ISSN 1851-1317

Fascículo 6 | Cursos de grado

Carlos Marcelo Sánchez

Lecciones de Álgebra

Departamento de Matemática

Facultad de Ciencias Exactas y Naturales

Universidad de Buenos Aires

2014

Cursos de grado

Fascículo 6

ISSN 1851-1317 (Versión Electrónica)
ISSN 1851-1295 (Versión Impresa)

Universidad de Buenos Aires.
Departamento de Matemática
Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires
Ciudad Universitaria – Pabellón I
(1428) Ciudad de Buenos Aires
Argentina.
http://www.dm.uba.ar
e-mail. secre@dm.uba.ar
tel/fax: (+54-11)-4576-3335

Lecciones de Algebra

Carlos Marcelo Sanchez

Departamento de Matemática
Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires

2014

A mi esposa y mis hijos

A Lucía, Trinidad y Matilda

Prólogo

A través de los quince capítulos que comprende este volumen −dirigido principalmente a estudiantes de los primeros años de carreras científicas, tecnológicas y de formación docente−, ofrecemos a la consideración del lector una exposición sistemática y detallada de los temas básicos del Algebra.

La obra es prácticamente autocontenida y su lectura no requiere mayores conocimientos previos, mas allá de aquellos que el estudiante seguramente ha adquirido en su paso por la escuela media o en algún curso preuniversitario. De cualquier modo, en la bibliografía incluida al final del libro el lector dispondrá de fuentes de consulta en las que podrá ampliar su aprendizaje y comprensión de las distintas materias tratadas.

A lo largo del dictado de muchos cursos universitarios de nivel inicial, hemos observado repetidamente que, en general, el mayor déficit de preparación que presentan los alumnos no es de tipo informativo −lo que supondría la carencia de ciertos conocimientos específicos−, sino formativo, en cuanto a que la enseñanza recibida ha sido básicamente descriptiva, escasamente fundamentada y sin poner en relieve las similitudes y relaciones existentes entre los diversos temas impartidos. Atendiendo a esta realidad nos propusimos escribir este libro, en la esperanza de poder ayudar amenamente al lector a superar sus dificultades de formación y por supuesto adentrarlo en el estudio de temas ya más avanzados.

Por supuesto que no seremos demasiado drásticos, introduciendo bruscamente un estilo excesivamente formal. Nuestro propósito es familiarizar paulativamente al estudiante con el indispensable manejo del lenguaje abstracto del Algebra, tratando siempre de aunar lo intuitivo con lo formal. Es así que los sucesivos capítulos −además de contener definiciones, teoremas, demostraciones, etc.−, están salpicados de comentarios y ejemplos, que ilustran y motivan las distintas situaciones y que muchas veces brindan una descripción más coloquial de las mismas. Siguiendo esa línea, hemos incluido una respetable cantidad de ejercicios al final de cada una de las secciones que integran los capítulos, con la idea de que el lector afiance al intentar resolverlos su dominio de la teoría.

Iniciando una breve reseña de los contenidos específicos del libro y del uso de la bibliografía sugerida, comencemos señalando que los tres primeros capítulos −además de brindar información pertinente− son esencialmente formativos y de acostumbramiento al lenguaje matemático, particularmen-

te al algebraico. Para abrir el texto, hemos considerado necesario incluir un capítulo introductorio de lógica proposicional (numerado 0), con la finalidad de que el lector entrene y desarrolle su capacidad natural de razonar correctamente. No dudamos de que lo hace en el transcurso de su vida cotidiana, pero la Matemática es la ciencia del razonamiento y la deducción, y requiere de un lenguaje preciso regido por las leyes de la lógica. Puesto que no nos ocupamos en este capítulo de temas estrictamente algebraicos, su lectura es optativa (aunque recomendable), y el lector podría omitirlo y valerse de él como fuente de consulta frente a ulteriores dificultades. Para los interesados en profundizar en el tema, sugerimos la lectura de las referencias [12] y [16] .

En el capítulo 1 desarrollamos desde un punto de vista intuitivo la teoría elemental de conjuntos e introducimos y estudiamos en detalle los conceptos de relación y función. Son temas básicos que luego usaremos intensivamente en el resto del libro, y como complemento a su estudio el lector podrá consultar por ejemplo [8] ó [12] .

En el capítulo 2, luego de un recordatorio informal de los distintos conjuntos numéricos, brindamos una presentación axiomática del cuerpo de números reales, con prescindencia inicial de la naturaleza de los mismos. El tema es de gran interés, porque aparece la idea de estructura algebraica y el estudiante se ve enfrentado, quizás por primera vez, al desafío de demostrar muchas propiedades númericas conocidas a partir de la validez de unos pocos axiomas, en particular el axioma de completitud, de notable importancia en el análisis matemático. A los interesados en una presentación constructiva de los números reales y en una prueba de la existencia de raíz enésima de un número real positivo, sugerimos consultar [18] y [21], respectivamente.

Contando con la estructura de cuerpo de los números reales, dedicamos el capítulo 3 a definir formalmente los conjuntos de números naturales, enteros y racionales, y principalmente a establecer e ilustrar en sus distintas versiones el principio de inducción, estrechamente ligado a la definición de número natural. Nos valemos para ello de una buena cantidad de ejemplos, en los que mostramos y explicamos el uso de de la inducción en la demostración de fórmulas numéricas, en la definición de secuencias, y en general, en el tratamiento y resolución de problemas recursivos. Para ampliar sus conocimientos sobre el tema, y casi seguro para entusiarmarse con él, sugerimos al lector consultar [7] .

En la primera sección del capítulo 4 exponemos las propiedades básicas de los cardinales finitos, que servirán de sustento teórico a las técnicas de conteo que desarrollamos luego en la segunda sección. En el plano algebraico −a raíz de la aparición natural de los números combinatorios−, demostramos la fórmula binomial de Newton y la fórmula multinomial de Leibniz. Por último cerramos el capítulo con una breve introducción a la teoría de la probabilidad, ciñéndonos al caso discreto. En cuanto a bibliografía auxiliar, sugerimos [14] para cuestiones de Combinatoria y [4] ó [7] para una lectura más avanzada de la teoría de la probabilidad.

Los capítulos 5, 6 y 7 están dedicados a la teoría elemental de números. Los dos primeros contienen los temas básicos de la misma, desde el concepto primordial de divisibilidad entera hasta los teoremas clásicos de la aritmética modular, mientras que en el último desarrollamos temas algo menos elementales, en general ausentes en un primer libro de álgebra y que el lector puede considerar de lectura optativa. Dentro de la muy amplia bibliografía disponible sobre la materia sugerimos unos cuantos títulos, apuntando a diversos objetivos. En [6] el lector hallará una buena lista de problemas interesantes, mientras que los que deseen realizar un estudio sistemático más completo pueden consultar [15] y [1], aclarando que la lectura de este último libro requiere conocimientos bastante avanzados de análisis matemático. Aquellos curiosos por conocer algo más de la historia de la conjetura de Fermat, citada en el texto, pueden consultar [17] ó [20] (obra de divulgación), mientras que recomendamos a los interesados en las modernas aplicaciones criptográficas de la Aritmética la lectura de [2] ó [13]. Mencionemos para finalizar la referencia [5], que brinda una interesante reseña del desarrollo histórico de la teoría de números hasta mediados del siglo pasado.

Partiendo de la motivación tradicional de ampliar la resolubilidad de las ecuaciones cuadráticas con coeficientes reales, introducimos en el capítulo 8 los números complejos. Una vez establecida la estructura de cuerpo del conjunto de los mismos, −pensados como pares ordenados de números reales−, adoptamos un enfoque más geométrico, pensando a los números complejos como puntos del plano y representándolos trigonométricamente mediante sus coordenadas polares. A partir de la fórmula de De Moivre, ello nos permitirá probar entre otras cosas que todo número complejo admite raíces enésimas complejas. Dentro de este contexto estudiamos especialmente los grupos de raíces de la unidad, incluyendo el concepto de raíz primitiva.

En un marco más analítico, el lector interesado podrá consultar por ejemplo [3] para ver cómo pueden extenderse al campo complejo ciertas funciones reales, como las trigonométricas, exponenciales y logarítmicas.

Los tres siguientes capítulos están dedicados a la teoría básica de los anillos de polinomios. Iniciamos el capítulo 9 estableciendo dicha estructura en el conjunto de polinomios con coeficientes en anillos numéricos, trasladando luego la situación al caso general de polinomios con coeficientes en un anillo cualquiera. Entre otras cuestiones, dedicamos especial atención al concepto de raíz de un polinomio, que desarrollamos con amplitud. Además de tratar las generalidades del tema, deducimos por ejemplo la fórmula de Cardano−Tartaglia para la resolución de ecuaciones cúbicas con coeficientes complejos y realizamos una breve exposición acerca del carácter algebraico o trascendente de un número complejo.

En el capítulo 10 desarrollamos la noción de divisibilidad en anillos de poliomios con coeficientes en un cuerpo, hasta arribar al concepto de irreducibilidad y al teorema de factorización única. Puesto que la situación es muy similar a la del caso entero obviamos algunas demostraciones, tratando de enfatizar en cambio la fuerte conexión existente entre factorización y

raíces. Finalmente, en la primera sección del capítulo 11 particularizamos los resultados obtenidos al caso de polinomios con coeficientes racionales, reales y complejos, mostrando qué tipo de factorización se alcanza en cada caso. En la segunda sección (de lectura optativa), extendemos la noción de polinomio a la de fracción racional, mostrando la estructura de cuerpo del conjunto de estas últimas. Obtenemos como colofón un resultado de utilidad en el cálculo integral, como es el hecho de que toda fracción racional se descompone como suma de fracciones simples.

Como bibliografía general de apoyo al aprendizaje de los temas citados sugerimos [11]. Además, recomendamos al estudiante interesado en conocer una prueba relativamente accesible del teorema fundamental del álgebra −que enunciamos pero no demostramos en el capítulo 11− la lectura de [19], aclarando que ésta requiere cierto manejo de la teoría de cuerpos.

Puesto que es materia presente en todos los campos de la Matemática, hemos incluido en la obra algunas nociones básicas de Algebra Lineal, a la que dedicamos los capítulos 12 y 13. Estudiamos en ellos la estructura de los espacios vectoriales de dimensión finita sobre un cuerpo cualquiera, los homomorfismos de dichas estructuras y los espacios de matrices, poniendo especial énfasis en la íntima conexión existente entre transformaciones lineales y matrices. Como temas prácticos pero de alcance teórico, exhibimos en el capítulo 13 sendos métodos de cálculo para la resolución de sistemas de ecuaciones lineales y para la determinación de la inversibilidad de una matriz cuadrada. El lector interesado en avanzar sobre otros temas del Algebra Lineal, como el espacio dual, la forma canónica de Jordan y los espacios vectoriales con producto interno, puede consultar por ejemplo [10] .

En el capítulo 14 ofrecemos una suerte de diccionario algebraico, en el que el lector podrá hallar referencias precisas de ciertas estructuras algebraicas abstractas (grupos, anillos, álgebras, etc.), algunas de las cuales fueron tratadas con anterioridad en situaciones más concretas. Sin ahondar en los resultados pero con abundancia de ejemplos, brindamos en cada caso una descripción general de los principales conceptos de la teoría, verbigracia los de subestructura y homomorfismo. Los contenidos de este capítulo deben interpretarse entonces como una introducción al lenguaje algebraico y como fuente primaria de consulta. Recomendamos a los estudiantes que deseen tener un panorama más amplio de estos temas la lectura de [9] ó [11] .

Agradecimientos

Deseo mencionar en primer término al profesor Natalio H. Guersenzvaig, que me sugirió la idea de escribir un libro de estas características. Juntos trazamos el plan general de la obra y él redactó una primera versión de los capítulos 3 y 4, muy cercana en espíritu y forma a la definitiva. Su capacidad intelectual y el empuje y entusiasmo que ha sabido transmitirme impulsaron notablemente la concreción de este trabajo. Vaya pues mi reconocimiento a este querido amigo.

Quiero también expresar mi agradecimiento al profesor Gustavo Piñeiro por su bosquejo inicial de los capítulos 12 y 13 de algebra lineal, cuyo enfoque se ve reflejado en el texto final ofrecido al lector. Extiende mi deuda con Gustavo su tarea de revisión de la totalidad del libro. Sus interesantes observaciones y sugerencias han contribuido sin duda a mejorarlo.

Mi gratitud y amor a mi esposa (y colega) María Elena Becker, que con paciencia atendió tantas veces mis dudas e interrogantes, y que leyó y corrigió cuando fue necesario las muchas páginas que sometí a su consejo.

Por último, un cálido agradecimiento al profesor Pablo Solernó, que con generosa amistad me ayudó a salvar diversas dificultades que fueron presentándose en la compilación de este libro.

Carlos M. Sanchez

Índice general

Capítulo 0

Nociones de Lógica

0.1. Lógica Proposicional

0.1.1. Proposiciones

Queremos en este breve capítulo introductorio familiarizar al lector con los conceptos básicos de la *lógica proposicional*, con el propósito de facilitarle en adelante la comprensión del lenguaje y el significado de los enunciados matemáticos. No seremos excesivamente formales y apelaremos frecuentemente al sentido natural de la lógica que sin duda posee el lector (no podría hacerse entender si no lo tuviera), pero sí trataremos de acostumbrarlo a través de estas sucintas lecciones de cálculo proposicional abstracto al lenguaje preciso y al pensamiento riguroso, condimentos indispensables del quehacer matemático.

Cualquiera que haya estudiado un poco de Matemática se ha encontrado con frases del tipo "toda función derivable es continua", "si dos rectas del plano no se cortan entonces son paralelas" o "no existe un número real mayor que su cuadrado". En todas ellas se afirma algo (de modo negativo en la última), y si bien podríamos en principio no saber si lo que se asevera en cada una de ellas es verdadero o falso, nuestro sentido de la lógica (innato o adquirido) nos lleva a convenir en que debe registrarse una, y solo una, de estas dos opciones.

Los enunciados anteriores son ejemplos matemáticos de un tipo más general de expresiones, que reúnen las características señaladas en el párrafo anterior y que son el objeto de estudio de la lógica proposicional. Introducimos pues la siguiente definición:

> Cualquier sentencia u oración del lenguaje de la que tenga sentido decir que es verdadera o falsa se dirá una *proposición* .

Cuando trabajemos con proposiciones abstractas las notaremos con letras del tipo p, q, r, etc., y dada una proposición p, las notaciones $v(p) = \text{V}$ y $v(p) = \text{F}$ indicarán respectivamente que p es verdadera o falsa. En cualquier

caso, $v(p)$ se llama el *valor de verdad* de p. Por ejemplo, el lector informado advertirá que las dos primeras proposiciones de nuestros ejemplos matemáticos son verdaderas, mientras que la última es falsa.

NOTA Hemos brindado una definición intuitiva de proposición, ya que para contar con una definición más técnica debiéramos precisar algunos términos, como por ejemplo qué entendemos por sentencia, qué significa verdadero o falso, etc. Nuestro plan es tomar a todos estos conceptos como primitivos, descontando que el bagaje cultural que trae el lector le permitirá asimilarlos sin grandes dificultades. Agreguemos además que el valor de verdad de una sentencia puede ser desconocido para una persona o incluso para todas las personas, lo cual no es obstáculo para que dicha sentencia sea una proposición. Un ejemplo matemático famoso de ello es la llamada última conjetura de Fermat, que fue planteada por este matemático francés alrededor de 1640. Más adelante hablaremos un poco de ella, pero digamos que luego de más de tres siglos de intentos, su validez fue demostrada por el matemático inglés Andrew Wiles en 1995. Dicha conjetura pasó entonces a ser una proposición verdadera (lo que en Matemática se llama un teorema), pero aún antes de eso ya era una proposición. Veamos ahora algunos ejemplos sencillos, extraídos en su mayor parte del lenguaje cotidiano.

Ejemplos 0.1.1 Las siguientes oraciones son proposiciones:

1) Los viernes sólo trabajo por la mañana.

2) Alberdi no fue presidente de la Argentina.

3) Hoy el día amaneció soleado pero ventoso.

4) Prepararé el examen junto con Andrea o Jorge.

5) Llueve, por lo tanto hoy iré a trabajar en auto.

6) Si $x > 8$ entonces $x > 5$. Recíprocamente, si $x > 5$ entonces $x > 8$.

   No son en cambio proposiciones las oraciones interrogativas, imperativas o exclamativas, como por ejemplo:

7) ¿ Cuál es tu nombre ?

8) Inscríbase hoy mismo.

9) ¡ Hola amigos ! $\quad \diamondsuit$

0.1.2. Conectivos lógicos

Con frecuencia, dentro de un enunciado proposicional podemos distinguir la existencia de otras proposiciones, relacionadas entre sí de diversas maneras. Así, tomando como referencia el ejemplo 0.1.1, la afirmación subyacente

en la proposición 2) consiste en negar la proposición "Alberdi fue presidente de la Argentina", mientras que la proposición 3) afirma simultáneamente que el día amaneció soleado y que el día amaneció ventoso, esto es, amalgama dos proposiciones. Notemos que la palabra "pero" es sobre todo un énfasis literario, ya que lógicamente vale por "y". Una situacion muy similar se produce en la proposición 4), con la diferencia de que el "y" se cambia por un "o". Señalemos finalmente que 5) indica una relación de implicancia entre dos proposiciones, a saber, "hoy llueve" y "hoy iré a trabajar en auto", mientras que en 6) se observa una doble implicancia.

En general, y tratando de utilizar siempre un lenguaje llano e informal, llamaremos *conectivos lógicos* a las diversas ligazones que se establecen entre proposiciones para crear otras.

Listaremos a continuación los principales conectivos. El primero de ellos actúa sobre una única proposición p, mientras que los restantes conectan dos proposiciones p y q para producir una tercera. Exhibiremos en cada caso algunos ejemplos y la correspondiente *tabla de verdad* del conectivo, que nos mostrará los valores de verdad que el mismo alcanza en función de los valores de verdad de las proposiciones intervinientes.

NEGACION Si p es una proposición, llamaremos *negación* de p a la proposición que consiste en negar lo que asevera p. La notaremos $\neg p$.

Por ejemplo, "Homero no fue el autor de La Odisea" es la negación de "Homero escribió La Odisea", mientras que si p designa la proposición "existe un número real menor que su cuadrado", entonces $\neg p$ es la proposición "todo número real es mayor o igual que su cuadrado". La tabla de verdad de la negación es evidente:

Tabla de verdad de la negación

p	$\neg p$
V	F
F	V

CONJUNCION Si p y q son proposiciones, la afirmación simultánea de lo que aseveran p y q se llama la *conjunción* de p y q. Está asociada con el conectivo linguístico *y* y la notaremos $p \wedge q$.

Por ejemplo, la proposición "Paula es inteligente y simpática", es la conjunción de las proposiciones "Paula es inteligente" y "Paula es simpática", mientras que como vimos en el ejemplo 0.1.1, la proposición 3) es la conjunción de las proposiciones "el día amaneció soleado" y "el día amaneció ventoso". Finalmente, si p es la proposición "f es continua" y q es la

proposición "f no es derivable", podemos enunciar $p \wedge q$ en la forma "f es continua aunque no derivable".

Nuestro sentido elemental de la lógica nos sugiere que $p \wedge q$ será verdadera sólo cuando p y q lo sean, por lo que su tabla de verdad es la siguiente:

Tabla de verdad de la conjunción

p	q	$p \wedge q$
V	V	V
V	F	F
F	V	F
F	F	F

DISYUNCION Si p y q son proposiciones, la proposición obtenida ligando sus respectivas aseveraciones mediante el conectivo linguístico *o* se llama la *disyunción* de p y q. La notaremos $p \vee q$.

Por ejemplo, la proposición "estudiaré Medicina o Biología" es la disyunción de las proposiciones "estudiaré Medicina" y "estudiaré Biología", y en el ejemplo 0.1.1 la proposición 4) es la disyunción de las proposiciones "prepararé el examen con Andrea" y "prepararé el examen con Jorge". Como ejemplo matemático citemos la disyunción "un número natural n es par o el resto de dividir n^2 por 4 es 1", que como veremos más adelante es una proposición verdadera.

Es razonable suponer que la disyunción es falsa sólo cuando las proposiciones que intervienen en ella lo son, por lo que la tabla de verdad de $p \vee q$ es la siguiente:

Tabla de verdad de la disyunción

p	q	$p \vee q$
V	V	V
V	F	V
F	V	V
F	F	F

CONDICIONAL Dadas proposiciones p y q, la proposición "si p entonces q" se llama el *condicional* de antecedente p y consecuente q. Terminologías alternativas para designarla son p *implica* q, p *es condición suficiente para* q ó q *es condición necesaria para* p. La notaremos $p \Rightarrow q$.

Por ejemplo, la proposición 5) del ejemplo anterior es el condicional de antecedente "hoy llueve" y consecuente "hoy iré a trabajar en auto", y la proposición "me siento bien cuando escucho música" es el condicional de antecedente "escucho música" y consecuente "me siento bien". Observe el lector, a través de estos simples ejemplos, cómo el lenguaje cotidiano se vale de distintas formas sintácticas para verbalizar proposiciones que tienen la misma estructura lógica.

Como hecho que nos interesa especialmente, señalemos también que una gran mayoría de los enunciados de los teoremas matemáticos son del tipo p implica q, como por ejemplo el teorema de Pitágoras: "si T es un triángulo rectángulo entonces el cuadrado de la longitud de la hipotenusa de T es igual a la suma de los cuadrados de las longitudes de sus catetos".

Si bien el condicional parece establecer una relación de causalidad o implicancia entre dos proposiciones, lo cierto es que el mismo está definido cualesquiera sean las proposiciones p y q. Por ejemplo, la sentencia (francamente disparatada) "hoy es lunes, luego me llamo Pablo" es lógicamente hablando una proposición condicional, y por lo tanto debe asignársele un valor de verdad. Es necesario entonces construir en general la tabla de verdad del condicional.

En el diseño de la misma, y usando un lenguaje profano dictado por la intuición, sólo declararemos falsa la proposición $p \Rightarrow q$ cuando *la inferencia sea incorrecta*, esto es, cuando de un antecedente verdadero p se concluya una consecuencia falsa q. Formalmente, la tabla del condicional es entonces la siguiente:

Tabla de verdad del condicional

p	q	$p \Rightarrow q$
V	V	V
V	F	F
F	V	V
F	F	V

BICONDICIONAL Dadas proposiciones p y q, la sentencia p *si y solo si* q, que representa la conjunción de los condicionales $p \Rightarrow q$ y $q \Rightarrow p$, es una proposición llamada el *bicondicional* de componentes p y q. También se la denomina doble implicación entre p y q, o se dice alternativamente que p *es condición necesaria y suficiente para* q. La notaremos $p \Leftrightarrow q$.

La expresión *si y solo si* es de carácter técnico, de uso permanente en la Matemática, aunque no se la emplea normalmente en el lenguaje cotidiano, donde se recurre a giros gramaticales que la sustituyen. Por ejemplo, la oración "me detendré a descansar si el viaje supera los 200 km, en caso

contrario no lo haré", es, desde el punto de vista de la lógica, el bicondicional de la proposiciones "me detendré a descansar" y "el viaje supera los 200 km".

Mucho más natural y sencillo es brindar ejemplos matemáticos, como los bicondicionales "un número natural es impar si y solo si su cuadrado es impar" o "una función es inyectiva si y solo si es creciente".

Dado que el bicondicional significa la conjunción de las implicaciones $p \Rightarrow q$ y $q \Rightarrow p$, sigue fácilmente que el mismo será verdadero sólo cuando p y q tengan el mismo valor de verdad. En consecuencia, su tabla de verdad es la siguiente:

Tabla de verdad del bicondicional

p	q	$p \Leftrightarrow q$
V	V	V
V	F	F
F	V	F
F	F	V

0.1.3. Fórmulas proposicionales

Una *fórmula proposicional* $\mathcal{F}$ es una secuencia de letras p, q, r, etc., (que llamaremos las variables de la fórmula), conectivos lógicos y paréntesis. Suele a veces notársela en la forma $\mathcal{F}(p, q, r, \ldots)$, y los símbolos deben estar dispuestos en la secuencia de tal manera que $\mathcal{F}(p, q, r, \ldots)$ resulte ser una proposición cuando las variables proposicionales son sustituidas por proposiciones. Por ejemplo, las secuencias $p \vee qq$ y $p \wedge (q \Rightarrow\Rightarrow r)$ no son fórmulas proposicionales. En cuanto a los paréntesis, los mismos cumplen la función de evitar posibles ambigüedades. Así, la fórmula $p \Rightarrow q \wedge r$ podría interpretarse de dos maneras, a saber: $(p \Rightarrow q) \wedge r$ ó $p \Rightarrow (q \wedge r)$. Veamos algunos ejemplos.

Ejemplos 0.1.2 Las siguientes son fórmulas proposicionales:

1) $p \wedge (q \vee r)$.

2) p.

3) $(p \Rightarrow q) \Leftrightarrow ((\neg p) \vee q)$.

4) $(q \vee (\neg p)) \Rightarrow (q \wedge s)$.

5) $p \wedge (\neg p)$. $\quad \Diamond$

Ejemplos 0.1.3 Recordemos, en el segundo de los siguientes ítems, que dados dos números enteros b y m, decimos que b es múltiplo de m (ó que b es divisible por m), si y solo si existe un número entero k tal que $b = mk$. Los múltiplos de 2 se llaman números pares.

1) La estructura lógica de la proposición

> "si el tiempo es bueno y el auto funciona bien llegaré antes de las 10, y en tal caso asistiré a la cena"

está dada por la fórmula proposicional

$$((p \wedge q) \Rightarrow r) \wedge (r \Rightarrow s)\,,$$

a través de las sustituciones $p \to$ "el tiempo es bueno", $q \to$ "el auto funciona bien", $r \to$ "llegaré antes de las 10" y $s \to$ "asistiré a la cena".

2) El enunciado aritmético

> "el cuadrado de un número natural n es múltiplo de 4 ó su resto al dividirlo por 8 es 1"

responde a la fórmula proposicional

$$p \Rightarrow (q \vee r),$$

bajo las sustituciones $p \to$ "n es un número natural", $q \to$ "n^2 es múltiplo de 4" y $r \to$ "el resto de dividir n^2 por 8 es 1". $\Diamond$

TAUTOLOGIAS Una fórmula proposicional $\mathcal{F}$ se dice una *tautología* si la misma es verdadera cualesquiera sean los valores de verdad asignados a sus variables. Por ejemplo, son tautologías las fórmulas proposicionales $p \Leftrightarrow (\neg(\neg p))$, $p \vee (\neg p)$ y

$$(p \Rightarrow q) \Leftrightarrow ((\neg q) \Rightarrow (\neg p)).$$

Ello es obvio en los dos primeros casos. Respecto de la tercera fórmula, que designaremos por $\mathcal{F}$, probamos nuestra afirmación examinando su tabla de verdad:

p	q	$p \Rightarrow q$	$(\neg q) \Rightarrow (\neg p)$	$\mathcal{F}$
V	V	V	V	V
V	F	F	F	V
F	V	V	V	V
F	F	V	V	V

CONTRADICCIONES Una fórmula proposicional $\mathcal{F}$ se dice una *contradicción* si la misma es falsa cualesquiera sean los valores de verdad asignados a sus

variables. Por ejemplo, la fórmula $p \wedge (\neg p)$ es una contradicción. Es claro que $\mathcal{F}$ es una contradicción si y solo si $\neg\mathcal{F}$ es una tautología.

En general, una fórmula proposicional no es ni una tautología ni una contradicción. En tal caso se dice que es *contingente*. Por ejemplo, la fórmula $\mathcal{F} = p \vee (q \wedge r)$ es verdadera o falsa según los valores de verdad que se asignen a sus variables, como vemos en los casos

p	q	r	$\mathcal{F}$
V	F	V	V
F	V	F	F

EQUIVALENCIAS Dos fórmulas proposicionales $\mathcal{F}_1$ y $\mathcal{F}_2$ se dicen *lógicamente equivalentes*, o simplemente equivalentes, si ambas tienen el mismo valor de verdad cualesquiera sean los valores de verdad asignados a sus variables. Notaremos la situación en la forma $\mathcal{F}_1 \approx \mathcal{F}_2$.

Por ejemplo, valen las equivalencias $\neg(\neg p) \approx p$ y $p \Rightarrow q \approx (\neg p) \vee q$, como el lector puede verificar sin dificultad. Como acotación importante, observemos que la equivalencia admite la siguiente definición alternativa:

$$\mathcal{F}_1 \approx \mathcal{F}_2 \text{ si y solo si } \mathcal{F}_1 \Leftrightarrow \mathcal{F}_2 \text{ es una tautología.}$$

Ejemplos 0.1.4 La siguiente es una lista de tautologías y equivalencias lógicas. Desde ya que la misma está lejos de ser exhaustiva, simplemente queremos mostrar al lector algunas herramientas lógicas corrientemente empleadas en las demostraciones y en el quehacer matemático en general.

1) $(p \Rightarrow q) \wedge (q \Rightarrow p) \approx p \Leftrightarrow q$

2) $p \Rightarrow (p \vee q)$

3) $(p \wedge q) \Rightarrow p$

4) $\neg(p \vee q) \approx (\neg p) \wedge (\neg q)$

5) $\neg(p \wedge q) \approx (\neg p) \vee (\neg q)$

6) $((p \Rightarrow r) \wedge (q \Rightarrow r)) \Rightarrow ((p \vee q) \Rightarrow r)$

7) $p \wedge (q \vee r) \approx (p \wedge q) \vee (p \wedge r)$

8) $p \vee (q \wedge r) \approx (p \vee q) \wedge (p \vee r)$

9) $((p \Rightarrow q) \wedge p) \Rightarrow q$

10) $p \Rightarrow q \approx (\neg q) \Rightarrow (\neg p)$

11) $p \Rightarrow q \approx (p \wedge (\neg q)) \Rightarrow (r \wedge (\neg r))$

12) $p \vee q \approx (\neg p) \Rightarrow q$.

Encargamos al lector la tarea de demostrar, a través del uso de las correspondientes tablas de verdad, el carácter tautológico de las fórmulas 2), 3), 6) y 9) y la validez de las restantes equivalencias. $\Diamond$

0.1.4. Razonamientos

El verbo razonar es de uso frecuente en el habla cotidiana, junto con otros términos derivados del mismo, como razón, razonamiento, racional, etc., aludiendo al empleo de líneas de pensamiento y argumentación regidas por las leyes de la lógica. Reconociendo la vaguedad de esta descripción, veamos cómo precisar el concepto de razonamiento en el marco de la lógica proposicional.

Un *razonamiento* consiste de una secuencia $\mathcal{F}_1, \mathcal{F}_2 \ldots, \mathcal{F}_n$ de fórmulas proposicionales, llamadas las *premisas*, y una fórmula proposicional $\mathcal{F}$ llamada la *conclusión*. Un razonamiento que contiene sólo dos premisas se dice un *silogismo*. Usualmente, para sugerir la idea de que la conclusión $\mathcal{F}$ se deduce de las premisas $\mathcal{F}_i$, los razonamientos se representan mediante un esquema del tipo

$$\begin{array}{l} \mathcal{F}_1 \\ \mathcal{F}_2 \\ \vdots \\ \mathcal{F}_n \\ \hline \mathcal{F}\,. \end{array}$$

Ahora bien, a lo largo de nuestra experiencia percibimos a veces que se nos ofrecen argumentos correctamente razonados, mientras que en otras ocasiones nos parece advertir fallas en algún razonamiento. Puesto que nuestras impresiones al respecto son básicamente intuitivas y no está muy claro qué significa "deducir", es necesario que formalicemos la noción de *validez* de un razonamiento, de manera que el lector disponga de un método sistemático que le permita distinguir el pensamiento lógicamente correcto del incorrecto. Las dos siguientes definiciones establecerán dicho concepto.

Diremos que un razonamiento de premisas $\mathcal{F}_1, \mathcal{F}_2 \ldots, \mathcal{F}_n$ y conclusión $\mathcal{F}$ es *válido* si y solo si, dada cualquier asignación de valores de verdad a las variables proposicionales respecto de la cual todas las $\mathcal{F}_i$ son verdaderas, resulta que $\mathcal{F}$ también es verdadera. Equivalentemente, la fórmula

$$(\mathcal{F}_1 \wedge \mathcal{F}_1 \wedge \ldots \wedge \mathcal{F}_n) \Rightarrow \mathcal{F}$$

es una tautología.

Conservando las notaciones de la definición anterior, y por simple negación de la misma, diremos que el razonamiento es *inválido* si y solo si existe una asignación de valores de verdad a las variables proposicionales respecto de la cual todas las $\mathcal{F}_i$ son verdaderas y la conclusión $\mathcal{F}$ es falsa.

Ejemplos 0.1.5 El razonamiento

$$\begin{array}{l} p \wedge q \\ q \Rightarrow (\neg r) \\ r \vee t \\ \hline t \end{array}$$

es válido, ya que $\upsilon(p) = \text{V}$, $\upsilon(q) = \text{V}$, $\upsilon(r) = \text{F}$ y $\upsilon(t) = \text{V}$ es la única asignación de valores de verdad a las variables que hacen verdaderas todas las premisas, resultando en particular que la conclusión t es verdadera.

En cambio, el razonamiento

$$\begin{array}{l} p \Rightarrow q \\ q \\ \hline p \end{array}$$

es inválido, ya que bajo la asignación $\upsilon(p) = \text{F}$ y $\upsilon(q) = \text{V}$ las premisas resultan verdaderas y la conclusión falsa.

Vale la pena comentar este último caso, ya que es la forma simbólica de una falla lógica en la que se incurre con frecuencia. Una forma coloquial de este silogismo inválido podría ser por ejemplo la siguiente:

Premisa: "si estudio por lo menos 4 horas por día aprobaré el examen"

Premisa: "aprobé el examen"

Conclusión: "estudié por lo menos 4 horas por día" . $\Diamond$

0.1.5. Funciones proposicionales

Las proposiciones "Julia es amable", "el profesor Fuentes es amable" y "mi madre es amable" tienen evidentemente algo en común: todas ellas expresan la propiedad de ser amable, y podrían simbolizarse mediante la expresión

$$x \text{ es amable.}$$

Observemos que esta última sentencia no es una proposición, ya que no podemos asignarle algún valor de verdad, pero da lugar a una proposición en cuanto reemplazamos el símbolo x por el nombre de una persona.

Similarmente, las proposiciones "8 es múltiplo de -2" y "27 es múltiplo de 6" admiten la forma genérica

$$y \text{ es múltiplo de } x,$$

donde en este caso x e y representan números enteros.

Motivados por estos ejemplos, llamaremos *función proposicional* a todo enunciado simbólico $\mathrm{P}(x, y, z, \ldots)$ que resulte ser una proposición cada vez que las variables $x, y, z, \ldots$ son sustituidas por constantes pertenecientes a un cierto universo o conjunto de referencia U.

Por ejemplo, los enunciados

$$\mathrm{P}(x, y): \text{ si } x \text{ e } y \text{ son números naturales entonces } xy > y$$

y

$$\mathrm{Q}(x, y): \text{ si } x \text{ es padre de } y \text{ entonces } y \text{ es parecido a } x$$

son funciones proposicionales. En el primer caso el conjunto de referencia es obviamente el de los números naturales, y en el segundo podríamos tomar el conjunto de seres humanos como conjunto de referencia para ambas variables, o el conjunto de seres humanos varones para la variable x y el de seres humanos para la variable y.

En cambio, el enunciado

$$\mathrm{R}(x, y): \text{ si } x \text{ es padre de } y, \text{ ¿ } y \text{ es parecido a } x \text{ ?}$$

no es una función proposicional, ya que ninguna sustitución de las variables produce una proposición, dada la forma interrogativa de la sentencia.

CUANTIFICADOR EXISTENCIAL Dada una función proposicional, que por simplicidad supondremos dependiente de una sola variable, digamos $\mathrm{P}(x)$, el enunciado

$$\text{existe algún } x \text{ tal que } \mathrm{P}(x) \text{ es verdadera}$$

es una proposición, que afirma que $\mathrm{P}(c)$ es verdadera para al menos un valor c del conjunto de referencia U. Se lo llama el *cuantificador existencial* de $\mathrm{P}(x)$ y se lo nota en la forma

$$\exists\, x : \mathrm{P}(x)\,.$$

Ocasionalmente, si se desea enfatizar cuál es el conjunto de referencia, se emplea también la notación

$$\exists\, x \in U : \mathrm{P}(x)\,,$$

donde $x \in U$ indica que x es un elemento del conjunto U.

Notemos que el cuantificador existencial está emparentado con el conectivo lógico $\vee$, ya que asevera que entre una cierta lista de proposiciones alguna de ellas es verdadera. Es razonable entonces asignarle el siguiente valor de verdad:

$v\ (\exists\, x : \mathrm{P}(x)) = \mathrm{V}$ si y solo si $v\ (\mathrm{P}(c)) = \mathrm{V}$ para algún c perteneciente a U .

En forma completamente análoga se define el cuantificador existencial de una función proposicional de dos o más variables.

Ejemplos 0.1.6 Ilustremos convenientemente la definición anterior:

1) El cuantificador existencial

$$\exists\, x : 1 < x < 2$$

ligado a la función proposicional $\mathrm{P}(x) : (1 < x) \wedge (x < 2)$, es verdadero si se toma como universo el conjunto de números racionales, ya que por ejemplo $\mathrm{P}(3/2)$ es verdadera, pero es falso si el conjunto referencial se restringe a los números naturales, ya que entre 1 y 2 no existe ningún número natural. Este caso nos muestra entonces que el valor de verdad de un cuantificador depende del conjunto de referencia.

2) Como ya sabemos, el lenguaje cotidiano emplea diversos giros gramaticales para referirse a proposiciones que responden a la misma estructura lógica. Por ejemplo, las oraciones "algunos miembros del grupo son impuntuales", "no todos los miembros del grupo son puntuales" y "no falta algún impuntual en el grupo" son tres formas posibles de enunciar el cuantificador existencial

$$\exists\, x : x \text{ es impuntual},$$

donde x varía sobre un cierto grupo de personas.

Naturalmente (y por suerte), no empleamos al hablar este lenguaje de robots, pero sobre todo cuando se trate de cuestiones matemáticas, el lector debe saber apreciar que a veces enunciados aparentemente distintos expresan esencialmente lo mismo. De ahí la importancia de aprender a manejar el lenguaje simbólico, que brinda una mayor precisión sobre aquello que afirmamos. Por ejemplo, es más conciso y claro escribir

$$\exists\, x : (x \text{ es primo}) \wedge (x^2 + 19 \text{ es primo})$$

que decir "el cuadrado de algunos números primos sumado a 19 da como resultado un número primo". De paso, y tomando como conjunto de referencia el de los números naturales, el lector verificará que se trata de una proposición verdadera.

3) El cuantificador existencial en tres variables

$$\exists\, x, y, z : x - (y - z) = (x - y) - z$$

es verdadero tomando como rango de variación de las variables el conjunto de números reales, ya que, por ejemplo, la afirmación es verdadera para $x = 7$, $y = 2$ y $z = 0$. En realidad, no es difícil verificar que vale la igualdad que señala la función proposicional si y solo si la terna de variables es de la forma $(a, b, 0)$, donde a y b son números reales cualesquiera. Se deduce de ello que la proposición es falsa si se toma a los números naturales como conjunto de referencia. $\Diamond$

CUANTIFICADOR UNIVERSAL Dada una función proposicional $\mathrm{P}(x)$, el enunciado

$$\mathrm{P}(x) \text{ es verdadera para todo } x$$

es una proposición, que afirma que $\mathrm{P}(c)$ es verdadera para todos los valores c del conjunto de referencia U. Se lo llama el *cuantificador universal* de $\mathrm{P}(x)$ y se lo nota en la forma

$$\forall\, x : \mathrm{P}(x)\,.$$

Con frecuencia, en vez de "para todo x" se usa la locución alternativa "cualquiera sea x", y como en el caso del cuantificador existencial, también se usa la notación

$$\forall\, x \in U : \mathrm{P}(x)\,.$$

De manera muy similar se define el cuantificador universal de una función proposicional de 2 o más variables.

El cuantificador universal está asociado con el conectivo lógico $\wedge$, ya que asegura que todas las proposiciones que componen una cierta lista son verdaderas. Es natural entonces asignarle al mismo el siguiente valor de verdad:

$\upsilon\ (\forall\, x : \mathrm{P}(x)) = \mathrm{V}$ si y solo si $\upsilon\ (\mathrm{P}(c)) = \mathrm{V}$ cualquiera sea c perteneciente a U.

Ejemplos 0.1.7 Veamos cómo trabajar con el cuantificador universal:

1) La proposición "cualquier secuencia de 7 letras consecutivas del alfabeto contiene al menos una vocal" es un cuantificador universal, cuya forma simbólica es

   $$\forall\, s \in U : s \text{ contiene alguna vocal}\,,$$

   donde U es el conjunto de todas las secuencias de 7 letras consecutivas del alfabeto. Para mostrar que la proposición es verdadera (efectivamente lo es), podemos proceder exhaustivamente efectuando 21 verificaciones, ya que existen 21 secuencias de tal tipo. De todos modos, la tarea puede facilitarse escribiendo en orden las 27 letras del alfabeto y estudiando cuál es la mayor distancia entre dos vocales consecutivas.

2) Tomando como universo el conjunto de números enteros, el cuantificador universal

   $$\forall\, x : \text{si } x \text{ es múltiplo de 4 entonces } x \text{ es par}$$

   es verdadero, mientras que el cuantificador

   $$\forall\, x : \text{si } x \text{ es par entonces } x \text{ es múltiplo de 4}$$

   es falso.

Respecto a la segunda cuestión, bastará probar que al menos una de las infinitas proposiciones involucradas es falsa, objetivo que logramos tomando por ejemplo $x = 6$, que es par pero no múltiplo de 4. En cuanto a la primera afirmación, esta vez debemos mostrar la veracidad de un número infinito de proposiciones, lo que nos obliga a proceder en general. Puesto que cada una de las proposiciones es un condicional, bastará probar en cada caso que el consecuente es verdadero si el antecedente lo es, tarea bastante sencilla, ya que escribiendo $x = 4t$ (t entero) resulta que $x = 2(2t)$, y por lo tanto x es par.

3) Similarmente a otras situaciones, los enunciados de tipo universal aparecen en el lenguaje corriente bajo diversas formas gramaticales. Por ejemplo, las tres proposiciones "todos los árboles pierden su follaje en invierno", "ninguna autopista tiene cruces a nivel" y "cualquier integrante del coro puede reemplazar al director si es necesario" son cuantificadores universales. Si bien no nos importa mucho dilucidar aquí sus valores de verdad, digamos que el primero es falso y el segundo es verdadero, mientras que el último, bueno, depende del coro, aunque parece poco probable que sea verdadero. $\Diamond$

NEGACION DE UN CUANTIFICADOR Informalmente hablando, la negación de un cuantificador también es un cuantificador. Para convencernos de tal afirmación podemos argumentar de la siguiente manera: el cuantificador existencial ligado a una función proposicional $\mathrm{P}(x)$ es falso si las proposiciones $\mathrm{P}(c)$ son falsas *para todos* los c en el conjunto referencial U, mientras que el cuantificador universal es falso si $\mathrm{P}(c)$ es falso *para algún* c en U. Vale decir, la negación de un cuantificador existencial se asocia a una afirmación de tipo universal y la negación de un cuantificador universal a una afirmación de tipo existencial.

Empleando un lenguaje más riguroso, resulta que son válidas las siguientes equivalencias lógicas:

$$\neg\,(\exists\, x : \mathrm{P}(x)) \;\approx\; \forall\, x : \neg\,\mathrm{P}(x)$$

y

$$\neg\,(\forall\, x : \mathrm{P}(x)) \;\approx\; \exists\, x : \neg\,\mathrm{P}(x)\,.$$

Naturalmente, $\neg\,\mathrm{P}(x)$ indica la función proposicional que asigna la negación de $\mathrm{P}(c)$ a todo c en U. Dejamos a cargo del lector la verificación de ambas equivalencias, y le sugerimos que las confronte con las equivalencias 4) y 5) del ejemplo 0.1.4.

Por ejemplo, la proposición "ningún profesor de este colegio es exigente" (cuantificador universal) es la negación de la proposición "algunos profesores de este colegio son exigentes" (cuantificador existencial), y la proposición "todas las rutas son inseguras" se niega en la forma "existe una ruta segura".

Aclaremos por si acaso que una manera equivocada de negarla sería decir "ninguna ruta es insegura", que es en realidad la negación de "algunas rutas son inseguras".

Como ejemplo matemático, consideremos la siguiente proposición, doblemente cuantificada:

$$\exists\, a : \forall\, b : a \text{ es múltiplo de } b\,,$$

donde a y b varían sobre el conjunto de números naturales. Se trata de una proposición (falsa) que afirma que un cierto número natural es múltiplo de todos los números naturales. De acuerdo con las reglas de negación de los cuantificadores, su negación es equivalente a la proposición

$$\forall\, a : \exists\, b : a \text{ no es múltiplo de } b\,,$$

que expresada en una forma más coloquial, afirma que no existe un número natural divisible por todos los números naturales.

0.2. Lógica Matemática

0.2.1. El método matemático

A diferencia de otras disciplinas, como por ejemplo las ciencias de la Naturaleza, la Matemática es exclusivamente una creación del hombre, y manipula objetos ideales (números, rectas, funciones, etc.) que en rigor sólo existen en la mente de los seres humanos. La distingue también de otros saberes la característica de que sus verdades son irrefutables, siempre y cuando éstas hayan sido obtenidas por medio de razonamientos matemáticos válidos, esto es, mediante la correcta aplicación de la leyes de la lógica a las cuestiones matemáticas.

De todos modos, las palabras anteriores no deben inducir al lector a pensar que la Matemática es una especie de juego o pasatiempo intelectual, ya que a través de los tiempos su incesante crecimiento siempre estuvo motivado por la necesidad de resolver situaciones y problemas concretos del quehacer humano, desde la introducción del concepto de número y las diversas formas de operar con ellos, desde el estudio de las formas geométricas aplicado a las construcciones, hasta su ineludible y vasta utilización actual en la Física, la Informática, la Economía, etc.

Quizás el título que lleva esta sección no es el más adecuado, ya que parecería sugerir que la Matemática utiliza un tipo especial de lógica. Ello no es así, la lógica del razonamiento matemático es exactamente la lógica proposicional que hemos expuesto en la primera sección, y desde ese punto de vista hubiera sido mejor titularla "La Lógica en la Matemática". Aclarada esta cuestión, sin duda no muy importante, dedicaremos el resto del capítulo a mostrarle al lector, sobre todo a través de ejemplos, la interacción entre ambas disciplinas, con el propósito fundamental de familiarizarlo con el lenguaje y el pensamiento matemático. Comenzaremos introduciendo algunos términos.

Axiomas y Teoremas La aplicación de la Lógica a las teorías matemáticas, como por ejemplo la Aritmética o la Geometría, se manifiesta a veces mediante el establecimiento de lo que se denomina un sistema axiomático. Se eligen como punto de partida ciertos enunciados de la teoría, a los que se da el nombre de *axiomas* o *postulados*, y a partir de ellos se generan otros enunciados, llamados *teoremas*, mostrando que son lógicamente implicados por los axiomas.

Empleando un lenguaje llano y simplificador, digamos que en la construcción de una teoría matemática se fijan primero las reglas del juego (los axiomas), y se la desarrolla luego generando proposiciones verdaderas (los teoremas) aplicando las leyes de la lógica. Son ejemplos de ello la geometría euclidiana, cuyo vasto cuerpo de conocimientos se edifica a partir de solo cinco axiomas (los famosos postulados de Euclides), y la teoría de números, que se inicia con la construcción axiomática de los números naturales mediante los postulados de Peano.

Con frecuencia se emplean términos sucedáneos al de teorema para designar enunciados matemáticos verdaderos, como proposición, lema, corolario, etc. Se los usa por costumbre, de acuerdo con su índole y su importancia relativa dentro de la teoría, pero aclaremos que técnicamente hablando todos ellos son teoremas, incluidos los axiomas, en cuanto también son enunciados verdaderos.

DEMOSTRACIONES Cada uno de los resultados (teoremas) obtenidos en el desarrollo de la Matemática requiere una *demostración* o *prueba*, entendiéndose por ello algún tipo de procedimiento que mediante la correcta aplicación de las leyes de la lógica y el uso de los axiomas y otros resultados ya conocidos asegure que el enunciado en cuestión es verdadero.

Para simplificar la discusión, digamos que la mayoría de los enunciados a demostrar son del tipo

$$p \Rightarrow q\,,$$

donde naturalmente p y q también son enunciados matemáticos. Suele decirse que p es la *hipótesis* del teorema y que q es la *tesis* del mismo, aunque en realidad nuestras hipótesis (todo aquello que podemos suponer verdadero) son mucho más amplias, ya que forman parte de ellas los axiomas y todo enunciado previamente demostrado.

Aclaremos de paso que cuando probamos un enunciado de tipo condicional *solo demostramos que la implicación es verdadera*, sin que tal hecho nos permita extraer conclusiones acerca de los valores de verdad de las premisas p y q. Por ejemplo, cuando postulamos (y demostramos) que si f es derivable en x_0 entonces f es continua en x_0, no afirmamos ni la derivabilidad ni la continuidad de f en x_0, simplemente establecemos una relación de implicancia entre ambos hechos

Por supuesto, lo anterior es solo un ejemplo aclaratorio, y en general, el rol de la Matemática es "deducir" teoremas a partir de otros teoremas. Respecto de ello, vale la pena traer a colación la tautología o *regla de inferencia* 9) del ejemplo 0.1.4:

$$((p \Rightarrow q) \wedge p) \Rightarrow q\,,$$

llamada *modus ponens*, que asegura que si $p \Rightarrow q$ es verdadera y p es un teorema entonces q es un teorema.

Exhibiremos en lo que sigue una breve lista (desde ya que no exhaustiva) de tipos de enunciado y posibles métodos de demostración. Iremos ilustrándola con ejemplos sencillos, y en el desarrollo de los mismos usaremos algunas propiedades matemáticas elementales, posiblemente conocidas por el lector, asumiendo que ellas fueron previamente demostradas a partir de algún sistema de axiomas.

No está demás aclarar que una prueba matemática puede ser muy simple o también extraordinariamente compleja, casos éstos en los que su obtención requiere mucho más que el buen empleo de la lógica. Expresado en un

lenguaje que acompañe al tema, digamos que razonar correctamente es una condición necesaria, aunque no suficiente, para demostrar un teorema.

I) METODO DIRECTO Para probar que $p \Rightarrow q$ es verdadero, suponemos que p es verdadero y mediante algún procedimiento válido mostramos que q también lo es. Observemos que ello asegura la verdad de la implicación, ya que de acuerdo con su tabla de verdad, el condicional sólo es falso cuando el antecedente es verdadero y el consecuente es falso. Por ejemplo, si a y b son números naturales, probemos por este método el enunciado

$$\text{Si } b \text{ es múltiplo de } a \text{ entonces } b^2 \text{ es múltiplo de } a^2\,.$$

Asumiendo que el antecedente es verdadero, podemos escribir $b = ac$, donde c es un número natural. Luego, elevando ambos de esta igualdad al cuadrado, obtenemos

$$b^2 = (ac)^2 = a^2c^2\,,$$

de donde se concluye que b^2 es múltiplo de a^2.

II) METODO CONTRARECIPROCO Consiste en demostrar que el enunciado

$$p \Rightarrow q$$

es verdadero probando que el enunciado

$$(\neg q) \Rightarrow (\neg p)$$

lo es, ya que de acuerdo con la fórmula 10) del ejemplo 0.1.4 ambas fórmulas son equivalentes. Procediendo como en el método directo, se supone entonces que q es falso (esto es, $\neg\, q$ es verdadero) y se demuestra que p es falso (o sea, $\neg\, p$ es verdadero).

Por ejemplo, probemos por el método contrarecíproco la proposición

$$(n^2 \text{ par}) \Rightarrow (n \text{ par}),$$

donde n designa un número natural. Suponiendo que el consecuente es falso y por lo tanto n es impar, escribamos $n = 2k - 1$, donde k es un cierto número natural. Elevando al cuadrado, tenemos:

$$n^2 = (2k-1)^2 = 4k^2 - 4k + 1 = 2(2k^2 - 2k) + 1,$$

de donde resulta que n^2 es impar y por lo tanto p es falso.

III) DEMOSTRACION DE EQUIVALENCIAS De acuerdo con la fórmula 1) de 0.1.4, para demostrar la equivalencia de dos proposiciones p y q, esto es, que el bicondicional

$$p \Leftrightarrow q$$

es verdadero, basta demostrar que los condicionales $p \Rightarrow q$ y $q \Rightarrow p$ (llamado *recíproco* del anterior) son verdaderos. Podemos por supuesto hacerlo en cualquier orden, y en ambos casos podemos aplicar los métodos descriptos en los ítems I) y II).

Por ejemplo, dado un número real positivo x, probemos la equivalencia

$$x^2 > x \Leftrightarrow x > 1\,.$$

La implicación $x > 1 \Rightarrow x^2 > x$ se obtiene por el método directo, suponiendo $x > 1$ y multiplicando ambos miembros por x (al ser x positivo no cambia el sentido de la desigualdad).

Respecto de la recíproca, siguiendo con el método directo asumamos la validez de la desigualdad $x^2 > x$. Puesto que

$$x(x-1) = x^2 - x > 0$$

y $x > 0$, concluimos que $x - 1 > 0$, vale decir, $x > 1$, como queríamos demostrar.

Observemos que empleando la validez de los correspondientes contrarecíprocos queda también probada la equivalencia

$$x^2 \leq x \Leftrightarrow x \leq 1\,,$$

suponiendo siempre que x es positivo.

IV) DEMOSTRACION POR DISCUSION DE CASOS Supóngase que queremos demostrar que una proposición r es verdadera y que la hipótesis comprende dos casos posibles: en uno de ellos es verdadera una cierta proposición p y en el otro una cierta proposición q. Bastará demostrar entonces la validez de las implicaciones $p \Rightarrow r$ y $q \Rightarrow r$, ya que la fórmula

$$((p \Rightarrow r) \wedge (q \Rightarrow r)) \Rightarrow ((p \vee q) \Rightarrow r)$$

es una tautología.

Por ejemplo, probemos que si a es un número entero impar entonces $a^2 - 1$ es divisible por 8. Para ello, escribamos $a = 2k+1$ y distingamos los casos i) k par ($k = 2t$) y ii) k impar ($k = 2t + 1$).

En el caso i) resulta $a = 2(2t) + 1 = 4t + 1$, de donde:

$$a^2 - 1 = (4t+1)^2 - 1 = 16t^2 + 8t + 1 - 1 = 8(2t^2 + t),$$

y por lo tanto $a^2 - 1$ es divisible por 8.

En el caso ii) tendremos $a = 4t + 3$, y entonces

$$a^2 - 1 = (4t + 3)^2 - 1 = 16t^2 + 24t + 9 - 1 = \\ = 16t^2 + 24t + 8 = 8(2t^2 + 3t + 1)$$

también es un múltiplo de 8.

Aclaremos que por razones de simplicidad hemos supuesto que nuestra hipótesis comporta sólo dos casos, pero es evidente que el mismo método puede aplicarse a situaciones que requieran el análisis de un número mayor de casos.

V) DEMOSTRACIONES POR EL ABSURDO Asumiendo nuevamente que deseamos probar que una implicación $p \Rightarrow q$ es verdadera, el método de demostración *por el absurdo* consiste en suponer que la hipótesis p es verdadera y la tesis q es falsa, y mediante procedimientos válidos arribar entonces a una contradicción. Justifica el metodo la fórmula 11) del ejemplo 0.1.4, a saber, la equivalencia

$$(p \Rightarrow q) \approx ((p \wedge (\neg q)) \Rightarrow (r \wedge (\neg r))\,,$$

donde r es una cierta proposición.

Demás está decir que no hay una regla universal que nos indique en cualquier situación cómo obtener la deseada contradicción, sino que cada caso requiere el uso de alguna técnica adecuada y un correcto manejo y conocimiento de los temas específicos en los que estamos trabajando. A manera de ejemplo más o menos sencillo, probemos por el absurdo que no es posible expresar a 30 como diferencia de los cuadrados de dos números naturales. Dicho más formalmente, probemos que la proposición

$$\text{si } a \text{ y } b \text{ son números naturales entonces } a^2 - b^2 \neq 30$$

es verdadera.

Habiendo demostrado anteriormente que k^2 es par si y solo si k es par (k un número natural), y aceptando como válido que la suma o diferencia de dos números naturales es impar si y solo si uno de ellos es par y el otro es impar (el lector puede intentar demostrar esto), supongamos por el absurdo que existen numeros naturales a y b tales que $a^2 - b^2 = 30$.

De acuerdo con nuestras consideraciones previas, resulta entonces que a^2 y b^2, y por lo tanto a y b, son ambos pares o ambos impares, de donde deducimos que $a + b$ y $a - b$ son pares. Escribiendo en tal caso $a + b = 2m$ y $a - b = 2n$, sigue luego que

$$30 = a^2 - b^2 = (a + b)(a - b) = (2m)(2n) = 4mn,$$

lo que es una contradicción, ya que dividiendo ambos miembros por 2 resultaría que $15 = 2mn$ es par.

VI) PRUEBA DE UNA DISYUNCION Si queremos probar por el método directo una implicación del tipo $p \Rightarrow (q \vee r)$, bastará suponer que alguna de las proposiciones q ó r es falsa y demostrar entonces que la otra es verdadera, ya que una disyunción es falsa sólo cuando las dos proposiciones intervinientes lo son. Como ilustración, probemos la validez de la implicación

$$x^2 > 1 \Rightarrow (x > 1) \ \vee \ (x < -1)\,,$$

donde x designa un número real.

Suponiendo por ejemplo que $x < -1$ es falsa, deducimos de nuestra hipótesis y de las propiedades elementales del orden que $x + 1 > 0$. Puesto que $(x-1)(x+1) = x^2 - 1 > 0$ y un producto es positivo si y solo si sus dos factores tienen el mismo signo, concluimos que $x-1 > 0$, o equivalentemente $x > 1$, como queríamos demostrar.

NOTA Naturalmente, la prueba de una implicación en la que el consecuente es una conjunción, digamos $p \Rightarrow (q \wedge r)$, requiere otro procedimiento. El método habitual, basado en la equivalencia

$$p \Rightarrow (q \wedge r) \approx (p \Rightarrow q) \wedge (p \Rightarrow r) \tag{1}$$

(pedimos al lector que verifique este hecho), es demostrar sucesivamente que las implicaciones $p \Rightarrow q$ y $p \Rightarrow r$ son verdaderas. Por ejemplo, probemos que si a y b son números pares consecutivos entonces su suma no es múltiplo de 4 y su producto es múltiplo de 8.

De acuerdo con la hipótesis, podemos escribir $a = 2k$ y $b = 2k + 2$, donde k es un número natural. Para probar la primera afirmación, supongamos por el absurdo que $a + b$ es múltiplo de 4, digamos $4t$. Tenemos entonces que

$$4t = a + b = 2k + (2k + 2) = 4k + 2,$$

de donde $2 = 4t - 4k = 4(t - k)$, lo que es una contradicción, pues obviamente 2 no es múltiplo de 4.

Respecto de la validez de la segunda afirmación, operando obtenemos:

$$ab = 2k(2k + 2) = 4k^2 + 4k = 4(k^2 + k).$$

Puesto que ya hemos visto que k^2 y k son ambos pares o ambos impares, resulta en cualquier caso que $k^2 + k$ es par, de la forma $2s$. En consecuencia, $ab = 4(2s) = 8s$, según nos proponíamos demostrar.

Con esta breve reseña de los procedimientos matemáticos usuales cerramos este capítulo preparatorio. Confiamos en que nuestra exposición de los conceptos básicos de la Lógica y su conexión con la Matemática ayuden al lector a lidiar con las inevitables (pero no insalvables) complicaciones que hallará en el desarrollo ulterior del libro, ya que muchas de ellas están directamente relacionadas con la natural dificultad que enfrenta el principiante al tomar contacto con el lenguaje abstracto del Algebra. De todos modos, y confiando en que el lector vaya asimilando paulatinamente la lógica del pensamiento matemático que hemos expuesto en este capítulo, lo instamos también a no desdeñar nunca sus propias ideas, por más ingenuas que puedan parecerle, ya que el desarrollo matemático requiere tanto del formalismo como de la intuición.

0.2.2. Ejercicios

1. Indicar cuáles de las siguientes sentencias son proposiciones:

   a) El algoritmo no funciona.

   b) ¿ Quedó rica la torta ?

   c) $x^2 + 3x - 6 = 0$.

   d) Hoy es 12 de abril, por lo tanto mañana es sábado.

   e) Todos debieran aprender algo de lógica.

   f) La guerra de Troya nunca ocurrió.

   g) La realidad y los sueños.

   h) Hay poblaciones de bacterias en Marte.

2. Escribir en forma simbólica cada una de las siguientes proposiciones:

   a) Para convencerme de que me quiere es suficiente que acepte salir conmigo.

   b) Viajaré a Japón con Andrea y a Alemania con Gustavo.

   c) No comeré mariscos ni pastas, a menos que éstas sean caseras.

   d) Ladran Sancho, señal que cabalgamos.

   e) Tomaré un café sólo si tú me acompañas.

   f) Tomaré un café si tú me acompañas.

   g) Aprobaré si toman un examen fácil, pero no si toman uno difícil.

   h) Hoy repasaré una sola materia, Matemática o Química.

3. Dadas las proposiciones:

$$\begin{aligned} p:&\ \text{Estamos en verano}\\ q:&\ \text{Hoy es un día fresco}\\ r:&\ \text{Vamos a la playa,}\end{aligned}$$

enunciar en lenguaje corriente las siguientes proposiciones:

a) $p \wedge q$
b) $q \Rightarrow (\neg r)$
c) $p \Leftrightarrow r$
d) $\neg r \Rightarrow (q \vee (\neg p))$.

4. Suponiendo verdadero el enunciado

Hoy es 30 de noviembre, hace calor y vamos a la playa,

analizar los valores de verdad de las proposiciones *a*) a *d*) del ejercicio 3.

5. Probar la validez de todas las fórmulas del ejemplo 0.1.4.

6. Empleando equivalencias, demostrar que la conjunción, el condicional y el bicondicional pueden expresarse en términos de la negación y la disyunción.

7. Demostrar las siguientes inequivalencias:

a) $(p \Rightarrow q) \not\approx (q \Rightarrow p)$.
b) $(p \Rightarrow q) \not\approx ((\neg p) \Rightarrow (\neg q))$.
c) $\neg(p \vee q) \not\approx (\neg p) \vee (\neg q)$.
d) $\neg(p \wedge q) \not\approx (\neg p) \wedge (\neg q)$.
e) $\neg(p \Rightarrow q) \not\approx (\neg p) \Rightarrow (\neg q)$.
f) $\neg(p \Leftrightarrow q) \not\approx (\neg p) \Leftrightarrow (\neg q)$.
g) $((p \Rightarrow q) \Rightarrow r) \not\approx (p \Rightarrow (q \Rightarrow r))$.
h) $((p \Leftrightarrow q) \Leftrightarrow r) \not\approx (p \Leftrightarrow (q \Leftrightarrow r))$.

8. Si p y q son proposiciones, construir la tabla de verdad de

$$\mathcal{F} = (p \wedge (\neg q)) \vee ((\neg p) \wedge q),$$

y observar que $\mathcal{F}$ es verdadera si y solo si exactamente una de las proposiciones p y q es verdadera. Debido a ello se la llama la *disyunción excluyente* de p y q, y se la nota $p \veebar q$.

9. Demostrar la validez de los siguientes razonamientos (comparar con el ejemplo 0.1.4):

   a) (modus ponens)

   $$\begin{array}{l} p \Rightarrow q \\ \underline{p} \\ q \end{array}$$

   b) (modus tollens)

   $$\begin{array}{l} p \Rightarrow q \\ \underline{\neg q} \\ \neg p \end{array}$$

   c) (silogismo hipotético)

   $$\begin{array}{l} p \Rightarrow q \\ \underline{q \Rightarrow r} \\ p \Rightarrow r \end{array}$$

   d) (silogismo disyuntivo)

   $$\begin{array}{l} p \vee q \\ \underline{\neg p} \\ q \end{array}$$

10. Expresar simbólicamente los siguientes razonamientos, analizando en cada caso su validez:

    a) La venta de autos se incrementa sólo si la situación económica es buena o el precio de la nafta se mantiene estable. La nafta aumentó y se incrementó la venta de autos. Por lo tanto la situación económica es buena.

    b) Si aumentamos los impuestos bajará nuestra popularidad. Si baja nuestra popularidad debemos iniciar una campaña de captación de adherentes. Si iniciamos una campaña de captación de adherentes es porque aumentamos los impuestos.

    c) Para ser feliz es necesario poseer dinero y no tener problemas de salud. Por otro lado, sólo se posee dinero si se trabaja mucho o se lo hereda. Luego es suficiente heredar dinero para ser feliz.

    d) Si disponemos del coche y el pronóstico es bueno, el fin de semana iremos a la playa o a la montaña. No es cierto que si el pronóstico es bueno iremos a la playa el fin de semana. Por lo tanto, si no vamos a la montaña es porque no disponemos del coche.

11. Determinar el valor de verdad de cada una de las siguientes proposiciones (se toma como conjunto de referencia el de los números naturales):

    a) $\forall\, x : \forall\, y : x + y \leq xy$.

    b) $\forall\, x : \exists\, y : x^2 < y < (x+1)^2$.

    c) $\exists\, x : \forall\, y : x + y$ es múltiplo de x .

    d) $\exists\, x : \forall\, y : x + y$ es múltiplo de y .

    e) $\exists\, x : \exists\, y : x^2 + 1 = 4y + 3$.

12. Negar, empleando cuantificadores, cada una de las proposiciones del ejercicio anterior.

13. Empleando notación matemática expresar simbólicamente los siguientes enunciados. Analizar en cada caso su verdad o falsedad.

    a) Todo número real es mayor que su cubo.

    b) Existe un número natural no divisible por sí mismo.

    c) El cociente entre dos números reales positivos es menor que su producto.

    d) Dos rectas distintas del plano son paralelas o se cortan en un punto.

    e) Entre dos números reales cualesquiera existe otro número real.

    f) Existe un número real positivo mínimo.

    g) Existe un número natural mínimo.

    h) Existe un número natural máximo.

    i) 18 puede expresarse como diferencia de los cuadrados de dos números naturales.

Capítulo 1

Conjuntos, relaciones y funciones

1.1. Conjuntos

1.1.1. Definición intuitiva

Un *conjunto* es una colección de objetos, cada uno de los cuales se dice un *elemento* del mismo. Usualmente, designaremos los conjuntos mediante letras mayúsculas del tipo A, B, etc, reservándonos el uso de algunos caracteres especiales para representar ciertos conjuntos numéricos. Así, a lo largo de estas páginas los símbolos $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ y $\mathbb{R}$ denotarán los conjuntos de números naturales, enteros, racionales y reales, respectivamente. Vale la pena consignar que la $\mathbb{Z}$ se debe a *zhalen* (número, en alemán) y $\mathbb{Q}$ se emplea por *quotient* (cociente).

Si X es un conjunto y x es un elemento de X diremos que x *pertenece* a X, situación que simbolizaremos por $x \in X$, mientras que $x \notin X$ significará que x no pertenece a X.

