocasión puede ser cualquiera, dada una transformación lineal T (cuya definición, después de todo, no depende de ninguna base) es natural que busquemos una base en que la matriz de T tenga una forma particularmente sencilla. Por ejemplo, si T es una transformación lineal sobre V , que es n dimensional sobre F , y si T tiene n raíces características distintas $\lambda_1, \dots, \lambda_n$ en F , entonces, de acuerdo con el corolario 2 al teorema 6.1, podemos encontrar una base v_1, \dots, v_n de V sobre F tal que $v_i T = \lambda_i v_i$. En esta base T tiene como matriz la de forma particularmente sencilla,

$$m(T) = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & \lambda_n \end{pmatrix}$$

Hemos visto que una vez que hemos escogido una base para V , a cada transformación lineal se le asocia una matriz. Recíprocamente, una vez que hemos escogido una base fija v_1, \dots, v_n de V sobre F , una matriz dada

$$\begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{pmatrix}, \quad \alpha_{ij} \in F,$$

da lugar a una transformación lineal T definida sobre V por $v_i T = \sum_j \alpha_{ij} v_j$ sobre esta base. Nótese que la matriz de la transformación lineal T que acabamos de construir en la base v_1, \dots, v_n es exactamente la matriz con la que comenzamos. Por tanto, toda posible ordenación en forma de cuadrado nos sirve como la matriz de alguna transformación lineal en la base v_1, \dots, v_n .

Es claro lo que quiere decir cada una de las expresiones primer renglón, segundo renglón, ..., de una matriz, como análogamente, lo que debe entenderse por primera columna, segunda columna, En la matriz

$$\begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{pmatrix}$$

el elemento α_{ij} está en el i -ésimo renglón y j -ésima columna; nos referimos a él como el elemento (i, j) (o la *entrada* (i, j)) de la matriz.

Escribir todo el arreglo cuadrado de la matriz es algo pesado; en lugar de ello escribiremos una matriz como (α_{ij}) ; esto indica que la entrada (i, j) de la matriz es α_{ij} .

Supongamos que V es un espacio vectorial de dimensión n sobre F y v_1, \dots, v_n es una base de V sobre F que quedará fija en toda la discusión que sigue. Supongamos que S y T son transformaciones lineales sobre V (y sobre F) con matrices $m(S) = (\sigma_{ij})$ y $m(T) = (\tau_{ij})$, respectivamente, en la base dada. Nuestro objetivo es aplicar la estructura algebraica de $A(V)$ al conjunto de matrices que tienen sus entradas en F .

Para comenzar, como $S = T$ si y sólo si $vS = vT$ para todo $v \in V$, se tiene que $S = T$ si y sólo si $v_i S = v_i T$ para todos los v_1, \dots, v_n que forman una base de V sobre F . O, lo que es equivalente, $S = T$ si y sólo si $\sigma_{ij} = \tau_{ij}$ para todo i y todo j .

Dadas $m(S) = (\sigma_{ij})$ y $m(T) = (\tau_{ij})$, ¿podemos escribir explícitamente $m(S+T)$? Como $m(S) = (\sigma_{ij})$, $v_i S = \sum_j \sigma_{ij} v_j$; análogamente, $v_i T = \sum_j \tau_{ij} v_j$, de donde $v_i(S+T) = v_i S + v_i T = \sum_j \sigma_{ij} v_j + \sum_j \tau_{ij} v_j = \sum_j (\sigma_{ij} + \tau_{ij}) v_j$.

Pero entonces, por lo que se entiende por matriz de una transformación lineal en una base dada, $m(S+T) = (\lambda_{ij})$ donde $\lambda_{ij} = \sigma_{ij} + \tau_{ij}$ para toda i y toda j . Un cálculo de la misma clase muestra que para $y \in F$, $m(yS) = (\mu_{ij})$ donde $\mu_{ij} = y\sigma_{ij}$ para toda i y toda j .

El cálculo más interesante, y también el más complicado, es el de $m(ST)$. Tenemos ahora $v_i(ST) = (v_i S)T = (\sum_k \sigma_{ik} v_k)T = \sum_k \sigma_{ik} (v_k T)$. Sin embargo, $v_k T = \sum_j \tau_{kj} v_j$; lo que sustituido en la fórmula anterior, nos da

$$v_i(ST) = \sum_k \sigma_{ik} \left(\sum_j \tau_{kj} v_j \right) = \sum_j \left(\sum_k \sigma_{ik} \tau_{kj} \right) v_j.$$

(Pruébese). Por tanto, $m(ST) = (\nu_{ij})$, donde para todo i y para toda j , $\nu_{ij} = \sum_k \sigma_{ik} \tau_{kj}$.

A primera vista, la regla para calcular la matriz del producto de dos transformaciones lineales en una base dada parece complicada. Sin embargo, nótese que la entrada (i, j) se obtiene como sigue: consideremos los renglones de S como vectores y las columnas de T como vectores; entonces la entrada (i, j) de $m(ST)$ es simplemente el producto punto de la i -ésima fila de S con la j -ésima columna de T .

Ilustremos esto con un ejemplo. Supongamos que

$$m(S) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

y

$$m(T) = \begin{pmatrix} -1 & 0 \\ 2 & 3 \end{pmatrix}$$

el producto punto del primer renglón de S con la primera columna de T es $(1)(-1)+(2)(2)=3$, de donde la entrada $(1, 1)$ de $m(ST)$ es 3; el producto punto de la primera fila de S con la segunda columna de T es $(1)(0)+(2)(3)=6$, de donde la entrada $(1, 2)$ de $m(ST)$ es 6; el producto punto del segundo renglón de S con la primera columna de T es $(3)(-1)+(4)(2)=5$, de donde la entrada $(2, 1)$ de $m(ST)$ es 5; finalmente, el producto punto de la segunda fila de S con la segunda columna de T es $(3)(0)+(4)(3)=12$, de donde la entrada $(2, 2)$ de $m(ST)$ es 12. Así pues,

$$m(ST) = \begin{pmatrix} 3 & 6 \\ 5 & 12 \end{pmatrix}$$

La anterior discusión se ha hecho pensando principalmente en que sirviera de motivación para las construcciones que estamos a punto de presentar.

Sea F un campo; una matriz $n \times n$ sobre F será un arreglo en forma de cuadrado de elementos en F ,

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{pmatrix}$$

(que representamos por (α_{ij})). Sea $F_n = \{(\alpha_{ij}) \mid \alpha_{ij} \in F\}$; en F_n queremos introducir la noción de igualdad entre sus elementos, una adición, una multiplicación escalar por elementos de F y una multiplicación de forma que se convierta en un álgebra sobre F . Usamos las propiedades de $m(T)$ para $T \in A(V)$ como nuestra guía en todo esto.

- 1) Afirmamos que $(\alpha_{ij}) = (\beta_{ij})$, cuando tenemos dos matrices en F_n , si y sólo si $\alpha_{ij} = \beta_{ij}$ para toda i y para toda j .
- 2) Definimos $(\alpha_{ij}) + (\beta_{ij}) = \lambda_{ij}$ donde $\lambda_{ij} = \alpha_{ij} + \beta_{ij}$ para toda i y para toda j .
- 3) Para $\gamma \in F$, definimos $\gamma(\alpha_{ij}) = (\mu_{ij})$ donde $\mu_{ij} = \gamma\alpha_{ij}$ para toda i y para toda j .
- 4) Definimos $(\alpha_{ij})(\beta_{ij}) = (\nu_{ij})$, donde para toda i y toda j $\nu_{ij} = \sum_k \alpha_{ik}\beta_{kj}$.

Sea V un espacio vectorial de dimensión n sobre F y sea v_1, \dots, v_n una base de V sobre F ; la matriz $m(T)$ en la base v_1, \dots, v_n asocia con $T \in A(V)$ un elemento $m(T)$ en F_n . Sin más preámbulo, afirmamos que la aplicación de

$A(V)$ en F_n definido al transformar T sobre $m(T)$ es un isomorfismo de álgebras de $A(V)$ sobre F_n . Por este isomorfismo F_n es un álgebra asociativa sobre F (como puede también verificarse directamente). Llamamos a F_n el álgebra de todas las matrices $n \times n$ sobre F .

Toda base de V nos provee de un isomorfismo de álgebras de $A(V)$ sobre F_n . Es un teorema que todo isomorfismo de álgebras de $A(V)$ sobre F_n es obtenible de tal forma.

A la luz de la misma naturaleza específica del isomorfismo entre $A(V)$ y F_n identificaremos a menudo una transformación lineal con su matriz, en alguna base, y $A(V)$ con F_n . En realidad, F_n puede considerarse como $A(V)$ actuando sobre el espacio vectorial $V = F^{(n)}$ de todos los n -tuples sobre F , donde para la base $v_1 = (1, 0, \dots, 0)$, $v_2 = (0, 1, 0, \dots, 0)$, ..., $v_n = (0, 0, \dots, 0, 1)$, $(\alpha_{ij}) \in F_n$ actúa como $v_i(\alpha_{ij}) = i$ -ésima fila de (α_{ij}) .

Resumimos lo que se ha hecho en el siguiente

TEOREMA 6.G. *El conjunto de todas las matrices $n \times n$ sobre F forma un álgebra asociativa F_n sobre F . Si V es un espacio vectorial de dimensión n sobre F , entonces $A(V)$ y F_n son isomorfos como álgebras sobre F . Dada una base cualquiera v_1, \dots, v_n de V sobre F , si para $T \in A(V)$, $m(T)$ es la matriz de T en la base v_1, \dots, v_n , la aplicación $T \rightarrow m(T)$ nos proporciona un isomorfismo de álgebras de $A(V)$ sobre F_n .*

El cero respecto a la adición en F_n es la *matriz cero* todas cuyas entradas son cero; a menudo la representaremos simplemente por 0. La *matriz uno*, que es el elemento unitario de F_n respecto a la multiplicación, es la matriz cuyas entradas están en la diagonal 1 y fuera de la diagonal 0; la representaremos por I , I_n (cuando queramos enfatizar las dimensiones de las matrices) o simplemente como 1. Para $\alpha \in F$, las matrices

$$\alpha I = \begin{pmatrix} \alpha & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \alpha \end{pmatrix}$$

(los espacios en blanco indican solamente entradas iguales a 0) se llaman *matrices escalares*. Por el isomorfismo entre $A(V)$ y F_n , es claro que $T \in A(V)$ es invertible si y sólo si $m(T)$, como matriz, tiene inversa en F_n .

Dada una transformación lineal $T \in A(V)$, si escogemos dos bases v_1, \dots, v_n y w_1, \dots, w_n de V sobre F , cada una da lugar a una matriz, a saber, $m_1(T)$ y $m_2(T)$, las matrices de T en las bases v_1, \dots, v_n y w_1, \dots, w_n , respectivamente. Como matrices, es decir, como elementos del álgebra de matrices F_n , ¿qué relación hay entre $m_1(T)$ y $m_2(T)$?

TEOREMA 6.H. *Si V es de dimensión n sobre F y si $T \in A(V)$ tiene la matriz $m_1(T)$ en la base v_1, \dots, v_n y la matriz $m_2(T)$ en la base w_1, \dots, w_n de V*

(ambas sobre F), entonces hay un elemento $C \in F_n$ tal que $m_2(T) = Cm_1(T)C^{-1}$. En realidad, si S es la transformación lineal de V definida por $v_iS = w_i$ para $i = 1, 2, \dots, n$, entonces podemos escoger como C a $m_1(S)$.

Prueba. Sea $m_1(T) = (\alpha_{ij})$ y $m_2(T) = (\beta_{ij})$; así pues $v_iT = \sum_j \alpha_{ij}v_j$, $w_iT = \sum_j \beta_{ij}w_j$.

Sea S la transformación lineal sobre V definida por $v_iS = w_i$. Como v_1, \dots, v_n y w_1, \dots, w_n son bases de V sobre F , S transforma V sobre V , de donde, según el teorema 6.d, S es invertible en $A(V)$.

Ahora bien, $w_iT = \sum_j \beta_{ij}w_j$; como $w_i = v_iS$, al sustituir esto en la expresión para w_iT obtenemos $(v_iS)T = \sum_j \beta_{ij}(v_jS)$. Pero entonces $v_i(ST) = (\sum_j \beta_{ij}v_j)S$; como S es invertible, esto se simplifica hasta obtener $v_i(STS^{-1}) = \sum_j \beta_{ij}v_j$. Por la misma definición de matriz de una transformación lineal en unas bases dadas, $m_1(STS^{-1}) = (\beta_{ij}) = m_2(T)$. Pero la aplicación $T \rightarrow m_1(T)$ es un isomorfismo de $A(V)$ sobre F_n ; por tanto, $m_1(STS^{-1}) = m_1(S)m_1(T)m_1(S^{-1}) = m_1(S)m_1(T)m_1(S)^{-1}$. Reuniendo todo lo que hemos estado estudiando, obtenemos $m_2(T) = m_1(S)m_1(T)m_1(S)^{-1}$, que es exactamente lo que se afirma en el teorema.

Ilustramos este último teorema con el ejemplo de la matriz de D que antes estudiamos, en varias bases. Para minimizar el cálculo, suponemos que V es el espacio vectorial de todos los polinomios sobre F de grado 3 o menor, y D será, como antes, el operador diferencial definido por $(\alpha_0 + \alpha_1x + \alpha_2x^2 + \alpha_3x^3)D = \alpha_1 + 2\alpha_2x + 3\alpha_3x^2$.

Como anteriormente vimos, en la base $v_1 = 1, v_2 = x, v_3 = x^2$ y $v_4 = x^3$, la matriz D es

$$m_1(D) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{pmatrix}$$

En la base $u_1 = 1, u_2 = 1+x, u_3 = 1+x^2, u_4 = 1+x^3$, la matriz de D es

$$m_2(D) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ -2 & 2 & 0 & 0 \\ -3 & 0 & 3 & 0 \end{pmatrix}$$

Sea S la transformación lineal de V definida por $v_1S = w_1 (= v_1)$, $v_2S = w_2 = 1+x = v_1+v_2$, $v_3S = w_3 = 1+x^2 = v_1+v_3$ y además $v_4S = w_4 = 1+x^3 = v_1+v_4$. La matriz de S en la base v_1, v_2, v_3, v_4 es

$$C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}$$

Un simple cálculo muestra que

$$C^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}$$

Entonces

$$\begin{aligned} Cm_1(D)C^{-1} &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ -2 & 2 & 0 & 0 \\ -3 & 0 & 3 & 0 \end{pmatrix} = m_2(D), \end{aligned}$$

como debía ser, de acuerdo con el teorema. (Verifíquense todos los cálculos usados.)

El teorema afirma que, si conocemos la matriz de una transformación lineal en una base cualquiera, podemos calcularla en cualquier otra base, siempre que conozcamos la transformación lineal (o matriz) del cambio de base.

Aún no hemos contestado la pregunta: dada una transformación lineal, ¿cómo se calculan sus raíces características? Esto llegará un poco más tarde. Partiendo de la matriz de una transformación lineal mostraremos

cómo construir un polinomio cuyas raíces sean precisamente las raíces características de la transformación lineal.

Problemas

1. Calcúlense los siguientes productos de matrices:

$$a) \begin{pmatrix} 1 & 2 & 3 \\ 1 & -1 & 2 \\ 3 & 4 & 5 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 3 \\ -1 & -1 & -1 \end{pmatrix}$$

$$b) \begin{pmatrix} 1 & 6 \\ -6 & 1 \end{pmatrix} \begin{pmatrix} 3 & -2 \\ 2 & 3 \end{pmatrix}$$

$$c) \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}^2$$

$$d) \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}^2$$

2. Verifíquense todos los cálculos hechos en el ejemplo que ilustra el teorema 6.h.

3. Pruébese directamente en F_n , usando las definiciones de suma y producto, que

$$a) A(B+C) = AB+AC;$$

$$b) (AB)C = A(BC);$$

para A , B y C pertenecientes a F_n .

4. Pruébese en F_2 para cualesquiera dos elementos A y B , que $(AB-BA)^2$ es una matriz escalar.

5. Sea V el espacio vectorial de los polinomios de grado menor o igual que 3 sobre F . Defínase T en V por $(\alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3)T = \alpha_0 + \alpha_1(x+1) + \alpha_2(x+1)^2 + \alpha_3(x+1)^3$. Calcúlese la matriz de T en las bases:

$$a) 1, x, x^2, x^3.$$

$$b) 1, 1+x, 1+x^2, 1+x^3.$$

c) Si la matriz de la parte (a) es A y la en parte (b) es B , encuéntrese una matriz C tal que $B = CAC^{-1}$.

6. Sea $V = F^{(3)}$ y supongamos que

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

es la matriz de $T \in A(V)$ en la base $v_1 = (1, 0, 0)$, $v_2 = (0, 1, 0)$ y $v_3 = (0, 0, 1)$. Encuéntrese la matriz de T en las bases:

- a) $u_1 = (1, 1, 1)$, $u_2 = (0, 1, 1)$, $u_3 = (0, 0, 1)$.
 b) $u_1 = (1, 1, 0)$, $u_2 = (1, 2, 0)$, $u_3 = (1, 2, 1)$.

7. Pruébese que dada la matriz

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 6 & -11 & 6 \end{pmatrix} \in F_3$$

(donde la característica de F no es 2), entonces:

- a) $A^3 - 6A^2 + 11A - 6 = 0$.
 b) Existe una matriz $C \in F_3$ tal que

$$CAC^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

8. Pruébese que es imposible encontrar una matriz $C \in F_2$ tal que

$$C \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} C^{-1} = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix},$$

para cualesquiera $\alpha, \beta \in F$.

9. Una matriz $A \in F_n$ se dice que es una matriz diagonal si todas las entradas fuera de la diagonal principal de A son 0, es decir, si $A = (\alpha_{ij})$ y $\alpha_{ij} = 0$ para $i \neq j$. Si A es una matriz diagonal tal que sus entradas sobre la diagonal principal son todas distintas, encuéntrense todas las matrices $B \in F_n$ que comutan con A , es decir, encuéntrense todas las matrices B tales que $BA = AB$.

10. Usando el resultado del problema 9, pruébese que solo las matrices en F_n que comutan con todas las matrices de F_n son matrices escalares.

11. Sea $A \in F_n$ la matriz

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & & \vdots & & \ddots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

todas cuyas entradas, excepto las de la superdiagonal, son 0, y cuyas entradas sobre la superdiagonal son todas iguales a 1. Pruébese que $A^n = 0$ pero $A^{n-1} \neq 0$.

*12. Si A es como en el problema 11, encuéntrense todas las matrices en F_n que conmutan con A y demuéstrese que deben ser de la forma $\alpha_0 + \alpha_1 A + \alpha_2 A^2 + \dots + \alpha_{n-1} A^{n-1}$ donde $\alpha_0, \alpha_1, \dots, \alpha_{n-1} \in F$.

13. Sea $A \in F_2$ y sea $C(A) = \{B \in F_2 \mid AB = BA\}$. Sea $C(C(A)) = \{G \in F_2 \mid GX = XG \text{ para todo } X \in C(A)\}$. Pruébese que si $G \in C(C(A))$ entonces G es de la forma $\alpha_0 + \alpha_1 A$, donde $\alpha_0, \alpha_1 \in F$.

14. Resuélvase el problema 13 para $A \in F_3$ probando que toda $G \in C(C(A))$ es de la forma $\alpha_0 + \alpha_1 A + \alpha_2 A^2$.

15. Definamos las matrices E_{ij} en F_n como sigue: E_{ij} es la matriz cuya única entrada distinta de cero es la (i, j) que es igual a 1. Pruébese que:

- a) Las E_{ij} forman una base de F_n sobre F .
- b) $E_{ij}E_{kl} = 0$ para $j \neq k$; $E_{ij}E_{jl} = E_{il}$.
- c) Dadas i y j , existe una matriz C tal que $CE_{ii}C^{-1} = E_{jj}$.
- d) Si $i \neq j$, existe una matriz C tal que $CE_{ij}C^{-1} = E_{12}$.
- e) Encuéntrense todas las $B \in F_n$ que conmutan con E_{12} .
- f) Encuéntrense todas las $B \in F_n$ que conmutan con E_{11} .

16. Sea F el campo de los números reales y sea C el campo de los números complejos. Para $a \in C$ sea $T_a: C \rightarrow C$ dada por $xT_a = xa$, para todo $x \in C$. Usando la base 1, i encuéntrese la matriz de la transformación lineal T_a y obténgase así una representación isomórfica de los números complejos como matrices 2×2 sobre el campo de los números reales.

17. Sea Q el anillo con división de los cuaternios sobre el campo real. Usando la base 1, i , j , k de Q sobre F , procédase como en el problema 16 para encontrar una representación isomórfica de Q por matrices 4×4 sobre el campo de los números reales.

*18. Combínense los resultados de los problemas 16 y 17 para encontrar una representación isomórfica de Q por matrices 2×2 sobre el campo de los números complejos.

19. Sea \mathfrak{M} el conjunto de todas las matrices $n \times n$ que tienen entradas 0 y 1 de tal forma que hay un único 1 en cada renglón y en cada columna. (Tales matrices se llaman *matrices de permutación*.)

a) Si $M \in \mathfrak{M}$ describáse AM en términos de los renglones y las columnas de A :

b) Si $M \in \mathfrak{M}$ describáse MA en términos de los renglones y las columnas de A .

20. Sea \mathfrak{M} como en el problema 19. Pruébese que :

a) \mathfrak{M} tiene $n!$ elementos.

b) Si $M \in \mathfrak{M}$, entonces es invertible y su inversa está también en \mathfrak{M} .

c) Proporcionése la forma explícita de la inversa de M .

d) Pruébese que \mathfrak{M} es un grupo respecto a la multiplicación de matrices.

e) Pruébese que \mathfrak{M} es isomorfo, como grupo, a S_n , el grupo simétrico de grado n .

21. Sea $A = (\alpha_{ij})$ tal que para todo i , $\sum_j \alpha_{ij} = 1$. Pruébese que 1 es una raíz característica de A (es decir, que $A - 1$ no es invertible).

22. Sea $A = (\alpha_{ij})$ tal que para todo j , $\sum_i \alpha_{ij} = 1$. Pruébese que 1 es una raíz característica de A .

23. Encuéntrense las condiciones necesarias y suficientes que α, β, γ y δ han de cumplir para que $A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ sea invertible. Para los casos en que A es invertible, escribáse A^{-1} explicitamente.

24. Si $E \in F_n$ es tal que $E^2 = E \neq 0$, pruébese que hay una matriz $C \in F_n$ tal que

$$CEC^{-1} = \left(\begin{array}{cccc|ccc} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdot & & \cdot \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ \hline 0 & \cdots & 0 & | & 0 & \cdots & 0 \\ 0 & \cdots & 0 & | & 0 & \cdots & 0 \end{array} \right),$$

donde la matriz unidad en la parte superior izquierda es $r \times r$, donde r es el rango de E .

25. Si F es el campo real, pruébese que es imposible encontrar matrices A, B pertenecientes a F_n tales que $AB - BA = 1$.

26. Si F es de característica 2, pruébese que en F_2 es posible encontrar matrices A, B tales que $AB - BA = 1$.

27. La matriz A se llama *triangular* si todas las entradas sobre la diagonal principal son 0. (Si todas las entradas debajo de la diagonal principal son 0 la matriz también se llama triangular.)

- a) Si A es triangular y ninguna entrada en la diagonal principal es 0, pruébese que A es invertible.
- b) Si A es triangular y una entrada en la diagonal principal es 0, pruébese que A es singular.

28. Si A es triangular, pruébese que sus raíces características son precisamente los elementos en su diagonal principal.

29. Si $N^k = 0$, $N \in F_n$, pruébese que $1 + N$ es invertible y encuéntrese su inversa como un polinomio en N .

30. Si $A \in F_n$ es triangular y todas las entradas en su diagonal principal son iguales a 0, pruébese que $A^n = 0$.

31. Si $A \in F_n$ es triangular y todas las entradas en su diagonal principal son iguales a $\alpha \neq 0 \in F$, encuéntrese A^{-1} .

32. Sean S, T transformaciones lineales sobre V tales que la matriz de S en una base es igual a la matriz de T en otra. Pruébese que existe una transformación lineal A sobre V tal que $T = ASA^{-1}$.

4. FORMAS CANÓNICAS: FORMA TRIANGULAR

Sea V un espacio vectorial n -dimensional sobre un campo F .

DEFINICIÓN. Las transformaciones lineales $S, T \in A(V)$ se dice que son *semejantes* si existe un elemento invertible $C \in A(V)$ tal que $T = CSC^{-1}$.

En vista de los resultados de la sección 3, esta definición se traduce en una acerca de las matrices. En realidad, como F_n actúa como $A(V)$ sobre $F^{(n)}$, la definición anterior define ya una semejanza entre matrices. Por ella, $A, B \in F_n$ son semejantes si existe una $C \in F_n$ invertible tal que $B = CAC^{-1}$.

La relación sobre $A(V)$ definida por la semejanza es una relación de equivalencia; la clase de equivalencia de un elemento se llamará su clase de *semejanza*. Dadas dos transformaciones lineales, ¿cómo podemos determinar si son o no semejantes? Desde luego, podíamos examinar la clase de semejanza de una de estas para ver si la otra se encuentra en ella,

pero este procedimiento no es realizable. En su lugar, intentaremos establecer alguna clase de señal en cada clase de semejanza y un camino para ir de cualquier elemento de la clase a su señal. Probaremos la existencia de transformaciones lineales en cada clase de semejanza cuya matriz, en alguna base, es de una forma particularmente conveniente. Estas matrices se llamarán *formas canónicas*. Para determinar si dos transformaciones lineales son semejantes no necesitaremos otra cosa que calcular una forma canónica particular para cada una y comprobar si estas son las mismas.

Hay muchas posibles formas canónicas; solo consideraremos nosotros tres de éstas, a saber, la forma triangular, la forma de Jordan y la forma canónica racional, en ésta y las siguientes dos secciones.

DEFINICIÓN. El subespacio W de V es *invariante bajo $T \in A(V)$* si $WT \subset W$.

LEMA 6.6. Si $W \subset V$ es invariante bajo T , entonces T induce una transformación lineal \bar{T} en V/W definida por $(v+W)\bar{T} = vT+W$. Si T satisface el polinomio $q(x) \in F[x]$, entonces también lo satisface \bar{T} . Si $p_1(x)$ es el polinomio mínimo para \bar{T} sobre F y si $p(x)$ es el polinomio mínimo para T , entonces $p_1(x) \mid p(x)$.

Prueba. Sea $\bar{V} = V/W$; los elementos de \bar{V} son, por supuesto, las clases laterales $v+W$ de W en V . Dados $\bar{v} = v+W \in \bar{V}$ definimos $\bar{v}\bar{T} = vT+W$. Verificar que \bar{T} tiene todas las propiedades formales de una transformación lineal sobre \bar{V} es una fácil tarea una vez que se ha establecido que \bar{T} está bien definida sobre \bar{V} . Nos contentaremos, pues, con probar este hecho.

Supongamos que $\bar{v} = v_1 + W = v_2 + W$ donde $v_1, v_2 \in V$. Debemos probar que $v_1 T + W = v_2 T + W$. Como $v_1 + W = v_2 + W$, $v_1 - v_2$ debe estar en W , y como W es invariante bajo T , $(v_1 - v_2)T$ debe estar también en W . Por consiguiente $v_1 T - v_2 T \in W$, de donde se sigue que $v_1 T + W = v_2 T + W$, como queríamos probar. Sabemos ahora que \bar{T} define una transformación lineal sobre $\bar{V} = V/W$.

Si $\bar{v} = v+W \in \bar{V}$, entonces $\bar{v}(\bar{T}^2) = vT^2 + W = (vT)T + W = (vT + W)\bar{T} = ((v+W)\bar{T})\bar{T} = \bar{v}(\bar{T})^2$; así pues $\bar{T}^2 = (\bar{T})^2$. Análogamente $\bar{T}^k = (\bar{T})^k$ para cualquier $k \geq 0$. Por consiguiente, para cualquier polinomio $q(x) \in F[x]$, $\bar{q}(\bar{T}) = q(\bar{T})$. Para cualquier $q(x) \in F[x]$ con $q(T) = 0$, como 0 es la transformación 0 sobre \bar{V} , $0 = \bar{q}(\bar{T}) = q(\bar{T})$.

Sea $p_1(x)$ el polinomio mínimo sobre F satisfecho por \bar{T} . Si $q(\bar{T}) = 0$ para $q(x) \in F[x]$, entonces $p_1(x) \mid q(x)$. Si $p(x)$ es el polinomio mínimo para T sobre F , entonces $p(T) = 0$, de donde $p(\bar{T}) = 0$; en consecuencia, $p_1(x) \mid p(x)$.

Como vimos en el teorema 6.f, todas las raíces características de T que se encuentran en F son raíces del polinomio mínimo de T sobre F . Decimos

que todas las raíces características de T están en F si todas las raíces del polinomio mínimo de T sobre F se encuentran en F .

En el problema 27 al final de la última sección, definimos como matriz *triangular* a toda aquella que tenga todas sus entradas sobre la diagonal principal iguales a 0. O lo que es lo mismo, si T es una transformación lineal de V sobre F , la matriz de T en la base v_1, \dots, v_n es triangular si

$$\begin{aligned}v_1 T &= \alpha_{11} v_1 \\v_2 T &= \alpha_{21} v_1 + \alpha_{22} v_2 \\&\vdots \\v_i T &= \alpha_{i1} v_1 + \alpha_{i2} v_2 + \dots + \alpha_{ii} v_i,\end{aligned}$$

es decir, si $v_i T$ es una combinación lineal solamente de v_i y sus predecesores en la base.

TEOREMA 6.J. Si $T \in A(V)$ tiene todas sus raíces características en F , entonces hay una base de V en que la matriz de T es triangular.

Prueba. La prueba se hace por inducción sobre la dimensión de V sobre F .

Si $\dim_F V = 1$ entonces todo elemento en $A(V)$ es un escalar y, por tanto, para tal caso el teorema es cierto.

Supongamos que el teorema es cierto para todos los espacios vectoriales sobre F de dimensión $n - 1$, y sea V de dimensión n sobre F .

La transformación lineal T sobre V tiene todas sus raíces características en F ; sea $\lambda_1 \in F$ una raíz característica de T . Existe en V un vector v_1 distinto de cero tal que $v_1 T = \lambda_1 v_1$. Sea $W = \{\alpha v_1 \mid \alpha \in F\}$; W es un subespacio unidimensional de V , y es invariante bajo T . Sea $\bar{V} = V/W$; por el lema 4.8, $\dim \bar{V} = \dim V - \dim W = n - 1$. De acuerdo con el lema 6.6, T induce una transformación lineal \bar{T} sobre \bar{V} cuyo polinomio mínimo sobre F divide al polinomio mínimo de T sobre F . Así pues, todas las raíces del polinomio mínimo de \bar{T} son raíces del polinomio mínimo de T , deben encontrarse en F . La transformación lineal \bar{T} en su acción sobre \bar{V} satisface la hipótesis del teorema; como \bar{V} es $(n - 1)$ -dimensional sobre F , por nuestra hipótesis de inducción, existe una base $\bar{v}_2, \bar{v}_3, \dots, \bar{v}_n$ de \bar{V} sobre F tal que:

$$\begin{aligned}\bar{v}_2 \bar{T} &= \alpha_{22} \bar{v}_2 \\ \bar{v}_3 \bar{T} &= \alpha_{32} \bar{v}_2 + \alpha_{33} \bar{v}_3 \\ &\vdots \\ \bar{v}_i \bar{T} &= \alpha_{i2} \bar{v}_2 + \alpha_{i3} \bar{v}_3 + \dots + \alpha_{ii} \bar{v}_i \\ &\vdots \\ \bar{v}_n \bar{T} &= \alpha_{n2} \bar{v}_2 + \dots + \alpha_{nn} \bar{v}_n.\end{aligned}$$

Sean v_2, \dots, v_n elementos de V que se transforman en $\bar{v}_2, \dots, \bar{v}_n$, respectivamente. Entonces v_1, v_2, \dots, v_n forman una base de V (ver el problema 3

al final de esta sección). Como $\bar{v}_2 T = \alpha_{22} \bar{v}_2$, $\bar{v}_2 T - \alpha_{22} \bar{v}_2 = 0$, de donde $v_2 T - \alpha_{22} v_2$ deben estar en W . Así pues, $v_2 T - \alpha_{22} v_2$ es un múltiplo de v_1 , digamos $\alpha_{21} v_1$, de donde tenemos, después de trasponer, $v_2 T = \alpha_{21} v_1 + \alpha_{22} v_2$. Análogamente, $v_i T - \alpha_{i2} v_2 - \alpha_{i3} v_3 - \dots - \alpha_{ii} v_i \in W$, de donde $v_i T = \alpha_{i1} v_1 + \alpha_{i2} v_2 + \dots + \alpha_{ii} v_i$. La base v_1, \dots, v_n de V sobre F nos proporciona una base respecto a la cual todo $v_i T$ es una combinación lineal de v_i y sus predecesores en la base. Por lo tanto, la matriz de T en esta base es triangular. Esto completa la inducción y prueba el teorema.

Queremos reformular el teorema 6.j para matrices. Supongamos que la matriz $A \in F_n$ tiene sus raíces características en F . A define una transformación lineal T sobre F^n cuya matriz en la base

$$v_1 = (1, 0, \dots, 0), v_2 = (0, 1, 0, \dots, 0), \dots, v_n = (0, 0, \dots, 0, 1),$$

es precisamente A . Las raíces características de T , siendo iguales a las de A , están todas en F , de donde, según el teorema 6.j, hay una base en $F^{(n)}$ en la que la matriz de T es triangular. Pero, de acuerdo con el teorema 6.h, este cambio de base varía simplemente la matriz de T , es decir, la A , en la primera base, en CAC^{-1} para una C adecuada $C \subset F_n$. Así pues

FORMA ALTERNADA DEL TEOREMA 6.J. *Si la matriz $A \in F_n$ tiene todas sus raíces características en F , entonces hay una matriz $C \in F_n$ tal que CAC^{-1} es una matriz triangular.*

El teorema 6.j (en cualquiera de sus formas) se describe usualmente diciendo que T (o A) puede ser llevada a una forma triangular sobre F .

Si volvemos nuestra mirada al problema 28, al final de la sección 3, veremos que después de que T se ha llevado a la forma triangular, los elementos de la diagonal principal de su matriz juegan el siguiente significativo papel: *son precisamente las raíces características de T .*

Concluimos la sección con el

TEOREMA 6.K. *Si V es n -dimensional sobre F y si $T \in A(V)$ tiene todas sus raíces características en F , entonces T satisface un polinomio de grado n sobre F .*

Prueba. De acuerdo con el teorema 6.j, podemos encontrar una base v_1, \dots, v_n de V sobre F tal que:

$$v_1 T = \lambda_1 v_1$$

$$v_2 T = \alpha_{21} v_1 + \lambda_2 v_2$$

⋮

$$v_i T = \alpha_{i1} v_1 + \dots + \alpha_{i, i-1} v_{i-1} + \lambda_i v_i$$

para $i = 1, 2, \dots, n$.

O lo que es equivalente:

$$\begin{aligned}v_1(T - \lambda_1) &= 0 \\v_2(T - \lambda_2) &= \alpha_{21} v_1 \\&\vdots \\v_i(T - \lambda_i) &= \alpha_{i1} v_1 + \dots + \alpha_{i, i-1} v_{i-1}\end{aligned}$$

para $i = 1, 2, \dots, n$.

¿Qué es $v_2(T - \lambda_2)(T - \lambda_1)$? Como resultado de $v_2(T - \lambda_2) = \alpha_{21} v_1$ y $v_1(T - \lambda_1) = 0$, obtenemos $v_2(T - \lambda_2)(T - \lambda_1) = 0$. Como

$$\begin{aligned}(T - \lambda_2)(T - \lambda_1) &= (T - \lambda_1)(T - \lambda_2), \\v_1(T - \lambda_2)(T - \lambda_1) &= v_1(T - \lambda_1)(T - \lambda_2) = 0.\end{aligned}$$

La continuación de este tipo de cálculo nos lleva a:

$$\begin{aligned}v_1(T - \lambda_i)(T - \lambda_{i-1}) \dots (T - \lambda_1) &= 0, \\v_2(T - \lambda_i)(T - \lambda_{i-1}) \dots (T - \lambda_1) &= 0, \\&\dots, \\v_i(T - \lambda_i)(T - \lambda_{i-1}) \dots (T - \lambda_1) &= 0.\end{aligned}$$

En particular, para $i = n$, la matriz $S = (T - \lambda_n)(T - \lambda_{n-1}) \dots (T - \lambda_1)$ satisface $v_1 S = v_2 S = \dots = v_n S = 0$. Como S suprime una base de V , S tiene que suprimir también a todo V . Por lo tanto, $S = 0$. Por consiguiente, T satisface el polinomio $(x - \lambda_1)(x - \lambda_2) \dots (x - \lambda_n)$ en $F[x]$ de grado n , con lo que el teorema queda probado.

Desgraciadamente está en la naturaleza de las cosas que no toda transformación lineal sobre un espacio vectorial sobre todo campo F tenga todas sus raíces características en F . Que tal ocurra depende totalmente del campo F . Por ejemplo, si F es el campo de los números reales, entonces la ecuación mínima de

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

sobre F es $x^2 + 1$ que no tiene raíz alguna sobre F . No tenemos, pues, ningún derecho a suponer que las raíces características se encuentren siempre en el campo en cuestión. Pero, podemos preguntarnos, ¿no podemos ampliar ligeramente F hasta un nuevo campo K de modo que todo trabaje muy bien sobre K ?

Haremos la discusión para matrices; lo mismo podría hacerse para transformaciones lineales. Lo que se necesitaría sería lo siguiente: dado un espacio vectorial V sobre un campo F de dimensión n , y dada una extensión K de F , entonces podemos sumergir V en un espacio vectorial V_K sobre K

de dimensión n sobre K . Una forma de hacer esto sería tomar una base v_1, \dots, v_n de V sobre F y considerar V_K como el conjunto de todos los $\alpha_1 v_1 + \dots + \alpha_n v_n$ con las $\alpha_i \in K$, considerando las v_i linealmente independientes sobre K . Este pesado uso de una base es antiestético; todo puede hacerse de modo independiente de toda base si introducimos el concepto de *producto tensorial* de espacios vectoriales. No lo haremos aquí; en su lugar argumentaremos con matrices (lo que es efectivamente el camino delineado anteriormente usando una base fija de V).

Consideremos el álgebra F_n . Si K es cualquier extensión del campo de F , entonces $F_n \subset K_n$, el conjunto de las matrices $n \times n$ sobre K . Así pues, cualquier matriz sobre el campo F puede considerarse como una matriz sobre K . Si $T \in F_n$ tiene el polinomio mínimo $p(x)$ sobre F , considerada como un elemento de K_n puede conceiblemente satisfacer a un polinomio diferente $p_0(x)$ sobre K . Pero entonces $p_0(x) | p(x)$, ya que $p_0(x)$ divide a todos los polinomios sobre K (y, por tanto, a todos los polinomios sobre F) que son satisfechos por T . Especializamos ahora a K . Por el teorema 5.h existe una extensión finita K , de F en la cual el polinomio mínimo $p(x)$, para T sobre F tiene todas sus raíces. Como elemento de K_n , ¿tiene T , para esta K , todas sus raíces características en K ? Como elemento de K_n el polinomio mínimo de T sobre K , $p_0(x)$, divide a $p(x)$ de modo que todas las raíces de $p_0(x)$ son raíces de $p(x)$ y, por tanto, se encuentran en K . Por consiguiente, como elemento de K_n , T tiene todas raíces características en K .

Así pues, dada T en F_n , al irnos al campo de descomposición K , de su polinomio mínimo llegamos a la situación en que las hipótesis de los teoremas 6.j y 6.k se satisfacen, no sobre F , sino sobre K . Por lo dicho, T puede, por ejemplo, ser llevada a la forma triangular sobre K y satisface un polinomio de grado n sobre K . A veces, cuando tenemos suerte, sabiendo que cierto resultado es cierto sobre K podemos limitarnos a F y saber que el resultado es también verdadero sobre F . Pero llegar hasta K no es ninguna panacea, pues hay situaciones frecuentes donde los resultados para K no implican nada para F . Es por esto por lo que tenemos dos tipos de teoremas de "formas canónicas", aquellos en que se supone en que todas las raíces características de T se encuentran en F y aquellos en que no se hace tal supuesto.

Una palabra final; si $T \in F_n$, por la frase "*una raíz característica de T* " entenderemos un elemento λ del campo de descomposición K del polinomio mínimo $p(x)$ de T sobre F tal que $\lambda - T$ no es invertible en K_n . Es un hecho (véase el problema 5) que toda raíz del polinomio mínimo de T sobre F es es una raíz característica de T .

Problemas

1. Pruébese que la relación de semejanza es una relación de equivalencia en $A(V)$.

2. Si $T \in F_n$ y si $K \supset F$, pruébese que como un elemento de K_n , T es invertible si y sólo si es ya invertible en F_n .

3. En la prueba del teorema 6.j pruébese que v_1, \dots, v_n es una base de V .

4. Proporciónese una prueba, usando cálculo matricial, que si A es una matriz triangular $n \times n$ con entradas $\lambda_1, \dots, \lambda_n$ sobre la diagonal, entonces

$$(A - \lambda_1)(A - \lambda_2) \dots (A - \lambda_n) = 0.$$

*5. Si $T \in F_n$ tiene $p(x)$ como polinomio mínimo sobre F , pruébese que toda raíz de $p(x)$ en su campo de descomposición K , es una raíz característica de T .

6. Si $T \in A(V)$ y si $\lambda \in F$ es una raíz característica de T en F , sea $U_\lambda = \{v \in V \mid vT = \lambda v\}$. Si $S \in A(V)$ comuta con T , pruébese que U_λ es invariante bajo S .

*7. Si \mathfrak{M} es un conjunto comutativo de elementos en $A(V)$ tales que toda $M \in \mathfrak{M}$ tiene todas sus raíces características en F , pruébese que hay un $C \in A(V)$ tal que toda CMC^{-1} , para $M \in \mathfrak{M}$ está en forma triangular.

8. Sea W un subespacio de V invariante bajo $T \in A(V)$. Cuando restringimos T a W , T induce una transformación lineal \bar{T} (definida por $w\bar{T} = wT$ para toda $w \in W$). Sea $\tilde{p}(x)$ el polinomio mínimo de \bar{T} sobre F .

- a) Pruébese que $\tilde{p}(x) \mid p(x)$, el polinomio mínimo de T sobre F .
- b) Si T induce \bar{T} sobre V/W , con \bar{T} satisfaciendo el polinomio mínimo $\bar{p}(x)$ sobre F , pruébese que $p(x) \mid \tilde{p}(x)\bar{p}(x)$.
- *c) Si $\tilde{p}(x)$ y $\bar{p}(x)$ son primos relativos, pruébese que $p(x) = \tilde{p}(x)\bar{p}(x)$.
- *d) Proporciónese un ejemplo de un T para el que $p(x) \neq \tilde{p}(x)\bar{p}(x)$.

9. Sea \mathfrak{M} un conjunto no vacío de elementos en $A(V)$; el subespacio $W \subset V$ se dice que es *invariante bajo* \mathfrak{M} si para todo $M \in \mathfrak{M}$, $WM \subset W$. Si W es invariante bajo \mathfrak{M} y es de dimensión r sobre F , pruébese que existe una base de V sobre F tal que todo $M \in \mathfrak{M}$ tiene una matriz, en esta base, de la forma

$$\left(\begin{array}{c|c} M_1 & 0 \\ \hline M_{12} & M_2 \end{array} \right)$$

donde M_1 es una matriz $r \times r$ y M_2 es una matriz $(n-r) \times (n-r)$.

10. En el problema 9 probamos que M_1 es la matriz de una transformación \tilde{M} inducida por M sobre W , y que M_2 es la matriz de la transformación lineal \bar{M} inducida por M en V/W .

*11. El conjunto no vacío \mathfrak{M} de transformaciones lineales en $A(V)$ se llama conjunto *irreducible* si los subespacios de V invariantes bajo \mathfrak{M} son

(0) y V . Si \mathfrak{M} es un conjunto irreducible de transformaciones lineales sobre V y si

$$D = \{T \in A(V) \mid TM = MT \text{ para toda } M \in \mathfrak{M}\},$$

pruébese que D es un anillo con división.

*12. Resuélvase el problema 11 usando el resultado (lema de Schur) del problema 14, final del capítulo 4.

*13. Si F es tal que todos los elementos de $A(V)$ tienen todas sus raíces características en F , pruébese que el D del problema 11 consiste solamente en escalares.

14. Sea F el campo de los números reales y sea

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in F_2.$$

a) Pruébese que el conjunto \mathfrak{M} consiste solamente en

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

es un conjunto irreducible.

b) Encuéntrese el conjunto D de todas las matrices que comutan con

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

y pruébese que D es isomorfo al campo de los números complejos.

15. Sea F el campo de los números reales.

a) Pruébese que el conjunto

$$\mathfrak{M} = \left\{ \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \right\}$$

es un conjunto irreducible.

b) Encuéntrense todas las $A \in F_4$ tales que $AM = MA$ para toda $M \in \mathfrak{M}$.

c) Pruébese que el conjunto de todas las A de la parte (b) es un anillo con división isomorfo al anillo con división de los cuaternios sobre el campo real.

16. Un conjunto de transformaciones lineales, $\mathfrak{M} \subset A(V)$, se llama *descomponible* si hay un subespacio $W \subset V$ tal que $V = W \oplus W_1$, $W \neq (0)$, $W \neq V$, y tanto W como W_1 son invariantes respecto a \mathfrak{M} . Si \mathfrak{M} no es descomponible se llama *indescomponible*.

- a) Si \mathfrak{M} es un conjunto descomponible de transformaciones lineales sobre V , pruébese que hay una base de V en que todo $M \in \mathfrak{M}$ tiene una matriz de la forma

$$\left(\begin{array}{c|c} M_1 & 0 \\ \hline 0 & M_2 \end{array} \right)$$

donde M_1 y M_2 son matrices cuadradas.

- b) Si V es un espacio vectorial n -dimensional sobre F y si $T \in A(V)$ satisface $T^n = 0$, pero $T^{n-1} \neq 0$, pruébese que el conjunto $\{T\}$ (consistente en T) es indescomponible.

17. Sea $T \in A(V)$ y supongamos que $p(x)$ es el polinomio mínimo para T sobre F .

- a) Si $p(x)$ es divisible por dos distintos polinomios irreducibles $p_1(x)$ y $p_2(x)$ en $F[x]$, pruébese que $\{T\}$ es descomponible.
 b) Si para algún $T \in A(V)$ es descomponible $\{T\}$, pruébese que el polinomio mínimo para T sobre F es la potencia de un polinomio irreducible.

18. Si $T \in A(V)$ es nilpotente, pruébese que T puede ser puesto en forma triangular sobre F y en esa forma todos los elementos de la diagonal son 0.

19. Si $T \in A(V)$ tiene solamente 0 como una raíz característica, pruébese que T es nilpotente.

5. FORMAS CANÓNICAS: TRANSFORMACIONES NILPOTENTES

Una clase de transformaciones lineales que tienen todas sus raíces características en F es la clase de las nilpotentes, pues como todas sus raíces características son 0, es evidente que todas están en F . Por tanto, por el resultado de la sección previa, una transformación lineal nilpotente puede siempre ser puesta en forma triangular sobre F . Para algunos propósitos, esto no es suficientemente agudo, y como veremos pronto, puede decirse bastante más.

Aunque la clase de las transformaciones lineales nilpotentes es bastante restringida, la verdad es que merece un estudio solo por sus propios méritos. Pero, lo que aún es más importante para nuestros propósitos, una vez que

hemos encontrado una buena forma canónica para ellas, nos es fácil encontrar una buena forma canónica para todas las transformaciones lineales que tienen todas sus raíces características en F .

Una palabra acerca del método que seguiremos en esta sección. Podríamos estudiar estos problemas "básicos" o podríamos basarnos en los resultados acerca de la descomposición de módulos que obtuvimos en el capítulo 4. Nos hemos decidido por un compromiso entre ambas posibilidades; estudiaremos el material en esta sección y en la siguiente (sobre formas de Jordan) independientemente de la noción de módulo y los resultados acerca de módulos desarrollados en el capítulo 4. Pero en la sección que trata de la forma canónica racional cambiaremos completamente de punto de vista, introduciendo por medio de una transformación lineal dada una estructura de módulo sobre el espacio vectorial bajo discusión; haciendo uso del teorema 4.j tendremos, entonces, una descomposición de un espacio vectorial, y la forma canónica resultante correspondiente a una transformación lineal dada.

Incluso aunque no usemos un enfoque basado en la teoría de módulos, por ahora, el lector debe darse cuenta de la analogía entre los argumentos usados en la prueba del teorema 4.j con los utilizados para probar el lema 6.10.

Antes de concentrar nuestros esfuerzos sobre transformaciones nilpotentes probemos un resultado de interés que verifica transformaciones lineales cualesquiera.

LEMA 6.7. Si $V = V_1 \oplus V_2 \oplus \dots \oplus V_k$, donde cada espacio V_i es de dimensión n_i y es invariante bajo T , un elemento de $A(V)$, entonces puede encontrarse una base de V tal que la matriz de T en esta base sea de la forma

$$\begin{pmatrix} A_1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & A_2 & \cdot & \cdot & \cdot & 0 \\ \vdots & \vdots & \cdot & \cdot & \cdot & \vdots \\ 0 & 0 & \cdot & \cdot & \cdot & A_k \end{pmatrix}$$

donde cada A_i es una matriz $n_i \times n_i$ y es la matriz de la transformación lineal inducida por T sobre V_i .

Prueba. Escojamos una base de V como sigue: $v_1^{(1)}, \dots, v_{n_1}^{(1)}$ es una base de V_1 , $v_1^{(2)}, \dots, v_{n_2}^{(2)}$ es una base de V_2 , y así sucesivamente. Como cada V_i es invariante bajo T , $v_j^{(i)} T \in V_i$, luego es una combinación lineal de $v_1^{(i)}, \dots, v_{n_i}^{(i)}$, y solamente de ellos. Así pues, la matriz de T en la base así escogida es de la forma deseada. Que cada A_i es la matriz de T_i , la transformación lineal inducida sobre V_i por T , es claro por la misma definición de matriz de una transformación lineal.

Limitamos ahora nuestra atención a las transformaciones nilpotentes.

LEMÁ 6.8. Si $T \in A(V)$ es nilpotente, entonces $\alpha_0 + \alpha_1 T + \dots + \alpha_m T^m$, donde las $\alpha_i \in F$, es invertible si $\alpha_0 \neq 0$.

Prueba. Si S es nilpotente y $\alpha_0 \neq 0 \in F$, un simple cálculo muestra que

$$(\alpha_0 + S) \left(\frac{1}{\alpha_0} - \frac{S}{\alpha_0^2} + \frac{S^2}{\alpha_0^3} - \dots + (-1)^{r-1} \frac{S^{r-1}}{\alpha_0^r} \right) = 1,$$

si $S^r = 0$. Ahora bien, si $T^r = 0$, $S = \alpha_1 T + \alpha_2 T^2 + \dots + \alpha_m T^m$ debe también satisfacer $S^r = 0$ (pruébese). Luego para $\alpha_0 \neq 0$ en F , $\alpha_0 + S$ es invertible.

NOTACIÓN. M_t denotará la matriz $t \times t$

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & \cdots & 0 & 1 & \\ 0 & 0 & & 0 & 0 & \end{pmatrix},$$

cuyas entradas son 0, excepto en la superdiagonal donde todas son 1.

DEFINICIÓN. Si $T \in A(V)$ es nilpotente, entonces a k le llamamos *índice de nilpotencia* de T si $T^k = 0$ pero $T^{k-1} \neq 0$.

El resultado clave respecto a transformaciones nilpotentes es

TEOREMA 6.L. Si $T \in A(V)$ es nilpotente, de índice de nilpotencia n_1 , entonces puede encontrarse una base de V tal que la matriz de T en esta base tenga la forma

$$\begin{pmatrix} M_{n_1} & 0 & \cdots & 0 \\ 0 & M_{n_2} & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & M_{n_r} \end{pmatrix},$$

donde $n_1 \geq n_2 \geq \dots \geq n_r$ y donde $n_1 + n_2 + \dots + n_r = \dim_F V$.

Prueba. La prueba será un poco detallada, de modo que al hacerla separaremos algunas sus de partes como lemas.

Como $T^{n_1} = 0$ pero $T^{n_1-1} \neq 0$, podemos encontrar un vector $v \in V$ tal que $vT^{n_1-1} \neq 0$. Afirmamos que los vectores v, vT, \dots, vT^{n_1-1} son linealmente independientes sobre F . En efecto, supongamos que $\alpha_1 v + \alpha_2 vT + \dots + \alpha_{n_1} vT^{n_1-1} = 0$ donde las $\alpha_i \in F$; sea α_s la primera α distinta de cero. Tenemos entonces

$$vT^{s-1}(\alpha_s + \alpha_{s+1}T + \dots + \alpha_{n_1}T^{n_1-s}) = 0.$$

Como $\alpha_s \neq 0$, por el lema 6.8, $\alpha_s + \alpha_{s+1}T + \dots + \alpha_{n_1}T^{n_1-s}$ es invertible y, por tanto, $vT^{s-1} = 0$. Pero $s < n_1$, luego esto contradice que $vT^{n_1-1} \neq 0$. Luego ningún α_s distinto de cero existe y v, vT, \dots, vT^{n_1-1} se ha mostrado son linealmente independientes sobre F .

Sea V_1 el subespacio de V generado por $v_1 = v, v_2 = vT, \dots, v_{n_1} = vT^{n_1-1}$; V_1 es invariante bajo T y, en la anterior base, la transformación lineal inducida por T sobre V_1 tiene como matriz M_{n_1} .

Hasta el momento hemos producido la esquina superior izquierda de la matriz del teorema. Debemos, de alguna forma, producir el resto de esta matriz.

LEMÁ 6.9. Si $u \in V_1$ es tal que $uT^{n_1-k} = 0$, donde $0 < k \leq n_1$, entonces $u = u_0 T^k$ para algún $u_0 \in V_1$.

Prueba. Como $u \in V_1$, $u = \alpha_1 v + \alpha_2 vT + \dots + \alpha_k vT^{k-1} + \alpha_{k+1} vT^k + \dots + \alpha_{n_1} vT^{n_1-1}$. Así pues, $0 = uT^{n_1-k} = \alpha_1 vT^{n_1-k} + \dots + \alpha_k vT^{n_1-k}$. Pero $vT^{n_1-k}, \dots, vT^{n_1-1}$ son linealmente independientes sobre F , de donde $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$, y por lo tanto, $u = \alpha_{k+1} vT^k + \dots + \alpha_{n_1} vT^{n_1-1} = u_0 T^k$, donde $u_0 = \alpha_{k+1} v + \dots + \alpha_{n_1} vT^{n_1-k-1} \in V_1$.

El argumento, hasta el momento, no ha sido nada complicado. Se hace ahora un poco más denso.

LEMÁ 6.10. Existe un subespacio W de V , invariante bajo T , tal que $V = V_1 \oplus W$.

Prueba. Sea W un subespacio de V , de la mayor dimensión posible, tal que:

- 1) $V_1 \cap W = \{0\}$
- 2) W es invariante bajo T .

Queremos ahora demostrar que $V = V_1 + W$. Supongamos que así no fuera; entonces existiría un elemento $z \in V$ tal que $Z \notin V_1 + W$. Como $T^{n_1} = 0$, existe un entero k , $0 < k \leq n_1$ tal que $zT^k \in V_1 + W$ y tal que $zT^i \notin V_1 + W$ para $i < k$. Así pues, $zT^k = u + w$ donde $u \in V_1$ y $w \in W$. Pero entonces $0 = zT^{n_1} = (zT^k)T^{n_1-k} = uT^{n_1-k} + wT^{n_1-k}$; pero, como tanto V_1 como W

son invariantes bajo T , $uT^{n_1-k} \in V_1$ y $wT^{n_1-k} \in W$. Como $V_1 \cap W = (0)$ esto nos dice que $uT^{n_1-k} = -wT^{n_1-k} \in V_1 \cap W = (0)$, de donde $uT^{n_1-k} = 0$. Según el lema 6.9, $u = u_0 T^k$ para algún $u_0 \in V_1$; por tanto $zT^k = u + w = u_0 T^k + w$. Sea $z_1 = z - u_0$; entonces $z_1 T^k = zT^k - u_0 T^k = w \in W$, y como W es un invariante bajo T esto implica $z_1 T^m \in W$ para toda $m \geq k$. Por otra parte, si $i < k$, $z_1 T^i = zT^i - u_0 T^i \notin W + W$, pues de otra forma zT^i debería estar en $V_1 + W$ en contra de la elección de k .

Sea W_1 el subespacio de V generado por W y $z_1, z_1 T, \dots, z_1 T^{k-1}$. Como $z_1 \notin W$, y como $W_1 \supset W$, la dimensión de W_1 debe ser mayor que la de W . Además como $z_1 T^k \in W$ y como W es invariante bajo T , W_1 debe ser invariante bajo T . Por la naturaleza máxima de W debe haber un elemento de la forma $w_0 + \alpha_1 z_1 + \alpha_2 z_1 T + \dots + \alpha_k z_1 T^{k-1} \neq 0$ en $W_1 \cap V_1$ donde $w_0 \in W$. No todos los $\alpha_1, \dots, \alpha_k$ pueden ser cero; de otra forma tendríamos $0 \neq w_0 \in W \cap V_1 = (0)$, una contradicción. Sea α_s el primer α distinto de cero; entonces $w_0 + z_1 T^{s-1}(\alpha_s + \alpha_{s+1} T + \dots + \alpha_k T^{k-s}) \in V_1$. Como $\alpha_s \neq 0$, por lema 6.8, $\alpha_s + \alpha_{s+1} T + \dots + \alpha_k T^{k-s}$ es invertible y su inversa R , es un polinomio en T . Por tanto, W y V_1 son invariantes bajo R ; pero, por lo anterior, $w_0 R + z_1 T^{s-1} \in V_1 R \subset V_1$, lo que obliga a que $z_1 T^{s-1} \in V_1 + WR \subset V_1 + W$. Como $s-1 < k$ esto es imposible; por tanto $V_1 + W = V$. Como $V_1 \cap W = (0)$, $V = V_1 \oplus W$, y el lema queda probado.

El trabajo pesado, por el momento, se terminó; ahora vamos a completar la prueba del teorema 6.1.

Según el lema 6.10, $V = V_1 \oplus W$ donde W es invariante bajo T . Usando las bases v_1, \dots, v_{n_1} de V_1 y cualquier base de W como una base de V , por el lema 6.7, la matriz de T en esta base tiene la forma

$$\begin{pmatrix} M_{n_1} & 0 \\ 0 & A_2 \end{pmatrix}$$

donde A_2 es la matriz de T_2 , la transformación lineal inducida sobre W por T . Como $T^{n_1} = 0$, $T_2^{n_2} = 0$ para algún $n_2 \leq n_1$. Repitiendo el argumento usado para T_2 sobre W podemos descomponer W como hicimos con V (o, aplicar inducción sobre la dimensión del espacio vectorial de que tratemos). Continuando este camino obtenemos una base de V en que la matriz de T es de la forma

$$\begin{pmatrix} M_{n_1} & 0 & \cdots & 0 \\ 0 & M_{n_2} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & M_{n_r} \end{pmatrix}.$$

Que $n_1 + n_2 + \dots + n_r = \dim V$ es claro, ya que la dimensión de la matriz es $n \times n$ donde $n = \dim V$.

DEFINICIÓN. Los enteros n_1, n_2, \dots, n_r se llaman los *invariantes de T*.

DEFINICIÓN. Si $T \in A(V)$ es nilpotente, el subespacio M de V , de dimensión m , que es invariante bajo T , se llama *cíclico con respecto a T*, si:

- 1) $MT^m = (0), MT^{m-1} \neq (0);$
- 2) hay un elemento $z \in M$ tal que z, zT, \dots, zT^{m-1} forma una base de M .

(Nota: La condición (1) está realmente implicada por la condición 2.)

LEMA 6.11. Si M , de dimensión m , es cíclica con respecto a T , entonces la dimensión de MT^k es $m-k$ para todo $k \leq m$.

Prueba. Podemos obtener una base de MT^k tomando la imagen de cualquier base de M bajo T^k . Usando la base z, zT, \dots, zT^{m-1} de M obtenemos una base $zT^k, zT^{k+1}, \dots, zT^{m-1}$ de MT^k . Como esta base tiene $m-k$ elementos, el lema queda probado.

El teorema 6.1 nos dice que dado un nilpotente T en $A(V)$ podemos encontrar enteros $n_1 \geq n_2 \geq \dots \geq n_r$ y subespacios V_1, \dots, V_r de V cílicos con respecto a T y de dimensiones n_1, n_2, \dots, n_r respectivamente, ya que $V = V_1 \oplus \dots \oplus V_r$.

¿Es posible que podamos encontrar otros enteros $m_1 \geq m_2 \geq \dots \geq m_s$ y otros subespacios U_1, \dots, U_s de V , cílicos respecto a T y de dimensiones m_1, \dots, m_s , respectivamente, tales que $V = U_1 \oplus \dots \oplus U_s$? Afirmamos que no es posible o, en otras palabras, que $s = r$ y $m_1 = n_1, m_2 = n_2, \dots, m_r = n_r$. Supongamos que éste no fuera el caso; entonces habría un primer entero i tal que $m_i \neq n_i$. Podemos suponer que $m_i < n_i$.

Consideremos VT^{m_i} . Por una parte, como $V = V_1 \oplus \dots \oplus V_r$, $VT^{m_i} = V_1 T^{m_i} \oplus \dots \oplus V_r T^{m_i}$. Como $\dim V_1 T^{m_i} = n_1 - m_i$, $\dim V_2 T^{m_i} = n_2 - m_i, \dots, \dim V_r T^{m_i} = n_r - m_i$ (según lema 6.11), $\dim VT^{m_i} \geq (n_1 - m_i) + (n_2 - m_i) + \dots + (n_r - m_i)$. Por otra parte, como $V = U_1 \oplus \dots \oplus U_s$ y como $U_j T^{m_i} = (0)$ para $j \geq i$, $VT^{m_i} = U_1 T^{m_i} \oplus U_2 T^{m_i} \oplus \dots \oplus U_{i-1} T^{m_i}$. Así pues

$$\dim VT^{m_i} = (m_1 - m_i) + (m_2 - m_i) + \dots + (m_{i-1} - m_i).$$

Por nuestra elección de i , $n_1 = m_1, n_2 = m_2, \dots, n_{i-1} = m_{i-1}$, de donde

$$\dim VT^{m_i} = (n_1 - m_i) + (n_2 - m_i) + \dots + (n_{i-1} - m_i).$$

Pero esto contradice el hecho anteriormente probado de que $\dim VT^{m_i} \geq (n_1 - m_i) + \dots + (n_{i-1} - m_i) + (n_i - m_i)$, ya que $n_i - m_i > 0$.

Así pues, hay un único conjunto de enteros $n_1 \geq n_2 \geq \dots \geq n_r$ tal que V es la suma directa de subespacios cílicos con respecto a T y de dimensiones n_1, n_2, \dots, n_r . Es decir, hemos demostrado que los invariantes de T son únicos.

Matricialmente, el argumento que acabamos de mostrar ha probado que si $n_1 \geq n_2 \geq \dots \geq n_r$, y $m_1 \geq m_2 \geq \dots \geq m_s$, entonces las matrices

$$\begin{pmatrix} M_{n_1} & \cdots & 0 \\ 0 & \ddots & \vdots \\ \vdots & \ddots & \vdots \\ 0 & \cdots & M_{n_r} \end{pmatrix} \text{ y } \begin{pmatrix} M_{m_1} & \cdots & 0 \\ 0 & \ddots & \vdots \\ \vdots & \ddots & \vdots \\ 0 & \cdots & M_{m_s} \end{pmatrix}$$

son semejantes solamente si $r = s$ y $n_1 = m_1, n_2 = m_2, \dots, n_r = m_r$.

Hasta el momento hemos probado la mitad más difícil del

TEOREMA 6.M *Dos transformaciones lineales nilpotentes son semejantes si y sólo si tienen las mismas variantes.*

Prueba. La discusión que precede al teorema ha demostrado que si dos transformaciones lineales nilpotentes tienen diferentes invariantes, entonces no pueden ser semejantes, pues sus respectivas matrices

$$\begin{pmatrix} M_{n_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & M_{n_r} \end{pmatrix} \text{ y } \begin{pmatrix} M_{m_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & M_{m_s} \end{pmatrix}$$

no pueden ser semejantes.

Pasemos a comprobar la parte del teorema en la otra dirección. Si las dos transformaciones lineales nilpotentes S y T tienen los mismos invariantes $n_1 \geq \dots \geq n_r$, por el teorema 6.1 hay bases v_1, \dots, v_n y w_1, \dots, w_n de V tales que la matriz de S en v_1, \dots, v_n y la de T en w_1, \dots, w_n son, ambas, iguales a

$$\begin{pmatrix} M_{n_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & M_{n_r} \end{pmatrix}.$$

Pero si A es la transformación lineal definida sobre V por $v_i A = w_i$, entonces $S = A T A^{-1}$ (*¡pruébese!*, compárese con el problema 32 al final de la sección 3), de donde S y T son semejantes.

Calculemos un ejemplo. Supongamos que

$$T = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in F_3$$

actúa sobre F^3 con base $u_1 = (1, 0, 0)$, $u_2 = (0, 1, 0)$ y $u_3 = (0, 0, 1)$. Sea $v_1 = u_1$, $v_2 = u_1 + u_3$, $v_3 = u_3$; en la base v_1, v_2, v_3 la matriz de T es

$$\left(\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array} \right),$$

de forma que los invariantes de T son 2, 1. Si A es la matriz del cambio de base, es decir

$$\left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{array} \right).$$

un simple cálculo muestra que

$$ATA^{-1} = \left(\begin{array}{ccc} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right).$$

Una observación final: los invariantes de T determinan una partición de n , la dimensión de V . Recíprocamente, una partición de n , $n_1 \geq \dots \geq n_r$, $n_1 + n_2 + \dots + n_r = n$, determina los invariantes de la transformación lineal nilpotente

$$\left(\begin{array}{ccc} M_{n_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & M_{n_r} \end{array} \right).$$

Así pues, el número de clases distintas de semejanza de las matrices nilpotentes $n \times n$ es precisamente $p(n)$, el número de particiones de n .

6. FORMAS CANÓNICAS. UNA DESCOMPOSICIÓN DE V: FORMA DE JORDAN

Sea V un espacio vectorial de dimensión finita sobre F y sea T un elemento arbitrario de $A_F(V)$. Supongamos que V_1 es un subespacio de V invariante bajo T . Por tanto, T induce una transformación lineal T_1 sobre V_1 , definida por $uT_1 = uT$ para toda $u \in V_1$. Dado un polinomio cualquiera $p(x) \in F[x]$, afirmamos que la transformación lineal inducida por $p(T)$ sobre V_1 es precisamente $p(T_1)$. (La prueba de esto se deja como ejercicio.) En par-

ticular, si $q(T) = 0$, entonces $q(T_1) = 0$. Así pues, T_1 satisface cualquier polinomio satisfecho por T sobre F . ¿Qué podemos decir en la dirección opuesta?

LEMA 6.12. *Supongamos que $V = V_1 \oplus V_2$ donde V_1 y V_2 son subespacios de V invariantes bajo T . Sean T_1 y T_2 las transformaciones lineales inducidas por T sobre V_1 y V_2 , respectivamente. Si el polinomio mínimo de T_1 sobre F es $p_1(x)$ mientras que el de T_2 es $p_2(x)$, entonces el polinomio mínimo para T sobre F es el mínimo común múltiplo de $p_1(x)$ y $p_2(x)$.*

Prueba. Si $p(x)$ es el polinomio mínimo para T sobre F , como hemos visto antes, tanto $p(T_1)$ como $p(T_2)$ son cero, de donde $p_1(x) | p(x)$ y $p_2(x) | p(x)$. Pero entonces el mínimo común múltiplo de $p_1(x)$ y $p_2(x)$ debe también dividir a $p(x)$.

Por otra parte, si $q(x)$ es el mínimo común múltiplo de $p_1(x)$ y $p_2(x)$, consideremos $q(T)$. Para $v_1 \in V_1$, como $p_1(x) | q(x)$, $v_1 q(T) = v_1 q(T_1) = 0$; análogamente, para $v_2 \in V_2$, $v_2 q(T) = 0$. Dada cualquier $v \in V$, v puede escribirse como $v = v_1 + v_2$ donde $v_1 \in V_1$ y $v_2 \in V_2$, en consecuencia de lo cual $vq(T) = (v_1 + v_2)q(T) = v_1 q(T) + v_2 q(T) = 0$. Así pues, $q(T) = 0$ y T satisface $q(x)$. Combinado con el resultado del primer párrafo, esto nos da el lema.

COROLARIO. *Si $V = V_1 \oplus \dots \oplus V_k$, donde todo V_i es invariante bajo T y si $p_i(x)$ es polinomio mínimo sobre F de T_i , la transformación lineal inducida por T sobre V_i , entonces el polinomio mínimo de T sobre F es el mínimo común múltiplo de $p_1(x), p_2(x), \dots, p_k(x)$.*

Dejamos la prueba del corolario al lector.

Sea $T \in A_F(V)$ y supongamos que $p(x)$ en $F[x]$ es el polinomio mínimo de T sobre F . Según lema 3.21, podemos factorizar $p(x)$ en $F[x]$ en forma única como $p(x) = q_1(x)^{l_1} q_2(x)^{l_2} \dots q_k(x)^{l_k}$, donde los $q_i(x)$ son polinomios irreducibles distintos en $F[x]$ y donde l_1, l_2, \dots, l_k son enteros positivos. Nuestro objetivo es descomponer V en suma directa de subespacios invariantes bajo T tales que sobre cada uno de éstos la transformación lineal inducida por T tiene como polinomio mínimo una potencia de un polinomio irreducible. Si $k = 1$, V mismo sirve a nuestro propósito. Supongamos pues que $k > 1$.

Sea $V_1 = \{v \in V \mid vq_1(T)^{l_1} = 0\}$, $V_2 = \{v \in V \mid vq_2(T)^{l_2} = 0\}$, ..., $V_k = \{v \in V \mid vq_k(T)^{l_k} = 0\}$. Es una trivialidad que cada V_i es un subespacio de V . Además, V_i es invariante bajo T , pues si $u \in V_i$, como T y $q_i(T)$ comutan, $(uT)q_i(T)^{l_i} = (uq_i(T)^{l_i})T = 0T = 0$. Por la definición de V_i esto sitúa a uT en V_i . Sea T_i la transformación lineal inducida por T sobre V_i .

TEOREMA 6.N. *Para cada $i = 1, 2, \dots, k$, $V_i \neq (0)$ y $V = V_1 \oplus V_2 \oplus \dots \oplus V_k$. El polinomio mínimo de T_i es $q_i(x)^{l_i}$.*

Prueba. Si $k = 1$ entonces $V = V_1$ y no hay nada que necesite probarse. Supongamos entonces que $k > 1$.

Primero necesitamos probar que todo $V_i \neq (0)$. Con este fin introducimos los k polinomios:

$$h_1(x) = q_2(x)^{l_2} q_3(x)^{l_3} \cdots q_k(x)^{l_k},$$

$$h_2(x) = q_1(x)^{l_1} q_3(x)^{l_3} \cdots q_k(x)^{l_k}, \dots,$$

$$h_i(x) = \prod_{j \neq i} q_j(x)^{l_j}, \dots,$$

$$h_k(x) = q_1(x)^{l_1} q_2(x)^{l_2} \cdots q_{k-1}(x)^{l_{k-1}}.$$

Como $k > 1$, $h_i(x) \neq p(x)$, de donde $h_i(T) \neq 0$. Así pues, dada i , hay un $r \in V$ tal que $w = rh_i(T) \neq 0$. Pero $wq_i(T)^{l_i} = r(h_i(T)q_i(T)^{l_i}) = rp(T) = 0$. En consecuencia, $w \neq 0$ está en V_i y por tanto $V_i \neq (0)$. En realidad hemos demostrado un poco más, a saber, que $Vh_i(T) \neq (0)$ está en V_i . Otra observación acerca de $h_i(x)$ viene ahora a cuenta, si $r_j \in V_j$ para $j \neq i$, como $q_j(x)^{l_j} | h_i(x)$, $r_j h_i(T) = 0$.

Los polinomios $h_1(x), h_2(x), \dots, h_k(x)$ son primos relativos. (*Pruébese!*) De aquí que según el lema 3.20 podemos encontrar polinomios $a_1(x), \dots, a_k(x)$ en $F[x]$ tales que $a_1(x)h_1(x) + \dots + a_k(x)h_k(x) = 1$. De donde tenemos, $a_1(T)h_1(T) + \dots + a_k(T)h_k(T) = 1$, de donde, dado $r \in V$, $r = r1 = r(a_1(T)h_1(T) + \dots + a_k(T)h_k(T)) = ra_1(T)h_1(T) + \dots + ra_k(T)h_k(T)$. Ahora bien, cada $ra_i(T)h_i(T)$ está en $Vh_i(T)$, y como hemos probado anteriormente que $Vh_i(T) \subset V_i$, hemos demostrado ahora r como $v = r_1 + \dots + r_k$, donde cada $r_i = ra_i(T)h_i(T)$ está en V_i . Luego $V = V_1 + V_2 + \dots + V_k$.

Debemos ahora verificar que esta suma es una suma directa. Para mostrar esto es suficiente probar que si $u_1 + u_2 + \dots + u_k = 0$ con cada $u_i \in V_i$, entonces cada $u_i = 0$. Supongamos, pues, que $u_1 + u_2 + \dots + u_k = 0$ y que algún u_i , digamos u_1 , no es 0. Multipliquemos esta relación por $h_1(T)$; obtenemos $u_1h_1(T) + \dots + u_kh_1(T) = 0h_1(T) = 0$. Además, $u_jh_1(T) = 0$ para $j \neq 1$ ya que $u_j \in V_j$; la ecuación se reduce así a $u_1h_1(T) = 0$. Pero $u_1q_1(T)^{l_1} = 0$ y como $h_1(x)$ y $q_1(x)$ son primos relativos, esto implica que $u_1 = 0$ (*pruébese!*), lo que es, desde luego, incompatible con la hipótesis de que $u_1 \neq 0$. Hasta el momento hemos conseguido probar que $V = V_1 \oplus V_2 \oplus \dots \oplus V_k$.

Para completar la prueba del teorema debemos todavía probar que el polinomio mínimo de T_i sobre V_i es $q_i(x)^{l_i}$. Por la definición de V_i , como $V_i q_i(T)^{l_i} = 0$, $q_i(T_i)^{l_i} = 0$, de donde la ecuación mínima de T_i debe ser un divisor de $q_i(x)^{l_i}$, luego de la forma $q_i(x)^{f_i}$ con $f_i \leq l_i$. Por el corolario al lema 6.12, el polinomio mínimo de T sobre F es el mínimo común múltiplo de $q_1(x)^{f_1}, \dots, q_k(x)^{f_k}$ y debe, por tanto, ser $q_1(x)^{f_1} \cdots q_k(x)^{f_k}$. Como este polinomio mínimo es en realidad $q_1(x)^{l_1} \cdots q_k(x)^{l_k}$ debemos tener que

$f_1 \geq l_1, f_2 \geq l_2, \dots, f_k \geq l_k$. Combinada con la desigualdad de sentido opuesto que antes probamos, esto nos da el resultado buscado, que $l_i = f_i$ para $i = 1, 2, \dots, k$, con lo que completamos la prueba del teorema.

Si todas las raíces características de T sucediera que estaban en F entonces el polinomio mínimo de T toma la forma particularmente sencilla $q(x) = (x - \lambda_1)^{l_1} \dots (x - \lambda_k)^{l_k}$ donde $\lambda_1, \dots, \lambda_k$ son las distintas raíces características de T . Los factores irreducibles $q_i(x)$ anteriores son simplemente $q_i(x) = x - \lambda_i$. Nótese que sobre V_i , T_i solamente tiene λ_i como raíz característica.

COROLARIO. Si todas las distintas raíces características $\lambda_1, \dots, \lambda_k$ de T se encuentran en F , entonces V puede escribirse como $V = V_1 \oplus \dots \oplus V_k$ donde $V_i = \{v \in V \mid v(T - \lambda_i)^{l_i} = 0\}$ y donde T_i tiene solamente una raíz característica, λ_i , sobre V_i .

Volvamos al teorema por un momento; usamos la misma notación T_i, V_i que en el teorema. Como $V = V_1 \oplus \dots \oplus V_k$, si la dimensión de V_i es n_i , por el lema 6.7 podemos encontrar una base de V tal que en esa base la matriz de T sea de la forma

$$\begin{pmatrix} A_1 \\ & A_2 \\ & & \ddots \\ & & & A_k \end{pmatrix},$$

donde cada A_i es una matriz $n_i \times n_i$ y es en realidad la matriz de T_i .