NOTA La noción de conjunto que introdujimos arriba es puramente intuitiva y se apoya en dos conceptos *primitivos* (indefinibles). Uno es el concepto mismo de conjunto, término sinónimo de otros como "colección", "lista", etc, y otro es el concepto de pertenencia de un elemento a un conjunto. Ello sin embargo no debe preocuparnos, pues ya veremos cómo esas definiciones intuitivas nos permitirán desarrollar perfectamente los aspectos básicos de la teoria de conjuntos y manejar fluidamente su lenguaje, siendo la última una de las principales finalidades de esta sección.

Ejemplos 1.1.1 Se acostumbra a describir un conjunto listando entre llaves todos sus elementos o bien estableciendo alguna regla o propiedad que deben satisfacer sus elementos. Nos referimos a esquemas del tipo $\mathcal{E} = \{x : p(x)\}$, donde x toma valores en un cierto conjunto de referencia y $p(x)$ es una forma proposicional que será verdadera si y sólo si x se sustituye por un elemento

de $\mathcal{E}$. Consideremos por ejemplo los conjuntos

$$A = \{1, 2, \alpha, \beta, \{\gamma\}\}$$
$$B = \text{conjunto de números naturales impares} = \{1, 3, 5, 7, \ldots\}$$
$$C = \text{conjunto de nombres de provincias argentinas}$$
$$D = \{x \in A : x \text{ es un número}\} = \{1, 2\}.$$

Son válidas por ejemplo las relaciones $\beta \in A$, $9 \in B$, $8 \notin B$ y Salta $\in C$. Como puede apreciarse, los elementos de un conjunto pueden ser de muy diversa índole, e incluso pueden ser conjuntos, como ocurre con el elemento $\{\gamma\}$ de A. Vemos también las distintas formas de presentar un conjunto. Por caso, A se define exhibiendo la lista completa de sus elementos, y lo mismo podría hacerse con los de C. Es claro en cambio que tal descripción no es posible en el caso del conjunto B, mientras que en D se emplean las dos formas de representación mencionadas. $\Diamond$

CONJUNTO VACIO Volviendo a C del ejemplo 1.1.1, consideremos ahora el conjunto

$$\mathcal{H} = \{x \in C : x \text{ comienza con H}\}.$$

Una breve incursión por la geografía nos muestra que ninguna provincia argentina posee un nombre que comience con H, y por lo tanto $\mathcal{H}$ es un conjunto que no tiene ningún elemento.

Si bien el ejemplo parece (y es) algo rebuscado, las frecuentes y naturales apariciones en Matemática de situaciones como ésta aconsejan −para disponer de un lenguaje cómodo− darle un nombre a la misma. Precisamente, diremos que un conjunto X es *vacío* si y sólo si no tiene ningún elemento. Notaremos en tal caso $X = \emptyset$.

En términos de lógica, $X = \emptyset$ si y sólo si la proposición

$$x \in X$$

es falsa cualquiera sea x. Más adelante apreciaremos la utilidad de esta caracterización del conjunto vacío.

INCLUSION Si A y B son conjuntos, diremos que A está *incluido* ó *contenido* en B si y sólo si todo elemento de A pertenece a B. Diremos también en tal caso que A es un *subconjunto* de B, y emplearemos la notación $A \subseteq B$. En lenguaje simbólico, $A \subseteq B$ si y sólo si la proposición

$$x \in A \Rightarrow x \in B$$

es verdadera para todo x. Alternativamente, se dice también que B incluye ó contiene a A, en cuyo caso notamos $B \supseteq A$.

Por ejemplo, $\{a, c\} \subseteq \{a, b, c, w, 6\}$ y el conjunto B del ejemplo 1.1.1 es un subconjunto del conjunto de números naturales.

Si X es un conjunto, sigue directamente de la definición que $X \subseteq X$, esto es, todo conjunto es subconjunto de sí mismo. Como otra propiedad importante, señalemos también que $\emptyset \subseteq X$ cualquiera sea el conjunto X. Sugerimos al lector demostrar esto teniendo en cuenta las definiciones en términos proposicionales del conjunto vacío y de la inclusión.

IGUALDAD Diremos que dos conjuntos A y B son iguales si y sólo si $A \subseteq B$ y $B \subseteq A$. Escribiremos por supuesto $A = B$. Simbólicamente, $A = B$ si y sólo si la proposición

$$x \in A \Leftrightarrow x \in B$$

es verdadera para todo x.

CONJUNTO DE PARTES Si A es un conjunto, la colección de todos los subconjuntos de A es un nuevo conjunto, que llamaremos el *conjunto de partes* de A, y que notaremos $\mathbb{P}(A)$. O sea,

$$\mathbb{P}(A) = \{\, T \,:\, T \subseteq A \,\}\,.$$

Por ejemplo, sea $X = \{a, b, c\}$. Entonces

$$\mathbb{P}(X) = \{\emptyset, \{a\}\,, \{b\}\,, \{c\}\,, \{a, b\}\,, \{a, c\}\,, \{b, c\}\,, X\}\,,$$

como el lector puede verificar fácilmente. Volviendo a la situación general, llamaremos subconjunto *propio* de A a todo elemento S de $\mathbb{P}(A)$ distinto de A. Suele emplearse en tal caso la notación $S \subset A$.

CONJUNTO REFERENCIAL Si bien no es absolutamente necesario, es cómodo suponer a veces que los conjuntos a considerar son subconjuntos de un conjunto $\mathcal{U}$, al que denominamos conjunto *referencial* ó *universal*. Esto no debe llevar al lector a pensar que existe una especie de conjunto total que contiene a todos los conjuntos. Simplemente fijamos $\mathcal{U}$ y trabajamos con elementos de $\mathbb{P}(\mathcal{U})$. Por ejemplo, trabajando con conjuntos numéricos es razonable tomar como conjunto de referencia al conjunto de números reales, o bien al de números complejos.

DIAGRAMAS DE VENN En ocasiones, es útil representar conjuntos genéricos mediante sencillos esquemas geométricos, llamados *diagramas de Venn*. Si bien pueden parecer algo ingenuos y carecen de sentido matemático preciso, configuran un método de visualización que puede ayudar al lector a comprender y familiarizarse con las nociones −sin duda abstractas− de la teoría de conjuntos.

Por ejemplo, en las figuras que siguen a continuación ilustramos las nociones de inclusión, pertenencia y conjunto universal:

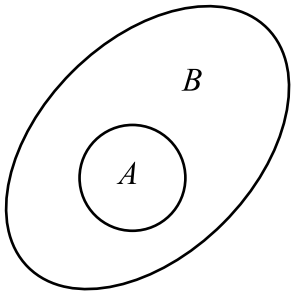


Figura 1.1: $A \subset B$

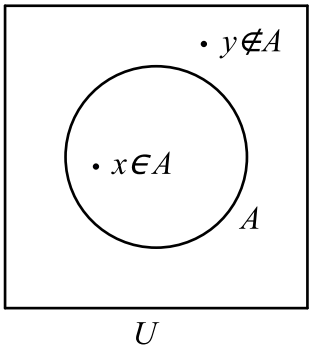


Figura 1.2: Pertenencia

1.1.2. Operaciones de conjuntos

Definiremos en lo que sigue las operaciones conjuntísticas elementales, estableciendo además sus propiedades básicas. Salvo la primera, que involucra un solo conjunto, las restantes son operaciones *binarias*, que a cada par de conjuntos le asocian otro (el resultado de la operación), formalmente similares a las operaciones numéricas, como la suma y el producto. Supondremos que todos los conjuntos en cuestión están contenidos en un conjunto de referencia $\mathcal{U}$ (en los ejemplos tomaremos $\mathcal{U} = \mathbb{N}$), y vizualizaremos cada una de las operaciones definidas a través del correspondiente diagrama de Venn.

COMPLEMENTO Si A es un conjunto, definimos su *complemento* en la forma

$$A^c = \{x \in \mathcal{U} \,:\, x \notin A\}\,.$$

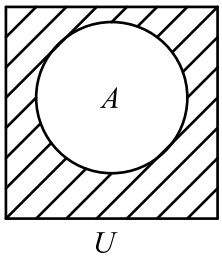


Figura 1.3: Complemento de A

Por ejemplo, si $A = \left\{n : n^2 \geq 15\right\}$ entonces $A^c = \{1, 2, 3\}$.

Propiedades del Complemento

i) $(X^c)^c = X$.

ii) $\mathcal{U}^c = \emptyset$ y $\emptyset^c = \mathcal{U}$.

iii) $X \subseteq Y \Leftrightarrow Y^c \subseteq X^c$.

DEMOSTRACION Probaremos iii), dejando las otras como ejercicio. Asumiendo primero que $X \subseteq Y$, resulta que

$$u \in Y^c \Rightarrow u \notin Y \Rightarrow u \notin X \Rightarrow u \in X^c,$$

y por lo tanto $Y^c \subseteq X^c$ (notemos que la proposición $u \notin Y \Rightarrow u \notin X$ es verdadera por ser contrarrecíproca de la proposición $u \in X \Rightarrow u \in Y$). Inversamente, supongamos que $Y^c \subseteq X^c$. Entonces, usando i) y la implicación que acabamos de demostrar obtenemos $X = (X^c)^c \subseteq (Y^c)^c = Y$. $\quad \Diamond$

INTERSECCION Si A y B son conjuntos, definimos su *intersección* en la forma

$$A \cap B = \{x \in \mathcal{U} \,:\, x \in A \ \text{ y } \ x \in B\}\,.$$

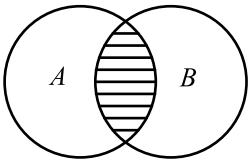


Figura 1.4: $A \cap B$

Por ejemplo, si $A = \{x < 20 : x \text{ es impar}\}$ y B el conjunto de múltiplos de 3 entonces $A \cap B = \{3, 9, 15\}$.

Propiedades de la Intersección

i) $(X \cap Y) \cap Z = X \cap (Y \cap Z)$ y $X \cap Y = Y \cap X$.

ii) $X \cap \mathcal{U} = X$ y $X \cap \emptyset = \emptyset$.

iii) $X \cap X = X$.

iv) $X \cap Y \subseteq X$ y $X \cap Y \subseteq Y$.

v) $X \cap X^c = \emptyset$. Más generalmente, $X \cap Y = \emptyset$ si y sólo si $Y \subseteq X^c$.

DEMOSTRACION Las demostraciones son inmediatas y quedan a cargo del lector, pero vale la pena hacer algunos comentarios. Las propiedades enunciadas en i) se llaman propiedad *asociativa* y propiedad *conmutativa*, respectivamente, mientras que debido al primer enunciado de ii) diremos que $\mathcal{U}$ es el elemento *neutro* de la operación intersección. Respecto de v), señalemos que dos conjuntos cuya intersección es vacía se dicen *disjuntos*. $\Diamond$

UNION Si A y B son conjuntos, definimos su *unión* en la forma

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

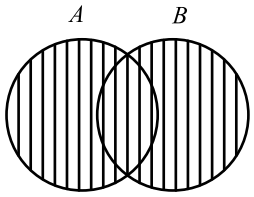


Figura 1.5: $A \cup B$

Por ejemplo, si $A = \{x : x \leq 8\}$ y B es el conjunto de múltiplos de 4 menores que 20 entonces

$$A \cup B = \{1, 2, 3, 4, 5, 6, 7, 8, 12, 16\}.$$

Propiedades de la Unión

i) $(X \cup Y) \cup Z = X \cup (Y \cup Z)$ y $X \cup Y = Y \cup X$.

ii) $X \cup \mathcal{U} = \mathcal{U}$ y $X \cup \emptyset = X$.

iii) $X \cup X = X$.

iv) $X \cup Y \supseteq X$ y $X \cup Y \supseteq Y$.

v) $X \cup X^c = \mathcal{U}$. Más generalmente, $X \cup Y = \mathcal{U}$ si y sólo si $Y \supseteq X^c$.

DEMOSTRACION Como antes, encargamos al lector la demostración de estos hechos. Se observa claramente cierto correlato ó *dualidad* con las correspondientes propiedades de la intersección. Valen por ejemplo las propiedades asociativa y conmutativa, siendo $\emptyset$ el elemento neutro de la unión. A diferencia de la intersección, el enunciado v) nos muestra que todo conjunto A admite un *inverso* respecto de la unión, a saber, su complemento A^c. $\Diamond$

Notemos la conexión existente entre las anteriores operaciones de conjunto y los conectivos lógicos de negación, conjunción y disyunción, relación que quedará aún más evidenciada en las siguientes propiedades, que vinculan entre sí la intersección y la unión. Las dos primeras son las propiedades distributivas de la intersección (la unión) con respecto a la unión (la intersección), mientras que las dos últimas son las leyes de De Morgan, donde veremos que cualquiera de estas dos operaciones puede obtenerse a partir de la otra y la complementación.

Propiedades distributivas y leyes de De Morgan

i) $X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$.

ii) $X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z)$.

iii) $(X \cap Y)^c = X^c \cup Y^c$.

iv) $(X \cup Y)^c = X^c \cap Y^c$.

DEMOSTRACION Para demostrar i), tomemos $u \in X \cap (Y \cup Z)$. Sigue entonces por definición que $u \in X$ y $u \in Y \cup Z$, resultando de esto último que $u \in Y$ ó $u \in Z$. Suponiendo por ejemplo que se verifica la primera situación, deducimos que $u \in X \cap Y$, y por lo tanto $u \in (X \cap Y) \cup (X \cap Z)$, por propiedad iv) de la unión. Claramente, lo mismo ocurre en el otro caso. Hemos probado luego que el conjunto del miembro de la izquierda está contenido en el de la derecha.

Para probar la otra inclusión, tomemos ahora $u \in (X \cap Y) \cup (X \cap Z)$, de donde $u \in X \cap Y$ ó $u \in X \cap Z$, por definición de unión. Puesto que $Y \cup Z$ contiene tanto a Y como a Z, en cualquiera de los dos casos resulta que $u \in X$ y $u \in Y \cup Z$, esto es, $u \in X \cap (Y \cup Z)$, como queríamos probar.

La prueba de ii) es similar, por lo que la proponemos como ejercicio, mientras que iii) es consecuencia de la validez de la siguiente cadena de equivalencias:

$$u \in (X \cap Y)^c \Leftrightarrow u \notin X \cap Y \Leftrightarrow (u \notin X) \vee (u \notin Y)$$
$$\Leftrightarrow (u \in X^c) \vee (u \in Y^c) \Leftrightarrow u \in X^c \cup Y^c.$$

Finalmente, podemos probar iv) usando iii) y las propiedades del complemento, ya que

$$(X \cup Y)^c = ((X^c)^c \cup (Y^c)^c)^c = ((X^c \cap Y^c)^c)^c = X^c \cap Y^c. \quad \Diamond$$

Observación. Las definiciones de intersección y unión de dos conjuntos se extienden a familias arbitrarias de conjuntos. En efecto, supongamos dado, para cada i perteneciente a una cierta familia I de índices, un subconjunto A_i de $\mathcal{U}$. Definimos entonces la intersección y la unión de la familia de conjuntos A_i en las formas

$$\bigcap_{i\in I} A_i = \{x \in \mathcal{U} \,:\, x \in A_j \text{ para todo } j \in I\}$$
$$\bigcup_{i\in I} A_i = \{x \in \mathcal{U} \,:\, x \in A_j \text{ para algún } j \in I\}\,,$$

respectivamente. Claramente, estas definiciones generalizan las nociones de intersección y unión dadas anteriormente. $\Diamond$

DIFERENCIA Si A y B son conjuntos, definimos su *diferencia* en la forma

$$A - B = \{x \in \mathcal{U} \,:\, x \in A \text{ y } x \notin B\}\,.$$

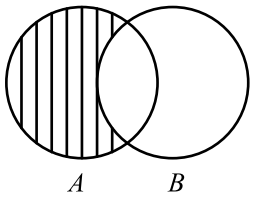


Figura 1.6: $A - B$

Por ejemplo, $A - B = \{1, 2, 3, 5, 6, 7\}$ si A y B son los conjuntos del ejemplo que sigue a la definición de unión. En el caso particular de que B esté incluido en A, la diferencia $A - B$ se llama el complemento de B *relativo* a A.

Observaciones Contrariamente a lo que ocurre con la intersección y la unión, la diferencia no es asociativa ni conmutativa. Encomendamos al lector la tarea de verificar que en general $(A-B)-C \neq A-(B-C)$ y $A-B \neq B-A$, tomando por ejemplo $A = \{1, 3, 6\}$, $B = \{2, 4, 6\}$ y $C = \{3\}$. Ya veremos que en la operación que definiremos a continuación desaparece la asimetría en los roles de A y B que acabamos de resaltar.

Como apunte final, notemos que la diferencia puede ser descripta en términos de las operaciones anteriores, ya que claramente $A - B = A \cap B^c$.

DIFERENCIA SIMETRICA Si A y B son conjuntos, definimos su *diferencia simétrica* en la forma

$$A \triangle B = (A - B) \cup (B - A)\,.$$

Por ejemplo, si $A = \{x : x \leq 8\}$ y $B = \{x : 4 \leq x < 11\}$ entonces

$$A \triangle B = \{1, 2, 3, 9, 10\}\,.$$

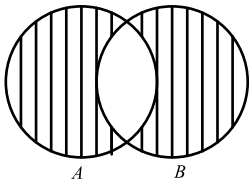


Figura 1.7: $A \triangle B$

Propiedades de la Diferencia Simétrica

i) $(X \triangle Y) \triangle Z = X \triangle (Y \triangle Z)$ y $X \triangle Y = Y \triangle X$.

ii) $X \triangle Y = (X \cup Y) - (X \cap Y)$.

iii) $X \triangle \emptyset = X$ y $X \triangle \mathcal{U} = X^c$.

iv) $X \triangle X = \emptyset$.

DEMOSTRACION Un rápido vistazo nos muestra que la diferencia simétrica goza de propiedades similares a las de la unión: es asociativa y conmutativa y admite un elemento neutro ($\emptyset$). Además, podemos interpretar la propiedad iv) diciendo que todo conjunto X admite inverso respecto a la diferencia simétrica, siendo éste el mismo X. Notemos por último que la propiedad ii) asegura que $X \triangle Y$ coincide con $X \cup Y$ cuando X e Y son disjuntos.

Sólo demostraremos la propiedad asociativa, ya que las restantes son inmediatas. Puesto que su prueba es algo engorrosa, procederemos mediante la siguiente *tabla de verdad*:

$u \in X$	$u \in Y$	$u \in Z$	$u \in (X \triangle Y) \triangle Z$	$u \in X \triangle (Y \triangle Z)$
V	V	V	V	V
V	V	F	F	F
V	F	V	F	F
V	F	F	V	V
F	V	V	F	F
F	V	F	V	V
F	F	V	V	V
F	F	F	F	F

Expliquemos en qué consiste este método de demostración. Lo que hemos hecho es considerar, en las tres columnas de la izquierda, los 8 casos que se presentan al plantear la pertenencia ó no de un elemento genérico u a

cada uno de los conjuntos involucrados, resultando que en todos ellos las proposiciones $u \in (X \triangle Y) \triangle Z$ y $u \in X \triangle (Y \triangle Z)$ tienen el mismo valor de verdad. En otras palabras, hemos probado que la proposición

$$u \in (X \triangle Y) \triangle Z \Leftrightarrow u \in X \triangle (Y \triangle Z)$$

es verdadera cualquiera sea u, lo que por definición prueba la igualdad de ambos conjuntos. $\Diamond$

PRODUCTO CARTESIANO. Si A y B son conjuntos, definimos el *producto cartesiano* de A y B como el conjunto

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

cuyos elementos (a, b) se denominan *pares ordenados*. Si bien podríamos definir formalmente la noción de par ordenado, dada su sencillez la consideraremos un concepto primitivo, cercano a la intuición del lector.

Por ejemplo, si $S = \{1, 2, \alpha\}$ y $T = \{1, 2, 3\}$ resulta que

$$S \times T = \{(1,1), (1,2), (1,3), (2,1), (2,2), (2,3), (\alpha,1), (\alpha,2), (\alpha,3)\}.$$

Notemos que cada objeto de $S \times T$ está determinado por dos elementos, uno de S y otro de T, dados en ese *orden*, y que por ejemplo consideramos distintos los pares $(1, 2)$ y $(2, 1)$. Respecto a este último comentario, establezcamos precisamente la definición de igualdad en cualquier producto cartesiano $A \times B$:

$$(a, b) = (a', b') \Leftrightarrow a = a' \text{ y } b = b',$$

donde $a, a' \in A$ y $b, b' \in B$.

Proponemos como un ejercicio sencillo demostrar que $A \times B = \emptyset$ si y sólo si $A = \emptyset$ ó $B = \emptyset$ y que $A \times B = B \times A$ si y sólo si $B = A$. En este último caso notaremos $A \times A = A^2$. Por ejemplo, $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ es la familiar representación de los puntos del plano como pares de números reales.

PARTICIONES DE UN CONJUNTO.

> Sea X un conjunto y sea $(X_i)_{i \in I}$ una familia de subconjuntos de X. Diremos que $(X_i)_{i \in I}$ es una *partición* de X si y sólo si se verifican las siguientes condiciones:
>
> $P_1)$ $X_i \neq \emptyset$ para todo $i \in I$
> $P_2)$ $\bigcup_{i \in I} X_i = X$
> $P_3)$ $X_i \cap X_j = \emptyset$ si $i \neq j$.

Resumidamente, una partición de un conjunto X es una familia de subconjuntos no vacíos de X tales que todo elemento de X pertenece a uno y solo a uno de dichos subconjuntos.

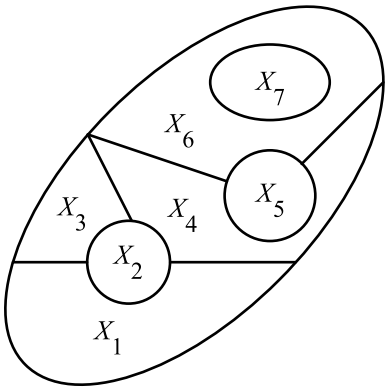


Figura 1.8: Partición de X

Por ejemplo, la familia $A_1 = \{a, c, f\}$, $A_2 = \{d\}$ y $A_3 = \{b, e\}$ es una partición del conjunto $A = \{a, b, c, d, e, f\}$, y los subconjuntos N_1 y N_2 de números naturales pares e impares, respectivamente, determinan una partición de $\mathbb{N}$.

Como otro ejemplo, algo más interesante, tomemos $A = \mathbb{R}^2$ y consideremos, para cada $b \in \mathbb{R}$, el subconjunto

$$\mathcal{L}_b = \{(x, y) \in A \,:\, y = x + b\}\,.$$

Veamos que la familia $(\mathcal{L}_b)_{b\in\mathbb{R}}$ es una partición de A. Es trivial que cada $\mathcal{L}_b$ es no vacío, pues por ejemplo $(0, b) \in \mathcal{L}_b$. Respecto a la condición P_2) de la definición, observemos que si $(u, v) \in A$, podemos escribir $v = u + (v - u)$, y por lo tanto $(u, v) \in \mathcal{L}_{v-u}$. Finalmente, y en cuanto a la pertenencia de (u, v) a un único miembro de la familia, notemos simplemente que

$$(u, v) \in \mathcal{L}_b \Leftrightarrow v = u + b \Leftrightarrow b = v - u.$$

La situación admite una interpretación geométrica muy sencilla, ya que los subconjuntos $\mathcal{L}_b$ de la partición son las rectas de pendiente 1 del plano, que claramente lo cubren y no se cortan entre sí por ser paralelas. $\Diamond$

1.1.3. Ejercicios

A lo largo de los siguientes ejercicios las letras mayúsculas denotan conjuntos, y las letras minúsculas elementos.

1. En cada uno de los siguientes casos exhibir un ejemplo en el que se satisfagan las condiciones planteadas:

   a) $\{a\} \subseteq A$ y $\{a\} \in A$.

b) $\{a\} \nsubseteq A$ y $\{a\} \in A$.

c) $\{a\} \subseteq A$ y $\{\{a\}\} \subseteq A$.

d) $\{a, b\} \subseteq A$,$\{c\} \in A$ y $\{a, b, c\} \nsubseteq A$.

e) $A \in B$ y $A \in \mathbb{P}(B)$.

f) $A \in B$ y $A \nsubseteq B$.

g) $\emptyset \in A$.

2. Demostrar que $\emptyset \subseteq A$ cualquiera sea el conjunto A.

3. Consideremos el conjunto referencial $\mathcal{U} = \{1, 2, 3, 4, 5, \{2\}, \{2, 5\}\}$ y sean $A = \{1, 2, 3, \{2\}\}$, $B = \{1, 5, \{2, 5\}\}$ y $C = \{1, 4, \{2\}\}$. Describir los conjuntos:

a) $B^c \cap C$

b) $(A - B) - C$

c) $A - (B - C)$

d) $(A \triangle B) \cap C$

e) $(A \triangle C) - B$

f) $(A \cup B^c) \cap C$.

4. Resolver las mismas cuestiones del ejercicio precedente tomando $\mathbb{N}$ como conjunto de referencia, A el conjunto de múltiplos de 12, B el conjunto $\left\{n : n^2 > 10000\right\}$ y C el conjunto de divisores de 336.

5. Representar las siguientes operaciones de conjuntos mediante diagramas de Venn:

a) $A \cap (B \cup C)$

b) $A \cup (B \cap C)$

c) $A^c \cup (B \cap C)$

d) $A \triangle (B \cup C)$

e) $A - (B^C \triangle C)$

f) $A \cup (B \triangle C)$.

6. Demostrar que todas las operaciones de conjuntos definidas en el texto pueden expresarse en términos de la unión y la complementación.

7. Demostrar las siguientes propiedades:

a) $A - \emptyset = A$.

b) $A \subseteq B \Leftrightarrow A \cap B = A$.

c) $A \subseteq B \Leftrightarrow A \cup B = B$.

d) $A = (A - B) \cup (A \cap B)$.

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

f) $A \cap (B \triangle C) = (A \cap B) \triangle (A \cap C)$.

g) $A - (A \triangle B) = A \cap B$.

8. Sean $A = \{1, 2, 3, 4\}$, $B = \{1, 4, 5\}$ y $C = \{x, y, z\}$. Listar los elementos de $A \times C$, B^2, $(A \cap B) \times A$, $(A \cap C) \times B$ y $(A \times C) \times B$.

9. Demostrar las siguientes propiedades del producto cartesiano:

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

b) $(A \cup B) \times C = (A \times C) \cup (B \times C)$.

c) $(A - B) \times C = (A \times C) - (B \times C)$.

d) $(A \triangle B) \times C = (A \times C) \triangle (B \times C)$.

10. Sea S un subconjunto de $A \times B$ y sean:

$$S_A = \{a \in A : (a, y) \in S \text{ para algún } y \in B\},$$
$$S_B = \{b \in B : (x, b) \in S \text{ para algún } x \in A\} .$$

Probar que $S \subseteq S_A \times S_B$. Mostrar un ejemplo en el que esta inclusión sea estricta y otro en el que valga la igualdad.

11. *a*) Demostrar que $A \subseteq B \Leftrightarrow \mathbb{P}(A) \subseteq \mathbb{P}(B)$.

b) Si $A = \{1, 2, 3, 4, 5\}$, ¿cuántos elementos tiene $\mathbb{P}(\mathbb{P}(A))$?

c) Si A es como en *b*), calcular el número de elementos $S \in \mathbb{P}(A)$ tales que $3 \in S$ y el número de elementos $S \in \mathbb{P}(A)$ tales que $3 \notin S$.

d) ¿Existe algún conjunto X que admita exactamente 27 subconjuntos? ¿Y 20 subconjuntos?

12. *a*) Determinar una partición de $\mathbb{N}$ en 8 subconjuntos infinitos.

b) Determinar una partición infinita de $\mathbb{N}$ tal que ninguno de sus miembros sea unitario (un conjunto se dice unitario si tiene exactamente un elemento).

c) Calcular el número de particiones del conjunto $\{1, 2, 3, 4, 5, 6\}$.

1.2. Relaciones

1.2.1. Definición y terminología

La palabra relación, y términos sucedáneos de ella como conexión, vínculo, correspondencia y otros, son de uso frecuente en nuestra vida cotidiana. Por ejemplo, hablamos (y oímos hablar) de relaciones humanas, de relaciones de causa y efecto, de la relación existente entre dos cantidades variables, y muchas otras de la más diversa índole. Haciendo un esfuerzo de abstracción para analizar lo que pueden tener en común los distintos usos del término, vemos que en todos los casos se establece un apareamiento entre ciertos elementos de un conjunto y algunos elementos de otro. Así, si nos referimos a la relación de amistad, distinguimos −entre todos los pares posibles de personas− aquellos cuyos integrantes son amigos entre sí, mientras que por ejemplo podemos relacionar el conjunto de habitantes de una ciudad con el conjunto de calles de la misma a través de los pares (vecino, calle en que vive). En la definición matemática que brindamos a continuación, completamente general y abstracta, se refleja la idea subyacente en toda relación, la de distinguir ciertos pares de elementos.

> Si A y B son conjuntos, cualquier subconjunto $\Re$ de $A \times B$ se dirá una *relación* de A en B. Si $(a, b) \in \Re$ diremos que a está *relacionado* con b, y emplearemos frecuentemente la notación alternativa $a \Re b$. En el caso $B = A$ diremos simplemente que $\Re$ es una relación en A.

Ejemplos 1.2.1 Si $A = \{1, 2, 3\}$ y $B = \{1, 2, 5, 6\}$, las siguientes son relaciones de A en B:

1) $\Re_1 = \{(1, 1), (1, 5), (2, 2), (2, 6), (3, 1), (3, 5)\}$.

2) $\Re_2 = \{(1, 1), (2, 1), (2, 5)\}$.

3) $\Re_3 = \emptyset$

Notemos que la primera de las relaciones admite la siguiente descripción: $a \,\Re_1\, b$ si y sólo si a y b son ambos impares ó ambos pares, mientras que no se intuye ningún patrón de formación de los pares del ejemplo 2). Lo hemos incluido para enfatizar el hecho de que *todo* subconjunto de $A \times B$ es una relación de A en B (quizás el lector piense que en el ejemplo 3) hemos abusado de esta libertad de elección). Como dato ilustrativo, que podremos corroborar más adelante, señalemos que hay 4096 relaciones distintas de A en B.

Ejemplos 1.2.2 Consideremos ahora las siguientes relaciones en $\mathbb{R}$:

4) $\Re_4 = \{(x, y) : 3x + 2y = 9\}$.

5) $\Re_5 = \{(x, y) : x^2 = y^2\}$.

6) $\Re_6 = \{(x, y) : x < 2\}$.

El lector informado no tendrá dificultades en advertir que los subconjuntos del plano determinados por estas relaciones consisten de una recta, de una unión de dos rectas y de un semiplano, respectivamente. Respecto del último ejemplo, notemos que la palabra relación tiene en Matemática un significado mucho más amplio que en el lenguaje corriente, ya que la pertenencia a $\Re_6$ de un par (x, y) no establece ningún tipo de lazo o asociación entre las componentes, en el sentido usual de estos términos. $\Diamond$

El dominio y la imagen de una relacion

Si $\Re$ es una relación de A en B, definimos el *dominio* y la *imagen* de $\Re$ en las formas

$$\begin{aligned} \mathrm{Dom}\,(\Re) &= \{x \in A \,:\, x\,\Re\, b \text{ para algún } b \in B\} \\ \mathrm{Im}\,(\Re) &= \{y \in B \,:\, a\,\Re\, y \text{ para algún } a \in A\}\,. \end{aligned}$$

Por ejemplo, volviendo a las relaciones de los ejemplos 1.2.1 y 1.2.2, es sencillo verificar que los dominios de las relaciones $\Re_1$, $\Re_3$ y $\Re_6$ son los conjuntos A, $\emptyset$ y $\{x \in \mathbb{R} : x < 2\}$, respectivamente, mientras que B, $\{1, 5\}$ y $\mathbb{R}$ son sus correspondientes imágenes. Sugerimos que el lector efectúe estas determinaciones en los restantes casos.

Representacion matricial Si A y B son conjuntos finitos, digamos de m y n elementos respectivamente, es cómodo describir las relaciones de A en B mediante esquemas matriciales (tablas) de ceros y unos. Precisamente, dada una relación $\Re$ de A en B, se la representa por una matriz de m filas y n columnas (las filas rotuladas por los elementos de A y las columnas por los elementos de B), de acuerdo con la siguiente regla: en la entrada correspondiente a la fila x ($x \in A$) y columna y ($y \in B$) colocamos un 1 si $(x, y) \in \Re$ y un 0 si $(x, y) \notin \Re$.

Recíprocamente, es claro que toda matriz de m filas por n columnas cuyas entradas son ceros o unos corresponde a una relacion de A en B.

Por ejemplo, sean $A = \{1, 2, 3, 4\}$ y $B = \{a, b, c\}$. Entonces la forma matricial de la relación $\Re_1 = \{(1, a), (3, c), (3, a), (2, b), (2, c), (4, a)\}$ es

	a	b	c
1	1	0	0
2	0	1	1
3	1	0	1
4	1	0	0

,

mientras que

	a	b	c
1	0	0	0
2	0	0	1
3	0	1	1
4	1	0	0

es la tabla de la relación $\Re_2 = \{(2, c), (3, b), (3, c), (4, a)\}$.

Observemos finalmente que el número de filas (columnas) no nulas indican la cantidad de elementos del dominio (de la imagen).

1.2.2. Propiedades especiales

Algunas relaciones bien conocidas entre objetos de la Matemática, como el paralelismo de rectas, la semejanza de triángulos, las desigualdades numéricas, etc., gozan de ciertas propiedades interesantes que vale la pena definir y estudiar en general. Suponiendo entonces que $\Re$ es una relación en un conjunto cualquiera A, destaquemos las siguientes propiedades:

REFLEXIVIDAD $\Re$ se dice *reflexiva* si y sólo si

$$a \,\Re\, a$$

para todo $a \in A$.

En la mayoría de los es sencillo decidir acerca de la validez de esta propiedad, ya que en términos de conjuntos sólo se trata de verificar la inclusión $D(A^2) \subseteq \Re$, donde

$$D(A^2) = \left\{(x, y) \in A^2 : y = x\right\}$$

es la usualmente llamada *diagonal* de A^2.

Por ejemplo, el paralelismo de rectas y la semejanza de triángulos son relaciones reflexivas, ya que toda recta (todo triángulo) es paralela (semejante) a sí misma (mismo). No lo es en cambio la relación de perpendicularidad entre rectas ni la relación $\Re_4$ del ejemplo 1.2.2, ya que se comprueba inmediatamente que $(9/5, 9/5)$ es el único par de la relación cuyas componentes son iguales.

SIMETRIA $\Re$ se dice *simétrica* si y sólo si la proposición

$$a \,\Re\, b \Rightarrow b \,\Re\, a$$

es verdadera cualesquiera sean $a, b \in A$.

Ejemplos de relaciones simétricas son el paralelismo y la perpendicularidad de rectas en el plano, la relación

$$(m,n)\mathcal{Z}(r,s) \Leftrightarrow m+s=n+r$$

en $\mathbb{N}^2$ y la relación $\mathcal{D}$ definida en el conjunto de partes de cualquier conjunto X por

$$S\,\mathcal{D}\,T \Leftrightarrow S\cap T=\emptyset,$$

como se verifica fácilmente. Respecto a relaciones no simétricas, mencionemos la relación $a\,\Re\,b \Leftrightarrow a<b$ en el conjunto de números reales y las relaciones $\Re_4$ y $\Re_6$ del ejemplo 1.2.2. En efecto, notemos que $1\,\Re_4\,3$ y $0\,\Re_6\,4$, mientras que las afirmaciones $3\,\Re_4\,1$ y $4\,\Re_6\,0$ son falsas.

ANTISIMETRIA $\Re$ se dice *antisimétrica* si y sólo si la proposición

$$(a\,\Re\,b \wedge b\,\Re\,a) \Rightarrow a=b$$

es verdadera cualesquiera sean $a,b\in A$.

Ejemplos típicos de antisimetría son la relación $a\,\Re\,b \Leftrightarrow a\leq b$ en el conjunto de números reales y la relación de inclusión en el conjunto de partes de cualquier conjunto. Por otro lado, la relación $\Re_6$ no lo es, ya que existen pares de elementos distintos (por ejemplo 0 y 1) mutuamente relacionados.

Intentando aclarar aún más el concepto, consideremos la relación

$$\mathcal{F}=\{(1,2),(2,5),(2,6)\},$$

definida en $\mathbb{N}$. En este caso, el antecedente $(a\,\mathcal{F}\,b \wedge b\,\mathcal{F}\,a)$ del condicional que define la antisimetría es falso cualesquiera sean a y b, y por lo tanto dicha proposición es verdadera, esto es, $\mathcal{F}$ es antisimétrica.

Como último comentario, observemos que la simetría y la antisimetría no son nociones antagónicas. De hecho, toda relación $\Re$ en un conjunto A tal que $\Re\subseteq D(A^2)$ satisface ambas propiedades.

TRANSITIVIDAD $\Re$ se dice *transitiva* si y sólo si la proposición

$$(a\,\Re\,b \wedge b\,\Re\,c) \Rightarrow a\,\Re\,c$$

es verdadera cualesquiera sean $a,b,c\in A$.

Por ejemplo, son transitivas la semejanza de triángulos, las relaciones numéricas de menor y mayor y las siguientes relaciones $\mathcal{T}_1$ y $\mathcal{T}_2$ en $\mathbb{N}$:

$$\mathcal{T}_1=\{(1,1),(1,2),(2,1),(2,2),(3,4),(3,6),(4,6)\}$$
$$\mathcal{T}_2=\{(1,1),(1,4),(2,3)\}.$$

Como ejemplos de relaciones no transitivas señalemos la perpendicularidad de rectas, la relación $\Re_4$ de 1.2.2 y la relación obtenida a partir de $\mathcal{T}_2$ agregando el par $(4,5)$. Dejamos a cargo del lector la verificación de estos hechos.

Ejemplo 1.2.3 Es claro que en cualquier conjunto X la relación de igualdad es reflexiva, simétrica, antisimétrica y transitiva. Veamos que ella es la única relación no vacía definida en X que satisface las cuatro propiedades. En efecto, sea $\Re$ una tal relación y sea $x \,\Re\, y$. Sigue entonces por simetría que $y \,\Re\, x$, y por lo tanto $y = x$, por ser $\Re$ antisimétrica. Hemos probado luego la inclusión $\Re \subseteq D(X^2)$. Siendo $\Re$ reflexiva también vale la inclusión contraria, de donde concluimos que

$$\Re = D(X^2),$$

esto es, $\Re$ es la relación de igualdad. Notemos que no fue necesario usar la hipótesis de transitividad, lo que significa que una relación que satisface las tres primeras propiedades es necesariamente la relación de igualdad, siendo entonces *a fortiori* transitiva. $\Diamond$

Relaciones de orden.

En el apartado anterior hemos hecho referencia a la relación $\leq$ definida en el conjunto de números reales, que satisface las propiedades de reflexividad, antisimetría y transitividad. Generalizando esta situación, introducimos la siguiente definición:

> Se llama *relación de orden* en un conjunto A a cualquier relación reflexiva, antisimétrica y transitiva $\mathcal{O}$ definida en A.

Alternativamente, diremos también que $\mathcal{O}$ es un orden en A ó que el par $(A, \mathcal{O})$ es un *conjunto ordenado*. Por ejemplo, la igualdad es una relación de orden en cualquier conjunto, y

$$\mathcal{O} = \{(\alpha, \alpha), (\alpha, \beta), (\beta, \beta), (\gamma, \beta), (\gamma, \gamma)\}$$

es una relación de orden en el conjunto $G = \{\alpha, \beta, \gamma\}$.

Volviendo al caso general, sea $(A, \mathcal{O})$ un conjunto ordenado y supongamos adicionalmente que para cualquier par de elementos $x, y \in A$ se satisface alguna de las relaciones $x \,\mathcal{O}\, y$ ó $y \,\mathcal{O}\, x$. Diremos entonces que $\mathcal{O}$ es un *orden total* en A, ó que $(A, \mathcal{O})$ es un conjunto totalmente ordenado. Por ejemplo (sin duda este hecho es familiar para el lector), la relación $\leq$ es un orden total en $\mathbb{R}$, mientras que el conjunto $(G, \mathcal{O})$ de arriba no es totalmente ordenado, ya que ninguno de los pares (α, γ) y (γ, α) pertenece a la relación $\mathcal{O}$.

Finalizamos esta breve presentación de la noción de conjunto ordenado exhibiendo algunos otros ejemplos. Retomaremos la cuestión en el próximo capítulo, donde estudiaremos en detalle las propiedades del orden usual en $\mathbb{R}$.

Ejemplos 1.2.4 En cada uno de los siguientes casos, es inmediato verificar que la relación $\mathcal{O}$ definida en el conjunto A es una relación de orden:

1) $A = \mathbb{P}(X)$ (X un conjunto cualquiera) y

$$S \mathcal{O} T \Leftrightarrow S \subseteq T.$$

Si X tiene al menos dos elementos, es muy sencillo mostrar dos subconjuntos de X de modo que no valga ninguna de las dos posibles relaciones de inclusión entre ellos. Deducimos entonces que el orden que acabamos de definir no es un orden total en A. Vale la pena señalar también que cambiando " $\subseteq$ " por " $\supseteq$ " se obtiene otra relación de orden en A.

2) $A = \mathbb{R}^2$. Dados $a = (a_1, a_2)$ y $b = (b_1, b_2)$ en A se define

$$a \mathcal{O} b \Leftrightarrow a = b \text{ ó } a_1 < b_1 \text{ ó } (a_1 = b_1 \text{ y } a_2 < b_2).$$

Examinando los diferentes casos que pueden presentarse, se comprueba fácilmente que $\mathcal{O}$ es un orden total en A, derivado del orden usual de los números reales. Se lo llama orden *lexicográfico*, debido a su analogía con la forma en que se ordenan las palabras en un diccionario.

3) $A = \mathbb{R}^2$. Empleando la misma notación que en 2) se define

$$a \mathcal{O} b \Leftrightarrow a_1 \leq b_1 \text{ y } a_2 \leq b_2.$$

A diferencia del anterior no se trata de un orden total, pues por ejemplo no existe ninguna relación con respecto a $\mathcal{O}$ entre los pares $(1, 2)$ y $(3, 0)$.

4) $A = \{\alpha, \beta, \gamma\}$ y

$$\mathcal{O} = \{(\alpha, \alpha), (\alpha, \gamma), (\beta, \alpha), (\beta, \beta), (\beta, \gamma), (\gamma, \gamma)\}.$$

Por simple inspección, deducimos inmediatamente que $(A, \mathcal{O})$ es totalmente ordenado. Digamos de paso, que pueden definirse 5 relaciones de orden esencialmente distintas en A, una sóla de las cuales determina un orden total. $\Diamond$

Relaciones de equivalencia.

Se llama *relación de equivalencia* en un conjunto A a cualquier relación reflexiva, simétrica y transitiva $\mathcal{R}$ definida en A.

Es frecuente usar determinados símbolos para designar una relación de equivalencia, tales como " $\sim$ ", " $\approx$ ", " $\equiv$ ", etc. También, si $\sim$ es una relación de equivalencia en A y u y v son elementos de A tales que $u \sim v$ (en cuyo caso $v \sim u$), suele decirse que u y v son *equivalentes*.

Ejemplos 1.2.5 Mencionemos en primer término algunas relaciones de tipo geométrico, como la congruencia de segmentos, la semejanza de triángulos y el paralelismo de rectas. Obviamente, la igualdad es una relación de equivalencia, siendo en general de equivalencia cualquier relación que pueda ser descripta a través de una igualdad. Por ejemplo, la relación de tener el mismo centro, en el conjunto de círculos del plano, ó la relación de *cumplir años en el mismo mes*, en el conjunto de miembros de una comunidad.

Ejemplos de otra índole son los siguientes (en todos los casos el lector deberá verificar que la relación definida es de equivalencia):

1) $A = \mathbb{N}$; $m \sim n \Leftrightarrow m + n$ es par.

2) $A = \mathbb{N}^2$; $(m, n) \sim (r, s) \Leftrightarrow ms = nr$.

3) $A = \mathbb{R}$; $x \sim y \Leftrightarrow x^2 + y = x + y^2$.

4) $A = \{1, 2, 3, 4\}$; $\Re = \{(1, 1), (2, 2), (2, 4), (3, 3), (4, 2), (4, 4)\}$. $\quad \Diamond$

Clases de equivalencia Informalmente hablando, una relación de equivalencia generaliza la noción de igualdad, identificando aquellos elementos que son equivalentes entre sí. Las propiedades que caracterizan la relación garantiza la consistencia de tal identificación. Por ejemplo, en la relación de los cumpleaños clasificamos a los miembros de la comunidad según su mes de nacimiento, y partimos así el conjunto en 12 clases: la clase de los que cumplen años en enero, la de los que cumplen en febrero, etc.

Para darle forma precisa a estos comentarios, introducimos la noción de clase de equivalencia:

> Sea $\sim$ una relación de equivalencia en un conjunto A y sea $a \in A$. Definimos entonces la *clase de equivalencia* de a en la forma
>
> $$Cl(a) = \{x \in A \,:\, x \sim a\}\,.$$

Por ejemplo, en las relaciones 2) y 3) del ejemplo 1.2.5 tenemos que

$$Cl\left((1, 2)\right) = \left\{(x, y) \in \mathbb{N}^2 : y = 2x\right\}$$

y $Cl(2) = \{2, -1\}$, respectivamente. Conservando la notación de la definición anterior, señalemos dos propiedades fundamentales de las clases de equivalencia.

Proposición 1.2.6 Son válidas las siguientes propiedades (las letras designan elementos de A):

1) $a \in Cl(a)$, lo que prueba en particular que $Cl(a) \neq \emptyset$.

2) $Cl(a) \cap Cl(b) = \emptyset$ ó $Cl(a) = Cl(b)$. Precisamente,

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

DEMOSTRACION. El primer enunciado es consecuencia de la reflexividad de la relación.

Para probar 2), supondremos que las clases no son disjuntas y probaremos entonces que son iguales. Tomando a tal efecto $z \in Cl(a) \cap Cl(b)$, deducimos primero por simetría y transitividad que $a \sim b$, ya que $a \sim z$ y $z \sim b$. Luego, razonando en forma muy similar arribamos a la igualdad de las clases, ya que

$$x \in Cl(a) \Leftrightarrow x \sim a \Leftrightarrow x \sim b \Leftrightarrow x \in Cl(b)\,.$$

En cuanto a la segunda afirmación, observemos que ya hemos probado que las clases de equivalencia de dos elementos relacionados entre sí son iguales. Recíprocamente, si $Cl(a) = Cl(b)$ resulta que $x \sim a \Leftrightarrow x \sim b$ cualquiera sea $x \in A$. Tomando en particular $x = a$ concluimos que $a \sim b$, como queríamos probar. $\Diamond$

SISTEMAS DE REPRESENTANTES Sea $\sim$ una relación de equivalencia en A y sea S un subconjunto de A con la siguiente propiedad:

$$\text{para todo } a \in A \text{ existe un único } u \in S \text{ tal que } a \sim u\,.$$

Diremos entonces que S es un *sistema de representantes* de las clases de equivalencia, y que el conjunto

$$A/\sim = \{\, Cl(u) : u \in S \,\}$$

es el *conjunto cociente* de A por la relación $\sim$.

Intuitivamente, un sistema de representantes se obtiene eligiendo un elemento (y sólo uno) en cada clase de equivalencia. Por ejemplo, tomemos el conjunto de palabras de una enciclopedia y consideremos en él la siguiente relación: dos palabras son equivalentes si y sólo si comienzan con la misma letra. En este caso, un sistema de representantes de la relación debe contener exactamente 27 palabras, una para cada posible letra inicial de nuestro idioma, y por lo tanto el conjunto cociente tendrá 27 elementos. Cada uno de ellos se obtiene identificando entre sí todas las palabras que comienzan con una cierta letra, a las que agrupamos y pensamos como un único elemento del conjunto cociente. Veamos cómo proceder en otros casos.

Ejemplos 1.2.7 Consideremos primero la relación $\parallel$ de paralelismo en el conjunto de rectas del plano. Puesto que dos rectas no verticales son paralelas si y sólo si tienen la misma pendiente, cada número real a determina una clase de equivalencia (las rectas de pendiente a), habiendo una clase adicional cuyos elementos son las rectas verticales. Tomando como representante de una clase del primer tipo a la recta $\mathcal{L}_a$ de ecuación $y = ax$, y como representante de la clase de rectas verticales a la recta $\mathcal{L}_\infty$ de ecuación $x = 0$, resulta que

$$S = \{\, \mathcal{L}_a \,:\, a \in \mathbb{R}\} \cup \mathcal{L}_\infty$$

es un sistema de representantes de las clases de $\|$.

Trabajaremos ahora con las relaciones del ejemplo 1.2.5. La situación en el caso 1) es muy sencilla, pues es bien sabido que la suma de dos números naturales es par si y sólo si ambos son pares ó ambos son impares. Por lo tanto, la clase de un elemento m consiste de todos los números que tienen la misma paridad que m. Hay en consecuencia 2 clases, la de los pares y la de los impares, y podemos elegir $\{1, 2\}$ como sistema de representantes.

En el ejemplo 3), la expresión que define la relación admite la forma equivalente

$$x \sim y \Leftrightarrow (x + y)(x - y) = x - y,$$

igualdad que claramente se verifica si y sólo si $y = x$ ó $y = 1-x$. Por lo tanto, la clase de equivalencia de cualquier $a \in \mathbb{R}$ consiste de los elementos a y $1-a$, siendo $Cl(1/2)$ la única clase puntual. En cualquier caso, exactamente uno de los números a y $1 - a$ es mayor ó igual que $1/2$, de donde deducimos que $S = \{u \in \mathbb{R} : u \geq 1/2\}$ es un sistema de representantes de las clases.

En cuanto a 4), es claro que sólo hay dos formas posibles de elegir un sistema de representantes, a saber: $S = \{1, 2, 3\}$ ó $S = \{1, 3, 4\}$. $\Diamond$

RELACIONES DE EQUIVALENCIA Y PARTICIONES El concepto de partición de un conjunto, que definimos en la sección anterior, está íntimamente ligado al de relación de equivalencia. Precisaremos esta aserción a través del siguiente resultado:

Teorema 1.2.8 Si A es un conjunto y $\sim$ es una relación de equivalencia en A, el conjunto cociente $A/\sim$ es una partición de A. Recíprocamente, toda partición $\mathcal{P}$ de A determina una relación de equivalencia $\Re$ en A tal que $A/\Re = \mathcal{P}$.

DEMOSTRACION. Para establecer la primera afirmación, escribamos

$$A/\sim = \{Cl(u) : u \in S\},$$

donde S es cualquier sistema de representantes de las clases de equivalencia. Si $a \in A$, sigue por definición que $a \sim u_a$ para algún $u_a \in S$, en cuyo caso $a \in Cl(u_a)$. Queda probado entonces que $A = \bigcup_{u\in S} Cl(u)$.

Por otro lado, dos elementos distintos u y v de S no son equivalentes entre sí, ya que $u \sim u$ y todo elemento de A es equivalente a un único elemento de S. Sigue luego por proposición 1.2.6 que $Cl(u) \cap Cl(v) = \emptyset$. Recordando por último que las clases de equivalencia son no vacías, resulta que $A/\sim$ es una partición de A.

Para demostrar la segunda parte del enunciado, tomemos una partición cualquiera $\mathcal{P} = (\mathcal{A}_i)_{i\in I}$ de A y consideremos la relación $\Re$ en A definida por

$$a \,\Re\, b \Leftrightarrow \text{existe } j \in I \text{ tal que } a \in \mathcal{A}_j \text{ y } b \in \mathcal{A}_j\,.$$

Es decir, dos elementos de A están relacionados si y sólo si pertenecen a un mismo miembro de la partición. Veamos que $\Re$ es una relación de equivalencia.

Puesto que $A = \bigcup_{i\in I} \mathcal{A}_i$, dado $a \in A$ existe un índice k tal que $a \in \mathcal{A}_k$, lo que asegura que $a\Re a$. Luego $\Re$ es reflexiva.

Igualmente sencillo resulta probar la simetría, ya que claramente las proposiciones $(a \in \mathcal{A}_j \wedge b \in \mathcal{A}_j)$ y $(b \in \mathcal{A}_j \wedge a \in \mathcal{A}_j)$ son lógicamente equivalentes.

Con respecto a la transitividad, sean a, b y c elementos de A tales que $a\Re b$ y $b\Re c$. Sigue entonces por definición que existen índices j y k tales que $\mathcal{A}_j$ contiene los elementos a y b y $\mathcal{A}_k$ contiene los elementos b y c, resultando en particular que $b \in \mathcal{A}_j \cap \mathcal{A}_k$. Puesto que dos miembros distintos de la partición son disjuntos, concluimos que $k = j$. En consecuencia a y c pertenecen a $\mathcal{A}_j$ y por lo tanto $a\Re c$, como queríamos demostrar.

Habiendo probado que $\Re$ es una relación de equivalencia, resta demostrar que las clases de equivalencia de $\Re$ son exactamente los miembros de la partición. Si $a \in A$ y j es el único índice tal que $a \in \mathcal{A}_j$, es claro por definición de $\Re$ que $x\Re a$ si y sólo si $x \in \mathcal{A}_j$, esto es, $Cl(a) = \mathcal{A}_j$. Por otro lado, si $\mathcal{A}_i$ es cualquier miembro de la partición, tomando cualquier elemento $z \in \mathcal{A}_i$ (recordemos que $\mathcal{A}_i \neq \emptyset$) resulta como arriba que $\mathcal{A}_i = Cl(z)$, y por lo tanto todo miembro de la partición es una clase de equivalencia. $\Diamond$

Para ilustrar el teorema anterior, tomemos por ejemplo $A = \{1, 2, 3, 4, 5\}$. Es inmediato verificar entonces que las particiones

$$A = \{1, 2\} \cup \{3, 5\} \cup \{4\}$$

y

$$A = \{1, 2, 4\} \cup \{3\} \cup \{5\}$$

corresponden respectivamente a las relaciones de equivalencia

$$\Re_1 = \{(1,1), (1,2), (2,1), (2,2), (3,3), (3,5), (4,4), (5,3), (5,5)\}$$

y

$$\Re_2 = \{(1,1), (1,2), (1,4), (2,1), (2,2), (2,4), (3,3), (4,1), (4,2), (4,4), (5,5)\} .$$

1.2.3. Ejercicios

1. En cada uno de los siguientes ítems se define una relación $\Re$ de un conjunto A en un conjunto B. Determinar en todos los casos el dominio y la imagen de $\Re$:

   a) $A = \{a, b, c\}$, $B = \{1, 2, b, 3\}$; $\Re = \{(a, 1), (c, b), (a, 3), (c, 1)\}$.

   b) A y B como en *a)* ; $\Re = \{(a, 2), (a, b), (a, 3), (a, 1)\}$.

   c) A y B como en *a)* ; $\Re = \{(b, b), (a, 2), (b, 3), (c, 1)\}$.

d) $A = \mathbb{R}\,, B = \mathbb{N}\,;\; a\,\Re\,b \Leftrightarrow b < a < b+1\,.$

e) $A = \mathbb{Z}\,, B = \mathbb{Q}\,;\; a\,\Re\,b \Leftrightarrow b = \dfrac{a}{2}\,.$

f) A el conjunto de rectas del plano$\,, B = A\,;$

$$a\,\Re\,b \Leftrightarrow a \text{ y } b \text{ se cortan en el origen}\,.$$

g) A y B como en $f)\,;$

$$a\,\Re\,b \Leftrightarrow a \text{ y } b \text{ se cortan en el segundo cuadrante}\,.$$

h) $A = B = \mathbb{R}\,;\; (a,b) \in \Re \Leftrightarrow 2a - b = 1\,.$

i) $A = B = \mathbb{R}\,;\; (a,b) \in \Re \Leftrightarrow a > b^2\,.$

j) $A = B = \mathbb{N}\,;\; a\,\Re\,b \Leftrightarrow a^2 < \dfrac{50}{3}\,.$

2. Sean A y B conjuntos y sea $\Re$ una relación de A en B. Demostrar que

$$\Re = \emptyset \Leftrightarrow \mathrm{Dom}(\Re) = \emptyset \Leftrightarrow \mathrm{Im}(\Re) = \emptyset\,.$$

3. Sea $X = \{a,b,c,d,e,f\}$. En cada uno de los siguientes casos, analizar la reflexividad, simetría, antisimetría y transitividad de la relación definida en X y representarla matricialmente:

a) $\Re = \{(a,a),(c,c)\}$.

b) $\Re = \{(a,a),(b,c),(d,f),(d,d),(f,d)\}$.

c) $\Re = \{(a,b),(b,c),(a,a)\}$.

d) $\Re = \{(a,b),(c,d),(f,e),(a,a),(b,b),(c,c),(d,d),(e,e)\}$.

e) $\Re = \emptyset\,.$

4. En cada uno de los ítems del ejercicio 3, determinar, cuando sea posible, el mínimo número de pares que deben agregarse a $\Re$ para que la nueva relación sea de orden o de equivalencia.

5. Una relación en un conjunto X se dice *circular* si y solo si la proposición

$$(a\,\Re\,b \,\wedge\, b\,\Re\,c) \;\Rightarrow\; c\,\Re\,a$$

es verdadera cualesquiera sean a, b y c en X .

a) Probar que toda relación de equivalencia es circular.

b) Probar que una relación reflexiva y circular es de equivalencia.

6. Verificar que las relaciones definidas en X en los siguientes ítems son de orden o de equivalencia. En el caso de las relaciones de orden analizar si el orden es total, y en el caso de las relaciones de equivalencia caracterizar las clases de equivalencia y determinar un sistema de representantes de las mismas:

   a) X un conjunto cualquiera ; $a \,\Re\, b \Leftrightarrow a = b$.

   b) $X = \mathbb{N}$; $a \,\Re\, b \Leftrightarrow b$ es múltiplo de a .

   c) $X = \{a, b, c, d, e, f\}$; $\Re$ definida por la tabla

	a	b	c	d	e	f
a	1	0	0	0	0	0
b	0	1	0	0	0	0
c	0	0	1	0	0	1
d	0	1	0	1	1	0
e	0	0	0	0	1	0
f	0	0	0	0	0	1

   d) X como en *c*) ; $\Re$ definida por la tabla

	a	b	c	d	e	f
a	1	0	1	0	1	0
b	0	1	0	1	0	1
c	1	0	1	0	1	0
d	0	1	0	1	0	1
e	1	0	1	0	1	0
f	0	1	0	1	0	1

   e) X es el conjunto de circunferencias del plano;
   $a \,\Re\, b \Leftrightarrow a$ y b son concéntricas .

   f) $X = \mathbb{N}^2$; $(a, b) \,\Re\, (c, d) \Leftrightarrow a + d = b + c$.

   g) $X = \mathbb{R}^2$; $(a, b) \,\Re\, (c, d) \Leftrightarrow (a < c) \vee [(a = c) \wedge (b \leq d)]$.

   h) $X = \mathbb{P}(\mathbb{N})$; $S \,\Re\, T \Leftrightarrow (S = T) \vee (S = \mathbb{N} - T)$.

   i) $X = \mathbb{R}$; $a \,\Re\, b \Leftrightarrow a^2 + b = a + b^2$.

   j) $X = \mathbb{N}$; $a \,\Re\, b \Leftrightarrow (a = b) \vee [(a \text{ es impar}) \wedge (b \text{ es par})]$.

7. Sea $A = \{a, b, c, d, e, f\}$.

a) Definir una relación de orden $\Re$ en A, que contenga exactamente 11 pares y que satisfaga las condiciones $b \,\Re\, d$, $d \,\Re\, c$ y $a \,\Re\, f$. Determinar cuántas distintas pueden definirse.

b) Definir una relación de equivalencia $\sim$ en A tal que $a \sim b$, $c \nsim b$, $c \notin Cl(e)$ y $d \in Cl(e)$ ¿Cuántas distintas pueden definirse?

8. *a*) Determinar cuántas relaciones de orden pueden definirse en un conjunto de 3 elementos.

b) Determinar cuántas relaciones de equivalencia pueden definirse en un conjunto de 4 elementos.

9. Sea $A = \{1, 2, \ldots, 10\}$. Representar matricialmente la relación de equivalencia en A asociada a la partición

$$A = \{1, 3, 8\} \cup \{2, 5, 7\} \cup \{9, 10\} \cup \{4\} \ .$$

¿De cuántas maneras puede elegirse un sistema de representantes?

1.3. Funciones

1.3.1. Definiciones

> Si $\Re$ es una relación de un conjunto A en un conjunto B, diremos que $\Re$ es una *función* ó *aplicación* de A en B si y sólo si para cada $a \in A$ existe un único $b \in B$ tal que $(a, b) \in \Re$.

Por ejemplo, sean $A = \{1, 2, 3, 4\}$ y $B = \{a, e, i, o, u\}$, y consideremos las relaciones

$$\begin{aligned}
\Re_1 &= \{(2, e), (3, i), (4, o), (1, u)\} \\
\Re_2 &= \{(3, e), (1, i), (2, i)\} \\
\Re_3 &= \{(3, a), (1, e), (4, e), (2, u)\} \\
\Re_4 &= \{(1, a), (3, e), (2, i), (1, i), (4, o)\}.
\end{aligned}$$

Una rápida inspección nos muestra que $\Re_1$ y $\Re_3$ son funciones de A en B, y $\Re_2$ y $\Re_4$ no lo son. En efecto, $\Re_2$ no es una función pues $4 \notin \mathrm{Dom}(\Re_2)$, mientras que $\Re_4$ no verifica la segunda condición de la definición, ya que $(1, a) \in \Re_4$ y $(1, i) \in \Re_4$.

En general, designaremos las funciones con letras del tipo $f, g, h \ldots$, y emplearemos lenguaje y notaciones especiales para referirnos a las relaciones funcionales. Así, si f es una función de A en B escribiremos

$$f : A \to B,$$

ó también

$$A \xrightarrow{f} B,$$

para reflejar la idea de que f *asigna* a cada elemento de A un elemento de B. Por otro lado, dado $a \in A$, la expresión $f(a)$ (se lee "f de a") indicará el único elemento b de B tal que $(a, b) \in f$. Diremos en tal caso que b es la *imagen* de a, o también que a es una *preimagen* de b. Finalmente, los conjuntos A y B se denominarán el *dominio* y el *codominio* de f, respectivamente.

IGUALDAD DE FUNCIONES. Es sencillo precisar la noción de igualdad entre dos funciones f y g de un conjunto A en un conjunto B, ya que siendo relaciones, las mismas serán iguales si y sólo si están determinadas por el mismo subconjunto del producto cartesiano $A \times B$. De acuerdo con las notaciones que introducimos arriba, la noción de igualdad se traduce entonces en la siguiente forma:

$$f = g \Leftrightarrow f(a) = g(a) \,\forall\, a \in A.$$

IMAGEN DE UNA FUNCION. Por supuesto, el concepto de imagen que definimos para relaciones también se aplica a funciones. Explícitamente, dada una función f de A en B tenemos:

$$\mathrm{Im}\,(f) = \{b \in B \,:\, b = f(a) \text{ para algún } a \in A\}.$$

RESTRICCION Y CORRESTRICCION. Si $f : A \to B$ es una función y $S \subseteq A$, la asignación $a \mapsto f(a)\,(a \in S)$ define una función de S en B. Se denomina la *restricción* de f a S, y la notaremos $f|_S$.

Por otro lado, sea T un subconjunto de B tal que $\mathrm{Im}(f) \subseteq T$. Es claro entonces que la misma relación f determina una función de A en T, que llamaremos la *correstricción* de f a T. Por ejemplo, la relación

$$\left\{\left(x, x^2\right) : x \in \mathbb{R}\right\}$$

determina una función de $\mathbb{R}$ en $\mathbb{R}$, y por correstricción también una función de $\mathbb{R}$ en el conjunto de números reales no negativos.

Ejemplos 1.3.1 Veamos algunos casos genéricos de aplicaciones, así como diversas formas de representar una función:

1) Una función se representa a veces mediante una tabla de valores. En las entradas de la columna de la izquierda aparecen los elemento del dominio, y en las de la derecha sus correspondientes imágenes. Por ejemplo, la función $f : \{1, 2, 3, 4, 5\} \to \{1, 2, 3, 4\}$ dada por la tabla

x	$f(x)$
1	3
2	1
3	4
4	2
5	1

2) Sean X e Y conjuntos. Una función $f : X \to Y$ se dice *constante* si $f(x) = c$ para todo $x \in X$, siendo c un elemento fijo de Y.

3) Sea A un conjunto y sea $S \in \mathbb{P}(A)$. La función $f : S \to A$ definida por

$$f(a) = a \;\forall\; a \in S$$

se llama la función *inclusión* de S en A, y la designaremos por $\iota_{S,A}$. Si $S = A$, la inclusión $\iota_{A,A}$ se nota I_A, y se denomina *función identidad* del conjunto A.

4) Es posible que el lector esté familiarizado con el manejo de funciones numéricas, es decir, funciones cuyo dominio y codominio son conjuntos numéricos. Dichas aplicaciones suelen representarse en la forma

$$y = f(x),$$

indicando la relación funcional existente entre la variable *independiente* x y la variable *dependiente* y. Casos típicos de tales funciones son

$y = 2x - 1$, $y = 1/x$, $y = \cos(x^2)$, $y = \log(e^x - 2)$, etc. De todos modos, debemos insistir en el hecho de que ninguna de las fórmulas anteriores define por sí sola una función, ya que deben especificarse siempre el dominio y el codominio. Por ejemplo, la relación funcional

$$y = x^2 + 4x + 3$$

define tanto una función $f_1 : \mathbb{N} \to \mathbb{N}$, como una función $f_2 : \mathbb{Z} \to \mathbb{Z}$.

5) Debido a que brinda una notación muy útil en muchas situaciones, mencionemos la función δ de Kronecker, que sólo toma los valores 0 y 1. Formalmente, δ es una función de $A \times A$ en $\{0, 1\}$, donde A es un conjunto cualquiera, y está definida por la fórmula

$$\delta(x, y) = \begin{cases} 1 \text{ si } x = y \\ 0 \text{ si } x \neq y\,. \end{cases}$$

En lugar de escribir $\delta(x, y)$ es costumbre usar el símbolo δ_{xy}, llamado familiarmente "delta"de Kronecker. $\Diamond$

1.3.2. Propiedades especiales

Por definición, una función asigna a cada elemento del dominio un único elemento del codominio, pero bien puede ocurrir que dos elementos distintos del dominio tengan la misma imagen, o que algunos elementos del codominio no admitan una preimagen. En relación con este tipo de situaciones, destacaremos a continuación ciertas características especiales atribuibles a cualquier función f de un conjunto A en un conjunto B:

INYECTIVIDAD Diremos que f es una función *inyectiva* si y sólo si

$$f(a) = f(a') \Leftrightarrow a = a'$$

cualesquieran sean $a, a' \in A$. Dicho en forma equivalente (y algo más coloquial), una función es inyectiva si y sólo si a elementos distintos del dominio corresponden elementos distintos del codominio. Debido a ello, se dice también que una función inyectiva es una función *uno a uno*. Ejemplos típicos de funciones inyectivas son las inclusiones.

SURYECTIVIDAD Diremos que f es una función *suryectiva* si y sólo si

$$\mathrm{Im}(f) = B,$$

es decir, para todo $b \in B$ existe $a \in A$ tal que $b = f(a)$. Como expresión alternativa, diremos que f es una aplicación de A *sobre* B.

BIYECTIVIDAD Diremos que f es una función *biyectiva*, o que f es una *biyección* de A en B, si y sólo si f es inyectiva y suryectiva. De acuerdo con

las definiciones precedentes, podemos resumir la situación en la siguiente forma: f es biyectiva si y sólo si

$$\text{para cada } b \in B \text{ existe un único } a \in A \text{ tal que } b = f(a).$$

Suele decirse también que f establece una correspondencia *biunívoca* entre los elementos de A y B. Una biyección de un conjunto X en sí mismo se dirá simplemente una biyección o *permutación* de X. Por ejemplo, la función identidad I_X es una permutación de cualquier conjunto X. Finalmente, dos conjuntos se dicen *coordinables* si y solo si existe una biyección entre ellos.

Ejemplos 1.3.2 Analicemos la validez de las propiedades que acabamos de definir para las siguientes funciones:

1) $f : \mathbb{N} \to \mathbb{Q}$ definida por $f(n) = 1 - 1/n$.

   Veamos que f es inyectiva pero no suryectiva. La inyectividad es clara, ya que

   $$f(a) = f(b) \Leftrightarrow 1 - 1/a = 1 - 1/b \Leftrightarrow 1/a = 1/b \Leftrightarrow a = b\,.$$

   Respecto a la suryectividad observemos que $mf(m) = m - 1$, resultando en particular que $f(m) \neq 1$ para todo $m \in \mathbb{N}$, ya que en caso contrario tendríamos $m = m-1$. Luego $1 \notin \mathrm{Im}(f)$ y f no es suryectiva.

2) $g : \mathbb{N} \to \mathbb{N}$ definida por

   $$g(n) = \begin{cases} n-1 & \text{si } n \text{ es par} \\ \frac{n+3}{2} & \text{si } n \text{ es impar y } n-1 \text{ es múltiplo de } 4 \\ \frac{n+1}{2} & \text{si } n \text{ es impar y } n-1 \text{ no es múltiplo de } 4\,. \end{cases}$$

   Por ejemplo, $g(2) = 1$, $g(5) = 4$ y $g(15) = 8$. Para estudiar la función en general, observemos en primer término que $\mathrm{Im}(g)$ contiene todos los impares, ya que $g(m+1) = m$ si m es impar. Si m es par, digamos $m = 2k$, consideremos los números naturales impares $a = 4k - 3$ y $b = 4k - 1$. Puesto que $a - 1 = 4(k-1)$ y $b - 1 = 4k - 2$, sigue por definición de g que

   $$g(a) = \frac{a+3}{2} = 2k = m = \frac{b+1}{2} = g(b)\,.$$

   Por lo tanto, g es suryectiva y no inyectiva.

3) $l : \mathbb{P}(X) \to \mathbb{P}(X)$ definida por $l(T) = T \cap X_0$, donde X es un conjunto y X_0 es un subconjunto propio de X.

   Si $S \subseteq X$, es claro por definición que $l(S) \subseteq X_0$, de donde deducimos que $X \notin \mathrm{Im}(l)$, por ser X_0 un subconjunto propio. Tenemos por otro lado que $l(X) = l(X_0) = X_0$, y en consecuencia l no es ni inyectiva ni suryectiva.

4) $h : \mathbb{Z} \to \mathbb{Z}$ definida por $h(m) = m + (-1)^m$.

Afirmamos que h es una biyección. En efecto, si $a \in \mathbb{Z}$ y $b = h(a)$, es evidente por la definición de h que a y b tienen distinta paridad, esto es, uno de ellos es par y el otro es impar. Por lo tanto $(-1)^a + (-1)^b = 0$, de donde sigue que

$$h(b) = b + (-1)^b = a + (-1)^a + (-1)^b = a.$$

En consecuencia $a \in \mathrm{Im}(h)$, lo que prueba que h es suryectiva.

Supongamos ahora que r y s son dos enteros tales que $h(r) = h(s)$. Aplicando h a esta igualdad, y de acuerdo con lo de arriba, obtenemos:

$$r = h\,(h(r)) = h\,(h(s)) = s.$$

Luego h es inyectiva y por lo tanto es una biyección. $\Diamond$

Composición de funciones. En muchos casos es posible aplicar sucesivamente dos funciones, y obtener de esta manera una nueva función. Precisamente, sean A, B, C y D conjuntos tales que $B \subseteq C$ y sean $f : A \to B$ y $g : C \to D$ funciones. Llamaremos *composición* de f y g a la función

$$g \circ f : A \to D$$

definida por $(g \circ f)(a) = g(f(a))$, cualquiera sea $a \in A$.

Por ejemplo, si $f(x) = 2x - 1$ y $g(x) = x^2 + 3$, tenemos:

$$(g \circ f)(x) = g(2x - 1) = (2x - 1)^2 + 3 = 4x^2 - 4x + 4,$$

mientras que

$$(f \circ g)(x) = f(x^2 + 3) = 2(x^2 + 3) - 1 = 2x^2 + 5.$$

Observemos que si bien las dos composiciones tienen sentido (ambas funciones tienen dominio y codominio en el conjunto de números reales), se obtienen funciones distintas, vale decir, la operación de componer no es conmutativa.

Nota Para la buena definición de $g \circ f$ sólo es necesario en rigor que pueda aplicarse g al elemento $f(a)$, por lo que la condición $B \subseteq C$ puede reemplazarse por la condición más débil $\mathrm{Im}(f) \subseteq C$. Por ejemplo, podemos componer las funciones $u(x) = e^{x^2/2}$ y $v(x) = 1/x$, obteniendo de esta manera la función $h(x) = e^{-x^2/2}$, pues $\mathrm{Im}(u) \subseteq \mathrm{Dom}(v) = \mathbb{R} - \{0\}$.

Veamos algunas propiedades elementales de la composición.

Proposición 1.3.3 Si $A \xrightarrow{f} B$, $B \xrightarrow{g} C$ y $C \xrightarrow{h} D$ son funciones, son válidas las siguientes afirmaciones:

(i) $h \circ (g \circ f) = (h \circ g) \circ f$.

(ii) $f \circ I_A = f$ y $I_B \circ f = f$.

(iii) Si f y g son inyectivas (resp. suryectivas) (resp. biyectivas), entonces $g \circ f$ es inyectiva (resp. suryectiva) (resp. biyectiva).

(iv) Si $g \circ f$ es inyectiva entonces f es inyectiva.

(v) Si $g \circ f$ es suryectiva entonces g es suryectiva.

DEMOSTRACION. Los ítems (i) y (ii) siguen directamente de las definiciones, asi que encargamos al lector sus demostraciones. Para probar (iii), asumamos primero que f y g son inyectivas y que $(g \circ f)(a) = (g \circ f)(a')$, donde $a, a' \in A$. Por definición de composición, esto significa que $g(f(a)) = g(f(a'))$, de donde deducimos que $f(a) = f(a')$, pues g es inyectiva. Puesto que también f es inyectiva, concluimos que $a = a'$, como queríamos probar.

Pasando al caso en que f y g son suryectivas, dado $c \in C$ podemos encontrar un elemento $b \in B$ y un elemento $a \in A$ tales que $g(b) = c$ y $f(a) = b$. Luego, $(g \circ f)(a) = g(f(a)) = g(b) = c$, y por lo tanto $g \circ f$ es suryectiva. Obviamente, la última afirmación sigue de las dos primeras.

Respecto de (iv), sean $x, x' \in A$ tales que $f(x) = f(x')$. En tal caso,

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

de donde sigue que $x = x'$, por hipótesis. Luego f es inyectiva.

Igualmente sencillo resulta probar (v). En efecto, siendo $g \circ f$ una función suryectiva, dado $y \in C$ existe $z \in A$ tal que $y = (g \circ f)(z) = g(f(z))$, y por lo tanto $y \in \mathrm{Im}(g)$. Resulta entonces que g es suryectiva. $\Diamond$

FUNCIONES INVERSIBLES. Sea $f : A \to B$ una función. Diremos que f es *inversible* si y sólo si existe una función $g : B \to A$ tal que $g \circ f = I_A$ y $f \circ g = I_B$. En términos de elementos, deben verificarse las relaciones

$$g(f(a)) = a \quad \text{y} \quad f(g(b)) = b$$

cualesquiera sean $a \in A$ y $b \in B$.

Es fácil probar que una tal función g, si existe, es única. En efecto, supongamos que g y g' son funciones de B en A satisfaciendo las condiciones de la definición. Tenemos entonces:

$$g' = g' \circ I_B = g' \circ (f \circ g) = (g' \circ f) \circ g = I_A \circ g = g.$$

Habiendo probado la unicidad de g, la llamaremos *función inversa* de f, y la designaremos por f^{-1}. Puesto que en la definición precedente los roles de una función y de su inversa son idénticos, es claro que f^{-1} también es inversible y que $\left(f^{-1}\right)^{-1} = f$.

Por ejemplo, resulta por verificación directa que la función $f : \mathbb{R} \to \mathbb{R}$ definida por

$$f(x) = \frac{x+3}{2}$$

es inversible, siendo $f^{-1}(x) = 2x - 3$. Consideremos en cambio las funciones

$$\gamma : \{1,2,3,4\} \to \{a,b,c,d,e\} \quad \text{y} \quad \delta : \{a,b,c,d,e\} \to \{1,2,3,4\}$$

definidas por las tablas

x	$\gamma(x)$
1	c
2	a
3	e
4	b

y

x	$\delta(x)$
a	3
b	2
c	1
d	2
e	4

Ninguna de las dos funciones es inversible. En efecto, si γ admitiera inversa, tendríamos $d = \gamma(\gamma^{-1}(d))$, de donde $d \in \mathrm{Im}(\gamma)$, lo que no es válido. De manera análoga, suponiendo la existencia de δ^{-1} obtendríamos $\delta^{-1}(2) = \delta^{-1}(\delta(b)) = b$, y también $\delta^{-1}(2) = \delta^{-1}(\delta(d)) = d$, obviamente una contradicción.

Los dos últimos ejemplos parecen indicar que una función inversible debe ser inyectiva y suryectiva. Probaremos a continuación la validez general de esta afirmación y de su recíproca.

Teorema 1.3.4 Una función $f : X \to Y$ es inversible si y sólo si es biyectiva.

DEMOSTRACION. Si f es inversible, supongamos que x y x' son elementos de X tales que $f(x) = f(x')$. Aplicando f^{-1} a esta igualdad obtenemos:

$$x = f^{-1}\left(f(x)\right) = f^{-1}\left(f(x')\right) = x'.$$

Luego f es inyectiva. Por otro lado, de la relación $y = f\left(f^{-1}(y)\right)$ sigue que $y \in \mathrm{Im}(f)$ para todo $y \in Y$, y por lo tanto f es suryectiva.

Recíprocamente, supongamos que f es biyectiva. Puesto que en una biyección cada elemento del codominio admite una única preimagen en el dominio, podemos definir una aplicación $h : Y \to X$ asignando a cada $y \in Y$ el único $x \in X$ tal que $f(x) = y$. En otras palabras, dados $u \in X$ y $v \in Y$ se verifica:

$$h(v) = u \Leftrightarrow f(u) = v.$$

Claramente, la relación anterior es una forma equivalente de expresar las igualdades funcionales $h \circ f = I_X$ y $f \circ h = I_Y$, y por lo tanto f es inversible, siendo $h = f^{-1}$. $\quad \Diamond$

Ejemplos 1.3.5 Veamos algunos otros ejemplos de funciones inversibles:

1) La aplicación h del ejemplo 1.3.2 es una biyección de $\mathbb{Z}$, por lo tanto es inversible. Más aún, puesto que probamos que $h\,(h(m)) = m\,\forall\, m \in \mathbb{Z}$, resulta que $h^{-1} = h$. Una tal función se dice una *involución*. Como otros ejemplos de involuciones mencionemos la biyección $x \mapsto -x$ de $\mathbb{Z}$ y la función identidad de cualquier conjunto.

2) Consideremos la función $t : \mathbb{N} \to \mathbb{Z}$ definida por:

$$t(n) = \begin{cases} \frac{n-1}{2} = t_1(n) & \text{si } n \text{ es impar} \\ -\frac{n}{2} = t_2(n) & \text{si } n \text{ es par.} \end{cases}$$

Probaremos que t es biyectiva, lo que nos mostrará que, en ciertos casos, es posible establecer una correspondencia biunívoca entre los elementos de un conjunto y los de un subconjunto propio del mismo.

Observando que $t(n) \geq 0$ si n es impar y $t(n) < 0$ si n es par, deducimos que t es inyectiva, ya que es trivial demostrar que cada una de las asignaciones $n \mapsto t_1(n)$ y $n \mapsto t_2(n)$ es uno a uno. En cuanto a la suryectividad, dado $a \in \mathbb{Z}$ sigue inmediatamente que $a = t(-2a)$ si $a < 0$ y $a = t(2a+1)$ si $a \geq 0$.

Habiendo probado que t es biyectiva (luego inversible), las relaciones anteriores nos muestran la forma de su inversa. Precisamente, t^{-1} es la aplicación de $\mathbb{Z}$ en $\mathbb{N}$ definida por:

$$t^{-1}(r) = \begin{cases} -2r & \text{si } r < 0 \\ 2r+1 & \text{si } r \geq 0. \end{cases}$$

3) Sea $\theta : \mathbb{R} - \{3\} \longrightarrow \mathbb{R}$ la función

$$\theta(x) = \frac{5x-14}{x-3}\ ,$$

y sean $u, v \in \mathrm{Dom}(\theta)$ tales que $\theta(u) = \theta(v)$. De la igualdad

$$\frac{5u-14}{u-3} = \frac{5v-14}{v-3}\ ,$$

operando elementalmente arribamos a la relación

$$5uv - 15u - 14v + 42 = 5uv - 14u - 15v + 42,$$

que claramente implica $v = u$. En consecuencia θ es inyectiva.

Sea ahora $b \in \mathbb{R}$, y veamos en qué casos es posible hallar $a \in \mathrm{Dom}(\theta)$ tal que $b = \theta(a)$. Planteando estas condiciones, tenemos:

$$b = \frac{5a-14}{a-3} \Leftrightarrow ab - 3b = 5a - 14 \Leftrightarrow ab - 5a = 3b - 14$$
$$\Leftrightarrow a(b-5) = 3b - 14.$$

Analizando esta última igualdad, vemos que b debe ser distinto de 5 (resultaría si no $0 = 1$), en cuyo caso

$$a = \frac{3b - 14}{b - 5} \ .$$

Observemos además que $a \in \mathrm{Dom}(\theta)$, ya que $a = 3$ implicaría

$$3b - 14 = 3b - 15 \, ,$$

obviamente una contradicción. Luego, $\mathrm{Im}(\theta) = \mathbb{R} - \{5\}$.

Deducimos entonces que la aplicación

$$\theta_0 : \mathbb{R} - \{3\} \longrightarrow \mathbb{R} - \{5\}$$

obtenida correstringiendo θ a su imagen es inyectiva (por serlo θ) y suryectiva. Por lo tanto es inversible, siendo su inversa la función

$$\theta_0^{-1} : \mathbb{R} - \{5\} \longrightarrow \mathbb{R} - \{3\}$$

definida por la fórmula

$$\theta_0^{-1}(x) = \frac{3x - 14}{x - 5} \ .$$

4) Si X es un conjunto, veamos que existe una correspondencia biunívoca entre el conjunto de funciones de X en $\{0, 1\}$, que notaremos $\{0, 1\}^X$, y el conjunto de partes de X. Consideremos para ello la aplicación

$$\mu : \{0, 1\}^X \to \mathbb{P}(X)$$

definida por

$$\mu(f) = \{x \in X \ : \ f(x) = 1\} \, .$$

Probaremos que μ es biyectiva, exhibiendo su inversa. Para ello, hagamos corresponder a cada $S \in \mathbb{P}(X)$ la aplicación $\alpha_S : X \to \{0, 1\}$ dada por

$$\alpha_S(x) = \begin{cases} 1 & \text{si } x \in S \\ 0 & \text{si } x \notin S \, . \end{cases}$$

Formalmente hablando, hemos definido una función

$$\nu : \mathbb{P}(X) \to \{0, 1\}^X \, ,$$

a través de la asignación $S \overset{\nu}{\to} \alpha_S$. Veamos que μ y ν son mutuamente inversas.

Tomemos en primer término $f \in \{0, 1\}^X$ y sea $T = \mu(f)$. Resulta entonces que

$$\{x \in X \ : \ f(x) = 1\} = T = \{x \in X \ : \ \alpha_T(x) = 1\} \, ,$$

por definición de ν, lo que implica que $\alpha_T = f$, pues ambas funciones sólo toman los valores 0 y 1. Por lo tanto, $\nu \circ \mu$ es la función identidad de $\{0,1\}^X$.

En cuanto a la otra composición, sea S un subconjunto de X. Entonces:

$$\mu\left(\nu(S)\right) = \{x \in X \,:\, \alpha_S(x) = 1\} = S,$$

esto es, $\mu \circ \nu = I_{\mathbb{P}(X)}$. Luego $\nu = \mu^{-1}$ y nuestra afirmación queda demostrada. $\Diamond$

1.3.3. Ejercicios

1. Decidir cuáles de las siguientes relaciones de A en B son funciones. En tales casos, determinar la imagen y analizar la inyectividad y suryectividad de la función:

   a) $A = \{1,2,3,4\}, B = \{a,b,c,x,2\}$;
   $\Re = \{(1,b),(2,c),(3,2),(4,c)\}$.

   b) A y B como en *a)*; $\Re = \{(1,a),(3,a),(4,b),(5,2)\}$.

   c) A y B como en *a)*; $\Re = \{(2,a),(1,2),(3,b),(2,d),(4,c),(5,c)\}$.

   d) $A = B = \mathbb{N}$; $a \,\Re\, b \Leftrightarrow b - a = 1$.

   e) $A = B = \mathbb{N}$; $a \,\Re\, b \Leftrightarrow a = b + 1$.

   f) $A = B = \mathbb{P}(\mathbb{N})$; $S \,\Re\, T \Leftrightarrow T = S \cap \{1,3,5\}$.

   g) $A = B = \mathbb{R}$; $\Re = \left\{(a,b) \in \mathbb{R}^2 : ab = 1\right\}$.

   h) $A = B = \mathbb{Z}$; $a \,\Re\, b \Leftrightarrow b - a^2 = 4a$.

   i) $A = B = \mathbb{R}$; $a \,\Re\, b \Leftrightarrow a = b^2 - 4$.

2. Analizar la inyectividad y suryectividad de las siguientes funciones. Hallar la inversa de aquellas que sean biyectivas:

   a) $f_1 : \mathbb{N} \to \mathbb{N}$; $f_1(n) = \begin{cases} \dfrac{n+3}{2} & \text{si } n \text{ es impar} \\ n-1 & \text{si } n \text{ es par}. \end{cases}$

   b) $f_2 : \mathbb{N} \to \mathbb{N}$; $f_2(n) = \begin{cases} 2n-3 & \text{si } n \text{ es par} \\ n+1 & \text{si } n \text{ es impar}. \end{cases}$

   c) $f_3 : \mathbb{R} \to \mathbb{R}$; $f_3(x) = 4x - 1$.

   d) $f_4 : \mathbb{N} \to \mathbb{Z}$; $f_4(n) = (-1)^n n + 3$.

   e) $f_5 : \mathbb{Z} \to \mathbb{Z}$; $f_5(n) = (-1)^n n + 3$.

   f) $f_6 : \mathbb{R}^2 \to \mathbb{R}$; $f_6\left((x,y)\right) = 2x + y$.

   g) $f_7 : \mathbb{R} - \{5\} \to \mathbb{R}$; $f_7(x) = \dfrac{2x+1}{x-5}$.

h) $f_8 : \mathbb{R} \to \mathbb{R}^2$; $f_8(x) = (3x, x^2 + 1)$.

i) $f_9 : \mathbb{Z} \to \mathbb{Z}$; $f_9(m) = m + (-1)^m$.

j) $f_{10} : \mathbb{Z} \to \mathbb{N}$; $f_{10}(m) = \begin{cases} 2m & \text{si } m > 0 \\ 1 - 2m & \text{si } m \leq 0 \, . \end{cases}$

k) $f_{11} : \mathbb{R} \to \mathbb{R}$; $f_{11}(x) = x^2 + 6x + 4$.

3. Sean f_1 y f_2 como en el ejercicio 2.

   a) Calcular $(f_1 \circ f_2)(5)$, $(f_1 \circ f_2)(12)$ y $(f_2 \circ f_1)(9)$.

   b) Hallar fórmulas generales para $(f_2 \circ f_1)(n)$ y $(f_1 \circ f_2)(n)$.

   c) ¿ Es inversible alguna de las dos funciones ?

4. Respecto de las funciones del ejercicio 2, hallar fórmulas para $f_3 \circ f_6$, $f_6 \circ f_8$, $f_8 \circ f_6$, $f_{11} \circ f_7$, $f_9 \circ f_9$, $f_2 \circ f_{10}$, $f_4 \circ f_{10}$, $f_{10} \circ f_5$, $f_{11} \circ f_3$ y $f_5 \circ f_5$.

5. Consideremos las funciones f, g y h de $\mathbb{R}$ en $\mathbb{R}$ definidas por las fórmulas $f(x) = ax + b$, $g(x) = cx + d$ y $h(x) = rx^2 + sx + t$, donde las letras designan números reales, siendo a, c y r no nulos.

   a) Demostrar que $f \circ h = h \circ f$ si y solo si $f = I_{\mathbb{R}}$.

   b) ¿ Es cierto en general que $f \circ g = g \circ f$?

   c) Hallar a, b, c y d no nulos tales que $f \circ g = g \circ f$.

6. Sea $\sigma : \mathbb{N} \to \mathbb{N}$ una función biyectiva tal que $\sigma(n + 1) = \sigma(n) + 1$ para todo $n \in \mathbb{N}$. Determinar σ .

7. En lo que sigue las letras mayúsculas denotan conjuntos y las minúsculas funciones:

   a) Sean $f : A \to B$ y $g : B \to C$ tales que $g \circ f$ es biyectiva. Demostrar que g es biyectiva si y solo si f lo es.

   b) Sean $f : A \to B$ y $g : B \to A$ tales que f es suryectiva y $g \circ f = I_A$. Probar que f es inversible y que $f^{-1} = g$.

   c) Sean f y g biyecciones de A tales que $g \circ f = f \circ g$. Probar que $g^{-1} \circ f^{-1} = f^{-1} \circ g^{-1}$.

8. En cada uno de los siguientes incisos se define una función de un conjunto A en un conjunto B. Determinar en todos los casos subconjuntos A' de A y B' de B , tan "grandes" como sea posible (en el sentido de la inclusión), de manera que $f|_{A'} : A' \to B'$ sea biyectiva:

a) $f : \mathbb{R} \to \mathbb{R}\,;\ f(x) = x^2 + 6x + 1\,.$

b) $f : \mathbb{N} \to \mathbb{N}\,;\ f(n) = \begin{cases} n - 6 & \text{si } n \geq 20 \\ 3n - 1 & \text{si } n < 20 \end{cases}\,.$

c) $f : \mathbb{N}^2 \to \mathbb{N}\,;\ f\left((m, n)\right) = m + n\,.$

d) $f : \mathbb{P}\left(\{1, 2, \cdots, 20\}\right) - \emptyset \longrightarrow \mathbb{Z}\,;\ f(S) = s - 1\,,$ donde s es el menor elemento de $S\,.$

9. Sean A y B conjuntos y sea f una función de A en $B\,.$

a) Probar que la relación

$$a \sim a' \Leftrightarrow f(a) = f(a')$$

es una relación de equivalencia en $A\,.$

b) Probar que la asignación

$$Cl(a) \longmapsto f(a)$$

determina correctamente una función $\hat{f} : A/\sim \to B\,,$ en el sentido de que el valor $f(a)$ asignado no depende del representante a elegido. Demostrar además que $\hat{f}$ es inyectiva.

c) Probar que si f es suryectiva entonce $\hat{f}$ es una biyección.

1.4. Operaciones binarias

1.4.1. Definicion

En páginas previas hemos empleado frecuentemente la expresión "operación binaria", para referirnos a diversas maneras de asociar a cada par de elementos de un conjunto otro elemento del conjunto. Por ejemplo, las operaciones conjuntísticas (unión, intersección, etc.), la composición de funciones, y por supuesto las familiares operaciones numéricas, como la adición, la multiplicación, la división, etc. Para abstraer el concepto, definamos formalmente la noción de operación binaria en un conjunto arbitrario.

Si A es un conjunto, cualquier aplicación

$$\vartheta : A \times A \to A$$

se dirá una *operación binaria* en A.

Notemos que esta definición responde a la idea que comentábamos arriba. Una operación hace corresponder a cada par ordenado (x, y) de elementos de A un único elemento $\vartheta\left((x, y)\right)$ de A, que suele llamarse el resultado de la operación.

Puesto que una operación es una función, puede describírsela en diversas formas. Por ejemplo, si el conjunto A consiste de unos pocos elementos es cómodo hacerlo a través de una tabla de doble entrada, con una fila y una columna para cada elemento de A. En dicha tabla se escribe, en el lugar correspondiente a la fila x y columna y, el resultado $\vartheta(x, y)$. Asímismo, es costumbre en la práctica usar otro tipo de símbolos para denotar tal resultado, tales como $x + y$, $x.y$, $x \times y$, $x \oplus y$, ..., ó cualquier otro que juzguemos adecuado, incluso la simple yuxtaposición xy.

Ejemplos 1.4.1 En los siguientes incisos, empleamos el símbolo $*$ para designar genéricamente una operación binaria, indicando en cada caso el conjunto A de definición:

1) $A = \mathbb{N}$; $x * y = \begin{cases} y - x + 1 & \text{si } x \leq y \\ x - y & \text{si } x > y\,. \end{cases}$

2) $A = \mathbb{Z}$; $x * y = x + y - xy$.

3) A el conjunto de funciones de $\mathbb{N}$ en $\mathbb{N}$; $f * g = f \circ g \circ f$.

4) $A = \{1, 2, 3, 4, 5, 6\}$ y la operación definida por la tabla

$*$	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	1	6	5	4	3
3	3	5	1	6	2	4
4	4	6	5	1	3	2
5	5	3	4	2	6	1
6	6	4	2	3	1	5

5) $A = \{u, v, w, z\}$ y la operación definida por la tabla

$*$	u	v	w	z
u	u	v	w	z
v	v	u	z	w
w	w	z	u	v
z	z	w	v	u

$\Diamond$

1.4.2. Propiedades básicas

En la primera sección del capítulo hemos señalado ciertas propiedades de las operaciones de conjuntos, que también satisfacen algunas de las operaciones numéricas. En lo que sigue, enunciaremos dichas propiedades en el caso general de una operación binaria $*$ definida en un conjunto cualquiera A:

ASOCIATIVIDAD Diremos que $*$ es *asociativa* si y sólo si

$$x * (y * z) = (x * y) * z$$

cualesquiera sean los elementos $x, y, z \in A$.

Ejemplos 1.4.2 El lector puede verificar sin mayor dificultad que las operaciones de los ítems 2) a 5) del ejemplo 1.4.1 son asociativas. En cambio la primera no lo es, ya que $(3*1)*5 = 2*5 = 4$, mientras que $3*(1*5) = 3*5 = 3$. Otros ejemplos de operaciones asociativas son la suma y el producto de números reales, la intersección, unión y diferencia simétrica de conjuntos y la composición de funciones. Mencionemos finalmente la diferencia y el cociente como ejemplos de operaciones numéricas no asociativas. $\Diamond$

CONMUTATIVIDAD Diremos que $*$ es *conmutativa* si y sólo si

$$x * y = y * x$$

cualesquiera sean los elementos $x, y \in A$.

Ejemplos 1.4.3 Similarmente a lo que ocurre con la asociatividad, la suma y el producto en los conjuntos numéricos son operaciones conmutativas, y no lo son la diferencia y el cociente. Como ya vimos, la composición de funciones y la diferencia de conjuntos son también ejemplos de operaciones no conmutativas. Respecto del ejemplo 1.4.1, sólo son conmutativas las operaciones de los casos 2) y 5). Por ejemplo, en el caso 1) tenemos $4 = 2 * 5 \neq 5 * 2 = 3$. Encargamos al lector hallar un contraejemplo para la operación del ítem 3), mientras que podrá deducir que la operación del ítem 4) no es conmutativa por simple inspección de la tabla. $\Diamond$

EXISTENCIA DE ELEMENTO NEUTRO Si $e \in A$, diremos que e es un *elemento neutro* respecto de $*$ si y sólo si

$$e * x = x * e = x$$

para todo $x \in A$.

Es muy sencillo probar en tal caso que e es único, ya que si $e' \in A$ también es un elemento neutro respecto de $*$ resulta

$$e = e * e' = e'.$$

Variantes más débiles de la definición anterior son las siguientes: e se dice un *elemento neutro a derecha* respecto de $*$ si y sólo si $x * e = x$ para todo $x \in A$, y un *elemento neutro a izquierda* respecto de $*$ si y sólo si $e * x = x$ para todo $x \in A$. Obviamente las tres definiciones coinciden si la operación es conmutativa.

Ejemplos 1.4.4 Ejemplos sin duda conocidos por el lector son los elementos neutros de la suma y el producto de números reales (0 y 1, respectivamente), mientras que si X es un conjunto cualquiera, ya vimos que $\emptyset$ es el elemento neutro respecto de la unión y de la diferencia simétrica en $\mathbb{P}(X)$, siendo X el elemento neutro respecto de la intersección. Volviendo a 1.4.1, no hay dificultad en verificar que las operaciones de los ítems 2), 4) y 5) admiten elemento neutro, siendo los mismos 0, 1 y u, respectivamente. En el caso 1), se ve fácilmente que 1 es un elemento neutro a izquierda, mientras que no existe elemento neutro a derecha. En efecto, supongamos que m lo fuera. Entonces:

$$m + 1 = (m + 1) * m = m + 1 - m = 1,$$

lo que es absurdo, pues $m \in \mathbb{N}$. En forma similar se prueba que $I_{\mathbb{N}}$ es un elemento neutro a izquierda respecto de la operación del ítem 3) y que ésta no admite elemento neutro a derecha. $\Diamond$

EXISTENCIA DE INVERSO Si $*$ admite elemento neutro e, un elemento $a \in A$ se dice *inversible* si y sólo si existe $b \in A$ tal que

$$a * b = b * a = e.$$

Diremos en tal caso que b es un *inverso* de a. Por simetría de los roles, es claro que entonces b también es inversible, siendo a un inverso de b. Notemos además que A admite por lo menos un elemento inversible, a saber el elemento neutro e (que coincide con su inverso).

Si $*$ es asociativa y a admite inverso, entonces éste es único. En efecto, supongamos que b' es un elemento de A satisfaciendo las mismas condiciones que b. Entonces:

$$b' = b' * e = b' * (a * b) = (b' * a) * b = e * b = b.$$

Más generalmente, diremos que a admite *inverso a derecha* (*a izquierda*) si y sólo si existe $b \in A$ tal que $a * b = e$ ($b * a = e$).

Digamos a titulo informativo que un conjunto G en el que está definida una operación binaria asociativa, que admite elemento neutro y en el cual todo elemento es inversible, se dice un *grupo* (con respecto a dicha operación). Si $S \subseteq G$ y la restricción de la operación a S también define una estructura de grupo en S, diremos que S es un *subgrupo* de G.

Ejemplos 1.4.5 Como estableceremos en detalle en el próximo capítulo, todo número real x admite inverso respecto a la suma, (se lo nota $-x$), y todo número real x distinto de 0 admite inverso respecto al producto (designado por x^{-1}). En el caso de las operaciones conjuntísticas definidas en las partes de un conjunto no vacío X, todo $S \subseteq X$ es su propio inverso respecto a la diferencia simétrica, $\emptyset$ es el único elemento inversible respecto a la unión y X es el único elemento inversible respecto a la intersección.

Otro ejemplo interesante ya estudiado es el de la operación composición, definida en el conjunto de funciones de un conjunto cualquiera C en sí mismo. Como vimos, en ese caso los elementos inversibles son exactamente las biyecciones de C.

En la situación del ítem 2) del ejemplo 1.4.1 los únicos elementos inversibles son 0 (el neutro) y 2, y ambos coinciden con su inverso. En los casos 4) y 5) todo elemento es inversible, como se advierte con sólo mirar la tabla. Por ejemplo, 5 y 6 son mutuamente inversos en el ítem 4), y en el 5) todo elemento coincide con su inverso. $\Diamond$

CANCELATIVIDAD Diremos que $*$ es *cancelativa a derecha* (resp. *a izquierda*) si y sólo si

$$x * u = y * u \implies x = y \quad (u * x = u * y \implies x = y)$$

cualesquiera sean $x, y, u \in A$. Diremos que $*$ es *cancelativa* si y sólo si lo es a derecha y a izquierda.

Si $*$ es asociativa y admite elemento neutro e, cualquier elemento inversible a derecha u puede cancelarse en una igualdad del tipo $x * u = y * u$. En efecto, si v es un inverso (a derecha) de u, tenemos:

$$x = x * e = x * (u * v) = (x * u) * v = (y * u) * v = y * (u * v) = y * e = y.$$

Es claro que se llega a una conclusión similar si u es inversible a izquierda y vale una relación de la forma $u * x = u * y$.

Ejemplos 1.4.6 Se deduce del comentario anterior que toda operación asociativa con elemento neutro respecto de la cual todo elemento es inversible es cancelativa. Por caso, las operaciones de los incisos 4) y 5) del ejemplo 1.4.1. Claro que la hipótesis de que todo elemento tenga inverso es suficiente pero no necesaria, ya que por ejemplo la operación de suma en $\mathbb{N}$ es cancelativa, a pesar de que ningún número natural admite inverso aditivo.

Observemos finalmente que las operaciones de los incisos 1) y 2) de 1.4.1 no son cancelativas, ya que $5 * 6 = 5 * 3 = 2$ y $3 * 5 = 8 * 5 = 3$ en 1), mientras que $x * 1 = y * 1$ cualesquiera sean $x, y \in \mathbb{Z}$ en 2). Encargamos al lector estudiar la situación en el inciso 3) de tal serie de ejemplos. $\Diamond$

1.4.3. Ejercicios

1. Sea $*$ una operación asociativa con elemento neutro definida en un conjunto A. Probar:

   a) Si cada elemento de A admite un inverso a derecha entonces todo elemento de A es inversible.

   b) Si la ecuación $a * x = b$ admite exactamente una solución en A cualesquiera sean a y b en A entonces todo elemento de A es inversible.

2. En cada uno de los siguientes ítems, estudiar la asociatividad, conmutatividad y cancelatividad de la operación $*$ definida en el conjunto A. En todos los casos analizar también la existencia de elementos neutros (a derecha y a izquierda) y de inversos (a derecha y a izquierda):

   a) $A = \mathbb{R}$; $a * b = a + b + ab$.

   b) La operación definida en el ítem 1) del ejemplo 1.4.1.

   c) A el conjunto de funciones de $\mathbb{N}$ en $\mathbb{N}$; $f * g = f \circ f \circ g$.

   d) $A = \mathbb{P}(\mathbb{N})$; $S * T = (S \triangle T) \triangle \{2, 3, 6\}$.

   e) $A = \mathbb{N}$; $a * b = \text{máx}(a, b)$.

   f) $A = \mathbb{R} - \{0\}$; $a * b = ab^{-1}$.

   g) $A = \mathbb{Q}^2$; $(a, b) * (c, d) = (ac, b + d)$.

Capítulo 2

El cuerpo de los números reales

2.1. Números reales

2.1.1. Introducción

Antes de entrar de lleno en el tratamiento que daremos al concepto de número real, haremos una breve recorrida por los diversos conjuntos numéricos y sus significados. Más que dar definiciones precisas de los diversos tipos de números, haremos una presentación intuitiva y cronológica de los mismos, intentando clarificar y resumir las nociones que el lector ha adquirido en sus estudios previos. Naturalmente, asumimos que conoce las operaciones elementales, y que no le son totalmente desconocidos términos como "infinito", "desarrollo decimal", "raíz cuadrada", etc.

El concepto primitivo de número corresponde sin duda al de número natural. Intuitivamente, son aquellos números que utiliza el hombre para contar o enumerar objetos de diversa índole. A pesar de su familiaridad, notemos que la idea de número encierra una abstracción. Por ejemplo, el número 3 es lo que tienen en común los conjuntos $\{a, b, c\}$ y $\{\text{Ana}, \text{Juan}, \text{Pedro}\}$. Este es el concepto original de número natural, el de representante ideal de todos los conjuntos que tienen una misma cantidad de elementos.

Lo anterior responde a la noción aún más primordial de unidad. La misma está simbolizada por el número 1, el primer número natural. La sencilla idea de que un conjunto es una colección de unidades se refleja matemáticamente en el hecho de que todo número natural se obtiene sumando 1 una cierta cantidad de veces. Así, la operación más simple que conocemos −la adición o suma de números naturales−, está directamente conectada con esta característica esencial de los mismos: sumar dos magnitudes naturales no es otra cosa que sumar reiteradas veces la unidad.

Es claro que la operación de sumar tiene su origen en la necesidad práctica de agregar cantidades. Cuando la cantidad a sumar es muchas veces la misma, aparece naturalmente la operación de multiplicación o producto de

números naturales. Así, multiplicar m por n no es otra cosa que sumar m veces n. En otras palabras, el producto de números naturales puede describirse a través de la suma.

Es claro que conocemos otra operación entre números naturales, la sustracción o resta. En realidad se deriva de la suma, ya que restar m de n consiste en resolver la ecuación $x + m = n$, que como sabemos puede no tener solución natural (se requiere que m sea menor que n). De lo anterior surge que la resta no es estrictamente hablando una operación entre números naturales, ya que no siempre puede efectuarse. Sin embargo, no es difícil imaginar hechos del quehacer humano que requieran la resolución de una ecuación como la anterior, por ejemplo el activo y el pasivo de una situación financiera. Esta necesidad condujo históricamente a la introducción de otra clase de números, los negativos u opuestos de los números naturales, así como el cero (ausencia de cantidad). Este nuevo conjunto ampliado de números se llama conjunto de los números enteros. Más aún, las operaciones de suma y producto de números naturales se extienden satisfactoriamente al conjunto de números enteros, en el sentido de que conservan sus propiedades básicas, amén de agregarse otras que permiten trabajar con mayor fluidez en el nuevo conjunto numérico.

Vayamos ahora al encuentro de otra operación bien conocida, como es la división o cociente. Primero debemos precisar qué entendemos por dividir, digamos dos números enteros. Si la pensamos como una operación inversa de la multiplicación, dividir n por m correspondería a encontrar un número entero q (el cociente) tal que $n = qm$. Por ejemplo, el cociente de dividir 96 por 8 es 12. Empero, dicho q no existe en general. Por ejemplo, no existe un número entero q tal que $3q = 7$. Por lo tanto, ninguna cantidad entera representa la tercera parte de 7. Esta necesidad práctica de representar una cantidad fraccionada en partes iguales llevó naturalmente a la consideración de un nuevo tipo de números, los números racionales, llamados así por ser razón o cociente de dos números enteros. Esto es, un número racional es de la forma m/n, donde m y n son números enteros ($n \neq 0$) y puede interpretarse como la única solución de la ecuación $nx = m$. Esta caracterización permite definir sin dificultad la suma y el producto de números racionales, que resultan ser también números racionales. Puesto que todo entero m es racional ($m = m/1$), el nuevo conjunto numérico extiende al de los enteros y sus operaciones conservan las propiedades que verificaban en el conjunto de números enteros.

Sin entrar por ahora en mayores detalles, observemos la similitud de ambos procesos de ampliación. En el caso de los enteros se agregan los opuestos de los números naturales, y en el de los racionales los inversos $1/n$ de los enteros no nulos. En las dos situaciones se logran objetivos análogos: la resta, que no siempre es posible entre números naturales es siempre factible entre números enteros, mientras que la división (que no es realizable en general entre números enteros), sí puede efectuarse siempre entre números racionales (exceptuado el caso de la división por cero, claro).

Hasta la época de los griegos sólo se conocían los números racionales. Por supuesto, nadie se puso a inventar nuevos números, sino que estos fueron apareciendo naturalmente ante alguna necesidad práctica que resolver. Por ejemplo, supongamos que queremos calcular la longitud de la diagonal de un cuadrado de lado 1. Una simple aplicación del teorema de Pitágoras muestra que dicha longitud x −que esperamos esté representada por un número positivo−, debe satisfacer la relación $x^2 = 2$. Como veremos más adelante, es fácil demostrar que ningún número racional la satisface. Aparecen así cantidades (en este caso la que conocemos como raíz cuadrada de 2) cuya existencia responde a situaciones concretas y que no pueden ser representadas por la razón de dos números enteros, motivo por el cual los números que las idealizan se llaman irracionales.

El conjunto de ambas clases de números, el de los racionales e irracionales se llama conjunto de números reales. Ya veremos más generalmente que la raíz cuadrada de n es irracional para todo natural n que no sea el cuadrado de otro número natural, lo que muestra que existen infinitos números irracionales. En realidad, en un sentido que no precisaremos aquí, hay "muchos más" números irracionales que racionales, y por cierto que no todos son del tipo indicado arriba, como por ejemplo el número π, que expresa la razón constante entre la longitud de una circunferencia y su diámetro.

Para finalizar esta breve introducción, cabría preguntarse cómo explicar sencillamente qué es un número real. Esto es algo más complicado que en el caso de los naturales, enteros y racionales, que admiten descripciones más o menos simples. Como veremos más adelante, esto puede hacerse a través del desarrollo decimal, que representa unívocamente cada número real por una secuencia infinita de números enteros, pero por ahora nos despreocuparemos del problema, aceptaremos que existe un conjunto que llamaremos de los números reales, y nos dedicaremos a señalar sus propiedades estructurales fundamentales. Ellas constituirán nuestras reglas de juego (axiomas), y tomándolas como base de partida iremos construyendo de manera precisa los diversos conjuntos numéricos que hemos descripto en una forma un tanto vaga en los párrafos anteriores.

2.1.2. Estructura de cuerpo de los números reales

Designaremos el conjunto de *números reales* con la letra $\mathbb{R}$. Supondremos definidas en él dos operaciones binarias, que llamaremos *suma* y *producto*, las cuales satisfacen las siguientes propiedades:

Propiedades de la suma.

S_1) Asociatividad

$$x + (y + z) = (x + y) + z, \text{ cualesquiera sean } x, y, z \in \mathbb{R}.$$

S_2) Conmutatividad

$$x + y = y + x, \text{ cualesquiera sean } x, y \in \mathbb{R}.$$

S_3) Existencia de elemento neutro

$$x + 0 = x \text{ cualquiera sea } x \in \mathbb{R},$$

donde 0 es un número real que llamamos *cero*.

S_4) Existencia de inverso aditivo

$$\text{Si } x \in \mathbb{R} \text{ existe } y \in \mathbb{R} \text{ tal que } x + y = 0.$$

El elemento y será notado $-x$ y lo llamaremos el *inverso aditivo* de x.

De aquí en más, designaremos por $\mathbb{R}^*$ el conjunto $\mathbb{R} - \{\, 0 \,\}$.

Propiedades del producto.

P_1) Asociatividad

$$x(yz) = (xy)z, \text{ cualesquiera sean } x, y, z \in \mathbb{R}.$$

P_2) Conmutatividad

$$xy = yx, \text{ cualesquiera sean } x, y \in \mathbb{R}.$$

P_3) Existencia de elemento neutro

$$x\,1 = x \text{ cualquiera sea } x \in \mathbb{R},$$

donde 1 es un número real distinto de 0 que llamamos *uno*.

P_4) Existencia de inverso multiplicativo

$$\text{Si } x \in \mathbb{R}^* \text{ existe } y \in \mathbb{R} \text{ tal que } xy = 1.$$

El elemento y será notado x^{-1} y lo llamaremos el *inverso multiplicativo* de x.

Propiedad distributiva.

D) Distributividad del producto con respecto a la suma

$$x(y + z) = xy + xz, \text{ cualesquiera sean } x, y, z \in \mathbb{R}\,.$$

Nota En general, un conjunto dotado de dos operaciones binarias (denominadas genéricamente suma y producto) que satisfacen las nueve propiedades anteriores se llama un *cuerpo*. Debido a ello, nos referiremos de aquí en más al cuerpo de los números reales. Vale la pena hacer un par de comentarios respecto de los axiomas que postulan la existencia de inversos. Como vimos

en ocasión de ocuparnos en general de las propiedades de una operación binaria, dichos inversos son únicos, ya que tanto la suma como el producto son operaciones asociativas. Además, como también hicimos notar anteriormente, la relación de un elemento con su inverso es simétrica, resultando que $-(-x) = x$ y $(x^{-1})^{-1} = x$. $\Diamond$

Dada su importancia, destacaremos especialmente la siguiente propiedad del producto:

Proposición 2.1.1 Sean $x, y \in \mathbb{R}$. Entonces

$$xy = 0 \iff x = 0 \text{ ó } y = 0.$$

DEMOSTRACION. Para probar que la condición es suficiente, supongamos por ejemplo que $x = 0$. Tenemos entonces:

$$0\,y = (0+0)y = 0\,y + 0\,y.$$

Sumando $-0\,y$ a ambos miembros de la igualdad anterior, obtenemos:

$$0 = -0\,y + 0\,y = -0\,y + (0\,y + 0\,y) = (-0\,y + 0\,y) + 0\,y = 0 + 0\,y = 0\,y,$$

como queríamos demostrar.

Recíprocamente, sea $xy = 0$ y supongamos que $x \neq 0$. Multiplicando por x^{-1} resulta que

$$0 = x^{-1}\,0 = x^{-1}(xy) = (x^{-1}x)y = 1\,y = y\,,$$

esto es, alguno de los dos factores debe ser 0. $\Diamond$

La proposición anterior afirma que el producto de números reales no nulos es no nulo, o dicho en otros términos, el conjunto $\mathbb{R}^*$ es cerrado para el producto. El lector podrá apreciar cómo prácticamente hemos usado todas las propiedades S_1) a D). En los ejercicios del final de la sección, encomendaremos al lector la tarea de demostrar, a partir de estas reglas básicas, otras propiedades familiares de las operaciones numéricas. Algunos hechos quizás le parezcan muy obvios, pero queremos que se vaya acostumbrando gradualmente a no temerle a las demostraciones y al correcto empleo de los axiomas. Por otro lado, es importante que conozca bien todas esas propiedades para operar sin inconvenientes.

Resta y cociente de números reales.

Nos hemos referido hasta aquí a la suma y al producto de números reales, dejando de lado dos de las llamadas operaciones elementales, como son la resta y cociente. La razón es que ambas pueden describirse en términos de las otras dos, y como consecuencia de ello sus propiedades se derivan fácilmente

de correspondientes propiedades de la suma y el producto. Pasemos entonces a definirlas.

RESTA. Si $x, y \in \mathbb{R}$, definimos su *resta* o *diferencia* en la forma

$$x - y = x + (-y).$$

COCIENTE. Si $x, y \in \mathbb{R}$ $(y \neq 0)$, definimos su *cociente* en la forma

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

En ocasiones también usaremos la notación x/y para referirnos al cociente.

El lector podrá advertir la analogía existente entre ambas definiciones: se resta sumando el inverso aditivo y se divide multiplicando por el inverso multiplicativo, resultando que $x - y = z$ si y solo si $x = y + z$ y $x/y = z$ si y solo si $x = yz$. Esto último explica la imposibilidad de dividir por 0, ya que $0\,z = 0$ cualquiera sea $z \in \mathbb{R}$.

En general, la diferencia y el cociente no conservan las propiedades estructurales de la suma y el producto, por ejemplo, no son ni asociativas ni conmutativas. Observemos además que de la definición de cociente sigue que $a^{-1} = 1/a$, que es la forma alternativa tradicional de designar el inverso multiplicativo de un número real no nulo.

2.1.3. Ejercicios

Para obtener las demostraciones requeridas en los siguientes ejercicios, el lector deberá utilizar los axiomas enunciados en el texto o propiedades previamente demostradas. Las letras designan números reales (no nulos cuando corresponda), y como es usual, x^2 indica el producto $x\,x$.

1. Demostrar las siguientes propiedades:

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

   b) $-ab = (-a)b = a(-b)$

   c) $(-1)a = -a$

   d) $(-a)(-b) = ab$

   e) $(a - b)c = ac - bc$

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

   g) $-\dfrac{a}{b} = \dfrac{-a}{b} = \dfrac{a}{-b}$

   h) $(ab)^{-1} = a^{-1}b^{-1}$

   i) $(ab)^2 = a^2b^2$.

2. Demostrar las siguientes propiedades:

   a) $\dfrac{a}{b}+\dfrac{c}{d}=\dfrac{ad+bc}{bd}$

   b) $\dfrac{a}{c}+\dfrac{b}{c}=\dfrac{a+b}{c}$

   c) $\dfrac{a}{b}\,\dfrac{c}{d}=\dfrac{ac}{bd}$

   d) $\left(\dfrac{a}{b}\right)^{-1}=\dfrac{b}{a}$

   e) Si $u=\dfrac{a}{b}$ y $v=\dfrac{c}{d}$ entonces $\dfrac{u}{v}=\dfrac{ad}{bc}$.

3. Demostrar las siguientes equivalencias:

   a) $a+c=b+c \Leftrightarrow a=b$

   b) $-a=-b \Leftrightarrow a=b$

   c) $a^{-1}=b^{-1} \Leftrightarrow a=b$

   d) $a^2=b^2 \Leftrightarrow a=b$ ó $a=-b$

   e) $ac=bc \Leftrightarrow c=0$ ó $a=b$

   f) $a/b=c/d \Leftrightarrow ad=bc$

   g) $a/b=a/c \Leftrightarrow a=0$ ó $b=c$

   h) $a/b=c/b \Leftrightarrow a=c$.

4. Analizar la resolubilidad en $\mathbb{R}$ de las siguientes ecuaciones:

   a) $(x+y)^2=x^2+y^2$

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

   c) $\dfrac{1}{x+y}=\dfrac{1}{x}+\dfrac{1}{y}$

   d) $x+yz=(x+y)(x+z)$

   e) $\dfrac{x}{y}=\dfrac{y}{z}=\dfrac{z}{x}$.

5. Comprobar que las operaciones de resta y cociente de números reales no son ni asociativas ni conmutativas.

2.2. Orden

2.2.1. La relación de orden en $\mathbb{R}$

Además de las operaciones que definen la estructura algebraica del conjunto de números reales, el lector conoce una manera de compararlos, en el sentido de poder decidir si un número es más "grande" o más "chico" que otro.

Estrictamente hablando, supondremos definida en $\mathbb{R}$ una relación de orden (es decir, una relación reflexiva, antisimétrica y transitiva), que notaremos con el símbolo $\leq$. Así, la expresión $x \leq y$ se leerá "x es menor o igual que y". También, para indicar que $x \leq y$ y $x \neq y$ escribiremos $x < y$ ("x es menor que y"). Notaciones y terminologías alternativas para ambas situaciones serán $y \geq x$ ("y es mayor o igual que x") e $y > x$ ("y es mayor que x"). Como casos particulares distinguidos, diremos que un número real x es *positivo* si y sólo si $x > 0$, y *negativo* si y solo si $x < 0$. Asimismo, emplearemos las notaciones $\mathbb{R}_{>0}$ y $\mathbb{R}_{\geq 0}$ para referirnos a los conjuntos de números reales positivos y números reales no negativos, respectivamente.

Agregaremos a nuestro sistema axiomático de definición de los números reales las siguientes propiedades fundamentales del orden. Debido a la validez de las mismas, diremos que $\mathbb{R}$ es un *cuerpo ordenado*.

Propiedades del orden.

O_1) Tricotomía del orden Dados $a, b \in \mathbb{R}$, se verifica una y sólo una de las siguientes relaciones:

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

O_2) Compatibilidad del orden con respecto a la suma

$$\text{Si } a < b \text{ y } c \in \mathbb{R} \text{ entonces } a + c < b + c\,.$$

O_3) Compatibilidad del orden con respecto al producto

$$\text{Si } a < b \text{ y } c \in \mathbb{R}_{>0} \text{ entonces } ac < bc\,. \quad \Diamond$$

Observemos que el axioma O_1) afirma que $\leq$ define una relación de orden total en $\mathbb{R}$. A partir de los doce axiomas que definen a un cuerpo ordenado es posible demostrar muchas otras propiedades del orden. Probaremos en la siguiente proposición algunas de ellas (las letras designan números reales).

Proposición 2.2.1 Se satisfacen en $\mathbb{R}$ las siguientes propiedades de orden:

1) $a > 0 \iff -a < 0$

2) $a < b \iff b - a > 0$

3) $a < b \iff -b < -a$

4) Si $a < b$ y $c < 0$ entonces $ac > bc$

5) $1 > 0$.

DEMOSTRACION La propiedad 1) sigue de O_2, ya que

$$-a = 0 + (-a) < a + (-a) = 0\,.$$

Para probar 2) basta sumar $-a$ a ambos miembros de la desigualdad $a < b$, mientras que 3) se obtiene sumando $-b$ a ambos miembros de la desigualdad $0 < b - a$.

La propiedad 4) se deduce de las propiedades 1) y 3) aplicando el axioma O_3. Finalmente, supongamos por el absurdo que $1 < 0$, en cuyo caso $-1 > 0$, por la propiedad 1). Entonces, aplicando O_3 obtenemos

$$1 = (-1)(-1) > (-1)0 = 0\,,$$

lo que resulta una contradicción. $\Diamond$

2.2.2. La recta real

Es costumbre establecer un modelo geométrico del cuerpo de los números reales, representando a éstos como puntos de una recta. La idea es asignar a cada número real un punto de una recta dada, y de manera que cada punto de ésta corresponda a un único número real. La correspondencia se establece de tal forma que la relación de orden existente entre dos números pueda visualizarse geométricamente a través de los puntos que los representan. Así, si $x < y$ el punto que representa a y estará a la derecha del que representa a x. Sin pretender ser demasiado rigurosos, ciertas propiedades de los números reales sugieren que el modelo es razonable. Por ejemplo, la noción intuitiva de que entre dos puntos de una recta existe otro (en realidad infinitos), tiene su correlato en el hecho de que entre dos números reales hay infinitos números reales, en el sentido del orden. En efecto, si $a < b$ es inmediato demostrar que vale la doble desigualdad $a < (a + b)/2 < b$. Naturalmente, iterando el procedimiento resulta que existen infinitos números reales entre a y b.

Respecto a la construcción en sí, comenzamos tomando un punto arbitrario O que corresponderá al cero. Convenimos entonces en representar los positivos a la derecha de O y los negativos a la izquierda. Si el punto P corresponde a un número positivo x entonces a $-x$ le corresponderá el único punto Q situado a la izquierda de O cuya distancia a O coincide con la longitud del segmento OP. Basta pues representar los reales positivos.

Para ello, comenzamos tomando un segmento OU de medida arbitraria, y asignamos al 1 su extremo derecho U. Replicando la medida del segmento OU cuanto sea necesario, vamos entonces representando los números $1, 2, 3, \ldots$.

Mediante construcciones geométricas elementales podemos a su vez subdividir el segmento OU en $2, 3, 4, \ldots$ segmentos de igual longitud, con lo cual tendremos representados a $1/2, 1/3, 1/4, \ldots$. Esto nos permite entonces hacer corresponder un punto a cada racional positivo. Por ejemplo, si queremos representar a $17/4$, tomamos el punto T que hicimos corresponder a $1/4$ y replicamos diecisiete veces el segmento OT.

Esperamos que el lector haya comprendido la idea de la construcción. No se trata de ser muy precisos, sino más bien tener una imagen geométrica del cuerpo de los números reales que en ocasiones nos ayude a pensar mejor. Quedarían algunas cuestiones delicadas que tratar, como por ejemplo entender por qué los números racionales no "llenan" la recta. Dicha cuestión transciende el contexto puramente algebraico dado por los doce axiomas que definen la estructura de cuerpo ordenado de $\mathbb{R}$, y entra en el terreno del análisis matemático y la topología. Retomaremos el tema en nuestra próxima sección, donde postularemos el axioma de completitud, que cerrará nuestra descripción de los números reales y nos posibilitará probar por ejemplo la existencia de números irracionales.

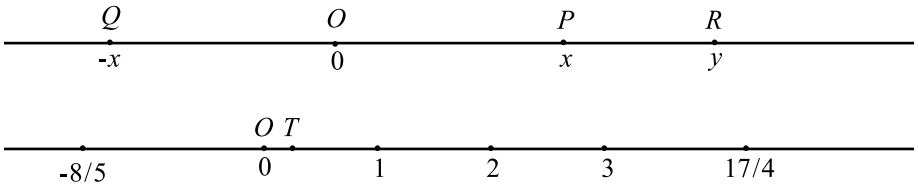


Figura 2.1: La recta real

Intervalos.

Por su frecuente aparición es conveniente distinguir ciertos subconjuntos de $\mathbb{R}$, genéricamente llamados *intervalos*. Supondremos en lo que sigue que a y b son números reales tales que $a \leq b$.

El conjunto

$$\{x \in \mathbb{R} \ : \ a < x < b\}$$

se designa por (a, b) y se llama *intervalo abierto* de extremos a y b. Observemos que $(a, b) = \emptyset$ si $a = b$.

El conjunto

$$\{x \in \mathbb{R} \ : \ a \leq x \leq b\}$$

se designa por $[a, b]$ y se llama *intervalo cerrado* de extremos a y b. De manera análoga se definen los intervalos semiabiertos $(a, b]$ y $[a, b)$. En cualquiera de los cuatro casos, diremos que la *longitud* del intervalo es $b - a$. Observemos que en nuestra representación lineal de los números reales los intervalos

definidos corresponden a segmentos de diversos tipos, lo que justifica la definición anterior.

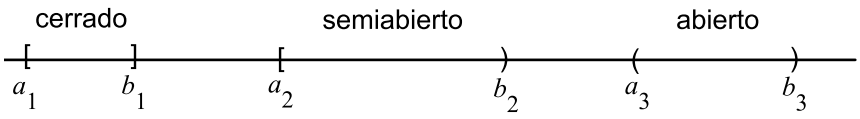


Figura 2.2: Intervalos

Es útil también introducir la noción de intervalo generalizado. Precisamente, si $c \in \mathbb{R}$, definimos los intervalos

$$[c, +\infty) = \{x \in \mathbb{R} \ : \ x \geq c\} \text{ y } (-\infty, c] = \{x \in \mathbb{R} \ : \ x \leq c\} \ .$$

Con las debidas modificaciones, definimos también los intervalos abiertos $(c, +\infty)$ y $(-\infty, c)$. Debe quedar perfectamente claro que el símbolo ∞ no designa un número y que su empleo tiene la finalidad de sugerir la situación que estamos definiendo. Por ejemplo, el intervalo $(0, +\infty)$ es el conjunto de números reales positivos, y avanzando un poco más en la notación, podemos pensar a $\mathbb{R}$ mismo como un intervalo, a saber, $\mathbb{R} = (-\infty, +\infty)$. Advirtamos finalmente que en la recta real los intervalos generalizados corresponden a semirrectas.

Valor absoluto.

Si $x \in \mathbb{R}$, definimos su *valor absoluto* o *módulo* en la forma

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

Tenemos por ejemplo que $|3| = 3$, $|-7/5| = 7/5$ y $|0| = 0$. El lector notará que el módulo es una función de $\mathbb{R}$ en $\mathbb{R}$ cuya imagen es el intervalo $[0, +\infty)$. Veamos a continuación algunas de sus propiedades más relevantes (las letras designan números reales).

Proposición 2.2.2 Son válidas las propiedades siguientes:

1) $|x| \geq 0$ y $|x| = 0 \iff x = 0$

2) $|-x| = |x|$.

3) $|x + y| \leq |x| + |y|$

4) $|xy| = |x|\,|y|$

5) Si $c > 0$ entonces $|x| \leq c \iff x \in [-c, c]$.

DEMOSTRACION. A excepción la tercera, las propiedades precedentes siguen inmediatamente de la definición. En cuanto a 3), observemos en primer término que vale la igualdad si alguno de los dos números es 0 o si ambos son no nulos y tienen el mismo signo. En efecto, esto último es claro si ambos son positivos, mientras que si los dos son negativos tenemos

$$|x + y| = -(x + y) = -x + (-y) = |x| + |y| .$$

Para estudiar el caso en que los números tienen distinto signo podemos suponer sin pérdida de generalidad que $y < 0 < x$. Analizamos entonces el signo de $x + y$. Si $x + y \geq 0$, obtenemos

$$|x + y| = x + y = x - (-y) = |x| - |y| < |x| + |y| ,$$

mientras que si $x + y < 0$ resulta

$$|x + y| = -(x + y) = -x - y = - |x| + |y| < |x| + |y| ,$$

lo cual completa la demostración. $\Diamond$

Distancia.

Si $a, b \in \mathbb{R}$, definimos la *distancia* entre a y b en la forma

$$\delta(a, b) = | b - a | .$$

Por ejemplo, $\delta(6, 2) = 4$, $\delta(-3, 5) = 8$ y $\delta(x, 0) = |x|$, cualquiera sea el número real x. Observemos que la distancia, que es un concepto geométrico, está emparentada con la representación de los números reales en una recta. En efecto, si pensamos a los números a y b como puntos de la recta, y suponiendo $b > a$, la distancia $b - a$ entre a y b es la longitud del intervalo (segmento) (a, b), lo que obviamente responde a nuestro concepto geométrico de distancia. Notemos además que la utilización del valor absoluto en la definición evita la necesidad de considerar las posiciones relativas de a y b, ya que

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

El anterior es seguramente el rol fundamental del valor absoluto, el de cuantificar la proximidad entre dos números. Las propiedades del módulo se traducen en las siguientes propiedades de la distancia, casi todas ellas obvias desde el punto de vista geométrico.

PROPIEDADES DE LA DISTANCIA

Proposición 2.2.3 Son válidas las siguientes propiedades (las letras designan números reales):

1) $\delta(a,b) \geq 0$ y $\delta(a,b) = 0 \iff a = b$

2) $\delta(a,b) = \delta(b,a)$

3) $\delta(a,c) \leq \delta(a,b) + \delta(b,c)$.

Demostracion. Como dijimos arriba, las afirmaciones anteriores son consecuencia inmediata de las correspondientes propiedades del módulo, por lo que encargamos sus demostraciones al lector. $\Diamond$

Para cerrar esta sección, y con la intención de ilustrar los temas desarrollados, vamos a resolver un par de *inecuaciones*, esto es, situaciones en las que queremos determinar qué números reales satisfacen una cierta desigualdad. Las inecuaciones pueden manejarse en forma relativamente similar a las ecuaciones, pero debemos tener ciertos cuidados. Por ejemplo, podemos realizar pasajes de términos sin inconvenientes, puesto que ellos consisten en sumar un cierto número a ambos miembros de la expresión, operación que no cambia el sentido de la desigualdad (propiedad O_2), pero distinta es la situación cuando efectuamos un pasaje de factor, ya que en ese caso la desigualdad se invierte si dicho factor es negativo. Veamos cómo trabajar.

Ejemplo 2.2.4 Determinemos el conjunto

$$\left\{ x \in \mathbb{R} \ : \ \frac{x}{x-1} < 2 \right\}.$$

Para librarnos del denominador lo "pasaremos" multiplicando al otro miembro, lo que nos obliga a discriminar su signo. Consideremos entonces primero el caso $x - 1 > 0$, o equivalentemente $x > 1$. Obtenemos en tal caso la condición $x < 2(x-1) = 2x - 2$, y mediante pasajes de términos arribamos a la desigualdad $2 < x$. Puesto que esta condición implica la anterior, concluimos que el conjunto de soluciones en este primer caso considerado es el intervalo $(2, +\infty)$.

Si $x - 1 < 0$, o equivalentemente $x < 1$, operamos como antes pero teniendo en cuenta que la desigualdad se invierte cuando multiplicamos por $x-1$. Ello nos conduce finalmente a la condición $x < 2$, y, por lo tanto, todo número real menor que 1 satisface este caso. En definitiva, el conjunto de soluciones es la unión de los intervalos $(-\infty, 1)$ y $(2, +\infty)$. $\Diamond$

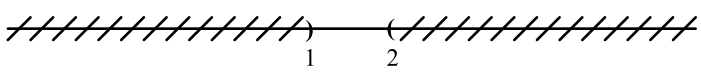


Figura 2.3: $(-\infty, 1) \cup (2, +\infty)$

Remarquemos que al realizar las operaciones algebraicas anteriores, en principio sólo determinamos condiciones *necesarias* para que x satisfaga la

inecuación propuesta. Sin embargo, es fácil ver que todos los pasos son reversibles, por lo que las condiciones halladas también son suficientes y la caracterización obtenida es completamente correcta. Insistimos en que este punto siempre debe ser tenido en cuenta.

Ejemplo 2.2.5 Resolvamos ahora la inecuación

$$|x+1| \geq |x-3| \ .$$

Lo haremos de dos maneras distintas, confiando en que el lector aprecie la ventaja de pensar geométricamente. En la primera de ellas distinguiremos cuatro casos, pues debemos considerar los signos de $x+1$ y $x-3$ para eliminar las barras de módulo.

a) $x+1 \geq 0$ y $x-3 \geq 0$. Estas dos condiciones se resumen en la condición $x \geq 3$, y la desigualdad a resolver es $x+1 \geq x-3$, que claramente es verificada por todo número real. Nuestro primer conjunto de soluciones es entonces el intervalo $[3, +\infty)$.

b) $x+1 \geq 0$ y $x-3 < 0$, o equivalentemente $x \in [-1, 3)(*)$. Por definición de módulo la desigualdad planteada deviene en $x+1 \geq 3-x$, y operando elementalmente arribamos a la condición $x \geq 1$. Teniendo en cuenta $(*)$ concluimos que el conjunto de soluciones en este caso es el intervalo $[1, 3)$.

c) $x+1 < 0$ y $x-3 \geq 0$. Estas condiciones no son verificadas simultáneamente por ningún número real, ya que la primera implica $x < -1$ y la segunda $x \geq 3$. Luego no hay soluciones en este caso.

d) $x+1 \geq 0$ y $x-3 \geq 0$. La desigualdad original se transforma entonces en $-(x+1) \geq -(x-3)$, inecuación que no tiene solución, pues es equivalente a la desigualdad $-1 \geq 3$.

En definitiva, nuestro conjunto de soluciones es

$$S = [3, +\infty) \cup [1, 3) = [1, +\infty) \, ,$$

vale decir, x es solución si y sólo si $x \geq 1$.

Resolvamos ahora la cuestión de una manera geométrica. Simplemente escribiendo $x+1 = x-(-1)$, podemos plantear la inecuación original en términos de distancias, en la forma

$$\delta(x, -1) \geq \delta(x, 3).$$

Para expresarlo geométricamente, estamos buscando los números reales x que están más "cerca" de 3 que de -1, asunto que puede resolverse por simple visualización de la recta real. En efecto, observemos que 1 es el punto

medio del intervalo $[-1, 3]$. Los que están a su derecha están más próximos a 3 que a -1, mientras que los que están a su izquierda se hallan más cerca de -1 que de 3. Por lo tanto, el conjunto de soluciones es, como no podía ser de otra manera, el intervalo $[1, +\infty)$. Notemos que la inclusión del 1 se debe a que la desigualdad planteada no es estricta.