¿Qué es exactamente lo que andamos buscando? Queremos encontrar un elemento en la clase de semejanza de T que pueda distinguirse en alguna forma. A la luz del teorema 6.h esto puede reformularse como sigue: buscamos una base de V en que la matriz de T tenga una forma especialmente sencilla (y reconocible).

De acuerdo con la anterior discusión, esta búsqueda puede quedar limitada a las transformaciones lineales T_i , con lo que el problema general puede reducirse de la discusión de transformaciones lineales generales a la de las transformaciones lineales especiales, cuyos polinomios mínimos son potencias de polinomios irreducibles. Para la situación especial en que todas las raíces características de T se encuentran en F es lo que vamos a hacer a continuación. El caso general en el que no ponemos restricción alguna sobre las raíces características de T lo estudiaremos en la próxima sección.

Estamos ahora en una feliz posición. Hemos construido todas las piezas y todo lo que tenemos que hacer es juntarlas. Resulta de ello el importan-

tísimo y utilísimo teorema en el que se exhibe lo que usualmente se llama *forma canónica de Jordan*. Pero demos primero una definición.

DEFINICIÓN. La matriz

$$\begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & & \cdot \\ \vdots & & & & \vdots \\ \vdots & & & & 1 \\ 0 & \cdots & & & \lambda \end{pmatrix}$$

con λ en la diagonal y 1 en la superdiagonal y 0 en las demás entradas es un *bloque básico de Jordan perteneciente a λ* .

TEOREMA 6.P. *Sea $T \in A_F(v)$ con todas sus distintas raíces características $\lambda_1, \dots, \lambda_k$, en F . Entonces puede encontrarse una base de V en que la matriz T sea de la forma*

$$\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{pmatrix}$$

donde cada

$$J_i = \begin{pmatrix} B_{i1} & & & \\ & B_{i2} & & \\ & & \ddots & \\ & & & B_{ir_i} \end{pmatrix}$$

y donde B_{i1}, \dots, B_{ir_i} son bloques básicos de Jordan pertenecientes a λ_i .

Prueba. Antes de comenzar notemos que un bloque básico $m \times m$ de Jordan perteneciente a λ es simplemente $\lambda + M_m$, donde M_m es la matriz que se ha definido al final del lema 6.8.

De acuerdo con la combinación del lema 6.7 y el corolario al teorema 6.n, podemos reducirnos al caso en que T tiene solamente una raíz característica λ , es decir, al caso en que $T - \lambda$ es nilpotente. Así pues, $T = \lambda + (T - \lambda)$, y

como $T - \lambda$ es nilpotente, según el teorema 6.1, hay una base en que su matriz es de la forma

$$\begin{pmatrix} M_{n_1} & & \\ & \ddots & \\ & & M_{n_r} \end{pmatrix}.$$

Pero entonces la matriz de T es de la forma

$$\begin{pmatrix} \lambda & & \\ & \lambda & \\ & & \ddots \\ & & & \lambda \end{pmatrix} + \begin{pmatrix} M_{n_1} & & \\ & \ddots & \\ & & M_{n_r} \end{pmatrix} = \begin{pmatrix} B_{n_1} & & \\ & \ddots & \\ & & B_{n_r} \end{pmatrix},$$

usando la primera observación hecha en esta prueba acerca de la relación entre un bloque de Jordan básico y las M_m . Y esto completa el teorema.

Usando el teorema 6.1 podríamos arreglar las cosas de forma que en cada J_i , la dimensión de $B_{i1} \geq$ la dimensión de $B_{i2} \geq \dots$. Cuando esto se ha hecho, entonces la matriz

$$\begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_k \end{pmatrix}$$

se llama la *forma de Jordan* de T . Nótese que el teorema 6.p para matrices nilpotentes se reduce al teorema 6.1.

Dejamos como ejercicio lo siguiente: *dos transformaciones lineales en $A_F(V)$ que tienen todas sus raíces características en F , son semejantes si y sólo si pueden llevarse a la misma forma de Jordan.*

Así pues, la forma de Jordan actúa como un “determinador” para clases de semejanza de este tipo de transformaciones lineales.

En términos de matrices, el teorema 6.p puede formularse como sigue: sea $A \in F_n$ y supongamos que K es el campo de descomposición del polinomio mínimo de A sobre F ; entonces puede encontrarse una matriz invertible $C \in K_n$ tal que CAC^{-1} esté en la forma de Jordan.

Dejamos los pocos puntos necesarios para hacer la traducción del teorema 6.p a su forma matricial como ejercicio para el lector.

Una observación final: si $A \in F_n$ y si K_n , donde K es el campo de des-

composición del polinomio mínimo de A sobre F ,

$$CAC^{-1} = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{pmatrix}$$

donde cada J_i corresponde a una raíz característica distinta λ_i de A , entonces la *multiplicidad de λ_i* como raíz característica de A es, por definición n_i , donde J_i es una matriz $n_i \times n_i$. Nótese que la suma de las multiplicidades es exactamente n .

Es claro que análogamente podríamos definir la multiplicidad de una raíz característica de una transformación lineal.

Problemas

1. Si S y T son transformaciones lineales nilpotentes que comutan, pruébese que ST y $S+T$ son transformaciones lineales nilpotentes.

2. Mediante un cálculo matricial directo, demuéstrese que

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

no son semejantes.

3. Si $n_1 \geq n_2$ y $m_1 \geq m_2$, pruébese, mediante un cálculo matricial directo que

$$\begin{pmatrix} M_{n_1} & \\ & M_{n_2} \end{pmatrix} \text{ y } \begin{pmatrix} M_{m_1} & \\ & M_{m_2} \end{pmatrix}$$

son semejantes si y sólo si $n_1 = m_1$ y $n_2 = m_2$.

*4. Si $n_1 \geq n_2 \geq n_3$ y $m_1 \geq m_2 \geq m_3$, pruébese, por medio de un cálculo matricial directo que

$$\begin{pmatrix} M_{n_1} & & \\ & M_{n_2} & \\ & & M_{n_3} \end{pmatrix} \text{ y } \begin{pmatrix} M_{m_1} & & \\ & M_{m_2} & \\ & & M_{m_3} \end{pmatrix}$$

son semejantes si y sólo si $n_1 = m_1$, $n_2 = m_2$ y $n_3 = m_3$.

5. a) Pruébese que la matriz

$$\begin{pmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 1 & 1 & 0 \end{pmatrix}$$

es nilpotente y encuéntrense sus invariantes y forma de Jordan.

b) Pruébese que la matriz de la parte (a) no es semejante a

$$\begin{pmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 1 & 0 & 0 \end{pmatrix}.$$

6. Pruébese el lema 6.12 y su corolario, incluso si las sumas que en él aparecen no son sumas directas.

7. Pruébese la afirmación hecha de que dos transformaciones lineales en $A_F(V)$ todas cuyas raíces características se encuentran en F son semejantes si y sólo si sus formas de Jordan son iguales (excepto por una permutación en la ordenación de las raíces características).

8. Complétense la prueba de la versión matricial del teorema 6.p, dada en el texto.

9. Pruébese que la matriz $n \times n$

$$\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & & \ddots & & \vdots & \vdots \\ 0 & 0 & 0 & & 1 & 0 \end{pmatrix},$$

con entradas 1 en la subdiagonal y 0 todas las demás, es semejante a M_n .

10. Si F tiene característica $p > 0$ pruébese que $A = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}$ satisface $A^p = 1$.

11. Si F tiene característica 0, pruébese que $A = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}$ satisface $A^m = 1$ para $m > 0$ solamente si $\alpha = 0$.

- 12.** Encuéntrense todas las formas de Jordan posibles para:
- todas las matrices 8×8 que tienen $x^2(x-1)^3$ como polinomio mínimo;
 - todas las matrices 10×10 sobre un campo de característica diferente de 2, que tiene $x^2(x-1)^2(x+1)^3$ como polinomio mínimo.
- 13.** Pruébese que la matriz $n \times n$

$$A = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix}$$

es semejante a

$$\begin{pmatrix} n & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix},$$

si la característica de F es 0 o si es p y $p \nmid n$. ¿Cuál es la multiplicidad de 0 como raíz característica de A ?

Una matriz $A = (\alpha_{ij})$ se dice que es una matriz *diagonal* si $\alpha_{ij} = 0$ para $i \neq j$, es decir, si todas las entradas aparte de las de la diagonal principal son 0. Una matriz, o transformación lineal, se dice que es *diagonalizable* si es semejante a una matriz diagonal (tiene una base en la que su matriz es diagonal).

***14.** Si T está en $A(V)$ entonces T es diagonalizable (si todas sus raíces características están en F) si y sólo si siempre que $v(T-\lambda)^m = 0$, para $v \in V$ y $\lambda \in F$, entonces $v(T-\lambda) = 0$.

15. Usando el resultado del problema 14, pruébese que si $E^2 = E$ entonces E es diagonalizable.

16. Si $E^2 = E$ y $F^2 = F$ pruébese que son semejantes si y sólo si tienen el mismo rango.

17. Si la multiplicidad de cada una de las raíces características de T es 1, y si todas las raíces características de T están en F , pruébese que T es diagonalizable sobre F .

***18.** Si la característica de F es 0 y si $T \in A_F(V)$ satisface $T^m = 1$, pruébese que si las raíces características de T están en F entonces T es diagonalizable. (Sugerencia: úsese la forma de Jordan de T .)

*19. Si $A, B \in F$ son diagonalizables y si comutan, pruébese que hay un elemento $C \in F_n$ tal que tanto CAC^{-1} como CBC^{-1} son diagonales.

20. Pruébese que el resultado del problema 19 es falso si A y B no comutan.

7. FORMAS CANÓNICAS: FORMA CANÓNICA RACIONAL

La forma de Jordan es la más comúnmente usada para probar teoremas acerca de las transformaciones lineales y las matrices. Desgraciadamente tiene un serio inconveniente en los requerimientos que impone sobre la localización de las raíces características. Es cierto que si $T \in A_F(V)$ (o $A \in F_n$) no tiene sus raíces características en F , no tenemos más que ir a una extensión finita K de F en que todas las raíces características de T se encuentran y luego llevan T a su forma de Jordan sobre K . En realidad, este es un procedimiento operativo estándar; pero prueba resultados en K_n , no en F_n . Muy a menudo el resultado en F_n puede deducirse del resultado en K_n , pero hay muchas ocasiones en que después que un resultado se ha establecido para $A \in F_n$ considerado como un elemento en K_n , no podemos volver de K_n para obtener la información deseada en F_n .

Así pues, necesitamos alguna forma canónica para elementos en $A_F(V)$ (o en F_n) que no presupongan nada sobre la localización de las raíces características de sus elementos, una forma canónica y un conjunto de invariantes creados en $A_F(V)$ mismo usando solamente sus elementos y operaciones. La *forma canónica racional*, que describimos a continuación en el teorema 6.q y su corolario, es una forma canónica de tal tipo.

Sea $T \in A_F(V)$; por medio de T nos proponemos hacer de V un módulo sobre $F[x]$, el anillo de los polinomios en x sobre F . Hacemos esto definiendo para cualquier polinomio $f(x)$ en $F[x]$, y cualquier $v \in V$, $f(x)v = vf(T)$. Dejamos la verificación al lector de que, bajo esta definición de multiplicación de elementos de V por elementos de $F[x]$, V se hace un $F[x]$ -módulo.

Como V es de dimensión finita sobre F , está finitamente generado sobre F , luego tanto más sobre $F[x]$ que contiene a F . Además, $F[x]$ es un anillo euclíadiano; luego como un módulo finitamente generado sobre $F[x]$, por el teorema 4.j, V es la suma directa de un número finito de submódulos cíclicos. Por la misma forma en que hemos introducido la estructura de módulo sobre V , cada uno de estos submódulos cíclicos es invariante bajo T ; además, hay un elemento m_0 , en un tal submódulo M , tal que todo elemento m en M es de la forma $m = m_0f(T)$ para algún $f(x) \in F[x]$.

Para determinar la naturaleza de T sobre V será, por tanto, bastante para nosotros conocer como parece T sobre un submódulo cíclico. Es esto precisamente lo que intentamos determinar.

Pero efectuemos primero una descomposición preliminar de V , como

hacemos en el teorema 6.n, de acuerdo con la descomposición del polinomio mínimo de T como producto de polinomios irreducibles.

Sea el polinomio mínimo $p(x)$ de T sobre F , $p(x) = q_1(x)^{e_1} \dots q_k(x)^{e_k}$ donde los $q_i(x)$ son polinomios irreducibles distintos en $F[x]$ y donde cada $e_i > 0$; entonces, como vimos en el teorema 6.n, $V = V_1 \oplus V_2 \oplus \dots \oplus V_k$ donde cada V_i es invariante bajo T y donde el polinomio mínimo de T sobre V_i es $q_i(x)^{e_i}$. Para resolver la naturaleza de un submódulo cíclico para un T arbitrario vemos, por esta discusión, que es suficiente establecerla para un T cuyo polinomio mínimo sea una potencia de uno irreducible.

Probamos el

LEMA 6.13. *Supongamos que T , en $A_F(V)$, tiene como polinomio mínimo sobre F el polinomio $p(x) = \gamma_0 + \gamma_1 x + \dots + \gamma_{r-1} x^{r-1} + x^r$. Supongamos, además, que V como módulo (de acuerdo con lo antes descrito), es un módulo cíclico (es decir, es cíclico respecto a T). Entonces hay una base de V sobre F tal que, en esta base, la matriz de T es*

$$\begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \dots & 1 \\ -\gamma_0 & -\gamma_1 & \ddots & \cdots & -\gamma_{r-1} \end{pmatrix}$$

Prueba. Como V es cíclico respecto a T , existe un vector v en V tal que todo elemento w en V es de la forma $w = vf(T)$ para algún $f(x)$ en $F[x]$.

Ahora bien, si para algún polinomio $s(x)$ en $F[x]$, $vs(T) = 0$, entonces para cualquier w en V , $ws(T) = (vf(T))s(T) = vs(T)f(T) = 0$; luego $s(T)$ aniquila a todo V y, por tanto, $s(T) = 0$. Pero entonces $p(x)|s(x)$ ya que $p(x)$ es el polinomio mínimo de T . Esta observación implica de inmediato que $v, vT, vT^2, \dots, vT^{r-1}$ son linealmente independientes sobre F , pues si así no fuera, entonces $\alpha_0 v + \alpha_1 vT + \dots + \alpha_{r-1} vT^{r-1} = 0$ con $\alpha_0, \dots, \alpha_{r-1}$ en F . Pero entonces $v(\alpha_0 + \alpha_1 T + \dots + \alpha_{r-1} T^{r-1}) = 0$, y de aquí, según la anterior discusión, $p(x)|(\alpha_0 + \alpha_1 x + \dots + \alpha_{r-1} x^{r-1})$, lo que es imposible ya que $p(x)$ es de grado r salvo si $\alpha_0 = \alpha_1 = \dots = \alpha_{r-1} = 0$.

Como $T^r = -\gamma_0 - \gamma_1 T - \dots - \gamma_{r-1} T^{r-1}$, es inmediato que T^{r+k} , para $k \geq 0$, es una combinación lineal de $1, T, \dots, T^{r-1}$ y, por tanto, que $f(T)$ para cualquier $f(x) \in F[x]$, es una combinación lineal de $1, T, \dots, T^{r-1}$ sobre F . Como cualquier w en V es de la forma $w = vf(T)$ tenemos que w es una combinación lineal de v, vT, \dots, vT^{r-1} .

Hemos probado, en los últimos dos párrafos, que los elementos v, vT, \dots, vT^{r-1} forman una base de V sobre F . En esta base, como puede verificarse de inmediato, la matriz de T es exactamente como afirmábamos.

DEFINICIÓN. Si $f(x) = \gamma_0 + \gamma_1 x + \dots + \gamma_{r-1} x^{r-1} + x^r$ está en $F[x]$ entonces la matriz $r \times r$

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & 1 \\ -\gamma_0 & -\gamma_1 & \ddots & \cdots & -\gamma_{r-1} \end{pmatrix}$$

se llama *matriz compañera* de $f(x)$. La representamos por $C(f(x))$.

Nótese que el lema 6.13 dice que si V es cíclico respecto a T y si el polinomio mínimo de T en $F[x]$ es $p(x)$, entonces para alguna base de V la matriz de T es $C(p(x))$.

Nótese, además que la matriz $C(f(x))$, para cualquier polinomio mónico $f(x)$ en $F[x]$, satisface $f(x)$ y tiene a $f(x)$ como su polinomio mínimo. (Véase el problema 4 al final de esta sección; véase también el problema 29 al final de la sección 1.)

Probamos ahora un importantísimo teorema.

TEOREMA 6.Q. Si T en $A_F(V)$ tiene un polinomio mínimo $p(x) = q(x)^e$ donde $q(x)$ es un polinomio mónico irreducible en $F[x]$, entonces puede encontrarse una base de V sobre F en que la matriz de T sea de la forma

$$\begin{pmatrix} C(q(x)^{e_1}) & & & \\ & C(q(x)^{e_2}) & & \\ & & \ddots & \\ & & & C(q(x)^{e_r}) \end{pmatrix}$$

donde $e = e_1 \geq e_2 \geq \dots \geq e_r$.

Prueba. Puesto que V como módulo sobre $F[x]$ está finitamente generado, y como $F[x]$ es euclíadiano, podemos descomponer V como $V = V_1 \oplus \dots \oplus V_r$ donde los V_i son módulos cíclicos. Los V_i son, entonces, invariantes bajo T ; si T_i es la transformación lineal inducida por T sobre V_i , su polinomio mínimo debe ser un divisor de $p(x) = q(x)^e$, luego de la forma $q(x)^{e_i}$. Podemos reordenar los espacios de forma que $e_1 \geq e_2 \geq \dots \geq e_r$.

Ahora bien, $q(T)^{e_i}$ aniquila a cada V_i , de donde suprime a V , de donde $q(T)^{e_i} = 0$. Luego $e_i \geq e$; como e_i es claramente cuando más igual a e , tenemos que $e_i = e$.

Según el lema 6.13, como cada V_i es cíclico respecto a T podemos encontrar una base tal que la matriz de la transformación lineal de T_i sobre V_i

es $C(q(x)^{e_i})$. Así pues, según el teorema 6.n, podemos encontrar una base de V tal que la matriz de T en esta base es

$$\begin{pmatrix} C(q(x)^{e_1}) & & & \\ & C(q(x)^{e_2}) & & \\ & & \ddots & \\ & & & C(q(x)^{e_r}) \end{pmatrix}.$$

COROLARIO. Si T en $A_F(V)$ tiene el polinomio mínimo $p(x) = q_1(x)^{l_1} \dots q_k(x)^{l_k}$ sobre F , donde $q_1(x), \dots, q_k(x)$ son polinomios irreducibles distintos en $F[x]$, entonces puede encontrarse una base de V en la que la matriz de T sea de la forma

$$\begin{pmatrix} R_1 & & & \\ & R_2 & & \\ & & \ddots & \\ & & & R_k \end{pmatrix}$$

donde cada

$$R_i = \begin{pmatrix} C(q_i(x)^{e_{i1}}) & & & \\ & \ddots & & \\ & & \ddots & \\ & & & C(q_i(x)^{e_{ir_i}}) \end{pmatrix}$$

donde $e_i = e_{i1} \geq e_{i2} \geq \dots \geq e_{ir_i}$.

Prueba. Por el teorema 6.1, V puede ser descompuesto en la suma directa $V = V_1 \oplus \dots \oplus V_k$, donde cada V_i es invariante bajo T y donde el polinomio mínimo de T_i , la transformación lineal inducida por T sobre V_i , es $q_i(x)^{e_i}$. Usando el lema 6.7 y el teorema anterior, obtenemos el corolario. Si el grado de $q_i(x)$ es d_i , nótese que la suma de todos los $d_i e_{ij}$ es n , la dimensión de V sobre F .

DEFINICIÓN. La matriz de T en el enunciado del corolario anterior se llama *forma canónica racional de T* .

DEFINICIÓN. Los polinomios $q_1(x)^{e_{11}}, q_1(x)^{e_{12}}, \dots, q_1(x)^{e_{1r_1}}, \dots, q_k(x)^{e_{kr_k}}$ en $F[x]$ se llaman *divisores elementales de T* .

¡Una definición más!

DEFINICIÓN. Si $\dim_F(V) = n$, entonces el *polinomio característico* de T , $p_T(x)$, es el producto de sus divisores elementales.

Podremos identificar el polinomio característico que acabamos de definir con otro polinomio que construiremos explícitamente en la sección 9. El polinomio característico de T es un polinomio de grado n que se encuentra en $F[x]$. Tiene muchas propiedades importantes, una de las cuales es la contenida en la siguiente

OBSERVACIÓN. *Toda transformación lineal $T \in A_F(V)$ satisface a su polinomio característico. Toda raíz característica de T es una raíz de $p_T(x)$.*

Nota 1. La primera parte de la anterior observación es el enunciado de un teorema muy famoso, el *teorema de Cayley-Hamilton*. Pero llamarlo así en la forma en que lo hemos expuesto resultaría un poco abusivo. El meollo del teorema de Cayley-Hamilton es el hecho de que T satisface $p_T(x)$ cuando a $p_T(x)$ se le da una forma concreta muy particular, fácilmente construible partiendo de T . Pero incluso en la forma en que aparece, la observación tiene un contenido bastante interesante, pues como el polinomio característico es un polinomio de grado n , hemos probado que todo elemento de $A_F(V)$ satisface un polinomio de grado n que se encuentra en $F[x]$. Hasta ahora, solo habíamos probado esto (en el teorema 6.k) para transformaciones lineales que tenían todas sus raíces características en F .

Nota 2. Tal como está formulada, la segunda parte no dice nada, pues siempre que T satisface un polinomio, entonces toda raíz característica de T satisface ese mismo polinomio; así pues, $p_T(x)$ no sería nada especial si lo que se enuncia en el teorema fuera todo lo que es válido en él. Pero la historia real es la siguiente: Toda raíz característica de T es una raíz de $p_T(x)$, y reciprocamente, *toda raíz de $p_T(x)$ es una raíz característica de T ; además, la multiplicidad de cualquier raíz de $p_T(x)$, como una raíz del polinomio, es igual a su multiplicidad como raíz característica de T .* Podríamos probar lo dicho ahora, pero diferimos la prueba hasta más tarde, cuando seamos capaces de hacerla de una forma más natural.

Prueba de la observación. Solamente tenemos que demostrar que T satisface a $p_T(x)$, pero esto es casi trivial. Como $p_T(x)$ es el producto de $q_1(x)^{e_{11}}, q_1(x)^{e_{12}}, \dots, q_k(x)^{e_{k1}}, \dots$, y como $e_{11} = e_1, e_{21} = e_2, \dots, e_{k1} = e_k$, $p_T(x)$ es divisible por $p(x) = q_1(e)^{e_1} \dots q_k(x)^{e_k}$, el polinomio mínimo de T . Como $p(T) = 0$ se sigue que $p_T(T) = 0$.

Hemos llamado al conjunto de los polinomios que aparecen en la forma racional canónica de T los divisores elementales de T . Sería muy conveniente que éstos determinasen una semejanza en $A_F(V)$, pues entonces las clases de semejanza en $A_F(V)$ estarían en una correspondencia biyectiva con

conjuntos de polinomios en $F[x]$. Nos proponemos hacer esto, pero primero establecemos un resultado que implica que dos transformaciones lineales tienen los mismos divisores elementales.

TEOREMA 6.R. *Sean V y W dos espacios vectoriales sobre F y supongamos que ψ es un isomorfismo de espacios vectoriales de V sobre W . Supongamos que $S \in A_F(V)$ y $T \in A_F(W)$ son tales que para cualquier $v \in V$, $(vS)\psi = (v\psi)T$. Entonces S y T tienen los mismos divisores elementales.*

Prueba. Comenzamos con un simple cálculo. Si $v \in V$, entonces $(vS^2)\psi = ((vS)S)\psi = ((v\psi)T)T = (v\psi)T^2$. Es claro que continuando análogo proceso tendremos que $(vS^m)\psi = (v\psi)T^m$ para cualquier entero $m \geq 0$, de donde para cualquier polinomio $f(x) \in F[x]$ y para cualquier $v \in V$, $(vf(S))\psi = (v\psi)f(T)$.

Si $f(S) = 0$, entonces $(v\psi)f(T) = 0$ para cualquier $v \in V$ y como ψ transforma V sobre W tendríamos que $Wf(T) = \{0\}$, a consecuencia de lo cual $f(T) = 0$. Recíprocamente, si $g(x) \in F[x]$ es tal que $g(T) = 0$, entonces para cualquier $v \in V$, $(vg(S))\psi = 0$ y como ψ es un isomorfismo, esto nos dice que $vg(S) = 0$. Esto desde luego implica que $g(S) = 0$. Así pues, S y T satisfacen el mismo conjunto de polinomios en $F[x]$, de donde *deben tener el mismo polinomio mínimo*

$$p(x) = q_1(x)^{\epsilon_1} q_2(x)^{\epsilon_2} \dots q_k(x)^{\epsilon_k}$$

donde $q_1(x), \dots, q_k(x)$ son polinomios irreducibles distintos en $F[x]$.

Si U es un subespacio de V invariante bajo S , entonces $U\psi$ es un subespacio de W invariante bajo T , pues $(U\psi)T = (US)\psi \subset U\psi$. Como U y $U\psi$ son isomorfos, el polinomio mínimo de S_1 , la transformación lineal inducida por S sobre U es la misma, de acuerdo con las observaciones anteriores que el polinomio mínimo de T_1 , la transformación lineal inducida sobre $U\psi$ por T .

Ahora bien, como el polinomio mínimo para S sobre V es $p(x) = q_1(x)^{\epsilon_1} \dots q_k(x)^{\epsilon_k}$, como hemos visto en el teorema 6.Q y su corolario, podemos tomar como el primer divisor elemental de S al polinomio $q_1(x)^{\epsilon_1}$ y podemos encontrar un subespacio de V_1 de V que es invariante bajo S tal que:

- 1) $V = V_1 \oplus M$ donde M es invariante bajo S ;
- 2) los únicos divisores elementales de S_1 , la transformación lineal inducida sobre V_1 por S , es $q_1(x)^{\epsilon_1}$;
- 3) los otros divisores elementales de S son los de la transformación lineal S_2 inducida por S sobre M .

Combinamos ahora las afirmaciones hechas anteriormente y afirmamos:

- 1) $W = W_1 \oplus N$ donde $W_1 = V_1\psi$ y $N = M\psi$ son invariantes bajo T .

- 2) El único divisor elemental de T_1 , la transformación lineal inducida por T sobre W_1 es $q_1(x)^{e_1}$ (*que es un divisor elemental de T ya que el polinomio mínimo de T es $p(x) = q_1(x)^{e_1} \dots q_k(x)^{e_k}$.*)
- 3) Los otros divisores elementales de T son los de la transformación lineal T_2 inducida por T sobre N .

Como $N = M\psi$, M y N son espacios vectoriales isomorfos sobre F bajo el isomorfismo ψ_2 inducido por ψ . Además, si $u \in M$ entonces $(uS_2)\psi_2 = (uS)\psi = (u\psi)T = (u\psi_2)T_2$, de donde S_2 y T_2 están en la misma relación con respecto a ψ_2 que S y T estaban respecto a ψ . Por inducción sobre la dimensión (o repitiendo el argumento) S_2 y T_2 tienen los mismos divisores elementales. Pero como los divisores elementales de S son simplemente $q_1(x)^{e_1}$ y los de S_2 mientras que los de T son simplemente $q_1(x)^{e_1}$ y los de T_2 , S y T deben tener los mismos divisores elementales, probando con ello el teorema.

El teorema 6.q y su corolario nos dieron la forma canónica racional y los divisores elementales. Nos gustaría apurar un poco más la situación y ser capaces de afirmar alguna propiedad de unicidad. Es lo que hacemos en el

TEOREMA 6.s. *Los elementos S y T en $A_F(V)$ son semejantes (en $A_F(V)$) si y sólo si tienen los mismos divisores elementales.*

Prueba. Probar esto es sencillo en una dirección, pues supongamos que S y T tienen los mismos divisores elementales. Entonces hay dos bases de V sobre F tales que la matriz de S en la primera base es igual a la matriz de T en la segunda (y cada una de ellas igual a la matriz de forma racional canónica). Pero como ya hemos visto varias veces antes, esto implica que S y T son semejantes.

Vamos ahora a ir en la otra dirección. También aquí el argumento se asemeja estrechamente al usado en la sección 5 en la prueba del teorema 6.m. Como allí fuimos muy cuidadosos con todos los detalles, creemos que aquí podemos permitirnos ser un poco más esquemáticos.

Observemos primero que en vista del teorema 6.n, podemos limitarnos al caso de la transformación lineal cuyo polinomio mínimo es una potencia de un polinomio irreducible. Así pues, sin pérdida de generalidad podemos suponer que el polinomio mínimo de T es $q(x)^e$ donde $q(x)$ es irreducible en $F[x]$ y de grado d .

La forma canónica racional nos dice que podemos descomponer V en la forma $V = V_1 \oplus \dots \oplus V_s$, donde los subespacios V_i son invariantes bajo T y donde la transformación lineal inducida por T sobre V_i tiene como matriz $C(q(x)^{e_i})$, la matriz compañera de $q(x)^{e_i}$. Suponemos que lo que realmente estamos intentando probar es lo siguiente: si $V = U_1 \oplus U_2 \oplus \dots \oplus U_s$ donde los U_j son invariantes bajo T y donde la transformación lineal inducida

por T sobre U_j tiene como matriz $C(q(x)^{f_j})$, $f_1 \geq f_2 \geq \dots \geq f_s$, entonces $r = s$ y $e_1 = f_1$, $e_2 = f_2, \dots, e_r = f_r$. (Pruébese que la demostración de esto es equivalente a la demostración del teorema.)

Supongamos entonces que tenemos las dos descomposiciones arriba descritas, $V = V_1 \oplus \dots \oplus V_s$, y $V = U_1 \oplus \dots \oplus U_s$ y que algún $e_i \neq f_i$. Entonces hay un primer entero m tal que $e_m \neq f_m$ mientras que $e_1 = f_1, \dots, e_{m-1} = f_{m-1}$. Podemos suponer que $e_m > f_m$.

Ahora bien, $g(T)^{f_m}$ suprime U_m, U_{m+1}, \dots, U_s , de donde

$$V_q(T)^{f_m} = U_1 q(T)^{f_m} \oplus \dots \oplus U_{m-1} q(T)^{f_m}.$$

Pero se puede demostrar que la dimensión de $U_i q(T)^{f_m}$ para $i \leq m$ es $d(f_i - f_m)$ (ipruébese!), de donde

$$\dim(V_q(T)^{f_m}) = d(f_1 - f_m) + \dots + d(f_{m-1} - f_m).$$

Por otra parte, $V_q(T)^{f_m} \supseteq V_1 q(T)^{f_m} \oplus \dots \oplus \dots \oplus V_m q(T)^{f_m}$ y como $V_i q(T)^{f_m}$ tiene dimensión $d(e_i - f_m)$ para $i \leq m$, tenemos que

$$\dim(V_q(T)^{f_m}) \geq d(e_1 - f_m) + \dots + d(e_m - f_m).$$

Como $e_1 = f_1, \dots, e_{m-1} = f_{m-1}$ y $e_m > f_m$ esto contradice la igualdad antes probada. Hemos, pues, probado el teorema.

COROLARIO 1. *Supongamos que las dos matrices A y B en F_n son semejantes en K_n donde K es una extensión de F . Entonces A y B son ya semejantes en F_n .*

Prueba. Supongamos que $A, B \in F_n$ son tales que $B = C^{-1}AC$ con $C \in K_n$. Consideramos a K_n como si actuara sobre $K^{(n)}$, el espacio vectorial de n -tuples sobre K . Así pues, $F^{(n)}$ está contenido en $K^{(n)}$ y aunque es un espacio vectorial sobre F no es un espacio vectorial sobre K . La imagen de $F^{(n)}$, en $K^{(n)}$, bajo C no necesariamente incidirá de nuevo en $F^{(n)}$ pero en cualquier caso $F^{(n)}C$ es un subconjunto de $K^{(n)}$ que es un espacio vectorial sobre F (pruébese). Sea V el espacio vectorial $F^{(n)}$ sobre F , W el espacio vectorial $F^{(n)}C$ sobre F y para $v \in V$ sea $v\psi = vC$. Ahora bien, $A \in A_F(V)$ y $B \in A_F(W)$ y para cualquier $v \in V$, $(vA)\psi = vAC = vCB = (v\psi)B$, de donde las condiciones del teorema 6.r se satisfacen. Así pues, A y B tienen los mismos divisores elementales; de acuerdo con el teorema 6.s, A y B deben ser semejantes en F_n .

Una palabra de advertencia: el corolario *no afirma* que si $A, B \in F_n$ son tales que $B = C^{-1}AC$ son $C \in K_n$ entonces C debe necesariamente estar en F_n ; esto es falso. Lo que afirma simplemente es que si $A, B \in F_n$ son tales que $B = C^{-1}AC$ con $C \in K_n$, entonces existe un $D \in F_n$ (posiblemente diferente a C) tal que $B = D^{-1}AD$.

Problemas

1. Verifíquese que V se hace un $F[x]$ módulo bajo la definición dada.
2. En la prueba del teorema 6.s proporcionense demostraciones completas de todos los puntos en que se señala (pruébese).
- *3. a) Pruébese que toda raíz del polinomio característico de T es una raíz característica de T .
b) Pruébese que la multiplicidad de cualquier raíz de $p_T(x)$ es igual a su multiplicidad como una raíz característica de T .
4. Pruébese que para $f(x) \in F[x]$, $C(f(x))$ satisface $f(x)$ y tiene a $f(x)$ como su polinomio mínimo. ¿Cuál es su polinomio característico?
5. Si F es el campo de los números racionales, encuéntrense todas las formas canónicas racionales posibles y todos los divisores elementales para:
 - a) Las matrices 6×6 en F_6 que tienen $(x-1)(x^2+1)^2$ como polinomio mínimo.
 - b) Las matrices 15×15 en F_{15} que tienen $(x^2+x+1)^2(x^3+2)^2$ como polinomio mínimo.
 - c) Las matrices 10×10 en F_{10} que tienen $(x^2+1)^2(x^3+1)$ como polinomio mínimo.
6. a) Si K es una extensión de F y si A está en K_n , pruébese que A puede escribirse como $A = \lambda_1 A_1 + \dots + \lambda_k A_k$, donde A_1, \dots, A_k están en F_n y donde $\lambda_1, \dots, \lambda_k$ están en K y son linealmente independientes sobre F .
b) Con igual notación que en la parte (a), pruébese que si $B \in F_n$ es tal que $AB = 0$ entonces $A_1 B = A_2 B = \dots = A_k B = 0$.
c) Si C en F_n commuta con A pruébese que C commuta con cada uno de los A_1, A_2, \dots, A_k .
- *7. Si A_1, \dots, A_k están en F_n y son tales que para ciertos $\lambda_1, \dots, \lambda_k$ en K , una extensión de F , $\lambda_1 A_1 + \dots + \lambda_k A_k$ es invertible en K_n , pruébese que si F tiene un número infinito de elementos podemos encontrar $\alpha_1, \dots, \alpha_k$ en F tales que $\alpha_1 A_1 + \dots + \alpha_k A_k$ es invertible en F_n .
- *8. Si F es un campo finito, pruébese que el resultado del problema 7 es falso.
- *9. Usando los resultados de los problemas 6 (a) y 7, pruébese que si F tiene un número infinito de elementos entonces siempre que $A, B \in F_n$ son semejantes en K_n , donde K es una extensión de F , entonces son semejantes en F_n . (Esto nos da una prueba, independiente de las formas canónicas del corolario 1 al teorema 6.s en el caso particular en que F es un campo infinito.)

10. Usando cálculos con matrices (pero siguiendo los lineamientos marcados en el problema 9), pruébese que si F es el campo de los números reales y K el de los números complejos, entonces dos elementos en F_2 que son semejantes en K_2 son ya semejantes en F_2 .

8. TRAZA Y TRANSPUESTA

Después de la dificultosa marcha en las últimas secciones, la falta de complicaciones del material sobre el que ahora vamos a tratar va a llegarnos como un agradable respiro.

Sea F un campo y A una matriz en F_n .

DEFINICIÓN. La *traza* de A es la suma de los elementos de la diagonal principal de A .

Representaremos a la traza de A por $\text{tr } A$; si $A = (\alpha_{ij})$, entonces $\text{tr } A = \sum_{i=1}^n \alpha_{ii}$.

Las propiedades fundamentales de la función traza están contenidas en el

LEMA 6.14. Para $A, B \in F_n$ y $\lambda \in F$,

- 1) $\text{tr}(\lambda A) = \lambda \text{tr } A$;
- 2) $\text{tr}(A + B) = \text{tr } A + \text{tr } B$;
- 3) $\text{tr}(AB) = \text{tr}(BA)$.

Prueba. Establecer (1) y (2) (que aseguran que la traza es una funcional lineal en F_n) es sencillo y se deja como ejercicio para el lector. Solamente presentamos la prueba de la parte (3) del lema.

Si $A = (\alpha_{ij})$ y $B = (\beta_{ij})$, entonces $AB = (\gamma_{ij})$ donde $\gamma_{ij} = \sum_{k=1}^n \alpha_{ik}\beta_{kj}$ y

$$BA = (\mu_{ij}) \quad \text{donde} \quad \mu_{ij} = \sum_{k=1}^n \beta_{ik}\alpha_{kj}.$$

Así pues, $(AB) = \sum_i \gamma_{ii} = \sum_i \left(\sum_k \alpha_{ik}\beta_{ki} \right)$; si intercambiamos el orden de sumación en la última suma, tenemos

$$\text{tr}(AB) = \sum_{k=1}^n \sum_{i=1}^n \alpha_{ik}\beta_{ki} = \sum_{k=1}^n \left(\sum_{i=1}^n \beta_{ki}\alpha_{ik} \right) = \sum_{k=1}^n \mu_{kk} = \text{tr}(BA).$$

COROLARIO. Si A es invertible entonces $\text{tr}(ACA^{-1}) = \text{tr } C$.

Prueba. Sea $B = CA^{-1}$; entonces $\text{tr}(ACA^{-1}) = \text{tr}(AB) = \text{tr}(BA) = \text{tr}(CA^{-1}A) = \text{tr } C$.

Este corolario es importante por dos razones; primero, nos permitirá definir la traza de una transformación lineal arbitraria; segundo, nos permitirá encontrar una expresión alternativa para la traza de A .

DEFINICIÓN. Si $T \in A(V)$ entonces $\text{tr } T$, la *traza* de T , es la traza de $m_1(T)$ donde $m_1(T)$ es la matriz de T en una base cualquiera de V .