No queremos dar la impresión de privilegiar el segundo método sobre el primero. Es cierto que resultó más sencillo y atractivo en este caso, pero no siempre es así. Ambos métodos son igualmente válidos y el estudiante debe manejar ambos recursos. $\Diamond$

2.2.3. Ejercicios

En los siguientes ejercicios las letras designan números reales (no nulos cuando corresponda).

1. Demostrar las siguientes propiedades:

   a) $ab > 0$ si y solo si a y b tienen el mismo signo.
   b) $ab < 0$ si y solo si a y b tienen distinto signo.
   c) $a > 0 \Leftrightarrow a^{-1} > 0$.
   d) $ab > 0 \Leftrightarrow \dfrac{a}{b} > 0$.
   e) Si $a < b$ y $c > 0$ entonces $\dfrac{a}{c} < \dfrac{b}{c}$.
   f) $a > 1 \Leftrightarrow a^{-1} < 1$ (a positivo).
   g) $a < b \Leftrightarrow a/b < 1$ (a y b positivos).
   h) $\dfrac{a}{b} < \dfrac{c}{d} \Leftrightarrow ad < bc$ (b y d positivos).
   i) $a < b \Leftrightarrow a^{-1} > b^{-1}$ (a y b del mismo signo).
   j) $a^2 < a \Leftrightarrow a < 1$ (a positivo).

2. Demostrar las siguientes propiedades:

   a) $|\,ab\,| = |\,a\,|\,|\,b\,|$.
   b) $\left|\,a^{-1}\,\right| = |\,a\,|^{-1}$.
   c) $\left|\,\dfrac{a}{b}\,\right| = \dfrac{|\,a\,|}{|\,b\,|}$.
   d) $a^2 \leq b^2 \Leftrightarrow |\,a\,| \leq |\,b\,|$.
   e) $|\,a - b\,| \geq |\,a\,| - |\,b\,|$.

3. Sean $a, b, c \in \mathbb{R}$ tales que $\delta(a,c) = \delta(a,b) + \delta(b,c)$. Probar que $b \in [a,c]$ ó $b \in [c,a]$.

4. Resolver en $\mathbb{R}$ las siguientes inecuaciones:

   a) $3x + 1 < 16 - 2x$.

   b) $\dfrac{x+3}{2x-1} < \dfrac{1}{4}$.

   c) $\dfrac{1}{x} \geq \dfrac{1}{1-x}$.

   d) $|\, x + 4 \,| \leq |\, 2x - 6 \,|$.

   e) $\dfrac{x+3}{x-5} < \dfrac{2x-1}{2x+4}$.

   f) $\dfrac{1}{x-1} < \dfrac{1}{x} < \dfrac{1}{x+1}$.

5. Sean u y v números reales $(0 < u < v)$ correspondientes a dos puntos P y Q de la recta real L .

   a) Determinar el número real correspondiente al punto medio del segmento PQ .

   b) Si R es el punto que corresponde a $u - v$, determinar los números reales correspondientes a los puntos medios de los segmentos RP y RO .

   c) Sea $S \in L$ tal que $3\,\ell(QS) = 2\,\ell(PS)$, donde en general $\ell(XY)$ indica la longitud del segmento XY. Determinar el número real correspondiente a S .

   d) Hallar los extremos del intervalo $I = \{\, x \in \mathbb{R} : |\, v - 2x \,| < u \}$.

2.3. Completitud del cuerpo de números reales

2.3.1. Axioma de completitud

Como se recordará, en nuestra breve introducción al capítulo hemos hablado vagamente de la necesidad de manipular cantidades (números) no racionales. Sin embargo, en nuestra presentación axiomática (y sin duda abstracta) de los números reales solo se mencionan específicamente dos números, el 0 y el 1. Claro que operando con ellos, y usando las doce reglas básicas, podemos asegurar la existencia de infinitos más. Así, podemos ir definiendo $2 = 1 + 1, 3 = 2 + 1$, etc., proceso que nos lleva a la construcción de los números naturales. Agregando sus inversos aditivos tenemos los enteros, y tomando todos los cocientes posibles obtenemos los números racionales. Pero no es mucho más lo que podemos lograr hasta aquí. Por ejemplo, no podríamos determinar usando dichos axiomas la existencia de un número real cuyo cuadrado sea 2. Necesitamos entonces algo más, un axioma adicional que completará nuestra descripción de los números reales.

Conjuntos acotados.

> Si $X \subseteq \mathbb{R}$, diremos que X es *acotado superiormente* si y sólo si existe $c \in \mathbb{R}$ tal que $c \geq x$ para todo $x \in X$. En tal caso c se dirá una cota superior de X. Similarmente, diremos que X es *acotado inferiormente* si y sólo si existe $c \in \mathbb{R}$ tal que $c \leq x$ para todo $x \in X$, en cuyo caso c se dirá una cota inferior de X. Si X es acotado superiormente e inferiormente diremos que X es *acotado*.

El lector puede probar sin demasiada dificultad que X es acotado si y sólo si existe un número real positivo M tal que $|x| \leq M$ para todo $x \in X$.

Por ejemplo, el conjunto de números positivos está acotado inferiormente por 0, pero no admite cota superior. En efecto, si existiera una tal cota c, resultaría en particular que $c \geq 1$, y c sería positiva, por transitividad. Tendríamos entonces que $c \geq c + 1$ (pues este número es positivo), lo que obviamente es falso. Inversamente, el conjunto de números negativos es acotado superiormente pero no inferiormente. Como ejemplo típico de conjunto acotado podemos tomar cualquier intervalo de extremos a y b ($a < b$), ya que claramente a es una cota inferior y b una cota superior del mismo. Observemos finalmente que el conjunto

$$\{\, x \in \mathbb{R} : |x| > 1 \,\} = (-\infty, -1) \cup (1, +\infty)$$

no es acotado ni superiormente ni inferiormente.

Supremo e ínfimo.

Comenzaremos con una observación. Si c es una cota superior de un conjunto X, es claro que cualquier número mayor que c también es cota

superior de X, de donde concluimos que X admite infinitas cotas superiores. Obviamente, un resultado análogo se obtiene para conjuntos acotados inferiormente. Por ejemplo, todo número mayor o igual que 4 es una cota superior del intervalo $(-\infty, 4)$. Ahora bien, ¿cuál es la "mejor" ? Sin duda que 4, pues es la que nos aporta la mejor información sobre el conjunto. Pensemos que decir que todos los elementos de un conjunto son menores o iguales que 6 no nos permite decidir por ejemplo si 5 pertenece o no al conjunto, cuestión que sí podemos elucidar si sabemos que 4 es una cota superior. En las dos definiciones siguientes precisaremos estas reflexiones.

Supremo de un conjunto.

> Sea X un subconjunto de $\mathbb{R}$ acotado superiormente y supongamos que existe una cota superior s de X tal que $s \leq c$ para toda cota superior c de X, esto es, s es la *menor* de las cotas superiores de X. Diremos entonces que s es el *supremo* de X y notaremos $s = \sup X$. Si además $s \in X$, diremos que s es el *máximo* de X y escribiremos $s = \max X$.

Observemos que la existencia de máximo de un conjunto X es equivalente a la existencia de un elemento s de X tal que $x \leq s$ para todo $x \in X$. Se demuestra inmediatamente que el supremo s de un conjunto X, si existe, es único. En efecto, supongamos que s' es un número real satisfaciendo las mismas condiciones que s. Resulta en particular que ambos son cotas superiores de X, por lo que deben verificarse las desigualdades $s \leq s'$ y $s' \leq s$. Por lo tanto, $s = s'$.

Existe una noción dual a la de supremo, referida a conjuntos acotados inferiormente, que definimos a continuación.

Infimo de un conjunto.

> Sea X un subconjunto de $\mathbb{R}$ acotado inferiormente y supongamos que existe una cota inferior t de X tal que $t \geq c$ para toda cota inferior c de X, esto es, t es la *mayor* de las cotas inferiores de X. Diremos entonces que t es el *ínfimo* de X y notaremos $t = \inf X$. Si además $t \in X$, diremos que t es el *mínimo* de X y escribiremos $t = \min X$.

Equivalentemente, el mínimo de X es un elemento $t \in X$ tal que $t \leq x$ para todo $x \in X$. En forma muy similar al caso del supremo, se prueba que el ínfimo de un conjunto, si existe, es único.

Ejemplo 2.3.1 Tomemos $X = [0, 2)$. Para ver que 2 es la menor de las cotas superiores de X, supongamos por el contrario que c es una cota superior menor que 2. Puesto que $c \geq 0$, por ser una cota superior de X, resulta que $c \in X$. Ahora bien, sabemos que entre dos números reales siempre existe

otro, luego podemos elegir a tal que $c < a < 2$. Pero entonces $a \in X$ (pues $a > c \geq 0$) y $a > c$, lo que contradice el hecho de que c es cota superior de X. En consecuencia $\sup X = 2$. Puesto que $2 \notin X$ concluimos que X no admite máximo. En cambio sí admite mínimo, ya que $0 \in X$ y $x \geq 0$ para todo $x \in X$, por definición. Luego, $\min X = 0$. $\Diamond$

Sin duda fue relativamente sencillo determinar el supremo y el ínfimo en el caso del ejemplo anterior. Ello se debió fundamentalmente a la forma muy particular del conjunto elegido, específicamente al hecho de ser un intervalo. La cuestión puede ser harto más complicada en otro tipo de situaciones, y en realidad no tenemos hasta aquí ninguna propiedad de los números reales que nos permita asegurar que un conjunto acotado superiormente o inferiormente admita supremo o ínfimo. Debemos pues enriquecer nuestra teoría, agregando el axioma al que aludimos al principio de la sección.

Axioma de completitud.

> Todo subconjunto no vacío y acotado superiormente de $\mathbb{R}$ admite supremo.

Debido a la validez de esta propiedad, diremos en adelante que $\mathbb{R}$ es un *cuerpo ordenado completo*. Notemos que la hipótesis de que el subconjunto sea no vacío es necesaria, ya que cualquier número real es cota superior del conjunto vacío (suponga lo contrario y arribará rápidamente a una contradicción). Observemos además que el enunciado del axioma de completitud de los números reales parece adolecer de cierta asimetría, ya que no postula la existencia de ínfimo para los subconjuntos acotados inferiormente. Sin embargo, veremos a continuación que dicha propiedad es igualmente válida y puede demostrarse a partir de dicho axioma.

Teorema 2.3.2 Todo subconjunto no vacío y acotado inferiormente de $\mathbb{R}$ admite ínfimo.

Demostracion. Sea X un subconjunto no vacío de $\mathbb{R}$ acotado inferiormente y sea

$$X^- = \{\, y \in \mathbb{R} : -y \in X \,\} ,$$

esto es, X^- es el conjunto de inversos aditivos de los elementos de X.

Claramente $X^- \neq \emptyset$, y dada una cota inferior c de X resulta que $-c$ es una cota superior de X^-, ya que $x \geq c$ y por lo tanto $-c \geq -x$ para todo $x \in X$. Luego, por el axioma de completitud X^- admite supremo, que designaremos por s. Probaremos que $-s$ es el ínfimo de X.

Si $x \in X$, de la desigualdad $-x \leq s$ sigue que $x \geq -s$, y por lo tanto $-s$ es una cota inferior de X. Sólo resta probar que es la mayor. Consideremos para ello cualquier cota inferior a de X. Al igual que arriba, $-a$ resulta ser entonces una cota superior de X^-, de donde sigue por definición de supremo

que $s \leq -a$, o equivalentemente, $-s \geq a$, como queríamos probar. Notemos de paso que hemos probado la igualdad $\inf X = -\sup X^-$, un resultado interesante en sí mismo y que justifica nuestra aseveración acerca de que el supremo y el ínfimo son conceptos duales. $\Diamond$

La idea que hemos desarrollado en los ejemplos de cálculo del supremo y el ínfimo se verá trasuntada en los siguientes lemas. Ofreceremos en ellos definiciones equivalentes de supremo e ínfimo, que nos permitirán trabajar con mayor fluidez y nos mostrarán una característica esencial de ambos.

Lema 2.3.3 Si $A \subset \mathbb{R}$ y u es un número real, son equivalentes las siguientes afirmaciones:

1) $\sup A = u$.

2) u es una cota superior de A y dado cualquier número real $\epsilon > 0$ existe $a \in A$ tal que $u - \epsilon < a \leq u$.

Demostracion. Comencemos probando que 1) implica 2). Por definición de supremo u es una cota superior de A, y dado $\epsilon > 0$ es claro que $u - \epsilon$ no lo es, ya que $u - \epsilon < u$ y u es la menor de las cotas superiores. Por lo tanto existe $a \in A$ tal que $u - \epsilon < a$, lo que asegura la conclusión, ya que obviamente $a \leq u$.

Recíprocamente, asumamos que se verifica 2) y probemos que $\sup A = u$. Puesto que por hipótesis u es cota superior de A, sólo debemos probar que es la menor de dichas cotas, para lo cual bastará demostrar que ningún número c menor que u es cota superior de A. En efecto, puesto que $u - c$ es positivo existe por hipótesis un elemento a de A tal que $u - (u - c) < a \leq u$, lo que nos muestra en particular $a > c$. Luego c no es cota superior de A, como nos proponíamos demostrar. $\Diamond$

Razonando de manera muy similar obtenemos la siguiente caracterización del ínfimo de un conjunto, por lo que dejaremos los detalles de la demostración a cargo del lector:

Lema 2.3.4 Si $A \subset \mathbb{R}$ y u es un número real, son equivalentes las siguientes afirmaciones:

1) $\inf A = u$.

2) u es una cota inferior de A y dado cualquier número real $\epsilon > 0$ existe $a \in A$ tal que $u \leq a < u + \epsilon$. $\Diamond$

Nota Observemos con cuidado el significado de los hechos que acabamos de establecer. Nos concentraremos en el caso del supremo, pero similares consideraciones valdrán para el ínfimo. Hemos probado concretamente que si u es el supremo de un conjunto A, para todo $\epsilon > 0$ existe algún elemento a

de A en el intervalo $(u-\epsilon, u]$. Pensando la cuestión en términos de distancias, obtenemos que

$$\delta(a,u) = |u-a| = u-a = \epsilon - (a-(u-\epsilon)) < \epsilon,$$

pues $a-(u-\epsilon) > 0$. Puesto que ϵ es arbitrario, esto indica que existen elementos de A *tan próximos al supremo como se quiera*. Naturalmente, lo mismo ocurre en el caso del ínfimo. Esta propiedad de densidad es distintiva de los números reales y está estrechamente ligada a la idea de límite, con la que el lector probablemente se haya familiarizado en algún curso de cálculo infinitesimal. $\Diamond$

Nos dedicaremos ahora a mostrarle una importante aplicación del axioma de completitud, como es la existencia de raíces cuadradas de números reales no negativos.

2.3.2. Raíces cuadradas

Recordemos que si $x \in \mathbb{R}$ el producto $x\,x$ se llama el *cuadrado* de x, que notamos x^2. Exhibiremos en el siguiente lema algunas propiedades elementales de los cuadrados que serán de utilidad en el resto de la sección (las letras designan números reales).

Lema 2.3.5 Son válidas las siguientes propiedades:

1) $(x \pm y)^2 = x^2 \pm 2xy + y^2$.

2) $x^2 - y^2 = (x-y)(x+y)$.

3) $x^2 \geq 0$.

4) Sean $x, y > 0$. Entonces $x < y \iff x^2 < y^2$.

Demostracion. 1) y 2) siguen por simple aplicación de la propiedad distributiva del producto respecto a la suma. En el ítem 3) podemos suponer que $x \geq 0$, ya que $x^2 = (-x)^2$, y el resultado sigue usando el axioma O_3). Finalmente, de la igualdad $x^2-y^2 = (x-y)(x+y)$ deducimos que los signos de x^2-y^2 y $x-y$ coinciden, pues $x+y$ es positivo, lo que obviamente prueba 4). $\Diamond$

Si bien el lector tiene cierta familiaridad con la noción de raíz cuadrada de un número real, la siguiente proposición nos permitirá definir con precisión el concepto.

Proposición 2.3.6 Dado un número real no negativo a existe un único número real no negativo b tal que $b^2 = a$.

DEMOSTRACION. El resultado es trivial si $a = 0$, pues basta tomar $b = 0$, con la unicidad asegurada por el hecho de que un producto de números reales es nulo si y sólo si alguno de los factores lo es. Podemos suponer entonces que a es positivo.

Respecto a la unicidad, supongamos que b y c son números reales positivos tales que $b^2 = c^2 = a$. En tal caso, obtenemos:

$$0 = b^2 - c^2 = (b - c)(b + c),$$

de donde sigue que alguno de los dos factores anteriores debe ser nulo. Puesto que $b + c > 0$, por ser ambos números positivos, concluimos que $b - c = 0$, esto es, $b = c$, como queríamos probar.

Probemos ahora que existe un número real positivo b tal que $b^2 = a$. A tal fin, consideraremos el conjunto

$$A = \{x \in \mathbb{R} : x > 0 \text{ y } x^2 \leq a\}.$$

Veamos en primer lugar que $A \neq \emptyset$. Para ello, tomemos cualquier número real $r > 0$. Si $r^2 \leq a$ la afirmación queda probada. En caso contrario, sigue de la desigualdad $a < r^2$ que $a/r^2 < 1$, de donde obtenemos

$$(a/r)^2 = a^2/r^2 = a(a/r^2) < a,$$

y por lo tanto $a/r \in A$.

Veamos por otro lado que A es acotado superiormente, siendo $a + 1$ una cota superior de A. En efecto, supongamos por el contrario que existe $x \in A$ tal que $x > a + 1$. Entonces

$$x^2 - a > (a + 1)^2 - a = a^2 + a + 1 > 0,$$

lo que contradice el hecho de que $x \in A$. Sigue luego del axioma de completitud que A admite supremo, digamos b (es claro que b es positivo, por ser cota superior de un conjunto de números positivos).

Nuestro objetivo es mostrar a continuación que cualquiera de las suposiciones $b^2 < a$ y $b^2 > a$ nos conduce a una contradicción. Resultará entonces por la ley de tricotomía que $b^2 = a$, y habremos completado la demostración.

Supongamos en primer término que $b^2 < a$ y sea $\delta = a - b^2$. Eligiendo h como el mínimo entre 1 y $\delta(2b + 1)^{-1}$, resulta que h es positivo y satisface las desigualdades $h \leq 1$ y $h(2b + 1) \leq \delta$. Sigue entonces que

$$(b + h)^2 - a = b^2 + 2bh + h^2 - a = h(2b + h) - \delta \leq h(2b + 1) - \delta \leq 0\,.$$

Por consiguiente $b + h \in A$ y $b + h > b$, lo cual es imposible por ser b una cota superior de A.

Supongamos ahora $b^2 > a$. Luego podemos escribir $b^2 = a + \gamma$, con $\gamma > 0$. Tomando ahora cualquier número positivo h menor que el mínimo entre b y $\gamma(2b)^{-1}$, probaremos que $b - h$ es una cota superior de A, y habremos

llegado a un absurdo, pues $b-h<b$ y b es la menor de las cotas superiores de A.

Sea x cualquier elemento de A. Dado que x y $b-h$ son positivos, bastará probar que $(b-h)^2>x^2$. Puesto que $x^2\leq a$, tenemos

$$(b-h)^2-x^2\geq(b-h)^2-a=b^2-2bh+h^2-a=\gamma-2bh+h^2>h^2>0\,,$$

como queríamos probar. El lector podrá apreciar, en la prueba que acabamos de finalizar, el uso reiterado del lema 2.3.5 y de las propiedades de consistencia del orden respecto a la suma y el producto. $\Diamond$

RAIZ CUADRADA

> Si $a\geq 0$, definimos la *raíz cuadrada* de a como el único número real no negativo b tal que $b^2=a$. La notaremos $\sqrt{a}$.

Por ejemplo, $\sqrt{1}=1$, $\sqrt{9}=3$ y $\sqrt{0}=0$.

NOTA Se imponen algunos comentarios. En primer lugar debe quedar perfectamente claro que el símbolo $\sqrt{a}$ sólo está definido si $a\geq 0$, y que por definición denota un número real no negativo. Podemos resumir lo antedicho mediante la siguiente caracterización de la raíz cuadrada:

$$b=\sqrt{a}\text{ si y sólo si }b\geq 0\text{ y }b^2=a\,.$$

Notemos además que no es lícito afirmar que la operación de tomar raíz cuadrada y la operación de elevar al cuadrado son mutuamente inversas (como suele pensarse), ya que tal relación sólo vale en un sentido. En efecto, si $x\geq 0$, sigue por definición que $(\sqrt{x})^2=x$, mientras que $\sqrt{x^2}=-x$ si $x<0$.

Queremos finalmente recalcar un punto para evitar confusiones en lo sucesivo, acerca del uso de expresiones del tipo $\sqrt{4}=\pm 2$. Las mismas son incorrectas, ya que la raíz cuadrada de un número es un único valor perfectamente determinado. Sin embargo, podemos usarlas en ciertos contextos si nos ponemos de acuerdo previamente sobre lo que significan. Por ejemplo, supongamos que t es un número positivo y consideremos la ecuación $x^2=t$. Operando, tenemos:

$$0=x^2-t=x^2-(\sqrt{t})^2=(x-\sqrt{t})(x+\sqrt{t}).$$

Puesto que un producto es cero si y sólo si alguno de los factores es nulo, concluimos que $x-\sqrt{t}=0$ ó $x+\sqrt{t}=0$, lo que equivale a decir que las soluciones son $x=\sqrt{t}$ y $x=-\sqrt{t}$. Podemos convenir entonces en decir abreviadamente que las soluciones son $x=\pm\sqrt{t}$. Confiamos en que el lector advertirá la diferencia de significado existente entre ambas situaciones. $\Diamond$

Ecuaciones cuadráticas.

En el párrafo anterior hemos tratado un caso particular de las denominadas *ecuaciones cuadráticas* o *ecuaciones de segundo grado*, es decir, ecuaciones de la forma

$$ax^2 + bx + c = 0,$$

donde a, b y c son números reales y $a \neq 0$. La expresión de arriba está asociada a la función $f : \mathbb{R} \to \mathbb{R}$ definida por $f(x) = ax^2 + bx + c$, y resolver la ecuación no es otra cosa que determinar los *ceros* de f, esto es, los valores de la variable en los que la función toma el valor nulo. No hay duda de que el lector tiene idea del asunto. Sabe por ejemplo que el gráfico de f es una parábola y casi seguro recuerda una fórmula para resolver la ecuación. Vamos sin embargo a deducirla en detalle, mostrando su estrecha conexión con la noción de raíz cuadrada. Para ello, vamos a manipular un poco la expresión de la función, de manera de llevarla a una forma más sencilla de manejar.

Por ejemplo, si multiplicamos por $4a$ obtenemos:

$$4af(x) = (2ax)^2 + 4abx + 4ac = (2ax+b)^2 + 4ac - b^2 = (2ax+b)^2 - (b^2 - 4ac).$$

El número $b^2 - 4ac$ se llamará el *discriminante* de la ecuación, y lo designaremos por Δ. Resulta por lo tanto que $4af(x) = (2ax + b)^2 - \Delta$. Esta expresión nos permite analizar más fácilmente el comportamiento de la función, y en particular estudiar sus ceros.

Observemos en primer lugar que $f(x) = 0$ si y sólo si $4af(x) = 0$, ya que $a \neq 0$. Luego, x es solución de la ecuación si y sólo si $(2ax + b)^2 - \Delta = 0$, o equivalentemente, $(2ax + b)^2 = \Delta$. Esto nos coloca en situación de resolver completamente la cuestión, por simple análisis de los siguientes casos:

1) $\Delta < 0$. La ecuación no admite entonces ninguna solución, ya que el cuadrado de un número real no puede ser negativo.

2) $\Delta = 0$. Puesto que $t^2 = 0$ si y sólo si $t = 0$, sigue que $2ax + b = 0$. Deducimos en consecuencia que la ecuación admite en este caso una única solución, precisamente, $x = -b/2a$.

3) $\Delta > 0$. Como vimos en la nota anterior, y de acuerdo con nuestra convención notacional, resulta entonces que $2ax + b = \pm\sqrt{\Delta}$. Despejando x, concluimos que en este último caso la ecuación admite exactamente dos soluciones, a saber:

$$x_1 = \frac{-b + \sqrt{\Delta}}{2a}$$
$$x_2 = \frac{-b - \sqrt{\Delta}}{2a} \ .$$

Resumiendo, una ecuación cuadrática admite ninguna, una o dos soluciones según que su discriminante sea negativo, cero o positivo. En los dos

últimos casos acostumbraremos a expresar las soluciones (incluye también el caso de solución única) en la forma más abreviada

$$x = \frac{-b \pm \sqrt{\Delta}}{2a}. \quad \lozenge$$

Ejemplo 2.3.7 Un simple cálculo nos muestra que -2 y 3 son las soluciones de la ecuación $x^2-x-6=0$, y que $1/2$ es la única solución de $4x^2-4x+1=0$. En cambio, la ecuación $9x^2-6x+4=0$ no admite ninguna solución, ya que en este caso $\Delta = (-6)^2 - 144 < 0$. $\lozenge$

Inecuaciones cuadráticas.

El análisis que hemos hecho en el parágrafo anterior nos permite también estudiar los cambios de signo de la función $f(x) = ax^2 + bx + c$. En efecto, supongamos en primer término que $\Delta < 0$. En tal caso, recordando que

$$4af(x) = (2ax + b)^2 - \Delta\,,$$

observemos que el segundo miembro de esta igualdad es positivo cualquiera sea x. Por lo tanto, $f(x)$ tiene signo constante, y éste coincide con el signo de a. Precisamente:

$f(x) > 0$ para todo $x \in \mathbb{R}$ si $a > 0$ y $f(x) < 0$ para todo $x \in \mathbb{R}$ si $a < 0$.

Si $\Delta \geq 0$ y u, v son los ceros de f, sigue inmediatamente de las fórmulas obtenidas para las soluciones que valen las relaciones

$$u + v = -b/a$$
$$uv = c/a\,.$$

Estas igualdades nos permiten factorizar la expresión de f, ya que entonces

$$(x-u)(x-v) = x^2 - (u+v)x + uv = x^2 + (b/a)x + c/a = f(x)/a,$$

de donde sigue que

$$f(x) = a(x-u)(x-v). \qquad (*)$$

Si $u = v$ (caso $\Delta = 0$), la igualdad anterior adopta la forma

$$f(x) = a(x + b/2a)^2,$$

resultando que $f(x)$ tiene el mismo signo que a para todo $x \neq -b/2a$. Por supuesto, $f(-b/2a) = 0$.

Si $u \neq v$ (caso $\Delta > 0$), supongamos primero que $a > 0$. En tal caso, usando que un producto es positivo si y sólo si sus factores son de igual signo, un sencillo análisis sobre (*) nos muestra que $f(x)$ es positivo si $x < u$ ó $x > v$, mientras que $f(x)$ es negativo si $u < x < v$ (hemos supuesto sin pérdida de generalidad que $u < v$). Naturalmente, $f(u) = f(v) = 0$. Finalmente, es inmediato verificar que estas relaciones de orden se invierten en el caso $a < 0$. $\lozenge$

Ejemplo 2.3.8 Caractericemos el conjunto

$$S = \left\{x \in \mathbb{R} \ : \ 2x^2 - 4 \geq 7x\right\} .$$

Podemos hacer uso de los resultados obtenidos arriba, ya que el problema es equivalente a resolver la inecuación cuadrática $f(x) = 2x^2 - 7x - 4 \geq 0$. Por directa aplicación de la fórmula de la ecuación de segundo grado, obtenemos que los ceros de f son $x = -1/2$ y $x = 4$. Puesto que en este caso a es positivo, obtenemos que $f(x) > 0$ si y sólo si $x < -1/2$ ó $x > 4$. Luego:

$$S = (-\infty, -1/2] \cup [4, +\infty). \quad \Diamond$$

2.3.3. Ejercicios

1. Demostrar las siguientes propiedades de la raíz cuadrada (salvo en el ítem a) las letras designan números reales no negativos):

   a) $\sqrt{a^2} = |\, a \,|$
   b) $a \leq b \Leftrightarrow \sqrt{a} \leq \sqrt{b}$
   c) $\sqrt{a} = \sqrt{b} \Leftrightarrow a = b$
   d) $\sqrt{ab} = \sqrt{a}\sqrt{b}$
   e) $\sqrt{a^{-1}} = (\sqrt{a})^{-1}$ $(a \neq 0)$
   f) $\sqrt{\dfrac{a}{b}} = \dfrac{\sqrt{a}}{\sqrt{b}}$ $(b \neq 0)$
   g) $\sqrt{ab} \leq \dfrac{a+b}{2}$. Probar que vale la igualdad si y solo si $a = b$.

2. En cada uno de los siguientes casos analizar si el conjunto S dado es acotado superiormente o inferiormente. Cuando corresponda, determinar el supremo (máximo) y el ínfimo (mínimo) de S.

   a) $S = \left\{x \in \mathbb{R} : x^2 + x < 1\right\}$
   b) $S = \left\{x \in \mathbb{R} : x^2 + 1 \geq -x\right\}$
   c) $S = \left\{x \in (-\infty, 0) : x^2 + x - 1 \geq 0\right\}$
   d) $S = \left\{\sqrt{a^2 + x} : x \in (0, +\infty)\right\}$ $(a \in \mathbb{R})$
   e) $S = \{x \in \mathbb{R} : (x-a)(x-b)(x-c)(x-d) < 0\}$ $(a < b < c < d)$
   f) $S = \left\{x^{-1} : x \in (0, +\infty)\right\}$
   g) $S = \left\{1 - x^{-1} : x \in (0, +\infty)\right\}$,

3. En los siguientes ítems A y B denotan subconjuntos no vacíos de $\mathbb{R}$.

   a) Sean A y B tales que $a \leq b$ para todo $(a, b) \in A \times B$. Probar que A admite supremo, B admite ínfimo y además $\sup A \leq \inf B$.

b) Exhibir un ejemplo de la situación anterior en el que la desigualdad entre el supremo de A y el ínfimo de B sea estricta y otro en el cual se verifique la igualdad.

c) Sean A y B acotados superiormente, y sea

$$A + B = \{x + y : x \in A\,, y \in B\}\,.$$

Probar que $A + B$ admite supremo y que

$$\sup(A + B) = \sup A + \sup B\,.$$

d) En la situación de *c*), probar que $A + B$ admite máximo si y solo si A y B admiten máximo.

4. Determinar los pares (α, β) de números reales tales que:

a) α y β son las soluciones de la ecuación cuadrática $x^2+\alpha x+\beta = 0$.

b) α y β son las soluciones de la ecuación cuadrática $\alpha x^2-x+\beta = 0$.

5. *a*) Demostrar que dos números reales quedan completamente determinados si se conocen su suma y su producto.

b) Hallar las dimensiones de un rectángulo cuyo perímetro es 32 y su área es 39.

c) Pablo le lleva 8 años a Andrea y el producto de sus edades es 609. ¿ Qué edades tienen ?

6. Sean las funciones $f(x) = 2x^2-4x+5$ y $g(x) = x^2+6x-4$. Determinar:

a) $\{x \in \mathbb{R} : f(x) \leq 0\}$

b) $\{x \in \mathbb{R} : g(x) \leq 3\}$

c) $\{x \in \mathbb{R} : g(x) \geq f(x)\}$

d) $\{x \in \mathbb{R} : g(x) < 5 < f(x)\}$.

7. Probar las siguientes desigualdades (las letras indican números reales):

a) $a + \dfrac{1}{a} \geq 2 \quad (a > 0)$

b) $a^2 + ab + b^2 \geq 0$

c) $2a^2 - 2ab + b^2 > 0 \quad (a \neq b)$

d) $(ac + bd)^2 \leq (a^2 + b^2)(c^2 + d^2)$.

Capítulo 3

Números enteros y racionales

3.1. Numeros naturales

3.1.1. Conjuntos inductivos y números naturales

Definiremos formalmente el conjunto $\mathbb{N}$ de números naturales de manera que se satisfagan las propiedades que intuitivamente lo caracterizan, a saber, que tiene un elemento mínimo (el 1) y que a partir de éste pueden obtenerse todos sus elementos sumando sucesivamente 1.

Consideraremos para ello una colección particular de subconjuntos de $\mathbb{R}$.

Conjuntos Inductivos.

> Si $A \subseteq \mathbb{R}$, diremos que A es un conjunto *inductivo* si y solo si se verifican las condiciones
>
> (1) $1 \in A$.
>
> (2) $x + 1 \in A$ para todo $x \in A$.

Ejemplos 3.1.1 Los conjuntos

$$A_1 = \mathbb{R} \quad , \quad A_2 = [-1, +\infty) \quad , \quad A_3 = \{1, 2\} \cup [3, +\infty)$$

son inductivos (verificación a cargo del lector), mientras que los conjuntos

$$B_1 = (1, +\infty),\ B_2 = \{x \in \mathbb{R} : |x - 2| > 1\},\ B_3 = (-\infty, \sqrt{101})$$

no lo son. En efecto, 1 no pertenece ni a B_1 ni a B_2, y B_3 contiene a 10 pero no a 11. $\Diamond$

Numeros naturales.

Atendiendo a la idea que esbozamos al principio, definimos los números naturales en la siguiente forma:

Un número real x se dice un *número natural* si y solo si x pertenece a todo subconjunto inductivo de $\mathbb{R}$.

Denotando por $\mathbb{N}$ el conjunto de números naturales (que también llamaremos *enteros positivos*), la definición precedente adopta la siguiente forma conjuntística:

$$\mathbb{N} = \bigcap_{\mathcal{I}} \mathcal{I}\,,$$

donde $\mathcal{I}$ recorre los subconjuntos inductivos de $\mathbb{R}$. El resultado siguiente establece que $\mathbb{N}$ es el "menor" subconjunto inductivo de $\mathbb{R}$.

Proposición 3.1.2 $\mathbb{N}$ es un conjunto inductivo y todo subconjunto inductivo de $\mathbb{R}$ contiene a $\mathbb{N}$.

Demostracion. La primera afirmación sigue del hecho (fácilmente demostrable) de que la intersección de conjuntos inductivos es un conjunto inductivo. La segunda es consecuencia de un hecho general: la intersección de una familia arbitraria de conjuntos está contenida en cada uno de sus miembros. $\Diamond$

Para cada $n \in \mathbb{N}$ el número natural $n+1$ se denomina *sucesor* o *siguiente* de n. También se dice que n es el *predecesor* de $n+1$. La siguiente proposición garantiza (como se esperaba) que

$$\mathbb{N} = \{1, 2, 3, \dots\}\,.$$

Proposición 3.1.3 Son válidas las siguientes afirmaciones:

1) 1 es el mínimo de $\mathbb{N}$.

2) Si $m \in \mathbb{N}$ y $m > 1$ entonces $m - 1 \in \mathbb{N}$.

3) Si $m \in \mathbb{N}$ no existe ningún número natural en el intervalo $(m, m+1)$.

Demostracion. El intervalo real $[1, +\infty)$ es claramente inductivo y por lo tanto $\mathbb{N} \subseteq [1, +\infty)$. Teniendo en cuenta que $1 \in \mathbb{N}$ deducimos la validez de 1).

Para demostrar 2) consideremos el conjunto

$$A = \{a + 1 : a \in \mathbb{N} \text{ ó } a = 0\}\ .$$

Observemos en primer término que $1 \in A$, pues $1 = 0 + 1$. Sea ahora $x \in A$, digamos $x = k + 1$, donde $k \in \mathbb{N}$ ó $k = 0$. Resulta entonces que

$$x + 1 = (k + 1) + 1 \in A\,,$$

pues $k+1$ es un número natural en cualquiera de los casos $k=0$ ó $k \in \mathbb{N}$.

Luego A es inductivo y por lo tanto $\mathbb{N} \subseteq A$. En particular $m \in A$, y puesto que $m > 1$, existe $s \in \mathbb{N}$ tal que $m = s+1$. Luego $m-1 = s \in \mathbb{N}$, como queríamos probar.

Finalmente, para obtener 3) bastará probar que el conjunto

$$B = [1, +\infty) - \bigcup_{k \in \mathbb{N}} (k, k+1)$$

es inductivo, ya que entonces $\mathbb{N} \subseteq B$.

Es claro que $1 \in B$. Tomemos $x \in B$ y supongamos por el absurdo que $x+1 \notin B$, en cuyo caso existe $l \in \mathbb{N}$ tal que $l < x+1 < l+1$, o equivalentemente $l-1 < x < l$. Puesto que $l \neq 1$, ya que en tal caso sería $x < 1$, resulta por 2) que $l-1 \in \mathbb{N}$, lo que contradice el hecho de que $x \in B$. Luego B es inductivo y nuestra afirmación queda probada.

Es útil registrar que 3) es equivalente a afirmar que si $m, n \in \mathbb{N}$ y $n > m$ entonces $n \geq m+1$. $\Diamond$

Inducción.

El hecho de que $\mathbb{N}$ sea el menor conjunto inductivo proporciona una manera peculiar de probar que un cierto subconjunto de $\mathbb{N}$ es igual a $\mathbb{N}$. Concretamente, es válido el siguiente enunciado:

Teorema 3.1.4 (de inducción) Si A es un un subconjunto inductivo de $\mathbb{N}$ entonces $A = \mathbb{N}$.

Demostracion Es inmediata, pues por 3.1.2 también es válida la inclusión $\mathbb{N} \subseteq A$. $\Diamond$

Nota. El hecho que postula el enunciado del teorema anterior puede razonarse de una manera sencilla e intuitiva, y conviene que el lector no pierda de vista esta idea. En efecto, si A es un subconjunto de $\mathbb{N}$ que contiene a 1 y al sucesor de cada uno de sus elementos, resulta que $1 \in A$ y por lo tanto $2 \in A$, por ser el sucesor de 1. Por la misma razón $3 \in A$, luego $4 \in A$, luego $5 \in A$, ..., y así sucesivamente, para concluir que todo número natural n está en A, ya que el proceso de ir tomando el sucesor nos conducirá a n en un número finito de pasos.

La argumentación anterior, si bien no es incorrecta es algo vaga, ya que debieran aclararse con mayor rigor algunas cuestiones, como qué significan los puntos suspensivos o la expresión "en un número finito de pasos". El procedimiento inductivo (es decir, el uso del teorema de inducción) permite salvar estas objeciones cada vez que intentemos probar que un cierto subconjunto A de los números naturales es igual a $\mathbb{N}$, reduciendo el problema de lidiar con un número infinito de situaciones a la tarea de verificar o demostrar sólo dos hechos, a saber:

- Probar que $1 \in A$.
- (Paso inductivo). Suponer $n \in A$ (hipótesis inductiva) y probar entonces que $n + 1 \in A$.

Lo antedicho de ningún modo significa que toda demostración inductiva sea sencilla. Ella puede ser bastante complicada a veces, sobre todo el paso inductivo. De cualquier forma, la inducción es una herramienta indispensable para establecer muchas propiedades importantes de los números naturales, y actúa en realidad en todos los campos de la Matemática. En lo que sigue brindaremos al lector varias pruebas concretas de la veracidad de nuestras palabras. $\Diamond$

Suma y producto de numeros naturales. Las operaciones de suma y producto definidas en el conjunto de números reales pueden por supuesto restringirse al conjunto de números naturales. Mostraremos a continuación que $\mathbb{N}$ es una parte cerrada de $\mathbb{R}$ respecto de dichas operaciones.

Proposición 3.1.5 Si $m, n \in \mathbb{N}$ entonces $m + n \in \mathbb{N}$ y $mn \in \mathbb{N}$.

Demostracion. Fijado $m \in \mathbb{N}$, bastará probar que los conjuntos

$$S = \{x \in \mathbb{N} : m + x \in \mathbb{N}\} \quad \text{y} \quad P = \{x \in \mathbb{N} : mx \in \mathbb{N}\}$$

son inductivos. Comencemos por S.

Dado que $m+1 \in \mathbb{N}$, por ser $\mathbb{N}$ inductivo, resulta que $1 \in S$. Suponiendo que $x \in S$, en cuyo caso $m + x \in \mathbb{N}$, resulta que $m + (x + 1) = (m + x) + 1$ pertenece a $\mathbb{N}$. Luego $x + 1 \in S$ y S es inductivo.

Probemos ahora que P es inductivo. De $m\,1 = m$ resulta que $1 \in P$. Por otro lado, si $x \in P$ tenemos que $mx \in \mathbb{N}$ y por lo tanto

$$m(x + 1) = mx + m \in \mathbb{N},$$

pues ya hemos probado que la suma de números naturales es natural. En consecuencia $x + 1 \in P$, como queríamos probar. $\Diamond$

Diferencia de numeros naturales. Por cierto, la diferencia de dos números naturales no es en general un número natural, ya que por ejemplo $n - m$ es negativo si $n > m$. Para precisar, estableceremos a continuación una condición necesaria y suficiente para que la diferencia de dos números naturales sea un número natural, lo que nos dará por añadidura una caracterización intrínseca del orden en $\mathbb{N}$.

Proposición 3.1.6 Sean $m, n \in \mathbb{N}$. Entonces $n - m \in \mathbb{N}$ si y solo si $m < n$.

DEMOSTRACION. La necesidad de la condición es inmediata, pues todo número natural es positivo. Emplearemos el teorema de inducción para probar la recíproca, considerando el conjunto

$$D = \{\, a \in \mathbb{N} : b - a \in \mathbb{N} \text{ para todo número natural } b > a\} \ .$$

Sigue del ítem 2) de la proposición 3.1.3 que $1 \in D$. Para resolver el paso inductivo, supongamos que $x \in D$ y sea $b \in \mathbb{N}$ tal que $b > x + 1$. Puesto que $b > x$ (por transitividad), resulta por definición de D que $b - x \in \mathbb{N}$. Como además $b - x > 1$, obtenemos usando nuevamente la propiedad citada arriba que

$$b - (x + 1) = b - x - 1 \in \mathbb{N}\,,$$

y por lo tanto $x + 1 \in D$.

Hemos probado así que D es inductivo, y en consecuencia $D = N$.

Particularizando, resulta que $n - m \in \mathbb{N}$ si m y n son números naturales tales que $n > m$, pues $m \in D$. $\Diamond$

Buena ordenación.

> Un conjunto ordenado X se dice *bien ordenado* si y solo si todo subconjunto no vacío de X admite mínimo (o primer elemento).

Nuestro primer ejemplo es el conjunto de números naturales:

Teorema 3.1.7 (Principio de buena ordenación.) $\mathbb{N}$ es bien ordenado.

DEMOSTRACION. Consideremos el subconjunto $\mathcal{B}$ de $\mathbb{N}$ definido por la siguiente propiedad:

$n \in \mathcal{B}$ si y solo si todo subconjunto de $\mathbb{N}$ que contiene a n tiene mínimo.

Vamos a probar que $\mathcal{B} = \mathbb{N}$, mostrando que $\mathcal{B}$ es inductivo. Esto asegurará la validez del principio de buena ordenación, ya que si A es un subconjunto no vacío de $\mathbb{N}$ y m es un cualquiera de sus elementos, podremos concluir que A admite mínimo por ser m un elemento de $\mathcal{B}$.

Es claro que $1 \in \mathcal{B}$, pues obviamente 1 es el mínimo de de cualquier subconjunto de $\mathbb{N}$ que lo contenga. Supongamos ahora que $x \in \mathcal{B}$ y que S es un subconjunto de $\mathbb{N}$ que contiene a $x + 1$.

Si $x \in S$ entonces S tiene mínimo, por hipótesis inductiva. Si $x \notin S$, por la misma razón el conjunto $T = S \cup \{x\}$ admite mínimo, digamos a. Se presentan entonces dos posibilidades, a saber:

(i) $a = x$. Es claro entonces que $x + 1$ es el mínimo de S.

(ii) $a \neq x$. En este caso $a \in S$ y sigue inmediatamente que $a = \min S$.

Hemos probado así que todo subconjunto de $\mathbb{N}$ que contiene a $x + 1$ admite mínimo, y por lo tanto $x + 1 \in \mathcal{B}$, lo que completa la demostración del paso inductivo. $\Diamond$

NOTA Hemos probado el principio de buena ordenación utilizando el teorema de inducción. Recíprocamente, probaremos el teorema de inducción a partir del principio de buena ordenación, lo que mostrará que ambos teoremas son equivalentes.

Procederemos por el absurdo, suponiendo que $\mathbb{N}$ contiene un subconjunto inductivo propio A. En tal caso $\mathbb{N} - A$ es un subconjunto no vacío de $\mathbb{N}$, y admite entonces por el principio de buena ordenación un primer elemento m.

Siendo $m > 1$, pues $1 \in A$, resulta que $m-1$ es un número natural menor que el mínimo de $\mathbb{N} - A$, de donde deducimos que $m - 1 \in A$. Pero entonces $m = (m - 1) + 1 \in A$, por ser A inductivo, lo que es una contradicción.

INDUCCION GENERALIZADA. Si $m \in \mathbb{N}$, el conjunto

$$\{\, k \in \mathbb{N} : k \geq m \,\}$$

será llamado la $m-$ésima sección final de $\mathbb{N}$ y lo notaremos $\mathbb{N}_m$ (observemos que $\mathbb{N}_1 = \mathbb{N}$). Asimismo, designaremos por $\mathbb{N}_0$ el conjunto $\mathbb{N} \cup \{\, 0 \,\}$, cuyos elementos se denominan *enteros no negativos*.

El teorema de inducción, válido para subconjuntos de $\mathbb{N}$, se puede generalizar a $\mathbb{N}_m$. Definamos primero una noción equivalente a la de conjunto inductivo.

> Dado $m \in \mathbb{N}_0$ y dado $A \subseteq \mathbb{R}$, diremos que A es $m-$inductivo si y solo si se verifican las siguientes condiciones:
>
> (1) $m \in A$.
>
> (2) $x + 1 \in A$ para todo $x \in A$.

El siguiente enunciado generaliza naturalmente el teorema de inducción:

Teorema 3.1.8 (de inducción generalizado) Si $m \in \mathbb{N}_0$ y A es un subconjunto $m-$inductivo de $\mathbb{N}_m$ entonces $A = \mathbb{N}_m$.

DEMOSTRACION. Puesto que $m \in A$, deberemos probar que A contiene a todo número natural mayor que m. Suponiendo falso este hecho, podemos asegurar por el principio de buena ordenación la existencia de un número natural $s > m$ tal que $s \notin A$ y mínimo con respecto a esta propiedad. Puesto que $s - m \in \mathbb{N}$ (aún en el caso $m = 0$), sigue en particular que $s - m \geq 1$, o equivalentemente, $s - 1 \geq m$, es decir, $s - 1 \in \mathbb{N}_m$.

Por otro lado, la minimalidad de s implica que $s - 1 \in A$, de donde sigue por la condición (2) que $s \in A$ (absurdo). Luego $A = \mathbb{N}_m$. $\quad \Diamond$

Teniendo en cuenta el teorema 3.1.8 y la equivalencia entre el teorema de inducción y el principio de buena ordenación, no es difícil imaginar que el principio de buena ordenación (PBO) se extiende a $\mathbb{N}_m$ cualquiera sea $m \in \mathbb{N}_0$. Efectivamente, tenemos:

Teorema 3.1.9 (PBO generalizado) Todo subconjunto no vacío de $\mathbb{N}_m$ admite mínimo.

DEMOSTRACION. Queda a cargo del lector. $\Diamond$

Ejemplo 3.1.10 Probemos que

$$n^3 > 4n^2 + n + 19 \qquad (*)$$

para todo número natural $n \geq 5$, para lo cual bastará demostrar que el conjunto A de números naturales n mayores o iguales que 5 que satisfacen la desigualdad anterior es $5-$inductivo.

Es inmediato verificar que $5 \in A$, ya que $5^3 = 125$ y $4 \times 5^2 + 5 + 19 = 124$. Suponiendo ahora que $n \in \mathbb{N}_5$ satisface la desigualdad, veamos que también la satisface $n+1$. Operando convenientemente, tenemos:

$$\begin{aligned}(n+1)^3 &= n^3 + 3n^2 + 3n + 1 > 4n^2 + n + 19 + 3n^2 + 3n + 1 = \\ &= 4(n+1)^2 + (n+1) + 19 + 3n^2 - 5n - 4 = \\ &= 4(n+1)^2 + (n+1) + 19 + n(3n-5) - 4 > \\ &> 4(n+1)^2 + (n+1) + 19,\end{aligned}$$

pues $n(3n-5) - 4 \geq 46 > 0$, por ser $n \geq 5$. Luego $n+1 \in A$.

Agreguemos que 5 es el menor número natural que satisface (*), como puede comprobarse fácilmente. $\Diamond$

3.1.2. Ejercicios

1. Determinar cuáles de los siguientes subconjuntos de $\mathbb{R}$ son inductivos:

   a) $A_1 = \{m/n : m, n \in \mathbb{N}\}$

   b) $A_2 = \{x : |\, x + 1 \,| \in \mathbb{N}\}$

   c) $A_3 = \{x \in (0, +\infty) : x - 0{,}25 \in \mathbb{N}_0\} \cup \mathbb{N}$

   d) $A_5 = \{x : x \text{ es raíz de una ecuación de la forma } X^2 + bX + 1\}$

   e) $A_4 = \bigcup_{k \in \mathbb{N}_0} [2k, 2k+1]\,.$

2. Sean A y B subconjuntos inductivos de $\mathbb{R}$.

   a) Probar que $A \cap B$ y $A \cup B$ son conjuntos inductivos.

   b) Probar que el conjunto $D = \{a - b : a \in A \wedge b \in B\}$ es inductivo.

   c) ¿Es $S = \{a + b : a \in A \wedge b \in B\}$ un conjunto inductivo? ¿ Y si se agrega la condición de que $x - 1 \in A$ para todo $x \in A$?

3. Demostrar que un conjunto acotado superiormente no es inductivo.

4. Sean m y n enteros no negativos.

   a) Probar que $\mathbb{N}_m$ es cerrado respecto a la suma y el producto.

   b) Exhibir una función biyectiva $f : \mathbb{N}_m \to \mathbb{N}_n$.

5. *a*) Sean $m, n \in \mathbb{N}$ tales que $m < n$. Probar que $m + 1 \leq n$.

   b) Probar que $m + n \leq mn + 1$ cualesquiera sean $m, n \in \mathbb{N}$.

6. *a*) Sea $f : \mathbb{N} \to \mathbb{N}$ una función tal que $f(k+1) = f(k)+1$ para todo $k \in \mathbb{N}$. Analizar la inyectividad y suryectividad de f.

   b) Una función $f : \mathbb{N} \to \mathbb{R}$ se dice *estrictamente decreciente* si y solo si $x < y \Rightarrow f(y) < f(x)$. Probar que f es estrictamente decreciente si y solo si $f(k + 1) < f(k)$ para todo $k \in \mathbb{N}$.

   c) ¿Existe una función estrictamente decreciente de $\mathbb{N}$ en $\mathbb{N}$?

7. Si $m \in \mathbb{N}$, sea $\mathbb{I}_m = \{k \in \mathbb{N} : k \leq m\}$. Demostrar que todo subconjunto no vacío de $\mathbb{I}_m$ admite máximo.

8. Probar que todo $m \in \mathbb{N}_0$ puede expresarse en una y solo una de las formas $m = 2k$ ó $m = 2k + 1$, con $k \in \mathbb{N}_0$. Análogamente, m puede expresarse en una y solo una de las formas $m = 3t$, $m = 3t + 1$ ó $m = 3t + 2$ $(t \in \mathbb{N}_0)$.

9. *a*) Sea $f : \mathbb{N}_0 \to \mathbb{N}_0$ una función tal que $f(x + y) = f(x) + f(y) - 2$. Calcular $f(50)$ sabiendo que $f(3) = 11$.

   b) Determinar todas las biyecciones f de $\mathbb{N}_0$ tales que

   $$f(x + y) = f(x) + f(y)$$

   cualesquiera sean $x, y \in \mathbb{N}_0$.

10. *a*) Determinar el máximo $m \in \mathbb{N}$ tal que $(m-1)^3 \leq 12m^2 - 3m - 9$.

    b) Probar que dado $m \in \mathbb{N}_{11}$ existen números naturales a y b tales que $m = 2a + 5b$.

11. Un número natural se dice *primo* si y solo si no puede descomponerse como producto de dos números naturales mayores que 1. Demostrar, usando el principio de buena ordenación, que todo elemento de $\mathbb{N}_2$ es divisible por algún número primo.

3.2. Definiciones inductivas y recursivas

3.2.1. Sucesiones

Comenzaremos introduciendo algunas notaciones útiles. Si $r, s \in \mathbb{N}_0$ y $r \leq s$, designaremos por $\mathbb{I}_{r,\,s}$ el conjunto

$$\{\, k \in \mathbb{N}_0 : r \leq k \leq s \,\},$$

que muchas veces representaremos más familiarmente en la forma

$$\mathbb{I}_{r,\,s} = \{\, r, r+1, \ldots, s \,\}.$$

Como ya dijimos, el caso especial $\mathbb{I}_{1,\,m}$ se notará simplemente $\mathbb{I}_m$. Por ejemplo,

$$\mathbb{I}_{2,\,6} = \{2,3,4,5,6\}\ ,\ \mathbb{I}_4 = \{1,2,3,4\}\ ,\ \mathbb{I}_{8,\,8} = \{8\}.$$

Tanto los conjuntos $\mathbb{I}_{r,\,s}$ como las secciones $\mathbb{N}_m$ serán denominados *intervalos enteros* de $\mathbb{N}_0$.

> Si X es un conjunto, toda función $f : D \mapsto X$, donde D es un intervalo entero de $\mathbb{N}_0$, se llama una *sucesión* en X (o sucesión de elementos de X). Emplearemos a veces la palabra *secuencia* como sinónimo de sucesión.

La sucesión se dice *finita* si D es un intervalo del tipo $\mathbb{I}_{r,\,s}$. Si $D = \mathbb{N}_m$ para algún m, se dice *infinita*.

Hablando informalmente, una sucesión en X puede pensarse como una serie ordenada de elementos de X; un primer *término*, imagen del primer elemento del dominio D, un segundo término, imagen del sucesor del primer elemento de D, y así siguiendo. Coherentemente con esta descripción, la sucesión (función) f suele notarse en la forma $(x_n)_{n\in D}$, donde $x_n = f(n)$ se llama el término de orden n o término $n-$ésimo de la misma.

Si $D = \mathbb{I}_{r,\,s}$ se escribe $(x_n)_{\,r\,\leq\, n \leq s}$ y se la representa en la forma más sugestiva

$$x_r,\ x_{r+1},\ \ldots,\ x_s\,,$$

mientras que en el caso infinito se utilizan las notaciones $(x_n)_{n\,\geq\, m}$ ó

$$x_m,\ x_{m+1},\ x_{m+2}, \cdots\cdots$$

Ejemplos 3.2.1 Los siguientes son ejemplos de sucesiones en $\mathbb{R}$ con dominio $\mathbb{N}$. Especificamos en cada caso el término $n-$ésimo de la sucesión y exhibimos sus primeros términos:

1) $x_n = n^3 \quad , \quad (x_n) = 1,\, 8,\, 27,\, 64,\, 125,\, \cdots\cdots$

2) $x_n = \frac{2n+1}{n} \quad , \quad (x_n) = 3\,,\, \frac{5}{2}\,,\, \frac{7}{3}\,,\, \frac{9}{4}\,,\, \cdots\cdots$

3) $x_n = (-1)^n \quad , \quad (x_n) = -1, 1, -1, 1, -1, \cdots\cdots$

4) $x_n = \{\frac{3n}{10}\} \quad , \quad (x_n) = 0.3, 0.6, 0.9, 0.2, 0.5, \cdots\cdots$

5) $x_n = 1 + \sqrt{2} \quad , \quad (x_n) = 1 + \sqrt{2}, 1 + \sqrt{2}, 1 + \sqrt{2}, \cdots\cdots$

6) $x_n = (2n-1)n^2 \quad , \quad (x_n) = 1, 12, 45, 112, 225, \cdots\cdots$

Puede observarse que en los casos 3) y 5) la sucesión sólo toma un número finito de valores distintos (la secuencia 5) se dice *constante*), y lo mismo ocurre en el caso 4), como el lector puede demostrar fácilmente. De todos modos, debe quedar claro que todas ellas tienen infinitos términos. ◊

Ejemplo 3.2.2 Si a y r son números reales, la sucesión $(x_n)_{n\geq 0}$ de números reales definida por

$$x_n = a + nr$$

se llama *progresión aritmética* de razón r y término inicial a. Observemos que

$$x_{n+1} = a + (n+1)r = a + nr + r = x_n + r\,,$$

vale decir, comenzando por $x_0 = a$ cada término de una progresión aritmética se obtiene sumando al anterior la razón r.

Como ejemplos sencillos de progresiones aritméticas mencionemos la sucesión de números naturales, de término inicial 1 y razón 1, y más generalmente las secciones $\mathbb{N}_m$, de término inicial m y razón 1. También las sucesiones constantes son progresiones ariméticas (de razón 0).

Como ejercicio sencillo, encargamos al lector la tarea de verificar que la secuencia

$$-17/6\,,\ 1/3\,,\ 17/6\,,\ 16/3\,,\ 47/6\,,\ \cdots\cdots$$

es una progresión aritmética. ◊

Definiciones inductivas.

En los ejemplos anteriores hemos definido algunas sucesiones de números reales en una forma que podemos llamar explícita, especificando para cada n el valor del término n-ésimo x_n. Cabe decir ahora que ésta no es la única ni la más importante de las formas de definición de una secuencia.

Por ejemplo, y apelando nuevamente a nuestra intuición sobre la validez de los procedimientos inductivos, para definir una secuencia de elementos en un cierto conjunto parece claro que bastaría especificar su primer término y establecer una regla que permita determinar cualquier término conocido su predecesor.

> Precisando un poco más esta idea, diremos que una sucesión no explícita (x_n) de elementos de un conjunto X con dominio $\mathbb{N}_m$ está definida *inductivamente* si se especifica el valor de x_m y, para cada $k \in \mathbb{N}_m$, el valor de x_{k+1} está unívocamente determinado por el valor de x_k.

Antes de mostrar algunos ejemplos instructivos reflexionemos nuevamente sobre la cuestión. Como en el caso del teorema de inducción, la idea subyacente es clara: se conoce x_m y por lo tanto su sucesor x_{m+1}; por la misma razón se conoce x_{m+2}, luego x_{m+3}, y así siguiendo. Naturalmente, el asunto de la existencia y unicidad de una tal secuencia requeriría una demostración rigurosa, para lo cual sería necesario precisar algunas de las expresiones usadas en la definición, por ejemplo aclarar qué significa exactamente que cualquier término determina unívocamente el siguiente. Sin embargo, no ofreceremos una tal prueba en estas páginas, ya que creemos que sus dificultades (sobre todo formales) confundirían al novel lector, con el riesgo de hacerle perder de vista el fondo de la cuestión, que insistimos, es completamente intuitivo. Nuestra aspiración básica es que comprenda el sentido de una definición inductiva, y que sea luego capaz de usarla eficazmente.

Ejemplo 3.2.3 Consideremos la sucesión de múltiplos de un número real c, esto es, la sucesión $(a_n)_{n\geq 1}$ definida explícitamente por $a_n = nc$. Notando que $a_1 = c$ y $a_{n+1} = (n+1)c = nc + c = a_n + c$, resulta que $(a_n)_{n\geq 1}$ satisface las condiciones

$$\begin{cases} a_1 = c \\ a_{n+1} = a_n + c \quad \text{si } n \geq 1. \end{cases}$$

Para completar la prueba de que $(a_n)_{n\geq 1}$ está definida inductivamente por estas condiciones, debemos probar que ella es la única sucesión en $\mathbb{R}$ que las satisface. Supongamos para ello que $(b_n)_{n\geq 1}$ es cualquier otra sucesión en $\mathbb{R}$ con las mismas propiedades y demostremos que $a_n = b_n$ para todo $n \in \mathbb{N}$, para lo cual bastará probar que el conjunto $A = \{n \in \mathbb{N} : b_n = a_n\}$ es inductivo.

Es inmediato que $1 \in A$, pues $b_1 = c = a_1$.

Para el paso inductivo, supongamos que $k \in A$. Entonces

$$b_{k+1} = b_k + c = a_k + c = a_{k+1},$$

y por lo tanto $k + 1 \in A$, como queríamos probar. $\Diamond$

Ejemplo 3.2.4 Las definiciones inductivas nos brindan una manera alternativa de presentar sucesiones ya conocidas, pero su mayor importancia reside en el hecho de a que a través de ellas podemos definir nuevas sucesiones mediante relaciones preestablecidas entre sus términos.

Por ejemplo, consideremos la secuencia $(x_n)_{n\geq 1}$ de números naturales definida inductivamente por las relaciones

$$\begin{cases} x_1 = 1 \\ x_{n+1} = x_n + (n+1) \quad \text{si } n \geq 1. \end{cases}$$

Si bien todavía no vamos a determinar en forma explícita el término general de esta sucesión, el cálculo de un término en particular, como por

ejemplo el cuarto, nos orientará respecto a dicha cuestión. Aplicando reiteradamente las condiciones dadas tenemos:

$$x_4 = x_3 + 4 = x_2 + 3 + 4 = x_1 + 2 + 3 + 4 = 1 + 2 + 3 + 4 = 10,$$

vale decir, x_4 es la suma de los primeros 4 números naturales, y parece razonable conjeturar que x_n es en general la suma de los primeros n números naturales. Probaremos más adelante este hecho, con el agregado de que brindaremos una fórmula cerrada para calcular dicha suma.

En forma similar podemos intentar definir inductivamente una secuencia cuyo término $n-$ésimo exprese el producto de los primeros n números naturales, a saber la secuencia definida en la forma

$$\begin{cases} y_1 = 1 \\ y_{n+1} = (n+1)\, y_n \quad \text{si } n \geq 1. \end{cases}$$

Por ejemplo,

$$y_4 = 4y_3 = (4\,.\,3)y_2 = (4\,.\,3\,.\,2)y_1 = 4\,.\,3\,.\,2\,.\,1 = 24\,.$$

Volveremos más tarde sobre esta sucesión, de gran relieve en Matemática. Su término general y_n (notado universalmente $n!$) se denomina el *factorial* de n. $\Diamond$

Definiciones recursivas.

Nos ocuparemos ahora de un método de definición de secuencias más general que el método inductivo, en el que cada término no necesariamente depende del predecesor sino de uno o varios de los anteriores. Precisamente:

> Diremos que una sucesión no explícita $(x_n)_{n \geq m}$ de elementos de un conjunto X está definida *recursivamente* si se especifican los valores de los términos iniciales $x_m, \ldots, x_r$ $(r \geq m)$ y, para cada $k \geq r$, el valor de x_{k+1} queda unívocamente determinado por uno o más de los valores anteriores $x_m, \ldots, x_k$. Los valores iniciales dados y la forma particular de determinar cada término en función de los anteriores se denominan las *reglas de recurrencia* de la sucesión.

Es claro que las definiciones inductivas constituyen un caso particular de las definiciones recursivas, y como en el caso de aquellas, tampoco probaremos formalmente en esta ocasión que existe una y solo una sucesión que satisface las condiciones impuestas (teorema de recursión). Exhibiremos en cambio un par de ejemplos, confiando en que ellos contribuyan a familiarizar al lector con esta forma de definición.

Ejemplos 3.2.5 Comencemos por una de las más famosas secuencias de la Matemática, la *sucesión* $(F_n)_{n\geq 0}$ *de Fibonacci* (1170-1250), definida recursivamente por las condiciones

$$\begin{cases} F_0 = 0\ ,\ F_1 = 1 \\ F_{k+1} = F_k + F_{k-1} \quad \text{si } k \geq 1\,, \end{cases}$$

vale decir, sus dos valores iniciales son 0 y 1 y a partir del siguiente cada término es suma de los dos inmediatamente anteriores. De acuerdo con esta regla, sus primeros nueve valores son

$$0\,,\,1\,,\,1\,,\,2\,,\,3\,,\,5\,,\,8\,,\,13\,,\,21\,.$$

Si bien sus términos parecen crecer lentamente, se trata de una impresión engañosa. Por ejemplo, un término de índice relativamente bajo, como F_{50}, supera los diez mil millones. En realidad, el crecimiento de la sucesión de Fibonacci es exponencial, ya que sus elementos responden (algo sorprendentemente) a la fórmula explícita

$$F_n = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right)\,.$$

Propondremos al lector que demuestre este hecho en uno de los ejercicios del final de esta sección, usando la noción de potencia $n-$ésima de un número real adquirida en la escuela media, que refiere la expresión a^n al producto

$$a\,.\,a\,.\,.\,.\,a \quad (n \text{ veces } a)\,.$$

Definiendo además $a^0 = 1$, es claro que dichas potencias verifican la regla de recurrencia

$$a^{k+1} = a\,.\,a^k\,,$$

cualquiera sea $k \in \mathbb{N}_0$. En la próxima sección volveremos con mayor detalle sobre el tema.

Mencionemos por último que los números de Fibonacci satisfacen una gran cantidad de propiedades interesantes y aparecen naturalmente en diversas cuestiones de la Matemática y aún de otras disciplinas. Prueba de ello es que existe una publicación periódica (el *Fibonacci Quarterly*) dedicada por entero a difundir hechos relevantes relacionados con los mismos.

Como otro ejemplo (menos interesante) de recursividad, consideremos la sucesión $(z_n)_{n\geq 1}$ definida recursivamente por las relaciones

$$\begin{cases} z_1 = 0\ ,\ z_2 = 2\ ,\ z_3 = 1\ ,\ z_4 = 3 \\ z_{k+1} = z_k - z_{k-1} + z_{k-2} - z_{k-3} \quad \text{si } k \geq 4. \end{cases}$$

Aplicando las reglas de recurrencia, podemos observar que la tabla de los primeros 14 términos de la sucesión es

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14
z_n	0	2	1	3	4	0	-2	-1	-3	-4	0	2	1	3

,

observándose que del undécimo al decimocuarto término se repiten en orden los cuatro primeros términos. Puesto que cada término depende de los cuatro anteriores, y la regla de formación es siempre la misma, concluimos que la secuencia es periódica, repitiéndose infinitamente el bloque de longitud diez

$$0\,,\,2\,,\,1\,,\,3\,,\,4\,,\,0\,,\,-2\,,\,-1\,,\,-3\,,\,-4\,. \quad \Diamond$$

3.2.2. Ejercicios

1. En cada uno de los siguientes ítems se indica el dominio D y los primeros términos de una sucesión (x_n) de números reales. Determinar en cada caso el término general x_n de la misma:

   a) $D = \mathbb{N}$; $(x_n) = 1, -2, 3, -4, 5, -6, \ldots$

   b) $D = \mathbb{N}_4$; $(x_n) = 2, 5, 8, 11, 14, 17, \ldots$

   c) $D = \mathbb{N}$; $(x_n) = 1, 5, 5, 9, 9, 13, \ldots$

   d) $D = \mathbb{N}_0$; $(x_n) = 2, 3, 13, 35, 97, \ldots$

   e) $D = \mathbb{N}$; $(x_n) = 3, 1/3, 1/3, 1/27, 1/27, 1/243, \ldots$

   f) $D = \mathbb{N}_2$; $(x_n) = 6, 12, 20, 30, 42, \ldots.$

2. Sea $p \in \mathbb{N}$. Una secuencia $(a_n)_{n \geq m}$ se dice *periódica* (de período p) si y solo si $a_{k+p} = a_k$ para todo $k \geq m$ y p es mínimo respecto a esta propiedad.

   a) Exhibir ejemplos, determinando su término general a_n, de secuencias de período 1, 2 y 4, respectivamente.

   b) Para cada $n \in \mathbb{N}_0$ sea b_n el dígito de las unidades de n^2. Probar que (b_n) es periódica y determinar su período.

   c) Como en *b*), donde ahora b_n es el resto de dividir F_n por 8.

3. *a*) ¿Es posible definir cinco progresiones aritméticas de números naturales de razón 4 de manera que sus respectivas imágenes sean disjuntas dos a dos?

   b) Determinar una progresión aritmética no constante en $\mathbb{N}_{40}$ cuyos términos sean múltiplos de 16 y el dígito de las unidades de cualquiera de ellos sea 2.

4. Brindar una definición inductiva de cada una de las siguientes sucesiones (en cada caso se indica el dominio y el término general, o bien se muestran los primeros términos de la secuencia):

   a) $D = \mathbb{N}\ ;\ a_n = n^2$

   b) $D = \mathbb{N}_0\ ;\ a_n = n^2 + n + 1$

   c) $D = \mathbb{N}\ ;\ (a_n) = 1\,, 2\,, 2/3\,, 8/3\,, 8/15\,, 16/5\,, \cdots\cdots$

   d) $D = \mathbb{N}_0\ ;\ a_n = \dfrac{2n - 3\left(1 + (-1)^{n+1}\right)}{4}$.

5. Brindar una definición recursiva de cada una de las siguientes sucesiones:

   a) $D = \mathbb{N}\ ;\ (a_n) = 5\,, 2\,, 3\,, -1\,, 4\,, -5\,, 9\,, -14\,, 23\,, \cdots\cdots$

   b) $D = \mathbb{N}_0\ ;\ (a_n) = 3\,, 3\,, 6\,, 12\,, 24\,, 48\,, \cdots\cdots$

   c) $D = \mathbb{N}_0\ ;\ a_n = n(b - c) + c\ \ (b, c \in \mathbb{R})$

   d) $D = \mathbb{N}_0\ ;\ a_n$ es el número de secuencias de $0's$ y $1's$ (de longitud n) que no contienen dos ceros consecutivos.

6. En cada uno de los siguientes casos determinar el término general de la sucesión definida por las reglas de recurrencia dadas:

   a) $a_0 = 2\ ,\ a_{n+1} = a_n + 4\left(1 + (-1)^{n+1}\right)$

   b) $a_0 = 1\ ,\ a_{n+1} = \left(1 + a_n^{-1}\right)^{-1}$

   c) $a_0 = 0\ ,\ a_{n+1} = n + a_n$

   d) $a_0 = 1\ ,\ a_1 = -2\ ,\ a_n + a_{n-1} + a_{n-2} = 0 \quad (n \geq 2)$.

7. Sean $a, b \in \mathbb{R}$ y sea $(c_n)_{n \geq 0}$ la sucesión definida por las siguientes reglas de recurrencia:

$$\begin{cases} c_0 = a,\ c_1 = b \\ c_k = c_{k-1} + c_{k-2} \quad \text{si } k \geq 2\,. \end{cases}$$

   Definiendo $F_{-1} = 1$, probar que $c_n = aF_{n-1} + bF_n$ para todo $n \geq 0$.

8. Sean α y β las soluciones reales de la ecuación $X^2 - X - 1 = 0$.

   a) Demostrar que la sucesión de potencias de α y la sucesión de potencias de β satisfacen reglas de recurrencia del tipo de las descriptas en el ejercicio 7.

   b) Demostrar la fórmula exhibida en el ejemplo 3.2.5 para el cálculo de los números de Fibonacci.

3.3. El Principio de inducción

3.3.1. Forma proposicional

Supongamos definida una función proposicional $\mathrm{P}(n)$ con conjunto de referencia $\mathbb{N}$, vale decir, $\mathrm{P}(n)$ es una proposición para cada $n \in \mathbb{N}$ (suele decirse en tal caso que P es una proposición *predicable* en $\mathbb{N}$). Por ejemplo, consideremos el enunciado

$$\mathrm{P}(n) : n^2 + n \text{ es par } .$$

Volviendo a la situación general, supongamos además que deseamos probar que $\mathrm{P}(n)$ es verdadera para todo $n \in \mathbb{N}$. Si bien esto nos obligaría en principio a demostrar la validez de infinitas proposiciones, contamos para hacerlo con una herramienta fundamental, llamada *principio de inducción*, que reduce la cuestión a la demostración de solo dos hechos. El mismo es en realidad la versión proposicional del teorema de inducción, y su enunciado es el siguiente:

Teorema 3.3.1 (Principio de Inducción) Sea P una proposición predicable en $\mathbb{N}$ que verifica las dos siguientes condiciones:

1) $\mathrm{P}(1)$ es verdadera.

2) La implicación $\mathrm{P}(k) \Rightarrow \mathrm{P}(k+1)$ es verdadera cualquiera sea $k \in \mathbb{N}$.

Entonces $\mathrm{P}(n)$ es verdadera para todo $n \in \mathbb{N}$.

Demostracion. La validez de la conclusión sigue inmediatamente del teorema de inducción, ya que las hipótesis aseguran que el conjunto

$$A = \{\, n \in \mathbb{N} : \mathrm{P}(n) \text{ es verdadera}\}$$

es inductivo, y por lo tanto $A = \mathbb{N}$. $\quad \Diamond$

Nota. Como ya señalamos, el uso del principio de inducción requiere dos pasos: la prueba o verificación de que $\mathrm{P}(1)$ es verdadera y la demostración de que la implicación

$$\mathrm{P}(k) \Rightarrow \mathrm{P}(k+1)$$

es verdadera, llamada el *paso inductivo*. Suele emplearse el método directo de demostración, esto es, se supone $\mathrm{P}(k)$ verdadera (*hipótesis inductiva*) y se demuestra a partir de ello que $\mathrm{P}(k+1)$ es verdadera. Notemos que este procedimiento de ninguna manera viola una regla de oro de las demostraciones, que consiste en no asumir nunca como cierto lo que se quiere probar. En el paso inductivo lo único que se prueba es la verdad de la implicación, estableciendo una relación de causalidad entre la verdad de $\mathrm{P}(k)$ y la de $\mathrm{P}(k+1)$, hecho que tomado aisladamente no brinda información sobre los valores de verdad de ninguna de las premisas. De todas maneras, y al igual que otros

resultados anteriores, la validez del principio de inducción es completamente intuitiva: puesto que P(1) es verdadera y P(1) implica P(2) entonces P(2) es verdadera; puesto que P(2) implica P(3) entonces P(3) es verdadera, y así *ad infinitum*.

Digamos también que en ciertas ocasiones, y por razones de comodidad, demostraremos en el paso inductivo que la implicación

$$\mathrm{P}(k-1) \Rightarrow \mathrm{P}(k)$$

es verdadera para todo $k > 1$, lo que es claramente equivalente a la condición 2) del enunciado del principio de inducción.

Puesto que brinda un lenguaje más fluido, en lo que sigue utilizaremos preferentemente el principio de inducción antes que el teorema de inducción, a pesar de que podríamos optar por cualquiera de los dos, ya que ambos resultados son lógicamente equivalentes. En efecto, hemos probado ya el principio de inducción usando el teorema de inducción. Recíprocamente, supongamos que A es un subconjunto inductivo de $\mathbb{N}$ y consideremos la función proposicional P definida sobre $\mathbb{N}$ por

$$\mathrm{P}(n) : n \in A\,.$$

Dado que A es inductivo, es claro que P verifica las condiciones 1) y 2) del principio de inducción. Luego $\mathrm{P}(n)$ es verdadera para todo $n \in \mathbb{N}$, lo que equivale a decir que $A = \mathbb{N}$, como queríamos probar.

Para finalizar estos comentarios, digamos que existen distintas versiones del principio de inducción (todas ellas lógicamente equivalentes entre sí), que iremos presentando a medida que las circunstancias lo requieran. Naturalmente, el lector deberá discernir según el caso cuál es la versión apropiada para el problema que debe tratar.

Ejemplo 3.3.2 Probemos por inducción que la proposición

$$\mathrm{P}(n) : n^2 + n \text{ es par}$$

es verdadera para todo $n \in \mathbb{N}$.

Puesto que el caso $n = 1$ es inmediato ($1^2 + 1 = 2$), asumamos que $\mathrm{P}(k)$ es verdadera para un cierto $k \in \mathbb{N}$ y demostremos que $\mathrm{P}(k+1)$ también lo es. En otros términos, suponiendo que $k^2 + k$ es par, debemos probar que $(k+1)^2 + (k+1)$ es par. Escribiendo $k^2 + k = 2m\,(m \in \mathbb{N})$, y efectuando las correspondientes operaciones, tenemos:

$$\begin{aligned}(k+1)^2 + (k+1) = k^2 + 2k + 1 + k + 1 = k^2 + k + 2k + 2 = \\ = 2m + 2k + 2 = 2(m + k + 1)\,,\end{aligned}$$

lo que prueba nuestro resultado, ya que $m + k + 1 \in \mathbb{N}$. $\quad \Diamond$

Los símbolos de sumatoria y productoria.

Si bien la suma y el producto de números reales son operaciones binarias, frecuentemente es necesario sumar o multiplicar cantidades arbitrarias de números. Por caso, en 3.2.4 de la sección anterior nos referimos a la suma y al producto de los primeros n números naturales, informalmente notados en las formas

$$1 + 2 + \cdots + n \qquad \text{y} \qquad 1 \,.\, 2 \ldots n \,.$$

Sin perder de vista lo que ellas realmente significan, y con el propósito de manipularlas efizcamente, vamos a dar definiciones inductivas rigurosas de tales operaciones. Introduciremos además simbologías especiales para notarlas, que eliminarán la ligera ambigüedad de los puntos suspensivos.

El simbolo de sumatoria. Sea $(a_n)_{n\geq 1}$ una sucesión en $\mathbb{R}$ y sea $(x_n)_{n\geq 1}$ la sucesión de números reales definida inductivamente en la forma

$$\begin{cases} x_1 = a_1 \\ x_{n+1} = x_n + a_{n+1} \quad \text{si } n \geq 1 \,. \end{cases}$$

Por ejemplo, $x_2 = a_1 + a_2$, $x_3 = x_2 + a_3 = a_1 + a_2 + a_3$, etc., resultando evidente que su término $n-$ésimo x_n es $a_1 + a_2 + \cdots + a_n$, lo que nos brinda una definición formal de esta suma. Alternativamente, una notación cómoda y sugerente para dicho término general es

$$x_n = \sum_{i=1}^{n} a_i \,,$$

que indica que x_n es la suma de los primeros n términos de la secuencia. Por ejemplo, la suma $1 + 2 + \cdots + n$ se representa como $\displaystyle\sum_{i=1}^{n} i$, mientras que

$$\sum_{i=1}^{5} (-1)^{i+1}(2i - 1) = 1 - 3 + 5 - 7 + 9 \,.$$

El símbolo $\sum$ (equivalente a la σ mayúscula en el alfabeto griego) se denomina *símbolo de sumatoria*. Puesto que lo usaremos continuamente, vale la pena remarcar su definición inductiva:

$$\begin{cases} \displaystyle\sum_{i=1}^{1} a_i = a_1 \\[2ex] \displaystyle\sum_{i=1}^{n+1} a_i = \left(\sum_{i=1}^{n} a_i \right) + a_{n+1} \quad \text{si } n \geq 1. \end{cases}$$

Ilustremos el uso de la definición inductiva de sumatoria en el caso de una sucesión constante. Concretamente, probemos la validez de la fórmula

$$\sum_{i=1}^{n} c = nc, \qquad (*)$$

hecho que probablemente le parezca evidente al lector.

Procediendo por indución, sigue inmediatamente que la igualdad es válida para $n = 1$, ya que $\sum_{i=1}^{1} c = c = 1\,c$. Suponiéndola válida para un número natural k cualquiera, tenemos:

$$\sum_{i=1}^{k+1} c = \sum_{i=1}^{k} c + c = kc + c = (k+1)c,$$

lo que muestra que también vale para $n = k + 1$. Concluimos entonces por el principio de inducción que (*) es válida para todo $n \in \mathbb{N}$.

Veamos otros ejemplos más interesantes.

Ejemplo 3.3.3 De acuerdo con lo prometido, mostremos una fórmula explícita para calcular la suma de los primeros n números naturales. Afirmamos que

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$

para todo $n \in \mathbb{N}$.

Designando por $\mathrm{P}(n)$ el enunciado cuya validez queremos probar, es inmediato verificar que $\mathrm{P}(1)$ es verdadera, ya que

$$\sum_{i=1}^{1} i = 1 = \frac{1(1+1)}{2}\ .$$

Para el paso inductivo, supongamos que $\mathrm{P}(k)$ es verdadera ($k \in \mathbb{N}$) y probemos que también lo es $\mathrm{P}(k+1)$. Aplicando la definición inductiva de sumatoria, y teniendo en cuenta que $a_i = i$ en este ejemplo, tenemos que

$$\sum_{i=1}^{k+1} i = \sum_{i=1}^{k} i + (k+1) = \frac{k(k+1)}{2} + k + 1 = \frac{(k+1)(k+2)}{2}\ ,$$

como se comprueba fácilmente (obsérvese que en la segunda igualdad hemos usado la hipótesis inductiva). Luego $\mathrm{P}(k+1)$ es verdadera y sigue por el principio de inducción que $\mathrm{P}(n)$ es verdadera para todo n. $\Diamond$

Ejemplo 3.3.4 Establezcamos una fórmula para la sumatoria

$$S_n = \sum_{i=1}^{n} (-1)^i i.$$

Si calculamos los primeros valores de estas sumas, buscando algún patrón de formación de las mismas, obtenemos

$$S_1 = -1 \ , \ S_2 = 1 \ , \ S_3 = -2 \ , \ S_4 = 2 \ , \ S_5 = -3 \text{ y } S_6 = 3 \, ,$$

lo que nos muestra que vamos obteniendo todos los números naturales y sus inversos aditivos, los primeros para valores pares de n y los negativos para valores impares de n. Observando con mayor detalle los resultados, conjeturamos que

$$S_n = \begin{cases} \frac{n}{2} & \text{si } n \text{ es par} \\ -\frac{n+1}{2} & \text{si } n \text{ es impar,} \end{cases}$$

fórmula que efectivamente se satisface para los primeros seis valores de n. Probaremos inductivamente que es válida para todo $n \in \mathbb{N}$.

Para $n = 1$ la fórmula ya ha sido verificada. Supongamos ahora que vale para un cierto $k \in \mathbb{N}$ y probemos que vale para $k + 1$. Dada la naturaleza de nuestra conjetura, es natural distinguir dos casos:

(i) k par; sigue entonces que

$$S_{k+1} = S_k + (-1)^{k+1}(k+1) = \frac{k}{2} - (k+1) = \frac{-k-2}{2} = -\frac{(k+1)+1}{2} \ ,$$

lo que muestra la validez de nuestra conjetura para $k + 1$.

(ii) k impar; en tal caso

$$S_{k+1} = S_k + (-1)^{k+1}(k+1) = -\frac{k+1}{2} + (k+1) = \frac{k+1}{2} \, ,$$

como queríamos probar. Luego, por el principio de inducción, la fórmula es válida para todo $n \in \mathbb{N}$. $\quad \Diamond$

Establecemos a continuación algunas propiedades relevantes del símbolo de sumatoria, que esencialmente se deducen de las propiedades de asociatividad, conmutatividad, distributividad, etc., de la suma y el producto en $\mathbb{R}$. Si bien las demostraremos formalmente, empleando la definición inductiva de sumatoria, la escritura en cada caso de un pocos términos de las sumas involucradas convencerá rápidamente al lector de la naturalidad de dichas propiedades (en lo que sigue a designa un número real y (u_n) y (v_n) sucesiones en $\mathbb{R}$).

Proposición 3.3.5 Las siguientes igualdades son válidas para todo $m \in \mathbb{N}$:

1) $\displaystyle\sum_{i=1}^{m}(u_i+v_i)=\sum_{i=1}^{m}u_i+\sum_{i=1}^{m}v_i$

2) $\displaystyle\sum_{i=1}^{m}au_i=a\sum_{i=1}^{m}u_i$

3) $\displaystyle\sum_{i=1}^{m}(u_{i+1}-u_i)=u_{m+1}-u_1$

4) $\displaystyle\sum_{i=1}^{m}u_i=\sum_{i=1}^{m}u_{m-i+1}$

5) $\displaystyle\sum_{i=1}^{r+m}u_i=\sum_{i=1}^{r}u_i+\sum_{i=1}^{m}u_{r+i}\ \ (r\in\mathbb{N})\,.$

DEMOSTRACION. Naturalmente, procederemos por inducción en m. Puesto que todas las igualdades son evidentes para $m=1$, nos ocuparemos directamente del paso inductivo, suponiendo que la propiedad en cuestión es válida para un cierto $k\in\mathbb{N}$ y probando entonces que también vale para $k+1$.

1)

$$\begin{aligned}\sum_{i=1}^{k+1}(u_i+v_i)&=\sum_{i=1}^{k}(u_i+v_i)+(u_{k+1}+v_{k+1})=\\&=\sum_{i=1}^{k}u_i+\sum_{i=1}^{k}v_i+(u_{k+1}+v_{k+1})=\\&=\sum_{i=1}^{k}u_i+u_{k+1}+\sum_{i=1}^{k}v_i+v_{k+1}=\sum_{i=1}^{k+1}u_i+\sum_{i=1}^{k+1}v_i\,.\end{aligned}$$

2)

$$\begin{aligned}\sum_{i=1}^{k+1}au_i&=\sum_{i=1}^{k}au_i+au_{k+1}=a\sum_{i=1}^{k}u_i+au_{k+1}=\\&=a\left(\sum_{i=1}^{k}u_i+u_{k+1}\right)=a\sum_{i=1}^{k+1}u_i\,.\end{aligned}$$

3)

$$\begin{aligned}\sum_{i=1}^{k+1}(u_{i+1}-u_i)&=\sum_{i=1}^{k}(u_{i+1}-u_i)+u_{k+2}-u_{k+1}=\\&=u_{k+1}-u_1+u_{k+2}-u_{k+1}=u_{k+2}-u_1\,.\end{aligned}$$

4) Consideremos la sucesión (w_n) definida por $w_n = u_{n+1}$ para todo $n \in \mathbb{N}$. Sigue entonces de 1), 2) y 3) que

$$\sum_{i=1}^{k+1} u_i = \sum_{i=1}^{k+1} u_{i+1} - u_{k+2} + u_1 = \sum_{i=1}^{k+1} w_i - u_{k+2} + u_1 =$$
$$= \sum_{i=1}^{k} w_i + w_{k+1} - u_{k+2} + u_1 = \sum_{i=1}^{k} w_i + u_1 = \sum_{i=1}^{k} w_{m-i+1} + u_1 =$$
$$= \sum_{i=1}^{k} u_{k-i+2} + u_{k+1-(k+1)+1} = \sum_{i=1}^{k+1} u_{(k+1)-i+1}\,.$$

5)

$$\sum_{i=1}^{r+k+1} u_i = \sum_{i=1}^{r+k} u_i + u_{r+k+1} = \sum_{i=1}^{r} u_i + \sum_{i=1}^{k} u_{r+i} + u_{r+k+1} =$$
$$= \sum_{i=1}^{r} u_i + \sum_{i=1}^{k+1} u_{r+i}\,. \quad \Diamond$$

El simbolo de productoria. Definiremos ahora un símbolo análogo al de sumatoria, correspondiente a la operación producto. Concretamente, dada una sucesión de números reales $(a_n)_{n\geq 1}$, sea $(y_n)_{n\geq 1}$ la sucesión en $\mathbb{R}$ definida inductivamente en la forma

$$\begin{cases} y_1 = a_1 \\ y_{n+1} = y_n\, a_{n+1} \quad \text{si } n \geq 1\,. \end{cases}$$

Así, $y_2 = a_1\, a_2$, $y_3 = a_1\, a_2\, a_3$, $y_4 = a_1\, a_2\, a_3\, a_4$, etc. Para designar los términos de la secuencia que acabamos de definir se emplea la notación

$$y_n = \prod_{i=1}^{n} a_i\,,$$

indicando de tal forma que y_n es el producto de los primeros n términos de la secuencia (a_i) (el símbolo $\prod$, equivalente a la π mayúscula en el alfabeto griego, se denomina *símbolo de productoria*). Por ejemplo, el producto de los primeros n números naturales impares se representa en la forma

$$\prod_{i=1}^{n} (2i-1)\,.$$

Antes de exhibir un par de ejemplos subrayemos la definición inductiva del símbolo de productoria:

$$\begin{cases} \displaystyle\prod_{i=1}^{1} a_i = a_1 \\ \\ \displaystyle\prod_{i=1}^{n+1} a_i = \left(\prod_{i=1}^{n} a_i\right) a_{n+1} \quad \text{si } n \geq 1. \end{cases}$$

Ejemplo 3.3.6 Como ya fue mencionado en la sección anterior, se denomina *factorial* de un número natural n al número

$$n\,! = \prod_{i=1}^{n} i\,,$$

producto de los primeros n números naturales. Equivalentemente, la sucesión $(n\,!)_{n\geq 1}$ está definida inductivamente por las condiciones

$$\begin{cases} 1\,! = 1 \\ (n+1)! = (n+1)n\,! \quad \text{si } n \geq 1\,. \end{cases}$$

Sus primeros términos son

$$1\,! = 1\ ,\ 2\,! = 2\ ,\ 3\,! = 6\ ,\ 4\,! = 24\ ,\ 5\,! = 120\ ,\ 6\,! = 720\ ,\ 7\,! = 5040\,,$$

y como se puede apreciar, crecen muy rápidamente. Por ejemplo, $20\,!$ es un número natural de 19 dígitos. $\Diamond$

Ejemplo 3.3.7 Como muestra de la utilidad de contar con una definición inductiva de factorial, probemos por ejemplo que

$$\frac{(2n)!}{n!} = \prod_{i=1}^{n}(4i-2)$$

para todo $n \in \mathbb{N}$. Designando por $\mathrm{P}(n)$ el enunciado dado por la fórmula de arriba, probemos inductivamente que $\mathrm{P}(n)$ es verdadera para todo $n \in \mathbb{N}$. El caso $n = 1$ es trivial (como casi siempre), ya que

$$\frac{2!}{1!} = 2 = 4\,.\,1 - 2 = \prod_{i=1}^{1}(4i-2)\,.$$

Asumiendo que $\mathrm{P}(k)$ es verdadera para un cierto $k \in \mathbb{N}$, y aplicando repetidamente la definición inductiva de factorial, tenemos que

$$\frac{(2(k+1))!}{(k+1)!} = \frac{2(k+1)(2k+1)!}{(k+1)k!} = \frac{2(2k+1)!}{k!} = \frac{2(2k+1)(2k)!}{k!} =$$
$$= \frac{(2k)!}{k!}(4k+2) = \left(\prod_{i=1}^{k}(4i-2)\right)(4(k+1)-2) = \prod_{i=1}^{k+1}(4i-2)\,,$$

esto es, $\mathrm{P}(k+1)$ es verdadera. Luego, por el principio de inducción, $\mathrm{P}(n)$ es verdadera para todo $n \in \mathbb{N}$. $\Diamond$

Examinaremos ahora algunas propiedades elementales del símbolo de productoria, similares a las demostradas en la proposición 3.3.5 para el símbolo de sumatoria. Como en tal caso, (u_n) y (v_n) denotan secuencias de números reales, distintos de cero cuando corresponda.

Proposición 3.3.8 Las siguientes igualdades son válidas para todo $m \in \mathbb{N}$:

1) $\displaystyle\prod_{i=1}^{m}(u_i v_i) = \prod_{i=1}^{m} u_i \prod_{i=1}^{m} v_i$

2) $\displaystyle\prod_{i=1}^{m}(u_{i+1}/u_i) = u_{m+1}/u_1$

3) $\displaystyle\prod_{i=1}^{m} u_i = \prod_{i=1}^{m} u_{m-i+1}$

4) $\displaystyle\prod_{i=1}^{r+m} u_i = \left(\prod_{i=1}^{r} u_i\right)\left(\prod_{i=1}^{m} u_{r+i}\right) \quad (r \in \mathbb{N})$.

Demostracion. Basta adecuar al caso del producto las demostraciones de los diversos incisos de la proposición 3.3.5. $\Diamond$

Extensiones de los símbolos de sumatoria y productoria. Las definiciones de los símbolos de sumatoria y productoria dadas hasta aquí, en las cuales el índice de la suma o el producto varía sobre un intervalo inicial de los números naturales, pueden extenderse sin dificultad para darles sentido preciso a expresiones del tipo $a_r + a_{r+1} + \cdots + a_s$ y $a_r a_{r+1} \ldots a_s$, en las que los índices varían sobre cualquier intervalo de enteros no negativos. Precisamente, dado $m \geq 0$ las definimos inductivamente en la forma

$$\begin{cases} \displaystyle\sum_{i=m}^{m} a_i = a_m \\ \\ \displaystyle\sum_{i=m}^{n+1} a_i = \sum_{i=m}^{n} a_i + a_{n+1} \quad \text{si } n \geq m \end{cases}$$

y

$$\begin{cases} \displaystyle\prod_{i=m}^{m} a_i = a_m \\ \displaystyle\prod_{i=m}^{n+1} a_i = \left(\prod_{i=m}^{n} a_i\right) a_{n+1} \quad \text{si } n \geq m\,. \end{cases}$$

Alternativamente, usaremos también para designarlas los símbolos

$$\sum_{m\leq i\leq n} a_i \quad \text{y} \quad \prod_{m\leq i\leq n} a_i\,.$$

Debido a su utilidad, vale la pena detenernos en las siguientes fórmulas de "cambio de variable" en sumatorias y productorias, que conectan los símbolos originales con la generalización que acabamos de efectuar:

Proposición 3.3.9 Sea $(a_i)_{i\geq m}$ una sucesión de números reales y sea n en $\mathbb{N}_m$. Entonces

$$\sum_{i=m}^{n} a_i = \sum_{i=1}^{n-m+1} a_{m+i-1} \quad \text{y} \quad \prod_{i=m}^{n} a_i = \prod_{i=1}^{n-m+1} a_{m+i-1}\,.$$

Demostracion. Ambos enunciados se prueban sin dificultad por inducción en $r = n - m + 1$, por lo que dejamos los detalles a cargo del lector. Se deduce fácilmente de estos hechos que las generalizaciones de los símbolos de sumatoria y productoria gozan de propiedades enteramente análogas a las demostradas en las proposiciones 3.3.5 y 3.3.8. $\Diamond$

Nota. Con bastante frecuencia se presentan casos en los que debemos sumar o multiplicar ciertos términos a_i de una sucesión de números reales, y en los cuales el índice i no varía sobre todo un intervalo de $\mathbb{N}_0$ sino sobre algún subconjunto J de un intervalo, circustancias que indicaremos en las formas

$$\sum_{k\in J} a_k \quad \text{y} \quad \prod_{k\in J} a_k\,,$$

permitiéndonos a veces, si ello fuera pertinente, reemplazar el enunciado $k \in J$ por cualquier otro lógicamente equivalente a él.

Para formalizar estas expresiones, sea $J \subseteq \{m, m+1, \ldots, n\}$ y consideremos las secuencias finitas $(s_k)_{m\leq k\leq n}$ y $(p_k)_{m\leq k\leq n}$ dadas por las fórmulas

$$s_k = \begin{cases} a_k & \text{si } k \in J \\ 0 & \text{si } k \notin J \end{cases} \qquad \text{y} \qquad p_k = \begin{cases} a_k & \text{si } k \in J \\ 1 & \text{si } k \notin J\,. \end{cases}$$

Con evidente razonabilidad definimos entonces:

$$\sum_{k\in J} a_k = \sum_{k=m}^{n} s_k \quad \text{y} \quad \prod_{k\in J} a_k = \prod_{k=m}^{n} p_k\,.$$

Sea por ejemplo $J = \{2, 3, 6\}$. Tomando $m = 2$ y $n = 6$ tenemos entonces

$$\sum_{k\in J} k^2 = 2^2 + 3^2 + 0^2 + 0^2 + 6^2 = 49$$

y

$$\prod_{k\in J} k^2 = 2^2\,.\,3^2\,.\,1^2\,.\,1^2\,.\,6^2 = 1296\,.$$

Análogamente, resulta que

$$\sum_{\substack{k=1\\ k\text{ par}}}^{9} k^2 = 2^2 + 4^2 + 6^2 + 8^2 = 120$$

y

$$\prod_{\substack{k=1\\ k\text{ primo}}}^{10} k = 2\,.\,3\,.\,5\,.\,7 = 210\,.$$

Como caso particular, observemos que

$$\sum_{k\in\emptyset} a_k = 0 \quad \text{y} \quad \prod_{k\in\emptyset} a_k = 1\,,$$

lo que justifica las siguientes convenciones, adoptadas fundamentalmente para evitar en ocasiones la necesidad de tratar por separado algún caso especial: si m y n son enteros no negativos y $n < m$ entonces

$$\sum_{k=m}^{n} a_k = 0 \quad \text{y} \quad \prod_{k=m}^{n} a_k = 1\,. \quad \Diamond$$

Potencias de exponente natural.

Como adelantamos anteriormente, si a es un número real arbitrario y $n \in \mathbb{N}$, introducimos el símbolo a^n para designar el producto $a\,.\,a\ldots a$ (n veces a), análogo multiplicativo de la suma $n\,a = a + a + \cdots + a$ (n veces a). Vale decir,

$$a^n = \prod_{i=1}^{n} a\,.$$

Diremos entonces que a^n es la *potencia n-ésima* de a. Sigue inmediatamente de la definición inductiva de productoria que la sucesión $(a^n)_{n\geq 1}$

de potencias de a es la única sucesión de números reales definida por las siguientes reglas de recurrencia:

$$\begin{cases} a^1 = a \\ a^{n+1} = a^n a \quad \text{si } n \geq 1. \end{cases}$$

Extendemos la definición al caso de exponente cero, definiendo $a^0 = 1$. Observemos que la misma es absolutamente coherente, ya que

$$\prod_{i=1}^{0} a = 1$$

y se preserva además la recurrencia, pues

$$a^1 = a = 1\, a = a^0 a\,.$$

Ejemplo 3.3.10 Si $a, q \in \mathbb{R}$, la sucesión $(x_n)_{n\geq 0}$ de números reales definida por

$$x_n = aq^n$$

se llama *progresión geométrica* de razón q y término inicial a. Observemos que

$$x_{n+1} = aq^{n+1} = aq^n q = x_n q\,,$$

vale decir, comenzando por $x_0 = a$ cada término de una progresión geométrica se obtiene multiplicando el anterior por la razón q.

Ejemplos de progresiones geométricas son las sucesiones constantes (de razón 1) y las sucesiones $(c^n)_{n\geq 0}$ de potencias de un número real c, de término inicial 1 y razón c. Adicionalmente, encargamos al lector la tarea de verificar que la secuencia

$$3,\ -2,\ 4/3,\ -8/9,\ 16/27,\ \cdots\cdots$$

es una progresión geométrica. $\Diamond$

Para establecer ciertas propiedades básicas de las potencias es imprescindible el manejo de la definición inductiva, como apreciaremos a continuación:

Proposición 3.3.11 Sean $a, b \in \mathbb{R}$ y sean $m, n \in \mathbb{N}$. Entonces:

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

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

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

4) $(a/b)^n = a^n/b^n$ $(b \neq 0)$

5) $a^n - b^n = (a-b)\sum_{i=1}^{n-1} a^{n-1-i}b^i$

6) $a^n < b^n \Leftrightarrow a < b$ (a y b positivos).

DEMOSTRACION. Probaremos inductivamente 1) y 5), encargando al lector la demostración de los restantes enunciados. También podrá verificar el amigo lector que los primeros cuatro enunciados resultan trivialmente ciertos si $m = 0$ ó $n = 0$.

Demostraremos 1) por inducción en n. Siendo el caso $n = 1$ una traducción literal de la definición inductiva de potencia, supongamos que el enunciado es verdadero para un cierto natural k, es decir, $a^{m+k} = a^m a^k$ para todo $m \in \mathbb{N}$. Entonces

$$a^{m+(k+1)} = a^{(m+k)+1} = a^{m+k}a = (a^m a^k)a = a^m(a^k a) = a^m a^{k+1},$$

y por lo tanto también es verdadero para $k+1$.

Para demostrar 5), también por inducción en n, designemos por $\mathrm{P}(n)$ su enunciado. Tenemos en primer término que

$$(a-b)\sum_{i=0}^{1-1} a^{1-1-i}b^i = (a-b)a^0b^0 = a - b = a^1 - b^1,$$

y por lo tanto $\mathrm{P}(1)$ es verdadera.

Respecto al paso inductivo, suponiendo que $\mathrm{P}(k)$ es verdadera resulta que

$$\begin{aligned} a^{k+1} - b^{k+1} &= a(a^k - b^k) + (a-b)b^k = \\ &= a(a-b)\sum_{i=0}^{k-1} a^{k-1-i}b^i + (a-b)b^k = \\ &= (a-b)\left(\sum_{i=0}^{k-1} a^{k-i}b^i + a^0b^k\right) = \\ &= (a-b)\sum_{i=0}^{k} a^{k-i}b^i, \end{aligned}$$

lo que prueba que $\mathrm{P}(k+1)$ es verdadera, y por lo tanto $\mathrm{P}(n)$ es verdadera para todo $n \in \mathbb{N}$. $\Diamond$

Ejemplos 3.3.12 A través de la demostración de algunos hechos sencillos ilustraremos la definición inductiva de potencia:

1) $7^n + 2$ es divisible por 3 para todo $n \in \mathbb{N}$.

La afirmación es cierta para $n = 1$, pues $7^1 + 2 = 9 = 3 \,.\, 3$.

Supongamos que el enunciado es verdadero para $k \in \mathbb{N}$ y sea $a \in \mathbb{N}$ tal que $7^k + 2 = 3a$. Entonces

$$7^{k+1}+2 = 7^k 7+2 = (3a-2)7+2 = 21a-14+2 = 21a-12 = 3(7a-4).$$

esto es, $7^{k+1} + 2$ es divisible por 3. Luego la afirmación es verdadera para $k + 1$, y en consecuencia es verdadera para todo $n \in \mathbb{N}$, por el principio de inducción.

2) Sea $x \in [-1, +\infty)$. Entonces $(1 + x)^n \geq 1 + nx$ para todo $n \in \mathbb{N}$.

El enunciado es verdadero para $n = 1$, ya que en tal caso ambos miembros de la desigualdad planteada son iguales. Suponiendo que la propiedad es válida para cierto natural k, resulta que

$$\begin{aligned}(1+x)^{k+1} &= (1+x)^k(1+x) \geq (1+kx)(1+x) = \\ &= 1 + (k+1)x + kx^2 \geq 1 + (k+1)x,\end{aligned}$$

y por lo tanto también es verdadera para $k + 1$, de donde concluimos que es verdadera para todo $n \in \mathbb{N}$. Notemos que la primera desigualdad de la secuencia de arriba se debe a la hipótesis inductiva y al hecho de que $1 + x \geq 0$.

3) $\left(1 + \dfrac{1}{m}\right)^m \geq 2$ para todo $m \in \mathbb{N}$.

Basta aplicar 2), tomando $x = \dfrac{1}{m}$ y $n = m$. $\quad \Diamond$

RAICES ENESIMAS. El hecho de que todo número real no negativo admita raíz cuadrada en $\mathbb{R}$, desarrollado en el capítulo 2, se extiende al hecho más general de existencia de raíz $n-$ésima para todo número natural n. Precisamente, vale el siguiente resultado:

Proposición 3.3.13 Sea $a \in \mathbb{R}\,(a \geq 0)$ y sea $n \in \mathbb{N}$. Existe entonces un único número real $b \geq 0$ tal que $b^n = a$.

DEMOSTRACION. No ofreceremos aquí una prueba detallada del enunciado, ya que la misma se obtiene procediendo en forma similar a la del caso $n = 2$ (proposición 2.3.6). Remarquemos de todas maneras que dicha prueba es técnicamente más complicada que la del caso citado, y utiliza, además del axioma de completitud de los números reales, la fórmula de Newton para la expansión de la potencia $n-$ésima de un binomio que demostraremos en el próximo capítulo.

Empleando terminología y notación familiares al lector, diremos que b es la *raíz $n-$ésima* de a y la designaremos por $\sqrt[n]{a}$. Esto es,

$$b = \sqrt[n]{a} \text{ si y sólo si } b \geq 0 \text{ y } b^n = a.$$

Por ejemplo, $\sqrt[3]{8} = 2$, $\sqrt[4]{81} = 3$ y $\sqrt[n]{1} = 1$ para todo $n \in \mathbb{N}$. $\quad \Diamond$

3.3.2. Generalizaciones del Principio de Induccion

A partir de la generalización del teorema de inducción establecida en 3.1.8 es posible extender la validez del principio de inducción a proposiciones predicables en $\mathbb{N}_m$, cualquiera sea el entero no negativo m. El enunciado preciso de dicha extensión es el siguiente:

Teorema 3.3.14 (Principio de inducción generalizado) Sea m un entero no negativo y sea P una proposición predicable en $\mathbb{N}_m$ que verifica las dos siguientes condiciones:

1) P(m) es verdadera.

2) La implicación P(k) $\Rightarrow$ P($k+1$) es verdadera cualquiera sea $k \in \mathbb{N}_m$.

Entonces P(n) es verdadera para todo $n \in \mathbb{N}_m$.

DEMOSTRACION. Basta observar que el conjunto

$$\{n \in \mathbb{N}_m : \mathrm{P}(n) \text{ es verdadera}\}$$

es $m-$inductivo. $\Diamond$

Ejemplos 3.3.15 Veamos dos ejemplos de aplicación.

1) Sea $\mathcal{D}_n$ el conjunto de subconjuntos de 2 elementos de $\mathbb{I}_n$ cuyos elementos no son consecutivos (consideraremos consecutivos a los elementos de $\{1, n\}$). Por ejemplo,

$$\mathcal{D}_5 = \{\{1,3\},\{1,4\},\{2,4\},\{2,5\},\{3,5\}\}\ .$$

Si x_n es el número de elementos de $\mathcal{D}_n$, probaremos por inducción generalizada que

$$x_n = \frac{n(n-3)}{2}$$

para todo número natural $n \geq 3$.

Designando por P(n) el enunciado que queremos probar, sigue inmediatamente que P(3) es verdadera, ya que $\mathcal{D}_3$ es vacío y

$$\frac{3(3-3)}{2} = 0\,.$$

Para efectuar el paso inductivo, supongamos que P(k) es verdadera para un cierto $k \geq 3$. Es claro que $\mathcal{D}_k \subseteq D_{k+1}$, faltando considerar para completar todos los elementos de $\mathcal{D}_{k+1}$ el subconjunto $\{1, k\}$ y los $k-2$ subconjuntos $\{2, k+1\}, \{3, k+1\}, \ldots, \{k-1, k+1\}$. Operando convenientemente y aplicando la hipótesis inductiva sigue entonces que

$$x_{k+1} = x_k + 1 + k - 2 = \frac{k(k-3)}{2} + k - 1 = \frac{k^2 - k - 2}{2} =$$
$$= \frac{(k+1)(k-2)}{2} = \frac{(k+1)(k+1-3)}{2}\,,$$

lo que prueba que P($k+1$) es verdadera. Luego P(n) es verdadera para todo $n \geq 3$, por el principio de induccion generalizado.

2) $2^n > n^2$ para todo número natural $n \geq 5$.

Supongamos que el enunciado es válido para un cierto $k \geq 5$ (compruebe el caso $n = 5$). Resulta entonces que

$$2^{k+1} = 2 \,.\, 2^k > 2k^2 = (k+1)^2 + k^2 - 2k - 1 > (k+1)^2,$$

ya que $k^2 - 2k - 1 = (k-1)2 - 2 \geq 14 > 0$, por ser $k \geq 5$. Esto completa el paso inductivo y el resultado es válido para todo $n \in \mathbb{N}_5$ (sería interesante que el lector observe qué ocurre si $n < 5$). $\Diamond$

Suma de los terminos de una progresion Probaremos a continuación dos fórmulas útiles referidas a la suma de los primeros términos de una progresión aritmética o geométrica.

Proposición 3.3.16 Sea $(x_n)_{n\geq 0}$ una progresión aritmética de término inicial a y razón r y sea $(y_n)_{n\geq 0}$ una progresión geométrica de término inicial b y razón $q\,(q \neq 1)$. Si m es un entero no negativo, valen las fórmulas

$$\sum_{i=0}^{m} x_i = \frac{(m+1)(2a+rm)}{2} \qquad \text{y} \tag{3.1}$$

$$\sum_{i=0}^{m} y_i = b\left(\frac{q^{m+1}-1}{q-1}\right). \tag{3.2}$$

Demostracion. Empleando la proposición 3.3.5 y la fórmula del ejemplo 3.3.3, resulta que

$$\begin{aligned}\sum_{i=0}^{m} x_i = \sum_{i=0}^{m}(a+ri) = \sum_{i=0}^{m} a + \sum_{i=0}^{m} ri = a + \sum_{i=1}^{m} a + \sum_{i=1}^{m} ri = \\ = a + ma + r\sum_{i=1}^{m} i = (m+1)a + \frac{rm(m+1)}{2} = \\ = \frac{2(m+1)a + rm(m+1)}{2} = \frac{(m+1)(2a+rm)}{2}\,,\end{aligned}$$

lo que prueba 1).

Como caso particular interesante podemos deducir una fórmula para la suma $\mathcal{I}_n$ de los primeros n números naturales impares, siendo la secuencia de los mismos una progresión aritmética de razón 2 y término inicial 1. Tomando $m = n-1$ en 1), tenemos:

$$\mathcal{I}_n \;=\; \sum_{i=0}^{n-1}(2i+1) \;=\; \frac{n(2+2(n-1))}{2} \;=\; n^2\,.$$

Para probar la segunda parte (obsérvese que la progresión es constante si $q = 1$), designemos por S la suma del enunciado y supongamos $m > 0$ (el caso $m = 0$ es trivial). Operando convenientemente, tenemos:

$$S = \sum_{i=0}^{m} y_i = \sum_{i=0}^{m} b\,q^i = b + \sum_{i=1}^{m} b\,q^i = b + q\sum_{i=1}^{m} b\,q^{i-1} = b + q\sum_{i=0}^{m-1} b\,q^i = \\ = b + q\,(S - bq^m) = b + q\,S - b\,q^{m+1}\,,$$

de donde

$$S(q-1) = b\,q^{m+1} - b = b\,(q^{m+1} - 1)\,,$$

esto es,

$$S \;=\; b\left(\frac{q^{m+1}-1}{q-1}\right),$$

según queríamos probar.

Como ejemplo de interés deducimos una fórmula para calcular la suma de los primeros términos de la sucesión de potencias de un número real $c \neq 1$, que es una progresión geométrica de razón c y término inicial 1. Aplicando 2) obtenemos en este caso:

$$\sum_{i=0}^{m} c^i \;=\; \frac{c^{m+1}-1}{c-1}\;. \quad \Diamond$$

El principio de inducción global.

En ciertas ocasiones, la hipótesis inductiva del principio de inducción (P(k) es verdadera) es insuficiente por sí sola para probar que P($k+1$) es verdadera, y puede ocurrir que para efectuar la prueba sea necesario asumir no sólo que P(1) y P(k) son enunciados verdaderos, sino también algunos de los enunciados (tal vez todos)

$$\mathrm{P}(2),\ \ldots\ ,\ \mathrm{P}(k-1).$$

Por ejemplo, si quisiéramos demostrar por inducción que todo número natural n puede expresarse en la forma

$$n = a^2 b,$$

donde a y b son números naturales y b no es divisible por 4, de poco nos serviría la suposición de que un cierto k admite una factorización como la del enunciado para probar que también $k+1$ se descompone de tal forma, ya que como veremos más adelante, las posibles factorizaciones de un número natural no se relacionan con las factorizaciones de su predecesor.

A fin de manejar este tipo de situaciones introduciremos el siguiente procedimiento inductivo llamado *principio de inducción global* (PIG) o *principio de inducción completa*:

Proposición 3.3.17 (Principio de induccion global) Sea P una proposición predicable en $\mathbb{N}$ que verifica las dos siguientes condiciones:

1) P(1) es verdadera.

2) La implicación

$$(\mathrm{P}(1) \wedge \mathrm{P}(2) \wedge \cdots \wedge \mathrm{P}(k)) \Rightarrow \mathrm{P}(k+1)$$

es verdadera cualquiera sea $k \in \mathbb{N}$.

Entonces P(n) es verdadera para todo $n \in \mathbb{N}$.

Demostracion. Remarquemos la diferencia con el principio de inducción ordinario: en el paso inductivo, para demostrar que P($k+1$) es verdadera no sólo se asume que P(k) es verdadera, sino que P(j) es verdadera para todo $j \leq k$.

En cuanto a la prueba, procedemos por el absurdo. Suponiendo que la conclusión del PIG es falsa, podemos considerar por el principio de buena ordenación el mínimo número natural r tal que P(r) es falsa. Observando que $r > 1$, pues P(1) es verdadera por hipótesis, sigue por la minimalidad de r que P(j) es verdadera para todo número natural $j \leq r-1$, de donde deducimos (aplicando la parte 2) de la hipótesis al caso $k = r-1$) que P(r) es verdadera, lo que es una contradicción. $\Diamond$

Nota Si bien la hipótesis inductiva en el principio de inducción global es aparentemente más fuerte que la hipótesis inductiva en el principio de inducción, ambos resultados son lógicamente equivalentes, esto es, se implican mutuamente. Puesto que en la proposición 3.3.17 hemos probado la validez del PIG a partir de la validez del principio de inducción, ya que empleamos en la demostración el principio de buena ordenación, equivalente a él, demostremos ahora la recíproca.

Supongamos para ello que la proposición P satisface las dos condiciones del principio de inducción y consideremos la proposición Q, predicable en $\mathbb{N}$, definida para todo n en la forma:

$$\mathrm{Q}(n) : \mathrm{P}(j) \text{ es verdadera para todo } j \leq n\,.$$

Es claro que Q(1) es verdadera, ya que P(1) lo es. Supongamos ahora que Q(j) es verdadera para todo natural j menor o igual que un cierto k, lo que implica por definición de Q que P(j) es verdadera para todo $j \leq k$. Resulta en particular que P(k) es verdadera, y por lo tanto también P($k+1$) (recordemos que P satisface la condición 2) del principio de inducción).

En consecuencia P(j) es verdadera para todo $j \leq k+1$ y Q($k+1$) es verdadera. Puesto que hemos probado que Q satisface las condiciones del PIG concluimos que Q(k) es verdadera para todo $n \in \mathbb{N}$, lo que obviamente significa que también P(n) es verdadera para todo n. $\Diamond$

Ejemplo 3.3.18 Probemos por inducción completa el enunciado con el que motivamos el tema, esto es, que todo número natural n se puede escribir como producto de un cuadrado perfecto y un número natural no divisible por 4.

Designando por $\mathrm{P}(n)$ el enunciado anterior, es trivial verificar que $\mathrm{P}(1)$ es verdadera, pues $1 = 1^2 \,.\, 1$. Asumamos ahora que $\mathrm{P}(j)$ es verdadera para todo j menor o igual que un cierto k en $\mathbb{N}$ y supongamos en primer término que $k+1$ no es divisible por 4, en cuyo caso basta escribir $k+1 = 1^2 \,.\, (k+1)$. Suponiendo por el contrario que $k+1$ es múltiplo de 4, sea $m \in \mathbb{N}$ tal que $k+1 = 4m$.

Puesto que $m < k$, ya que

$$k - m = k - \frac{k+1}{4} = \frac{3k-1}{4} \geq \frac{1}{2} > 0\,,$$

resulta por hipótesis inductiva que $\mathrm{P}(m)$ es verdadera, digamos $m = c^2 h$ con h no divisible por 4. Reemplazando, obtenemos $k+1 = 4c^2 h = (2c)^2 h$, y por lo tanto $k+1$ es de la forma indicada.

Habiendo probado que $\mathrm{P}(k+1)$ es verdadera cualquiera sea el caso, sigue por el principio de inducción global que $\mathrm{P}(n)$ es verdadera para todo $n \in \mathbb{N}$, como queríamos demostrar. $\Diamond$

Como es fácil imaginar a esta altura de nuestro desarrollo, el principio de inducción global es válido en general para proposiciones predicables en cualquier sección final del conjunto de números enteros no negativos. No brindaremos aquí una prueba detallada de tal hecho, ya que basta adecuar la demostración de la proposición 3.3.17 para obtener una, habida cuenta de la validez del principio de inducción ordinario en dichas secciones. Nos limitaremos a enunciarlo con cuidado y directamente lo aplicaremos cuando nos parezca necesario. Precisamente, el mismo afirma:

Sea $m \geq 0$ y sea P una proposición predicable en $\mathbb{N}_m$ que verifica las dos condiciones siguientes:

1) $\mathrm{P}(m)$ es verdadera.

2) La implicación

$$(\mathrm{P}(m) \wedge \mathrm{P}(m+1) \wedge \cdots \wedge \mathrm{P}(k)) \Rightarrow \mathrm{P}(k+1)$$

es verdadera cualquiera sea $k \in \mathbb{N}_m$.

Entonces $\mathrm{P}(n)$ es verdadera para todo $n \in \mathbb{N}_m$.

Ejemplo 3.3.19 Sea $(x_n)_{n\geq 0}$ la sucesión de números reales definida por las reglas de recurrencia

$$\begin{cases} x_0 = x_1 = -1 \\ \\ x_n = 5x_{n-1} - 6x_{n-2} \quad \text{para todo } n \geq 2\,, \end{cases}$$

cuyos primeros términos son $-1, -1, 1, 11, 49, 179\ldots$

Probaremos por inducción completa que $x_n = 3^n - 2^{n+1}$ para todo n en $\mathbb{N}_0$.

La afirmación es cierta si $n = 0$ ó $n = 1$, ya que

$$3^0 - 2^1 = 3^1 - 2^2 = -1.$$

Para el paso inductivo, supongamos que la conjetura es válida para todo entero no negativo j menor o igual que un cierto $k \geq 1$, y probemos su validez para $k+1$. Empleando la regla de recurrencia y aplicando la hipótesis inductiva, tenemos:

$$\begin{aligned} x_{k+1} &= 5x_k - 6x_{k-1} = 5(3^k - 2^{k+1}) - 6(3^{k-1} - 2^k) = \\ &= (5\,.\,3^k - 6\,.\,3^{k-1}) + (6\,.\,2^k - 5\,.\,2^{k+1}) = \\ &= 3^{k-1}(15-6) - 2^k(10-6) = 3^{k-1}3^2 - 2^k 2^2 = 3^{k+1} - 2^{k+2}, \end{aligned}$$

lo que prueba que el enunciado es verdadero para $k+1$. Luego es verdadero para todo $n \in \mathbb{N}_0$, por el principio de inducción global. $\Diamond$

3.3.3. Ejercicios

1. En cada uno de los siguientes incisos se define una función proposicional P predicable en $\mathbb{N}$. Determinar en cada caso para qué valores de k es verdadera la implicación $\mathrm{P}(k) \Rightarrow \mathrm{P}(k+1)$:

   a) $\mathrm{P}(n) : n$ es impar

   b) $\mathrm{P}(n) : n > 2$

   c) $\mathrm{P}(n) : n^2 > 7n - 10$

   d) $\mathrm{P}(n) : n > n + 1$

   e) $\mathrm{P}(n) : n^2 + 1$ no es divisible por 3 .

2. Expresar las siguientes sumas usando el símbolo de sumatoria:

   a) $2 + 4 + 6 + 8 + \cdots + 52$

   b) $2 - 4 + 8 - 16 + 32 - 64$

   c) $1/2 - 1/6 + 1/12 - 1/20 + 1/30$

   d) $1 + 1/3 + 9 + 1/27 + 81$

   e) $-1 + 4 + 48 + 320 + 1792$.

3. Demostrar la validez, para todo $n \in \mathbb{N}$, de las siguientes fórmulas:

   a) $\displaystyle\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$

   b) $\displaystyle\sum_{k=1}^{n} k^3 = \left(\sum_{k=1}^{n} k\right)^2$

   c) $\displaystyle\sum_{k=1}^{n} (-1)^k k^2 = \frac{(-1)^n n(n+1)}{2}$

   d) $\displaystyle\sum_{k=1}^{n} \frac{1}{k(k+1)} = \frac{n}{n+1}$

   e) $\displaystyle\sum_{i=1}^{n} \frac{1}{4k^2-1} = \frac{n}{2n+1}$

   f) $\displaystyle\prod_{k=1}^{n} \frac{1}{2k-1)} = \frac{2^n n!}{(2n)!}$.

4. Calcular:

   a) $\displaystyle\sum_{k=0}^{n} 6k - 3$

   b) $\displaystyle\prod_{k=1}^{n} 3k$

   c) $\displaystyle\prod_{k=0}^{n} 5^k$

   d) $\displaystyle\sum_{k=1}^{n} \frac{(k-1)^2}{2}$

   e) $\displaystyle\sum_{k=m}^{n} q^k \quad (q \neq 1\ ,\ 0 \leq m \leq n)\,.$

5. *a)* Sean $(a_i)_{i\geq 1}$ y $(b_i)_{i\geq 1}$ sucesiones en $\mathbb{R}$ tales que $b_i = a_{i+1} - a_i$ para todo $i \geq 1$. Probar que $\displaystyle\sum_{i=1}^{n} b_i = a_{n+1} - a_1$.

   b) Hallar una fórmula cerrada para la suma $\displaystyle\sum_{i=1}^{n} \frac{1}{(3i-1)(3i+1)}$.

6. *a)* Sea $(a_k)_{k\geq 1}$ una sucesión en $\mathbb{R}$. Probar que

$$\sum_{k=1}^{n} k a_k = \sum_{i=1}^{n} \left(\sum_{k=i}^{n} a_k\right).$$

b) Hallar fórmulas cerradas para las sumas $\sum_{k=1}^{n} k2^k$ y $\sum_{k=1}^{n} k3^k$.

7. Determinar, justificando el resultado, fórmulas cerradas para el cálculo de las siguientes sumas ($n \in \mathbb{N}_0$):

   a) $\sum_{k=0}^{n}(-1)^k$

   b) $\sum_{k=0}^{n}(-1)^k k$

   c) $\sum_{k=0}^{n} kk!$

   d) $\sum_{k=0}^{n} \frac{5^{k-1}}{2^{k+1}}$

   e) $\sum_{k=0}^{n} \frac{(-1)^{k+1}}{4^k}$

   f) $\sum_{k=0}^{n} k^4$.

8. Si $(A_i)_{i\geq 1}$ es una sucesión de conjuntos, definir inductivamente las operaciones $A_1 \cap A_2 \cap \cdots \cap A_n$, $A_1 \cup A_2 \cup \cdots \cup A_n$ y $A_1 \times A_2 \times \cdots \times A_n$.

9. Probar las siguientes propiedades ($a, b \in \mathbb{R}\,, n \in \mathbb{N}$):

   a) $a \leq 1 \Leftrightarrow a^n \geq a^{n+1} \quad (a > 0)$.

   b) $a^n - b^n = (a+b)\sum_{i=0}^{n-1}(-1)^i a^{n-i-1}b^i \quad (n \text{ par})$.

   c) $a^n + b^n = (a+b)\sum_{i=0}^{n-1}(-1)^i a^{n-i-1}b^i \quad (n \text{ impar})$.

10. Sea $n \geq 3$ y sea P un polígono convexo de n lados (P se dice convexo si el segmento que une dos cualesquiera de sus puntos está contenido en P).

    a) Probar que P tiene $\frac{n(n-3)}{2}$ diagonales.

    b) Probar que la suma de las medidas de los ángulos interiores de P es $(n-2)\pi$.

11. Demostrar las siguientes propiedades de los números de Fibonacci (m y n designan números naturales):

    a) $\displaystyle\sum_{k=1}^{n} F_k \;=\; F_{n+2} - 1$

    b) $F_{n-1}F_{n+1} = F_n^2 + (-1)^n$

    c) $F_{m+n} = F_{m-1}F_{n+1} + F_m F_{n+1}$

    d) $F_{2n} = F_n(F_{n-1} + F_{n+1})$.

12. En cada uno de los siguientes incisos se define por recurrencia una sucesión de números reales. Demostrar en cada caso la fórmula indicada para su término general:

    a) $a_1 = 0$, $a_{i+1} = a_i + 2i$; $a_n = n^2 - n$

    b) $a_1 = \dfrac{1}{2}$, $a_{i+1} = \left(1 + \dfrac{2}{i}\right) a_i^{-1}$; $a_n = \dfrac{n}{n+1}$

    c) $a_3 = 1$, $a_{i+1} = 3a_i + 2^{i-1} - 1$; $a_n = \dfrac{3^{n-1} - 2^n + 1}{2}$

    d) $a_0 = 3$, $a_{i+1} = a_i^2 - 2a_i + 2$; $a_n = 2^{2^n} + 1$

    e) $a_0 = 1$, $a_1 = 1$, $a_{i+2} = 3a_{i+1} + (i^2 - 1)a_i$; $a_n = n!$.

13. En cada uno de los siguientes incisos se define por recurrencia una sucesión de números reales. En cada caso conjeturar una fórmula para el término general a_n y probar su validez:

    a) $a_1 = 1$, $a_{i+1} = 3a_i + 4$

    b) $a_1 = 2$, $a_{i+1} = (i+1)a_i$

    c) $a_0 = 0$, $a_1 = 1$, $a_{i+2} = 2a_{i+1} - a_i$

    d) $a_0 = 0$, $a_1 = 1$, $a_{i+2} = a_{i+1} + 2\sqrt{a_i} + 3$

    e) $a_0 = 2$, $a_1 = 1$, $a_{i+2} = a_{i+1} + 6a_i$

    f) $a_0 = -1$, $a_1 = 0$, $a_{i+2} = 5a_{i+1} - 6a_i$.

14. Demostrar las siguientes desigualdades ($n \in \mathbb{N}$):

    a) $2n + 1 < 2^n$ para todo $n \geq 3$

    b) $2^n + 5^n \geq 2 \cdot 3^{n+1}$ para todo $n \geq 4$

    c) $\displaystyle\sum_{k=1}^{n} \frac{1}{k!} \;<\; \frac{2n+1}{n+1}$ para todo $n \in \mathbb{N}$.

15. Sea $(a_n)_{n\geq 0}$ la sucesión de números reales definida recursivamente en la forma

$$a_0 = 0\ ,\ a_1 = 5/2\ ,\ 2a_i = 7a_{i-1} - 3a_{i-2} \quad (i \geq 2)\,.$$

Probar que $3^n - \dfrac{1}{2} \leq a_n < 3^n$ para todo $n > 0\,$.

16. Sean $(a_i)_{i\geq 0}$ y $(b_i)_{i\geq 0}$ las sucesiones de números reales respectivamente definidas por las reglas de recurrencia

$$\begin{cases} a_0 = 1\ ,\ a_1 = 1\ ,\ a_{i+2} = 2a_{i+1} + a_i \text{ si } i \geq 0 \\ b_0 = 1\ ,\ b_1 = 1\ ,\ b_{i+2} = 2b_{i+1} + b_i \text{ si } i \geq 0\,. \end{cases}$$

a) Probar que $a_{n+1} = a_n + 2b_n$ y $b_{n+1} = a_n + b_n$ para todo $n \in \mathbb{N}_0\,$.

b) Probar que $\left(1+\sqrt{2}\right)^n = a_n + b_n\sqrt{2}$ para todo $n \in \mathbb{N}_0\,$.

3.4. Numeros enteros y racionales

3.4.1. Números enteros

Comencemos precisando la noción de número entero:

> Un número real x se dice un *número entero* si y solo si $x \in \mathbb{N}_0$ ó x es el inverso aditivo de un número natural. En forma equivalente y más concisa, x es entero si y solo si $|x| \in \mathbb{N}_0$.

Designaremos el conjunto de números enteros por la letra $\mathbb{Z}$. Conjuntísticamente, la definición anterior se expresa en la forma

$$\mathbb{Z} = \mathbb{N} \cup \{0\} \cup \mathbb{N}^- ,$$

donde $\mathbb{N}^-$ denota el conjunto de inversos aditivos de los números naturales. Más familiarmente, suele escribirse también

$$\mathbb{Z} = \{\ldots , -3 , -2 , -1 , 0 , 1 , 2 , 3 , \ldots\} .$$

Es claro por definición que $\mathbb{N} \subset \mathbb{Z}$.

Estructura de los numeros enteros. La estructura algebraica del conjunto de números enteros es más rica que la del conjunto de números naturales. Como lo hicimos en el caso de éstos, consideremos en primer lugar las restricciones a $\mathbb{Z}$ de las operaciones de adición y multiplicación definidas para los reales:

Proposición 3.4.1 Si $m, n \in \mathbb{Z}$ entonces $m + n \in \mathbb{Z}$ y $mn \in \mathbb{Z}$.

Demostracion. Sean $r, s \in \mathbb{Z}$ y consideremos en primer término el caso de la suma. Si alguno de los dos es nulo es resultado es obvio, y puesto que la suma de números naturales es un número natural, también es sencillo tratar el caso en que ambos tienen igual signo. En efecto, si ambos son positivos se trata de la suma de dos números naturales, mientras que si ambos son negativos, tenemos que $r + s = -((-r) + (-s))$ es el opuesto de un número natural, y por lo tanto entero.

Podemos suponer entonces sin pérdida de generalidad que $r > 0$ y $s < 0$, en cuyo caso el resultado sigue de la proposición 3.1.6 ya que

$$r + s = r - (-s) \in \mathbb{N}_0$$

si $r \geq -s$, mientras que

$$-(r + s) = (-s) - r \in \mathbb{N}$$

si $r < -s$.

La demostración para el producto es todavía más sencilla. Es claro que podemos suponer que ambos números son distintos de 0 (si no el producto

es 0), y teniendo en cuenta las igualdades $rs = (-r)(-s) = -r(-s)$ resulta que $rs \in \mathbb{N}$ si r y s tienen el mismo signo y $rs \in -N$ en otro caso. Hemos usado por supuesto que el producto de dos números naturales es un número natural. $\Diamond$

Como se recordará, hemos puesto especial énfasis en la estructura de cuerpo de los números reales, dada por sus operaciones de adición y multiplicación. El teorema precedente muestra que $\mathbb{Z}$ es una parte cerrada para dichas operaciones, lo cual determina una estructura algebraica en dicho conjunto. En el siguiente teorema enumeramos sus propiedades.

Proposición 3.4.2 Las operaciones de suma y producto en $\mathbb{Z}$ satisfacen las siguientes propiedades (las letras designan números enteros):

S_1) Asociatividad de la suma: $x + (y + z) = (x + y) + z$.

S_2) Conmutatividad de la suma: $x + y = y + x$.

S_3) Existencia de elemento neutro para la suma: $0 \in \mathbb{Z}$.

S_4) Existencia de inverso aditivo: si $x \in \mathbb{Z}$ entonces $-x \in \mathbb{Z}$.

P_1) Asociatividad del producto: $x(yz) = (xy)z$.

P_2) Conmutatividad del producto: $xy = yx$.

P_3) Existencia de elemento neutro para el producto: $1 \in \mathbb{Z}$.

D) Distributividad del producto respecto a la suma: $x(y + z) = xy + xz$.

Demostracion. Las propiedades estructurales de la suma y el producto son consecuencia de la validez de las mismas en el cuerpo de los números reales. Las propiedades S_3, S_4 y P_3 siguen inmediatamente de la definición de número entero. $\Diamond$

Nota. En general, un conjunto A con dos operaciones (denominadas genéricamente suma y producto) que satisfacen las ocho propiedades anteriores se llama un *anillo*, o más propiamente un anillo conmutativo con identidad. Si se verifica la propiedad adicional de que todo elemento no nulo admite un inverso multiplicativo (dado $x \neq 0$ existe y tal que $xy = 1$), A se dice un cuerpo, que es el caso de los números reales. Ciertamente $\mathbb{Z}$ no es un cuerpo, ya que 1 y -1 son los únicos enteros que admiten inverso multiplicativo entero. En efecto, supongamos que a y b son elementos de $\mathbb{Z}$ tales que $ab = 1$. Tomando valores absolutos resulta entonces que

$$1 = |\, ab \,| = |\, a \,|\, |\, b \,| \geq |\, a \,|\,,$$

pues $|\, b \,| \geq 1$ (por ser un número natural). En consecuencia $|\, a \,| = 1$ y por lo tanto $a = 1$ ó $a = -1$. Naturalmente, todo entero de módulo mayor que 1 admite inverso multiplicativo en $\mathbb{R}$, pero éste no es entero. $\Diamond$

Potencias de exponente entero.

Hasta aquí, hemos definido inductivamente las potencias de exponente entero no negativo de un número real a cualquiera y hemos establecido sus propiedades básicas. Nos proponemos ahora extender la definición al caso de exponente entero negativo. Precisamente, dado $a \in \mathbb{R}\,(a \neq 0)$ y $m \in \mathbb{N}$, definimos:

$$a^{-\,m} = (a^m)^{-1}\,.$$

Por ejemplo, $(2/3)^{-2} = 9/4$ y $(-1/5)^{-3} = -125$. Esta definición de potencia de exponente negativo no es caprichosa, y está guiada por el propósito de conservar la validez las propiedades de la potencia enunciadas en la proposición 3.3.11, ya que en tal caso debería cumplirse:

$$a^{\,-m} = a^{m\,.\,(-1)} = (a^m)^{-1} = (a^{-1})^m.$$

Enunciaremos a continuación las extensiones de las diversas propiedades de la potencia al caso de exponentes enteros, dejando las demostraciones a cargo del lector, ya que las mismas comportan un sencillo examen de las diversas situaciones que se presentan (en casos de exponente negativo suponemos que las bases son no nulas).

Proposición 3.4.3 Sean $m, n \in \mathbb{Z}$ y $a, b \in \mathbb{R}$. Entonces

1. $a^m a^n = a^{m+n}$
2. $(a^m)^n = a^{mn}$
3. $(ab)^m = a^m b^m\,.$ $\quad \Diamond$

Arquimedianidad.

Con el objetivo de probar que el conjunto de números enteros no está acotado superiormente en $\mathbb{R}$, comenzaremos demostrando el siguiente resultado:

Proposición 3.4.4 Todo subconjunto no vacío de $\mathbb{Z}$ acotado superiormente en $\mathbb{R}$ admite máximo.

Demostracion. Si A es un tal conjunto, sea $s = \sup A$ (axioma de completitud). Si $s \notin A$, existe $m \in A$ tal que $s - 1 < m < s$, o equivalentemente, $s < m + 1 < s + 1$. Puesto que por 3.1.3 no existen enteros en el intervalo $(m, m+1)$, resulta que m es una cota superior de A menor que s, lo que es una contradicción. Luego $s \in A$ y por lo tanto s es el máximo de A. $\quad \Diamond$

Corolario 3.4.5 $\mathbb{N}$ no está acotado superiormente (luego tampoco $\mathbb{Z}$).

DEMOSTRACION Es claro que $\mathbb{N}$ no tiene un elemento máximo, ya que todo número natural a es menor que su sucesor $a+1$, que también es natural. Sigue entonces de 3.4.4 que $\mathbb{N}$ no está acotado superiormente. $\Diamond$

El hecho que acabamos de demostrar admite la siguiente formulación equivalente:

Teorema 3.4.6 (Propiedad Arquimedeana) Sea a un número real positivo y sea b un número real cualquiera. Existe entonces $n \in \mathbb{N}$ tal que $n\,a > b$.