Afirmamos que la definición tiene sentido y depende solamente de T y no de cual sea la base de V que se emplee. En efecto, si $m_1(T)$ y $m_2(T)$ son matrices semejantes, entonces, según el corolario al lema 6.14, ambas tienen la misma traza.

LEMA 6.15. *Si $T \in A(V)$, entonces $\text{tr } T$ es la suma de las raíces características de T (usando cada raíz característica tantas veces como su multiplicidad).*

Prueba. Podemos suponer que T es una matriz en F_n ; si K es el campo de descomposición para el polinomio mínimo de T sobre F , entonces en K_n , por el teorema 6.p, T puede llevarse a su forma de Jordan, J . J es una matriz sobre cuya diagonal aparecen las raíces características de T , cada raíz que aparece tantas veces como unidades tiene su multiplicidad. Así pues, $\text{tr } J =$ suma de las raíces características de T ; pero como J es de la forma ATA^{-1} , $\text{tr } J = \text{tr } T$, y esto prueba el lema.

Si T es nilpotente, entonces todas sus raíces características son 0, de donde, de acuerdo con el lema 6.15, $\text{tr } T = 0$. Pero si T es nilpotente, entonces también lo son T^2, T^3, \dots , luego $\text{tr } T^i = 0$ para todo $i \geq 1$.

¿Y qué podemos decir en la otra dirección, es decir, si $\text{tr } T^i = 0$, para $i = 1, 2, \dots$? ¿Se sigue de ello que T es nilpotente? Con esta generalidad la contestación es no, pues si F es un campo de característica 2, entonces la matriz unidad

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

en F_2 tiene traza 0 (pues $1+1=0$) al igual que todas sus potencias, pero es claro que la matriz unidad no es nilpotente. Pero si restringimos la característica de F a 0, el resultado es verdaderamente cierto.

LEMA 6.16. *Si F es un campo de característica 0 y si $T \in A_F(V)$ es tal que $\text{tr } T^i = 0$ para todo $i \geq 1$, entonces T es nilpotente.*

Prueba. Como $T \in A_F(V)$, T satisface algún polinomio mínimo $p(x) = x^m + \alpha_1 x^{m-1} + \dots + \alpha_m$; como $T^m + \alpha_1 T^{m-1} + \dots + \alpha_{m-1} T + \alpha_m = 0$, tomando trazas de ambos lados, tenemos

$$\operatorname{tr} T^m + \alpha_1 \operatorname{tr} T^{m-1} + \dots + \alpha_{m-1} \operatorname{tr} T + \operatorname{tr} \alpha_m = 0.$$

Pero por hipótesis, $\operatorname{tr} T^i = 0$ para $i \geq 1$, luego tenemos $\operatorname{tr} \alpha_m = 0$; si $\dim V = n$, $\operatorname{tr} \alpha_m = n\alpha_m$, de donde $n\alpha_m = 0$. Pero la característica de F es 0; luego $n \neq 0$, de donde se sigue que $\alpha_m = 0$. Como el término constante del polinomio mínimo de T es 0, por el teorema 6.b T es singular y por tanto 0 es una raíz característica de T .

Podemos considerar a T como una matriz en F_n y, por tanto, también como una matriz en K_n , donde K es una extensión de F que, a su vez, contiene todas las raíces características de T . En K_n , según el teorema 6.j, podemos poner T en forma triangular, y como 0 es una raíz característica de T , podemos realmente llevarla a la forma

$$\left(\begin{array}{c|ccc} 0 & 0 & \cdots & 0 \\ \hline \beta_2 & \alpha_2 & 0 & \cdot & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \beta_n & * & & & \alpha_n \end{array} \right) = \left(\begin{array}{c|c} 0 & 0 \\ \hline * & T_2 \end{array} \right),$$

donde

$$T_2 = \left(\begin{array}{ccc} \alpha_2 & 0 & 0 \\ & \ddots & \vdots \\ & * & \alpha_n \end{array} \right)$$

es una matriz $(n-1) \times (n-1)$ (los * indican partes en cuyas entradas no estamos interesados). Ahora

$$T^k = \left(\begin{array}{c|c} 0 & 0 \\ \hline * & T_2^k \end{array} \right)$$

de donde $0 = \operatorname{tr} T^k = \operatorname{tr} T_2^k$. Luego T_2 es una matriz $(n-1) \times (n-1)$ con la propiedad de que $\operatorname{tr} T_2^k = 0$ para todo $k \geq 1$. O bien usando inducción sobre n , o repitiendo el argumento sobre T_2 que usamos para T , tenemos, como $\alpha_2, \dots, \alpha_n$ son las raíces características de T_2 , que $\alpha_2 = \dots = \alpha_n = 0$. Luego cuando T se pone en forma triangular todas sus entradas en la diagonal principal son 0, lo que implica que T sea nilpotente (pruébese).

Este lema, aunque pueda parecer particular, nos servirá en una gran cantidad de casos. Hacemos uso inmediato de él para probar un resultado usualmente conocido como el *lema de Jacobson*.

LEMÁ 6.17. Si F es de característica 0 y si S y T , de $A_F(V)$, son tales que $ST - TS$ conmuta con S , entonces $ST - TS$ es nilpotente.

Prueba. Para cualquier $k \geq 1$, calculamos $(ST - TS)^k$. Ahora bien, $(ST - TS)^k = (ST - TS)^{k-1}(ST - TS) = (ST - TS)^{k-1}ST - (ST - TS)^{k-1}TS$. Como $ST - TS$ conmuta con S , el término $(ST - TS)^{k-1}ST$ puede escribirse en la forma $S((ST - TS)^{k-1})T$. Si hacemos $B = (ST - TS)^{k-1}T$ vemos que $(ST - TS)^k = SB - BS$; de donde $\text{tr}((ST - TS)^k) = \text{tr}(SB - BS) = \text{tr}(SB) - \text{tr}(BS) = 0$ según el lema 6.14. El lema anterior nos dice ahora que $ST - TS$ debe ser nilpotente.

La traza nos provee de una funcional lineal sobre F_n (y, por tanto, sobre $A_F(V)$) en F , extremadamente útil. Introducimos ahora una importante transformación de F_n en sí mismo.

DEFINICIÓN. Si $A = (\alpha_{ij}) \in F_n$, entonces la *transpuesta de A*, escrita como A' , es la matriz $A' = (\gamma_{ij})$ donde $\gamma_{ij} = \alpha_{ji}$ para todas las i y j .

La transpuesta de A es la matriz que se obtiene intercambiando los renglones de A con las columnas de A . Las propiedades formales básicas de la transpuesta, están contenidas en

LEMÁ 6.18. Para cualesquiera $A, B \in F_n$,

- 1) $(A')' = A$;
- 2) $(A + B)' = A' + B'$;
- 3) $(AB)' = B'A'$.

Prueba. Las pruebas de las partes (1) y (2) son muy sencillas y se dejan como ejercicio para el lector; nos contentamos nosotros con la prueba de la parte (3).

Supongamos que $A = (\alpha_{ij})$ y $B = (\beta_{ij})$; entonces $AB = (\lambda_{ij})$ donde

$$\lambda_{ij} = \sum_{k=1}^n \alpha_{ik} \beta_{kj}.$$

Por tanto, por definición, $(AB)' = (\mu_{ij})$, donde $\mu_{ij} = \lambda_{ji} = \sum_{k=1}^n \alpha_{jk} \beta_{ki}$. Por otra parte, $A' = (\gamma_{ij})$ donde $\gamma_{ij} = \alpha_{ji}$ y $B' = (\xi_{ij})$ donde $\xi_{ij} = \beta_{ji}$, de donde el elemento (i, j) de $B'A'$ es $\sum_{k=1}^n \xi_{ik} \gamma_{kj} = \sum_{k=1}^n \beta_{ki} \alpha_{jk} = \sum_{k=1}^n \alpha_{jk} \beta_{ki} = \mu_{ij}$. Es decir, $(AB)' = B'A'$, con lo que hemos verificado la parte (3) del lema.

En la parte (3), si nos fijamos en el caso particular en que $A = B$, obtenemos $(A^2)' = (A')^2$. Continuando obtenemos $(A^k)' = (A')^k$ para todo entero positivo k . Cuando A es invertible, entonces $(A^{-1})' = (A')^{-1}$.

Existe otra propiedad de la transpuesta, a saber, si $\lambda \in F$ entonces $(\lambda A)' = \lambda A'$ para toda $A \in F_n$. Ahora bien, si $A \in F_n$ satisface un polinomio $\alpha_0 A^m + \alpha_1 A^{m-1} + \dots + \alpha_m = 0$, obtenemos $(\alpha_0 A^m + \dots + \alpha_m)' = 0' = 0$. Calculando explícitamente $(\alpha_0 A^m + \dots + \alpha_m)'$ usando las propiedades de la transpuesta, obtenemos $\alpha_0 (A')^m + \alpha_1 (A')^{m-1} + \dots + \alpha_m = 0$, es decir, A' satisface cualquier polinomio sobre F al que satisfaga A . Como $A = (A')$, por el mismo razonamiento, A satisface cualquier polinomio sobre F al que satisfaga A' . En particular, A y A' tienen el mismo polinomio mínimo sobre F y, por tanto, *tienen las mismas raíces características*. Puede demostrarse que todas las raíces tienen la misma multiplicidad en A que en A' . Esto es evidente una vez que se establece que A y A' son realmente semejantes (véase el problema 14).

DEFINICIÓN. La matriz A se dice que es una *matriz simétrica* si $A' = A$.

DEFINICIÓN. La matriz A se dice que es una *matriz antisimétrica* si $A' = -A$.

Cuando la característica de F es 2, como $1 = -1$, no podemos distinguir entre matrices simétricas y antisimétricas. Para lo que resta de esta sección, convenimos de una vez por todas que la característica de F es diferente de 2.

Tenemos procedimientos muy sencillos para producir matrices simétricas y matrices antisimétricas. Por ejemplo, si A es una matriz arbitraria, entonces $A + A'$ es simétrica y $A - A'$ es antisimétrica. Si pensamos que $A = \frac{1}{2}(A + A') + \frac{1}{2}(A - A')$, vemos que toda matriz resulta ser la suma de una matriz simétrica y otra antisimétrica. Esta descomposición es única (véase el problema 19). Otro método de producir matrices simétricas es el que sigue: si A es una matriz arbitraria, entonces tanto AA' como $A'A$ son simétricas. (Nótese que no tienen porqué ser iguales.)

Está en la naturaleza de todo matemático que, una vez que se ha dado un concepto interesante surgido de una situación particular, ha de intentar despojarlo de las particularidades de su origen y emplear las propiedades claves del concepto como medio de hacerlo más abstracto. Procedemos a seguir tal camino con la transpuesta. Tomamos, como propiedades formales de mayor interés, aquellas que aparecen en el enunciado del lema 6.18 que afirma que sobre F_n la transpuesta define un antiautomorfismo de periodo 2. Nos lleva esto a la siguiente

DEFINICIÓN. Una aplicación de F_n en F_n se llama *adjunta* sobre F_n si

- 1) $(A^*)^* = A$;
- 2) $(A + B)^* = A^* + B^*$;
- 3) $(AB)^* = B^*A^*$;

para cualesquiera $A, B \in F_n$.

Nótese en que no insistimos en que $(\lambda A)^* = \lambda A^*$ para $\lambda \in F$. En realidad, en algunas de las adjuntas más interesantes este no es el caso. Pasamos a discutir una tal. Sea F el campo de los números complejos; para $A = (x_{ij}) \in F_n$, sea $A^* = (\gamma_{ij})$ donde $\gamma_{ij} = \bar{x}_{ji}$, el conjugado complejo de x_{ij} . En este caso * suele llamarse *adjunta hermitiana* sobre F_n . Dentro de unas pocas secciones haremos un estudio bastante extensivo de las matrices bajo la adjunta hermitiana.

Todo lo que hemos dicho acerca de la transpuesta como, por ejemplo, los conceptos de simetría y antisimetría, puede ser aplicado a las adjuntas generales, y hablamos de elementos simétricos bajo * (es decir, de aquellos A tales que $A^* = A$), de elementos antisimétricos bajo *, etc. En los ejercicios del final de esta sección, hay muchos ejemplos y problemas que se refieren a adjuntas en general.

Pero ahora, como diversión, juguemos un poco con la adjunta hermitiana. No llamamos a nada de lo que obtenemos un teorema, no porque no se merezcan tal título, sino más bien porque los volveremos a hacer más tarde (y los designaremos entonces propiamente) partiendo de un punto de vista central.

Así pues, supongamos que F es el campo de los números complejos y que la adjunta * sobre F_n es la adjunta hermitiana. La matriz A se llama *hermitiana* si $A^* = A$.

Primera observación: si $A \neq 0 \in F_n$ entonces $\text{tr}(AA^*) > 0$. Segunda observación: Como una consecuencia de la primera observación, si $A_1, \dots, A_k \in F_n$ y si $A_1A_1^* + A_2A_2^* + \dots + A_kA_k^* = 0$, entonces $A_1 = A_2 = \dots = A_k = 0$. Tercera observación: Si λ es una matriz escalar, entonces $\lambda^* = \bar{\lambda}$, el conjugado complejo de λ .

Supongamos que $A \in F_n$ es hermitiana y que el número complejo $\alpha + \beta i$, donde α y β son reales e $i^2 = -1$, es una raíz característica de A . Tenemos, pues, que $A - (\alpha + \beta i)$ no es invertible; pero entonces $(A - (\alpha + \beta i))(A - (\alpha - \beta i)) = (A - \alpha)^2 + \beta^2$ no es invertible. Pero si una matriz es singular debe eliminar una matriz distinta de cero (teorema 6.b, corolario 2). Debe haber, por tanto, una matriz $C \neq 0$ tal que $C((A - \alpha)^2 + \beta^2) = 0$. Multiplicamos esto a la derecha por C^* y obtenemos:

$$1) \quad C(A - \alpha)^2 C^* + \beta^2 CC^* = 0.$$

Sea $D = C(A - \alpha)$ y $E = \beta C$. Como $A^* = A$ y α es real, $C(A - \alpha)^2 C^* = DD^*$; como β es real, $\beta^2 CC^* = EE^*$. Luego la ecuación (1) toma la forma $DD^* + EE^* = 0$; por las observaciones antes hechas esto implica $D = 0$ y $E = 0$. Solamente vamos a usar la relación $E = 0$. Como $0 = E = \beta C$ y como $C \neq 0$, debemos tener $\beta = 0$. ¿Qué es exactamente lo que hemos probado? En realidad, hemos probado el bello e importante resultado de que si un número complejo λ es una raíz característica de una matriz hermitiana, entonces λ debe ser real. Aprovechando las propiedades del campo de los

números complejos se puede, realmente, reformular esto como sigue:
Las raíces características de una matriz hermitiana son, todas, reales.

Continuamos con esta vena un poco más adelante. Para $A \in F_n$, sea $B = AA^*$; B es una matriz hermitiana. Si el número real α es una raíz característica de B , ¿puede α ser un número real arbitrario o debe estar restringido de algún modo? Afirmamos que α debe ser no negativo. Pues si α fuera negativo entonces $\alpha = -\beta^2$, donde β es un número real. Pero entonces $B - \alpha = B + \beta^2 = AA^* + \beta^2$ no es invertible, de donde hay un $C \neq 0$ tal que $C(AA^* + \beta^2) = 0$. Multiplicando por C^* a la derecha y razonando como anteriormente, tenemos $\beta = 0$, una contradicción. Hemos demostrado que cualquier raíz característica real de AA^* debe ser no negativa. En realidad, lo de "real" en la anterior afirmación es superfluo y podríamos decir: para cualquier $A \in F_n$ todas las raíces características de AA^* son no negativas.

Problemas

1. Pruébese que $\text{tr}(A+B) = \text{tr} A + \text{tr} B$ y que para $\lambda \in F$, $\text{tr}(\lambda A) = \lambda \text{tr} A$.

2. a) Usando un argumento basado en la traza pruébese que si la característica de F es 0 entonces es imposible encontrar $A, B \in F_n$ tales que $AB - BA = 1$.
 b) En la parte (a), pruébese que en realidad $1 - (AB - BA)$ no puede ser nilpotente.

3. a) Sea f una función definida sobre F_n con valores en F tales que:

- 1) $f(A+B) = f(A)+f(B)$,
- 2) $f(\lambda A) = \lambda f(A)$,
- 3) $f(AB) = f(BA)$,

para todo $A, B \in F_n$ y para todo $\lambda \in F$. Pruébese que hay un elemento $\alpha_0 \in F$ tal que $f(A) = \alpha_0 \text{tr} A$ para todo A en F_n .

- b) Si la característica de F es 0 y si la f de la parte (a) satisface la propiedad adicional de que $f(1) = n$, pruébese que $f(A) = \text{tr} A$ para todo $A \in F_n$.

Nótese que el problema 3 caracteriza la función "traza".

*4. a) Si el campo F tiene un número infinito de elementos, pruébese que todo elemento en F_n puede escribirse como la suma de matrices regulares.

- b) Si F tiene un número infinito de elementos y si f , definido sobre F_n y con sus valores en F , satisface

- 1) $f(A+B) = f(A)+f(B)$,
- 2) $f(\lambda A) = \lambda f(A)$,
- 3) $f(BAB^{-1}) = f(A)$,

para toda $A \in F_n$, $\lambda \in F$ y todo elemento invertible B en F_n , pruébese que $f(A) = \alpha_0 \operatorname{tr} A$ para un $\alpha_0 \in F$ determinado y toda $A \in F_n$.

5. Pruébese que el lema de Jacobson para elementos A , B en F_n si n es menor que la característica de F .

- 6. a)** Si $C \in F_n$, definamos la aplicación d_C sobre F_n por $d_C(X) = XC - CX$ para toda $X \in F_n$. Pruébese que $d_C(XY) = (d_C(X))Y + X(d_C(Y))$. (¿No le recuerda esto al lector la derivada?)
- b)** Usando la parte (a), pruébese que si $AB - BA$ commuta con A , entonces para cualquier polinomio $q(x) \in F[x]$, $q(A)B - Bq(A) = q'(A)(AB - BA)$, donde $q'(x)$ es la derivada de $q(x)$.

***7.** Úsese la parte (b) del problema 6 para dar una prueba del lema de Jacobson. (*Sugerencia:* Sea $p(x)$ el polinomio mínimo para A y considérese $0 = p(A)B - Bp(A)$.)

- 8. a)** Si A es una matriz triangular, pruébese que las entradas sobre la diagonal de A son exactamente todas las raíces características de A .
- b)** Si A es triangular y los elementos en su diagonal principal son 0, pruébese que A es nilpotente.
- 9.** Para cualquier A , $B \in F_n$ y $\lambda \in F$ pruébese que $(A')' = A$, $(A + B)' = A' + B'$ y $(\lambda A)' = \lambda A'$.

10. Si A es invertible, pruébese que $(A^{-1})' = (A')$.

11. Si A es antisimétrica, pruébese que los elementos en su diagonal principal son, todos, cero.

12. Si A y B son matrices simétricas, pruébese que AB es simétrica si y sólo si $AB = BA$.

13. Proporciónese un ejemplo de una A tal que $AA' \neq A'A$.

***14.** Demuéstrese que A y A' son semejantes.

15. Los elementos simétricos en F_n forman un espacio vectorial; encuéntrese su dimensión y exhibase una de sus bases.

***16.** Denotemos por S el conjunto de los elementos simétricos de F_n ; pruébese que el subanillo de F_n generado por S es, todo, F_n .

***17.** Si la característica de F es 0 y $A \in F_n$ tiene traza 0 ($\operatorname{tr} A = 0$) pruébese que hay una $C \in F_n$ tal que CAC^{-1} tiene solamente 0 en su diagonal principal.

***18.** Si F es de característica 0 y $A \in F_n$ tiene traza 0, pruébese que existen B , $C \in F_n$ tales que $A = BC - CB$. (*Sugerencia:* Primer paso, supóngase, por el resultado del problema 17, que todos los elementos diagonales de A son 0.)

19. a) Si $*$ es cualquier adjunto sobre F_n , sea $S = \{A \in F_n : A^* = A\}$ y sea $K = \{A \in F_n | A^* = -A\}$. Pruébese que $S + K = F_n$.

b) Si $A \in F_n$ y $A = B + C$ donde $B \in S$ y $C \in K$, pruébese que B y C son únicos y determiníñense.

20. a) Si $A, B \in S$ pruébese que $AB + BA \in S$.

b) Si $A, B \in K$ pruébese que $AB - BA \in K$.

c) Si $A \in S$ y $B \in K$ pruébese que $AB - BA \in S$ y que $AB + BA \in K$.

21. Si ϕ es un automorfismo del campo F definimos la aplicación Φ sobre F_n por: si $A = (\alpha_{ij})$ entonces $\Phi(A) = (\phi(\alpha_{ij}))$. Pruébese que $\Phi(A+B) = \Phi(A) + \Phi(B)$ y que $\Phi(AB) = \Phi(A)\Phi(B)$ para toda $A, B \in F_n$.

22. Si $*$ y \odot definen dos adjuntos sobre F_n , pruébese que la aplicación $\psi : A \rightarrow (A^*) \odot$ para todo $A \in F_n$ satisface $\psi(A+B) = \psi(A) + \psi(B)$ y $\psi(AB) = \psi(A)\psi(B)$ para cualesquiera $A, B \in F_n$.

23. Si $*$ es un adjunto cualquiera sobre F_n y λ es una matriz escalar en F_n , pruébese que λ^* debe también ser una matriz escalar.

***24.** Supongamos que conocemos el siguiente teorema: si ψ es un automorfismo de F_n (es decir, ψ transforma F_n sobre él mismo, de tal modo que $\psi(A+B) = \psi(A) + \psi(B)$ y $\psi(AB) = \psi(A)\psi(B)$) tal que $\psi(\lambda) = \lambda$ para toda matriz escalar λ , entonces hay un elemento $P \in F_n$ tal que $\psi(A) = PAP^{-1}$ para todo $A \in F_n$. Basándose en este teorema, pruébese que: si $*$ es un adjunto de F_n tal que $\lambda^* = \lambda$ para toda matriz escalar λ , entonces existe una matriz $P \in F_n$ tal que $A^* = PA'P^{-1}$ para toda $A \in F_n$. Además, $P^{-1}P'$ debe ser un escalar.

25. Si $P \in F_n$ es tal que $P^{-1}P' \neq 0$ es un escalar, pruébese que la aplicación definida por $A^* = PA'P^{-1}$ es un adjunto sobre F_n .

***26.** Basándose en el teorema acerca de automorfismo enunciado en el problema 24, pruébese lo siguiente: Si $*$ es un adjunto sobre F_n hay un automorfismo ϕ de F de periodo 2 y un elemento $P \in F_n$ tales que $A^* = P(\Phi(A))'P^{-1}$ para todo $A \in F_n$ (para notación, véase el problema 21). Además, P , debe satisfacer $P^{-1}\Phi(P)'$ es un escalar.

Los problemas 24 y 26 indican que una adjunta general sobre F_n no está tan alejada de la transpuesta como se habría creído a primera vista.

****27.** Si ψ es un automorfismo de F_n tal que $\psi(\lambda) = \lambda$ para todos los escalares, pruébese que hay un $P \in F_n$ tal que $\psi(A) = PAP^{-1}$ para todo $A \in F_n$.

En el resto de los problemas, F será el campo de los números complejos y $$ la adjunta hermitiana sobre F_n .*

28. Si $A \in F_n$, pruébese que hay matrices hermitianas únicas B y C tales que $A = B + iC$ ($i^2 = -1$).

29. Pruébese que $\text{tr } AA^* > 0$ si $A \neq 0$.

30. Por cálculo directo de las entradas de las matrices, pruébese que si $A_1 A_1^* + \dots + A_k A_k^* = 0$, entonces $A_1 = A_2 = \dots = A_k = 0$.

31. Si A está en F_n y si $BAA^* = 0$, pruébese que $BA = 0$.

32. Si $A \in F_n$ es hermitiana y $BA^k = 0$, pruébese que $BA = 0$.

*33. Si $A \in F_n$ es hermitiana y si λ, μ son dos raíces características reales distintas de A y si $C(A - \lambda) = 0$ y $D(A - \mu) = 0$, pruébese que $CD = DC = 0$. (Sugerencia: Considérese primero el caso en que C y D son hermitianos y luego aplíquese el resultado del problema 31).

*34. a) Suponiendo que todas las raíces características de la matriz hermitiana A están en el campo de los números complejos, combinando los resultados de los problemas 32 y 33, y el hecho de que las raíces deben, por tanto, ser todas reales, y el resultado del corolario del teorema 6.n, pruébese que A puede ser puesta en forma diagonal; es decir, que hay una matriz P tal que PAP^{-1} es diagonal.

b) En la parte (a) pruébese que P puede escogerse de forma que $PP^* = 1$.

35. Sea $V_n = \{A \in F_n \mid AA^* = 1\}$. Pruébese que V_n es un grupo bajo la multiplicación de matrices.

36. Si A conmuta con $AA^* - A^*A$, pruébese que $AA^* = A^*A$.

9. DETERMINANTES

La traza define una función importante y útil del anillo de las matrices F_n (y de $A_F(V)$) en F ; sus propiedades se relacionan en su mayor parte con las propiedades aditivas de las matrices. Introduciremos ahora la función, aún más importante, conocida como el determinante, que transforma F_n en F . Sus propiedades están estrechamente ligadas con las propiedades multiplicativas de las matrices.

Aparte de su efectividad como argumento para probar teoremas, el determinante es valioso para usos "prácticos". Dada una matriz T , podemos construir en términos de determinantes explícitos un polinomio concreto cuyas raíces son las raíces características de T ; aún más, la multiplicidad de una raíz de este polinomio es igual a su multiplicidad como raíz característica de T . En realidad, el polinomio característico de T , definido anteriormente, puede exhibirse como este polinomio determinante explícitamente.

Los determinantes juegan también un papel fundamental en la solución de sistemas de ecuaciones lineales. Por esta dirección es por la que motivaremos su definición.

Hay muchas formas de desarrollar la teoría de determinantes, algunas muy elegantes y otras muy aburridas. Nosotros hemos escogido un camino distinto del de cualquiera de estos extremos, pero que para nosotros tiene la ventaja de que podemos alcanzar los resultados necesarios para nuestra discusión de las transformaciones lineales con la mayor rapidez posible.

En lo que sigue, F será un campo arbitrario, F_n el anillo de las matrices $n \times n$ sobre F , y $F^{(n)}$ el espacio vectorial de n -adas sobre F . Por una matriz entenderemos tácitamente un elemento en F_n . Como es usual, las letras griegas indicarán elementos de F (salvo advertencia en contra).

Consideremos el sistema de ecuaciones

$$\alpha_{11}x_1 + \alpha_{12}x_2 = \beta_1$$

$$\alpha_{21}x_1 + \alpha_{22}x_2 = \beta_2.$$

Nos preguntamos: ¿bajo qué condiciones sobre las α_{ij} podemos resolver para x_1 y x_2 con β_1 y β_2 dadas cualesquiera? O, lo que es equivalente, dada la matriz

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix},$$

¿cuándo esta matriz transforma $F^{(2)}$ sobre sí mismo?

Procediendo como en secundaria, eliminamos x_1 entre las dos ecuaciones; el criterio de solubilidad resulta, entonces, ser que $\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21} \neq 0$.

Pasamos ahora al sistema de tres ecuaciones lineales

$$\alpha_{11}x_1 + \alpha_{12}x_2 + \alpha_{13}x_3 = \beta_1$$

$$\alpha_{21}x_1 + \alpha_{22}x_2 + \alpha_{23}x_3 = \beta_2$$

$$\alpha_{31}x_1 + \alpha_{32}x_2 + \alpha_{33}x_3 = \beta_3,$$

y de nuevo nos preguntamos sobre las condiciones de solubilidad para β_1 , β_2 y β_3 arbitrarias. Eliminando x_1 entre estas dos a la vez, y luego x_2 de las restantes dos ecuaciones, obtenemos como criterio de solubilidad

$$\alpha_{11}\alpha_{22}\alpha_{33} + \alpha_{12}\alpha_{23}\alpha_{31} + \alpha_{13}\alpha_{21}\alpha_{32} - \alpha_{12}\alpha_{21}\alpha_{33}$$

$$- \alpha_{11}\alpha_{23}\alpha_{32} - \alpha_{13}\alpha_{22}\alpha_{31} \neq 0.$$

Usando estos dos como modelo (y con el presentimiento de que esto va a funcionar) daremos el gran salto hasta el caso general y definiremos el determinante de una matriz arbitraria $n \times n$ sobre F . Pero fíjémonos antes un poco en la notación.

Sea S_n el grupo simétrico de grado n ; consideramos que los elementos de S_n están actuando sobre el conjunto $\{1, 2, \dots, n\}$. Para $\sigma \in S_n$, $\sigma(i)$ denotará la imagen de i bajo σ . (Cambiamos la notación escribiendo la permutación como si actuara a la izquierda en lugar de, como previamente, a la derecha. Lo hacemos para facilitar la escritura de los subíndices.) El símbolo $(-1)^\sigma$ para $\sigma \in S_n$ indica $+1$ si σ es una permutación *par*, y -1 si es una permutación *impar*.

DEFINICIÓN. Si $A = (\alpha_{ij})$, entonces el *determinante de A*, lo que se escribe: $\det A$, es el elemento de $F \sum_{\sigma \in S_n} (-1)^\sigma \alpha_{1\sigma(1)} \alpha_{2\sigma(2)} \dots \alpha_{n\sigma(n)}$.

Usaremos a veces la notación

$$\left| \begin{array}{cccc} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{array} \right|$$

para el determinante de la matriz

$$\left(\begin{array}{cccc} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{array} \right).$$

Nótese que el determinante de una matriz A es la suma (si prescindimos, por el momento, de los signos) de todos los productos posibles de entradas de A en los que aparezcan uno y solo uno de cada renglón y columna. En general es una labor pesada desarrollar el determinante de una matriz—fíjémonos que hay nada menos que $n!$ términos en la expansión—mas para al menos un tipo de matriz podemos hacer este desarrollo visualmente, a saber

LEMA 6.19. *El determinante de una matriz triangular es el producto de sus entradas en la diagonal principal.*

Prueba. Ser triangular implica dos posibilidades, a saber, o todos los elementos por encima de la diagonal principal son 0, o todos los elementos por debajo de la diagonal principal son 0. Probaremos aquí el resultado para A de la forma

$$\left(\begin{array}{cccc} \alpha_{11} & 0 & \cdots & 0 \\ \alpha_{22} & \ddots & & \vdots \\ * & \ddots & \ddots & \vdots \\ & & & \alpha_{nn} \end{array} \right)$$

e indicaremos el pequeño cambio en el argumento a emplear para la otra clase de matrices triangulares.

Como $\alpha_{1i} = 0$ salvo si $i = 1$, en la expansión de $\det A$, la única contribución no nula viene de aquellos términos donde $\sigma(1) = 1$. Así pues, como σ es una permutación, $\sigma(2) \neq 1$; pero si $\sigma(2) > 2$, $\alpha_{2\sigma(2)} = 0$; luego, para obtener una contribución no nula a $\det A$, $\sigma(2) = 2$. Continuando de esta forma, debemos tener $\sigma(i) = i$ para todo i , lo que es lo mismo que decir que en la expansión de $\det A$ el único término distinto de cero se presenta cuando σ es el elemento identidad de S_n . De aquí que la suma de los $n!$ términos se reduce a exactamente uno solo, $\alpha_{11}\alpha_{22} \dots \alpha_{nn}$, que es lo que el teorema afirma.

Si A es una triangular inferior comenzamos con el extremo opuesto probando que para una contribución distinta de cero $\sigma(n) = n$, luego que $\sigma(n-1) = n-1$, etc.

Algunos casos especiales son de interés:

1) Si

$$A = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

es diagonal, $\det A = \lambda_1 \lambda_2 \dots \lambda_n$.

2) Si

$$A = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix},$$

la matriz identidad, entonces $\det A = 1$.

3) Si

$$A = \begin{pmatrix} \lambda & & & \\ & \lambda & & \\ & & \ddots & \\ & & & \lambda \end{pmatrix},$$

la matriz escalar, entonces $\det A = \lambda^n$.

Obsérvese también que si un renglón o columna de una matriz está compuesta solo de ceros, entonces el determinante es 0, pues cada término del desarrollo del determinante será un producto en el que al menos uno de los factores es 0, de donde cada término es 0.

Dada la matriz $A = (\alpha_{ij})$ en F_n podemos considerar su primera fila $v_1 = (\alpha_{11}, \alpha_{12}, \dots, \alpha_{1n})$ como un vector en $F^{(n)}$, y análogamente para su segunda fila, v_2 , y las restantes. Podemos considerar entonces $\det A$ como una función de los n vectores v_1, \dots, v_n . Muchos resultados se pueden enunciar más sucintamente en estos términos, por lo que a menudo consideraremos $\det A = d(v_1, \dots, v_n)$; en este caso *la notación siempre se entiende que implica que v_1 es el primer renglón, v_2 el segundo, y así sucesivamente, de A* .

Una observación más: aunque estamos trabajando sobre un campo, podríamos sin la menor dificultad suponer que estábamos trabajando sobre un anillo conmutativo, excepto en las obvias ocasiones en que dividimos por elementos. Esta observación solamente vendrá a cuenta cuando discutamos determinantes de matrices que tengan entradas polinomiales, lo que haremos dentro de poco en esta misma sección.

LEMA 6.20. Si $A \in F_n$ y $\gamma \in F$, entonces $d(v_1, \dots, v_{i-1}, \gamma v_i, v_{i+1}, \dots, v_n) = \gamma d(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_n)$.

Nótese que el lema dice que si todos los elementos de un renglón de A son multiplicados por un elemento fijo γ de F , entonces el determinante de A queda también multiplicado por γ .

Prueba. Como solamente las entradas de la i -ésima fila han cambiado, el desarrollo de $d(v_1, \dots, v_{i-1}, \gamma v_i, v_{i+1}, \dots, v_n)$ es

$$\sum_{\sigma \in S_n} (-1)^\sigma \alpha_{1\sigma(1)} \cdots \alpha_{i-1,\sigma(i-1)} (\gamma \alpha_{i\sigma(i)}) \alpha_{i+1,\sigma(i+1)} \cdots \alpha_{n\sigma(n)};$$

como esto es igual a $\gamma \sum_{\sigma \in S_n} (-1)^\sigma \alpha_{1\sigma(1)} \cdots \alpha_{i\sigma(i)} \cdots \alpha_{n\sigma(n)}$, es claro que es igual a $\gamma d(v_1, \dots, v_n)$.

LEMA 6.21. $d(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_n) + d(v_1, \dots, v_{i-1}, u_i, v_{i+1}, \dots, v_n) = d(v_1, \dots, v_{i-1}, v_i + u_i, v_{i+1}, \dots, v_n)$.

Antes de probar el resultado, veamos qué es lo que dice y lo que no dice. No dice que $\det A + \det B = \det(A + B)$; esto es falso como puede verse en el ejemplo

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

donde $\det A = \det B = 0$ mientras que $\det(A + B) = 1$. Dice que si A y B son matrices iguales en todas partes salvo en el i -ésimo renglón entonces

la nueva matriz obtenida de A y B usando todos los renglones de A excepto el i -ésima, y usando como i -ésimo renglón la suma de los i -ésimos renglones de A y B , tiene un determinante igual a $\det A + \det B$. Si

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

y

$$B = \begin{pmatrix} 1 & 1 \\ 3 & 4 \end{pmatrix},$$

entonces

$$\det A = -2, \quad \det B = 1, \quad \det \begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix} = -1 = \det A + \det B.$$

Prueba. Si $v_1 = (\alpha_{11}, \dots, \alpha_{1n}), \dots, v_i = (\alpha_{i1}, \dots, \alpha_{in}), \dots, v_n = (\alpha_{n1}, \dots, \alpha_{nn})$ y si $u_i = (\beta_{i1}, \dots, \beta_{in})$, entonces

$$\begin{aligned} d(v_1, \dots, v_{i-1}, u_i + v_i, v_{i+1}, \dots, v_n) \\ &= \sum_{\sigma \in S_n} (-1)^\sigma \alpha_{1\sigma(1)} \cdots \alpha_{i-1,\sigma(i-1)} (\alpha_{i\sigma(i)} + \beta_{i\sigma(i)}) \alpha_{i+1,\sigma(i+1)} \cdots \alpha_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} (-1)^\sigma \alpha_{1\sigma(1)} \cdots \alpha_{i-1,\sigma(i-1)} \alpha_{i\sigma(i)} \cdots \alpha_{n\sigma(n)} \\ &\quad + \sum_{\sigma \in S_n} (-1)^\sigma \alpha_{1\sigma(1)} \cdots \alpha_{i-1,\sigma(i-1)} \beta_{i\sigma(i)} \cdots \alpha_{n\sigma(n)} \\ &= d(v_1, \dots, v_i, v_n) + d(v_1, \dots, u_i, \dots, v_n). \end{aligned}$$

Las propiedades que aparecen en los lemas 6.19, 6.20 y 6.21, junto con las que aparecen en el próximo lema, puede demostrarse que caracterizan a la función determinante (véase el problema 13 al final de esta sección). Así pues, la propiedad formal exhibida en el siguiente lema es básica en la teoría de determinantes.

LEMA 6.22. *Si dos renglones de A son iguales (es decir, si $v_r = v_s$ para $r \neq s$), entonces $\det A = 0$.*

Prueba. Sea $A = (\alpha_{ij})$ y supongamos que para ciertos r, s con $r \neq s$ $\alpha_{rj} = \alpha_{sj}$ para todo j . Consideremos el desarrollo

$$\det A = \sum_{\sigma \in S_n} (-1)^\sigma \alpha_{1\sigma(1)} \cdots \alpha_{r\sigma(r)} \cdots \alpha_{s\sigma(s)} \cdots \alpha_{n\sigma(n)}.$$

En el desarrollo, apareamos los términos como sigue: Para $\sigma \in S_n$ apareamos el término $(-1)^\sigma \alpha_{1\sigma(1)} \cdots \alpha_{n\sigma(n)}$ con el término $(-1)^{\sigma'} \alpha_{1\tau\sigma(1)} \cdots \alpha_{n\tau\sigma(n)}$,

donde τ es la transposición $(\sigma(r), \sigma(s))$. Como τ es una transposición y $\tau^2 = 1$, esto nos da, ciertamente, un aparejamiento. Pero como $\alpha_{r\sigma(r)} = \alpha_{s\sigma(r)}$, por hipótesis, y $\alpha_{s\sigma(r)} = \alpha_{s\tau\sigma(s)}$, tenemos que $\alpha_{r\sigma(r)} = \alpha_{s\tau\sigma(s)}$. Análogamente, $\alpha_{s\sigma(s)} = \alpha_{r\tau\sigma(r)}$. Por otra parte, para $i \neq r$ y $i \neq s$, como $\tau\sigma(i) = \sigma(i)$, $\alpha_{i\sigma(i)} = \alpha_{i\tau\sigma(i)}$. Luego los términos $\alpha_{1\sigma(1)} \dots \alpha_{n\sigma(n)}$ y $\alpha_{1\tau\sigma(1)} \dots \alpha_{n\tau\sigma(n)}$ son iguales. El primero aparece con el signo $(-1)^{\sigma}$ y el segundo con el signo $(-1)^{\tau\sigma}$ en la expansión de $\det A$. Como τ es una transposición y por tanto una permutación impar, $(-1)^{\tau\sigma} = -(-1)^{\sigma}$. Por tanto, en el aparejamiento, los términos apareados se cancelan mutuamente en la suma, de donde $\det A = 0$. (La prueba no depende de la característica de F y es igualmente válida incluso en el caso de característica 2.)