DEMOSTRACION. Siendo a positivo, la desigualdad que postula el enunciado es equivalente a la desigualdad $b/a < n$, que necesariamente debe verificarse para algún n, ya que en caso contrario b/a resultaría ser una cota superior de $\mathbb{N}$.

Recíprocamente, suponiendo válido el enunciado que acabamos de probar podemos demostrar que $\mathbb{N}$ no está acotado superiormente. En efecto, tomando $a = 1$ y aplicando la propiedad arquimediana concluimos inmediatamente que existen números naturales mayores que cualquier número real dado. $\Diamond$

> Debido a la propiedad que acabamos de probar, se dice que $\mathbb{R}$ es un *cuerpo arquimediano completo*.

PARTE ENTERA Y MANTISA DE UN NÚMERO REAL. En la siguiente proposición daremos forma precisa al hecho intuitivo de que todo número real se "encuentra" entre dos números enteros consecutivos:

Proposición 3.4.7 Todo número real x se expresa unívocamente en la forma

$$x = a + u\,, \tag{3.3}$$

donde $a \in \mathbb{Z}$ y u es un número real tal que $0 \leq u < 1$.

DEMOSTRACION. Comencemos probando que x admite una descomposición como la del enunciado. Puesto que el hecho es obvio si $x \in \mathbb{Z}$, ya que entonces basta tomar $a = x$ y $u = 0$, supongamos que $x \notin \mathbb{Z}$ y analizemos primero el caso en que x es positivo. Bajo esa hipótesis, el conjunto $A = \{k \in \mathbb{Z} : k < x\}$ es no vacío ($0 \in A$) y está acotado superiormente por x, por lo que admite un elemento máximo m (proposición 3.4.4). Teniendo en cuenta la definición de A, es claro que valen las desigualdades

$$m < x < m + 1,$$

o equivalentemente las desigualdades $0 < x - m < 1$. Obtenemos luego el resultado tomando $a = m$ y $u = x - m$.

Considerando por último el caso $x < 0$, sean $c \in \mathbb{Z}$ y $w \in (0,1)$ tales que $-x = c + w$. Entonces

$$x = -(-x) = -c - w = -c - 1 + 1 - w = -(c+1) + 1 - w\,,$$

y arribamos a la expresión deseada tomando $a = -(c+1)$ y $u = 1 - w$.

Para demostrar la unicidad de la representación, supongamos que x admite también la descomposición

$$x = b + v\,, \tag{3.4}$$

donde $b \in \mathbb{Z}$ y $v \in [0,1)$, y asumamos sin pérdida de generalidad que $a \leq b$. De esta suposición y de las igualdades 3.3 y 3.4 deducimos las relaciones

$$0 \leq b - a = u - v \leq u < 1\,,$$

lo que permite concluir que $b - a = 0$, pues no existen otros enteros en el intervalo $[0,1)$. Luego $b = a$ y $v = u$, como queríamos demostrar. $\Diamond$

> El número entero a se denomina la *parte entera* de x y el número real u se llama la *mantisa* de x.

Los notaremos $[\,x\,]$ y $\{\,x\}$, respectivamente. Remarquemos que $[\,x\,]$ es el mayor entero menor o igual que x y que vale la fórmula

$$x = [\,x\,] + \{\,x\}\,.$$

Por ejemplo, si $x = 3,5$ y $z = -7/5$ tenemos $[x] = 3$, $\{x\} = 0,5$, $[z] = -2$ y $\{z\} = 3/5$.

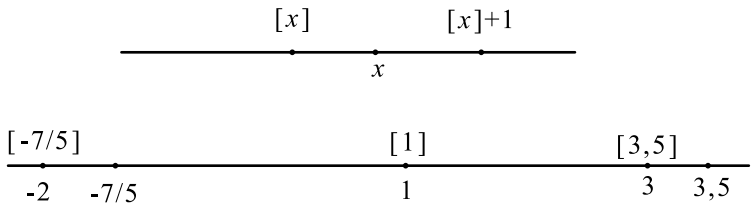


Figura 3.1: Parte entera

Enunciemos algunas propiedades de la parte entera y la mantisa. Todas ellas se deducen fácilmente de las definiciones precedentes, por lo que dejamos las demostraciones a cargo del lector.

Proposición 3.4.8 Valen las siguientes afirmaciones ($x, y \in \mathbb{R}$ y $m \in \mathbb{Z}$):

1) $[x] = x \iff x \in \mathbb{Z}$

2) $[x] = m \iff m \leq x < m + 1$

3) $m \leq x \iff m \leq [x]$

4) $\{x\} = \{y\} \iff x - y \in \mathbb{Z}$

5) $[m + x] = m + [x]$

6) $[x] + [y] \leq [x + y] \leq [x] + [y] + 1$

7) $[-x] = -[x] - 1$ si $x \notin \mathbb{Z}$. $\quad \Diamond$

3.4.2. Números racionales

Vamos ahora al encuentro de otro importante subconjunto de $\mathbb{R}$, el de los números racionales, llamados así por ser cociente o razón de números enteros. Su precisa definición es la siguiente:

> Un número real x se dice *racional* si y sólo si existen números enteros a y b $(b \neq 0)$ tales que $x = a/b$. En otro caso, diremos que x es un número *irracional*.

Designaremos por $\mathbb{Q}$ el conjunto de números racionales.

Observaciones Destaquemos algunos hechos estrechamente relacionados con la definición de número racional. El lector comprobará fácilmente la validez de todas las afirmaciones.

i) Todo número entero a es racional, ya que $a = a/1$. Luego $\mathbb{Z} \subset \mathbb{Q}$.

ii) Un número racional admite infinitas representaciones como cociente de números enteros, ya que $a/b = ka/kb$ para todo entero $k \neq 0$.

iii) Sean $u = a/b$ y $v = c/d$ en $\mathbb{Q}$. Entonces $u = v \Leftrightarrow ad = bc$. Sigue en particular que $u = 0 \iff a = 0$.

iv) Se deduce de la igualdad $a/b = (-a)/(-b)$ que todo racional puede representarse como un cociente de enteros con denominador positivo.

v) Sean u y v como en iii), con denominadores positivos. Entonces

$$u < v \Longleftrightarrow ad < bc. \quad \Diamond$$

Estructura de los numeros racionales Veremos en lo que sigue que $\mathbb{Q}$ tiene una estructura algebraica similar a la de $\mathbb{R}$.

Proposición 3.4.9 $\mathbb{Q}$ es un cuerpo ordenado.

DEMOSTRACION. Comencemos probando que la suma y el producto de dos números racionales son números racionales, usando el hecho de que $\mathbb{Z}$ es cerrado respecto a dichas operaciones. En efecto, dados números racionales a/b y c/d es inmediato verificar la validez de las igualdades

$$\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}$$

y

$$\frac{a}{b} \times \frac{c}{d} = \frac{ac}{bd} \ ,$$

lo que muestra que en ambos casos el resultado es un cociente de enteros.

Nos encontramos entonces nuevamente con una situación que ya empieza a ser familiar. Tenemos un conjunto ($\mathbb{Q}$) con dos operaciones (las restricciones a $\mathbb{Q}$ de la suma y el producto en $\mathbb{R}$) que satisfacen todas las propiedades estructurales de asociatividad, conmutatividad, etc, y contiene además al 0 y al 1 (por ser enteros), que son los respectivos elementos neutros. Observemos por último que los inversos aditivo y multiplicativo de un número racional a/b también son números racionales, ya que sigue de las identidades de arriba que dichos inversos son respectivamente $(-a)/b$ y b/a, ambos cocientes de enteros (naturalmente, suponemos $a \neq 0$ en el segundo caso).

Hemos probado así que $\mathbb{Q}$ es un cuerpo. Puesto que los axiomas que definen la relación de orden en $\mathbb{R}$ son obviamente válidos cuando se los aplica a números racionales, concluimos que $\mathbb{Q}$ es un cuerpo ordenado. $\Diamond$

Potencias de exponente racional.

Vamos ahora a extender la definición de potencia de exponente entero al caso de exponente racional. Consideremos para ello un número real positivo a y sean $m, n \in \mathbb{Z}$ $(n > 0)$. Definimos entonces:

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

Observará el lector que el carácter positivo de a asegura la existencia de la raíz $n-$ésima de a^m.

Por ejemplo, $4^{3/2} = 8$ y $25^{-1/2} = 1/5$. Esta definición de potencia de exponente racional, que claramente generaliza la noción de potencia de exponente entero, apunta además a preservar la validez de las propiedades enunciadas en la proposición 3.4.3, ya que en particular debiera verificarse entonces la igualdad

$$\left(a^{m/n}\right)^n = a^{(m/n)n} = a^m \, .$$

Enumeraremos a continuación las correspondientes propiedades de las potencias de exponente racional, quedando las demostraciones a cargo del lector.

Proposición 3.4.10 Sean $x, y \in \mathbb{Q}$ y sean a y b números reales positivos. Entonces

1) $a^x a^y = a^{x+y}$

2) $(a^x)^y = a^{xy}$

3) $(ab)^x = a^x b^x$. $\quad \Diamond$

Bien podría creer el lector a esta altura que no hay grandes diferencias entre $\mathbb{Q}$ y $\mathbb{R}$, ya que en ambas estructuras se satisfacen los doce axiomas de cuerpo ordenado descriptos en el capítulo 2. Pronto veremos sin embargo que existe una diferencia esencial: el axioma de completitud no es válido en $\mathbb{Q}$. A manera de prolegómeno a esta afirmación mostraremos un primer ejemplo de número irracional, lo que probará de paso que la inclusión $\mathbb{Q} \subset \mathbb{R}$ es estricta.

Ejemplo 3.4.11 $\sqrt{2} \notin \mathbb{Q}$.

Supongamos que la afirmación es falsa y elijamos una representación de $\sqrt{2}$ como cociente de números naturales cuyo denominador sea mínimo (principio de buena ordenación), digamos $\sqrt{2} = a/b$. Observemos entonces que $a^2 = 2b^2$ y que valen además las desigualdades $b < a < 2b$, pues $1 < \sqrt{2} < 2$.

Tenemos por otro lado que

$$a(a-b) = a^2 - ab = 2b^2 - ab = (2b-a)b,$$

lo que equivale a la igualdad

$$\frac{a}{b} = \frac{2b-a}{a-b} \ .$$

Puesto que $0 < a - b < b$, resulta que la fracción de la derecha es una representación de $\sqrt{2}$ como cociente de números naturales cuyo denominador es menor que b, lo que es una contradicción. Por lo tanto $\sqrt{2}$ no es racional.

De una manera muy similar puede demostrarse que $\sqrt{m}$ es irracional si m es natural y no es un cuadrado perfecto. Esto por supuesto prueba que hay infinitos números irracionales. $\quad \Diamond$

El siguiente lema, de interés propio, servirá de antesala a la demostración de la incompletitud de $\mathbb{Q}$.

Lema 3.4.12 Dados $x, y \in \mathbb{R}$ $(x < y)$ existe $q \in \mathbb{Q}$ tal que $x < q < y$.

Demostracion. Sin ser completamente rigurosos, contando solamente con la noción intuitiva de conjunto infinito que seguramente posee el lector,

observemos que se deduce del enunciado anterior que entre dos números reales existen infinitos números racionales. Vamos ahora a la prueba.

Puesto que $y - x > 0$, sigue por la propiedad arquimediana que existe $b \in \mathbb{N}$ tal que $b(y-x) > 2$. Designando por a la parte entera de by tenemos:

$$by > a - 1 = bx + a + 1 - (bx + 2) > bx + a + 1 - by > bx,$$

es decir, valen las desigualdades $bx < a - 1 < by$, de donde se obtiene dividiendo por b que

$$x < \frac{a-1}{b} < y \ .$$

Basta tomar entonces $q = \dfrac{a-1}{b}$. $\quad \Diamond$

Proposición 3.4.13 El axioma de completitud no es válido en $\mathbb{Q}$.

Demostracion. Mostraremos que el conjunto

$$A = \left\{ x \in \mathbb{Q} \ : x < \sqrt{2} \right\}$$

está acotado superiormente en $\mathbb{Q}$ y no admite supremo racional.

La primera afirmación es obvia, ya que todo número natural mayor que 1 es una cota superior de A. Supongamos que A admite supremo en $\mathbb{Q}$, digamos s, y comparémoslo con $\sqrt{2}$ (sabemos que $s \neq \sqrt{2}$).

Si $s > \sqrt{2}$, sigue de 3.4.12 que existe $q \in \mathbb{Q}$ tal que $\sqrt{2} < q < s$, resultando entonces que q es una cota superior de A menor que s, lo que contradice el hecho de que s es la menor de las cotas superiores de A. Luego este caso no es posible y debe ser $s < \sqrt{2}$.

Nuevamente por 3.4.12, existe $r \in \mathbb{Q}$ tal que $s < r < \sqrt{2}$, lo que también nos lleva a una contradicción, ya que tendríamos en tal caso que $r \in A$ y $r > \sup A$. Por lo tanto A no admite supremo racional, lo que demuestra el enunciado. El lector podrá observar de paso que la demostración anterior asegura que $\sup A = \sqrt{2}$. $\quad \Diamond$

3.4.3. Ejercicios

1. Demostrar las propiedades enunciadas en la proposición 3.4.3

2. Sea $a \in \mathbb{R}\,(a \geq 1)$ y sea $q \in \mathbb{N}_2$. Probar que existe $n \in \mathbb{N}_0$ tal que

$$q^n \leq a < q^{n+1}.$$

3. Sea $x \in \mathbb{R}$ y sea $a = [x]$. Expresar en función de a los posibles valores de $[\,x/2\,]$, $[\,x - 7/3\,]$, $[\,x + a\,]$, $[\,x - a\,]$, $[\,ax\,]$ y $[\,5x + 6a\,]$.

4. Demostrar las propiedades enunciadas en la proposición 3.4.8

5. Sea $a \in \mathbb{R}$ y sea $n \in \mathbb{N}$.

   a) Si n es impar, probar que existe un único $b \in \mathbb{R}$ tal que $b^n = a$.

   b) Si n es par y $a > 0$, probar que la ecuación $x^n = a$ admite exactamente dos soluciones en $\mathbb{R}$.

6. *a*) Sea $a \in \mathbb{R}\,(a > 0)$ y sea $x \in \mathbb{Q}$. Probar que la definición de a^x brindada en el texto es independiente de la forma particular elegida para representar a x como cociente de enteros.

   b) Demostrar las propiedades enunciadas en la proposición 3.4.10.

7. Sea F un subconjunto de $\mathbb{Q}$, cerrado para la suma, que contiene los inversos aditivo y multiplicativo de cada uno de sus elementos. Probar que $1 \in F$ si y solo si $F = \mathbb{Q}$.

8. Sea n un número natural que no es un cuadrado perfecto (no existe $m \in \mathbb{Z}$ tal que $m^2 = n$). Probar que $\sqrt{n}$ es un número irracional.

9. Analizar la validez de las siguientes afirmaciones ($a, b \in \mathbb{R}$):

   a) a es irracional si y solo si $-a$ es irracional

   b) Si a y b son irracionales entonces $a + b$ es irracional

   c) Si a es irracional y b es racional entonces $a + b$ es irracional

   d) a es irracional si y solo si a^{-1} es irracional

   e) Si a y b son irracionales entonces ab es irracional.

   f) Si a es irracional y b es racional entonces ab es irracional

   g) Si a es irracional y b es un racional no nulo entonces ab es irracional.

10. Sean $a, b, c, d \in \mathbb{Q}$ y sea n un número natural que no es un cuadrado perfecto. Probar que $a + b\sqrt{n} = c + d\sqrt{n}$ si y solo si $c = a$ y $d = b$.

11. Hallar $x \in \mathbb{Q}$, con denominador mínimo, tal que $\sqrt{2} < x < \sqrt{3}$. Resolver la misma cuestión respecto a las relaciones $\sqrt{65} < x < \sqrt{67}$.

12. Sea A un subconjunto no vacío de $\mathbb{Q}$, acotado superiormente, y sea s el supremo de A en $\mathbb{R}$. Probar que A admite supremo en $\mathbb{Q}$ si y solo si $s \in \mathbb{Q}$.

Capítulo 4

Cardinalidad

4.1. Conjuntos finitos

4.1.1. Coordinabilidad

Recordemos (capítulo 1) que dos conjuntos A y B se dicen coordinables si y solo si existe una biyección entre ellos. Empleando en tal caso la notación $A \approx B$, es inmediato probar que la relación $\approx$ es una relación de equivalencia en cualquier colección no vacía de conjuntos.

> Diremos que un conjunto A es *finito* si y solo si $A = \emptyset$ ó A es coordinable con $\mathbb{I}_n$ para algún $n \in \mathbb{N}$. Sigue en particular que cada sección inicial $\mathbb{I}_n$ es un conjunto finito. Si A no es finito, diremos que A es un conjunto *infinito*.

Probaremos que el número natural n de la definición anterior es único. Puesto que la relación de coordinabilidad es transitiva, bastará demostrar que $\mathbb{I}_n$ no es coordinable con $\mathbb{I}_m$ si $n \neq m$, para lo cual será suficiente probar el siguiente resultado, algo más específico:

Teorema 4.1.1 No existen funciones inyectivas de $\mathbb{I}_m$ en $\mathbb{I}_n$ si $m > n$.

Demostracion. Procediendo por inducción en n vemos que el caso $n = 1$ es trivial, ya que $f(1) = f(m) = 1$ para toda función $f : \mathbb{I}_m \mapsto \mathbb{I}_1$.

Asumiendo ahora que la afirmación es cierta para un cierto $k \in \mathbb{N}$, supongamos por el absurdo que existe una función inyectiva $f : \mathbb{I}_m \mapsto \mathbb{I}_{k+1}$ con $m > k + 1$. Notemos entonces que $k + 1 \in Im(f)$, ya que en caso contrario obtendríamos por correstricción una función inyectiva de $\mathbb{I}_m$ en $\mathbb{I}_k$, contradiciendo la hipótesis inductiva.

Una contradicción semejante tiene lugar si $k + 1 = f(m)$, ya que en tal caso la restricción de f a $\mathbb{I}_{m-1}$ determina una inyección de $\mathbb{I}_{m-1}$ en $\mathbb{I}_k$, lo que no es posible por ser $m - 1 > k$. Podemos suponer por lo tanto que existe $j \neq m$ tal que $f(j) = k + 1$.

Definimos entonces la función $g : \mathbb{I}_m \mapsto \mathbb{I}_m$ por

$$g(x) = \begin{cases} x & \text{si } x \neq j \text{ y } x \neq m \\ m & \text{si } x = j \\ j & \text{si } x = m. \end{cases}$$

Es fácil ver que g es biyectiva, de donde deducimos que la composición

$$h = f \circ g : \mathbb{I}_m \mapsto \mathbb{I}_{k+1}$$

también es inyectiva, sólo que ahora $h(m) = f(g(m)) = f(j) = k+1$, lo cual ya vimos, conduce a una contradicción. Esto completa la demostración. $\Diamond$

Corolario 4.1.2 $\mathbb{I}_m \approx \mathbb{I}_n \iff m = n.$ $\Diamond$

Cardinales finitos.

El corolario precedente garantiza la consistencia de la siguiente definición:

> Sea A un conjunto finito y sea $n \in \mathbb{N}$ tal que A es coordinable con $\mathbb{I}_n$. Diremos entonces que A es de *cardinal* n (o también que A tiene n elementos), y emplearemos la notación $\#(A) = n$. Si $A = \emptyset$ diremos que A tiene cardinal 0 y escribiremos $\#(A) = 0$. Se deduce de 4.1.2 que dos conjuntos finitos son coordinables si y solo si tienen el mismo cardinal.

Observemos que las definiciones de conjunto finito y de cardinal o número de elementos del mismo reproducen fielmente nuestra intuición respecto a dichos conceptos. En efecto, cuando contamos los elementos de un conjunto finito X lo que hacemos es numerarlos desde 1 hasta un cierto número natural n, acción que corresponde en términos matemáticos a establecer una biyección de X en $\mathbb{I}_n$.

Iniciaremos ahora un desarrollo teórico en el que nos ocuparemos de diversas propiedades de los conjuntos finitos, comenzando por estudiar la cardinalidad de sus subconjuntos:

Teorema 4.1.3 Sea A un conjunto finito de n elementos ($n \in \mathbb{N}$) y sea S un subconjunto de A. Entonces S es finito y $\#(S) \leq n$. Además, $\#(S) = n$ si y solo si $S = A$.

Demostracion. Ambas afirmaciones son obvias si $S = A$, por lo que supondremos que S es un subconjunto propio de A. Asimismo, la restricción a S de cualquier biyección de A en $\mathbb{I}_n$ induce una biyección de S en un subconjunto propio de $\mathbb{I}_n$, por lo que podemos suponer −usando las propiedades de la relación de coordinabilidad− que S es un subconjunto propio de

$\mathbb{I}_n$. Probaremos entonces por inducción en n que $S = \emptyset$ o existe un número natural $m < n$ tal que $S \approx \mathbb{I}_m$.

El caso $n = 1$ es inmediato, ya que el único subconjunto propio de $\mathbb{I}_1$ es el conjunto vacío. Supongamos ahora que el resultado vale para $n = k$ y sea S un subconjunto propio de $\mathbb{I}_{k+1}$.

Si $k + 1 \notin S$ entonces $S = \mathbb{I}_k$ ó S es un subconjunto propio de $\mathbb{I}_k$. En el primer caso $\#(S) = k < k + 1$, y en el segundo el resultado sigue por hipótesis inductiva.

Suponiendo que $k+1 \in S$, resulta que $T = S-\{k+1\}$ es un subconjunto propio de $\mathbb{I}_k$, y podemos emplear nuevamente la hipótesis inductiva. Si $T = \emptyset$ es claro que $S = \{k+1\}$, lo que prueba nuestra afirmación, ya que entonces $S \approx \mathbb{I}_1$ (notemos que $1 < k + 1$).

Si $T \neq \emptyset$ existe un natural $r < k$ y una biyección $g : T \mapsto \mathbb{I}_r$, en cuyo caso la función h definida por

$$h(x) = \begin{cases} g(x) & \text{si } x \in T \\ r+1 & \text{si } x = k+1 \end{cases}$$

establece una biyección entre S y $\mathbb{I}_{r+1}$, como se prueba inmediatamente. Puesto que $r + 1 < k + 1$, nuestra aserción queda demostrada. $\Diamond$

Obtenemos como corolario de 4.1.3 la versión funcional del Principio de Dirichlet, familiarmente llamado *Principio del Palomar* (volveremos luego sobre él):

Teorema 4.1.4 Sean A y B conjuntos finitos tales que $\#(A) > \#(B)$. No existe entonces ninguna función inyectiva de A en B.

Demostracion. Si $f : A \mapsto B$ es inyectiva entonces A es coordinable con $\mathrm{Im}(f) \subseteq B$, de donde sigue que $\#(\mathrm{Im}(f)) = \#(A) > \#(B)$, en contradicción con el resultado precedente. $\Diamond$

Veamos cómo la suma de cardinales se corresponde con una de las operaciones de conjuntos.

Teorema 4.1.5 (Principio aditivo) Sean A y B conjuntos finitos disjuntos. Entonces $A \cup B$ es finito y

$$\#(A \cup B) = \#(A) + \#(B).$$

Más generalmente, sea $r \in \mathbb{N}$ y sean $A_1, A_2, \ldots, A_r$ conjuntos finitos disjuntos dos a dos . Entonces $A_1 \cup A_2 \cup \ldots \cup A_r$ es finito y

$$\#(A_1 \cup A_2 \cup \ldots \cup A_r) = \sum_{i=1}^{m} \#(A_i).$$

DEMOSTRACION. El enunciado es obvio si alguno de los conjuntos dados es vacío. En otro caso, existen funciones biyectivas $f : \mathbb{I}_m \mapsto A$ y $g : \mathbb{I}_n \mapsto B$, donde $m = \#(A)$ y $n = \#(B)$. Encomendamos al lector la tarea de probar entonces que la función $h : \mathbb{I}_{n+m} \mapsto A \cup B$, definida por

$$h(k) = \begin{cases} f(k) & \text{si } k \leq m \\ g(k-m) & \text{si } k > m, \end{cases}$$

es biyectiva, lo que prueba nuestra afirmación. En cuanto a la generalización, la misma sigue fácilmente por inducción en r. $\Diamond$

Aplicando convenientemente el principio aditivo se obtienen fórmulas para calcular el cardinal de diversos conjuntos finitos, como enunciamos a continuación:

Proposición 4.1.6 Sean A y B conjuntos finitos. Entonces $A - B$, $A \,\Delta\, B$ y $A \cup B$ son finitos y valen las fórmulas:

$$\#(A - B) = \#(A) - \#(A \cap B) = \#(A \cup B) - \#(B) \tag{4.1}$$
$$\#(A \,\Delta\, B) = \#(A) + \#(B) - 2\#(A \cap B) \tag{4.2}$$
$$\#(A \cup B) = \#(A) + \#(B) - \#(A \cap B)\,. \tag{4.3}$$

DEMOSTRACION. A cargo del lector. $\Diamond$

Con el objeto de extender lo establecido para la unión de dos conjuntos finitos a la unión de una cantidad finita de conjuntos finitos, supongamos que A, B y C son conjuntos finitos cualesquiera, de donde deducimos en primer lugar que $A \cup B \cup C$ es finito, por ser unión de los conjuntos finitos $A \cup B$ y C. En relación con su cardinal, aplicando reiteradamente la fórmula 4.3 y la propiedad distributiva de la intersección respecto a la unión, obtenemos:

$$\begin{aligned} \#\left(A \cup B \cup C\right) &= \#(A) + \#(B \cup C) - \#\left(A \cap (B \cup C)\right) \\ &= \#(A) + \#(B \cup C) - \#\left((A \cap B) \cup (A \cap C)\right) \\ &= \#(A) + \#(B) + \#(C) - \#(B \cap C) - \#(A \cap B) \\ &\quad - \#(A \cap C) + \#\left((A \cap B) \cap (A \cap C)\right)\,, \end{aligned}$$

de donde resulta la fórmula

$$\begin{aligned} \#(A \cup B \cup C) \;=\; & \#(A) + \#(B) + \#(C) - \\ & - (\#(A \cap B) + \#(A \cap C) + \#(B \cap C)) + \\ & + \#(A \cap B \cap C). \end{aligned}$$

Comparando esta fórmula con la dada para $\#(A \cup B)$, es posible que el lector advierta un patrón de formación de las mismas. Con un procedimiento análogo (aunque con mayor complejidad notacional) no es difícil probar por inducción una fórmula general para calcular el cardinal de la unión de un número finito de conjuntos finitos. Su enunciado completo es el siguiente:

Teorema 4.1.7 Sea $n \in \mathbb{N}$ y sea $A_1, A_2, \ldots, A_n$ una familia de conjuntos finitos. Entonces $A_1 \cup A_2 \cup \ldots \cup A_n$ es un conjunto finito y vale la fórmula:

$$\begin{aligned}\#(A_1 \cup \ldots \cup A_n) = &\sum_{1\le i\le n} \#(A_i) - \sum_{1\le i_1<i_2\le n} \#(A_{i_1} \cap A_{i_2}) + \ldots + \\ &+ (-1)^{k-1} \sum_{1\le i_1<\ldots<i_k\le n} \#(A_{i_1} \cap \ldots \cap A_{i_k}) + \ldots + \\ &+ (-1)^{n-1} \#(A_1 \cap \ldots \cap A_n)\,. \quad \Diamond \end{aligned} \tag{4.4}$$

DEMOSTRACION La dejamos al cuidado del lector interesado (en la próxima sección ofreceremos una prueba combinatoria del teorema, como aplicación de la fórmula del binomio), aunque vale la pena aclarar cómo son los términos de la suma del enunciado: primero se suman los cardinales de los conjuntos A_i, luego se restan los cardinales de todas las intersecciones de dos cualesquiera de ellos, luego se suman los cardinales de todas las intersecciones de tres cualesquiera de ellos y así sucesivamente, para finalizar sumando (o restando, según sea la paridad de n) el cardinal de la intersección de todos los A_i. $\Diamond$

En ocasiones, y contrariamente a la situación anterior, puede interesarnos conocer el número de elementos de un conjunto A que no pertenecen a ninguno de ciertos subconjuntos A_i de A. En tal caso resulta útil la siguiente fórmula:

Corolario 4.1.8 Sea $n \in \mathbb{N}$ y sean $A_1, \ldots, A_n$ subconjuntos de un conjunto finito S. Entonces

$$\#(A_1^c \cap \ldots \cap A_n^c) = \#(S) - \left(\sum_{k=1}^{n} (-1)^{k-1} \sum_{1\le i_1<\ldots<i_k\le n} \#(A_{i_1} \cap \ldots \cap A_{i_k} \right) .$$

DEMOSTRACION. La fórmula es consecuencia immediata del teorema 4.1.7 y de la igualdad

$$\#(A_1^c \cap \ldots \cap A_n^c) = \#(S) - \#(A_1 \cup \ldots \cup A_n). \quad \Diamond$$

Existe un resultado dual del principio del palomar (teorema 4.1.4), referido a la existencia de funciones suryectivas de un conjunto finito en otro:

Teorema 4.1.9 Sean A y B conjuntos finitos tales que $\#(A) < \#(B)$. No existe entonces ninguna función suryectiva de A en B.

DEMOSTRACION. Asumamos por el contrario que existe una función suryectiva $f : A \to B$. Si para cada $b \in B$ definimos $S_b = \{a \in A : f(a) = b\}$, sigue que $\{S_b : b \in B\}$ es una colección de subconjuntos no vacíos de A (por

ser f suryectiva), disjuntos dos a dos y cuya unión es A, ya que $a \in S_{f(a)}$ para todo $a \in A$. Luego, por el principio aditivo tenemos:

$$\#(A) = \sum_{b \in B} \#(S_b) \geq \sum_{b \in B} 1 = \#(B)\,,$$

lo que contradice la hipótesis. $\Diamond$

Se registra un hecho sumamente importante en conexión con funciones cuyo dominio y codominio son conjuntos finitos del mismo cardinal. Precisamente, demostraremos en el siguiente teorema que en tal caso los conceptos de inyectividad, suryectividad y biyectividad coinciden.

Teorema 4.1.10 Sean A y B conjuntos finitos de igual cardinal y sea f una función de A en B. Entonces,

$$f \text{ es inyectiva} \iff f \text{ es suryectiva} \iff f \text{ es biyectiva.}$$

Demostracion. Ciertamente bastará probar la primera de las equivalencias, para lo cual comenzaremos suponiendo que f es inyectiva. Queda inducida entonces por correstricción una biyección entre A y $\mathrm{Im}(f)$, de donde resulta tomando cardinales que

$$\#(A) = \#(\mathrm{Im}(f)) \leq \#(B) = \#(A),$$

lo que implica que $\#(\mathrm{Im}(f)) = \#(B)$. Aplicando el teorema 4.1.3 concluimos que $\mathrm{Im}(f) = B$, esto es, f es suryectiva.

Supongamos ahora que f es suryectiva y consideremos como en 4.1.9 la partición de A determinada por los conjuntos S_b. Por el principio aditivo tenemos ahora

$$\#(A) = \sum_{b \in B} \#(S_b) \geq \sum_{b \in B} 1 = \#(B) = \#(A)\,,$$

relaciones que sólo pueden verificarse si $\#(S_b) = 1$ para todo $b \in B$. Luego, para cada $b \in B$ existe un único $a \in A$ tal que $f(a) = b$, lo que significa que f es inyectiva. $\Diamond$

Usando el teorema 4.1.5 podemos determinar el cardinal de un producto cartesiano de conjuntos finitos.

Teorema 4.1.11 (Principio multiplicativo) Sean A y B conjuntos finitos. Entonces $A \times B$ es finito y

$$\#(A \times B) = \#(A)\#(B)\,.$$

Más generalmente, sean $A_1, A_2, \cdots, A_m$ conjuntos finitos ($m \in \mathbb{N}$). Entonces $A_1 \times A_2 \times \cdots \times A_m$ es finito y

$$\#\left(A_1 \times A_2 \times \cdots \times A_m\right) = \prod_{i=1}^{m} \#(A_i)\,.$$

DEMOSTRACION. Podemos suponer que A y B son no vacíos, ya que en otro caso el resultado es trivial.

Si $a \in A$, sea T_a el subconjunto de $A \times B$ definido por

$$T_a = \{(a, y) : y \in B\},$$

es decir, T_a es el conjunto de elementos de $A \times B$ cuya primera componente es a. Es inmediato probar que T_a es coordinable con B, a través de la biyección $(a, y) \mapsto y$.

Por otra parte, la familia de conjuntos $\{T_a : a \in A\}$ determina una partición de $A \times B$. En efecto, cada T_a es no vacío (pues $B \neq \emptyset$), $T_a \cap T_c = \emptyset$ si $a \neq c$ (lo que es obvio), y finalmente, todo elemento de $A \times B$ está en alguno de los miembros de la familia, a saber, $(x, y) \in T_x$. Luego $A \times B$ es finito por ser unión de un número finito de conjuntos finitos, y aplicando el principio aditivo tenemos

$$\#(A \times B) = \sum_{a \in A} \#(T_a) = \sum_{a \in A} \#(B) = \#(A)\#(B),$$

como queríamos probar. La segunda parte se demuestra sin dificultad por inducción en m. $\quad \Diamond$

Como aplicación interesante obtenemos el cardinal del conjunto de partes de cualquier conjunto finito.

Corolario 4.1.12 Sea A un conjunto finito cualquiera. Entonces $\mathbb{P}(A)$ es finito y

$$\#(\mathbb{P}(A)) = 2^{\#(A)}.$$

DEMOSTRACION. Puesto que el caso $A = \emptyset$ es trivial, asumamos que A tiene un número natural n de elementos y consideremos una numeración cualquiera $\{a_1, \ldots, a_n\}$ de los elementos de A. Definimos entonces una función

$$f : \mathcal{P}(A) \mapsto \{0, 1\} \times \{0, 1\} \times \cdots \times \{0, 1\} = \{0, 1\}^n,$$

asignando a cada $S \subseteq A$ la $n-$upla $(e_1, e_2, \ldots, e_n)$ de ceros y unos definida por

$$e_i = \begin{cases} 1 & \text{si } a_i \in S \\ 0 & \text{si } a_i \notin S \end{cases}$$

Por ejemplo, $f(\emptyset) = (0, 0, \ldots, 0)$ y $f(A) = (1, 1, \ldots, 1)$. Es tarea sencilla probar que la función f es biyectiva, obteniéndose así el resultado deseado por aplicación del principio multiplicativo, ya que entonces

$$\#(\mathbb{P}(A)) = \#(\{0, 1\}^n) = (\#(\{0, 1\})^n = 2^n. \quad \Diamond$$

Número de funciones.

Si A y B son conjuntos, designaremos por B^A el conjunto de funciones de A en B. Puesto que cualquiera de ellas es una relación de A en B, vale decir un subconjunto del producto cartesiano $A \times B$, es claro a partir de 4.1.11 y 4.1.12 que B^A es finito si A y B lo son. En la siguiente proposición determinaremos en tal caso el cardinal de B^A, así como también el número de funciones inyectivas y biyectivas de A en B, cantidades que denotaremos por $iny(A, B)$ y $biy(A, B)$, respectivamente.

Proposición 4.1.13 Sean A y B conjuntos finitos tales que $\#(A) = m$ y $\#(B) = n$. Valen entonces las siguientes fórmulas (en el segundo caso suponemos $m \leq n$ y en el tercero $m = n$):

$$\#\left(B^A\right) = n^m \tag{4.5}$$

$$iny(A, B) = \frac{n!}{(n-m)!} \tag{4.6}$$

$$biy(A, B) = m!\,. \tag{4.7}$$

DEMOSTRACION Para probar (4.5) observemos que B^A es coordinable con B^m, a través de la biyección $\sigma : B^A \to B^m$ definida por

$$\sigma(f) = (f(a_1), f(a_2), \cdots, f(a_m))\,,$$

siendo $\{a_1, a_2, \cdots, a_m\}$ una numeración cualquiera de A. Luego:

$$\#(B^A) = \#(B^m) = \#(B)^m = n^m.$$

Respecto de (4.6), recordando que no existen funciones inyectivas de A en B si $m > n$ (teorema 4.1.4), efectuaremos la prueba por inducción en m.

Si $m = 1$ toda función de A en B es inyectiva, resultando entonces

$$iny(A, B) = \#\left(B^A\right) = n = \frac{n!}{(n-1)!}\ .$$

Tomemos ahora $m > 1$ y asumamos que el resultado es válido para $m-1$. Si X es el conjunto de funciones inyectivas de A en B, consideremos, para cada $b \in B$, el conjunto X_b de elementos de X que aplican a_m en b. Una tal función define por restricción una función inyectiva de $A' = \{a_1, \ldots, a_{m-1}\}$ en $B' = B - \{b\}$, mientras que toda función inyectiva $h : A' \to B'$ se extiende de manera obvia a un elemento de X_b, simplemente definiendo $h(a_m) = b$.

Es trivial verificar que las asignaciones anteriores de restricción y extensión son mutuamente inversas, lo que muestra que X_b es coordinable con el conjunto de funciones inyectivas de A' en B'. Puesto que $\#\left(A'\right) = m - 1$ y $\#\left(B'\right) = n - 1$, sigue usando la hipótesis inductiva que

$$\#(X_b) = iny(A', B') = \frac{(n-1)!}{(n-1-(m-1))!} = \frac{(n-1)!}{(n-m)!}\ .$$

Por otro lado, es claro que $X = \bigcup_{b\in B} X_b$ (unión disjunta), de donde obtenemos:

$$\#(X) = \sum_{b\in B} \#(X_b) = \sum_{b\in B} \frac{(n-1)!}{(n-m)!} = n\frac{(n-1)!}{(n-m)!} = \frac{n!}{(n-m)!} \ ,$$

como queríamos probar.

Finalmente, usando (4.6) y el teorema 4.1.10 deducimos inmediatamente la validez de (4.7). $\Diamond$

4.1.2. Ejercicios

1. Respecto a la primera parte de la proposición 4.1.13, probar que σ es efectivamente una biyección.

2. A una función de teatro asisten 400 espectadores. Demostrar que entre ellos hay 2 que cumplen años el mismo día.

3. *a*) Sea $n \in \mathbb{N}$ y sea A un conjunto finito tal que $\#(A) \geq n+1$. Si $\mathcal{C}$ es una partición de A en n subconjuntos, probar que algún miembro de $\mathcal{C}$ tiene al menos 2 elementos.

   b) Sean $m, n \in \mathbb{N}$ y sea A un conjunto finito tal que $\#(A) \geq mn+1$. Si $\mathcal{C}$ es una partición de A en n subconjuntos, probar que algún miembro de $\mathcal{C}$ tiene al menos $m+1$ elementos.

   NOTA El enunciado *a*), del cual *b*) es una generalización, es la versión conjuntista del principio de Dirilchet, conocido también como principio del palomar, ya que por tradición se lo describe coloquialmente de la siguiente manera: si $n+1$ o más palomas se distribuyen en n nidos, algún nido deberá alojar por lo menos 2 palomas. En general, se aplica a situaciones en las que $n+1$ objetos deben ser distribuidos o clasificados en n categorías diferentes.

4. Dado un conjunto S de 51 números naturales menores que 101, probar que existen dos elementos de S cuya suma es 101.

5. Para un congreso se han contratado 5 traductores, que deberán cubrir 6 lenguas diferentes. Si cada una de éstas requiere el empleo de 3 traductores, demostrar que alguno de los intérpretes deberá traducir por lo menos 4 idiomas.

6. ¿Cuántos números de 4 cifras pueden formarse con los dígitos 1, 2, 3 y 4, si cada uno de ellos puede aparecer a lo sumo 2 veces?

7. De un grupo de 60 estudiantes, 40 cursan Algebra, 20 cursan Cálculo y 9 Computación. Sabiendo que 18 sólo cursan Algebra, 10 sólo Cálculo, 6 sólo Computación, y que un único estudiante cursa las 3 materias, determinar cuántos alumnos no cursan ninguna de ellas.

8. Una cierta cantidad de equipos de fútbol han disputado un torneo por el sistema de "todos contra todos". Probar que algún par de equipos ha empatado la misma cantidad de partidos.

9. Sean $m, n \in \mathbb{N}$. Determinar:

   a) El número de funciones $f : \mathbb{I}_m \to \mathbb{I}_n$ tales que $f(1) \geq f(2)$.

   b) El número de funciones inyectivas $f : \mathbb{I}_6 \to \mathbb{I}_{10}$ tales que $f(4) = 2$ y $f(6) < 5$.

   c) El número de funciones suryectivas de $\mathbb{I}_n$ en $\mathbb{I}_2$ y de $\mathbb{I}_n$ en $\mathbb{I}_3$.

10. Sea $m \in \mathbb{N}$ y sea A un conjunto de m elementos. Calcular:

    a) El número de relaciones en A.

    b) El número de relaciones reflexivas en A.

    c) El número de relaciones reflexivas y simétricas en A.

11. Determinar el número de relaciones de equivalencia $\sim$ en $\mathbb{I}_9$, con exactamente 3 clases de equivalencia, que verifican las condiciones $1 \sim 4$, $2 \sim 5$, $3 \nsim 6$ y $9 \sim 5$.

4.2. Combinatoria

4.2.1. Técnicas para contar

Con el apoyo teórico de los resultados expuestos en la sección anterior acerca de cardinales finitos, desarrollaremos en esta sección los temas básicos de la *Combinatoria* elemental, a través de un enfoque principalmente dirigido a la resolución de problemas. Nuestro propósito es que el lector se adiestre en el reconocimiento de ciertas situaciones, abstrayendo sus características principales y descubriendo sus similitudes (o diferencias) con respecto a otras que puedan servirle de modelo y sea capaz de manejar. Por lo tanto, además de brindarle una serie de resultados y fórmulas, intentaremos familiarizarlo con algunas ideas y formas de pensar.

¿Qué es la Combinatoria ? Sin pretender encuadrarla en una definición rígida, podríamos describirla como la técnica o habilidad de *contar sin enumerar*, esto es, el desarrollo de métodos que nos permitan, por ejemplo, calcular el número de elementos de un conjunto, el número de formas posibles de ordenar un grupo de objetos, la cantidad total de resultados que puede arrojar una experimento, etc., sin confeccionar la lista exhaustiva de los mismos. Aunque todo esto suene algo ambiguo no debemos preocuparnos, ya que a través de unos pocos ejemplos captaremos rápidamente la naturaleza del tema. Abordemos entonces nuestro primer punto.

Principio de multiplicación.

Comenzaremos estableciendo un principio muy general, que no obstante su sencillez servirá de sustento a la mayoría de los temas que trataremos de aquí en adelante. Lo llamaremos *Principio general de multiplicación* (PGM), y su enunciado, de tipo coloquial, es el siguiente:

> Sean m y n números naturales y supongamos que la realización de una cierta experiencia E_1 puede arrojar m resultados, y que ante la ocurrencia de cualquiera de éstos otra experiencia E_2 puede arrojar n resultados. Entonces la realización conjunta de E_1 y E_2 puede arrojar mn resultados.

Hagamos notar que las palabras "experiencia" y "resultado" deben interpretarse en sentido amplio, y su significado varía con las aplicaciones. Por ejemplo, usemos el PGM para contar la cantidad de vocablos de 2 letras terminados en vocal, sin excluir los que no tienen aparición en el lenguaje. Podemos pensar que E_1 consiste en ubicar la letra inicial, pudiendo ser su resultado cualquiera de las 27 letras de nuestro alfabeto, y similarmente E_2 consistirá en la elección de la segunda letra, que deberá ser cualquiera de las 5 vocales. De acuerdo con el PGM, el número de tales digrafos es entonces $27 \,.\, 5 = 135$. Como otro ejemplo, supongamos que un edificio tiene 6 puertas y nos preguntamos por el número de formas de entrar por una de

ellas y salir por otra diferente. Aqui la primera experiencia puede arrojar 6 resultados (cualquiera de las puertas) mientras que la segunda 5 (cualquiera de las restantes). Luego, la respuesta es $30 = 6 \,.\, 5$ formas.

A pesar de lo intuitivo que resulta el PGM, brindaremos una prueba formal del mismo. A tal efecto, designemos por $R_1, R_2, \ldots, R_m$ los posibles resultados de E_1, y para cada i, notemos por A_i el conjunto de resultados que puede arrojar la realización de ambas experiencias habiéndose obtenido R_i en la primera. Puesto que claramente los A_i determinan una partición del conjunto X de realizaciones sucesivas de ambas experiencias, aplicando 4.1.5 obtenemos:

$$\#(X) = \sum_{i=1}^{m} \#(A_i) = \sum_{i=1}^{m} n = mn\,,$$

según queríamos demostrar $\quad \Diamond$

Como se puede apreciar, la validez del principio de multiplicación reside en el simple hecho de que sumar m veces n no es otra cosa que multiplicar m por n. Por ello, tengamos claro que para poder aplicarlo es fundamental que el número de resultados que puede arrojar la segunda experiencia sea siempre el mismo, *cualquiera sea el resultado de la primera*, exigencia que claramente se verifica en los dos ejemplos anteriores (nótese que si bien en el segundo la elección de la segunda puerta depende de cuál haya sido la primera, el número de elecciones posibles es siempre 5). Por ejemplo, no podríamos usar el PGM en la cuestión de los vocablos si agregásemos la condición de que la vocal preceda en el alfabeto a la letra inicial.

En muchas ocasiones es necesario aplicar reiteradamente el principio anterior, por lo que conviene establecer una versión más general del mismo. Su formulación precisa es la siguiente:

> Sea $r \geq 2$ y supongamos que se realizan sucesivamente ciertas experiencias $E_1, E_2, \ldots, E_r$. Asumamos que E_1 puede arrojar m_1 resultados, que por cada uno de ellos la experiencia E_2 puede arrojar m_2 resultados, y en general, que por cada realización conjunta de las experiencias $E_1, E_2, \ldots, E_{i-1}$ la experiencia E_i puede arrojar m_i resultados $(2 \leq i \leq r)$. Entonces la realización sucesiva de $E_1, \ldots, E_r$ puede arrojar $m_1 m_2 \ldots m_r$ resultados.

De ahora en más, emplearemos la denominación de principio general de multiplicación para referirnos tanto a la versión original como a esta generalización. Una prueba de la misma se obtiene fácilmente por inducción en r, por lo que dejamos los detalles a cargo del lector.

Ejemplo 4.2.1 Un viajero debe trasladarse de la ciudad A a la ciudad D, debiendo detenerse en su viaje en las ciudades B y C. Supongamos que hay 4 rutas que unen A con B, 3 que unen B con C y que hay 3 caminos posibles para ir de C a D. Bajo estas premisas, respondamos las siguientes preguntas:

a) ¿De cuántas formas puede el viajero realizar su itinerario?

b) ¿De cuántas formas puede realizar el recorrido de ida y vuelta?

c) Como en *b*), pero con la condición de que a la vuelta alguno de los tramos no sea recorrido por la misma ruta que a la ida.

d) ¿De cuántas maneras puede ir y volver si ningún camino puede ser usado más de una vez?

Con variantes, todas estas cuestiones se resuelven aplicando el principio de multiplicación. En el caso de la pregunta *a*) se trata de 3 "experiencias", que consisten en elegir un camino para cada uno de los tramos, por lo que el número posible de itinerarios es $4 \,.\, 3 \,.\, 3 = 36$.

Una vez elegida cualquiera de las 36 formas de llegar de A a D, es claro que hay también 36 formas de efectuar el recorrido de vuelta. Luego, aplicando el PGM para el caso de dos experiencias, resulta que hay $36 \,.\, 36 = 1296$ maneras de realizar el camino de ida y vuelta.

Respecto de *c*), observemos que la condición del enunciado determina que de los 36 itinerarios posibles de vuelta sólo debemos descartar uno (el que emplea en cada tramo la misma ruta que a la ida). Luego, la respuesta es $36 \,.\, 35 = 1260$.

Finalmente, podemos resolver *d*) aplicando el PGM al caso de la realización conjunta de 6 experiencias, pudiendo arrojar éstas 4, 3, 3, 2, 2 y 3 resultados, respectivamente, ya que en la elección de cada tramo de la vuelta una de las rutas está vedada. Por lo tanto, la cantidad de itinerarios sujetos a la condición de *d*) es $4 \,.\, 3 \,.\, 3 \,.\, 2 \,.\, 2 \,.\, 3 = 432$. $\Diamond$

Variaciones, permutaciones y combinaciones.

Examinemos los siguientes problemas: si una clave bancaria consiste de 4 dígitos, ¿de cuántas maneras distintas puede elegirse una? ¿De cuántas maneras si los dígitos deben ser todos distintos?

Se trata de dos situaciones claramente asimilables al caso general de aplicación del principio de multiplicación. En el primer interrogante cada uno de los dígitos de la clave puede ser seleccionado de 10 maneras distintas, mientras que en el segundo debemos descartar en cada elección los dígitos elegidos previamente. En consecuencia, y de acuerdo con el PGM, el número total de claves es $10000 = 10^4$ y el número de claves con todos sus dígitos distintos es $10 \,.\, 9 \,.\, 8 \,.\, 7 = 5040$.

Reflexionemos sobre las dos situaciones que acabamos de resolver, señalando sus similitudes y diferencias. En ambas debimos contar todas las formas posibles de seleccionar ordenadamente 4 objetos (dígitos en este caso) elegidos entre 10, difiriendo entre ellas en el hecho de que en la primera los objetos podían repetirse y en la segunda no. Generalizaremos estas situaciones, dada su frecuente aparición en combinatoria, y les asignaremos

terminología y notación específicas. En todos los casos m y n designarán números naturales.

VARIACIONES CON REPETICIÓN

> Dados n objetos distintos, cualquier elección *ordenada* de m de los mismos (distintos o no), se llamará una *variación con repetición* de n elementos tomados de a m.

Razonando en forma idéntica a la del problema anterior, obtenemos por aplicación directa del PGM la siguiente fórmula para contar el número de tales variaciones, que designaremos por $(VR)^n_m$:

Fórmula 4.2.2

$$(VR)^n_m = n^m \,.$$

VARIACIONES

> Suponiendo $m \leq n$ y dados n objetos distintos, cualquier elección *ordenada* de m distintos de ellos se llamará una *variación* de n elementos tomados de a m. El número de las mismas será notado V^n_m.

Como en el ejemplo de las claves, V^n_m se calcula en general aplicando el principio de multiplicación: tenemos n formas de elegir el primer elemento, $n-1$ formas de elegir el segundo, pues debe ser distinto del primero, $n-2$ formas de elegir el tercero, y así sucesivamente. Observando que luego de elegir los primeros $m-1$ elementos el último debe ser seleccionado entre los $n-(m-1)$ elementos restantes, obtenemos la fórmula:

Fórmula 4.2.3

$$V^n_m = n(n-1)(n-2)\ldots(n-m+1) = \frac{n!}{(n-m)!} \,.$$

Alternativamente, nos referiremos a las variaciones (con o sin repetición) de n elementos tomados de a m por el nombre de variaciones $m-$arias de n objetos.

Ejemplo 4.2.4 Para participar en un torneo de tenis de dobles mixtos (parejas de un hombre y una mujer), es necesario presentar un equipo de 3 parejas, debiéndose elegir los jugadores entre los integrantes de un grupo constituido por 6 hombres y 3 mujeres. ¿ De cuántas maneras puede seleccionarse el equipo?

Puesto que ineludiblemente todas las mujeres deben formar parte del equipo, solo deberá decidirse quiénes serán sus compañeros de juego. Por lo tanto, cada plantel quedará determinado por una selección ordenada de 3

hombres (obviamente distintos) entre 6, resultando entonces que el número de formas de armar el equipo es $V_3^6 = 120$. Vale la pena aclarar por qué la elección debe ser ordenada: no solo importa qué hombres van a jugar, sino también con cuál de las damas va a formar pareja cada uno de ellos. $\Diamond$

Ejemplo 4.2.5 Con los dígitos $1, 2, \ldots, 9$, ¿cuantos números de 3 cifras pueden formarse con la condición de que la suma de las mismas sea par?

Si no impusiéramos la restricción sobre la suma de las cifras, la solución sería sencillamente $(VR)_3^9 = 729$. Existiendo tal restricción, conviene comenzar con un breve análisis aritmético, separando los números a contar en dos clases: los que tienen sus 3 cifras pares y los que tienen exactamente una cifra par. Contaremos entonces cuántos podemos formar en cada uno de los dos casos, para luego sumar los resultados, ya que ambas situaciones se excluyen mutuamente (principio aditivo).

Primer caso (los tres pares). Debemos elegir ordenadamente 3 elementos del conjunto $\{2, 4, 6, 8\}$, por lo que la cantidad de elecciones es $(VR)_3^4 = 64$.

Segundo caso (un par y dos impares). Aplicaremos el PGM a la realización conjunta de 3 experiencias, a saber: elección de un dígito par, elección de la posición que él ocupará en el número, y por último, elección ordenada de 2 dígitos impares, que ocuparán las otras dos posiciones. Ellas pueden arrojar $V_1^4 = 4$, $V_1^3 = 3$ y $(VR)_2^5 = 20$ resultados, respectivamente, y es claro que éstos son independientes entre sí. Luego, usando el PGM resulta que la cantidad de números del segundo tipo que podemos formar es $4 \,.\, 3 \,.\, 20 = 240$.

En definitiva, la respuesta a nuestro problema es $64 + 240 = 284$. $\Diamond$

NOTA Si X es un conjunto de n elementos, cualquier elección ordenada de m de ellos determina una función f de $\mathbb{I}_m$ en X, definiendo $f(i)$ como el elemento de X que ocupa la $i-$ésima posición en dicho ordenamiento. Recíprocamente, toda función g de $\mathbb{I}_m$ en X corresponde a una única selección ordenada de m elementos de X, a saber, la secuencia $g(1)g(2)\cdots g(m)$, resultando que existe una biyección entre las variaciones m-arias de elementos de X y el conjunto de funciones de $\mathbb{I}_m$ en X, bajo la cual las variaciones sin repetición se corresponden con las funciones inyectivas.

Por ejemplo, volviendo al problema con el que iniciamos el tema, la clave bancaria 2702 determina la función $f(1) = 2$, $f(2) = 7$, $f(3) = 0$ y $f(4) = 2$, de $\mathbb{I}_4$ en el conjunto X de dígitos, mientras que la función $g : \mathbb{I}_4 \to X$ definida por $g(i) = 2i + 1$ corresponde a la clave 3579.

Esta interpretación de las variaciones como funciones de un cierto conjunto en otro nos brinda entonces una forma alternativa de obtener las fórmulas 4.2.2 y 4.2.3, ya que se tiene:

$$(VR)_m^n = \#\left(X^{\mathbb{I}_m}\right) = n^m$$

y

$$V_m^n = iny\,(\mathbb{I}_m, X) = \frac{n!}{(n-m)!}\ ,$$

por proposición 4.1.13.

PERMUTACIONES

> Dados n objetos distintos $x_1, x_2, \ldots, x_n$, cualquier secuencia de la forma
>
> $$x_{\pi(1)} x_{\pi(2)} \cdots x_{\pi(n)} \,,$$
>
> donde π es una biyección de I_n, se dirá una *permutación* de los mismos. Designaremos por P_n el número de tales arreglos, a los que también llamaremos permutaciones n-arias.

Por ejemplo, el lector puede comprobar que las siguientes son todas las permutaciones de los elementos $1, 2, 3, 4$:

1234	1243	1324	1342
1423	1432	2134	2143
2314	2341	2413	2431
3124	3142	3214	3241
3412	3421	4123	4132
4213	4231	4312	4321

En cuanto al número de permutaciones, observemos que una permutación n-aria no es otra cosa que una variación n-aria sin repetición de n elementos, por lo que usando 4.2.3 en el caso $m = n$ deducimos la validez de la siguiente fórmula para calcular P_n :

Fórmula 4.2.6

$$P_n = n! \,.$$

Antes de ilustrar con algunos ejemplos, y similarmente al caso de las variaciones, notemos que toda permutación $y_1 y_2 \ldots y_n$ de los elementos del conjunto $X = \{x_1, x_2, \ldots, x_n\}$ determina una biyección de X, definida a través de la asignación

$$x_i \mapsto y_i \,,$$

y que recíprocamente, toda biyección σ de X corresponde a una permutación de sus elementos, a saber, la permutación

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

Dejando los detalles a cargo del lector, resulta entonces que el conjunto de permutaciones de los elementos de X es coordinable con el conjunto de biyecciones de X, hecho que unido a la fórmula (4.7) de la proposición 4.1.13 nos brinda otra demostración de la validez de la fórmula 4.2.6.

Ejemplo 4.2.7 ¿En cuántas permutaciones de las cifras de 123456789 el 4 y el 7 conservan sus posiciones y el 1 ocupa una posición par? ¿En cuántas los múltiplos de 3 aparecen consecutivamente?

Respecto a la primera pregunta, observemos que podemos ubicar al 1 en el segundo, sexto u octavo lugar, ya que en la cuarta posición debe quedar el 4. Una vez ubicado aquél, resta ubicar los números 2, 3, 5, 6, 8 y 9, para lo cual bastará permutar estos de cualquier manera y ubicarlos de izquierda a derecha en los lugares vacíos, según el orden dado por la permutación elegida. Así, si 1 ocupase el segundo lugar, la permutacion 592683 de esos seis elementos determinaría el ordenamiento 519426783.

Habiendo 3 opciones para ubicar el 1 y $720 = 6!$ permutaciones de 6 elementos, sigue por el PGM que hay $3 \,.\, 720 = 2160$ ordenamientos de las 9 cifras que satisfacen los requerimientos del enunciado.

Respecto a la segunda cuestión, conviene pensar que son siete los elementos a permutar, considerando los múltiplos de 3 (3, 6 y 9) como un único bloque. Hay entonces $5040 = 7!$ formas de ordenarlos, resultando que cada una de ellas da lugar a su vez a $6 = 3!$ arreglos distintos de las 9 cifras, obtenidos permutando de cualquier manera los integrantes del bloque de múltiplos de 3. Nuevamente por el PGM, concluimos que la respuesta al problema es $5040 \,.\, 6 = 30240$. $\Diamond$

Ejemplo 4.2.8 Con cuentas de 9 colores diferentes, ¿cuántos collares distintos pueden diseñarse?

Si bien podría parecer de entrada que la solución consiste en una rutinaria aplicación del caso general (número de permutaciones de 9 objetos), la disposición circular de éstos introduce una variante, En efecto, supongamos que

$$C_1C_2C_3 \ldots C_8C_9$$

es una permutación cualquiera de las cuentas y consideremos la permutación

$$C_9C_1C_2 \ldots C_7C_8\,,$$

obtenida de la anterior desplazando cada cuenta un lugar a la derecha (la última pasa al primer lugar). A pesar de ser distintos, los dos ordenamientos determinan el mismo collar, ya que si los pensamos en forma circular vemos que el orden relativo entre las cuentas no varía de uno a otro. Naturalmente, lo mismo ocurre si cada cuenta se desplaza $2, 3, \ldots$, hasta 8 lugares hacia la derecha, resultando entonces que las 9 permutaciones así obtenidas deben contarse como si fueran la misma.

Inversamente, es muy sencillo demostrar que dos permutaciones determinan el mismo collar (el mismo orden relativo entre las cuentas) solo si una de ellas se obtiene de la otra a través de alguno de los desplazamientos que hemos descripto arriba, de donde concluimos que el número de posibles diseños es

$$\frac{P_9}{9} = \frac{9!}{9} = 8! = 40320. \quad \Diamond$$

COMBINACIONES Para ilustrar el tema planteamos la siguiente cuestión: dados 10 puntos en un plano tales que ninguna terna de ellos yace sobre una recta, ¿cuántos triángulos con vértices en dichos puntos quedan determinados?

La situación no parece muy alejada de otras que hemos resuelto. Simplemente debemos contar todas las formas de seleccionar 3 puntos distintos entre los 10, ya que la hipótesis asegura que cualquiera de esas elecciones determina un triángulo. Sin embargo, existe en este caso una importante diferencia: *el orden de elección es irrelevante*, ya que por ejemplo, la secuencia de puntos PQR determina el mismo triángulo que la secuencia QPR. Dejando en suspenso la respuesta a este problema particular, veamos cómo resolver en general situaciones de este tipo. Introducimos para ello la siguiente definición:

> Suponiendo $m \leq n$ y dados n objetos distintos, cualquier elección *no ordenada* de m distintos de ellos se llamará una *combinación* de n elementos tomados de a m. El número de las mismas será notado C_m^n.

Aclaremos el sentido de la expresión "elección no ordenada": significa que sólo importa cuáles son los elementos elegidos, sin atender ningún ordenamiento entre ellos. Vemos por ejemplo que $C_2^4 = 6$, pues dados los objetos x_1, x_2, x_3, x_4 resulta que x_1x_2, x_1x_3, x_1x_4, x_2x_3, x_2x_4 y x_3x_4 son todas las formas posibles de elegir 2 elementos distintos entre ellos, sin tener en cuenta ningún orden de elección. Notemos por caso que no distinguimos entre x_2x_3 y x_3x_2.

Parece evidente que el número de variaciones m-arias de n objetos distintos $x_1, x_2, \ldots, x_n$ es mayor que el número de combinaciones m-arias de los mismos. El siguiente lema nos brindará la relación exacta entre ambas cantidades:

Lema 4.2.9 $V_m^n = m!\, C_m^n$.

DEMOSTRACION Si $\mathcal{V}$ es el conjunto de variaciones m-arias de $x_1, x_2, \ldots, x_n$, definimos en $\mathcal{V}$ la siguiente relación:

$$y_1y_2 \ldots y_m \sim z_1z_2 \ldots z_m \Leftrightarrow \{y_1, y_2, \ldots, y_m\} = \{z_1, z_2, \ldots, z_m\},$$

esto es, dos variaciones están relacionadas si y solo si consisten de sendos ordenamientos (posiblemente distintos) de los mismos m elementos.

Es claro que $\sim$ es una relación de equivalencia en $\mathcal{V}$, y que la clase de equivalencia de cualquier variación m-aria $v = v_1v_2 \ldots v_m$ consiste de todas

las permutaciones de los elementos $\{v_1, v_2, \ldots, v_m\}$, resultando en particular que $\#(Cl(v)) = m!$. Finalmente, notemos que una clase de equivalencia está determinada por la elección no ordenada de m elementos distintos entre los x_i, ya que, fijados éstos, dos ordenamientos distintos de ellos definen la misma clase de equivalencia. En consecuencia, el número c de clases de equivalencia coincide con el número de combinaciones m-arias de los x_i.

De acuerdo con las consideraciones de arriba, y recordando que las clases de equivalencia determinan una partición (que en este caso tiene todos sus miembros de cardinal $m!$), obtenemos:

$$V_m^n = \#(\mathcal{V}) = m!\, c = m!\, C_m^n\,,$$

como queríamos. $\Diamond$

Como consecuencia directa de la fórmula 4.2.3 y el lema anterior obtenemos la siguiente fórmula para calcular el número de combinaciones m-arias de n elementos:

Fórmula 4.2.10

$$C_m^n = \frac{n!}{m!(n-m)!}$$

Aplicándola por ejemplo a nuestro problema introductorio, resulta que 10 puntos no alineados de a tres determinan

$$C_3^{10} = \frac{10!}{3!7!} = 120$$

triángulos.

NOTA. Notemos la diferencia esencial entre los conceptos de variación y combinación. El primero está referido a secuencias (elementos dados en un cierto orden), mientras que el segundo está ligado a la noción de subconjunto (una colección de elementos). Resulta así que C_m^n expresa *el número de subconjuntos de m elementos contenidos en un conjunto de n elementos.*

Observemos además que la fórmula 4.2.10 tiene sentido también en el caso $m = 0$, resultando que $C_0^n = 1$ cualquiera sea $n \geq 0$. Esto se corresponde perfectamente con la interpretación combinatoria de la fórmula, ya que todo conjunto admite un único subconjunto de 0 elementos (el subconjunto vacío).

Como un último apunte, veamos que similarmente a los casos de variaciones y permutaciones, también las combinaciones m-arias de n elementos están conectadas con una cierta clase de funciones, a saber, las funciones *estrictamente crecientes* de $\mathbb{I}_m$ en $\mathbb{I}_n$ (vale decir, funciones $f : \mathbb{I}_m \to \mathbb{I}_n$ tales que $f(i) < f(j)$ si $i < j$).

Designando por $E_c\,(m, n)$ el conjunto de dichas funciones, es obvio que toda $f \in E_c(m, n)$ es inyectiva, por lo que la misma determina un subconjunto de m elementos distintos de $\mathbb{I}_n$, a saber, $\{f(1), f(2), \ldots, f(m)\}$.

Recíprocamente, cualquier subconjunto $\{x_1, x_2, \ldots, x_m\}$ de m elementos de $\mathbb{I}_n$ (cuyos elementos podemos suponer escritos en orden creciente), corresponde a un elemento de $E_c(m,n)$, precisamente la función g definida por $g(i) = x_i$ para $i = 1, 2, \ldots, m$. Puesto que es trivial probar que estas asignaciones son mutuamente inversas, resulta que $E_c(m,n)$ es coordinable con la colección de subconjuntos de m elementos de $\mathbb{I}_n$, ó equivalentemente, con el conjunto de combinaciones m-arias de $\{1, 2, \ldots, n\}$.

Deducimos en particular que

$$\#\left(E_c(m,n)\right) = C_m^n\,,$$

esto es, existen C_m^n funciones estrictamente crecientes de $\mathbb{I}_m$ en $\mathbb{I}_n$. $\quad\Diamond$

Cerraremos este apartado resolviendo un par de problemas sobre el tema. Resultará claro en el primero que se trata de una cuestión de combinaciones, mientras que en el segundo el uso de las mismas es menos rutinario y hasta un tanto inesperado.

Ejemplo 4.2.11 Para efectuar un rescate una brigada de 13 socorristas decide separarse en grupos, uno de 4 personas y otros tres de 3 miembros cada uno. ¿De cuántas maneras pueden hacerlo si los dos coordinadores de la brigada deben estar en diferentes grupos?

Procederemos por descarte, contando en primer término el número de formas de distribuir 13 personas en 4 grupos como los del enunciado, sin tener en cuenta la restricción impuesta, y restando luego de la cantidad obtenida el número de casos en los cuales los coordinadores integran el mismo grupo.

Para formar los diferentes grupos debemos realizar selecciones no ordenadas de personas, ya que no se menciona ningún tipo de ordenamiento dentro de ellos. Se trata por lo tanto de un problema de combinaciones, resultando que hay C_4^{13} formas de constituir el grupo de 4, y luego, sucesivamente, C_3^9, C_3^6 y C_3^3 formas de integrar los tres grupos de 3 socorristas (obsérvese que $C_3^3 = 1$, lo cual expresa un hecho evidente: una vez armados 3 grupos hay una única manera de formar el último, ya que solo quedan 3 personas disponibles para ello).

Podríamos concluir entonces, usando el principio de multiplicación, que el número total de equipos que podemos formar es el producto de las cuatro cantidades anteriores. Sin embargo, estaríamos en ese caso cometiendo un error. En efecto, supongamos por ejemplo que una vez elegido un grupo H de 4 personas, hubiéramos formado grupos G_1, G_2 y G_3 de 3 personas cada uno, y consideremos ahora la secuencia de grupos H, G_3, G_1 y G_2. Si bien es claro que ambas secuencias producen la misma subdivisión en grupos, ésta sería contada dos veces, ya que en el PGM se presupone que las distintas elecciones se realizan en un orden determinado. Naturalmente, lo mismo ocurre si permutamos de cualquier manera los grupos G_i, de donde resulta que para contar correctamente debemos dividir el producto de las

citadas cantidades por el número de permutaciones de 3 objetos distintos. En definitiva, el número de equipos que pueden formarse es

$$\frac{C_4^{13}C_3^9C_3^6C_3^3}{3!} = \frac{715\,.\,84\,.\,20}{6} = 200200\,.$$

Para calcular el número de situaciones en las que los coordinadores forman parte de un mismo grupo G, dividimos la cuestión en dos casos, según sea G de 3 o de 4 personas. En el primero de ellos, las restantes 11 personas se distribuirán formando un grupo de 4, dos grupos de 3 y un grupo de un solo integrante (el que acompaña a los coordinadores). Razonando en forma muy similar a la anterior, resulta que dicha subdivisión puede hacerse de

$$\frac{C_4^{11}C_3^7C_3^4C_1^1}{2!} = \frac{330\,.\,35\,.\,4}{2} = 23100$$

maneras. Análogamente, en el segundo caso los otros 11 socorristas forman tres grupos de 3 personas y un grupo de 2, siendo

$$\frac{C_3^{11}C_3^8C_3^5C_2^2}{3!} = \frac{165\,.\,56\,.\,10}{6} = 15400$$

el número de formas de llevar a cabo una subdivisión de este tipo. Puesto que los dos casos que hemos considerado son mutuamente excluyentes, la respuesta final a nuestro problema es

$$200200 - (23100 + 15400) = 161700\,. \quad \Diamond$$

Ejemplo 4.2.12 En la grilla de la figura de abajo, ¿de cuántas maneras podemos trasladarnos de la casilla situada en el ángulo inferior izquierdo a la casilla situada en el ángulo superior derecho moviéndonos siempre o bien una casilla hacia la derecha o bien una casilla hacia arriba?

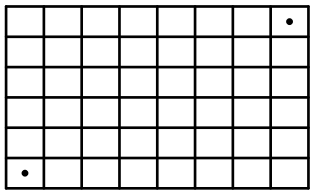


Figura 4.1: Grilla

Observemos que para realizar el recorrido es necesario desplazarse, en algún orden, 7 casillas hacia la derecha y 5 casillas hacia arriba. Si convenimos en simbolizar con la letra H cualquier movimiento horizontal y con la letra V cualquier movimiento vertical, resulta entonces que cada camino se representa unívocamente por una secuencia de 7 letras H y 5 letras V. Por ejemplo, la secuencia $HHHHHHHVVVVV$ corresponde a ir por la fila

de abajo hasta la última columna y luego siempre hacia arriba, mientras que el camino que consiste en moverse en "zigzag" hasta la fila de arriba está representado por la secuencia $HVHVHVHVHVHH$.

Ahora que hemos reducido la cuestión a la tarea de contar el número de tales secuencias, notemos que cualquiera de ellas queda completamente determinada decidiendo cuáles de las 12 posiciones ocuparán las 5 letras V, por lo que el número de las mismas coincide con la cantidad de formas de elegir 5 elementos distintos entre 12, siendo irrelevante el orden de elección, ya que se trata de posiciones. Por ejemplo, si elegimos las posiciones 2, 5, 6, 8 y 12 obtenemos la secuencia $HVHHVVHVHHHV$.

De acuerdo con nuestro análisis, resulta en definitiva que el número de caminos que unen las casillas señaladas con puntos es

$$C_5^{12} = 792\,. \quad \Diamond$$

Permutaciones con repeticion Si bien hasta aquí hemos empleado el término permutaciones para referirnos a los ordenamientos de una cierta cantidad de objetos distintos, nos proponemos ahora ampliar nuestro lenguaje y nuestros conocimientos, ya que ciertos problemas requieren manejar situaciones en las que los elementos a ordenar no son todos distintos, como podrían ser el contar la cantidad de formas de ordenar las cifras de un número o las letras de una palabra. En esa dirección, y a manera de ejemplo ilustrativo, calculemos el número de *anagramas* de la palabra $LAVARROPA$, vale decir, contemos el número de permutaciones de sus letras.

Puesto que entre éstas hay algunas repetidas no podemos aplicar la fórmula 4.2.6, pero podemos resolver el problema sin mayor dificultad aplicando el principio general de multiplicación, ya que cualquier anagrama puede obtenerse mediante la realización sucesiva de las siguientes tres experiencias: se elige primero −entre las 9 posiciones a llenar− las 3 que ocuparán las letras A, luego se elige entre las 6 posiciones restantes las 2 que ocuparán las letras R, y finalmente, en las 4 posiciones que aún falta cubrir se permutan de cualquier manera las letras L, V, O y P. Por lo tanto, el número de anagramas es

$$C_3^9\, C_2^6\, P_4 = \frac{9!}{3!\,6!} \cdot \frac{6!}{2!\,4!}\,.\,4! = \frac{9!}{3!\,2!} = 30240\,.$$

Visto el ejemplo anterior vayamos ahora a la situación general, en la que consideraremos secuencias de n objetos de r tipos distintos, habiendo t_k objetos idénticos entre sí de cada tipo k, donde $t_1, t_2, \ldots, t_r$ son números naturales y $n = t_1 + t_2 + \ldots + t_r$.

> Designaremos por $P_n(t_1, t_2, \cdots, t_r)$ el número de tales secuencias, a las que llamaremos *permutaciones con repetición* de n objetos con parámetros $t_1, t_2, \ldots, t_r$.

Por ejemplo, en el caso de los anagramas de $LAVARROPA$ tenemos $n = 9$, $r = 6$, $t_1 = 3$, $t_2 = 2$ y $t_i = 1$ para todo $i > 2$. Notemos asimismo que las secuencias de letras H y V del ejemplo 4.2.12 también son un caso particular de este tipo de arreglos, siendo $n = 12$, $r = 2$, $t_1 = 7$ y $t_2 = 5$.

En cuanto al cálculo de $P_n(t_1, t_2, \ldots, t_r)$ podemos razonar en forma completamente análoga a la del caso particular de los anagramas, resultando que hay $C_{t_1}^n$ formas de ubicar los objetos de tipo 1, por cada una de estas hay $C_{t_2}^{n-t_1}$ formas de ubicar los de tipo 2, por cada una de estas hay $C_{t_3}^{n-t_1-t_2}$ formas de ubicar los de tipo 3, y así siguiendo. Por lo tanto, aplicando el PGM obtenemos:

$$P_n(t_1, t_2, \ldots, t_r) = \prod_{k=1}^{r} C_{t_k}^{n-\sum_{i=1}^{k-1} t_i} .$$

Ahora bien, en la productoria del segundo miembro de la igualdad de arriba hay una gran cantidad de cancelaciones, ya que el producto de dos factores consecutivos de la misma es de la forma

$$C_b^a \, C_c^{a-b} = \frac{a!}{(a-b)!\, b!} \times \frac{(a-b)!}{(a-b-c)!\, c!} = \frac{a!}{(a-b-c)!\, b!\, c!} \quad (a, b, c \in \mathbb{N}).$$

Resulta por lo tanto que todos los numeradores a partir del segundo factor se cancelan, los que nos permite establecer la siguiente fórmula para el cálculo del número de permutaciones con repetición de n objetos con parámetros $t_1, t_2, \ldots, t_r$:

Fórmula 4.2.13

$$P_n(t_1, t_2, \ldots, t_r) = \frac{n!}{t_1!\, t_2! \cdots t_r!} .$$

La argumentación que precede a la fórmula anterior es solo un esbozo de la demostración de su validez. El lector interesado en una prueba más rigurosa puede obtener una bastante sencilla por inducción en r.

Vale la pena consignar por último que las permutaciones sin repetición configuran un caso particular de esta situación general, ya que ellas corresponden a los valores $r = n$ y $t_i = t_2 = \cdots = t_n = 1$. Es inmediato verificar que en tal caso las fórmulas 4.2.6 y 4.2.13 coinciden.

Ejemplo 4.2.14 El diseño de un cierto juego de ingenio consiste en un damero de 5×5, con 2 casillas blancas, 2 casillas negras y una casilla gris en cada fila. Si no puede haber dos filas iguales, ¿cuántos diseños distintos pueden lograrse?

Veamos en primer término de cuántas maneras puede diseñarse una fila. Simbolizando cada color con la letra adecuada, podemos pensar a cualquiera de ellas como una cierta secuencia de los símbolos b, b, n, n, g, esto es, como

una permutación con repetición de 5 objetos con parámetros $t_1 = t_2 = 2$ y $t_3 = 1$. De acuerdo con 4.2.13, el número de diseños posibles para una fila es entonces

$$P_5(2,2,1) = \frac{5!}{2!\,2!\,1!} = 30\,.$$

Finalmente, para calcular la cantidad de dameros distintos que pueden lograrse, observemos que las condiciones del problema exigen elegir ordenadamente 5 modelos distintos de fila entre los 30 existentes, por lo que el número de diseños coincide con el número de variaciones de 30 elementos tomados de a 5. En consecuencia, la respuesta al problema es

$$V_5^{30} = \frac{30!}{25!} = 17100720\,. \quad \Diamond$$

Números combinatorios.

Estudiaremos en este apartado algunas particularidades de los números designados por los símbolos C_k^n, que como vimos, expresan la cantidad de formas de seleccionar k elementos en un conjunto de n elementos. Los mismos, a los que por razones evidentes se denomina *números combinatorios*, tienen muchas propiedades interesantes y aparecen frecuentemente en diversas cuestiones del Algebra y la Matemática en general, como por ejemplo en la fórmula de Newton para el desarrollo de potencias binomiales.

Comencemos introduciendo una notación alternativa −muy familiar por otra parte− para dichos números, y que será la que utilizaremos en adelante. Precisamente, dados enteros k y n, con $0 \leq k \leq n$, escribiremos:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = C_k^n\,.$$

El símbolo $\binom{n}{k}$ se leerá *"número combinatorio n k"*, y simplemente para unificar la notación cuando ello convenga, extendemos su rango de definición estableciendo que $\binom{n}{k} = 0$ si $k < 0$ ó $k > n$. Observemos que la nueva definición es compatible con el significado combinatorio de los $\binom{n}{k}$. En la siguiente proposición exhibiremos algunas de sus propiedades básicas.

Proposición 4.2.15 Si $n \in \mathbb{N}_0$ y $k \in \mathbb{Z}$, valen las siguientes propiedades:

1) $\binom{n}{0} = \binom{n}{n} = 1$

2) $\binom{n}{1} = \binom{n}{n-1} = n$

3) $\binom{n}{k} = \binom{n}{n-k}$

4) $\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}\,.$

DEMOSTRACION Las propiedades 1) a 3) se obtienen por aplicación directa de la definición, por lo que obviamos los detalles. Remarquemos además el sentido combinatorio de 3): en un conjunto de n elementos hay tantos subconjuntos de k elementos como subconjuntos de $n-k$ elementos, ya que por cada elección de k elementos hay $n-k$ que se descartan.

En cuanto a 4), notemos que la fórmula brinda una forma recurrente de calcular los números combinatorios mediante sumas, sin necesidad de multiplicar, lo que obviamente representa una ventaja computacional. Observemos también que dicha *fórmula de adición* permite probar algebraicamente −por inducción en n− que los números combinatorios son números enteros, con prescindencia de su interpretación combinatoria. Su demostración se obtiene aplicando las definiciones y operando, por lo que también la dejamos a cargo del lector. $\Diamond$

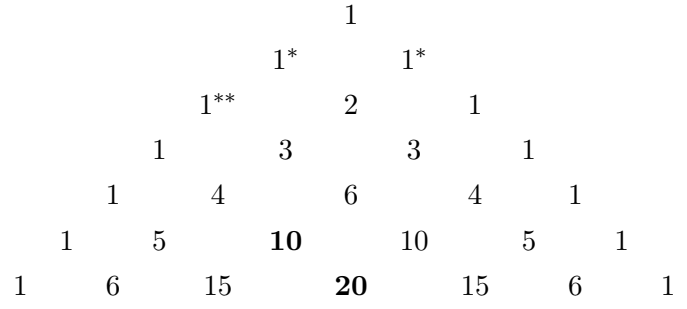
TRIANGULO DE PASCAL

Las propiedades que acabamos de enunciar pueden visualizarse a través de un esquema triangular, disponiendo los números combinatorios en sucesivas filas horizontales. La primera (fila 0) tiene como único término a $\binom{0}{0}$, la segunda (fila 1) a los combinatorios $\binom{1}{0}$ y $\binom{1}{1}$, y, en general, la fila n consiste de los números combinatorios de "nivel" n, a saber, $\binom{n}{0}, \binom{n}{1}, \ldots, \binom{n}{n}$. El matemático y filósofo francés Blaise Pascal (1623-1662) hizo un fecundo uso del mismo al sentar las bases de la teoría de la probabilidad, debido a lo cual se lo conoce con el nombre de *triángulo de Pascal*, aunque también se lo conoce como *triángulo de Tartaglia*.

Para su construcción basta ceñirse a las siguientes reglas, que reflejan las propiedades demostradas en 4.2.15:

- T_1 La fila n tiene $n+1$ términos.
- T_2 Todas las filas comienzan y terminan en 1 (propiedad 1)).
- T_3 Cada término interior de una fila es la suma de los dos términos de la fila anterior ubicados inmediatamente por encima de él (propiedad 4)).

Ilustremos la situación mostrando las primeras 7 filas del triángulo de Pascal (idealmente infinito), o sea, hasta los combinatorios de nivel 6:



Vemos así por ejemplo (elementos en negrita) que $\binom{5}{2} = 10$ y $\binom{6}{3} = 20$. Notemos además que cada fila se lee igual de izquierda a derecha que de derecha a izquierda, lo cual es consecuencia de la propiedad 3).

En realidad hay una gran cantidad de simetrías y regularidades en el triángulo de Pascal, algunas detectables a simple vista y otras que requieren un análisis más minucioso. Entre las primeras observemos las dos diagonales idénticas que parten de los números señalados con un asterisco. En ambos casos se va obteniendo la sucesión de números naturales, hecho que corresponde a las igualdades $\binom{n}{1} = n = \binom{n}{n-1}$. Asimismo, es posible que el lector reconozca la secuencia de números situados en la diagonal que comienza en el término destacado con doble asterisco. En tal caso, lo invitamos a explicar la razón de su presencia en el triángulo.

FORMULA DE NEWTON Seguramente el lector conoce las fórmulas

$$(a+b)^2 = a^2 + 2ab + b^2,$$
$$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3,$$

correspondientes al desarrollo del cuadrado y del cubo de una suma (*binomio*) de números reales a y b. Las mismas se obtienen por simple aplicación de la propiedad distributiva, y podemos observar en ellas varias similitudes. En efecto, ambas consisten de una serie de términos (monomios), cada uno de los cuales es el producto de una potencia de a por una potencia de b, multiplicado a su vez por un coeficiente numérico. En todos ellos la suma de los exponentes es constante −dos en el primer caso y tres en el segundo−, siendo las potencias de a decrecientes y las de b crecientes. Las secuencias de coeficientes son $1, 2, 1$ y $1, 3, 3, 1$, respectivamente, aunque la existencia de un patrón de formación en ambas situaciones se hace más evidente si observamos que los elementos de la primera secuencia son $\binom{2}{0}, \binom{2}{1}, \binom{2}{2}$ y los de la segunda $\binom{3}{0}, \binom{3}{1}, \binom{3}{2}, \binom{3}{3}$.

Vistos estos ejemplos, vayamos ahora al caso general de la expansión de

$$(a+b)^n = \underbrace{(a+b)(a+b)\dots(a+b)}_{n \text{ factores}},$$

donde $a, b \in \mathbb{R}$ y $n \in \mathbb{N}$. Mostraremos que dicha expansión responde a una fórmula esencialmente idéntica a las obtenidas en los casos $n = 2$ y $n = 3$, la célebre *fórmula binomial de Newton*. Su enunciado preciso es el siguiente:

Teorema 4.2.16 (Fórmula binomial de Newton)

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^{n-i} b^i \ .$$

DEMOSTRACION Procediendo por inducción en n ($n \geq 0$), observemos que el resultado es inmediato si $n = 0$, ya que en tal caso ambos miembros de la fórmula son iguales a 1.

Supongamos ahora que la fórmula es válida para un cierto $h \in \mathbb{N}_0$. Aplicando la hipótesis inductiva, y distribuyendo convenientemente obtenemos entonces:

$$\begin{aligned}
(a+b)^{h+1} = (a+b)(a+b)^h = (a+b)\sum_{k=0}^{h}\binom{h}{k}a^{h-k}b^k = \\
= \sum_{k=0}^{h}\binom{h}{k}a^{h+1-k}b^k + \sum_{k=0}^{h}\binom{h}{k}a^{h-k}b^{k+1} = \\
= \sum_{i=0}^{h}\binom{h}{i}a^{h+1-i}b^i + \sum_{i=1}^{h+1}\binom{h}{i-1}a^{h+1-i}b^i = \\
= \sum_{i=1}^{h}\left(\binom{h}{i} + \binom{h}{i-1}\right)a^{h+1-i}b^i + a^{h+1} + b^{h+1} = \\
= \sum_{i=1}^{h}\binom{h+1}{i}a^{h+1-i}b^i + a^{h+1} + b^{h+1} = \\
= \sum_{i=0}^{h+1}\binom{h+1}{i}a^{h+1-i}b^i\,,
\end{aligned}$$

esto es, la fórmula es válida para el exponente $h+1$, y por lo tanto es válida para todo $n \in \mathbb{N}_0$, como queríamos demostrar. Vale la pena aclarar un poco algunos de los pasos anteriores: en la segunda sumatoria de la tercera línea hemos efectuado el cambio de índices $k+1 = i$, mientras que en la primera simplemente cambiamos el nombre del índice de suma; en la cuarta línea escribimos separadamente los términos a^{h+1} y b^{h+1}, correspondientes al valor $i = 0$ en la primera suma y al valor $i = h+1$ en la segunda, de manera de poder agrupar las sumatorias restantes y aplicar finalmente en la penúltima línea la fórmula de adición de los números combinatorios. $\Diamond$

Nota. Se impone hacer algunos comentarios. Puesto que la suma es conmutativa, es claro que en el segundo miembro de la fórmula de Newton los roles de a y b deben ser intercambiables. Para justificar dicha simetría basta efectuar el cambio de índices $j = n - i$, ya que entonces resulta

$$(a+b)^n = \sum_{i=0}^{n}\binom{n}{i}a^{n-i}b^i = \sum_{j=0}^{n}\binom{n}{n-j}a^jb^{n-j} = \sum_{j=0}^{n}\binom{n}{j}a^jb^{n-j}\,.$$

Observemos que la primera y la última suma de arriba recorren los mismos términos, pero en orden inverso.

La fórmula de Newton puede demostrarse algo más informalmente mediante argumentos combinatorios. En efecto, aplicando reiteradamente la propiedad distributiva resulta que un término genérico de la expansión de

$$(a+b)(a+b)\ldots(a+b)$$

consiste del producto de n factores, cada uno de los cuales es una a ó una b, dando como resultado un monomio del tipo $a^{n-i}b^i$ $(0 \leq i \leq n)$. Por otro lado, obtendremos tal monomio cada vez que seleccionemos i veces el factor b (y por descarte $n-i$ veces el factor a) entre los n factores disponibles. Como sabemos, existen $\binom{n}{i}$ formas de realizar dicha elección, y es debido a ello que el monomio $a^{n-i}b^i$ aparece $\binom{n}{i}$ veces en el desarrollo de $(a+b)^n$. Digamos de paso que a raíz de su presencia en la fórmula de Newton los números combinatorios son también llamados *coeficientes binomiales*.

La fórmula binomial puede aplicarse también a restas, teniendo en cuenta que $a - b = a + (-b)$. Se obtiene de esa manera la fórmula

$$(a-b)^n = \sum_{i=0}^{n} \binom{n}{i} (-1)^i a^{n-i} b^i. \quad \Diamond$$

Ejemplos 4.2.17 Veamos otro par de ejemplos de relación fructífera entre el álgebra y la combinatoria. Observando nuevamente las primeras 7 filas del triángulo de Pascal, verificamos inmediatamente que al sumar los elementos de cualquiera de sus filas obtenemos una potencia de 2, cuyo exponente coincide con el nivel de la fila. Por ejemplo, la suma de los elementos de la última fila (combinatorios de nivel 6) es:

$$1 + 6 + 15 + 20 + 15 + 6 + 1 = 64 = 2^6.$$

Como es de imaginar, estas coincidencias no son casuales. Se trata de una situación general que admite la siguiente y sencilla prueba:

$$2^n = (1+1)^n = \sum_{k=0}^{n} \binom{n}{k} 1^{n-k} 1^k = \sum_{k=0}^{n} \binom{n}{k},$$

cualquiera sea $n \geq 0$.

El hecho de que la suma de los combinatorios de nivel n sea igual a 2^n, que acabamos de demostrar algebraicamente, tiene su correlato combinatorio en la propiedad que afirma que un conjunto A de n elementos tiene exactamente 2^n subconjuntos, ya mencionada anteriormente. En efecto, puesto que para cada k el número combinatorio $\binom{n}{k}$ cuenta el número de subconjuntos de k elementos de A, al sumar sobre k estamos contando *todos* los subconjuntos de A.

Asimismo, el lector podrá verificar que la suma alternada de los elementos de cada fila (a partir de la segunda) es nula, como por ejemplo

$$0 = 1 - 5 + 10 - 10 + 5 - 1$$

en la sexta. Para demostrar esta afirmación en general basta aplicar la fórmula de la resta al caso $a = b = 1$, ya que entonces

$$0 = (1-1)^n = \sum_{k=0}^{n} (-1)^k \binom{n}{k}, \tag{4.8}$$

para todo $n \in \mathbb{N}$.

Esta última igualdad tiene también una interpretación combinatoria digna de mención, ya que escribiéndola en la forma equivalente

$$\sum_{k \text{ par}} \binom{n}{k} = \sum_{k \text{ impar}} \binom{n}{k} \quad (0 \leq k \leq n)$$

concluimos que todo conjunto finito no vacío *admite tantos subconjuntos de cardinal par como subconjuntos de cardinal impar.* $\Diamond$

Ejemplo 4.2.18 Como otra aplicación, usaremos 4.8 para delinear una demostración del teorema 4.1.7, referido al cardinal de la unión de una familia $A_1, A_2, \ldots, A_n$ de conjuntos finitos. Podemos razonar de la siguiente forma: los elementos de la intersección $A_r \cap A_s$ de dos cualesquiera de los conjuntos se cuentan 2 veces en la suma $\sum_i \#(A_i)$, luego deben ser descontados una vez. Análogamente, los elementos de cualquiera de las intersecciones $A_r \cap A_s \cap A_t$ de tres de los conjuntos se cuentan 3 veces en la suma $\sum_i \#(A_i)$ y se descuentan también 3 veces, a saber, en $\#(A_r \cap A_s)$, $\#(A_r \cap A_t)$ y $\#(A_s \cap A_t)$. Por lo tanto, deben ser agregados una vez.

Siguiendo con este argumento y procediendo por recurrencia, sea $k \leq n$ y consideremos los elementos pertenecientes a k de los conjuntos (digamos $A_1, \ldots, A_k$, para simplificar la notación). Los mismos se cuentan k veces en $\sum_i \#(A_i)$, se descuentan $\binom{k}{2}$ veces en $\sum_{1 \leq i < j \leq k} \#(A_i \cap A_j)$, y en general, dado $l < k$, se cuentan o descuentan $\binom{k}{l}$ veces (según que l sea impar o par) en la suma de todos los términos de la forma $\#(A_{i_1} \cap A_{i_2} \cap \ldots \cap Ai_l)$, tomada sobre todos los índices $i_1, \ldots, i_l$ tales que $1 \leq i_1 < i_2 < \ldots < i_l \leq k$. Por lo tanto, el número (quizás negativo) de veces que dichos elementos han sido considerados es

$$\begin{aligned} \sum_{l=1}^{k-1} (-1)^{l-1} \binom{k}{l} &= -\left(\sum_{l=1}^{k} (-1)^l \binom{k}{l} \right) = -\left(\sum_{l=0}^{k} (-1)^l \binom{k}{l} - 1 - (-1)^k \right) \\ &= 1 + (-1)^k \,, \end{aligned}$$

vale decir, hasta aquí los elementos en cuestión han sido contados 2 veces si k es par y ninguna vez si k es impar, y por lo tanto debemos descontarlos o agregarlos una vez, según el caso. Englobando ambas situaciones, y teniendo en cuenta que el análisis anterior es válido cualquiera sea la elección de k miembros de $\{A_1, A_2, \ldots, A_n\}$, resulta que para ajustar la cuenta debe agregarse el término

$$(-1)^{k-1} \sum_{(i_1, i_2, \ldots, i_k)} \#(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) \,,$$

como queríamos demostrar. Naturalmente, la suma varía sobre todas las k-uplas $(i_1, i_2, \ldots, i_k)$ tales que $1 \leq i_1 < i_2 < \ldots < i_k \leq n$. $\Diamond$

FORMULA DE LEIBNIZ La fórmula de Newton se generaliza al caso del desarrollo de expresiones del tipo $(a_1+a_2+\ldots+a_m)^n$, donde los a_i son números reales y m y n son números naturales.

En efecto, escribiendo como antes

$$(a_1+a_2+\ldots+a_m)^n = \underbrace{(a_1+a_2+\ldots+a_m)\ldots\ldots(a_1+a_2+\ldots+a_m)}_{n \text{ factores}},$$

y razonando de manera combinatoria, resulta que un término genérico de la expansión del producto del miembro de la derecha es de la forma

$$a_1^{k_1} a_2^{k_2} \ldots a_m^{k_m},$$

donde los exponentes k_i son enteros no negativos cuya suma es n. Para completar el panorama, debemos averiguar cuántas veces aparece cada uno de estos monomios en el desarrollo, lo que nos indicará el coeficiente por el que debe ser multiplicado. A tal efecto, observemos que cada una de sus apariciones corresponde a una permutación con repetición de los símbolos $a_1, a_2, \cdots, a_m$, cada a_i repetido k_i veces, de donde concluimos que dicho coeficiente es $P_n(k_1,\ldots,k_m)$.

Obtenemos así la llamada *fórmula multinomial de Leibniz*, cuyo enunciado preciso es el siguiente:

Fórmula 4.2.19 (Fórmula multinomial de Leibniz)

$$(a_1+a_2+\ldots+a_m)^n = \sum_{k_1+k_2+\ldots+k_m=n} \frac{n!}{k_1!k_2!\ldots k_m!}\; a_1^{k_1} a_2^{k_2} \ldots a_m^{k_m}\;,$$

donde la notación indica que la suma varía sobre las m-uplas $(k_1,k_2,\ldots,k_m)$ de enteros no negativos tales que $\sum_i k_i = n$. Se comprueba fácilmente que para $m=2$ la fórmula de Leibniz coincide con la fórmula binomial de Newton. ◊

Ejemplo 4.2.20 En el caso $n=2$ la fórmula adopta la forma

$$(a_1+a_2+\ldots+a_m)^2 = \sum_{r=1}^{m} a_r^2 + 2\sum_{r<s} a_r\, a_s\,,$$

ya que las únicas m-uplas de enteros no negativos que satisfacen la condición $\sum_i k_i = 2$ son aquellas en las que $k_r = 2$ para algún índice r y $k_i = 0$ para todo $i \neq r$ y aquellas en las que $k_r = k_s = 1$ para cierto par de índices distintos r y s y $k_i = 0$ para todo otro índice.

Similarmente, encargamos al lector verificar que para $n=3$ se tiene

$$(a_1+a_2+\cdots+a_m)^3 = \sum_{r=1}^{m} a_r^3 + 3\sum_{r\neq s} a_r\, a_s^2 + 6\sum_{r<s<t} a_r\, a_s\, a_t\,. \quad ◊$$

Distribuciones. Combinaciones con repetición.

Veamos cómo resolver un problema del siguiente tipo: un ascensor que contiene 12 bultos iguales ascenderá 5 pisos. ¿De cuántas maneras podrán ser descargados los bultos?

La dificultad del problema reside en el hecho de que los bultos sean idénticos, ya que si fueran distintos bastaría contar el número de funciones de un conjunto de 12 elementos en otro de 5, lo que nos daría 5^{12} formas de descargarlos. No siendo ese el caso, debemos examinar un poco la cuestión.

Observemos para ello que cada posible distribución está asociada a una forma de descomponer a 12 como suma de 5 enteros no negativos, aunque una tal descomposición corresponderá en general a muchas distribuciones distintas. Por ejemplo, la partición $12 = 4 + 3 + 3 + 2 + 0$ corresponde a descargar 4 bultos en alguno de los 5 pisos, 3 bultos en cada uno de otro par de pisos, y los 2 últimos bultos en alguno de los 2 pisos restantes. Una sencilla aplicación del principio de multiplicación nos muestra que el número de formas de realizar dicha selección de pisos es $5 \,.\, C_4^2 \,.\, 2 = 60$, esto es, la partición $12 = 4 + 3 + 3 + 2 + 0$ da lugar a 60 distribuciones distintas.

Parecería entonces que debemos contar el número de particiones de 12 como suma de 5 términos y multiplicar luego por la cantidad de distribuciones asociadas a cada una de ellas, pero desafortunadamente esta empresa no es viable. En primer lugar porque no existe una forma razonable de contar el número de descomposiciones de un número natural r como suma de n enteros no negativos (son 47 en el caso $r = 12$ y $n = 5$), y por otro lado porque no existe regularidad, en el sentido de que la cantidad de distribuciones que corresponden a cada partición depende de la forma de ésta. Por ejemplo, el lector puede comprobar que la partición $12 = 3 + 3 + 3 + 2 + 1$ genera 20 distribuciones.

Exhibidas las dificultades, resolveremos ahora en general el problema de *distribuir* r objetos *indistinguibles* (los bultos en el ejemplo anterior) en lugares o categorías *distinguibles* (los pisos). Esquematizaremos la situación pensando que debemos distribuir r bolillas idénticas en n cajas distintas, designando por D_r^n el número de dichas distribuciones.

Puesto que las cajas son distinguibles las numeraremos de 1 a n. Acordaremos en designar cada caja con una barra y cada bolilla con un asterisco, y representaremos cada distribución con una secuencia de n símbolos $|$ y r símbolos $*$, conviniendo en que el número de asteriscos situados a la izquierda de cada barra indica cuántas bolillas se alojan en la correspondiente caja. Por ejemplo, si $r = 6$ y $n = 4$, la distribución que consiste en alojar 2 bolillas en las cajas 1 y 3 y una bolilla en las cajas 2 y 4 se representa mediante la secuencia

$$* \ * \ | \ * \ | \ * \ * \ | \ * \ |,$$

mientras que la secuencia

$$* \ * \ * \ | \ | \ * \ | \ * \ * \ |$$

está asociada a la distribución que deja vacía la caja 2 y aloja 3, 1 y 2 bolillas en las cajas 1, 3 y 4 respectivamente.

En principio, debemos contar entonces el número de tales secuencias, pero teniendo en cuenta que por nuestra convención el último símbolo de cualquier sucesión siempre resultará una |, podemos olvidarnos de él. Por lo tanto, las distribuciones de r bolillas indistinguibles en n cajas distinguibles están en correspondencia biunívoca con las permutaciones con repetición de $n-1$ símbolos | y r símbolos $*$, de lo que deducimos la fórmula

Fórmula 4.2.21

$$D_r^n = P_{n+r-1}(n-1, r) = \binom{n+r-1}{r}$$

Así, en el caso de nuestro problema introductorio la solución es

$$D_{12}^5 = \binom{16}{12} = 1820\,.$$

Ejemplo 4.2.22 Una promotora de ventas debe obsequiar 25 muestras idénticas de un producto a 10 personas, con la condición de que todas reciban al menos 2 muestras. ¿De cuántas formas puede hacerlo?

El problema responde a la situación general de distribuir 25 bolillas indistinguibles (las muestras) en 10 cajas distinguibles (las personas), pero bajo la restricción de que todas las cajas alojen por lo menos 2 bolillas. Podemos manejar fácilmente esta variante, colocando de entrada 2 bolillas en cada una de las 10 cajas (20 en total) y distribuyendo luego de cualquier manera las 5 bolillas restantes. Es claro que cada una de estas distribuciones corresponden a una, y a solo una, de las del tipo buscado, resultando que el número de formas de repartir las muestras es

$$D_5^{10} = \binom{14}{5} = 2002\,. \quad \Diamond$$

Hemos usado hasta aquí el término combinación para referirnos a selecciones no ordenadas de objetos distintos, pero no es difícil imaginar situaciones en las cuales no se requiera que los objetos sean distintos. Por ejemplo, supongamos que queremos determinar el número de formas de seleccionar 3 vocales, sin que importe el orden de las mismas. La siguiente es la lista exhaustiva de tales arreglos, que resultan ser 35:

aaa	*aae*	*aai*	*aao*	*aau*	*aee*	*aei*
aeo	*aeu*	*aii*	*aio*	*aiu*	*aoo*	*aou*
auu	*eee*	*eei*	*eeo*	*eeu*	*eii*	*eio*
eiu	*eoo*	*eou*	*euu*	*iii*	*iio*	*iiu*
ioo	*iou*	*iuu*	*ooo*	*oou*	*ouu*	*uuu*

Visto el ejemplo anterior, y en la búsqueda de una fórmula que nos permita calcular el número de resultados sin necesidad de listarlos, pasemos a describir la situación general a estudiar:

> Dados objetos pertenecientes a n clases distintas (los objetos de cada clase indistinguibles entre sí), cualquier elección no ordenada de m de ellos (distintos o no) se denominará una *combinación con repetición* de los mismos. Designaremos por $(CR)_m^n$ el número de dichos arreglos.

Por ejemplo, en el caso de las vocales tenemos $n = 5$ y $m = 3$, resultando por tanto que $(CR)_3^5 = 35$. Ahora bien, veremos que los arreglos que nos ocupan pueden interpretarse como distribuciones, por lo que no es difícil hallar una fórmula que resuelva el caso general. En efecto, puesto que en una combinación con repetición lo único que interesa es *cuántos* objetos de cada clase van a seleccionarse, podemos pensar que debemos distribuir m unidades en n cajas (una por cada clase de objeto). Por ejemplo, la combinación de vocales *aoo* corresponde a colocar una unidad en la "caja" de las letras a, dos unidades en la de las letras o y ninguna en las cajas correspondientes a las letras e, i y u. Siendo inmediato que a cada combinación le corresponde una y solo una de tales distribuciones, obtenemos el siguiente resultado:

Fórmula 4.2.23

$$(CR)_m^n = D_m^n = \binom{m+n-1}{m}$$

Así, aplicando la fórmula al problema anterior resulta que el número de arreglos no ordenados de 3 vocales es $\binom{7}{3} = 35$. Vayamos a otro ejemplo, en el que nos encontraremos con una variante interesante.

Ejemplo 4.2.24 Un florista prepara para la venta ramos de 12 rosas, pudiendo éstas ser rojas, amarillas, blancas o rosadas. Si nunca coloca en un ramo más de 4 rosas amarillas, ¿cuántos ramos distintos puede ofrecer?

Podemos pensar a cada ramo como una combinación con repetición de 12 objetos pertenecientes a 4 clases distintas, por lo que en principio podríamos intentar aplicar la fórmula 4.2.23. Sin embargo, la restricción impuesta nos enfrenta con una situación novedosa, ya que en el planteo general se ha supuesto tácitamente que se dispone de una cantidad ilimitada de objetos de cada clase. Manejaremos dicha restricción trabajando por descarte, restando del total de distribuciones aquellas que alojan 5 o más unidades en la "caja" de las rosas amarillas. Para calcular esta última cantidad colocamos 5 unidades en dicha caja, y luego repartimos de cualquier manera las 7 unidades restantes en las 4 cajas, resultando entonces que el número de ramos que contienen a lo sumo 4 rosas amarillas es

$$(CR)_{12}^4 - (CR)_7^4 = D_{12}^4 - D_7^4 = \binom{15}{12} - \binom{10}{7} = 335\,. \quad \Diamond$$

PRINCIPIO DE INCLUSION Y EXCLUSION. En muchas situaciones de conteo suele ser más provechoso proceder por descarte, restando de la cantidad total de casos −que suponemos conocida−, el número de aquellos que no satisfacen ciertas condiciones dadas. Esta tarea requiere algún cuidado, por ejemplo para no excluir dos veces un mismo caso, y es necesario para su recta ejecución el uso del corolario 4.1.8, referido al cardinal del complemento de una unión de conjuntos. En combinatoria, una versión útil del mismo, llamada *principio de inclusión y exclusión*, es la siguiente:

Sea S un conjunto finito y sean $P_1, P_2, \ldots, P_n$ ciertas propiedades aplicables a los elementos de S. Designando por A_i el conjunto de elementos de S que verifican la propiedad $P_i\,(1 \leq i \leq n)$, resulta que el número c de elementos de S que *no verifican ninguna* de las propiedades P_i es

$$c = \#(S) + \sum_{k=1}^{n} (-1)^k \sum_{1 \leq i_1 < \ldots < i_k \leq n} \#\left(A_{i_1} \cap \ldots \cap A_{i_k}\right),$$

donde la notación indica que la sumatoria del k-ésimo término varía sobre todas las formas posibles de elegir k índices distintos entre 1 y n.

NOTA Es frecuente en las aplicaciones que para cada k todas las intersecciones $A_{i_1} \cap \cdots \cap A_{i_k}$ tengan el mismo cardinal, digamos c_k, en cuyo caso la fórmula anterior se simplifica gratamente. En efecto, teniendo en cuenta que existen $\binom{n}{k}$ formas de elegir k índices distintos entre 1 y n, resulta que

$$c = \#(S) + \sum_{k=1}^{n} (-1)^k \binom{n}{k} c_k \,. \tag{4.9}$$

La resolución de los siguientes problemas ilustrará el uso del principio.

Ejemplo 4.2.25 Sea S_n el conjunto de permutaciones de los elementos de $\mathbb{I}_n$. ¿Cuántos elementos de S_n no dejan ningún número en su posición natural?

Designaremos por D_n el número de tales permutaciones, que denominaremos *desarreglos*. Para fijar las ideas, el lector puede comprobar que hay exactamente 9 desarreglos de los elementos de $\mathbb{I}_4$, a saber:

2143	2341	2413
3142	3412	3421
4123	4312	4321

Para tratar la cuestión en general, dado cualquier $i \in \mathbb{I}_n$ y dada cualquier permutación $\pi = x_1 x_2 \ldots x_n$ en S_n, diremos que π satisface la propiedad P_i si y solo si $x_i = i$. Debemos por lo tanto contar cuántas permutaciones no satisfacen ninguna de las propiedades P_i, por lo que resulta natural usar el principio de inclusión y exclusión. Se trata además del caso simple que

comentamos arrriba, ya que dados índices $1 \leq i_1 < \ldots < i_k \leq n\,(1 \leq k \leq n)$, es claro que el número de elementos de S_n que satisfacen las k propiedades P_{i_j} es $(n-k)!$, y por lo tanto no depende de los índices elegidos sino solamente de k. En consecuencia, aplicando 4.9 obtenemos:

$$D_n = n! + \sum_{k=1}^{n} (-1)^k \binom{n}{k} (n-k)! = \sum_{k=0}^{n} (-1)^k \, \frac{n!}{k!} \ . \tag{4.10}$$

Por ejemplo, si $n = 5$ resulta

$$D_5 = 120 - 120 + 60 - 20 + 5 - 1 = 44\,.$$

Para redondear el tema, supongamos que nos interese estimar el porcentaje de desarreglos, el decir, la razón $D_n/n!$ entre la cantidad de los mismos y el número total de permutaciones. De acuerdo con la fórmula obtenida resulta

$$\frac{D_n}{n!} = \sum_{k=0}^{n} \frac{(-1)^k}{k!} \ .$$

Ahora bien, los lectores con conocimientos básicos de análisis matemático posiblemente reconozcan la expresión del miembro de la derecha, ya que se trata de la n-ésima suma parcial del desarrollo en serie de e^{-1}, donde $e = 2,71828\ldots$ es la base de los logaritmos neperianos. En términos de convergencia resulta entonces que

$$\lim_{n\to\infty} \frac{D_n}{n!} = e^{-1} = 0,36789\ldots,$$

hecho que admite la siguiente interpretación: si n es grande, aproximadamente el 36 % de los elementos de S_n no fijan ningún elemento de $\mathbb{I}_n$. $\quad \Diamond$

Ejemplo 4.2.26 Determinemos la cantidad de números naturales menores que un millón cuyas cifras suman 35. Observando que un número menor que 10^6 tiene a lo sumo 6 cifras, podemos unificar todos los casos representándolos en la forma $c_1c_2c_3c_4c_5c_6$, donde $0 \leq c_i \leq 9$ para todo i. Por ejemplo, dos números que satisfacen la condición requerida son 9989 y 91799, que representamos por 009989 y 091799, respectivamente.

Tenemos que contar entonces el número de soluciones de la ecuación

$$c_1 + c_2 + c_3 + c_4 + c_5 + c_6 = 35,$$

que es en realidad un problema de distribuciones, ya que podemos pensar que debemos repartir 35 unidades (bolillas) entre las 6 variables c_i (cajas). Aplicando la correspondiente fórmula, resulta entonces que el número de soluciones en enteros no negativos es $D_{35}^6 = C_{35}^{40}$.

Puesto que el problema tiene restricciones (las cajas deben contener a lo sumo 9 bolillas) usamos el principio de inclusión y exclusión, designando

por P_i la propiedad "la i-ésima caja contiene 10 o más bolillas". Habrá entonces que sumar o restar los casos de distribuciones que cumplen una cierta cantidad k de dichas propiedades, observando que solo debemos considerar los casos $k = 1, 2$ ó 3, ya que al repartir 35 bolillas es imposible que haya 10 o más bolillas en 4 cajas distintas. Por otra parte, es claro que al analizar los diversos casos es irrelevante cuáles son las cajas consideradas, por lo que nos encontramos nuevamente con la versión simplificada del principio.

Para contar genéricamente el número de distribuciones en las cuales k cajas determinadas alojan 10 o más bolillas, colocamos de entrada 10 bolillas en cada una de ellas, y distribuimos luego de cualquier manera las $35 - 10k$ restantes en las 6 cajas, lo que arroja el resultado

$$D^6_{35-10k} = \binom{40 - 10k}{35 - 10k}.$$

Finalmente, aplicando 4.9 sigue que la cantidad x buscada es

$$x = \binom{40}{35} - \binom{6}{1}\binom{30}{25} + \binom{6}{2}\binom{20}{15} - \binom{6}{3}\binom{10}{5} = 30492. \quad \Diamond$$

4.2.2. Ejercicios

1. El código de patente de un automóvil consiste de una secuencia de 3 letras seguida de una secuencia de 3 dígitos. Determinar el número de patentes distintas que pueden adjudicarse. ¿En cuántas patentes la secuencia de letras contiene al menos 2 vocales y la secuencia de números es capicúa?

2. Un mensaje telegráfico consiste de una secuencia de puntos y rayas ¿Cuántos mensajes distintos de hasta 20 símbolos pueden enviarse?

3. ¿Cuántos números naturales menores o iguales que 100000 no contienen ningún par de cifras consecutivas iguales?

4. Consideremos el triángulo de Pascal hasta los números combinatorios de nivel $n\,(n \in \mathbb{N})$. Por movimientos en diagonal, ¿cuántos caminos de n tramos conducen del vértice a la base?

5. Un código para cifrar mensajes asigna a cada una de las 27 letras del alfabeto un número distinto entre 1 y 28, empleándose el número restante para representar el espacio entre dos palabras. ¿Cuántos códigos distintos pueden confeccionarse? ¿Y si las vocales deben corresponder a múltiplos de 5?

6. Una comisión directiva de 10 miembros sesiona alrededor de una mesa redonda. ¿De cuántas maneras pueden ubicarse si entre el presidente y el secretario debe haber exactamente 2 personas?

7. De un grupo formado por 8 hombres y 5 mujeres hay que elegir una comisión de 6 personas, que deberá incluir al menos 3 mujeres y al menos 2 hombres. Determinar el número de formas de seleccionarlas, teniendo en cuenta que Alejandro y Verónica no quieren estar juntos en la comisión.

8. Dados 20 números enteros consecutivos, ¿de cuántas formas pueden elegirse 3 de ellos de manera que la suma de éstos sea par?

9. Doce personas jugarán un campeonato de truco por parejas. ¿De cuántas maneras pueden integrarse los equipos?

10. Sean L y L' dos rectas paralelas y distintas del plano, y sean $A \subset L$ y $A' \subset L'$ conjuntos de 15 y 20 puntos, respectivamente. Calcular el número de triángulos determinados por los elementos de $A \cup A'$.

11. Determinar el número de anagramas de la palabra combinación que no contienen ningún par de vocales seguidas.

12. Una persona adquiere en un gimnasio un abono para efectuar 10 sesiones durante 10 días distintos del mes de junio. ¿De cuántas maneras puede programar su actividad, si desea tener por lo menos un día de descanso luego de cada sesión?

13. ¿En cuántas permutaciones de la palabra probabilidad las consonantes aparecen en orden alfabético? ¿ Cuántas empiezan o terminan en i?

14. Probar la validez de las siguientes fórmulas (las letras designan enteros no negativos):

    a) $m^n = \displaystyle\sum_{k_1+k_2+\cdots+k_m=n} \frac{n!}{k_1!\,k_2!\cdots k_m!}$

    b) $\displaystyle\sum_{k=r}^{n} \binom{k}{r} = \binom{n+1}{r+1}$

    c) $\displaystyle\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$

    d) $\displaystyle\binom{m+n}{k} = \sum_{j=0}^{k} \binom{m}{j}\binom{n}{k-j}$

e) $\binom{2n}{n} = \sum_{j=0}^{n} \binom{n}{j}^2$.

15. Una editorial donará a una biblioteca pública 20 ejemplares de 5 libros diferentes. ¿De cuántas maneras puede hacerlo, si la institución debe recibir al menos un ejemplar de cada texto y en ningún caso más de 7 de ellos?

16. Para rendir un examen 7 estudiantes se sientan en una fila de 20 asientos. ¿De cuántas maneras pueden ubicarse si entre dos cualesquiera de ellos debe quedar por lo menos un asiento libre?

17. Una panadería elabora 15 clases de facturas. ¿De cuántas maneras puede elegirse una docena?

18. Calcular el número de funciones crecientes de I_m en I_n (f se dice creciente si $x < y \Rightarrow f(x) \leq f(y)$).

19. Calcular el número de resultados distintos que pueden obtenerse al arrojar simultáneamente 5 dados idénticos.

20. Calcular el número de términos del desarrollo de $(x+y+z)^{20}$. Determinar en cuántos de ellos no aparece x.

21. Calcular la cantidad de números naturales menores que 10000 cuyas cifras suman 24.

22. Al arrojar 8 veces un dado se han obtenido los 6 resultados posibles, no habiendo salido 2 veces seguidas un mismo número. Calcular cúantas secuencias de resultados pueden haberse registrado.

23. *a*) Calcular el número de desarreglos de 1234567 en los que el 1 ocupa una posición par.

    b) Calcular el número de desarreglos de 123456 en los que el 1 y el 2 ocupan posiciones contiguas.

24. Sean A y B conjuntos finitos de m y n elementos, respectivamente. Determinar una fórmula para el número de funciones suryectivas de A en B.

25. *a*) Un joven subirá por una escalera de 16 escalones, ascendiendo 1 ó 2 peldaños por vez. ¿De cuántas maneras puede hacerlo?

b) ¿ Y bajo la restricción de que en el tercer movimiento debe saltar 2 escalones ?

26. Si se listan los números naturales hasta 10000, ¿ cuántas veces se escribe la cifra 8 ?

4.3. Nociones de Probabilidad

4.3.1. Introducción

En un estilo más bien coloquial, nos proponemos brindar en esta sección un esbozo de la *Teoría Matemática de la Probabilidad*. Dejando de lado algunos remotos y aislados antecedentes previos, podemos ubicar el inicio de la teoría a mediados del siglo XVII, cuando a requerimiento del caballero Le Mère, y a través del intercambio de una serie de cartas, Blaise Pascal y Pierre de Fermat analizaron cuáles eran las formas más convenientes de apostar en diversos juegos de azar, en función de las chances de ganar de los apostadores. El estudio de tales problemas, estrechamente ligado al desarrollo del cálculo combinatorio, fue continuado posteriormente por Bernoulli, de Moivre, Lagrange y otros, y fue Laplace en el siglo XIX quien estableció la primera teoría general de la probabilidad. Con el correr del tiempo la misma amplió fuertemente su base matemática, y sus alcances superaron largamente la motivación original relacionada con el juego. Así es como en la actualidad la teoría se emplea fructíferamente en campos tan diversos como la física, la medicina, la economía, las ciencias sociales, etc.

La palabra probabilidad y otras afines con ella se usan con frecuencia y de manera informal en el lenguaje cotidiano. Por ejemplo, nos resultan familiares expresiones del tipo: "es probable que la próxima semana viaje a Mendoza", "es improbable que el aumento del índice de precios supere los 2 puntos este mes", "se espera vender por lo menos la mitad de las entradas", o "pronostican un 30 % de probabilidades de lluvia". Aceptando que todas estas frases adolecen de cierta vaguedad, observemos algunas características comunes: en todos los casos se hace referencia a un hecho o suceso cuya ocurrencia es incierta, estimándose además las chances de que el mismo ocurra. Siguiendo esta idea, la teoría matemática de la probabilidad cuantifica esa incertidumbre, asignando una cantidad numérica a la posible ocurrencia del hecho. Pasemos a describir con mayor precisión los lineamientos básicos de la teoría y los conceptos probabilísticos elementales.

4.3.2. Elementos básicos de la probabilidad

Experimento aleatorio Todo modelo probabilístico está asociado a un *experimento aleatorio*, entendiendo por tal a cualquier acción o proceso cuyo *resultado* no puede predecirse con certeza. Como ejemplos de dichos experimentos podemos mencionar el arrojar un dado, extraer una carta de un mazo de cartas, registrar el grupo sanguíneo de un individuo elegido al azar en una población, arrojar repetidamente una moneda hasta obtener dos caras, etc. El término proviene del latín *aleatoriŭs*, que significa propio del juego de dados.

Espacio muestral Dado un experimento aleatorio, el conjunto de sus resultados, que puede ser finito o infinito, se denomina *espacio muestral*. Por

ejemplo, en los tres primeros casos anteriores el espacio muestral es finito (6, 40 y 4 resultados, respectivamente), mientras que en el último es infinito, ya que consiste de las secuencias

$$CC, CSC, SCC, CSSC, SCSC, SSCC, \ldots\ldots$$

SUCESOS Dado un experimento aleatorio con espacio muestral E, cualquier subconjunto de E será llamado un *suceso*, o también un *evento*. Dicho de otro modo, un suceso es una colección de resultados del experimento.

Por ejemplo, al arrojar un dado el subconjunto $\{2, 4, 6\}$ es el suceso "salió un número par", y en el caso de la moneda arrojada hasta obtener 2 caras, el suceso "salieron exactamente 3 secas" corresponde al subconjunto de resultados

$$\{CSSSC, SCSSC, SSCSC, SSSCC\}.$$

Asimismo, el suceso "la moneda se arrojó 2 veces" corresponde al subconjunto unitario $\{CC\}$. En general, los sucesos que consisten exactamente de un resultado del experimento se llaman *simples*. Así, en el caso de los grupos sanguíneos los eventos simples son $\{0\}$, $\{A\}$, $\{B\}$ y $\{AB\}$.

Si dos sucesos corresponden a subconjuntos disjuntos del espacio muestral E, diremos que los mismos son *mutuamente excluyentes*. Por ejemplo, supongamos que el experimento consiste en arrojar simultáneamente un dado rojo y otro verde y observar los números obtenidos. Entonces los sucesos "la suma de los números que salieron es 8" y "uno de los dos números es un 1" son mutuamente excluyentes, mientras que los sucesos "uno de los números es el doble del otro" y "la suma de ambos números es par" no lo son.

ESPACIO DE PROBABILIDAD Supongamos dado un experimento aleatorio con espacio muestral E, que por razones de simplicidad supondremos finito. Intuitivamente hablando, el objetivo de la teoría de la probabilidad es asignar a cada suceso S un número real no negativo $p(S)$, llamado la *probabilidad* de S, de manera que el mismo brinde una medida precisa de la chance de ocurrencia de S. En términos más formales, dicha asignación debe ser una función

$$p : \mathbb{P}(E) \to \mathbb{R}_{\geq 0}$$

que verifique las condiciones:

$P_1)$ $p(E) = 1$.

$P_2)$ $p(S \cup T) = p(S) + p(T)$ si S y T son sucesos mutuamente excluyentes.

Diremos en tal caso que p es una *función de probabilidad* y que el par (E, p) es un *espacio de probabilidad*.

En la siguiente proposición exhibiremos algunas propiedades básicas de los espacios de probabilidad, de interés práctico y teórico. Probaremos entre

otros hechos que la probabilidad de cualquier evento A satisface las desigualdades

$$0 \leq p\,(A) \leq 1\,,$$

por lo que la función de probabilidad p resulta ser en realidad una aplicación del conjunto de partes del espacio muestral E en el intervalo $[0,1]$.

Proposición 4.3.1 Si (E,p) es un espacio de probabilidad, valen las siguientes propiedades (las letras mayúsculas denotan sucesos):

1) $p\,(S^c) = 1 - p\,(S)$, donde S^c denota el complemento de S (suceso *opuesto* a S). Deducimos entonces que $p\,(S) \leq 1$ cualquiera sea S.

2) $p\,(\emptyset) = 0\,.$

3) Si $S \subseteq T$ entonces $p\,(S) \leq p\,(T)\,.$

4) Sean $S_1, S_2, \ldots, S_r$ sucesos tales que S_i y S_j son mutuamente excluyentes si $i \neq j$. Entonces

$$p\left(\bigcup_{i=1}^{r} S_i\right) = \sum_{i=1}^{r} p\,(S_i).$$

5) $p\,(S \cup T) = p\,(S) + p\,(T) - p\,(S \cap T)\,.$

Demostracion El ítem 1) se obtiene aplicando sucesivamente los axiomas P_1) y P_2), ya que

$$p\,(S^c) = (p\,(S^c) + p\,(S)) - p\,(S) = p\,(E) - p\,(S) = 1 - p\,(S).$$

Claramente, 2) sigue de 1) tomando $S = E$, mientras que 3) se obtiene en forma análoga, pues tenemos

$$p\,(T) = p\,\left(S \cup (T - S)\right) = p\,(S) + p\,(T - S) \geq p\,(S),$$

por ser $p\,(T - S) \geq 0$. Notemos de paso que vale la fórmula

$$p\,\left(T - S\right) = p\,\left(T\right) - p\,\left(S\right).$$

La propiedad 4) es una generalización de P_2) y su demostración sigue sin dificultad por inducción en r, por lo que la dejamos a cargo del lector. Finalmente, para probar 5) comencemos observando que valen las igualdades

$$p\,(S) = p\,(S - T) + p\,(S \cap T)$$

y

$$p\,(T) = p\,(T - S) + p\,(S \cap T),$$

de donde obtenemos

$$\begin{aligned} p\,(S) + p\,(T) &= p\,(S \cap T) + (p\,(S \cap T) + p\,(S - T) + p\,(T - S)) = \\ &= p\,(S \cap T) + p\,(S \cup T), \end{aligned}$$

lo que implica el enunciado. Notemos que la última igualdad es consecuencia de 4), ya que $S \cup T$ es unión disjunta de $S \cap T$, $S - T$ y $T - S$. $\quad \Diamond$

Asignación de probabilidades.

En general, un punto delicado en la construcción de un espacio de probabilidad (E, p) es la asignación de probabilidades a los distintos eventos. Sin embargo, en el caso de un espacio finito la cuestión es bastante sencilla. En efecto, supongamos que E tiene n elementos, digamos $E = \{r_1, r_2, \ldots, r_n\}$, y sea S un subconjunto cualquiera de E. Considerando que todo conjunto es unión disjunta de sus subconjuntos unitarios, y usando la propiedad 4) de la proposición 4.3.1 tenemos:

$$p(S) = p\left(\bigcup_{r_i \in S} \{r_i\}\right) = \sum_{r_i \in S} p\left(\{r_i\}\right),$$

esto es, la probabilidad de cualquier suceso es suma de probabilidades de sucesos simples. Por lo tanto, para definir la función de probabilidad p basta asignar una probabilidad a cada resultado r_i. Así, escribiendo $p\left(\{r_i\}\right) = p_i$, resulta que cualquier modelo probabilístico asociado al espacio muestral E viene dado por la elección de n números reales no negativos $p_1, p_2, \ldots, p_n$ tales que

$$p_1 + p_2 + \ldots + p_n = 1\,, \tag{4.11}$$

ya que debe satisfacerse el axioma P_1). Notemos además que una tal asignación de probabilidades a los eventos simples garantiza la validez del axioma P_2), puesto que dados sucesos A y B mutuamente excluyentes tenemos:

$$p\left(A \cup B\right) = \sum_{z \in A \cup B} p\left(\{z\}\right) = \sum_{x \in A} p\left(\{x\}\right) + \sum_{y \in B} p\left(\{y\}\right) = p\left(A\right) + p\left(B\right).$$

Frecuencias relativas Visto que para diseñar un espacio de probabilidad finito es suficiente asignar probabilidades a cada uno de los resultados, es natural preguntarse acerca de la forma de efectuar dichas asignaciones. Ella descansa en general en el concepto de frecuencia relativa, que ya pasamos a explicar.

Asumamos que es posible realizar repetidamente en idénticas condiciones un cierto experimento aleatorio, siendo cada realización independiente de las otras. Dada una serie de m de tales realizaciones, y dado cualquier suceso A asociado al experimento, supongamos que A ocurre m_A veces durante la serie. En tal caso, el cociente

$$f_A = m_A/m,$$

que brinda la proporción de casos de ocurrencia del suceso A, se dirá la *frecuencia relativa* de A en dicha serie. La evidencia empírica muestra que al aumentar el número de realizaciones del experimento la frecuencia relativa de cualquier suceso tiende a estabilizarse, aproximándose a un valor límite.

Es natural entonces identificar la probabilidad teórica del suceso A con dicho valor, vale decir,

$$p\,(A) = \lim_{m\to\infty} m_A/m\,.$$

Así, si las probabilidades son asignadas de acuerdo con las frecuencias relativas observadas, un enunciado del tipo "la probabilidad de obtener un 4 al arrojar un cierto dado es $0,2$" deberá interpretarse en el sentido de que en cualquier larga secuencia de tiradas de dicho dado aproximadamente la quinta parte de las veces saldrá un 4.

Señalemos además otro hecho que apuntala la razonabilidad de asignar probabilidades a través del registro de las frecuencias relativas. Sea como antes

$$E = \{r_1, r_2, \ldots, r_n\}$$

y supongamos que en una serie de m realizaciones del experimento aleatorio el resultado r_i ocurrió m_i veces $(1 \leq i \leq n)$. Entonces, designando por f_i las correspondientes frecuencias relativas, tenemos:

$$\sum_{i=1}^{n} f_i = \sum_{i=1}^{n} \frac{m_i}{m} = \frac{1}{m}\sum_{i=1}^{n} m_i = \frac{m}{m} = 1\,,$$

esto es, los f_i verifican la relación 4.11 que deben satisfacer las probilidades p_i asignadas a los resultados r_i. Examinemos un par de ejemplos.

Ejemplo 4.3.2 En una localidad el 80 % de los habitantes está suscripto al periódico local, el 60 % a un periódico editado en una gran ciudad cercana, y un 50 % está suscripto a ambos periódicos. Si se elige un ciudadano al azar, ¿ cuál es la probabilidad de que esté suscripto por lo menos a un periódico ? ¿ Y la probabilidad de que esté suscripto a exactamente un periódico ?

Si designamos por S el suceso "el ciudadano está suscripto al periódico local", y por T el suceso "el ciudadano está suscripto al periódico de la ciudad vecina", resulta de los datos del problema que $p\,(S) = 0,8$, $p\,(T) = 0,6$ y $p\,(S \cap T) = 0,5$. Usando entonces la propiedad 5) de la proposición 4.3.1 tenemos:

$$p\,(S \cup T) = p\,(S) + p\,(T) - p\,(S \cap T) = 0,8 + 0,6 - 0,5 = 0,9\,,$$

lo que claramente da respuesta al primer interrogante. Respecto a la segunda cuestión, debemos calcular la probabilidad del suceso $S \triangle T$. Puesto que vale la relación

$$S \cup T = (S \triangle T) \cup (S \cap T)$$

y los sucesos $S \triangle T$ y $S \cap T$ son mutuamente excluyentes, obtenemos:

$$p\,(S \triangle T) = p\,(S \cup T) - p\,(S \cap T) = 0,9 - 0,5 = 0,4\,. \quad \Diamond$$

Ejemplo 4.3.3 Un médico examinará muestras de sangre tomadas al azar hasta hallar una con factor RH negativo. Si las probabilidades de que una muestra de sangre tomada al azar tenga factor RH positivo o factor RH negativo son $0,85$ y $0,15$, respectivamente, ¿cuál es la probabilidad p de que tenga que examinar por lo menos 3 muestras?

Si bien la situación no parece corresponder en principio al caso de un espacio muestral finito, ya que el suceso que nos interesa es unión de infinitos eventos simples (3 muestras, 4 muestras, etc.), el mismo puede ser descripto equivalentemente en la forma "las 2 primeras muestras tienen factor RH positivo", por lo que bastará estudiar la probabilidad de este suceso.

Para ello, consideremos un espacio muestral de 4 resultados, a saber:

$$E = \{(++), (+-), (-+), (--)\}\,.$$

Para asignar las correspondientes probabilidades pensemos en términos de frecuencias relativas. Puesto que al elegir al azar una muestra de sangre obtendremos factor RH positivo aproximadamente el $85\,\%$ de las veces, y al repetir el experimento, por cada una de ellas también debemos esperar alrededor de un $85\,\%$ de muestras con factor RH positivo, lo razonable es asignar al resultado $(++)$ la probabilidad

$$p_1 = \frac{85}{100} \times \frac{85}{100} = 0,7225\,.$$

Razonando en forma idéntica, asignamos a los resultados $(+-)$, $(-+)$ y $(--)$ las probabilidades $p_2 = 0,1275$, $p_3 = 0,1275$ y $p_4 = 0,0225$, respectivamente. Observando que se satisface la condición $\sum p_i = 1$, concluimos que la respuesta al problema es $p = p_1 = 0,7225$. $\Diamond$

Resultados equiprobables.

Por su importancia, consideremos aquel caso en el cual los n resultados del espacio muestral son *equiprobables*, esto es, tienen todos la misma probabilidad p de ocurrencia. Ello sucede por ejemplo al arrojar un dado o una moneda perfectamente equilibrados, y es en general la situación bajo la cual se analizan las chances en cualquier juego de azar.

Usando nuestras notaciones habituales, puesto que $p_i = p\,(r_i) = p$ para todo i, sigue de la fórmula 4.11 que

$$p = 1/n\,.$$

Por ejemplo, al arrojar un dado o una moneda equilibrados las probabilidades de obtener cualquiera de los resultados posibles son $1/6$ y $1/2$, respectivamente, mientras que la probabilidad de acertar un pleno en una ruleta con perfecto balance es $1/37$.

En el caso de equiprobabilidad de los resultados, el cálculo de probabilidades es una cuestión esencialmente combinatoria. En efecto, si S es un

suceso, procediendo como antes obtenemos:

$$p(S) = p\left(\bigcup_{r_i \in S} \{r_i\}\right) = \sum_{r_i \in S} p(r_i) = \sum_{r_i \in S} \frac{1}{n} = \frac{\#(S)}{n},$$

vale decir, la probabilidad de un suceso (que es un subconjunto del espacio muestral) es la razón entre su cardinal y el cardinal del espacio.

Tradicionalmente, esta última relación suele expresarse en la forma

$$\text{probabilidad} = \frac{\text{CF}}{\text{CP}},$$

donde CF significa casos favorables (los elementos de S) y CP casos posibles (cualquiera de los resultados del espacio muestral). Con riesgo de ser reiterativos, recalquemos que este lenguaje, posiblemente inspirado en la conexión original existente entre la teoría de la probabilidad y los juegos de azar, solo se aplica a espacios finitos con resultados equiprobables.

Por ejemplo, calculemos la probabilidad de obtener suma 7 al arrojar dos dados equilibrados, uno verde y otro rojo. Puesto que los dados son distintos, el número de casos posibles es 36, mientras que una simple inspección nos muestra que los casos favorables son $(1,6)$, $(2,5)$, $(3,4)$, $(4,3)$,$(5,2)$ y $(6,1)$. Por lo tanto, la probabilidad de que la suma de los resultados obtenidos sea 7 es $6/36 = 1/6$.

Ejemplo 4.3.4 En una mano de truco, se reparten a cada uno de los cuatro jugadores tres cartas de una baraja española. Calculemos la probabilidad de que el jugador X reciba: a) el as de espadas, b) dos cartas del mismo palo, c) tres figuras.

Aclaremos que las 40 cartas que contiene una baraja española pertenecen a 4 palos distintos, habiendo 10 cartas de cada palo. Naturalmente, supondremos que todos los resultados que puede obtener X son equiprobables. Puesto que ellos consisten en la elección de 3 cartas distintas entre 40, resulta que el número de casos posibles es C_3^{40}. Comenzando con la cuestión a), es claro que el número de ternas favorables es C_2^{39}, ya que la misma debe incluir el as de espadas y otras 2 cartas cualesquiera elegidas entre las 39 restantes. Por la tanto, la probabilidad que tiene X de recibir el as de espadas es

$$C_2^{39}/C_3^{40} = 0,075.$$

Una observación destinada a los entusiastas del truco. El resultado hallado significa que en una jornada de suerte moderada el jugador X debiera recibir el as de espadas (carta de máxima valoración en el truco) aproximadamente una vez cada 13 manos.

Respecto al problema b), relacionado con el lance del envido, es más sencillo contar el número de casos no favorables, en los que las 3 cartas son de distintos palos. Para hacerlo, elijamos primero 3 palos entre los 4

de la baraja, y luego seleccionemos dentro de cada uno de ellos una carta cualquiera entre 10. Resulta así que el número de casos no favorables es $10^3 C_3^4$, por lo que la probabilidad buscada es

$$1 - \frac{10^3 C_3^4}{C_3^{40}} = 0,595\ldots$$

Esto nos muestra que alrededor del 60 % de las veces X tendrá alguna chance de ganar el envido. Finalmente, el suceso de recibir 3 figuras (nada anhelado por los jugadores de truco) contiene C_3^{12} casos favorables, ya que hay 12 figuras (3 de cada palo), de las cuales hay que seleccionar 3. En consecuencia, su probabilidad (bastante baja) es

$$\frac{C_3^{12}}{C_3^{40}} = 0,022\ldots \quad \Diamond$$

Realizaciones sucesivas de un experimento.