De acuerdo con los resultados hasta ahora obtenidos, podemos determinar el efecto sobre un determinante de una matriz dada de una permutación de sus renglones.

LEMA 6.23. *El intercambio de dos renglones de A cambia el signo de su determinante.*

Prueba. Como hay dos renglones iguales, según el lema 6.22, $d(v_1, \dots, v_{i-1}, v_i + v_j, v_{i+1}, \dots, v_{j-1}, v_i + v_j, v_{j+1}, \dots, v_n) = 0$. Usando el lema 6.21 varias veces podemos desarrollar esto para obtener $d(v_1, \dots, v_{i-1}, v_i, \dots, v_{j-1}, v_j, \dots, v_n) + d(v_1, \dots, v_{i-1}, v_j, \dots, v_{j-1}, v_i, \dots, v_n) + d(v_1, \dots, v_{i-1}, v_i, \dots, v_{j-1}, v_i, \dots, v_n) + d(v_1, \dots, v_{i-1}, v_j, \dots, v_{j-1}, v_j, \dots, v_n) = 0$. Pero cada uno de los últimos dos términos tiene en él dos renglones iguales. de donde, según el lema 6.22, cada uno es 0. La anterior relación se reduce entonces a $d(v_1, \dots, v_{i-1}, v_i, \dots, v_{j-1}, v_j, \dots, v_n) + d(v_1, \dots, v_{i-1}, v_j, \dots, v_{j-1}, v_i, \dots, v_n) = 0$, que es precisamente lo que el lema afirma.

COROLARIO. *Si la matriz B se obtiene de la A mediante una permutación de los renglones de A , entonces $\det A = \pm \det B$, siendo el signo +1 si la permutación es par, y -1 si la permutación es impar.*

Estamos ahora en posición de unir piezas para probar la propiedad algebraica básica de la función determinante, a saber, que preserva los productos. Como un homomorfismo de la estructura multiplicativa de F_n en F el determinante adquirirá ciertas características importantes.

TEOREMA 6.1. *Para $A, B \in F_n$, $\det(AB) = (\det A)(\det B)$.*

Prueba. Sea $A = (\alpha_{ij})$ y $B = (\beta_{ij})$; sean las filas de B los vectores u_1, u_2, \dots, u_n . Introducimos los n vectores w_1, \dots, w_n como sigue:

$$w_1 = \alpha_{11}u_1 + \alpha_{12}u_2 + \dots + \alpha_{1n}u_n,$$

$$w_2 = \alpha_{21}u_1 + \alpha_{22}u_2 + \dots + \alpha_{2n}u_n, \dots, w_n = \alpha_{n1}u_1 + \dots + \alpha_{nn}u_n.$$

Consideremos $d(w_1, \dots, w_n)$; desarrollando este determinante y haciendo un uso múltiple de los lemas 6.20 y 6.21, obtenemos

$$d(w_1, \dots, w_n) = \sum_{i_1, i_2, \dots, i_n} \alpha_{1i_1} \alpha_{2i_2} \cdots \alpha_{ni_n} d(u_{i_1}, u_{i_2}, \dots, u_{i_n}).$$

En esta suma múltiple i_1, \dots, i_n van tomando independientemente todos los valores desde 1 hasta n . Pero, si cualesquiera dos $i_r = i_s$ entonces $u_{i_r} = u_{i_s}$, de donde $d(u_{i_1}, \dots, u_{i_r}, \dots, u_{i_s}, \dots, u_{i_n}) = 0$ por el lema 6.22. En otras palabras, los únicos términos en la suma que pueden dar una contribución distinta de cero son aquellos para los que todo los i_1, i_2, \dots, i_n son distintos, es decir, aquellos para los que la aplicación

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix}$$

es una permutación de $1, 2, \dots, n$. Además, cualquier permutación tal es posible.

Observemos finalmente que según el corolario del lema 6.23, cuando

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix}$$

es una permutación, entonces $d(u_{i_1}, u_{i_2}, \dots, u_{i_n}) = (-1)^\sigma d(u_1, \dots, u_n) = (-1)^\sigma \det B$. Tenemos así

$$\begin{aligned} d(w_1, \dots, w_n) &= \sum_{\sigma \in S_n} \alpha_{1\sigma(1)} \cdots \alpha_{n\sigma(n)} (-1)^\sigma \det B \\ &= (\det B) \sum_{\sigma \in S_n} (-1)^\sigma \alpha_{1\sigma(1)} \cdots \alpha_{n\sigma(n)} \\ &= (\det B) (\det A). \end{aligned}$$

Deseamos ahora identificar ahora $d(w_1, \dots, w_n)$ como $\det(AB)$. Pero como $w_1 = \alpha_{11}u_1 + \dots + \alpha_{1n}u_n$,

$$w_2 = \alpha_{21}u_1 + \dots + \alpha_{2n}u_n, \dots, w_n = \alpha_{n1}u_1 + \dots + \alpha_{nn}u_n$$

tenemos que $d(w_1, \dots, w_n)$ es $\det C$, donde el primer renglón de C es w_1 , la segunda es w_2 , etc.

Pero si desarrollamos w_1 en términos de coordenadas obtenemos

$$\begin{aligned} w_1 &= \alpha_{11}u_1 + \dots + \alpha_{1n}u_n = \alpha_{11}(\beta_{11}, \beta_{12}, \dots, \beta_{1n}) \\ &\quad + \dots + \alpha_{1n}(\beta_{n1}, \dots, \beta_{nn}) \\ &= (\alpha_{11}\beta_{11} + \alpha_{12}\beta_{12} + \dots + \alpha_{1n}\beta_{1n}, \alpha_{11}\beta_{12} + \dots \\ &\quad + \alpha_{1n}\beta_{n2}, \dots, \alpha_{11}\beta_{1n} + \dots + \alpha_{1n}\beta_{nn}) \end{aligned}$$

que es el primer renglón de AB . Análogamente w_2 es el segundo renglón de AB , y así sucesivamente, para el resto de los renglones. Luego $C = AB$. Como $\det(AB) = \det C = d(w_1, \dots, w_n) = (\det A)(\det B)$, hemos probado el teorema.

COROLARIO 1. Si A es invertible entonces $\det A \neq 0$ y $\det(A^{-1}) = (\det A)^{-1}$

Prueba. Como $AA^{-1} = I$, $\det(AA^{-1}) = \det I = 1$. Luego según el teorema, $1 = \det(AA^{-1}) = (\det A)(\det A^{-1})$. Esta relación afirma entonces que $\det A \neq 0$ y $\det A^{-1} = \frac{1}{\det A}$.

COROLARIO 2. Si A es invertible, entonces para toda B , $\det(ABA^{-1}) = \det B$.

Prueba. Usando el teorema en la forma en que se aplicó a $(AB)A^{-1}$ tenemos $\det((AB)A^{-1}) = \det(AB) \det(A^{-1}) = \det A \det B \det(A^{-1})$. Aplicando el corolario 1 esto se reduce a $\det B$. Luego $\det(ABA^{-1}) = \det B$.

El corolario 2 nos permite definir el determinante de una transformación lineal. Pues si $T \in A(V)$ y $m_1(T)$ es la matriz de T en alguna base de V , para otra base, si $m_2(T)$ es la matriz en esta segunda base, entonces, según el teorema 6.h, $m_2(T) = Cm_1(T)C^{-1}$, de donde $\det(m_2(T)) = \det(m_1(T))$ según el anterior corolario 2. Es decir, la matriz de T en cualquier base tiene el mismo determinante. Luego la definición: $\det T = \det m_1(T)$ es en realidad independiente de la base y provee a $A(V)$ de una función determinante.

En uno de los primeros problemas, la finalidad del problema era la de probar que A' , la matriz transpuesta de la A , es semejante a A . Si esto fuera cierto (y lo es), entonces A' y A de acuerdo con el corolario 2 anterior tendrían el mismo determinante. No es, pues, motivo de asombro que podamos dar una prueba directa de este hecho.

LEMA 6.24. $\det A = \det A'$.

Prueba. Sea $A = (\alpha_{ij})$ y $A' = (\beta_{ij})$; desde luego, $\beta_{ij} = \alpha_{ji}$. Ahora bien

$$\det A = \sum_{\sigma \in S_n} (-1)^\sigma \alpha_{1\sigma(1)} \cdots \alpha_{n\sigma(n)}$$

mientras que

$$\det A' = \sum_{\sigma \in S_n} (-1)^\sigma \beta_{1\sigma(1)} \cdots \beta_{n\sigma(n)} = \sum_{\sigma \in S_n} (-1)^\sigma \alpha_{\sigma(1)1} \cdots \alpha_{\sigma(n)n}.$$

Pero el término $(-1)^\sigma \alpha_{\sigma(1)} \cdots \alpha_{\sigma(n)}$ es igual a $(-1)^\sigma \alpha_{1\sigma^{-1}(1)} \cdots \alpha_{n\sigma^{-1}(n)}$. (pruébese). Pero σ y σ^{-1} son de la misma paridad, es decir, si σ es impar, entonces también lo es σ^{-1} , mientras que si σ es par entonces σ^{-1} es par.

Luego

$$(-1)^\sigma \alpha_{1\sigma^{-1}(1)} \cdots \alpha_{n\sigma^{-1}(n)} = (-1)^{\sigma^{-1}} \alpha_{1\sigma^{-1}(1)} \cdots \alpha_{n\sigma^{-1}(n)}.$$

Finalmente, como σ recorre S_n , σ^{-1} recorre también por ello S_n . Luego

$$\begin{aligned} \det A' &= \sum_{\sigma^{-1} \in S_n} (-1)^{\sigma^{-1}} \alpha_{1\sigma^{-1}(1)} \cdots \alpha_{n\sigma^{-1}(n)} \\ &= \sum_{\sigma \in S_n} (-1)^\sigma \alpha_{1\sigma(1)} \cdots \alpha_{n\sigma(n)} \\ &= \det A. \end{aligned}$$

A la luz del lema 6.24, el intercambio de los renglones y las columnas de una matriz no cambia su determinante. *Pero entonces los lemas 6.20, 6.21, 6.22 y 6.23 que son válidos para operaciones con renglones de la matriz, se verifican igualmente para las columnas de la matriz.*

Hacemos un uso inmediato de la observación para derivar la *regla de Cramer* para la resolución de un sistema de ecuaciones lineales.

Dado el sistema de ecuaciones lineales:

$$\begin{aligned} \alpha_{11}x_1 + \dots + \alpha_{1n}x_n &= \beta_1 \\ \vdots \\ \alpha_{n1}x_1 + \dots + \alpha_{nn}x_n &= \beta_n, \end{aligned}$$

llamamos a $A = (\alpha_{ij})$ la matriz del sistema y a $\Delta = \det A$ el *determinante del sistema*.

Supongamos que $\Delta \neq 0$; es decir, que

$$\Delta = \begin{vmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{vmatrix} \neq 0.$$

De acuerdo con el lema 6.20 (en su forma modificada para columnas en lugar de para renglones),

$$x_i \Delta = \begin{vmatrix} \alpha_{11} & \cdots & \alpha_{1i}x_i & \cdots & \alpha_{1n} \\ \vdots & & \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{ni}x_i & \cdots & \alpha_{nn} \end{vmatrix}.$$

Pero como una consecuencia de los lemas 6.21 y 6.22, podemos añadir

cualquier múltiplo de una columna a otra sin cambiar el determinante (véase el problema 5). Añádase a la i -ésima de $x_i \Delta$, x_i veces la primera columna, x_2 veces la segunda, ..., x_j veces la j -ésima (para todo $j \neq i$). Así pues

$$x_i \Delta = \begin{vmatrix} \alpha_{11} & \cdots & \alpha_{1,i-1} & (\alpha_{11}x_1 + \alpha_{12}x_2 + \cdots + \alpha_{1n}x_n)x_{1,i+1} & \cdots & \alpha_{1n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \cdots & \alpha_{n,i-1} & (\alpha_{n1}x_1 + \alpha_{n2}x_2 + \cdots + \alpha_{nn}x_n)x_{n,i+1} & \cdots & \alpha_{nn} \end{vmatrix}$$

y usando $\alpha_{k1}x_1 + \cdots + \alpha_{kn}x_n = \alpha_k$, vemos finalmente que

$$x_i \Delta = \begin{vmatrix} \alpha_{11} & \cdots & \alpha_{1,i-1} & \beta_1 & \alpha_{1,i+1} & \cdots & \alpha_{1n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \alpha_{in} & \cdots & \alpha_{n,i-1} & \beta_n & \alpha_{n,i+1} & \cdots & \alpha_{nn} \end{vmatrix} = \Delta_i,$$

De donde, $x_i = \frac{\Delta_i}{\Delta}$. Esto es

TEOREMA 6.U. (REGLA DE CRAMER). Si es determinante Δ del sistema de ecuaciones lineales

$$\begin{aligned} \alpha_{11}x_1 + \cdots + \alpha_{1n}x_n &= \beta_1 \\ &\vdots \\ \alpha_{n1}x_1 + \cdots + \alpha_{nn}x_n &= \beta_n \end{aligned}$$

es diferente de cero, entonces la solución del sistema viene dada por $x_i = \frac{\Delta_i}{\Delta}$, donde Δ_i es el determinante obtenido de Δ al reemplazar en la i -ésima columna por $\beta_1, \beta_2, \dots, \beta_n$.

Ejemplo. El sistema

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= -5 \\ 2x_1 + x_2 + x_3 &= -7 \\ x_1 + x_2 + x_3 &= 0 \end{aligned}$$

tiene determinante

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

de donde

$$x_1 = \frac{\begin{vmatrix} -5 & 2 & 3 \\ -7 & 1 & 1 \\ 0 & 1 & 1 \end{vmatrix}}{\Delta}, \quad x_2 = \frac{\begin{vmatrix} 1 & -5 & 3 \\ 2 & -7 & 1 \\ 1 & 0 & 1 \end{vmatrix}}{\Delta}, \quad x_3 = \frac{\begin{vmatrix} 1 & 2 & -5 \\ 2 & 1 & -7 \\ 1 & 1 & 0 \end{vmatrix}}{\Delta}.$$

Podemos relacionar la invertibilidad de una matriz (o transformación lineal) con el valor de su determinante. El determinante nos provee, por tanto, de un criterio de invertibilidad.

TEOREMA 6.v. *A es invertible si y sólo si $\det A \neq 0$.*

Prueba. Si A es invertible, hemos visto en el corolario 1 del teorema 6.t, que $\det A \neq 0$.

Supongamos ahora que el $\det A \neq 0$ donde $A = (a_{ij})$. Según la regla de Cramer, podemos resolver el sistema

$$\begin{aligned} \alpha_{11}x_1 + \dots + \alpha_{1n}x_n &= \beta_1 \\ &\vdots \\ \alpha_{n1}x_1 + \dots + \alpha_{nn}x_n &= \beta_n \end{aligned}$$

para x_1, \dots, x_n dando β_1, \dots, β_n arbitrarios. Como una transformación lineal sobre $F^{(n)}$, A' es pues suprayectiva, en realidad el vector $(\beta_1, \dots, \beta_n)$ es la imagen bajo A' de $\left(\frac{\Delta_1}{\Delta}, \dots, \frac{\Delta_n}{\Delta}\right)$. Por ser suprayectiva, según el teorema 6.d, A' es invertible, de donde A es invertible (pruébese).

Podemos ver el teorema 6.v desde un punto de vista alternativo y probablemente más interesante. Dada $A \in F_n$ podemos sumergirla en K_n donde K es una extensión de F escogida de modo tal que en K_n , A pueda ser puesta en forma triangular. Hay, por tanto, un $B \in K_n$ tal que

$$BAB^{-1} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ * & \lambda_2 & \ddots & \vdots \\ & * & \ddots & \vdots \\ & & & \lambda_n \end{pmatrix};$$

aquí $\lambda_1, \dots, \lambda_n$ son todas las raíces características de A , cada una apareciendo tantas veces como unidades tiene su multiplicidad como raíces características de A . Así pues, $\det A = \det(BAB^{-1}) = \lambda_1 \lambda_2 \dots \lambda_n$ según el lema 6.19. Pero A es invertible si y sólo si ninguna de sus raíces características es cero;

pero $\det A \neq 0$ si y sólo si $\lambda_1, \lambda_2, \dots, \lambda_n \neq 0$, es decir, si ninguna de las raíces características de A es 0. Luego A es invertible si y sólo si $\det A \neq 0$.

Este argumento alternativo tiene algunas ventajas, pues al efectuarlo probamos realmente un subresultado interesante por sí mismo, a saber

LEMA 6.25. *$\det A$ es el producto, contando las multiplicidades, de las raíces características de A .*

DEFINICIÓN. Dada $A \in F_n$, la *ecuación secular* de A es el polinomio $\det(x - A)$ en $F[x]$.

Generalmente lo que hemos llamado la ecuación secular de A se suele llamar polinomio característico de A . Pero hemos definido ya el polinomio característico de A como el producto de sus divisores elementales. Es un hecho (véase el problema 8) que el polinomio característico de A es igual a su ecuación secular, pero como nosotros no necesitamos desarrollar esto explícitamente en el texto, introducimos el término de ecuación secular.

Calculemos un ejemplo. Si

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix},$$

entonces

$$x - A = \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix} - \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix} = \begin{pmatrix} x-1 & -2 \\ -3 & x \end{pmatrix};$$

de donde $\det(x - A) = (x-1)x - (-2)(-3) = x^2 - x - 6$. Así pues, la ecuación secular de

$$\begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix}$$

es $x^2 - x - 6$.

Unas cuantas observaciones acerca de la ecuación secular: Si λ es una raíz de $\det(x - A)$, entonces $\det(\lambda - A) = 0$; de donde, según el teorema 6.v, $\lambda - A$ no es invertible. Así pues, λ es una raíz característica de A . Recíprocamente, si λ es una raíz característica de A , $\lambda - A$ no es invertible, de donde $\det(\lambda - A) = 0$ y, por tanto, λ es una raíz de $\det(x - A)$. Así pues, el polinomio explícito y computable “ecuación secular de A ”, nos proporciona un polinomio cuyas raíces son exactamente las raíces características de A . Necesitamos subir un escalón más y probar que una raíz dada entra como una raíz de la ecuación secular precisamente tantas veces como su multiplici-

dad como raíz característica de A . En efecto, si λ_i es la raíz característica de A con multiplicidad m_i , podemos poner A en forma triangular de modo que

$$BAB^{-1} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ & \ddots & & \\ & & \lambda_1 & \\ & & & \lambda_2 \\ & & & & \ddots \\ & & & & & \lambda_2 \\ & & * & & & \\ & & & \lambda_k & 0 & \\ & & & & \ddots & \\ & & & & & \lambda_k \end{pmatrix}.$$

donde cada λ_i aparece en la diagonal m_i veces. Pero

$$B(x - A)B^{-1} = x - BAB^{-1} =$$

$$= \begin{pmatrix} x - \lambda_1 & 0 & \cdots & 0 \\ & \ddots & & \\ & & x - \lambda_1 & \\ & & & x - \lambda_2 \\ & & & & \ddots \\ & & & & & x - \lambda_2 \\ & & * & & & \\ & & & x - \lambda_k & & \\ & & & & \ddots & \\ & & & & & x - \lambda_k \end{pmatrix}$$

de modo que $\det(x - A) = \det(B(x - A)B^{-1}) = (x - \lambda_1)^{m_1}(x - \lambda_2)^{m_2} \dots (x - \lambda_k)^{m_k}$, y, por tanto, cada λ_i , cuya multiplicidad como raíz característica de A es m_i , es una raíz del polinomio $\det(x - A)$ de multiplicidad exactamente igual a m_i . Y hemos probado el

TEOREMA 6.w. *Las raíces características de A son las raíces, con la multiplicidad correcta, de la ecuación secular, $\det(x - A)$, de A .*

Damos término a esta sección con el significativo e histórico *teorema de Cayley-Hamilton*.

TEOREMA 6.x. *Toda $A \in F_n$ satisface su ecuación secular.*

Prueba. Dada cualquier matriz invertible $B \in K_n$, donde K es una extensión cualquiera de F , $A \in F$ y BAB^{-1} satisfacen los mismos polinomios. Además, como $\det(x - BAB^{-1}) = \det(B(x - A)B^{-1}) = \det(x - A)$, BAB^{-1} y A tienen la misma ecuación secular. Si podemos demostrar que algún BAB^{-1} satisface su ecuación secular, se seguirá de ello entonces que A también la satisface. Pero podemos escoger $K \supset F$ y $B \in K_n$ de modo que BAB^{-1} sea triangular; en tal caso ya vimos bastante antes (teorema 6.k) que una matriz triangular satisface su ecuación secular. Luego el teorema queda probado.

Problemas

1. Si F es el campo de los números complejos, evalúense los siguientes determinantes:

$$a) \begin{vmatrix} 1 & i \\ 2-i & 3 \end{vmatrix}. \quad b) \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}. \quad c) \begin{vmatrix} 5 & 6 & 8 & -1 \\ 4 & 3 & 0 & 0 \\ 10 & 12 & 16 & -2 \\ 1 & 2 & 3 & 4 \end{vmatrix}$$

2. ¿Para qué características de F son 0 los siguientes determinantes?

$$a) \begin{vmatrix} 1 & 2 & 3 & 0 \\ 3 & 2 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 2 & 4 & 5 & 6 \end{vmatrix} ? \quad b) \begin{vmatrix} 3 & 4 & 5 \\ 4 & 5 & 3 \\ 5 & 3 & 4 \end{vmatrix} ?$$

3. Si A es una matriz con entradas enteras tales que A^{-1} es también una matriz con entradas enteras, ¿cuáles pueden ser los valores de $\det A$?

4. Pruébese que si se suma el múltiplo de un renglón a otro no se cambia el valor del determinante.

*5. Dada la matriz $A = (a_{ij})$, sea A_{ij} la matriz obtenida de la A quitando el i -ésimo renglón y la j -ésima columna. Sea $M_{ij} = (-1)^{i+j} \det A_{ij}$. A M_{ij} se le suele llamar *cofactor* de a_{ij} . Pruébese que $\det A = a_{11}M_{11} + \dots + a_{nn}M_{nn}$.

6. a) Si A y B son submatrices cuadradas, pruébese que

$$\det \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = (\det A)(\det B).$$

b) Generalícese la parte (a) a

$$\det \begin{pmatrix} A_1 & & * & & \\ & A_2 & & & \\ & & \ddots & & \\ & & & A_n & \end{pmatrix},$$

donde cada A_i es una submatriz cuadrada.

7. Si $C(f)$ es la matriz compañera del polinomio $f(x)$, pruébese que la ecuación secular de $C(f)$ es $f(x)$.

8. Usando los problemas 6 y 7 pruébese que la ecuación secular de A es su polinomio característico. (Véase la sección 1; esto prueba la observación que antes hicimos de que las raíces de $p_T(x)$ aparecen con multiplicidades iguales a sus multiplicidades como raíces características de T .)

9. Usando el problema 8, proporcionese una prueba alternativa del teorema de Cayley-Hamilton.

10. Si F es el campo de los números racionales, calcúlense la ecuación secular y las raíces características con sus multiplicidades de

$$a) \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \quad b) \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 4 \\ 3 & 4 & 7 \end{pmatrix}. \quad c) \begin{pmatrix} 4 & 1 & 1 & 1 \\ 1 & 4 & 1 & 1 \\ 1 & 1 & 4 & 1 \\ 1 & 1 & 1 & 4 \end{pmatrix}.$$

11. Para cada una de las matrices de problema 10, verifíquese, por cálculo matricial directo, que satisface su ecuación secular.

*12. Si el rango de A es r , pruébese que hay una submatriz cuadrada $r \times r$ de A de determinante distinto de 0, y si $r < n$, que no hay ninguna submatriz $(r+1) \times (r+1)$ de A con esta propiedad.

*13. Sea f una función de n variables de $F^{(n)}$ a F tal que:

- $f(v_1, \dots, v_n) = 0$ para $v_i = v_j \in F^{(n)}$ con $i \neq j$.
- $f(v_1, \dots, \alpha v_i, \dots, v_n) = \alpha f(v_1, \dots, v_n)$ para toda i y $\alpha \in F$.
- $f(v_1, \dots, v_i + u_i, v_i + 1, \dots, v_n) = f(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_n) + f(v_1, \dots, v_{i-1}, u_i, v_{i+1}, \dots, v_n)$.
- $f(e_1, \dots, e_n) = 1$, donde $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0)$, ..., $e_n = (0, 0, \dots, 0, 1)$.

Pruébese que $f(v_1, \dots, v_n) = \det A$ para cualquier $A \in F_n$, donde v_i es la primera fila de A , v_2 la segunda, etc.

14. Úsese el problema 13 para probar que $\det A' = \det A$.

15. a) Pruébese que AB y BA tienen la misma ecuación secular (característica).

- b) Proporcionese un ejemplo en donde AB y BA no tengan el mismo polinomio mínimo.

16. Si A es triángular pruébese, por un cálculo directo, que A satisface su ecuación secular.

17. Úsese la regla de Cramer para calcular las soluciones en el campo real de los sistemas:

$$\begin{array}{ll} a) x+y+z=1 & b) x+y+z+w=1 \\ 2x+3y+4z=1 & x+2y+3z+4w=0 \\ x-y-z=0 & x+y+4z+5w=1 \\ & x+y+5z+6w=0. \end{array}$$

18. a) Sea $GL(n, F)$ el conjunto de todos elementos de F_n cuyo determinante es diferente de 0. Pruébese que $GL(n, F)$ es un grupo bajo la multiplicación de matrices.

b) Sea $D(n, F) = \{A \in GL(n, F) \mid \det A = 1\}$. Pruébese que $D(n, F)$ es un subgrupo normal de $GL(n, F)$.

c) Pruébese que $GL(n, F)/D(n, F)$ es isomorfo al grupo de elementos distintos de cero de F bajo la multiplicación.

19. Si K es un campo extensión de F , sea $E(n, K, F) = \{A \in GL(n, K) \mid \det A \in F\}$.

a) Pruébese que $E(n, K, F)$ es un subgrupo normal de $GL(n, K)$.

*b) Determíñese $GL(n, K)/E(n, K, F)$.

*20. Si F es el campo de los números racionales, pruébese que cuando N es un subgrupo normal de $D(2, F)$ entonces o $N = D(2, F)$ o N consiste solamente en matrices escalares.

10. TRANSFORMACIONES HERMITIANAS, UNITARIAS Y NORMALES

En nuestras consideraciones previas acerca de las transformaciones lineales, la naturaleza específica del campo F ha jugado un papel relativamente insignificante. Cuando se hizo sentir fue usualmente respecto a la presencia o ausencia de raíces características. Ahora, por primera vez, restringiremos el campo F —generalmente será el campo de los números complejos, pero a veces será el de los números reales—y haremos un gran uso de las propiedades de los números complejos y reales. *A menos que explícitamente se diga lo contrario, en toda esta sección F representará al campo de los números complejos.*

Haremos también un uso extensivo y constante de las nociones y resultados de la sección 4, capítulo 4, sobre espacios con producto interior. Aconsejamos al lector que revise y asimile por completo tal material antes de seguir más adelante.

Una observación más acerca de los números complejos: hasta ahora hemos evitado usar resultados que no hubieran sido probados en el libro. Ahora, sin embargo, nos vemos forzados a desviarnos de esta norma y a emplear un hecho básico referente al campo de los números complejos, el llamado “*teorema fundamental del álgebra*”, sin que aquí lo demostremos. Nos desagrada sacar tal resultado como quien dice del aire, enunciarlo como un hecho y pasar sin más a hacer uso de él. Desgraciadamente, es esencial para lo que sigue y hacer aquí una digresión para probarlo nos llevaría demasiado lejos. Esperamos que la mayoría de los lectores habrán estudiado ya su demostración en un curso sobre variables complejas.

HECHO 1. *Un polinomio con coeficientes que son números complejos tiene todas sus raíces en el campo complejo.*

El hecho 1 puede reformularse diciendo que los únicos polinomios irreducibles no constantes sobre el campo de los números complejos son los de grado 1.

HECHO 2. *Los únicos polinomios irreducibles no constantes sobre el campo de los números reales son los de grado 1 o grado 2.*

La fórmula para las raíces de una ecuación cuadrática nos permite probar fácilmente la equivalencia de los hechos 1 y 2.

La implicación inmediata, para nosotros, del hecho 1, será que *toda transformación lineal de las que aquí consideraremos tendrá sus raíces características en el campo de los números complejos*.

En lo que sigue, V será un espacio vectorial de dimensión finita con producto interior sobre F , el campo de los números complejos; el producto interior de dos elementos de V se escribirá, como antes se hizo, como (v, w) ,

LEMA 6.26. *Si $T \in A(V)$ es tal que $(vT, v) = 0$ para todo $v \in V$, entonces $T = 0$.*

Prueba. Como $(vT, v) = 0$ para $v \in V$, dados $u, w \in V$, $((u+w)T, u+w) = 0$. Desarrollando esto y haciendo uso de que $(iT, u) = (wT, w) = 0$, obtenemos

$$1) (uT, w) + (wT, u) = 0 \text{ para cualesquiera } u, w \in V.$$

Como la ecuación (1) se verifica para w arbitrario en V , debe aun verificarse si reemplazamos en ella w por iw , donde $i^2 = -1$; pero $(uT, iw) = -i(uT, w)$, mientras que $((iw)T, u) = i(wT, u)$. Sustituyendo estos valores en (1) y cancelando i tenemos

$$2) -(uT, w) + (wT, u) = 0.$$

Sumando (1) y (2) tenemos $(wT, u) = 0$ para cualesquiera u y w de V , de donde, en particular, $(wT, wT) = 0$. Por las propiedades definitorias de un producto interior esto implica que $wT = 0$ para todo $w \in V$, de donde $T = 0$. (*Nota:* si V es un espacio con producto interior sobre el campo real, el lema puede ser falso. Por ejemplo, sea $V = \{(\alpha, \beta) | \alpha, \beta \text{ reales}$, donde el producto interior es el producto punto. Sea T la transformación lineal que manda (α, β) en $(-\beta, \alpha)$. Una simple comprobación nos dice que $(vT, v) = 0$ para cualquier $v \in V$, sin embargo, $T \neq 0$.)

DEFINICIÓN. La transformación lineal $T \in A(V)$ se dice que es *unitaria* si $(uT, vT) = (u, v)$ para cualesquiera $u, v \in V$.

Una transformación unitaria es una que preserva toda la estructura de V , su suma, su multiplicación por escalares y su *producto interior*. Nótese que una transformación unitaria preserva la longitud, puesto que

$$\|v\| = \sqrt{(v, v)} = \sqrt{(vT, vT)} = \|vT\|.$$

¿Es lo recíproco cierto? La contestación nos la da el

LEMA 6.27. *Si $(vT, vT) = (v, v)$ para toda $v \in V$, entonces T es unitaria.*

Prueba. La prueba tiene el mismo estilo que la del lema 6.26. Sean $u, v \in V$; por hipótesis $((u+v)T, (u+v)T) = (u+v, u+v)$. Desarrollando y

simplificando, tenemos

$$1) (uT, vT) + (vT, uT) = (u, v) + (v, u),$$

para cualesquiera $u, v \in V$. Reemplazando en (1) v por iv y calculando esto nos da

$$2) -(uT, vT) + (vT, uT) = -(u, v) + (v, u).$$

Sumando (1) y (2) obtenemos $(uT, vT) = (u, v)$ para cualesquiera $u, v \in V$, de donde T es unitario.

Caracterizamos la propiedad de ser unitario en términos de la acción sobre una base de V .

TEOREMA 6.Y. *La transformación lineal T sobre V es unitaria si y sólo si lleva una base ortonormal de V en una base ortonormal de V .*

Prueba. Supongamos que $\{v_1, \dots, v_n\}$ es una base ortonormal de V ; por tanto, $(v_i, v_j) = 0$ si $i \neq j$, mientras que $(v_i, v_i) = 1$. Queremos demostrar que si T es unitario, entonces $\{v_1 T, \dots, v_n T\}$ es también una base ortonormal de V . Pero $(v_i T, v_j T) = (v_i, v_j) = 0$ para $i \neq j$ y $(v_i T, v_i T) = (v_i, v_i) = 1$, de donde ciertamente $\{v_1 T, \dots, v_n T\}$ es una base ortonormal de V .

Por otra parte, si $T \in A(V)$ es tal que tanto $\{v_1, \dots, v_n\}$ como $\{v_1 T, \dots, v_n T\}$ son bases ortonormales de V , para $u, w \in V$ tenemos entonces

$$u = \sum_{i=1}^n \alpha_i v_i, \quad w = \sum_{i=1}^n \beta_i v_i,$$

de donde por la ortonormalidad de las v_i , $(u, w) = \sum_{i=1}^n \alpha_i \beta_i$. Pero $uT = \sum_{i=1}^n \alpha_i v_i T$ y $wT = \sum_{i=1}^n \beta_i v_i T$, de donde por la ortonormalidad de las $v_i T$, $(uT, wT) = \sum_{i=1}^n \alpha_i \beta_i = (u, w)$, lo que prueba que T es unitaria.

El teorema 6.y nos dice que un cambio de base de una base ortonormal a otra base también ortonormal es precisamente lo que produce una transformación lineal unitaria.

LEMÁ 6.28. *Si $T \in A(V)$, entonces dada cualquier $v \in V$, existe un elemento $w \in V$, que depende de v y de T , tal que $(uT, v) = (u, w)$ para toda $u \in V$. Este elemento queda únicamente determinado por v y T .*

Prueba. Para probar el lema es suficiente exhibir un $w \in V$ que trabaje para todos los elementos de una base de V . Sea $\{u_1, \dots, u_n\}$ una base

ortonormal de V ; definimos $= \sum_{i=1}^n \overline{(u_i T, v)} u_i$. Un fácil cálculo muestra que $(u_i, w) = (u_i T, v)$, de donde el elemento w tiene la propiedad deseada. Que w es único puede verse como sigue: Supongamos que $(u T, v) = (u, w_1) = (u, w_2)$; entonces $(u, w_1 - w_2) = 0$ para toda $u \in V$, lo que obliga al poner $u = w_1 - w_2$ a que $w_1 = w_2$.

El lema 6.28 nos permite dar la siguiente

DEFINICIÓN. Si $T \in A(V)$, entonces el *adjunto hermitiano* de T , al que representaremos por T^* , se define por $(u T, v) = (u, v T^*)$ para cualesquiera $u, v \in V$.

Dado $v \in V$ acabamos de obtener una expresión explícita para $v T^*$ (como w) y prodríamos usar esta expresión para probar las distintas propiedades que deseamos tenga T^* . Pero preferimos hacerlo de modo que no tengamos que depender de una base determinada.

LEMA 6.29. Si $T \in A(V)$, entonces $T^* \in A(V)$.

Además:

- 1) $(T^*)^* = T$,
- 2) $(S + T)^* = S^* + T^*$,
- 3) $(\lambda S)^* = \bar{\lambda} S^*$,
- 4) $(ST)^* = T^* S^*$,

para $S, T \in A(V)$ cualesquiera y todo λ .

Prueba. Debemos primero probar que T^* es una transformación lineal sobre V . Si u, v, w están en V , entonces $(u, (v+w) T^*) = (u T, v+w) = (u T, v) + (u T, w) = (u, v T^*) + (u, w T^*) = (u, v T^* + w T^*)$, en consecuencia de lo cual $(v+w) T^* = v T^* + w T^*$. Análogamente, para $\lambda \in F$, $(u, (\lambda v) T^*) = (u T, \lambda v) = \bar{\lambda}(u T, v) = \bar{\lambda}(u, v T^*) = (u, \bar{\lambda}(v T^*))$, de donde $(\lambda v) T^* = \bar{\lambda}(v T^*)$. Con lo que hemos probado que T^* es una transformación lineal sobre V .

Para ver que $(T^*)^* = T$, notemos que $(u, v(T^*)^*) = (u T^*, v) = \overline{(v, u T^*)} = \overline{(v T, u)} = (u, v T)$ para todo $u, v \in V$, de donde $v(T^*)^* = v T$ lo que implica que $(T^*)^* = T$. Dejamos las pruebas de $(S+T)^* = S^* + T^*$ y de $(\lambda T)^* = \bar{\lambda} T^*$ para el lector. Finalmente, $(u, v(ST)^*) = (u ST, v) = (u S, v T^*) = (u, v T^* S^*)$ para $u, v \in V$ cualesquiera; esto implica $v(ST)^* = v T^* S^*$ para cualquier $v \in V$ lo que nos dice que $(ST)^* = T^* S^*$.

Como consecuencia del lema tenemos que la adjunta hermitiana define una adjunta, en el sentido de la sección 8, sobre $A(V)$.

La adjunta hermitiana nos permite dar una descripción alternativa para las transformaciones unitarias en términos de la relación de T y T^* .

LEMA 6.30. $T \in A(V)$ es unitaria si y sólo si $TT^* = 1$.

Prueba. Si T es unitaria, entonces para todo $u, v \in V$, $(u, vTT^*) = (uT, vT) = (u, v)$, de donde $TT^* = I$. Por otra parte, si $TT^* = I$, entonces $(u, v) = (u, vTT^*) = (uT, vT)$, lo que implica que T es unitaria.

TEOREMA 6.z. *Si $\{v_1, \dots, v_n\}$ es una base ortonormal de V y si la matriz de $T \in A(V)$ en esta base es (α_{ij}) , entonces la matriz de T^* en esta base es (β_{ij}) , donde $\beta_{ij} = \bar{\alpha}_{ji}$.*

Prueba. Como las matrices de T y T^* en esta base son, respectivamente, (α_{ij}) y (β_{ij}) , entonces $v_i T = \sum_{j=1}^n \alpha_{ij} v_j$ y $v_i T^* = \sum_{j=1}^n \beta_{ij} v_j$. Ahora bien, $\beta_{ij} = (v_i T^*, v_j) = (v_i, v_j T) = (v_i, \sum_{k=1}^n \alpha_{jk} v_k) = \bar{\alpha}_{ji}$ por la ortonormalidad de las v_i . Lo que prueba el teorema.

Este teorema nos interesa muy en particular a la luz de lo que hicimos anteriormente en la sección 8. Pues la adjunta hermitiana abstracta definida sobre el espacio con producto interior V , cuando es trasladado a matrices en una base ortonormal de V , no se hace otra cosa que la adjunta hermitiana concreta explícita que definimos para las matrices.

Usando la representación matricial en una base ortonormal, afirmamos que $T \in A(V)$ es unitaria si y sólo si siempre que (α_{ij}) es la matriz de T en esta base ortonormal, entonces $\sum_{i=1}^n \alpha_{ij} \bar{\alpha}_{ik} = 0$ para $j \neq k$, mientras que $\sum_{i=1}^n |\alpha_{ij}|^2 = 1$. En términos de productos punto sobre espacios vectoriales complejos, esto nos dice que los renglones de la matriz de T forman un conjunto ortonormal de vectores en $F^{(n)}$ bajo el producto punto.

DEFINICIÓN. $T \in A(V)$ se llama *autoadjunta* o *hermitiana* si $T^* = T$.

Si $T^* = -T$, entonces llamamos a T *antihermitiana*. Dada cualquier $S \in A(V)$,

$$S = \frac{S+S^*}{2} + i \left(\frac{S-S^*}{2i} \right),$$

y como $\frac{S+S^*}{2}$ y $\frac{S-S^*}{2i}$ son hermitianas, $S = A + iB$, donde tanto A como B son hermitianas.

En la sección 8, usando el cálculo de matrices, probamos que cualquier raíz característica compleja de una matriz hermitiana es real; a la luz del hecho 1, esto puede cambiarse para que diga: Toda raíz característica de una matriz hermitiana es real. Volvemos ahora a probar esto desde el punto de vista, más uniforme, de un espacio con producto interior.

TEOREMA 6.21. Si $T \in A(V)$ es hermitiana, entonces todas sus raíces características son reales.

Prueba. Sea λ una raíz característica de T ; hay pues una $v \neq 0$ en V tal que $vT = \lambda v$. Calculamos: $\lambda(v, v) = (\lambda v, v) = (vT, v) = (v, vT^*) = (v, vT) = (v, \lambda v) = \lambda(v, v)$; como $(v, v) \neq 0$, como tenemos $\lambda = \bar{\lambda}$, λ es real.

Deseamos describir formas canónicas para transformaciones lineales unitarias, hermitianas e incluso de tipos más generales que serán también más sencillas que las formas de Jordan. Es por esto por lo que aparecen los siguientes lemas que, aunque de interés independiente, son en su mayor parte de naturaleza más bien técnica.

LEMA 6.31. Si $S \in A(V)$ y si $vSS^* = 0$, entonces $vS = 0$.

Prueba. Consideremos (vSS^*, v) ; como $vSS^* = 0$, $0 = (vSS^*, v) = (vS, v(S^*)^*) = (vS, vS)$ según el lema 6.29. En un espacio con producto interior esto implica que $vS = 0$.

COROLARIO. Si T es hermitiana y $vT^k = 0$ para $k \geq 1$, entonces $vT = 0$.

Prueba. Mostramos que si $vT^{2^m} = 0$, entonces $vT = 0$; pues si $S = T^{2^{m-1}}$, entonces $S^* = S$ y $SS^* = T^{2^m}$, de donde $(vSS^*, v) = 0$ implica que $0 = vS = vT^{2^{m-1}}$. Continuando hacia abajo en esta forma, obtenemos $vT = 0$. Si $vT^k = 0$, entonces $vT^{2^m} = 0$ para $2^m > k$, de donde $vT = 0$.

Introducimos una clase de transformaciones lineales que contiene como casos especiales las transformaciones unitarias, hermitianas y antihermitianas.

DEFINICIÓN. $T \in A(V)$ se dice que es *normal* si $TT^* = T^*T$.

En lugar de probar los teoremas que siguen para transformaciones unitarias y para transformaciones hermitianas separadamente, lo que haremos será probarlos para transformaciones normales y derivar, como corolarios, los resultados deseados para las unitarias y hermitianas.

LEMA 6.32. Si N es una transformación lineal normal y si $vN = 0$ para $v \in V$, entonces $vN^* = 0$.

Prueba. Consideremos (vN^*, vN^*) ; por definición, $(vN^*, vN^*) = (vN^*N, v) = (vNN^*, v)$, ya que $NN^* = N^*N$. Pero $vN = 0$, de donde ciertamente $vNN^* = 0$. De esta forma obtenemos que $(vN^*, vN^*) = 0$, de donde forzosamente ha de tenerse $vN^* = 0$.

COROLARIO 1. Si λ es una raíz característica de la transformación normal N y si $vN = \lambda v$, entonces $vN^* = \bar{\lambda}v$.

Prueba. Como N es normal, $NN^* = N^*N$, de donde tenemos $(N - \lambda)(N - \bar{\lambda})^* = (N - \lambda)(N^* - \bar{\lambda}) = NN^* - \lambda N^* - \bar{\lambda}N + \lambda\bar{\lambda} = N^*N - \lambda N^* - \bar{\lambda}N + \lambda\bar{\lambda} = (N^* - \bar{\lambda})(N - \lambda) = (N - \lambda)^*(N - \lambda)$, es decir, $N - \lambda$ es normal. Como $v(N - \lambda) = 0$, por la normalidad de $N - \lambda$ se tiene del lema: $v(N - \lambda)^* = 0$, de donde $vN^* = \bar{\lambda}v$.

El corolario enuncia el interesante hecho de que si λ es una raíz característica de la transformación normal N , no solamente es $\bar{\lambda}$ una raíz característica de N^* , sino que cualquier vector característico de N perteneciente a λ es un vector característico de N^* perteneciente a $\bar{\lambda}$ y viceversa.

COROLARIO 2. Si T es unitaria y si λ es una raíz característica de T , entonces $|\lambda| = 1$.

Prueba. Como T es unitaria, es normal. Sea λ una raíz característica de T y supongamos que $vT = \lambda v$ con $v \neq 0$ en V . Por el corolario 1, $vT^* = \bar{\lambda}v$, luego $v = vTT^* = \lambda vT^* = \lambda\bar{\lambda}v$, ya que $TT^* = I$. Luego tenemos $\lambda\bar{\lambda} = 1$, lo que nos dice $|\lambda| = 1$.

Hacemos una pausa para ver adonde vamos. Nuestro objetivo inmediato es probar que una transformación normal N puede llevarse a la forma diagonal por una unitaria. Si $\lambda_1, \dots, \lambda_k$ son raíces características distintas de V , usando el teorema 6.n podemos descomponer V como $V = V_1 \oplus \dots \oplus V_k$, donde para $v_i \in V_i$, $v_i(N - \lambda_i)^{n_i} = 0$. De acuerdo con esto, necesitamos estudiar dos cosas, a saber: la relación entre vectores que se encuentran en distintos V_i , y la naturaleza característica de cada V_i . Cuando estas dos cosas hayan sido estudiadas, seremos capaces de reunirlas para probar el teorema deseado.

LEMA 6.33 Si N es normal y $vN^k = 0$, entonces $vN = 0$.

Prueba. Sea $S = NN^*$; S es hermitiana, y según la normalidad de N , $vS^k = v(NN^*)^k = vN^k(N^*)^k = 0$. De acuerdo con el corolario al lema 6.31, deducimos que $vS = 0$, es decir, $vNN^* = 0$. Aplicando el lema 6.31 se tiene que $vN = 0$.

COROLARIO. Si N es normal y si para $\lambda \in F$, $v(N - \lambda)^k = 0$, entonces $vN = \lambda v$.

Prueba. De la normalidad de N se sigue que $N - \lambda$ es normal, de donde, al aplicar el lema que acabamos de probar a $N - \lambda$ obtenemos el corolario.

Siguiendo con la discusión que precedía al último lema, este corolario demuestra que todo vector en V_i es un vector característico de N perteneciente a la raíz característica λ_i . Hemos determinado la naturaleza de V_i ; ahora procederemos a investigar la interrelación entre dos distintas V_i .

LEMA 6.34. *Sea N una transformación normal y supongamos que λ y μ son dos raíces características distintas de N . Si v, w son de V y tales que $vN = \lambda v$, $wN = \mu w$, entonces $(v, w) = 0$.*

Prueba. Calculamos (vN, w) de dos formas diferentes. Como una consecuencia de que $vN = \lambda v$, $(vN, w) = (\lambda v, w) = \lambda(v, w)$. Como $wN = \mu w$, usando el lema 6.32 obtenemos que $wN^* = \bar{\mu}w$, de donde $(vN, w) = (v, wN^*) = (v, \bar{\mu}w) = \mu(v, w)$. La comparación de los dos cálculos, nos da $\lambda(v, w) = \mu(v, w)$, y como $\lambda \neq \mu$, de ello resulta que $(v, w) = 0$.

Todo el trabajo preliminar ya ha sido hecho para que podamos probar este básico y bello teorema:

TEOREMA 6.z₂. *Si N es una transformación lineal normal sobre V , entonces existe una base ortonormal, consistente en vectores característicos de N , en la cual la matriz de N es diagonal. Equivalentemente, si N es una matriz normal, existe una matriz unitaria U tal que UNU^{-1} ($= UNU^*$) es diagonal.*

Prueba. Completamos el esquema informal que hemos hecho de la prueba antes de la demostración del lema 6.33.

Sea N normal y sea $\lambda_1, \dots, \lambda_k$ las distintas raíces características de N . Por el corolario al teorema 6.n, podemos descomponer V en la forma $V = V_1 \oplus \dots \oplus V_k$ donde toda $v_i \in V_i$ es aniquilada por $(N - \lambda_i)^{n_i}$. Por el corolario al lema 6.33, V_i consiste solamente en vectores característicos de N pertenecientes a la raíz característica λ_i . El producto interior de V induce un producto interior sobre V_i ; por el teorema 4.h, podemos encontrar una base de V_i ortonormal respecto a su producto interior.

Por el lema 6.34, los elementos pertenecientes a distintos V_i son ortogonales. Luego la unión de las bases ortonormales de las V_i nos proporciona una base ortonormal de V . Esta base consiste en los vectores característicos de N , de donde en esta base la matriz de N es diagonal.

No probamos el equivalente matricial dejándolo como un problema; solamente señalamos que dos hechos son necesarios:

- 1) Una transformación unitaria (teorema 6.y) cambia una base ortonormal por una base también ortonormal.
- 2) En un cambio de base, la matriz de una transformación lineal se cambia por conjugación por la matriz del cambio de base (teorema 6.h).

Los dos corolarios que siguen son casos muy particulares del teorema 6.z₂, pero como cada uno de ellos es tan importante por sí mismo, los enunciamos como corolarios para subrayarlos.

COROLARIO 1. *Si T es una transformación unitaria, entonces hay una base ortonormal en la que la matriz de T es diagonal; equivalentemente, si T es una matriz unitaria, entonces hay una matriz unitaria U tal que UTU^{-1} ($= UTU^*$) es diagonal.*

COROLARIO 2. *Si T es una transformación lineal hermitiana, entonces existe una base ortonormal en la que la matriz de T es diagonal; equivalentemente, si T es una matriz hermitiana, entonces existe una matriz unitaria U tal que UTU^{-1} ($= UTU^*$) es diagonal.*

El teorema probado es el resultado básico para las transformaciones normales, pues las caracteriza en forma neta como precisamente aquellas transformaciones que pueden llevarse a la forma diagonal por unitarias. También muestra que la distinción entre transformaciones normales, hermitianas y unitarias es solamente una distinción causada por la naturaleza de sus raíces características. Precisamos esto en el

LEMA 6.35. *La transformación normal N es:*

- 1) *Hermitiana si y solo si sus raíces características son reales;*
- 2) *Unitaria si y solo si sus raíces características son todas de valor absoluto 1.*

Prueba. Argumentamos usando matrices. Si N es hermitiana, entonces es normal y todas sus raíces características son reales. Si N es normal y tiene solamente raíces características reales, entonces para alguna matriz unitaria U , $UNU^{-1} = UNU^* = D$ donde D es una matriz diagonal con elementos reales en la diagonal. Así pues, $D^* = D$; como $D^* = (UNU^*)^* = UNU^*$, la relación $D^* = D$, implica $UNU^* = UNU^*$, y como U es invertible obtenemos $N^* = N$. Luego N es hermitiana.

Dejamos al lector la prueba de la parte referente a las transformaciones unitarias.

Si A es una transformación lineal cualquiera sobre V , entonces $\text{tr}(AA^*)$ puede calcularse usando la representación matricial de A en cualquier base de V . Escogemos una base ortonormal de V ; en esta base si la matriz de A es (α_{ij}) entonces la de A^* es (β_{ij}) donde $\beta_{ij} = \bar{\alpha}_{ji}$. Un cálculo simple nos muestra entonces que $\text{tr}(AA^*) = \sum_{i,j} |\alpha_{ij}|^2$ y esto es cero si y sólo si todo $\alpha_{ij} = 0$, es decir, si y sólo si $A = 0$. En una palabra, $\text{tr}(AA^*) = 0$ si y sólo

si $A = 0$. Este es un criterio útil para mostrar que una transformación lineal dada es 0. Ilustramos esto en el siguiente

LEMA 6.36. *Si N es normal y $AN = NA$, entonces $AN^* = N^*A$.*

Prueba. Queremos demostrar que $X = AN^* - N^*A$ es 0; lo que haremos es probar que $\text{tr } XX^* = 0$, y deducir de esto que $X = 0$.

Como N commuta con A y con N^* , debe commutar con $AN^* - N^*A$, así pues $XX^* = (AN^* - N^*A)(NA^* - A^*N) = (AN^* - N^*A)NA^* - (AN^* - N^*A)A^*N = N\{(AN^* - N^*A)A^*\} - \{(AN^* - N^*A)A^*\}N$. Como XX^* es de la forma $NB - BN$, la traza de XX^* es 0. Luego $X = 0$, y $AN^* = N^*A$.

Acabamos de ver que N^* commuta con todas las transformaciones lineales que commutan con N , cuando N es normal; esto es suficiente para que forzosamente N^* sea una expresión polinomial en N . Pero esto puede demostrarse directamente como una consecuencia del teorema 6.z₂ (véase el problema 14).

La transformación lineal T es hermitiana si y sólo si (vT, v) es real para todo $v \in V$ (véase el problema 19). De especial interés son aquellas transformaciones lineales hermitianas para las que $(vT, v) \geq 0$ para todo $v \in V$. Las llamamos transformaciones lineales *no negativas* y denotamos el hecho de que una transformación lineal sea no negativa escribiendo $T \geq 0$. Si $T \geq 0$ y además $(vT, v) > 0$ para $v \neq 0$ entonces llamamos a T *positiva* (*o positivamente definida*) y escribimos $T > 0$. Queremos distinguir a estas transformaciones lineales por sus raíces características.

LEMA 6.37. *La transformación lineal hermitiana T es no negativa (positiva) si y sólo si todas sus raíces características son no negativas (positivas).*

Prueba. Supongamos que $T \geq 0$; si λ es una raíz característica de T , entonces $vT = \lambda v$ para algún $v \neq 0$. Luego $0 \leq (vT, v) = (\lambda v, v) = \lambda(v, v)$; como $(v, v) > 0$, se deduce que $\lambda \geq 0$.

Recíprocamente, si T es hermitiana con raíces características no negativas, entonces podemos encontrar una base ortonormal $\{v_1, \dots, v_n\}$ consistente en vectores características de T . Para cada v_i , $v_i T = \lambda_i v_i$, donde $\lambda_i \geq 0$. Dado $v \in V$, $v = \sum \alpha_i v_i$ de donde $vT = \sum \alpha_i v_i T = \sum \lambda_i \alpha_i v_i$. Pero entonces $(vT, v) = (\sum \lambda_i \alpha_i v_i, \sum \alpha_i v_i) = \sum \lambda_i \alpha_i \alpha_i$ por la ortonormalidad de las v_i . Como $\lambda_i \geq 0$ y $\alpha_i \alpha_i \geq 0$, se tiene $(vT, v) \geq 0$, de donde $T \geq 0$.

Los resultados correspondientes para el caso “positivo” se dejan como ejercicio.

LEMA 6.38. *$T \geq 0$ si y sólo si $T = AA^*$ para alguna A .*

Prueba. Demostramos primero que $AA^* \geq 0$. Dado $v \in V$, $(vAA^*, v) = (vA, vA) = 0$, de donde $AA^* \geq 0$.

Por otra parte, si $T \geq 0$ podemos encontrar una matriz unitaria U tal que

$$UTU^* = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

donde cada λ_i es una raíz característica de T , luego toda $\lambda_i \geq 0$. Sea

$$S = \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix};$$

como cada $\lambda_i \geq 0$, cada $\sqrt{\lambda_i}$ es real, luego S hermitiana. Por tanto U^*SU es hermitiana; pero

$$(U^*SU)^2 = U^*S^2U = U^* \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U = T.$$

Hemos representado a T en la forma AA^* , donde $A = U^*SU$.

Nótese que realmente hemos probado un poco más; a saber, si al construir S hubiéramos escogido la raíz no negativa $\sqrt{\lambda_i}$ para cada λ_i , entonces S , y U^*SU , habría sido no negativa. Luego $T \geq 0$ es el cuadrado de una transformación lineal no negativa; es decir, toda $T \geq 0$ tiene una raíz cuadrada no negativa. Esta raíz cuadrada no negativa puede demostrarse que es única (véase el problema 24).

Cerramos esta sección con una discusión sobre las matrices unitarias y hermitianas *sobre el campo real*. En este caso, las matrices unitarias se llaman *ortogonales*, y satisfacen $QQ' = I$. Las hermitianas son en este caso exactamente simétricas.

Afirmamos que una matriz real simétrica puede llevarse a la forma diagonal por una matriz ortogonal. Sea A una matriz real simétrica. Podemos considerar a A actuando sobre un espacio real V con producto interior. Considerada como una matriz compleja, A es hermitiana y por tanto todas sus raíces características son reales. Si estas son $\lambda_1, \dots, \lambda_k$ entonces V puede descomponerse en $V = V_1 \oplus \dots \oplus V_k$, donde $v_i(A - \lambda_i)^n = 0$ para $v_i \in V_i$. Como en la prueba del teorema 6.33 esto trae como consecuencia obligada $v_i A = \lambda_i v_i$. Usando exactamente la misma prueba que la que usamos en el lema 6.34, mostramos que para $v_i \in V_i$, $v_j \in V_j$ con $i \neq j$,

$(v_i, v_j) = 0$. Podemos, pues, encontrar una base ortonormal de V todos cuyos elementos sean vectores característicos de A . El cambio de bases, de la base ortonormal $\{(1, 0, \dots, 0), (0, 1, 0, \dots, 0), \dots, (0, \dots, 0, 1)\}$ a esta nueva base se efectúa mediante una matriz unitaria real, es decir, por una ortogonal. Así pues, A puede llevarse a forma diagonal por una matriz ortogonal, probando nuestra afirmación.

Determinar formas canónicas para las matrices ortogonales reales sobre el campo real es un poco más complicado, tanto en su respuesta como en su ejecución. Pasamos ahora a estudiar este problema; pero antes vamos a hacer una observación general acerca de todas las transformaciones unitarias.

Si W es un subespacio de V invariante bajo la transformación unitaria T , ¿es cierto que W' , el complemento ortogonal de W , es también invariante bajo T ? Sea $w \in W$ y $x \in W'$; tenemos entonces: $(wT, xT) = (w, x) = 0$; como W es invariante bajo T y T es regular, $WT = W$, de donde xT , para $x \in W'$, es ortogonal para todo W . Luego es cierto que $(W')T \subset W'$. Recuérdese que $V = W \oplus W'$.

Sea Q una matriz ortogonal real; entonces $T = Q + Q^{-1} = Q + Q'$ es simétrica, de donde tiene raíces características reales. Si estas son $\lambda_1, \dots, \lambda_k$, entonces V puede descomponerse en $V = V_1 \oplus \dots \oplus V_k$, donde $v_i \in V_i$ implica $v_i T = \lambda_i v_i$. Las V_i son mutuamente ortogonales. Afirmamos que cada V_i es invariante bajo Q (pruébese). Luego, para discutir la acción de Q sobre V , es suficiente describirla sobre cada V_i .

Sobre V_i , como $\lambda_i v_i = v_i T = v_i(Q + Q^{-1})$, multiplicando por Q tenemos $v_i(Q^2 - \lambda_i Q + I) = 0$. Se presentan dos casos particulares, a saber: $\lambda_i = 2$ y $\lambda_i = -2$ (que pueden, desde luego, no ocurrir), pues entonces $v_i(Q + I)^2 = 0$, lo que nos lleva a $v_i(Q \pm I) = 0$. Sobre estos espacios, Q actúa como I o como $-I$.

Si $\lambda_i \neq 2, -2$, entonces Q no tiene ningún vector característico sobre V_i , de donde para $v \neq 0 \in V_i$, v, vQ son linealmente independientes. El subespacio que generan, W , es invariante bajo Q , ya que $vQ^2 = \lambda_i vQ - v$. Ahora bien, $V_i = W \oplus W'$ con W' invariante bajo Q . Luego podemos presentar a V_i como la suma directa de dos subespacios bidimensionales mutuamente ortogonales invariantes bajo Q . Para encontrar formas canónicas de Q sobre V_i (de donde sobre V), solamente debemos resolver el problema para matrices ortogonales reales 2×2 .

Sea Q una matriz ortogonal real 2×2 que satisface $Q^2 - \lambda Q + I = 0$: supongamos que $Q = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$. La ortogonalidad de Q implica:

- 1) $\alpha^2 + \beta^2 = 1$,
- 2) $\gamma^2 + \delta^2 = 1$,
- 3) $\alpha\gamma + \beta\delta = 0$;

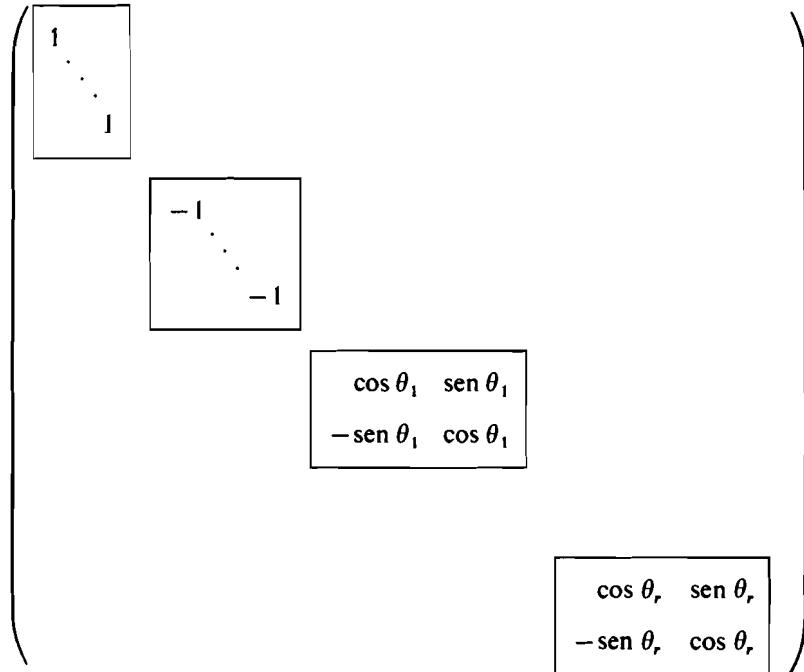
como $Q^2 - \lambda Q + I = 0$, el determinante de Q es 1, de donde

$$4) \alpha\delta - \beta\gamma = 1.$$

Afirmamos que las ecuaciones 1, ..., 4, implican que $\alpha = \delta$, $\beta = -\gamma$. Como $\alpha^2 + \beta^2 = 1$, $|\alpha| \leq 1$, de donde podemos escribir $\alpha = \cos \theta$ para algún ángulo real θ ; en estos términos $\beta = \sin \theta$. Por tanto, la matriz Q toma la forma

$$\begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}.$$

Todos los espacios usados en todas nuestras decomposiciones eran mutuamente ortogonales, luego eligiendo bases ortogonales en cada uno de ellos obtenemos una base ortonormal de V . En esta base la matriz de Q es



Como hemos ido de una base ortonormal a otra también ortonormal, y como esto se ha conseguido por una matriz ortogonal, dada una matriz ortogonal real Q podemos encontrar una matriz *ortogonal* T tal que TQT^{-1} ($= TQT^*$) es de la forma que acabamos de describir.

Problemas

1. Determinese cuáles de las siguientes matrices son unitarias, cuáles hermitianas, cuáles normales.

$$a) \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \quad b) \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad c) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

$$d) \begin{pmatrix} 1 & 2-i \\ 2-i & i \end{pmatrix}, \quad e) \begin{pmatrix} 3 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} - \frac{i}{\sqrt{2}} & 0 \\ 0 & \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

2. Para aquellas matrices del problema 1 que sean normales, encuéntrense sus raíces características y llévense a la forma diagonal por una matriz unitaria.

3. Si T es unitaria, pruébese usando tan solo la definición $(vT, uT) = (v, u)$ que T es no singular.

4. Si Q es una matriz ortogonal real, pruébese que $\det Q = \pm 1$.

5. Si Q es una matriz real simétrica que satisface $Q^k = I$ para $k \geq 1$, pruébese que $Q^2 = I$.

6. Complétense la prueba del lema 6.29 mostrando que $(S+T)^* = S^* + T^* y (\lambda T)^* = \bar{\lambda} T^*$.

7. Pruébense las propiedades de $*$ en el lema 6.29 haciendo uso de la forma explícita de $w = vT^*$ dada en la prueba del lema 6.28.

8. Si T es antihermitiana, pruébese que todas sus raíces características son imaginarias puras.

9. Si T es una matriz real antisimétrica $n \times n$, pruébese que si n es impar entonces $\det T = 0$.

10. Por un cálculo matricial directo, pruébese que una matriz real simétrica 2×2 puede ser puesta en forma diagonal por una ortogonal.

11. Complétense la prueba delineada para la parte de equivalencia de matrices del teorema 6.22.

12. Pruébese que una transformación normal es unitaria si y sólo si las raíces características son todas de valor absoluto igual a 1.

13. Si N_1, \dots, N_k es un número finito de transformaciones normales que comutan, pruébese que existe una transformación unitaria T tal que todas las $TN_i T^{-1}$ son diagonales.

14. Si N es normal, pruébese que $N^* = p(N)$ para algún polinomio $p(x)$.

15. Si N es normal y si $AN = 0$, pruébese que $AN^* = 0$.

16. Pruébese que A es normal si y sólo si A conmuta con AA^* .

17. Si N es normal pruébese que $N = \sum \lambda_i E_i$ donde $E_i^2 = E_i$, $E_i^* = E_i$, y las λ_i son las raíces características de N . (A ésta se le llama la *resoluciónpectral de N*.)

18. Si N es una transformación normal sobre V y si $f(x)$ y $g(x)$ son dos polinomios primos relativos con coeficientes reales, pruébese que si $vf(N) = 0$ y $wg(N) = 0$, para v, w en V , entonces $(v, w) = 0$.

19. Pruébese que una transformación lineal T sobre V es hermitiana si y sólo si (rT, r) es real para todo $r \in V$.

20. Pruébese que $T > 0$ si y sólo si T es hermitiana y tiene todas sus raíces características positivas.

21. Si $A \geq 0$ y $B \geq 0$ y $AB = BA$, pruébese que $AB \geq 0$.

22. Pruébese que si $A \geq 0$, entonces A tiene una raíz cuadrada no negativa única.

23. Si $A \geq 0$ y $(rA, v) = 0$, pruébese que $vA = 0$.

24. a) Si $A \geq 0$ y A^2 conmuta con la transformación hermitiana B , entonces A conmuta con B .

b) Pruébese la parte (a) sin exigir que B sea hermitiano.

25. Sea $A = (\alpha_{ij})$ una matriz $n \times n$ real simétrica.

Sea

$$A_s = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1s} \\ \vdots & & \vdots \\ \alpha_{s1} & \cdots & \alpha_{ss} \end{pmatrix}.$$

a) Si $A > 0$, pruébese que $A_s > 0$ para $s = 1, 2, \dots, n$.

b) Si $A > 0$, pruébese que $\det A_s > 0$ para $s = 1, 2, \dots, n$.

c) Si $\det A_s > 0$ para $s = 1, 2, \dots, n$, pruébese que $A > 0$.

d) Si $A \geq 0$, pruébese que $A_s \geq 0$ para $s = 1, 2, \dots, n$.

- e) Si $A \geq 0$, pruébese que $\det A_s \geq 0$ para $s = 1, 2, \dots, n$.
f) Proporcionese un ejemplo de una A tal que $\det A_s \geq 0$ para toda $s = 1, 2, \dots, n$ y, sin embargo, A no sea no negativo.

26. Pruébese que cualquier matriz compleja puede ser llevada a la forma triangular por una matriz unitaria.

11. FORMAS CUADRÁTICAS REALES

Cerramos el capítulo con una breve discusión sobre formas cuadráticas sobre el campo de los números reales.

Sea V un espacio real con producto interior y supongamos que A es una transformación lineal (real) simétrica sobre V . La función valuada en el campo real $Q(v)$ definida sobre V por $Q(v) = (vA, v)$ se llama la *forma cuadrática asociada con A* .

Si consideramos, cómo podemos hacer sin pérdida de generalidad, que A es una matriz simétrica real $n \times n$, (α_{ij}) actuando sobre $F^{(n)}$ y que el producto interior para $(\delta_1, \dots, \delta_n)$ y $(\gamma_1, \dots, \gamma_n)$ en $F^{(n)}$ es el número real $\delta_1\gamma_1 + \delta_2\gamma_2 + \dots + \delta_n\gamma_n$, para un vector arbitrario $v = (x_1, \dots, x_n)$ en $F^{(n)}$, un simple cálculo muestra que $Q(v) = (vA, v) = \alpha_{11}x_1^2 + \dots + \alpha_{nn}x_n^2 + 2 \sum_{i < j} \alpha_{ij}x_i x_j$.

Por otra parte, dada una función cuadrática cualquiera en n variables $\gamma_{11}x_1^2 + \dots + \gamma_{nn}x_n^2 + 2 \sum_{i < j} \gamma_{ij}x_i x_j$, con coeficientes reales γ_{ij} , es claro que podemos realizarla como la forma cuadrática asociada con la matriz real simétrica $C = (\gamma_{ij})$.

En el espacio euclíadiano n -dimensional una función cuadrática sirve para definir las superficies cuádricas. Por ejemplo, en el plano real la forma $\alpha x^2 + \beta xy + \gamma y^2$ da lugar a una sección cónica (posiblemente con su eje mayor inclinado). No es ilógico pensar que las propiedades geométricas de esta sección cónica deben estar ligadas íntimamente con la matriz simétrica

$$\begin{pmatrix} \alpha & \beta/2 \\ \beta/2 & \gamma \end{pmatrix},$$

con la que su forma cuadrática está asociada.

Recordemos que en geometría analítica elemental se prueba que por una rotación de ejes adecuada la ecuación $\alpha x^2 + \beta xy + \gamma y^2$ puede, en el nuevo sistema de coordenadas, tomar la forma $\alpha_1(x')^2 + \gamma_1(y')^2$. Recordemos que $\alpha_1 + \gamma_1 = \alpha + \gamma$ y $\alpha\gamma - \beta^2/4 = \alpha_1\gamma_1$. Luego α_1 y γ_1 son las raíces características de la matriz

$$\begin{pmatrix} \alpha & \beta/2 \\ \beta/2 & \gamma \end{pmatrix};$$

la rotación de ejes es tan sólo un cambio de bases por una transformación ortogonal, y lo que hicimos en la geometría fue simplemente llevar la matriz simétrica a su forma diagonal por una matriz ortogonal. La naturaleza de $\alpha x^2 + \beta xy + \gamma y^2$ como cónica estaba básicamente determinada por la magnitud y signo de sus raíces características α_1 y γ_1 .

Una discusión análoga puede llevarse a efecto para clasificar las superficies cuádricas en el espacio tridimensional y, ciertamente, para superficies cuádricas en espacios de dimensión n . Lo que esencialmente determina la naturaleza geométrica de la superficie cuádrica asociada con $\alpha_{11}x_1^2 + \dots + \alpha_{nn}x_n^2 + 2 \sum_{i < j} \alpha_{ij}x_i x_j$ es la magnitud y signo de las raíces características de la matriz (α_{ij}) . Si no estuviésemos interesados en el achatamiento relativo de la superficie cuádrica (por ejemplo, si consideramos una elipse como una circunferencia aplastada), entonces podríamos ignorar la magnitud de las raíces características distintas de cero y el factor determinante de la forma de la superficie cuádrica sería el número de raíces características 0 y el número de positivas (y de negativas).

Estas cosas motivan, y al mismo tiempo se clarifican en ella, la discusión que sigue, que culmina en la *ley de inercia de Sylvester*.

Sea A una matriz real simétrica y consideremos su forma cuadrática asociada $Q(v) = (vA, v)$. Si T es una transformación lineal real no singular cualquiera, dado $v \in F^{(n)}$, $v = wT$ para algún $w \in F^{(n)}$, de donde $(vA, v) = (wTA, wT) = (wTAT', w)$. Luego A y TAT' definen, efectivamente, la misma forma cuadrática. Sugiere esto la siguiente

DEFINICIÓN. Dos matrices simétricas reales A y B son *congruentes* si hay una matriz real no singular T tal que $B = TAT'$.

LEMA 6.39. *La congruencia es una relación de equivalencia.*

Prueba. Escribamos, cuando A es congruente a B , $A \cong B$.

- 1) $A \cong A$ pues $A = 1A1'$.
- 2) Si $A \cong B$ entonces $B = TAT'$ donde T es no singular, de donde $A = SBS'$ donde $S = T^{-1}$. Luego $B \cong A$.
- 3) Si $A \cong B$ y $B \cong C$, entonces $B = TAT'$ mientras que $C = RBR'$, de donde $C = RTAT'R' = (RT)A(RT)',$ y por tanto $A \cong C$.

Como la relación satisface las condiciones definitorias para una relación de equivalencia, el lema queda probado.

El principal teorema que concierne a las congruencias es su caracterización, contenida en la *ley de Sylvester*.

TEOREMA 6.2.3. *Dada la matriz real simétrica A hay una matriz invertible T tal que*

$$TAT' = \begin{pmatrix} I_r & & \\ & -I_s & \\ & & 0_t \end{pmatrix}$$

donde I_r e I_s son respectivamente las matrices unitarias $r \times r$ y $s \times s$ y donde 0_t es la 0 matriz $t \times t$. Los enteros $r+s$, el rango de A , y $r-s$, la signatura de A , caracterizan la clase de congruencia de A . Es decir, dos matrices simétricas reales son congruentes si y sólo si tienen el mismo rango y la misma signatura.

Prueba. Como A es real simétrica sus raíces características son todas reales; sean $\lambda_1, \dots, \lambda_r$ sus raíces características positivas y $-\lambda_{r+1}, \dots, -\lambda_{r+s}$ sus raíces negativas. Por la discusión al final de la sección 10 podemos encontrar una matriz ortogonal real C tal que

$$CAC^{-1} = CAC' = \begin{pmatrix} \lambda_1 & & & & \\ & \ddots & & & \\ & & \lambda_r & & \\ & & & -\lambda_{r+1} & \\ & & & & \ddots \\ & & & & -\lambda_{r+s} \\ & & & & & 0_t \end{pmatrix}$$

donde $t = n - r - s$. Sea D la matriz diagonal real

$$D = \begin{pmatrix} \frac{1}{\sqrt{\lambda_1}} & & & & \\ & \ddots & & & \\ & & \frac{1}{\sqrt{\lambda_r}} & & \\ & & & \frac{1}{\sqrt{\lambda_{r+1}}} & \\ & & & & \ddots \\ & & & & & \frac{1}{\sqrt{\lambda_{r+s}}} \\ & & & & & & I_t \end{pmatrix};$$

un simple cálculo muestra que

$$DCAC'D' = \begin{pmatrix} I_r & & \\ & -I_s & \\ & & 0_t \end{pmatrix}.$$

Luego hay una matriz de la forma requerida en la clase de congruencia de A .

Nuestra tarea es ahora demostrar que ésta es la única matriz en la clase de congruencia de A de esta forma o, lo que es equivalente, que

$$L = \begin{pmatrix} I_r & & \\ & -I_s & \\ & & 0_t \end{pmatrix} \quad \text{y} \quad M = \begin{pmatrix} I_r' & & \\ & -I_s' & \\ & & 0_t' \end{pmatrix}$$

son congruentes solamente si $r = r'$, $s = s'$ y $t = t'$.

Supongamos que $M = TLT'$ donde T es invertible. Por el lema 6.3 el rango de M es igual al de L : como el rango de M es $n-t'$ mientras que el de L es $n-t$, tenemos, $t = t'$.

Supongamos que $r < r'$; como $n = r+s+t = r'+s'+t'$, y como $t = t'$, debemos tener $s > s'$. Sea U el subespacio de $F^{(n)}$ de todos los vectores que tienen las primeras s y las últimas t coordenadas iguales a 0; U es de dimensión s y para $u \neq 0$ en U , $(uL, u) < 0$.

Sea W el subespacio de $F^{(n)}$ para el que los componentes $r'+1, \dots, r'+s$ son todos 0; sobre W , $(wM, w) \geq 0$ para cualquier $w \in W$. Como T es invertible, y como W es $(n-s')$ -dimensional, WT es $(n-s')$ -dimensional. Para $w \in W$, $(wM, w) \geq 0$; de donde $(wTLT', w) \geq 0$; es decir, $(wTL, wT) \geq 0$. Por tanto, sobre WT , $(wTL, wT) \geq 0$ para todos los elementos. Ahora bien, $\dim(WT) + \dim(U) = (n-s') + r = n + s - s' > n$; luego según el corolario al lema 4.8, $WT \cap U \neq 0$. Pero esto no tiene sentido, pues si $x \neq 0 \in WT \cap U$, por una parte, estando en U , $(xL, x) < 0$, mientras que por la otra, estando en WT , $(xL, x) \geq 0$. Luego $r = r'$ y $s = s'$.

El rango $r+s$, y la signatura $r-s$, determinan desde luego r y s , y por lo tanto $t = (n-r-s)$, de donde determinan la clase de congruencia.

Problemas

1. Determínense el rango y la signatura de cada una de las siguientes formas cuadráticas reales:

a) $x_1^2 + 2x_1x_2 + x_2^2$.

b) $x_1^2 + x_1x_2 + 2x_1x_3 + 2x_2^2 + 4x_2x_3 + 2x_3^2$.

2. Si A es una matriz simétrica con entradas complejas, pruébese que podemos encontrar una matriz invertible compleja B tal que

$$BAB' = \begin{pmatrix} I_r \\ & 0_{t-r} \end{pmatrix}$$

y que r , el rango de A , determina la clase de congruencia de A respecto a la congruencia compleja.

3. Si F es un campo de característica diferente de 2, dada $A \in F^n$, pruébese que existe una $B \in F^n$ tal que BAB' es diagonal.

4. Pruébese que el resultado del problema 3 es falso si la característica de F es 2.

Lecturas supplementarias

HALMOS, PAUL R. *Finite Dimensional Vector Spaces*, segunda edición.
D. Van Nostrand Company, Inc., Princeton, Nueva Jersey, 1958.

CAPITULO

7


Tópicos selectos

EN ESTE último capítulo nos hemos marcado dos objetivos. El primero de ellos es presentar algunos resultados matemáticos que penetren más profundamente que la mayor parte del material que hasta ahora hemos visto, resultados que sean más sofisticados y un poco apartados del desarrollo general que hemos seguido. Nuestro segundo objetivo es escoger resultados de esta clase cuya discusión, además, haga uso de una gran sección transversa de ideas y teoremas de los anteriormente expuestos en este libro. Con estas finalidades en mente hemos escogido tres temas como puntos focales de este capítulo.

El primero de estos es un teorema famoso probado por Wedderburn en 1905 ("A Theorem on Finite Algebras", *Transactions of the American*

Mathematical Society, vol. 6 (1905), páginas 349-352) que afirma que un anillo con división que tiene solamente un número finito de elementos debe ser un campo conmutativo. Daremos dos pruebas de este teorema, totalmente diferentes una de otra. La primera seguirá fielmente la prueba original de Wedderburn y usará un argumento tipo conteo; se apoyará en gran medida sobre resultados que desarrollamos en el capítulo sobre teoría de grupos. La segunda usará una mezcla de argumentos de la teoría de grupos y de la teoría de campos, y sacará un gran partido del material que estudiamos en estas dos teorías. La segunda prueba tiene la evidente ventaja de que en su curso de ejecución obtendremos ciertos resultados colaterales que nos permitirán proceder a la prueba, en el caso de los anillos con división, de un bello teorema debido a Jacobson ("Structure Theory for Algebraic Algebras of Bounded Degree", *Annals of Mathematics*, vol. 46 (1945), páginas 695-707) que es una generalización de gran alcance del teorema de Wedderburn.

Nuestro segundo gran tema es un teorema debido a Frobenius ("Über lineäre Substitutionen und bilineären Formen", *Revue für die reine und angewandte Mathematik*, vol. 84 (1877), especialmente las páginas 59-63) que afirma que los únicos anillos con división algebraicos sobre el campo de todos los números reales son el campo de los números reales, el campo de los números complejos y el anillo con división de los cuaternios reales. El teorema señala un papel único para los cuaternios y es sorprendente, en cierto modo, que Hamilton los descubriera en su forma, podríamos decir, un poco *ad hoc*. Nuestra prueba del teorema de Frobenius, ahora completamente elemental, es una variación de un enfoque marcado por Dickson y Albert; empleará resultados de la teoría de polinomios y de la teoría de campos.

Nuestro tercer objetivo es el teorema de que todo entero positivo puede representarse como la suma de cuatro cuadrados. Este famoso resultado parece que fue conjecturado ya por el primitivo matemático griego Diofantos. Fermat trabajó en su demostración sin éxito y anunció con tristeza su derrota (en un escrito donde él, sin embargo, resolvió el teorema de los dos cuadrados que nosotros probamos en la sección 8 del capítulo 3). Euler abrió grandes brechas que, aprovechadas por Lagrange, permitieron que éste, en 1770, diera una primera prueba completa. Nuestro enfoque será completamente distinto del de Lagrange. Tiene sus raíces en el trabajo de Adolfo Hurwitz y empleará una generalización de los anillos euclidianos. Usando nuestras técnicas de teoría de anillos sobre un cierto anillo de cuaternios, el teorema de Lagrange caerá como una consecuencia.

En nuestra marcha hacia el establecimiento de estos teoremas, cosecharemos muchas ideas y resultados interesantes de por sí. Esto es característico de un buen teorema — su prueba invariablemente conduce a resultados colaterales de casi igual interés.

1. CAMPOS FINITOS

Antes que podamos entrar en una discusión del teorema de Wedderburn y de los anillos finitos con división, es esencial que investiguemos la naturaleza de los campos que tienen solo un número finito de elementos. Tales campos se llaman *campos finitos*. Es claro que existen campos finitos, pues el anillo J_p de los enteros módulo cualquier primo p nos da un ejemplo de tal campo. En esta sección determinaremos todos los posibles campos finitos y muchas de las importantes propiedades que poseen.

Comenzamos con el

LEMA 7.1. *Sea F un campo finito con q elementos y supongamos que $F \subset K$ donde K es también un campo finito. Entonces K tiene q^n elementos donde $n = [K:F]$.*

Prueba. K es un espacio vectorial sobre F y como K es finito es ciertamente de dimensión finita como espacio vectorial sobre F . Supongamos que $[K:F] = n$; entonces K tiene una base de n elementos sobre F . Sea v_1, \dots, v_n una tal base. Entonces todo elemento en K tiene una representación única en la forma $\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n$ donde $\alpha_1, \alpha_2, \dots, \alpha_n$ están todas en F . Así pues, el número de elementos en K es el número de $\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n$ que se producen cuando las $\alpha_1, \alpha_2, \dots, \alpha_n$ van tomando valores sobre F . Como cada coeficiente puede tomar q valores, K debe tener q^n elementos.

COROLARIO 1. *Sea F un campo finito; entonces F tiene p^m elementos donde el número primo p es la característica de F .*

Prueba. Como F tiene un número finito de elementos, el corolario 2 al teorema 2.a, $f! = 0$ donde f es el número de elementos de F . Así pues, F tiene característica p para algún número primo p . Por tanto F contiene un campo F_0 isomorfo a J_p . Como F_0 tiene p elementos, F tiene p^m elementos donde $m = [F:F_0]$ según el lema 7.1.

COROLARIO 2. *Si el campo finito F tiene p^m elementos, entonces todo $a \in F$ satisface $a^{p^m} = a$.*

Prueba. Si $a = 0$, la afirmación del corolario es trivialmente cierta.

Por otra parte, los elementos distintos de cero de F forman un grupo bajo la multiplicación de orden $p^m - 1$, luego, según el corolario 2 al teorema 2.a, $a^{p^m-1} = 1$ para todo $a \neq 0$ en F . Multiplicando esta relación por a obtenemos $a^{p^m} = a$.

De este último corolario podemos fácilmente pasar al

LEMA 7.2. *Si el campo finito F tiene p^m elementos, entonces el polinomio $x^{p^m} - x$ en $F[x]$ se factoriza en $F[x]$ como $x^{p^m} - x = \prod_{\lambda \in F} (x - \lambda)$.*

Prueba. De acuerdo con el lema 5.2, el polinomio $x^{p^m} - x$ tiene cuando más p^m raíces en F . Pero, según el corolario 2 al lema 7.1, conocemos p^m de tales raíces, a saber, todos los elementos de F . Por el corolario al lema 5.1 podemos concluir que $x^{p^m} - x = \prod_{\lambda \in F} (x - \lambda)$.

COROLARIO. *Si el campo F tiene p^m elementos, entonces F es el campo de descomposición del polinomio $x^{p^m} - x$.*

Prueba. Por el lema 7.2, $x^{p^m} - x$ se descompone en F . Pero no puede descomponerse en un campo más pequeño, porque ese campo tendría que tener todas las raíces de este polinomio y, por tanto, tendría que tener al menos p^m elementos. De esta manera, F es el campo de descomposición de $x^{p^m} - x$.

Como vimos en el capítulo 5 (teorema 5.j) cualesquiera dos campos de descomposición sobre un campo dado de un polinomio dado son isomorfos. A la luz del corolario al lema 7.2 podemos enunciar

LEMA 7.3. *Cualesquiera dos campos finitos que tienen el mismo número de elementos son isomorfos.*

Prueba. Si estos campos tienen p^m elementos, por el anterior corolario ambos son campos de descomposición del polinomio $x^{p^m} - x$, sobre J_p , luego ambos son isomorfos.

Así pues, para cualquier entero m y cualquier número primo p hay, salvo isomorfismo, cuando más un campo que tiene p^m elementos. El propósito del próximo lema es demostrar que para cualquier número primo p y cualquier entero m hay un campo que tiene p^m elementos. Cuando hayamos hecho esto, sabremos que hay exactamente un campo con p^m elementos, donde p es un primo arbitrario y m entero arbitrario.

LEMA 7.4. *Para todo número primo p y todo entero positivo m existe un campo con p^m elementos.*

Prueba. Consideremos el polinomio $x^{p^m} - x$ en $J_p[x]$, el anillo de polinomios en x sobre J_p , el campo de los enteros mod p . Sea K el campo de descomposición de este polinomio. En K sea $F = \{a \in K \mid a^{p^m} = a\}$. Los elementos de F son, pues, las raíces de $x^{p^m} - x$ que, según el corolario 2 al lema 5.6 son distintas, de donde F tiene p^m elementos. Afirmando ahora que

F es un campo. Si $a, b \in F$, entonces $a^{p^m} = a, b^{p^m} = b$, y así $(ab)^{p^m} = a^{p^m}b^{p^m} = ab$; luego $ab \in F$. Además, como la característica es p , $(a \pm b)^{p^m} = a^{p^m} \pm b^{p^m} = a \pm b$, de donde $a \pm b \in F$. Por consiguiente F es un subcampo de K y, por tanto, un campo. Al mostrar que el campo F tiene p^m elementos, hemos probado el lema 7.4.

Combinando los lemas 7.3 y 7.4, tenemos

TEOREMA 7.A. *Para todo número primo p y todo entero positivo m , hay un campo único que tiene p^m elementos.*

Volvamos ahora, por un momento, a la teoría de los grupos. El resultado de la teoría de los grupos que buscamos, determinará la estructura de cualquier subgrupo multiplicativo finito del grupo de elementos distintos de cero de un campo y, en particular, determinará la estructura multiplicativa de cualquier campo finito.

LEMA 7.5. *Sea G un grupo abeliano finito con la propiedad de que la relación $x^n = e$ se satisface por, a lo más, n elementos de G , para todo entero n . Entonces G es un grupo cíclico.*

Prueba. Si el orden de G es una potencia de algún número primo q entonces el resultado es muy sencillo. Supongamos, en efecto, que $a \in G$ es un elemento cuyo orden es todo lo grande que sea posible; su orden debe ser q^r para algún entero r . Los elementos $e, a, a^2, \dots, a^{q^r-1}$ nos dan q^r soluciones distintas de la ecuación $x^{q^r} = e$ que, por nuestra hipótesis, implica que estas son todas las soluciones de la ecuación. Ahora bien, si $b \in G$, su orden es q^s donde $s \leq r$, de donde $b^{q^r} = (b^{q^s})^{q^{r-s}} = e$. Por la observación anteriormente hecha, esto obliga a que $b = a^i$ para algún i , y por lo tanto G es cíclico.

El grupo abeliano finito general G , puede realizarse como $G = S_{q_1} S_{q_2} \dots S_{q_k}$ donde las q_i son los distintos divisores primos de $o(G)$ y donde los S_{q_i} son los subgrupos de Sylow de G . Además, todo elemento $g \in G$ puede escribirse de forma única como $g = s_1 s_2 \dots s_k$, donde $s_i \in S_{q_i}$ (véase la sección 7, capítulo 2). Cualquier solución de $x^n = e$ en S_{q_i} es una de $x^n = e$ en G , de forma que todo S_{q_i} hereda la hipótesis que hemos impuesto sobre G . Por las observaciones del primer párrafo de la prueba, cada S_{q_i} es un grupo cíclico; sea a_i un generador de S_{q_i} . Afirmamos que $c = a_1 a_2 \dots a_k$ es un generador cíclico de G . Para verificar esto, todo lo que tenemos que hacer es probar que $o(G)$ divide a m , el orden de c . Como $c^m = e$, tenemos que $a_1^m a_2^m \dots a_k^m = e$. Por la unicidad de la representación de un elemento de G como un producto de elementos en las S_{q_i} , concluimos que $a_i^m = e$ para toda i . Luego $o(S_{q_i}) \mid m$ para toda i . Luego $o(G) = o(S_{q_1}) o(S_{q_2}) \dots o(S_{q_k}) \mid m$. Pero $m \mid o(G)$, luego $o(G) = m$. Lo que prueba que G es cíclico.

El lema 7.5 tiene una consecuencia importante.

LEMA 7.6. *Sea K un campo y sea G un subgrupo finito del grupo multiplicativo de elementos distintos de cero de K . Entonces G es un grupo cíclico.*

Prueba. Como K es un campo, cualquier polinomio de grado n en $K[x]$ tiene cuando más n raíces en K . Luego, en particular, para cualquier entero n , el polinomio $x^n - 1$ tiene cuando más n raíces en K , y también cuando más, n raíces en G , evidentemente. La hipótesis del lema 7.5 se satisface, luego G es cíclico.

Aun cuando la situación de un campo finito es un caso particular tan solo del lema 7.6, es de interés en tantos campos que lo subrayamos enunciándolo como un

TEOREMA 7.B. *El grupo multiplicativo de elementos distintos de cero de un campo finito es cíclico.*

Prueba. Sea F un campo finito. Aplicando simplemente el lema 7.6 con $F = K$ y $G = \text{grupo de elementos distintos de cero de } F$, tenemos el resultado.

Concluimos esta sección usando un argumento de conteo para probar la existencia de soluciones de ciertas ecuaciones en un campo finito. Necesitaremos el resultado en una demostración del teorema de Wedderburn.

LEMA 7.7. *Si F es un campo finito y $\alpha \neq 0, \beta \neq 0$ son dos elementos de F , entonces podemos encontrar elementos a y b en F tales que $1 + \alpha a^2 + \beta b^2 = 0$.*

Prueba. Si la característica de F es 2, F tiene 2^n elementos y cada elemento x en F satisface $x^{2^n} = x$. Así pues, cada elemento en F es un cuadrado. En particular $\alpha^{-1} = a^2$ para alguna $a \in F$. Usando esta a y $b = 0$ tenemos $1 + \alpha a^2 + \beta b^2 = 1 + \alpha a^{-1} + 0 = 1 + 1 = 0$, en donde la última igualdad es una consecuencia del hecho de que la característica de F es 2.

Si la característica de F es un número impar primo p , F tiene p^n elementos. Sea $W_\alpha = \{1 + \alpha x^2 | x \in F\}$. ¿Cuántos elementos hay en W_α ? Debemos comprobar cuantas veces $1 + \alpha x^2 = 1 + \alpha y^2$. Pero esta relación obliga a que $\alpha x^2 = \alpha y^2$ y, por tanto, como $\alpha \neq 0$, a que $x^2 = y^2$. Finalmente, esto nos lleva a que $x = \pm y$. Luego para $x \neq 0$ tenemos de cada par x y $-x$ un elemento en W_α y para $x = 0$ obtenemos $1 \in W_\alpha$. Luego W_α tiene $1 + \frac{p^n - 1}{2} = \frac{p^n + 1}{2}$ elementos. Análogamente $W_\beta = \{-\beta x^2 | x \in F\}$ tiene $\frac{p^n + 1}{2}$ elementos. Como tanto W_α como W_β tiene más de la mitad de los

elementos de F deben tener una intersección no vacía. Sea $c \in W_\alpha \cap W_\beta$. Como $c \in W_\alpha$, $c = 1 + \alpha a^2$ para algún $a \in F$; como $c \in W_\beta$, $c = -\beta b^2$ para algún $b \in F$. Por tanto, $1 + \alpha a^2 = -\beta b^2$, que por transposición nos da el resultado deseado, $1 + \alpha a^2 + \beta b^2 = 0$.

Problemas

1. De acuerdo con el teorema 7.b, los elementos distintos de cero de J_p forman un grupo cíclico bajo la multiplicación. Cualquier generador de este grupo se llama una *raíz primitiva de p* .
 - a) Encuéntrense las raíces primitivas de: 17, 23, 31.
 - b) ¿Cuántas raíces primitivas tiene un primo p ?
2. Usando el teorema 7b pruébese que $x^2 \equiv -1 \pmod{p}$ es soluble si y sólo si el primo impar p es de la forma $4n+1$.
3. Si a es un entero no divisible por el primo impar p pruébese que $x^2 \equiv a \pmod{p}$ es soluble para algún entero x si y sólo si $a^{(p-1)/2} \equiv 1 \pmod{p}$. (Se llama a esto el *criterio de Euler* para que a sea un residuo cuadrático mod p .)
4. Usando el resultado del problema 3 determínese si:
 - a) 3 es un cuadrado mód 17.
 - b) 10 es un cuadrado mód 13.
5. Si el campo F tiene p^n elementos pruébese que los automorfismos de F forman un grupo cíclico de orden n .
6. Si F es un campo finito, por los cuaternios sobre F entenderemos el conjunto de todos los $\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$ donde $\alpha_0, \alpha_1, \alpha_2, \alpha_3 \in F$ y donde la suma y la multiplicación se efectúan como en los cuaternios reales (es decir, $i^2 = j^2 = k^2 = ijk = -1$, etc.). Pruébese que los cuaternios sobre un campo finito *no* forman un anillo con división.

2. TEOREMA DE WEDDERBURN SOBRE ANILLOS FINITOS CON DIVISIÓN

En 1905, Wedderburn probó el teorema, considerado ahora como clásico, de que un anillo finito con división debía ser un campo comutativo. Este resultado ha captado la imaginación de la mayoría de los matemáticos, por lo inesperado de su contenido en que dos cosas aparentemente tan ajenas como son el número de elementos de un cierto sistema algebraico y la multiplicación en ese sistema, aparecen de pronto en una estrecha interrelación. Aparte de su intrínseca belleza el resultado ha sido muy importante

y útil, pues surge en los más variados contextos. Para citar solo un ejemplo; la única prueba conocida del hecho puramente geométrico de que en una geometría finita la configuración de Desargues implica la de Pappus (para la definición de estos términos véase cualquier buen texto de geometría proyectiva) consiste en reducir el problema geométrico a uno algebraico, y este problema algebraico tiene una solución basada en el teorema de Wedderburn. Para los algebristas, el teorema de Wedderburn ha servido como trampolín para saltar a una gran área de investigación, durante algunas décadas, concerniente a la comutatividad de anillos.

TEOREMA 7.C. *Un anillo finito con división es necesariamente un campo conmutativo.*

Primera prueba. Sea K un anillo finito con división y sea $Z = \{z \in K | zx = xz \text{ para todo } x \in K\}$ su centro. Si Z tiene q elementos entonces, como en la prueba del lema 7.1, se sigue que K tiene q^n elementos. Nuestro objetivo es probar que $Z = K$, o lo que es equivalente, que $n = 1$.

Si $a \in K$, sea $N(a) = \{x \in K | xa = ax\}$. $N(a)$ claramente contiene a Z , y, como una simple comprobación revela, $N(a)$ es un subanillo con división de K . Así pues, $N(a)$ contiene $q^{n(a)}$ elementos para algún entero $n(a)$. Afirmamos que $n(a)$ divide a n . En efecto, los elementos distintos de cero de $N(a)$ forman un subgrupo de orden $q^{n(a)} - 1$ del grupo, bajo la multiplicación, de elementos distintos de cero de K , que tiene $q^n - 1$ elementos. De acuerdo con el teorema de Lagrange (teorema 2.a) $q^{n(a)} - 1$ es un divisor de $q^n - 1$; pero esto obliga a que $n(a)$ sea un divisor de n (véase el problema 1 al final de esta sección).

En el grupo de elementos distintos de cero de K tenemos la relación de conjugación usada en el capítulo 2, a saber, a es conjugado de b si $a = x^{-1}bx$ para algún $x \neq 0$ en K .

Según el teorema 2.h el número de elementos de K conjugados de a es el índice del normalizador de a en el grupo de elementos distintos de cero de K . Por tanto, el número de conjugados de a en K es $\frac{q^n - 1}{q^{n(a)} - 1}$. Ahora bien, $a \in Z$ si y sólo si $n(a) = n$, luego, por la ecuación de clase (véase el corolario al teorema 2.h)

$$1) \quad q^n - 1 = q - 1 + \sum_{\substack{n(a)|n \\ n(a) \neq n}} \frac{q^n - 1}{q^{n(a)} - 1}$$

donde la suma es efectuada sobre una a en cada clase conjugada para a no en el centro.

El problema se ha reducido a probar que ninguna ecuación tal como la (1) puede verificarse en los enteros. Hasta este punto hemos seguido la prueba del artículo original de Wedderburn con casi absoluta fidelidad. Wedderburn

prosigue hasta desechar la posibilidad de la ecuación (1) haciendo uso del siguiente resultado de la teoría de números debido a Birkhoff y Vandiver: para $n > 1$ existe un número primo que es un divisor de $q^n - 1$, pero no es un divisor de ningún $q^m - 1$ donde m es un divisor propio de n , con las excepciones de $2^6 - 1 = 63$ cuyos factores primos ya se presentaron como divisores de $2^2 - 1$ y $2^3 - 1$, y $n = 2$, y q un primo de la forma $2^k - 1$. Si admitimos este resultado, ¿cómo acabaríamos la prueba? Este número primo sería un divisor del primer miembro de (1) y también un divisor de cada término de la suma que aparece en el segundo miembro pues divide a $q^n - 1$, pero no a $q^{n(a)} - 1$; luego este primo dividiría también a $q - 1$ dándonos una contradicción. El caso $2^6 - 1$ se tendría también que desechar, pero esto es sencillo. En el caso $n = 2$, la otra posibilidad no cubierta por el anterior argumento, no hay ningún subcampo entre Z y K lo que obliga a que $Z = K$. (¡Pruébese! Véase el problema 2.)

Pero no queremos aplicar el resultado de Birkhoff y Vandiver sin probarlo y su prueba nos llevaría a una digresión demasiado larga. Buscamos, pues, otro artificio. Nuestra finalidad es encontrar un entero que divida a $\frac{q^n - 1}{q^{n(a)} - 1}$, para todos los divisores $n(a)$ de n , pero que no divida a $q - 1$. Una vez hecho esto, la ecuación (1) será imposible salvo para $n = 1$ y, por tanto, el teorema de Wedderburn habrá sido probado. El medio que emplearemos con este propósito es la teoría de polinomios ciclotómicos. (Los hemos mencionado en los problemas al final de la sección 6, capítulo 5.)

Consideremos el polinomio $x^n - 1$ como elemento de $C[x]$ donde C es el campo de los números complejos. En $C[x]$

$$2) \quad x^n - 1 = \prod(x - \lambda),$$

donde este producto se toma sobre todos los λ que satisfacen $\lambda^n = 1$.

Un número complejo θ se dice que es *una raíz primitiva n-ésima de la unidad si $\theta^n = 1$ pero $\theta^m \neq 1$ para cualquier entero positivo $m < n$* . Los números complejos que satisfacen $x^n = 1$ forman un subgrupo finito, bajo la multiplicación, de los números complejos, de donde, según el teorema 7.b, este grupo es cíclico. Cualquier generador cíclico de este grupo debe, entonces, ser una raíz n -ésima primitiva de la unidad, de donde sabemos que tales raíces primitivas existen. (Alternativamente, $\theta = e^{2\pi i/n}$ nos da una raíz primitiva de la unidad.)

Sea $\Phi_n(x) = \prod(x - \theta)$ donde este producto se toma sobre todas las raíces n -ésimas primitivas de la unidad. Este polinomio se llama polinomio *ciclotómico*. Enumeramos los primeros polinomios ciclotómicos: $\Phi_1(x) = x - 1$, $\Phi_2(x) = x + 1$, $\Phi_3(x) = x^2 + x + 1$, $\Phi_4(x) = x^2 + 1$, $\Phi_5(x) = x^4 + x^3 + x^2 + x + 1$, $\Phi_6(x) = x^2 - x + 1$.

Nótese que todos ellos son polinomios mónimos con coeficientes enteros.

Nuestro primer objetivo es probar que, en general, $\Phi_n(x)$ es un polinomio mónico con coeficientes enteros. Reagrupamos la forma factorizada de $x^n - 1$ como se nos da en (2), y obtenemos

$$3) \quad x^n - 1 = \prod_{d|n} \Phi_d(x).$$

Por inducción, suponemos que $\Phi_d(x)$ es un polinomio mónico con coeficientes enteros para $d|n$, $d \neq n$. Luego $x^n - 1 = \Phi_n(x)g(x)$ donde $g(x)$ es un polinomio mónico con coeficientes enteros. Por tanto

$$\Phi_n(x) = \frac{x^n - 1}{g(x)},$$

que, al dividirse realmente (o por comparación de coeficientes), nos dice que $\Phi_n(x)$ es un polinomio mónico con coeficientes enteros.

Afirmamos ahora que para cualquier divisor d de n , donde $d \neq n$,

$$\Phi_n(x) \left| \frac{x^n - 1}{x^d - 1} \right.$$

en el sentido de que el cociente es un polinomio con coeficientes enteros. Para ver esto observemos primero que $x^d - 1 = \prod_{\substack{k|d \\ k \neq d}} \Phi_k(x)$, y como todo divisor de d es también un divisor de n , reagrupando términos en el segundo miembro de (3) obtenemos $x^d - 1$ sobre el segundo miembro; además, como $d < n$, $x^d - 1$ no envuelve a $\Phi_n(x)$. Por tanto, $x^n - 1 = \Phi_n(x)(x^d - 1)f(x)$ donde $f(x) = \prod_{\substack{k|n \\ k+d}} \Phi_k(x)$ tiene coeficientes enteros y por tanto

$$\Phi_n(x) \left| \frac{x^n - 1}{x^d - 1} \right.$$

en el sentido de que el cociente es un polinomio con coeficientes enteros. Y esto establece nuestra afirmación.

Para cualquier entero t , $\Phi_n(t)$ es un entero y, por lo anteriormente dicho, como un entero divide a $(t^n - 1)|(t^d - 1)$. En particular, volviendo a la ecuación (1),

$$\Phi_n(q) \left| \frac{q^n - 1}{q^{n(a)} - 1} \right.$$

y $\Phi_n(q)|(q^n - 1)$; luego por (1), $\Phi_n(q)|(q - 1)$. Afirmamos, sin embargo, que si $n > 1$ entonces $|\Phi_n(q)| > q - 1$. Pues $\Phi_n(q) = \prod (q - \theta)$ donde θ toma los valores de todas las raíces primitivas n -ésimas de la unidad y $|q - \theta| > q - 1$ para todo $\theta \neq 1$ una raíz de la unidad (pruébese) de donde $|\Phi_n(q)| =$

$\prod |q - \theta| > q - 1$. Es claro entonces que $\Phi_n(q)$ no puede dividir a $q - 1$, lo que nos lleva a una contradicción. Debemos por tanto suponer que $n = 1$, lo que obliga a admitir el teorema de Wedderburn.

Segunda prueba. Antes de examinar explícitamente los anillos finitos con división una vez más, probamos algunos lemas preliminares.

LEMA 7.8. *Sea R un anillo y sea $a \in R$. Sea T_a la aplicación de R en sí mismo definida por $xT_a = xa - ax$. Entonces*

$$\begin{aligned} xT_a^m &= xa^m - maxa^{m-1} + \frac{m(m-1)}{2} a^2 xa^{m-2} \\ &\quad - \frac{m(m-1)(m-2)}{3!} a^3 xa^{m-3} + \dots \end{aligned}$$

Prueba. ¿Qué es xT_a^2 ? $xT_a^2 = (xT_a)T_a = (xa - ax)T_a = (xa - ax)a - a(xa - ax) = xa^2 - 2axa + a^2x$. ¿Qué podemos decir acerca de xT_a^3 ? $xT_a^3 = (xT_a^2)T_a = (xa^2 - 2axa + a^2x)a - a(xa^2 - 2axa + a^2x) = xa^3 - 3axa^2 + 3a^2xa - a^3x$. Continuando de esta forma o por inducción, obtenemos el resultado del lema 7.8.

COROLARIO. *Si R es un anillo en el que $px = 0$ para toda $x \in R$, donde p es un número primo, entonces $xT_a^{pm} = xa^{pm} - a^{pm}x$.*

Prueba. De acuerdo con la fórmula del lema 7.8, si $p = 2$, $xT_a^2 = xa^2 - a^2x$, ya que $2axa = 0$. Así pues, $xT_a^4 = (xa^2 - a^2x)a^2 - a^2(xa^2 - a^2x) = xa^4 - a^4x$, y así sucesivamente hasta xT_a^{2m} .

Si p es un primo impar, de nuevo, según la fórmula del lema 7.8,

$$xT_a^p = xa^p - paxa^{p-1} + \frac{p(p-1)}{2} a^2 xa^{p-2} + \dots - a^p x,$$

y como

$$p \left| \frac{p(p-1) \cdots (p-i+1)}{i!} \right.$$

para $i < p$, todos los términos medios desaparecen y nos quedamos con $xT_a^p = xa^p - a^p x = xT_{ap}$. Ahora bien, $xT_a^{p^2} = x(T_{ap})^p = xT_{ap^2}$, y así sucesivamente por las potencias más altas de p .

LEMA 7.9. *Sea D un anillo con división de característica $p > 0$ con centro Z , y sea $P = \{0, 1, 2, \dots, (p-1)\}$ el subcampo de Z isomorfo a J_p . Supongamos*

que $a \in D$, $a \notin Z$ es tal que $a^{p^n} = a$ para algún $n \geq 1$. Entonces existe una $x \in D$ tal que

$$1) \quad xax^{-1} \neq a.$$

2) $xax^{-1} \in P(a)$, el campo obtenido por la adjunción de a a P .

Prueba. Definamos la aplicación T_a de D en sí mismo por $yT_a = ya - ay$ para todo $y \in D$.

$P(a)$ es un campo finito, ya que a es algebraico sobre P y tiene, digamos, p^m elementos. Todos ellos satisfacen $u^{p^m} = u$. De acuerdo con el corolario al lema 7.8, $yT_a^{p^m} = ya^{p^m} - a^{p^m}y = ya - ay = yT_a$, luego $T_a^{p^m} = T_a$.

Ahora bien, si $\lambda \in P(a)$, $(\lambda x)T_a = (\lambda x)a - a(\lambda x) = \lambda xa - \lambda ax = \lambda(xa - ax) = \lambda(xT_a)$, ya que λ commuta con a . Así pues, la aplicación λI de D en sí mismo definida por $\lambda I: y \rightarrow \lambda y$ commuta con T_a para todo $\lambda \in P(a)$. Ahora bien, el polinomio $u^{p^m} - u = \prod_{\lambda \in P(a)} (u - \lambda)$ por el lema 7.2. Como T_a commuta con λI para todo $\lambda \in P(a)$, y como $T_a^{p^m} = T_a$, tenemos que $0 = T_a^{p^m} - T_a = \prod_{\lambda \in P(a)} (T_a - \lambda I)$.

Si para todo $\lambda \neq 0$ en $P(a)$, $T_a - \lambda I$ no aniquila a ningún elemento distinto de cero en D (si $y(T_a - \lambda I) = 0$ implica $y = 0$), como $T_a(T_a - \lambda_1 I) \dots (T_a - \lambda_k I) = 0$, donde $\lambda_1, \dots, \lambda_k$ son los elementos distintos de cero de $P(a)$, tendríamos, $T_a = 0$. Es decir, $0 = yT_a = ya - ay$ para todo $y \in D$, lo que obligaría a que $a \in Z$ en contra de la hipótesis. Luego hay un $\lambda \neq 0$ en $P(a)$ y un $x \neq 0$ en D tales que $x(T_a - \lambda I) = 0$. Escribiendo esto explícitamente, $xa - ax - \lambda x = 0$; de donde $xax^{-1} = a + \lambda$ está en $P(a)$ y no es igual a a ya que $\lambda \neq 0$. Lo que prueba el lema.

COROLARIO. En el lema 7.9, $xax^{-1} = a^i \neq a$ para algún entero i .

Prueba. Sea a de orden s ; entonces en el campo $P(a)$ todas las raíces del polinomio $u^s - 1$ son $1, a, a^2, \dots, a^{s-1}$ ya que estas son, todas, raíces distintas y son s en total. Como $(xax^{-1})^s = xa^s x^{-1} = 1$, y como $xax^{-1} \in P(a)$, xax^{-1} es una raíz en $P(a)$ de $u^s - 1$, de donde $xax^{-1} = a^i$.

Tenemos ya todas las piezas que necesitábamos para efectuar nuestra segunda prueba del teorema de Wedderburn.

Sea D un anillo finito con división y sea Z su centro. Por inducción, podemos suponer que cualquier anillo con división que tenga menos elementos que D es un campo comutativo.

Hagamos notar primero que si $a, b \in D$ son tales que $b'a = ab'$ pero $ba \neq ab$, entonces $b' \in Z$. En efecto, consideremos $N(b') = \{x \in D \mid b'x = xb'\}$. $N(b')$ es un subanillo con división de D ; si no fuera D , por nuestra hipótesis de inducción sería comutativo. Pero tanto a como b se encuentran en $N(b')$ y no commutan, luego $N(b')$ no es comutativo, luego debe ser todo D . Luego $b' \in Z$.

Todo elemento distinto de cero en D tiene orden finito; luego alguna potencia positiva de él cae en Z . Dado $w \in D$ sea el orden de w relativo a Z el entero positivo mínimo $m(w)$ tal que $w^{m(w)} \in Z$. Escojamos un elemento a en D , pero no en Z , que tenga el mínimo orden relativo a Z posible, y sea tal orden r . Afirmamos que r es un número primo, pues si $r = r_1 r_2$ con $1 < r_1 < r_2 < r$, entonces a^{r_1} no está en Z , pero $(a^{r_1})^{r_2} = a^r \in Z$, luego a^{r_1} tiene un orden relativo a Z menor que el de a .

Por el corolario al lema 7.9, hay un $x \in D$ tal que $xax^{-1} = a^i \neq a$; luego $x^2ax^{-2} = x(xax^{-1})x^{-1} = xa^ix^{-1} = (xax^{-1})^i = (a^i)^i = a^{i^2}$. Análogamente, tenemos $x^{r-1}ax^{-(r-1)} = a^{i(r-1)}$. Pero r es un número primo, luego por el pequeño teorema de Fermat (corolario al teorema 2.a), $i^{r-1} = 1 + u_0r$, de donde $a^{i(r-1)} = a^{1+u_0r} = aa^{u_0r} = \lambda a$ donde $\lambda = a^{u_0r} \in Z$. Así pues, $x^{r-1}a = \lambda ax^{r-1}$. Como $x \notin Z$, por la naturaleza mínima de r , $x^{r-1}a \neq ax^{r-1}$ y por lo tanto $\lambda \neq 1$. Sea $b = x^{r-1}$; entonces $bab^{-1} = \lambda a$; por consiguiente, $\lambda^r a^r = (bab^{-1})^r = ba^rb^{-1} = a^r$, ya que $a^r \in Z$. Esta relación obliga a que $\lambda^r = 1$.

Afirmamos que si $y \in D$, entonces siempre que $y^r = 1$ ha de tenerse $y = \lambda^i$ para algún i , pues en el campo $Z(y)$ hay cuando más r raíces del polinomio $u^r - 1$; los elementos $1, \lambda, \lambda^2, \dots, \lambda^{r-1}$ de Z son todos distintos, ya que λ es del orden primo r y todos ellos constituyen las r raíces de $u^r - 1$ en $Z(y)$, en consecuencia de lo cual $y = \lambda^i$.

Como $\lambda^r = 1$, $b^r = \lambda^r b^r = (\lambda b)^r = (a^{-1}ba)^r = a^{-1}b^r a$, de donde obtenemos $ab^r = b^r a$. Como a commuta con b^r , pero no commuta con b por la observación antes hecha, b^r debe estar en Z . Según el teorema 7.b, el grupo multiplicativo de los elementos distintos de cero de Z es cíclico; sea $\gamma \in Z$ un generador. Entonces $a^r = \gamma^j$, $b^r = \gamma^k$; si $j = sr$, entonces $a^r = \gamma^s$, de donde $(a|\gamma^s)^r = 1$; esto implicaría que $a|\gamma^s = \lambda^i$, lo que implica $a \in Z$ en contra de $a \notin Z$. De donde $r \nmid j$; análogamente $r \nmid k$. Sea $a_1 = a^k$ y $b_1 = b^j$; un cálculo directo partiendo de $ba = \lambda ab$ nos lleva a $a_1 b_1 = \mu b_1 a_1$ donde $\mu = \lambda^{-jk} \in Z$. Como el número primo r que es el orden de λ no divide ni a j ni a k , $\lambda^{jk} = 1$, de donde $\mu \neq 1$. Nótese que $u^r = 1$.

Veamos donde estamos. Hemos producido dos elementos a_1 y b_1 tales que:

- 1) $a_1^{-r} = b_1^{-r} = x \in Z$.
- 2) $a_1 b_1 = \mu b_1 a_1$, con $\mu \neq 1$ en Z .
- 3) $\mu^r = 1$.

Calculamos $(a_1^{-1}b_1)^r$; $(a_1^{-1}b_1)^2 = a_1^{-1}b_1 a_1^{-1}b_1 = a_1^{-1}(b_1 a_1^{-1})b_1 = a_1^{-1}(\mu a_1^{-1}b_1)b_1 = \mu a_1^{-2}b_1^2$. Si calculamos $(a_1^{-1}b_1)^3$ encontramos que es igual a $\mu^{1+2}a_1^{-3}b_1^3$. Continuando de esta forma obtenemos $(a_1^{-1}b_1)^r = \mu^{1+2+\dots+(r-1)}a_1^{-r}b_1^r = \mu^{1+2+\dots+(r-1)} = \mu^{r(r-1)/2}$. Si r es un primo impar, como $\mu^r = 1$, tenemos $\mu^{r(r-1)/2} = 1$, de donde $(a_1^{-1}b_1)^r = 1$. Siendo una solución de $y^r = 1$, $a_1^{-1}b_1 = \lambda^i$ de modo que $b_1 = \lambda^i a_1$; pero

entonces $\mu b_1 a_1 = a_1 b_1 = b_1 a_1$, lo que contradice $\mu \neq 1$. Luego si r es un número primo impar, el teorema está probado.

Debemos ahora descartar el caso $r = 2$. En esta situación especial tenemos dos elementos $a_1, b_1 \in D$ tales que $a_1^2 = b_1^2 = x \in Z$, $a_1 b_1 = \mu b_1 a_1$ donde $\mu^2 = 1$ y $\mu \neq 1$. Así pues, $\mu = -1$ y $a_1 b_1 = -b_1 a_1 \neq b_1 a_1$; como consecuencia, la característica de D no es 2. De acuerdo con el lema 7.7 podemos encontrar elementos $\zeta, \eta \in Z$ tales que $1 + \zeta^2 - x\eta^2 = 0$. Consideraremos $(a_1 + \zeta b_1 + \eta a_1 b_1)^2$; al computar esto encontramos que $(a_1 + \zeta b_1 + \eta a_1 b_1)^2 = a_1(1 + \zeta^2 - x\eta^2) = 0$. Estando en un anillo con división esto implica que $a_1 + \zeta b_1 + \eta a_1 b_1 = 0$; luego $0 \neq 2a_1^2 = a_1(a_1 + \zeta b_1 + \eta a_1 b_1) + (a_1 + \zeta b_1 + \eta a_1 b_1)a_1 = 0$. Esta contradicción termina la prueba y el teorema de Wedderburn queda establecido.

Esta segunda prueba tiene la ventaja de que podemos utilizar partes de ella para establecer un resultado notable debido a Jacobson, a saber,

TEOREMA 7.D. (JACOBSON). *Sea D un anillo con división tal que para todo $a \in D$ existe un entero positivo $n(a) > 1$, dependiente de a , tal que $a^{n(a)} = a$. Entonces D es un campo conmutativo.*

Prueba. Si $a \neq 0$ está en D , entonces $a^n = a$ y $(2a)^m = 2a$ para algunos enteros $n, m > 1$. Sea $s = (n-1)(m-1)+1$; $s > 1$ y un simple cálculo muestra que $a^s = a$ y $(2a)^s = 2a$. Pero $(2a)^s = 2^s a^s = 2^s a$, de donde $2^s a = 2a$ de lo que se obtiene $(2^s - 2)a = 0$. Así pues, D tiene característica $p > 0$. Si $P \subset Z$ es el campo que tiene p elementos (isomorfo a J_p), como a es algebraico sobre P , $P(a)$ tiene un número finito de elementos, en realidad p^h elementos para algún entero h . Así pues, como $a \in P(a)$, $a^{p^h} = a$. Por tanto, si $a \notin Z$ todas las condiciones del lema 7.9 se satisfacen, de donde existe una $b \in D$ tal que

$$1) \quad bab^{-1} = a^n \neq a.$$

Por el mismo argumento, $b^{p^k} = b$ para algún entero $k > 1$. Sea $W = \{x \in D \mid x = \sum_{i=1}^{p^h} \sum_{j=1}^{p^k} p_{ij} a^i b^j \text{ donde } p_{ij} \in P\}$. W es finito y es cerrado respecto a la adición. Por virtud de (1) es también cerrado respecto a la multiplicación (¡Verifíquese!). Luego W es un anillo finito con división; por el teorema de Wedderburn es conmutativo. Pero a y b están ambas en W ; por tanto, $ab = ba$ en contra de que $a^n b = ba$. Y esto prueba el teorema.

El teorema de Jacobson realmente se verifica para *cualquier* anillo R que satisface $a^{n(a)} = a$ para todo $a \in R$, no solamente para anillos con división. La transición del caso de anillo con división al caso general aunque no

difícil exige la aplicación del axioma de elección, y discutirlo nos llevaría demasiado lejos.

Problemas

1. Si $t > 1$ es un entero y $(t^m - 1)|(t^n - 1)$, pruébese que $m|n$.
2. Si D es un anillo con división, pruébese que su dimensión (como espacio vectorial) sobre su centro no puede ser mayor que 2.
3. Pruébese que cualquier subanillo finito de un anillo con división es un anillo con división.
4. a) Sea D un anillo con división de característica $p \neq 0$ y sea G un subgrupo finito del grupo de elementos distintos de 0 de D bajo la multiplicación. Pruébese que G es abeliano. (*Sugerencia:* considérese el subconjunto $\{x \in D \mid x = \sum \lambda_i g_i, \lambda_i \in P, g_i \in G\}$.)
b) Pruébese en la parte (a) que G es realmente cíclico.
- *5. a) Si R es un anillo finito en el que $x^n = x$, para todo $x \in R$ donde $n > 1$, pruébese que R es comutativo.
b) Si R es un anillo finito en el que $x^2 = 0$ implica que $x = 0$, pruébese que R es comutativo.
- *6. Sea D un anillo con división y supongamos que $a \in D$ solamente tiene un número finito de conjugados (es decir, solamente un número finito de elementos $x^{-1}ax$). Pruébese que a tiene solamente un conjugado y debe estar en el centro de D .
7. Úsese el resultado del problema 6 para probar que si un polinomio de grado n con coeficientes en el centro de un anillo con división tiene $n+1$ raíces en el anillo con división, entonces tiene un número infinito de raíces en ese anillo con división.
- *8. Sea D un anillo con división y K un subanillo con división de D tal que $xKx^{-1} \subset K$ para todo $x \neq 0$ de D . Pruébese que o $K \subset Z$, el centro de D , o $K = D$. (Este resultado se conoce como el *teorema de Brauer-Cartan-Hua*.)
- *9. Sea D un anillo con división y K un subanillo con división de D . Supongamos que el grupo de elementos distintos de cero de K es un subgrupo de índice finito en el grupo (bajo la multiplicación) de elementos distintos de cero de D . Pruébese que entonces o D es finito o $K = D$.
10. Si $\theta \neq 1$ es una raíz de la unidad y si q es un entero positivo, pruébese que $|q - \theta| > q - 1$.

3. TEOREMA DE FROBENIUS

En 1877, Frobenius clasificó todos los anillos que tienen el campo de los números reales en su centro y que satisfacen, además, una condición que describiremos posteriormente. La finalidad de esta sección es presentar este trabajo de Frobenius.

En el capítulo 6 señalamos dos importantes hechos acerca del campo de los números complejos. Los recordamos aquí:

HECHO 1. Todo polinomio de grado n sobre el campo de los números complejos tiene todas sus n raíces en el campo de los números complejos.

HECHO 2. Los únicos polinomios irreducibles sobre el campo de los números reales son de grado 1 o 2.

DEFINICIÓN. Una álgebra con división D se dice que es *algebraica sobre un campo F* si:

- 1) F está contenido en el centro de D ;
- 2) todo $a \in D$ satisface un polinomio no trivial con coeficientes en F .

Si D , como espacio vectorial, es de dimensión finita sobre el campo F que está contenido en su centro, se puede mostrar fácilmente que D es algebraico sobre F (véase el problema 1 al final de esta sección). Pero puede suceder que D sea algebraico sobre F y, sin embargo, no sea de dimensión finita sobre F .

Comenzamos nuestra discusión sobre anillos algebraicos con división sobre el campo real investigando, en primer lugar, cuáles son los algebraicos sobre el campo complejo.

LEMA 7.10. *Sea C el campo de los números complejos y supongamos que el anillo con división D es algebraico sobre C . Entonces $D = C$.*

Prueba. Supongamos que $a \in D$. Como D es algebraico sobre C , $a^n + \alpha_1 a^{n-1} + \dots + \alpha_{n-1} a + \alpha_n = 0$ para algunas $\alpha_1, \alpha_2, \dots, \alpha_n$ en C .

Ahora bien, el polinomio $p(x) = x^n + \alpha_1 x^{n-1} + \dots + \alpha_{n-1} x + \alpha_n$ en $C[x]$ puede, por el hecho 1, factorizarse en $C[x]$ en un producto de factores lineales; es decir, $p(x) = (x - \lambda_1)(x - \lambda_2) \dots (x - \lambda_n)$ donde $\lambda_1, \lambda_2, \dots, \lambda_n$ están todos en C . Como C está en el centro de D , todo elemento de C commuta con a , de donde $p(a) = (a - \lambda_1)(a - \lambda_2) \dots (a - \lambda_n)$. Pero, por hipótesis, $p(a) = 0$, luego $(a - \lambda_1)(a - \lambda_2) \dots (a - \lambda_n) = 0$. Como un producto en un anillo con división es solo cero en el caso de que uno de los factores, al menos, sea cero, concluimos que $a - \lambda_k = 0$ para algún k , de donde $a = \lambda_k$, entonces se tiene que $a \in C$. Por tanto, todo elemento de D es de C ; como $C \subset D$, se obtiene $D = C$.

Estamos ahora en posición de probar el clásico resultado de Frobenius, a saber

TEOREMA 7.E (FROBENIUS). *Sea D un anillo con división algebraico sobre F , el campo de los números reales. Entonces D es isomorfo a uno de los siguientes: el campo de los números reales, el campo de los números complejos, o el anillo con división de los cuaternios reales.*

Prueba. La prueba consta de tres partes. En la primera, y más sencilla, resolvemos la cuestión para el caso conmutativo; en la segunda, suponiendo que D no es conmutativo, construimos una réplica de los cuaternios reales en D ; en la tercera parte mostramos que esta réplica de los cuaternios satisface completamente a D .

Supongamos que $D \neq F$ y que a está en D , pero no en F . De acuerdo con nuestras hipótesis, a satisface algún polinomio sobre F , de donde algún polinomio irreducible sobre F . Si esta ecuación es lineal, a debe estar en F en contra de lo supuesto. Así que podemos suponer que $a^2 - 2\alpha a + \beta = 0$ donde $\alpha, \beta \in F$. Luego $(a - \alpha)^2 = \alpha^2 - \beta$; afirmamos que $\alpha^2 - \beta < 0$, pues, de otra forma tendría una raíz cuadrada δ y tendríamos $a - \alpha = \pm \delta$ y, por tanto, a estaría en F . Como $\alpha^2 - \beta < 0$ se puede escribir como $-\gamma^2$ donde $\gamma \in F$. En consecuencia $(a - \alpha)^2 = -\gamma^2$, de donde $\left(\frac{a - \alpha}{\gamma}\right)^2 = -1$. Así pues si $a \in D$, $a \notin F$ podemos encontrar reales α, γ tales que $\left(\frac{a - \alpha}{\gamma}\right)^2 = -1$.

Si D es conmutativo, escogamos $a \in D$, $a \notin F$ y sea $i = \frac{a - \alpha}{\gamma}$ donde α, γ en F se escogen de modo que hagan $i^2 = -1$. Por tanto, D contiene a $F(i)$, un campo isomorfo al campo de los números complejos. Como D es conmutativo y algebraico sobre F es evidentemente también algebraico sobre $F(i)$. Según el lema 7.10 concluimos que $D = F(i)$. Luego si D es conmutativo entonces es F o $F(i)$.

Supongamos, entonces, que D no es conmutativo. Afirmamos que el centro de D debe ser exactamente F . Si así no fuera, habría un a en el centro que no sería de F . Pero entonces para algunas $\alpha, \gamma \in F$, $\left(\frac{a - \alpha}{\gamma}\right)^2 = -1$ de forma que el centro contendría un campo isomorfo al de los números complejos. Pero, de acuerdo con el lema 7.10, si los números complejos (o un campo isomorfo a ellos) estuviese en el centro de D entonces $D = C$, luego D sería conmutativo. Por tanto, F es el centro de D .

Sea $a \in D$, $a \notin F$; para algunas $\alpha, \gamma \in F$, $i = \frac{a - \alpha}{\gamma}$ satisface $i^2 = -1$. Como $i \notin F$, i no está en el centro de D . Por lo tanto hay un elemento $b \in D$ tal que

$c = bi - ib \neq 0$. Calculamos $ic + ci; ic + ci = i(bi - ib) + (bi - ib)i = ibi - i^2b + bi^2 - ibi = 0$, ya que $i^2 = -1$. Así pues, $ic = -ci$; deducimos de esto que $ic^2 = -c(ic) = -c(-ci) = c^2i$, es decir, c^2 commuta con i . Ahora bien, c satisface alguna ecuación cuadrática sobre F , $c^2 + \lambda c + \mu = 0$. Como c^2 y μ commutam con i , λc debe también commutar con i ; es decir, $\lambda ci = \lambda ic = -\lambda ci$, de donde $2\lambda ci = 0$, y como $2ci \neq 0$, tenemos que $\lambda = 0$. Luego $c^2 = -\mu$; como $c \notin F$ (pues $ci = -ic \neq ic$) podemos decir, como antes hicimos, que μ es positivo y por lo tanto $\mu = v^2$ donde $v \in F$. Por lo tanto, $c^2 = -v^2$; sea $j = \frac{c}{v}$. Entonces j satisface:

$$1) j^2 = \frac{c^2}{v^2} = -1.$$

$$2) ji + ij = \frac{c}{v}i + i\frac{c}{v} = \frac{ci + ic}{v} = 0.$$

Sea $k = ij$. Las i, j, k que hemos construido se comportan como las de los cuaternios, de donde $T = \{\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k \mid \alpha_0, \alpha_1, \alpha_2, \alpha_3 \in F\}$ forma un subanillo con división de D isomorfo a los cuaternios reales. ¡Hemos construido una réplica T , del anillo con división de los cuaternios reales en D !

Nuestro último objetivo es demostrar que $T = D$.

Si $r \in D$ satisface $r^2 = -1$, sea $N(r) = \{x \in D \mid xr = rx\}$. $N(r)$ es un subanillo con división de D ; además, r , y por lo tanto todos los $\alpha_0 + \alpha_1 r$, $\alpha_0, \alpha_1 \in F$, están en el centro de $N(r)$. Según el lema 7.10, de ello se sigue que $N(r) = \{\alpha_0 + \alpha_1 r \mid \alpha_0, \alpha_1 \in F\}$. Luego si $xr = rx$ entonces $x = \alpha_0 + \alpha_1 r$ para algunas α_0, α_1 en F .

Supongamos que $u \in D$, $u \notin F$. Para algunos $\alpha, \beta \in F$, $w = \frac{u - \alpha}{\beta}$ satisface $w^2 = -1$. Afirmamos que $wi + iw$ commuta tanto con i como con w ; para $i(wi + iw) = iwi + i^2w = iwi + wi^2 = (iw + wi)i$ ya que $i^2 = -1$. Análogamente, $w(wi + iw) = (wi + iw)w$. Por la observación del párrafo anterior, $wi + iw = \alpha'_0 + \alpha'_1 i = \alpha_0 + \alpha_1 w$. Si $w \notin T$ esta última relación implica $\alpha_1 = 0$ (pues de otra forma podríamos resolver para w en términos de i). Luego $wi + iw = \alpha_0 \in F$. Análogamente, $wj + jw = \beta_0 \in F$ y $wk + kw = \gamma_0 \in F$. Sea

$$z = w + \frac{\alpha_0}{2}i + \frac{\beta_0}{2}j + \frac{\gamma_0}{2}k.$$

Entonces

$$zi + iz = wi + iw + \frac{\alpha_0}{2}(i^2 + i^2) + \frac{\beta_0}{2}(ji + ij) + \frac{\gamma_0}{2}(ki + ik)$$

$$= \alpha_0 - \alpha_0 = 0;$$

análogamente $zj + jz = 0$ y $zk + kz = 0$. Afirmamos que estas relaciones obligan a z a ser 0. En efecto, $0 = zk + kz = z(ij + ij) = (zi + iz)j + i(jz - zj) = i(jz - zj)$, pues $zi + iz = 0$. Pero $i \neq 0$, y como estamos en un anillo con división de ello se sigue que $jz - zj = 0$. Pero $jz + zj = 0$. Luego $2jz = 0$, y como $2j \neq 0$, tenemos $z = 0$. Volviendo a la expresión para z tenemos

$$w + \frac{\alpha_0}{2}i + \frac{\beta_0}{2}j + \frac{\gamma_0}{2}k = 0,$$

de donde $w \in T$, en contradicción con $w \notin T$. Luego, ciertamente, $w \in T$. Como $w = \frac{u - \alpha}{\beta}$, $u = \beta w + \alpha$ y por lo tanto, $u \in T$. Hemos probado que cualquier elemento de D está en T . Como $T \subset D$ concluimos que $D = T$; como T es isomorfo a los cuaternios reales tenemos que D es isomorfo al anillo con división de los cuaternios reales. Pero esto es, exactamente, el enunciado del teorema.

Problemas

1. Si el anillo con división D es de dimensión finita como espacio vectorial sobre el campo F contenido en el centro de D , pruébese que D es algebraico sobre F .
2. Proporcíñese un ejemplo de un campo K algebraico sobre otro campo F , pero no finito dimensional sobre F .
3. Si A es un anillo algebraico sobre un campo F y A no tiene divisores de 0 pruébese que A es un anillo con división.

4. CUATERNIOS ENTEROS Y EL TEOREMA DE LOS CUATRO CUADRADOS

En el capítulo 3 consideramos cierta clase particular de dominios enteros, la de los dominios euclidianos. Cuando los resultados de esta clase de anillos se aplicaban al anillo de los enteros gaussianos obteníamos, como una consecuencia, el famoso resultado de Fermat de que todo número primo de la forma $4n + 1$ es la suma de dos cuadrados.

Consideraremos ahora un subanillo particular del de los cuaternios que en todos los aspectos, salvo en el de su falta de commutatividad, parecerá un anillo eucliano. A causa de ello será posible caracterizar explícitamente a sus ideales izquierdos. Esta caracterización de los ideales izquierdos nos llevará rápidamente a una prueba del teorema clásico de Lagrange, de que todo entero positivo es una suma de cuatro cuadrados.

Sea Q el anillo con división de los cuaternios reales. Procedemos a introducir una operación adjunta en Q , $*$, por la siguiente

DEFINICIÓN. Para $x = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$ en Q , el *adjunto* de x , al que denotaremos por x^* , está definido por $x^* = \alpha_0 - \alpha_1 i - \alpha_2 j - \alpha_3 k$.

LEMA 7.11. *El adjunto en Q satisface*

- 1) $x^{**} = x$
- 2) $(\delta x + \gamma y)^* = \delta x^* + \gamma y^*$
- 4) $(xy)^* = y^* x^*$

para todo x, y en Q y cualesquiera reales δ y γ .

Prueba. Si $x = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$, entonces $x^* = \alpha_0 - \alpha_1 i - \alpha_2 j - \alpha_3 k$. luego $x^{**} = (x^*)^* = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$, lo que prueba (1).

Sean $x = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$ y $y = \beta_0 + \beta_1 i + \beta_2 j + \beta_3 k$ elementos de Q y sean δ y γ números reales arbitrarios. Entonces $\delta x + \gamma y = (\delta\alpha_0 + \gamma\beta_0) + (\delta\alpha_1 + \gamma\beta_1)i + (\delta\alpha_2 + \gamma\beta_2)j + (\delta\alpha_3 + \gamma\beta_3)k$, luego, por la definición de $*$, $(\delta x + \gamma y)^* = (\alpha_0\delta + \beta_0\gamma) - (\alpha_1\delta + \beta_1\gamma)i - (\alpha_2\delta + \beta_2\gamma)j - (\alpha_3\delta + \beta_3\gamma)k = \delta(\alpha_0 - \alpha_1 i - \alpha_2 j - \alpha_3 k) + \gamma(\beta_0 - \beta_1 i - \beta_2 j - \beta_3 k) = \delta x^* + \gamma y^*$. Lo que es claro que prueba (2).

A la luz de (2), para probar (3) es suficiente hacerlo para una base de Q sobre los reales. Lo probamos para la base $1, i, j, k$. Ahora bien, $ij = k$, de donde $(ij)^* = k^* = -k = ji = (-j)(-i) = j^*i^*$. Análogamente $(ik)^* = k^*i^*, (jk)^* = k^*j^*$. Además, $(i^2)^* = (-1)^* = -1 = (i^*)^2$, y análogamente para j y k . Como (3) es cierto para los elementos de la base y (2) se verifica, (3) es cierto para todas las combinaciones lineales de los elementos de la base con coeficientes reales, de donde (3) se verifica para x y y de Q arbitrarios.

DEFINICIÓN. Si $x \in Q$ entonces la *norma* de x , a la que representaremos por $N(x)$, está definida por $N(x) = xx^*$.

Nótese que si $x = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$, entonces $N(x) = xx^* = (\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k)(\alpha_0 - \alpha_1 i - \alpha_2 j - \alpha_3 k) = \alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2$; por tanto, $N(0) = 0$ y $N(x)$ es un número real *positivo* para $x \neq 0$ en Q . En particular, para cualquier número real α , $N(\alpha) = \alpha^2$. Si $x \neq 0$, nótese

que $x^{-1} = \frac{1}{N(x)}x^*$.

LEMA 7.12. *Para todo $x, y \in Q$, $N(xy) = N(x)N(y)$.*

Prueba. Por la misma definición de la norma, $N(xy) = (xy)(xy)^*$; por parte (3) del lema 7.11, $(xy)^* = y^*x^*$ y por lo tanto $N(xy) = xyx^*y^*$. Pero $yy^* = N(y)$ es un número real y por tanto está en el centro de Q ; en

particular debe commutar con x^* . Por consiguiente, $N(xy) = x(yy^*)x^* = (xx^*)(yy^*) = N(x)N(y)$.

Como una consecuencia inmediata del lema 7.12 se tiene.

LEMA 7.13 (IDENTIDAD DE LAGRANGE). *Si $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ y $\beta_0, \beta_1, \beta_2, \beta_3$ son números reales, entonces $(\alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2)(\beta_0^2 + \beta_1^2 + \beta_2^2 + \beta_3^2) = (\alpha_0\beta_0 - \alpha_1\beta_1 - \alpha_2\beta_2 - \alpha_3\beta_3)^2 + (\alpha_0\beta_1 + \alpha_1\beta_0 + \alpha_2\beta_3 - \alpha_3\beta_2)^2 + (\alpha_0\beta_2 - \alpha_1\beta_3 + \alpha_2\beta_0 + \alpha_3\beta_1)^2 + (\alpha_0\beta_3 + \alpha_1\beta_2 - \alpha_1\beta_2 + \alpha_3\beta_0)^2$.*

Prueba. Hay desde luego una prueba obvia de este resultado, la de efectuar las multiplicaciones en ambos miembros y comparar los resultados.

Pero una forma más fácil de reconstruir el resultado y al mismo tiempo probarlo, es observar que el primer miembro es $N(x)N(y)$, mientras que el segundo miembro es $N(xy)$ con $x = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$ y $y = \beta_0 + \beta_1 i + \beta_2 j + \beta_3 k$. De acuerdo con el lema 7.12, $N(x)N(y) = N(xy)$, luego la identidad de Lagrange.

La identidad de Lagrange nos dice que la suma de cuatro cuadrados por la suma de cuatro cuadrados es, de nuevo, de una forma muy específica, la suma de cuatro cuadrados. Un resultado muy impresionante de Adolf Hurwitz dice que si la suma de n cuadrados por la suma de n cuadrados es de nuevo una suma de n cuadrados, donde esta última suma tiene términos bilinealmente calculables partiendo de las otras dos sumas, entonces $n = 1, 2, 4$ u 8 . Hay, en realidad, una identidad para el producto de sumas de ocho cuadrados, pero es demasiado largo y complicado para transcribirlo en este lugar.

Veamos ahora por qué es oportuno introducir el anillo de Hurwitz de cuaternios reales. Sea $\zeta = \frac{1}{2}(1+i+j+k)$ y

$$H = \{m_0\zeta + m_1i + m_2j + m_3k \mid m_0, m_1, m_2, m_3 \text{ enteros}\}.$$

LEMA 7.14. *H es un subanillo de Q . Si $x \in H$, entonces $x^* \in H$ y $N(x)$ es un entero positivo para todo elemento distinto de cero x de H .*

Dejamos la prueba del lema 7.14 para el lector. No ofrece dificultad alguna.

En cierto modo, H podría parecernos un anillo un poco extraño, arbitrario. ¿Por qué usar los cuaternios ζ ? ¿Por qué no considerar simplemente el anillo más natural $Q_0 = \{m_0 + m_1i + m_2j + m_3k \mid m_0, m_1, m_2, m_3 \text{ enteros}\}$? La contestación es que Q_0 no es suficientemente grande, mientras que H es, según el lema clave que sigue algo que parece suficiente. Necesitamos este lema por que nos va a permitir caracterizar los ideales izquierdos del anillo. Esta posibilidad quizás fue la razón por la que Hurwitz se inclinó a trabajar en H en lugar de en Q_0 .

LEMÁ 7.15 (ALGORITMO DE LA DIVISIÓN IZQUIERDA). Sean a y b elementos de H con $b \neq 0$. Entonces existen dos elementos c y d en H , tales que $a = cb + d$ y $N(d) < N(b)$.

Prueba. Antes de probar el lema, veamos qué es lo que nos dice. Si observamos la sección del capítulo 3 que trata de los anillos euclidianos, podemos ver que el lema 7.15 nos asegura que, excepto por su falta de conmutatividad, H tiene todas las propiedades de un anillo eucliano. El hecho de que los elementos de H puedan fallar en cuanto a conmutatividad se refiere, no nos preocupa. Ciertamente, debemos tener un poco de cuidado para no saltar a conclusiones erróneas; por ejemplo, $a = cb + d$, pero no tenemos ningún derecho a suponer que a es también igual a $bc + d$, pues b y c es posible que no comuten. Pero esto no influirá en ningún argumento de los que usemos.

Debemos comenzar por probar el lema en un caso muy particular, a saber, aquél en que a es un elemento arbitrario de H , pero b es un entero positivo n . Supongamos que $a = t_0\zeta + t_1i + t_2j + t_3k$ donde t_0, t_1, t_2 y t_3 son enteros y que $b = n$ donde n es un entero positivo. Sea $c = x_0\zeta + x_1i + x_2j + x_3k$ donde x_0, x_1, x_2 y x_3 son enteros aún por determinar. Queremos escogerlos en tal forma que hagan que obligadamente $N(a - cn) < N(n) = n^2$. Pero

$$\begin{aligned} a - cn &= \left(t_0 \left(\frac{1+i+j+k}{2} \right) + t_1 i + t_2 j + t_3 k \right) - nx_0 \left(\frac{1+i+j+k}{2} \right) \\ &\quad - nx_1 i - nx_2 j - nx_3 k \\ &= \frac{1}{2}(t_0 - nx_0) + \frac{1}{2}(t_0 + 2t_1 - n(t_0 + 2x_1))i \\ &\quad + \frac{1}{2}(t_0 + 2t_2 - n(t_0 + 2x_2))j + \frac{1}{2}(t_0 + 2t_3 - n(t_0 + 2x_3))k. \end{aligned}$$

Si pudiésemos escoger los enteros x_0, x_1, x_2, x_3 de tal forma que se tuviera $|t_0 - nx_0| \leq \frac{1}{2}n, |t_0 + 2t_1 - n(t_0 + 2x_1)| \leq n, |t_0 + 2t_2 - n(t_0 + 2x_2)| \leq n$ y $|t_0 + 2t_3 - n(t_0 + 2x_3)| \leq n$ entonces tendríamos

$$\begin{aligned} N(a - cn) &= \frac{(t_0 - nx_0)^2}{4} + \frac{(t_0 + 2t_1 - n(t_0 + 2x_1))^2}{4} + \dots \\ &\leq \frac{1}{16}n^2 + \frac{1}{4}n^2 + \frac{1}{4}n^2 + \frac{1}{4}n^2 < n^2 = N(n), \end{aligned}$$

que es el resultado deseado. Pero afirmamos que esto siempre puede hacerse:

I) Hay un entero x_0 tal que $t_0 = x_0n + r$ donde $-\frac{n}{2} \leq r \leq \frac{n}{2}$; para

este $x_0, |t_0 - x_0n| = |r| \leq \frac{n}{2}$.

- 2) Hay un entero k tal que $t_0 + 2t_1 = kn + r$ y $0 \leq r < n$. Si $k - t_0$ es par, póngase $2x_1 = k - t_0$; entonces $t_0 + 2t_1 = (2x_1 + t_0)n + r$ y $|t_0 + 2t_1 - (2x_1 + t_0)n| = r < n$. Si, por otra parte, $k - t_0$ es impar, hagamos $2x_1 = k - t_0 + 1$; entonces $t_0 + 2t_1 = (2x_1 + t_0 - 1)n + r = (2x_1 + t_0)n + r - n$, de donde $|t_0 + 2t_1 - (2x_1 + t_0)n| = |r - n| \leq n$ ya que $0 \leq r < n$. Podemos, pues, encontrar un entero x_1 que satisfaga $|t_0 + 2t_1 - (2x_1 + t_0)n| \leq n$.
- 3) Como en (2), podemos encontrar enteros x_2 y x_3 que satisfacen $|t_0 + 2t_2 - (2x_2 + t_0)n| \leq n$ y $|t_0 + 2t_3 - (2x_3 + t_0)n| \leq n$ respectivamente.

En el caso especial en que a es un elemento arbitrario de H y b es un entero positivo, hemos mostrado que el lema es cierto.

Vamos ahora al caso general en que a y b son elementos arbitrarios de H y $b \neq 0$. Según el lema 7.14, $n = bb^*$ es un entero positivo, luego existe un $c \in H$ tal que $ab^* = cn + d$, donde $N(d) < N(n)$. Luego $N(ab^* - cn) < N(n)$; pero $n = bb^*$ de donde tenemos $N(ab^* - cbb^*) < N(n)$ y, por tanto, $N((a - cb)b^*) < N(n) = N(bb^*)$. De acuerdo con el lema 7.12, esto se reduce a $N(a - cb)N(b^*) < N(b)N(b^*)$; como $N(b^*) > 0$ tenemos $N(a - cb) < N(b)$. Haciendo $d = a - cb$ tenemos $a = cb + d$ donde $N(d) < N(b)$. Y esto completa la prueba del lema.

Como en el caso conmutativo, podemos deducir del lema 7.15 el

LEMA 7.16. *Sea L un ideal izquierdo de H . Entonces existe un elemento $u \in L$ tal que todo elemento en L es un múltiplo izquierdo de u ; en otras palabras, existe un $u \in L$ tal que todo $x \in L$ es de la forma $x = ru$ donde $r \in H$.*

Prueba. Si $L = (0)$ nada hay que probar, simplemente hacemos $u = 0$.

Podemos, pues, suponer que L tiene elementos distintos del cero. Las normas de los elementos distintos de cero son enteros positivos (lema 7.14) de donde hay un elemento $u \neq 0$ en L cuya norma es mínima entre las de los elementos distintos de cero de L . Si $x \in L$, según el lema 7.15, $x = cu + d$ donde $N(d) < N(u)$. Pero d está en L porque x y u , y por tanto cu , están en L que es un ideal izquierdo. Luego $N(d) = 0$ y, por tanto, $d = 0$. De donde es una consecuencia que $x = cu$.

Antes de que podamos probar el teorema de los cuatro cuadrados, que es la finalidad de esta sección, necesitamos un lema más, a saber

LEMA 7.17. *Si $a \in H$ entonces $a^{-1} \in H$ si y solo si $N(a) = 1$.*

Prueba. Si tanto a como a^{-1} están en H , entonces, según el lema 7.14, tanto $N(a)$ como $N(a^{-1})$ son enteros positivos. Pero $aa^{-1} = 1$, de donde, de acuerdo con el lema 7.12, $N(a)N(a^{-1}) = N(aa^{-1}) = N(1) = 1$. Luego ha de tenerse $N(a) = 1$.

Por otra parte, si $a \in H$ y $N(a) = 1$, entonces $aa^* = N(a) = 1$ y $a^{-1} = a^*$. Pero según el lema 7.14, como $a \in H$ tenemos $a^* \in H$, de donde $a^{-1} = a^*$ está también en H .

Hemos determinado bastante de la estructura de H para usarlo en forma efectiva en el estudio de las propiedades de los enteros. Probamos ahora el clásico teorema de Lagrange.

TEOREMA 7.F. *Todo entero positivo puede expresarse como la suma de los cuadrados de cuatro enteros.*

Prueba. Dado un entero positivo n afirmamos en el teorema que $n = x_0^2 + x_1^2 + x_2^2 + x_3^2$ para cuatro enteros x_0, x_1, x_2 y x_3 . Como todo entero se factoriza en un producto de números primos, si todo número primo fuera realizable como una suma de cuatro cuadrados, teniendo en cuenta la identidad de Lagrange (lema 7.13), todo entero sería expresable como una suma de cuatro cuadrados. Hemos reducido el problema para poder considerar tan solo números primos n . El número primo 2 es claro que puede escribirse como la suma de cuatro cuadrados: $2 = 1^2 + 1^2 + 0^2 + 0^2$.

Luego, sin pérdida de generalidad, podemos suponer que n es un *número primo impar*. Como es costumbre, lo denotamos por p .

Consideremos los cuaternios Wp sobre Jp , los enteros mod p ; $Wp = \{x_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k \mid \alpha_0, \alpha_1, \alpha_2, \alpha_3 \in Jp\}$. Wp es un anillo finito; además, como $p \neq 2$ no es conmutativo, pues $ij = -ji \neq ji$. Luego, según el teorema de Wedderburn, no puede ser un anillo con división, de donde según el problema 1 al final de la sección 5 del capítulo 3, debe tener un ideal izquierdo que no sea ni (0) ni Wp .

Pero entonces el ideal bilateral V en H definido por $V = \{x_0 \zeta + x_1 i + x_2 j + x_3 k \mid p \text{ divide a } x_0, x_1, x_2 \text{ y } x_3\}$ no puede ser un ideal izquierdo máximo de H , ya que H/V es isomorfo a W_p . (Pruébese!) (Si V fuera un ideal máximo izquierdo en H , H/V , y por tanto W_p , no tendría otros ideales izquierdos que (0) y H/V).

Hay, pues, un ideal izquierdo L de H que satisface: $L \neq H$, $L \neq V$, y $L \supset V$. De acuerdo con el lema 7.16, hay un elemento $u \in L$ tal que todo elemento de L es un múltiplo izquierdo de u . Como $p \in V$, $p \in L$, de donde $p = cu$ para algún $c \in H$. Como $u \notin V$, c no puede tener un inverso en H , pues de otra forma $u = c^{-1}p$ estaría en V . Luego $N(c) > 1$ por lema 7.17. Como $L \neq H$, u no puede tener un inverso en H , de donde $N(u) > 1$. Luego $p = cu$, $p^2 = N(p) = N(cu) = N(c)N(u)$. Pero $N(c)$ y $N(u)$ son enteros, pues tanto c como u están en H , ambos son mayores que 1 y ambos dividen a p^2 . La única forma de que esto sea posible es que $N(c) = N(u) = p$.

Como $u \in H$, $u = m_0 \zeta + m_1 i + m_2 j + m_3 k$ donde m_0, m_1, m_2, m_3 son enteros; luego $2u = 2m_0 \zeta + 2m_1 i + 2m_2 j + 2m_3 k = (m_0 + m_0)i + (m_0 + m_0)j + (m_0 + m_0)k + 2m_1 i + 2m_2 j + 2m_3 k = m_0 + (2m_1 + m_0)i + (2m_2 + m_0)j + (2m_3 + m_0)k$. Por tanto, $N(2u) = m_0^2 + (2m_1 + m_0)^2 + (2m_2 + m_0)^2 + (2m_3 + m_0)^2$. Pero

$N(2u) = N(2)N(u) = 4p$ ya que $N(2) = 4$ y $N(u) = p$. Hemos demostrado que $4p = m_0^2 + (2m_1 + m_0)^2 + (2m_2 + m_0)^2 + (2m_3 + m_0)^2$. Y ya casi hemos terminado.

Para terminar la prueba introducimos un viejo truco de Euler: Si $2a = x_0^2 + x_1^2 + x_2^2 + x_3^2$, donde a, x_0, x_1, x_2 y x_3 son enteros, entonces $a = y_0^2 + y_1^2 + y_2^2 + y_3^2$ para algunos enteros y_0, y_1, y_2, y_3 . Para ver esto nótese que como $2a$ es par, los x son todos pares, todos impares, o dos pares y dos impares. En cualquiera de los tres casos podemos renombrar los x y aparearlos de forma que

$$y_0 = \frac{x_0 + x_1}{2}, \quad y_1 = \frac{x_0 - x_1}{2}, \quad y_2 = \frac{x_2 + x_3}{2}, \quad y_3 = \frac{x_2 - x_3}{2}$$

sean todos enteros. Pero

$$\begin{aligned} y_0^2 + y_1^2 + y_2^2 + y_3^2 &= \left(\frac{x_0 + x_1}{2}\right)^2 + \left(\frac{x_0 - x_1}{2}\right)^2 + \left(\frac{x_2 + x_3}{2}\right)^2 + \left(\frac{x_2 - x_3}{2}\right)^2 \\ &= \frac{1}{2}(x_0^2 + x_1^2 + x_2^2 + x_3^2) \\ &= \frac{1}{2}(2a) \\ &= a. \end{aligned}$$

Como $4p$ es una suma de cuatro cuadrados, según la observación que acabamos de hacer $2p$ también lo es; como $2p$ es una suma de cuatro cuadrados, p también debe ser igual a una tal suma. Luego $p = a_0^2 + a_1^2 + a_2^2 + a_3^2$ para algunos enteros a_0, a_1, a_2, a_3 , y el teorema de Lagrange ha quedado establecido.

Este teorema es el punto de partida de una gran área de investigación en teoría de números, la del llamado *problema de Waring*. Se pregunta en éste si todo entero puede escribirse como una suma de un número fijo de potencias k -ésimas. Por ejemplo, puede demostrarse que todo entero es la suma de nueve cubos, diecinueve cuartas potencias, etc. En el presente siglo el gran matemático Hilbert demostró que el problema de Waring tiene una respuesta afirmativa.

Problemas

1. Pruébese el lema 7.14.
2. Encuéntrense todos los elementos a de Q_0 tales que a^{-1} está también en Q_0 .
3. Pruébese que hay exactamente 24 elementos a en H tales que a^{-1} está también en H . Determínense todos ellos.

4. Proporcionese un ejemplo de una a y una b , $b \neq 0$, en Q_0 tales que sea imposible encontrar c y d en Q_0 que satisfagan $a = cb + d$ donde $N(d) < N(b)$.
5. Pruébese que si $a \in H$ entonces existen enteros α , β tales que $a^2 + 2a + \beta = 0$.
6. Pruébese que hay un entero positivo que no puede escribirse como la suma de tres cuadrados.
- *7. Exhíbase un número infinito de enteros positivos que no puedan escribirse como la suma de tres cuadrados.

Lecturas supplementarias

Para una discusión más profunda de campos finitos: ALBERT, A. A., *Fundamental Concepts of Higher Algebra*. University of Chicago Press, Chicago, 1956.

Para muchas pruebas del teorema de los cuatro cuadrados y una discusión del problema de Waring: HARDY, G. H. y WRIGHT, E. M., *An Introduction to the Theory of Numbers*, segunda edición. Clarendon Press, Oxford, Inglaterra, 1945.

Para otra prueba del teorema de Wedderburn: ARTIN, E., "Über einen Satz von Herrn J. H. Wedderburn", *Abhandlungen, Hamburg, Mathematisches Seminar*, vol. 5 (1928). págs. 245-250.