Si (E, p) es un espacio de probabilidad correspondiente a un cierto experimento aleatorio, y s es un número natural, la realización sucesiva de s ensayos de dicho experimento determina un nuevo espacio de probabilidad, con espacio muestral E^s y función de probabilidad q que definimos por

$$q\left((z_1, z_2, \ldots, z_s)\right) = \prod_{i=1}^{s} p\left(z_i\right), \tag{4.12}$$

para toda $s-$upla $(z_1, z_2, \cdots, z_s)$ de resultados de nuestro experimento (confrontar con el ejemplo 4.3.3 acerca de los factores sanguíneos).

Por razones de simplicidad notacional, sólo probaremos que esta asignación de probabilidades satisface la condición 4.11 en el caso $s = 2$, dejando el caso general a cargo del lector. Escribiendo como de costumbre $E = \{r_1, r_2, \ldots, r_n\}$ y $p(r_i) = p_i$ para todo i, tenemos:

$$\begin{aligned}\sum_{1\le i,j\le n} q\left((r_i, r_j)\right) = \sum_{1\le i,j\le n} p_i p_j = \sum_{1\le i\le n}\sum_{1\le j\le n} p_i p_j = \\ = \sum_{1\le i\le n} p_i \sum_{1\le j\le n} p_j = \sum_{1\le i\le n} p_i = 1\,,\end{aligned}$$

como queríamos demostrar.

El lector puede comprobar fácilmente que los resultados del espacio E^s son equiprobables si los resultados r_i lo son. Por caso, al arrojar 5 veces seguidas una moneda equilibrada la probabilidad de cualquiera de los 32 resultados posibles es 1/32. A través de los siguientes ejemplos ilustraremos la nueva situación descripta.

Ejemplo 4.3.5 Supongamos que se pregunta la fecha de nacimiento a 23 personas elegidas al azar. ¿Le parece probable que haya entre ellas dos personas que cumplan años el mismo día?

En principio estamos usando el término "probable" en un sentido no muy técnico, como sinónimo de "bastante posible", "esperable", etc. A efectos de precisar algo más la cuestión, preguntémonos si la probabilidad de tal suceso S es razonablemente alta, por ejemplo, si es mayor que $1/2$. Para calcularla efectivamente, advirtamos que la encuesta consiste en 23 realizaciones sucesivas de un experimento que puede arrojar 365 resultados equiprobables (excluyamos el 29 de febrero), por lo que ella puede arrojar 365^{23} resultados posibles también equiprobables.

En realidad es más sencillo calcular la probabilidad del suceso opuesto a S, por lo que consideraremos caso favorable a cualquier secuencia de 23 fechas distintas del año. Cada una de ellas corresponde a una función inyectiva de un conjunto de 23 elementos en otro de 365, y en consecuencia el número de casos favorables es $365!/342!$. Luego:

$$p(S) = 1 - p(S^c) = 1 - \frac{365!}{342! \times 365^{23}} = 0,507\ldots > \frac{1}{2}\,.$$

Expresado en lenguaje coloquial, hemos demostrado el siguiente hecho, no demasiado intuitivo: dado un grupo de 23 personas elegidas al azar, hay por lo menos un 50 % de chances de que dos de ellas cumplan años el mismo día. Es interesante comentar que si planteamos el mismo problema con menos de 23 personas obtenemos una probabilidad menor que $1/2$, esto es, 23 es mínimo respecto a la propiedad matemática que hemos demostrado. $\Diamond$

Ejemplo 4.3.6 Para aprobar un examen de elección múltiple, los estudiantes deben contestar correctamente por lo menos 8 de las 15 preguntas planteadas. Joaquín, que sabe responder perfectamente 4 de las preguntas, decide contestar las otras 11 eligiendo al azar en cada caso alguna de las 5 opciones ofrecidas, de las cuales exactamente una es correcta. ¿Qué probabilidad tiene de aprobar?

El problema se inserta en el marco teórico que acabamos de desarrollar, pues podemos pensar a la elección múltiple que debe efectuar Joaquín como la realización sucesiva de 11 ensayos de un experimento aleatorio que puede arrojar 2 resultados, a saber: $\mathcal{C}$ (por respuesta correcta) e $\mathcal{I}$ (por respuesta incorrecta). Resulta así que un elemento genérico de nuestro espacio muestral es una secuencia de k letras $\mathcal{C}$ y $11-k$ letras $\mathcal{I}$, cuya probabilidad de ocurrencia es $(1/5)^k(4/5)^{11-k}$, de acuerdo con la fórmula 4.12, ya que en este caso tenemos $p_1 = p(\mathcal{C}) = 1/5$ y $p_2 = p(\mathcal{I}) = 4/5$.

Puesto que en esta elección por azar Joaquín debe sumar por lo menos 4 respuestas correctas, debemos calcular la probabilidad q del suceso formado por secuencias que contienen 4 o más letras $\mathcal{C}$, o alternativamente, calcular la probabilidad complementaria q' del suceso determinado por secuencias

conteniendo a lo sumo 3 letras $\mathcal{C}$. En general, el número de secuencias con exactamente k letras $\mathcal{C}$ es C_k^{11}, por lo que resulta

$$q = 1 - q' = 1 - \sum_{k=0}^{3} \binom{11}{k} \left(\frac{1}{5}\right)^k \left(\frac{4}{5}\right)^{11-k} = 0,161\ldots$$

El número obtenido es muy cercano a 1/6, lo que significa que la expectativa de aprobar que tiene Joaquín es similar a la que tendría alguien que arroja un dado con la esperanza de obtener un as. Todo parece indicar que conviene estudiar un poco más. $\Diamond$

4.3.3. Ejercicios

1. Determinar el espacio muestral correspondiente a los siguientes experimentos aleatorios:

   a) Arrojar una moneda hasta que se repita un resultado.

   b) Lanzar una moneda hasta obtener dos caras o dos secas consecutivas.

   c) Extraer sucesivamente una carta de un mazo de baraja española hasta que se repita algún palo.

   d) Elegir al azar un número real en el intervalo $(0, 1)$.

2. Determinar el subconjunto del espacio muestral que corresponde al suceso "la suma de las caras es múltiplo de 3", si el experimento consiste en arrojar dos dados distintos.

3. Si se arrojan dos dados distintos, ¿cuál es la probabilidad de obtener algún seis?

4. Se extraen 3 bolas de una urna que contiene 3 bolas blancas y 2 negras. Describir el espacio muestral y determinar los sucesos simples de mayor y menor probabilidad.

5. Se arroja 9 veces una moneda equilibrada. Calcular la probabilidad de que

   a) Se obtenga una cantidad par de caras.

   b) La primera y la quinta tirada sean secas.

   c) Se obtengan tantas caras como secas.

6. Una urna contiene 4 bolas blancas y 4 bolas negras. Si se extraen 3 bolas al azar, ¿cuál es la probabilidad de que exactamente una sea blanca? ¿Y de que a lo sumo una sea negra?

7. Se elige al azar una permutación de 2345789. Calcular la probabilidad de que

   a) El 2 y el 3 ocupen posiciones pares.

   b) Ningún número conserve su posición original.

   c) El número determinado sea impar y comience en 1.

8. Para tomarse una fotografía, Paula, Lucía y otras cuatro personas se disponen al azar en fila. Calcular la probabilidad de que Paula se encuentre a la izquierda de Lucía.

9. En una ciudad hay 5 hoteles. Si 3 viajeros buscan hospedaje, ¿cuál es la probabilidad de que ningún par de ellos se aloje en el mismo hotel?

10. Una mujer tiene 6 llaves y sólo una de ellas abre su puerta. Si prueba al azar con las llaves, apartando cada vez las ya utilizadas, ¿ cuál es la probabilidad de que acierte en el tercer intento? ¿Y en el cuarto? Resolver las mismas cuestiones suponiendo que no descarta las llaves ya probadas.

11. En un curso de Cálculo el 40 % de los alumnos reprobó el primer examen parcial, el 30 % reprobó el segundo y el 20 % no aprobó ninguno de los dos. Si un estudiante elegido al azar no aprobó el primer examen, ¿cuál es la probabilidad de que haya reprobado los dos?

12. Una urna contiene 4 bolas rojas y 4 bolas azules. Se extraen al azar 4 bolas de la urna, y a menos que 2 sean rojas y 2 sean azules, se las devuelve a la urna, repitiéndose el procedimiento hasta que se extraigan dos rojas y dos azules. Determinar, para cada $n \in \mathbb{N}$, la probabilidad de que ello suceda en la $n-$ésima extracción.

13. Calcular la probabilidad de sumar 7 al arrojar dos veces un dado no equilibrado, si las probabilidades de sus caras son $p_1 = 0{,}2$, $p_2 = 0{,}16$, $p_3 = 0{,}18$, $p_4 = 0{,}16$, $p_5 = 0{,}17$, y $p_6 = 0{,}13$.

14. Un examen de elección múltiple ofrece 4 respuestas incorrectas y una respuesta correcta para cada una de sus 10 preguntas. Si un alumno elige todas las respuestas al azar, calcular (para cada k entre 0 y 10) la probabilidad de que responda bien exactamente k preguntas.

15. Una empresa manufactura discos compactos y los vende en cajas de 10 unidades. La probabilidad de que un disco resulte defectuoso es 0,02. Si la empresa garantiza el cambio de una caja en caso de que contenga más de un disco fallado, ¿ cuál es la probabilidad de que una caja elegida al azar deba ser reemplazada ?

Capítulo 5

Aritmética

5.1. Divisibilidad

5.1.1. Divisores y múltiplos

El conjunto $\mathbb{Z}$ de números enteros constituye el objeto central de estudio de la *Aritmética* (del griego *arithmos*, número), también llamada teoría de números. Su punto de partida es el concepto de *divisibilidad*, que pasamos a definir:

> Dados $a, b \in \mathbb{Z}$, diremos que a *divide* a b si y sólo si existe $c \in \mathbb{Z}$ tal que $b = ac$. Emplearemos en tal caso la notación $a \mid b$, mientras que si a no divide a b escribiremos $a \nmid b$.

Alternativamente, también nos referiremos a la situación anterior diciendo que a es *divisor* de b o que b es *múltiplo* de a. Por ejemplo, 12 es un divisor de 252, pues $252 = 12 \,.\, 21$, mientras que 252 es múltiplo de -7, ya que $252 = (-7).(-36)$. Observemos además que los roles de a y c en la definición de divisibilidad son intercambiables, de donde resulta que también c es un divisor de b.

Señalemos un par de hechos salientes referidos a la definición anterior. En primer término, es muy fácil analizar la situación cuando alguno de los números es cero. En efecto, de la igualdad $0 = a \,.\, 0$ deducimos simultáneamente que $a \mid 0$ cualquiera sea $a \in \mathbb{Z}$ y que $0 \mid b$ si y solo si $b = 0$. Resumidamente, 0 es múltiplo de todos los enteros y sólo es divisor de sí mismo.

Por otro lado, y suponiendo $a \neq 0$, observemos que el factor de divisibilidad c de la definición está completamente determinado, a saber, $c = b/a$. Tenemos luego una forma alternativa de definir la relación de divisibilidad en el caso $a \neq 0$:

$$a \text{ divide a } b \text{ si y sólo si } b/a \in \mathbb{Z}\,.$$

Enunciaremos y probaremos a continuación algunas propiedades básicas de la divisibilidad. En todos los casos, las letras designan números enteros.

Proposición 5.1.1 Valen las siguientes propiedades:

1) $a \mid b \Leftrightarrow |\,a\,| \mid |\,b\,|$

2) Si $b \neq 0$ y $a \mid b$ entonces $|\,a\,| \leq |\,b\,|$. Deducimos en particular que todo entero no nulo admite un número finito de divisores

3) Todo entero no nulo admite infinitos múltiplos

4) La relación de divisibilidad es reflexiva y transitiva

5) $a \mid b$ y $b \mid a \Leftrightarrow b = \pm a$

6) Si $a \mid b_1$ y $a \mid b_2$ entonces $a \mid b_1x_1 + b_2x_2$. Más generalmente, si $n \in \mathbb{N}$ y $a \mid b_i$ para $i = 1, 2, \cdots, n$ entonces $a \mid b_1x_1 + b_2x_2 + \cdots + b_nx_n$

7) Si $a \mid b_1 + b_2$ y $a \mid b_1$ entonces $a \mid b_2$. Más generalmente, sea $n \in \mathbb{N}$ tal que $a \mid b_1 + b_2 + \cdots + b_n$ y $a \mid b_i$ para todo $i < n$. Entonces $a \mid b_n$.

Demostracion.

1) Una igualdad del tipo $b = ac$ es equivalente a cualquiera de las igualdades $-b = a(-c)$, $b = (-a)(-c)$ y $-b = (-a)c$. Por lo tanto, $a \mid b$ si y sólo si $a \mid -b$, $-a \mid b$ y $-a \mid -b$. Esto claramente prueba nuestro requerimiento y nos muestra de paso que es suficiente estudiar relaciones de divisibilidad entre números naturales.

2) Sigue de $b = ac$ que $c \neq 0$. Luego, tomando módulos obtenemos

$$|\,b\,| = |\,ac\,| = |\,a\,|\,|\,c\,| \geq |\,a\,|,$$

pues $|\,c\,| \geq 1$. La segunda afirmación del enunciado sigue entonces del hecho de que sólo hay un número finito de enteros de módulo acotado.

3) Si $a \in \mathbb{Z}$, notaremos por $a\mathbb{Z}$ el conjunto de sus múltiplos. Esto es:

$$a\mathbb{Z} = \{ak : k \in \mathbb{Z}\}.$$

Si $a \neq 0$, resulta que $ak = ak' \Leftrightarrow k = k'$. Luego, la aplicación $x \mapsto ax$ es una biyección del conjunto de números enteros en el conjunto de múltiplos de a. En particular, a admite infinitos múltiplos. Como caso importante observemos que todo entero es múltiplo de 1, ya que $k = 1\,k$ para todo k. Usando 1) y 2) sigue fácilmente que 1 y -1 son los únicos enteros que satisfacen tal propiedad.

4) La reflexividad es inmediata, ya que $u = u\,1$ y por lo tanto $u \mid u$ cualquiera sea u. En cuanto a la transitividad, supongamos que $u \mid v$ y $v \mid w$, digamos $v = uk$ y $w = vt$. Entonces $w = (uk)t = u(kt)$, lo que muestra que $u \mid w$, pues $kt \in \mathbb{Z}$. Luego la relación de divisibilidad es transitiva.

Aunque sólo sea otra manera de decir lo mismo, conviene tener presente la siguiente descripción de la propiedad transitiva: si b es múltiplo de a entonces bk es múltiplo de a cualquiera sea $k \in \mathbb{Z}$.

5) Sigue inmediatamente de las propiedades anteriores.

6) Como señalamos en nuestro comentario acerca de la transitividad, la hipótesis del enunciado asegura que b_1x_1 y b_2x_2 son múltiplos de a, digamos $b_1x_1 = ak_1$ y $b_2x_2 = ak_2$. Entonces:

$$b_1x_1 + b_2x_2 = ak_1 + ak_2 = a(k_1 + k_2),$$

y por lo tanto $a \mid b_1x_1 + b_2x_2$. Eligiendo convenientemente x_1 y x_2 obtenemos como casos particulares que $a \mid b_1 + b_2$ y $a \mid b_1 - b_2$.

En cuanto a la generalización aludida en el enunciado, sigue fácilmente por inducción en n.

7) Como antes, bastará probar el caso $n = 2$ y usar luego un argumento inductivo. Para ello, observemos simplemente que $b_2 = (b_1 + b_2) - b_1$ es un múltiplo de a por ser diferencia de dos múltiplos de a. $\Diamond$

5.1.2. Algoritmo de división

En nuestra etapa escolar, todos hemos aprendido, junto con las restantes operaciones elementales, a efectuar la división entera de dos números naturales. Descontamos que el lector conoce el algoritmo de obtención del cociente y el resto de la división de un número natural b por otro a, método de cálculo que permite en particular decidir si a es un divisor de b. Ahora bien, no obstante la sencillez de la idea que encierra, la división entera es un hecho de gran importancia dentro de la Aritmética, tanto por sus consecuencias prácticas como teóricas. Por lo tanto, además de saber hacer las cuentas debemos comprender perfectamente su significado y sus alcances.

Una breve descripción del algoritmo podría ser la siguiente: se trata de aproximar de la mejor manera posible a b por un múltiplo de a. La diferencia entre b y dicho múltiplo es lo que llamamos el resto de la división, que será nulo exactamente en el caso en que b sea un múltiplo de a. Es evidente que todo esto es algo impreciso, y que surgen inmediatamente algunos interrogantes. Por ejemplo, ¿qué significa "de la mejor manera posible"? Veamos cómo clarificar esta cuestión y cómo establecer las características del cociente y el resto.

Teorema 5.1.2 (Algoritmo de división) Sean $a, b \in \mathbb{Z}\,(a \neq 0)$. Existe entonces un único par (q, r) de enteros, a los que llamaremos respectivamente *cociente* y *resto* de la división entera de b por a, satisfaciendo las siguientes condiciones:

$$\begin{aligned} D_1) &\quad b = qa + r \\ D_2) &\quad 0 \leq r < |\,a\,|\,. \end{aligned}$$

DEMOSTRACION. Supongamos en primer término $b \geq 0$ y consideremos el conjunto

$$R = \{b - ka : k \in \mathbb{Z}\} \cap \mathbb{N}_0.$$

Tomando $k = 0$ sigue por nuestra suposición que $b \in R$, y por lo tanto $R \neq \emptyset$. Siendo un subconjunto de enteros no negativos resulta que R tiene mínimo, digamos r. Luego, por definición, existe $q \in \mathbb{Z}$ tal que $r = b - qa$, ó equivalentemente, $b = qa+r$. Para completar este tramo de la demostración sólo nos resta probar que $r < |\, a \,|$.

Para ello, escribamos $|\, a \,| = \epsilon a$, con $\epsilon = \pm 1$. Entonces $r - |\, a \,| < r$ y además

$$r - |\, a \,| = r - \epsilon a = b - qa - \epsilon a = b - (q + \epsilon)a\,,$$

esto es, $r - |\, a \,|$ es de la forma $b - ka$ con $k \in \mathbb{Z}$. Siendo $r = \min R$, deducimos que $r - |\, a \,|$ es negativo, y por lo tanto $r < |\, a \,|$, como queríamos probar.

Supongamos ahora $b < 0$. Por lo demostrado arriba, sabemos que existen enteros p y s tales que $-b = pa + s$, siendo $0 \leq s < |\, a \,|$. Obtenemos entonces la igualdad $b = (-p)a - s$, lo que nos acerca a nuestra meta. En efecto, si $s = 0$ basta tomar $q = -p$ y $r = 0$, mientras que si $s > 0$ tenemos:

$$b = (-p)a - s + |\, a \,| - |\, a \,| = (-p - \epsilon)a + |\, a \,| - s,$$

y obtenemos el resultado tomando $q = -p - \epsilon$ y $r = |\, a \,| - s$ (notemos que $0 < |\, a \,| - s < |\, a \,|$).

Para probar la unicidad, supongamos que el par (q_1, r_1) también satisface las condiciones D_1) y D_2) del enunciado y asumamos sin pérdida de generalidad que $r_1 \leq r$. Restando las correspondientes igualdades obtenemos

$$0 = b - b = (q_1 a + r_1) - (qa + r) = (q_1 - q)a + (r_1 - r),$$

de donde

$$r - r_1 = (q_1 - q)a, \tag{5.1}$$

esto es, $r - r_1$ es múltiplo de a, y por lo tanto también de $|\, a \,|$. Siendo

$$0 \leq r - r_1 \leq r < |\, a \,|\,,$$

la única posibilidad es $r - r_1 = 0$, de donde $r_1 = r$ y $q_1 = q$, como deducimos de la igualdad 5.1. $\Diamond$

NOTA. Las siguientes observaciones nos ayudarán a mejorar nuestra percepción del cociente y el resto.

Conservando las notaciones del teorema y suponiendo $a > 0$, por simplicidad, se advierte fácilmente que valen las desigualdades

$$qa \leq b < (q + 1)a\,,$$

esto es, la idea de la división entera es ubicar a b (el *dividendo*) entre dos múltiplos consecutivos de a (el *divisor*). De tal manera, qa es el múltiplo de a más próximo a b por defecto, y la diferencia entre ambos es el resto de la división. En la figura 5.1 visualizaremos la situación.

Vale la pena comentar también que el algoritmo de división es rápido y fácilmente programable, ya que el cociente y el resto pueden obtenerse simplemente restando. En efecto, suponiendo $a, b \in \mathbb{N}$ y $a \leq b$, basta ir efectuando las sucesivas diferencias

$$b - a\,, b - 2a\,, b - 3a\,, \ldots\,,$$

hasta obtener una en el intervalo $[0, a)$, hecho que ocurre exactamente una vez, ya que nos movemos decrecientemente a partir de b con un "paso" de longitud a. Esta última diferencia es el resto de la división, y el número de veces que restamos a es el cociente. Notemos además que en el caso $a < b$ basta tomar $q = 0$ y $r = b$.

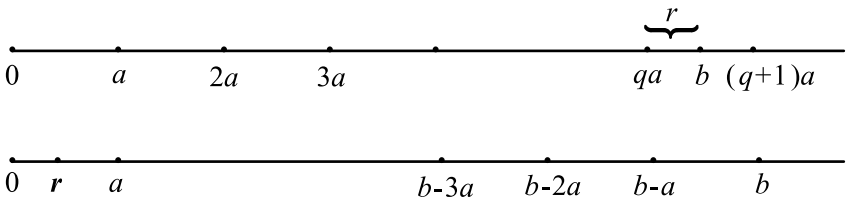


Figura 5.1: Algoritmo de división

Si los números involucrados no son positivos caben interpretaciones similares. Digamos de paso que para hallar el cociente y el resto de la división entera de dos números enteros cualesquiera basta efectuar la división entera de sus valores absolutos, hecho cuyos detalles dejaremos a cargo del lector en los ejercicios del final de la sección. Finalmente, y siempre suponiendo $a > 0$, veamos de qué otra forma puede interpretarse el cociente. Si dividimos ambos miembros de la igualdad D_1) por a obtenemos

$$\frac{b}{a} = q + \frac{r}{a}\,,$$

resultando que

$$0 \leq \frac{r}{a} < 1\,,$$

por D_2). Concluimos entonces que el cociente de la división entera de b por a es *la parte entera del número racional* b/a .

De ahora en más, para agilizar la notación designaremos por $r_m(x)$ el resto de dividir un entero x por cualquier entero no nulo m. $\Diamond$

Como señalamos anteriormente, el algoritmo de división permite decidir efectivamente si un cierto número es múltiplo de otro. Precisamente:

Corolario 5.1.3 Sean $a, b \in \mathbb{Z}\,(a \neq 0)$. Entonces $a \mid b$ si y sólo si $r_a(b) = 0$.

DEMOSTRACION. La condición sobre el resto es claramente suficiente, ya que en tal caso $b = qa$, y por lo tanto $a \mid b$. Recíprocamente, supongamos que a divide a b, digamos,

$$b = ka = ka + 0\,.$$

Por la unicidad de cociente y resto, concluimos entonces que k es el cociente de la división entera de b por a y $r_a(b) = 0$, como queríamos demostrar. $\Diamond$

En la demostración anterior nos hemos referido a la unicidad del cociente y el resto. En el siguiente ejemplo ilustraremos nuevamente esta cuestión, a la que dada su importancia debe prestarse especial atención.

Ejemplo 5.1.4 Calculemos los restos de dividir un entero m por 4, 6 y 12, respectivamente, sabiendo que m excede en 47 a un cierto múltiplo de 24.

Puesto que m no está determinado, no hay en principio ningún cálculo concreto que efectuar. Sin embargo podemos contestar las preguntas. Para ello, observemos que la condición del enunciado significa que existe $k \in \mathbb{Z}$ tal que $m = 24k + 47$, igualdad que podemos reescribir de las tres siguientes formas:

$$\begin{aligned} m &= 4(6k + 11) + 3 \\ m &= 6(4k + 7) + 5 \\ m &= 12(2k + 3) + 11\,, \end{aligned}$$

como es inmediato verificar. Si a es cualquiera de los números 4, 6 ó 12, notemos que en todos los casos hemos expresado a m como suma de un múltiplo de a y un número que está en el rango de los restos de dividir por a. Luego, por unicidad del resto, resulta que los resultados pedidos son 3, 5 y 11, respectivamente. $\Diamond$

Analizando brevemente la resolución anterior, advertimos que la cuestión resultó sencilla debido a que 24 es múltiplo de 4, 6 y 12, y todo lo que tuvimos que hacer es dividir 47 por dichos números y luego agrupar convenientemente. Esta situación se inserta en un marco teórico más general, que ya pasamos a describir.

Proposición 5.1.5 Sean $m, u, v \in \mathbb{Z}$ $(m \neq 0)$. Entonces $r_m(u) = r_m(v)$ si y sólo si u y v difieren en un múltiplo de m.

DEMOSTRACION. Si $r_m(u) = r_m(v) = r$, restando las igualdades $u = km + r$ y $v = tm + r$ obtenemos que $u - v = (k - t)m$. Luego $m \mid u - v$.

Recíprocamente, sea $h \in \mathbb{Z}$ tal que $u - v = hm$ y sean q y s el cociente y el resto de la división entera de v por m, de donde

$$u = u - v + v = hm + qm + s = (h + q)m + s.$$

Resulta entonces que $r_m(u) = s$, por unicidad de cociente y resto, lo que concluye la prueba. $\Diamond$

CONGRUENCIA. Idealmente, supongamos que recorremos los números enteros y vamos anotando los restos de dividirlos por un número natural m. Resulta entonces que dichos restos forman el ciclo infinito

$$\ldots, 0, 1, 2, 3, \ldots, (m-2), (m-1), 0, 1, 2, \ldots ,$$

en el que cada valor se repite exactamente cada m pasos, como indica el corolario 5.1.5. La situación que ella caracteriza es de gran importancia en teoría de números (Aritmética) y emplearemos una terminología especial para referirnos a ella. Con precisión, si $r_m(u) = r_m(v)$, o equivalentemente, si $v - u$ es múltiplo de m, diremos que u es *congruente* con v módulo m. Emplearemos en tal caso la notación

$$u \equiv v \ (m) . \quad \Diamond$$

Cálculo de restos.

A partir de 5.1.5 es fácil describir cómo se comportan los restos respecto a las operaciones de anillo de los números enteros, como veremos a continuación.

Proposición 5.1.6 Sean $x, y \in \mathbb{Z}$ y sean $s = r_m(x)$ y $t = r_m(y)$, donde m es un entero no nulo. Valen entonces las siguientes propiedades:

1) $r_m(x + y) = r_m(s + t)$

2) $r_m(x - y) = r_m(s - t)$

3) $r_m(xy) = r_m(st)$

4) $r_m(x^j) = r_m(s^j)$ para todo $j \in \mathbb{N}_0$.

DEMOSTRACION. De acuerdo con las hipótesis, podemos escribir

$$x = qm + s \tag{5.2}$$
$$y = km + t, \tag{5.3}$$

para ciertos enteros q y k. Entonces, sumando o restando 5.2 y 5.3 obtenemos

$$x \pm y = (q \pm k)m + (s \pm t),$$

esto es, $x \pm y$ y $s \pm t$ difieren en un múltiplo de m, de lo que deducimos 1) y 2), por proposición 5.1.5.

En forma análoga se prueba 3), ya que multiplicando 5.2 y 5.3 resulta que

$$xy = qkm^2 + qtm + ksm + st = (qkm + qt + ks)m + st,$$

y nuevamente el resultado sigue de 5.1.5. En cuanto a 4), sigue fácilmente por inducción en j. $\Diamond$

NOTA. Mediante sencillos argumentos inductivos puede probarse más generalmente la fórmula

$$r_m\left(\sum_{i=1}^{n} x_i y_i\right) = r_m\left(\sum_{i=1}^{n} s_i t_i\right),$$

donde las letras representan números enteros, siendo s_i y t_i los restos de dividir x_i e y_i por m, respectivamente. Como antes, m es no nulo, mientras que n es un número natural.

Notemos el significado de todas estas propiedades: si sólo nos interesa calcular el resto de dividir por m el resultado de una operación, podemos reemplazar todos los números involucrados por los restos de dividirlos por m y operar luego con ellos. Más aún, dichas propiedades siguen siendo válidas si reemplazamos cualquier número z no ya por $r_m(z)$ sino por cualquier u congruente con z módulo m, hecho que el lector puede comprobar fácilmente. Es claro que todas estas acciones de *reducción* módulo m −conocidas en conjunto como *Aritmética modular*− facilitan los cálculos, ya que permiten operar con números mucho más pequeños. Veamos un ejemplo.

Ejemplo 5.1.7 Los restos de dividir 77, 69 y 126 por 13 son 12, 4 y 9, respectivamente. Calculemos el resto de dividir $77^{45} \,.\, 69^2 + 126$ por 13. Aplicando reiteradamente las propiedades 5.1.6 y operando, obtenemos:

$$\begin{aligned} r_{13}(77^{45} \,.\, 69^2 + 126) &= r_{13}(12^{45} \,.\, 4^2 + 9) = r_{13}((-1)^{45} \,.\, 4^2 + 9) = \\ &= r_{13}((-1) \,.\, 16 + 9) = r_{13}(-16 + 9) = \\ &= r_{13}(-7) = r_{13}(6) = 6\,. \end{aligned}$$

Adviértase que es lícito reemplazar el resto 12 por -1 en la segunda igualdad, ya que $12 \equiv -1 \ (13)$. ◊

Sistemas de numeración.

Desde muy temprana edad nos hemos acostumbrado a reconocer los números por su forma de expresión. Es tal la familiaridad que tenemos con el sistema decimal de numeración que prácticamente no distinguimos entre el número (como idea abstracta) y la secuencia de símbolos que lo representa. Sin embargo su uso es relativamente reciente, ya que proveniente de la India y de Arabia fue introducido en Europa por Leonardo de Pisa en el siglo XI. El mismo es un sistema posicional, en el que cada símbolo (dígito) cuenta tanto por su valor intrínseco como también por la posición que ocupa. Esta última es su característica más notable e ingeniosa, ya que permite representar los infinitos números naturales con apenas 10 símbolos.

La adopción de 10 como *base* del sistema de numeración obedeció seguramente a razones culturales (la costumbre de contar con los dedos), pero es perfectamente posible usar otras bases de numeración, como de hecho

ocurre actualmente en los sistemas informáticos. Veamos entonces qué es un sistema de numeración.

Generalización del sistema decimal Sea b un número natural mayor que 1. A través del siguiente resultado mostraremos en qué consiste el sistema de numeración en base b:

Proposición 5.1.8 Todo número natural m se expresa de manera única en la forma

$$m = \sum_{i=0}^{k} m_i\, b^i, \tag{5.4}$$

donde $k \in \mathbb{N}_0$ y los m_i son enteros tales que $0 \leq m_i < b$ para todo i, siendo $m_k \neq 0$.

Demostracion. Probaremos la existencia de un desarrollo como el del enunciado por inducción en m. El resultado es trivial si $m = 1$, ya que en tal caso basta tomar $k = 0$ y $m_0 = 1$ (recordemos que $b > 1$). Supongamos ahora que $m > 1$ y que todo número natural menor que m admite un desarrollo del tipo (5.4). Si $m < b$, como en el caso $m = 1$ tomamos $k = 0$ y $m_0 = m$. Luego podemos asumir que $m \geq b$.

Dividiendo m por b, obtenemos una expresión de la forma

$$m = qb + r_b(m), \tag{5.5}$$

de donde sigue que

$$q = \frac{m - r_b(m)}{b} \leq \frac{m}{b} < m,$$

pues $b > 1$. Por otro lado, siendo $r_b(m) < b \leq m$ resulta que $q > 0$, y podemos aplicar entonces la hipótesis inductiva a q, digamos

$$q = \sum_{i=0}^{s} q_i b^i,$$

con las restricciones correspondientes para s y los enteros q_i. Reemplazando en (5.5) obtenemos:

$$\begin{aligned} m &= \left(\sum_{i=0}^{s} q_i b^i \right) b + r_b(m) = \sum_{i=0}^{s} q_i b^{i+1} + r_b(m) \\ &= \sum_{i=1}^{s+1} q_{i-1} b^i + r_b(m)\,. \end{aligned}$$

Luego basta tomar $k = s + 1$, $m_0 = r_b(m)$ y $m_i = q_{i-1}$ si $1 \leq i \leq k$.

Con respecto a la unicidad de la expresión (5.4), comencemos observando que valen las desigualdades

$$m = \sum_{i=0}^{k} m_i\, b^i \geq m_k b^k \geq b^k$$

y

$$m = \sum_{i=0}^{k} m_i b^i \leq \sum_{i=0}^{k} (b-1) b^i = (b-1) \sum_{i=0}^{k} b^i = \\ = (b-1) \left(\frac{b^{k+1}-1}{b-1} \right) = b^{k+1} - 1 < b^{k+1}\,.$$

Por lo tanto, k está completamente determinado por m, ya que la sucesión de potencias de b es estrictamente creciente y existe un único entero no negativo k tal que $b^k \leq m < b^{k+1}$. Supongamos ahora que m admite dos desarrollos del tipo (5.4) (necesariamente de igual longitud), digamos

$$m = \sum_{i=0}^{k} m_i b^i = \sum_{i=0}^{k} c_i b^i. \tag{5.6}$$

Si $k = 0$ la igualdad de arriba afirma simplemente que $c_0 = m_0$ y no hay nada más que probar. Si $k > 0$, podemos reescribir (5.6) en la forma

$$m = b \left(\sum_{i=0}^{k-1} m_{i+1} b^i \right) + m_0 = b \left(\sum_{i=0}^{k-1} c_{i+1} b^i \right) + c_0,$$

de donde deducimos por unicidad de cociente y resto que $c_0 = m_0 = r_b(m)$ y

$$\sum_{i=0}^{k-1} m_{i+1} b^i = \sum_{i=0}^{k-1} c_{i+1} b^i.$$

Por un argumento inductivo sobre k sigue entonces que $c_{i+1} = m_{i+1}$ para $0 \leq i \leq k-1$, ó equivalentemente, $c_i = m_i$ para $1 \leq i \leq k$, como queríamos demostrar. $\Diamond$

La expresión (5.4) se llama el *desarrollo b-ádico* de m o desarrollo en base b de m. Por supuesto, el desarrollo decimal corresponde al caso $b = 10$. Los $k+1$ coeficientes $m_0, m_1, \ldots, m_k$, unívocamente determinados por m y b, se llaman las *cifras b-ádicas* de m, y en forma análoga al caso del sistema decimal emplearemos la notación

$$m = (m_k m_{k-1} \ldots m_1 m_0)_b\,.$$

Por ejemplo, 252 se escribe $(2002)_5$ en base 5 y $(501)_7$ en base 7, como es inmediato comprobar. Por una simple razón de uniformidad, le asignaremos a cero el desarrollo b-ádico $(0)_b$.

NOTA. Dado un número natural m, podemos obtener su expansión en base b mediante sucesivas aplicaciones del algoritmo de división, como lo sugiere la prueba de 5.1.8. En efecto, dividiendo en primer término m por b hallaremos una expresión del tipo:

$$m = q_0 b + x_0 \ (0 \leq x_0 < b),$$

donde como ya vimos $0 \leq q_0 < m$. Si $q_0 = 0$, entonces m coincide con x_0 y su desarrollo b-ádico consiste de esta única cifra. En otro caso iniciamos un proceso iterativo, dividiendo ahora q_0 por b y escribiendo

$$q_0 = q_1 b + x_1 \ (0 \leq x_1 < b),$$

de donde resulta que

$$m = q_1 b^2 + x_1 b + x_0.$$

Continuando en este plan (ahora dividimos q_1 por b, etc.), vamos obteniendo dos secuencias x_i y q_i de enteros, donde q_i es el cociente de dividir q_{i-1} por b y x_i es el resto, verificándose las relaciones

$$m > q_0 > q_1 > \ldots$$

y

$$m = q_r b^r + x_{r-1} b^{r-1} + \ldots + x_1 b + x_0$$

para $r \geq 1$.

Puesto que la sucesión de cocientes es estrictamente decreciente, es claro que existe $k \in \mathbb{N}$ tal que $q_{k-1} \geq b$ y $0 < q_k < b$. Tomando $x_k = q_k$ y teniendo en cuenta el rango de variación de los x_i deducimos finalmente que

$$m = x_k b^k + x_{k-1} b^{k-1} + \ldots + x_1 b + x_0$$

es el desarrollo b-ádico de m.

Por ejemplo, hallemos el desarrollo en base 9 de 2276. Procediendo como en el caso general obtenemos $2276 = 252 \,.\, 9 + 8$, $252 = 28 \,.\, 9$ y $28 = 3 \,.\, 9 + 1$. Luego, $2276 = (3108)_9$.

En cuanto al número de cifras, recordemos que k está caracterizado por satisfacer la doble desigualdad $b^k \leq m < b^{k+1}$. Si tomamos logaritmos en base b, resulta que

$$k \leq \log_b m < k + 1,$$

esto es, k es la parte entera del logaritmo en base b de m. En definitiva, si designamos por $N_b(m)$ el número de cifras b-ádicas de m, vale la fórmula:

$$N_b(m) = [\log_b m] + 1. \quad \Diamond$$

Un párrafo especial merece el *sistema binario* (base 2) de numeración. El mismo utiliza sólo los símbolos (*bits*) 0 y 1, y es, debido a su sencillez, el lenguaje que emplean las computadoras para almacenar y manejar sus datos. El hecho es que las operaciones en este sistema responden a instrucciones muy fáciles de implementar electrónicamente: simplemente, la suma de dos cifras binarias es 1 si son distintas y 0 si son iguales (naturalmente, el caso $1+1$ exige un transporte de bit), mientras que su producto es 1 sólo si ambas son iguales a 1.

Otros dos sistemas de numeración utilizados en informática son el *octal* (base 8) y el *hexadecimal* (base 16), que usa como símbolos los diez dígitos y las seis letras $\mathrm{A}, \mathrm{B}, \cdots, \mathrm{F}$. Por tratarse de potencias de 2, la conversión de binario a octal ó a hexadecimal es particularmente simple y no requiere un pasaje intermedio a base 10. Por ejemplo:

$$(1011011001)_2 = (1331)_8 = (2\mathrm{D}9)_{16}\,.$$

Lo que hemos hecho es agrupar las cifras de a 3, en el primer caso, y de a 4 en el segundo, siempre de derecha a izquierda. Cada uno de esos bloques determina un número menor que 8 ó 16, según el caso.

Criterios de divisibilidad.

La expansión decimal de un número −más generalmente su desarrollo en cualquier base− no sólo sirve para reconocerlo, sino que en algunos casos también nos informa rápidamente sobre sus relaciones de divisibilidad con otros números. Por ejemplo, sabemos que es divisible por 2 si su última cifra es par, que lo es por 5 si la misma es 0 ó 5, y que los múltiplos de 10 son aquellos cuyo desarrollo decimal termina en 0. Naturalmente, todas estas afirmaciones pueden demostrarse, y cada una de ellas configura lo que llamamos un *criterio de divisibilidad.*

En rigor, un criterio de divisibilidad no es otra cosa que un teorema, que establece una condición necesaria y suficiente para que un número a divida a otro b. Se trata usualmente de una condición sobre las cifras de b, que deberá poder examinarse fácilmente para que el criterio resulte eficaz. Si bien los modernos recursos de cálculo han tornado algo obsoletos a los criterios de divisibilidad, hemos decidido incluirlos brevemente en estas páginas por dos razones: en primer lugar siguen siendo útiles, por ejemplo para descartar rápidamente a ciertos números como posibles divisores de otro cuya primalidad se está estudiando, y por otro lado queremos mostrarle cómo se justifican algunos de los más conocidos, con la finalidad didáctica de ilustrar los conceptos que venimos desarrollando. Manos a la obra pues.

Si

$$c = \sum_{i=0}^{n} c_i 10^i$$

es el desarrollo decimal de un número natural c, consideremos cualquier índice r entre 1 y n. Agrupando convenientemente, tenemos:

$$c = 10^r \sum_{i=r}^{n} c_i 10^{i-r} + \sum_{i=0}^{r-1} c_i 10^i.$$

El segundo término del miembro de la derecha, que designaremos por c_r, difiere de c en un múltiplo de 10^r y es menor que 10^r (tiene a lo sumo r cifras decimales), luego es el resto de dividir c por 10^r. O sea, hemos probado que el número determinado por las últimas r cifras de un número es el resto de dividirlo por 10^r. Este resultado, que tiene interés en sí mismo, implica a su vez el familiar criterio:

(C_1) $10^r \mid c$ si y sólo el desarrollo decimal de c termina en r ceros.

Teniendo en cuenta que $10^r = 2^r \,.\, 5^r$, empleando propiedades elementales de la divisibilidad obtenemos asimismo los siguientes criterios:

(C_2) $2^r \mid c$ si y sólo si $2^r \mid c_r$.
(C_3) $5^r \mid c$ si y sólo si $5^r \mid c_r$.

Por ejemplo, 13456 es múltiplo de 8, pues 456 lo es, mientras que 7085 es múltiplo de 5 pero no de 25, ya que $25 \nmid 85$.

Es claro por qué es posible obtener criterios de divisibilidad por potencias de 2, 5 y 10 tan simples: 10 es la base del sistema de numeración y 2 y 5 son divisores de 10. Si trabajáramos en otra base b, estos números perderían su posición privilegiada, que sería ocupada por b y sus divisores. Los resultados de arriba se trasladan sin ninguna dificultad a la nueva situación. Por ejemplo, en base 16 un número es múltiplo de 4 si y sólo si su última cifra es 0, 4, 8 ó C, mientras que en el sistema binario un número es múltiplo de 64 si y sólo si sus últimos 6 bits son nulos.

Existen otros criterios de uso corriente basados en el sistema decimal, como los de divisibilidad por 3, 9 y 11, que desarrollaremos a continuación.

Puesto que el resto de dividir 10 por 9 es 1, deducimos de la proposición 5.1.6 la validez de la fórmula

$$r_9(c) = r_9(c_n + c_{n-1} + \cdots + c_1 + c_0),$$

esto es, para hallar el resto de dividir cualquier número por 9, basta calcular el resto de dividir por 9 la suma de sus cifras. Por ejemplo, el resto de dividir 174218 por 9 es 5, ya que la suma de sus dígitos es 23 y éste tiene resto 5 al dividirlo por 9. Notemos que sólo hemos usado que el resto de dividir 10 por 9 es 1, por lo que vale idéntico resultado para la división por 3. En particular, obtenemos los conocidos criterios de divisibilidad por 3 y por 9:

(C_4) $3 \mid c$ si y sólo si 3 divide a la suma de las cifras de c .
(C_5) $9 \mid c$ si y sólo si 9 divide a la suma de las cifras de c .

Similarmente, puesto que $10 \equiv -1 \ (11)$ y por lo tanto $10^k \equiv (-1)^k \ (11)$ para todo $k \in \mathbb{N}$, podemos deducir un criterio de divisibilidad por 11. En efecto, aplicando nuevamente 5.1.6 obtenemos en este caso

$$r_{11}(c) = r_{11}\left(\sum_{k=0}^{n}(-1)^k c_k\right),$$

vale decir, el resto de dividir un número por 11 coincide con el resto de dividir por 11 el número obtenido sumando y restando alternativamente sus cifras decimales. Por ejemplo, el resto de dividir 3452 por 11 es 9, ya que la suma alternada de sus cifras es -2. Como en los casos anteriores, esto brinda un criterio de divisibilidad por 11:

(C_6) $11 \mid c$ si y sólo si 11 divide a la suma alternada de las cifras de c.

Observando que 1001 es múltiplo de 7 y de 13, se pueden obtener criterios similares de divisibilidad por 7 y por 13. En ambos casos, no es la suma alternada de las cifras la que ejerce el control, sino que se separa el número en bloques de a 3 cifras, y luego se los suma alternadamente. Por ejemplo, 12322843 es múltiplo de 13, ya que

$$843 - 322 + 12 = 533 = 41 \,.\, 13\,.$$

En un futuro ejercicio, encomendaremos al lector la justificación de estos criterios. Por supuesto que pueden establecerse muchos otros criterios, pero tengamos en claro que son pocos los realmente útiles, ya que la aplicación de algunos podría resultar tanto ó más complicada que la división efectiva. Imaginemos por un momento la ineficacia del siguiente criterio (válido) de divisibilidad por 99990001: se procede como en el caso de 7 ó 13, pero separando al número en bloques de ¡12 cifras! $\Diamond$

Desarrollo de un número real.

Suponiendo siempre que b es un número natural mayor que uno, veamos que la idea de desarrollo b-ádico se extiende a números reales.

Proposición 5.1.9 Si $x \in [0, 1)$, existe un único par sucesiones $(x_n)_{n\geq 1}$ de números enteros y $(\alpha_n)_{n\geq 1}$ de números reales satisfaciendo, para todo $k \in \mathbb{N}$, las siguientes condiciones:

i) $0 \leq x_k < b$

ii) $0 \leq \alpha_k < 1$

iii) $x = \sum_{i=1}^{k} x_i b^{-i} + \alpha_k b^{-k}$.

DEMOSTRACION. Definiremos inductivamente ambas sucesiones. Para ello, comencemos por tomar $x_1 = [bx]$ y $\alpha_1 = \{bx\}$. Dividiendo por b la igualdad $bx = x_1 + \alpha_1$, obtenemos entonces

$$x = x_1 b^{-1} + \alpha_1 b^{-1} .$$

Por otro lado, notemos que $0 \leq x_1 \leq bx < b$, por hipótesis, y $0 \leq \alpha_1 < 1$ por definición de mantisa. Luego, x_1 y α_1 satisfacen los requerimientos del enunciado para $k = 1$.

Supongamos ahora que $m \in \mathbb{N}$ y que han sido definidas secuencias $(x_i)_{i \leq m}$ y $(\alpha_i)_{i \leq m}$ de tal manera que se verifican las condiciones i), ii) y iii) para todo $k \leq m$. Definimos entonces $x_{m+1} = [b\alpha_m]$ y $\alpha_{m+1} = \{b\alpha_m\}$. En forma similar a la anterior, resulta inmediatamente que $0 \leq x_{m+1} < b$, $0 \leq \alpha_{m+1} < 1$ y

$$\alpha_m = x_{m+1} b^{-1} + \alpha_{m+1} b^{-1} .$$

En consecuencia,

$$\begin{aligned} x &= \sum_{i=1}^{m} x_i b^{-i} + \alpha_m b^{-m} = \sum_{i=1}^{m} x_i b^{-i} + x_{m+1} b^{-(m+1)} + \alpha_{m+1} b^{-(m+1)} = \\ &= \sum_{i=1}^{m+1} x_i b^{-i} + \alpha_{m+1} b^{-(m+1)} , \end{aligned}$$

y por lo tanto también se satisfacen las condiciones i), ii) y iii) para $k = m+1$, como queríamos demostrar.

Respecto de la unicidad, supongamos que las secuencias (t_n) y (β_n) satisfacen las mismas condiciones que las secuencias (x_n) y (α_n) que hemos construido. Probaremos por inducción que $t_n = x_n$ y $\beta_n = \alpha_n$ para todo $n \in \mathbb{N}$. Como primer paso, observemos que de la igualdad

$$x = t_1 b^{-1} + \beta_1 b^{-1}$$

se deduce que $bx = t_1 + \beta_1$. Siendo $t_1 \in \mathbb{Z}$ y $0 \leq \beta_1 < 1$, resulta que t_1 es la parte entera de bx y β_1 su mantisa. Vale decir, $t_1 = x_1$ y $\beta_1 = \alpha_1$.

Para el paso inductivo, supongamos que $t_i = x_i$ y $\beta_i = \alpha_i$ para todo $i \leq m$. Ahora bien, considerando las condiciones iii) que satisfacen ambos pares de secuencias para el caso $k = m + 1$, tenemos:

$$\begin{aligned} x &= \sum_{i=1}^{m} t_i b^{-i} + t_{m+1} b^{-(m+1)} + \beta_{m+1} b^{-(m+1)} \\ &= \sum_{i=1}^{m} x_i b^{-i} + t_{m+1} b^{-(m+1)} + \beta_{m+1} b^{-(m+1)} \\ &= x - x_{m+1} b^{-(m+1)} - \alpha_{m+1} b^{-(m+1)} + t_{m+1} b^{-(m+1)} + \beta_{m+1} b^{-(m+1)} , \end{aligned}$$

de donde

$$x_{m+1} b^{-(m+1)} + \alpha_{m+1} b^{-(m+1)} = t_{m+1} b^{-(m+1)} + \beta_{m+1} b^{-(m+1)} ,$$

ó equivalentemente,

$$x_{m+1} + \alpha_{m+1} = t_{m+1} + \beta_{m+1}\,.$$

Siendo x_{m+1} y t_{m+1} números enteros y $0 \leq \alpha_{m+1}, \beta_{m+1} < 1$, sigue que $t_{m+1} = x_{m+1}$ y $\beta_{m+1} = \alpha_{m+1}$, lo que concluye la prueba por inducción. $\Diamond$

Conservando las notaciones anteriores, resulta sencillo establecer a partir de la proposición 5.1.9 el siguiente resultado (suponemos que el lector está familiarizado con los conceptos básicos del cálculo infinitesimal):

Proposición 5.1.10

$$x = \sum_{n=1}^{\infty} x_n b^{-n}\,. \tag{5.7}$$

DEMOSTRACION. Si $m \in \mathbb{N}$, designemos por s_m la m-ésima suma parcial de la serie. Resulta entonces que

$$0 \leq x - s_m = \alpha_m b^{-m} < b^{-m} \leq 2^{-m}\,.$$

Por lo tanto

$$\lim_{m\to\infty} s_m = x,$$

lo que prueba nuestro enunciado. $\Diamond$

Hagamos notar que $x_i = 0$ para todo i si $x = 0$ y, más generalmente, que $x_i = 0$ para todo $i > m$ si $\alpha_m = 0$. En tal caso, la expresión (5.7) se reduce a una suma finita.

Hemos trabajado hasta aquí con números reales pertenecientes al intervalo $[\,0, 1)\,$. Para extender la situación a cualquier número real θ no negativo, designemos por q y x la parte entera y la mantisa de θ, respectivamente. Como sabemos, q admite una expansión finita de la forma

$$q = \sum_{i=0}^{r} q_i b^i \ (r \in \mathbb{N}_0),$$

mientras que x admite el desarrollo

$$x = \sum_{n=1}^{\infty} x_n b^{-n}\,.$$

Por lo tanto, sumando ambas igualdades obtenemos:

$$\theta = q + x = \sum_{i=0}^{r} q_i b^i + \sum_{n=1}^{\infty} x_n b^{-n}\,. \tag{5.8}$$

La expresión (5.8) se llama el desarrollo b-ádico de θ, o también el desarrollo en base b de θ. Como en el caso bien conocido del desarrollo decimal, emplearemos la notación

$$\theta = (q_r\, q_{r-1}\, \ldots\, q_1\, q_0\, , x_1\, x_2\, x_3\, \ldots\ldots)_b\,.$$

Mostraremos algunos casos particulares, encomendando al lector la tarea de verificar que los cálculos son correctos.

Ejemplos 5.1.11 Tomemos primero $\theta = \pi$. Entonces

$$\begin{aligned}\pi &= (3, 1\,4\,1\,5\,9\,2\,6\,5\,3\,5\ldots)_{10}\\ \pi &= (1\,1, 0\,0\,1\,0\,0\,1\,0\,0\,0\ldots)_2\\ \pi &= (3, 0\,6\,6\,3\,6\,5\,1\,4\,3\,2\ldots)_7\,.\end{aligned}$$

Naturalmente sólo hemos calculado las primeras cifras. Los puntos suspensivos indican que el desarrollo continúa.

Los siguientes son los desarrollos de 997/45 en base 15 y de 533/1008 en base 4:

$$\begin{aligned}997/45 &= (17, 2\,5)_{15}\\ 533/1008 &= (0, 2\,0\ \langle 1\,3\,1\rangle)_4\,.\end{aligned}$$

Como se ve, hemos introducido algunas convenciones de notación. En el primer caso, todas las cifras a partir de la tercera después de la coma son nulas, por lo que las obviamos, mientras que en el segundo la notación $\langle 1\,3\,1\rangle$ indica que el bloque de cifras 1 3 1 se repite infinitamente a partir del tercer lugar. $\Diamond$

Lo anterior sugiere la siguiente definición: diremos que el desarrollo de θ es *periódico* si y sólo si existen m y p en $\mathbb{N}$ tales que $x_{k+p} = x_k$ para todo $k \geq m$ y p es mínimo con respecto a esta propiedad. Más gráficamente, el desarrollo en base b de θ es del tipo

$$\theta = (q_r\, \ldots\, q_0, x_1\, x_2\, \ldots\, x_{m-1}\ \langle x_m\, \ldots\, x_{m+p-1}\rangle)_b\,.$$

La longitud p del bloque $x_m\, x_{m+1}\, \ldots\, x_{m+p-1}$ que se repite cíclicamente en el desarrollo se llama el *período* de θ. Por ejemplo, $c = 533/1008$ tiene período 3 en base 4. Obviamente, todo desarrollo finito es periódico.

La importancia de los desarrollos periódicos quedará claramente reflejada en el siguiente teorema. Como antes, θ designará un número real no negativo y b cualquier base de numeración.

Teorema 5.1.12 El desarrollo en base b de θ es periódico si y sólo si $\theta \in \mathbb{Q}$.

DEMOSTRACION Puesto que un número real es racional si y sólo si su mantisa lo es, podemos asumir que $0 < \theta < 1$, y por lo tanto su desarrollo b-ádico es del tipo $\theta = 0, x_1\, x_2 \ldots$

Supongamos en primer término que el desarrollo b-ádico de θ es periódico y de período t, digamos

$$\theta = (0, x_1\, x_2\, \ldots\, x_{r-1}\, \langle x_r\, x_{r+1}\, \ldots\, x_{r+t-1}\rangle)_b\,.$$

Separando la parte no periódica podemos escribir $\theta = \gamma + z$, donde

$$\gamma = \sum_{i=1}^{r-1} x_i b^{-i} \quad \text{y} \quad z = (0, \langle x_r\, x_{r+1}\, \ldots\, x_{r+t-1}\rangle)_b\,.$$

Puesto que claramente γ es racional, bastará probar que z también lo es. En efecto, de acuerdo con su definición y usando propiedades elementales de las series convergentes, tenemos:

$$\begin{aligned} z &= \sum_{i=1}^{t} x_{i+r-1} b^{-i} + \sum_{i=t+1}^{\infty} x_{i+r-1} b^{-i} = \\ &= \sum_{i=1}^{t} x_{i+r-1} b^{-i} + b^{-t} z\,, \end{aligned}$$

de donde resulta que

$$z = \left(\frac{b^t}{b^t - 1}\right) \sum_{i=1}^{t} x_{i+r-1} b^{-i} \in \mathbb{Q}\,.$$

Recíprocamente, supongamos ahora que $\theta \in \mathbb{Q}$, digamos $\theta = u/v$, donde u y v son enteros tales que $0 < u < v$. Empleando las notaciones de la proposición 5.1.9, resulta en este caso que x_1 es el cociente de dividir bu por v y $\alpha_1 = r_1/v$, donde r_1 es el resto de dicha división. Argumentando en forma inductiva, se prueba fácilmente en general que α_n es de la forma r_n/v, donde r_n es el resto de dividir $b\, r_{n-1}$ por v y x_n es el cociente de dicha división.

Puesto que sólo hay un número finito de restos de dividir por v, existirán números naturales $i < j$ tales que $\alpha_j = \alpha_i$, en cuyo caso tendremos

$$\alpha_{j+1} = \alpha_{i+1}\,,\ \alpha_{j+2} = \alpha_{i+2}\,, \ldots$$

Por lo tanto el desarrollo de θ en base b es periódico. $\Diamond$

Ejemplo 5.1.13 En el siguiente esquema mostraremos los pasos del desarrollo en base 5 de $454/775$ (por comodidad notacional tomamos $r_0 = u$.)

$r_0 = 454$	$5r_0 = 2270$	$x_1 = 2$
$r_1 = 720$	$5r_1 = 3600$	$x_2 = 4$
$r_1 = 500$	$5r_2 = 2500$	$x_3 = 3$
$r_3 = 175$	$5r_3 = 875$	$x_4 = 1$
$r_4 = 100$	$5r_4 = 500$	$x_5 = 0$
$r_5 = \mathbf{500}$	$5r_5 = 2500$	$x_6 = 3$

Como puede verse hemos destacado la primera repetición de un resto, que marca el inicio del período. Resulta entonces que

$$\frac{454}{775} = (0, 2\,4\ \langle 3\,1\,0\rangle)_5\,.$$

NOTA. En general, dada cualquier secuencia $(t_i)_{i\geq 1}$ de enteros no negativos tales que $t_i < b$ para todo i, la serie

$$\sum_{i=1}^{\infty} t_i b^{-i} \tag{5.9}$$

es convergente, ya que sus sumas parciales están acotadas. En efecto, usando la fórmula de la suma de los términos de una progresión geométrica resulta que

$$\sum_{i=1}^{n} t_i b^{-i} \leq (b-1)\sum_{i=1}^{n}\left(\frac{1}{b}\right)^i = (b-1)\frac{b^n-1}{(b-1)b^n} = \frac{b^n-1}{b^n} < 1$$

para todo $n \in \mathbb{N}$, lo que nos permite concluir además que la misma converge a un número real x perteneciente al intervalo $[0, 1]$.

Sin embargo, la secuencia de cifras del desarrollo $b-$ádico de x no necesariamente coincide con la secuencia (t_i). Por ejemplo, tomemos $b = 2$, $t_1 = 0$ y $t_i = 1$ para $i \geq 2$. Entonces, empleando resultados elementales de series geométricas obtenemos:

$$x = \sum_{i=2}^{\infty}\left(\frac{1}{2}\right)^i = \frac{1}{4}\sum_{i=0}^{\infty}\left(\frac{1}{2}\right)^i = \frac{1}{4}\,\frac{1}{1-\frac{1}{2}} = \frac{1}{2}\,,$$

mientras que claramente $1/2 = (0, 1)_2$. En uno de los ejercicios que siguen encomendaremos al lector la tarea de estudiar en qué casos dos series distintas del tipo (5.9) convergen al mismo número real ♢

5.1.3. Ejercicios

1. En cada uno de los siguientes casos decidir si el enunciado es verdadero o falso (las letras designan números enteros):

a) $a \mid b+c \Rightarrow a \mid b$ y $a \mid c$

b) $a \mid b \Leftrightarrow ac \mid bc$

c) $a \mid bc \Rightarrow a \mid b$ ó $a \mid c$

d) $a \mid c$ y $b \mid c \Rightarrow ab \mid c$

e) $a \mid b$ y $c \mid d \Rightarrow ac \mid bd$

f) $a \mid b \Rightarrow a^m \mid b^m$ para todo $m \geq 0$.

2. Resolver en $\mathbb{Z}^2$ las ecuaciones $a^2 + ab + 1 = 0$ y $a^2 - ab + 2 = 0$.

3. Sean $a, b \in \mathbb{Z}$ y sea $m \in \mathbb{N}$. Probar:

   a) $a - b \mid a^m - b^m$

   b) $a + b \mid a^m - b^m$ si m es par

   c) $a + b \mid a^m + b^m$ si m es impar.

4. En cada uno de los siguientes casos determinar los $n \in \mathbb{N}$ que satisfacen la relación planteada:

   a) $n \mid n + 1$

   b) $n - 2 \mid 3n + 2$

   c) $n - 3 \mid n^2 + 4$.

5. *a*) Dados $a, b \in \mathbb{Z}$ $(a \leq b)$, determinar el número de elementos enteros del intervalo $[a, b)$.

   b) Resolver la misma cuestión que en *a*) para $a, b \in \mathbb{Q}$.

   c) Dados a y b como en *a*) y dado $n \in \mathbb{N}$, determinar cuántos múltiplos de n hay en el intervalo $[a, b)$.

   d) ¿Cuántos números de 4 cifras son divisibles por 13?

6. Probar que todo número entero es de la forma $2k$ o de la forma $2k+1$, para algún $k \in \mathbb{Z}$.

7. Demostrar las siguientes propiedades:

   a) La suma de 3 enteros consecutivos es múltiplo de 3.

   b) La suma de 2 enteros impares consecutivos es divisible por 4.

   c) La suma de 2 enteros pares consecutivos no es divisible por 4.

   d) El producto de dos números pares consecutivos es divisible por 8.

8. Si $n \in \mathbb{N}$, probar que el producto de n enteros consecutivos cualesquiera es divisible por $n!$. Probar que $\binom{2n}{n}$ es par.

9. Hallar el cociente y el resto de dividir por 96 y por -96 los números 527, -714, 29, 0 y -14.

10. Sean $a, m, n \in \mathbb{N}\,(a \geq 2)$. Probar que $a^m - 1 \mid a^n - 1$ si y solo si $m \mid n$.

11. Las edades de 3 personas son números consecutivos, que divididos por 9 arrojan cocientes 5, 5 y 6, respectivamente. ¿Cuántos años tiene el menor?

12. Sea $n \in \mathbb{N}$ tal que el cociente de dividir n por 29 es 5 y el resto de dividir $2n + 9$ por 29 es 3. Determinar n.

13. Sean $a, b, n \in \mathbb{N}$. Calcular el cociente y el resto de la división de b por a en cada uno de los siguientes casos:

    a) $b = 5a + 4$

    b) $b = a^3 - 5$

    c) $a = n + 1$, $b = 2n^2 + 3$

    d) $a = n^2 + 1$, $b = n + 3$.

14. Dado $n \in \mathbb{N}$, analizar la validez de las siguientes afirmaciones:

    a) La suma de n números enteros consecutivos es divisible por n.

    b) El producto de n números enteros consecutivos es divisible por n.

15. Sea $x \in \mathbb{Z}$ tal que $r_{12}(x) = 7$. Calcular:

    a) $r_{12}\left(x^2 - 2x + 5\right)$

    b) $r_4(-x)$

    c) $r_6(2 - 3x)$

    d) $r_{24}\left(x^2 + 6\right)$

    e) $r_{60}(5x^2 - 11)$

    f) $r_{42}\left(42x^{90} - 33\right)$.

16. Determinar $r_7\left(\sum_{i=0}^{100} 2^i\right)$. ¿El cociente es par o impar?

17. *a*) Calcular $r_{13}\left(4^{651}\right)$.
    b) Sea $k \in \mathbb{N}$ tal que $r_{13}\left(4^k\right) = 9$. Determinar $r_6(k)$.

18. Sean $a, b, c \in \mathbb{Z}$ tales que $a^2 + b^2 = c^2$. Probar:
    a) $2 \mid a$ ó $2 \mid b$.
    b) $4 \mid a$ ó $4 \mid b$.
    c) $5 \mid abc$.
    d) $7 \mid c$ si y solo si $7 \mid a$ y $7 \mid b$.

19. Desarrollar en base 2, 7 y 16 los números 254, 1023 y 2401.

20. Sea $b > 1$ y sea $n \in \mathbb{N}$. Determinar las cifras del desarrollo $b-$ádico de
$$\text{i) } b^n \ , \ \text{ii) } b^n - 1 \ , \ \text{iii) } \frac{b^n - 1}{b - 1} \ , \ \text{iv) } \sum_{i=0}^{2n} (-1)^i b^i .$$

21. Sea $(F_k)_{k \geq 0}$ la sucesión de Fibonacci.
    a) Si $m \in \mathbb{N}$, sea $S = (s_k)_{k \geq 0}$ la sucesión de Fibonacci módulo m, es decir, $s_k = r_m(F_k)$. Probar que S es periódica.
    b) Deducir de *a*) que (F_k) contiene infinitos múltiplos de m.
    c) ¿Para qué valores de k termina en 0 el desarrollo decimal de F_k?

22. Justificar los criterios de divisibilidad por 7 y 13 mencionados en el texto y hallar, demostrando su validez, un criterio de divisibilidad por 101.

23. Establecer criterios de divisibilidad por 3, 7 y 9 a partir del desarrollo binario de un número.

24. Hallar las primeras 7 cifras de los desarrollos de $\sqrt{3}$ y de $\sqrt{5}$ en bases 2, 8 y 16.

25. Si $b > 1$ y $m \in \mathbb{N}$, hallar los desarrollos $b-$ádicos de b^{-1}, b^{-m} y $(b-1)^{-1}$. Probar que la serie
$$\sum_{i=2}^{\infty} (b-1) b^{-i}$$
es convergente y calcular su suma.

26. Hallar el desarrollo de $2/3$ en bases 2 y 3, y de $19/62$ y $43/62$ en base 5.

27. Dado $b > 1$, hallar $x \in \mathbb{Q}$ tal que:

$$\text{i) } x = (0, 0\ \langle 1 \rangle)_b \quad , \quad \text{ii) } x = (0, \langle 0\,1 \rangle)_b \quad , \quad \text{iii) } x = (0, \langle 1\,0 \rangle)_b \ .$$

28. Sea $b > 1$.

   a) Probar que $x \in \mathbb{Q}$ tiene desarrollo $b-$ádico finito si y solo si x es de la forma a/b^m, con $a, m \in \mathbb{N}_0$.

   b) Caracterizar los números racionales x cuyo desarrollo en base b tiene período 1.

29. Sea $b > 1$ y sea $x \in [\,0, 1\,]$ tal que x se expresa de dos maneras distintas como suma de una serie de tipo (5.9), digamos

$$x \ = \ \sum_{i=1}^{\infty} u_i\, b^{-i} \ = \ \sum_{i=1}^{\infty} v_i\, b^{-i} \, .$$

   a) Sea $s = \min\,\{i \in \mathbb{N} \,:\, u_i \neq v_i\}$ y supongamos sin pérdida de generalidad que $v_s < u_s$. Probar que $u_s = v_s + 1$.

   b) Demostrar que $u_i = 0$ y $v_i = b - 1$ para todo $i > s$.

   c) Deducir que x es racional y admite una representación como cociente de enteros cuyo denominador es una potencia de b .

   d) Recíprocamente, probar que dos series de tipo (5.9) cuyos coeficientes satisfacen las condiciones *a*) y *b*) tienen igual suma.

5.2. Máximo común divisor

5.2.1. Definición y método de cálculo

> Si a y b son números enteros y $a \neq 0$ ó $b \neq 0$, definimos el *máximo común divisor* de a y b como el mayor de sus divisores positivos comunes, que notaremos $(a : b)$. Para completar todos los casos definimos $(0 : 0) = 0$.

Obsérvese la buena definición del máximo común divisor (abreviadamente mcd), habida cuenta de que todo entero no nulo admite un número finito de divisores. Calculemos por ejemplo $(24 : 132)$. Una simple prueba por ensayo y error nos muestra que los divisores positivos de 24 son 1, 2, 3, 4, 6, 8, 12 y 24, siendo 1, 2, 3, 4, 6 y 12 los que también dividen a 132. Luego, $(24 : 132) = 12$. Veamos algunas propiedades elementales del mcd :

Proposición 5.2.1 Si $a, b \in \mathbb{Z}$, son válidas las siguientes propiedades:

1) $(a : b) = (b : a)$

2) $(a : b) \geq 0$ y $(a : b) = 0 \Leftrightarrow a = b = 0$

3) $(a : b) = (|\, a \,| : |\, b \,|)$

4) $a \mid b$ si y sólo si $(a : b) = |\, a \,|$

5) $(a : b) = (a : b - a)$

6) Si $a \neq 0$ entonces $(a : b) = (a : r_a(b))$.

Demostracion. Las propiedades 1) y 2) son consecuencia inmediata de la definición, mientras que 3) sigue del hecho de que todo número entero tiene los mismos divisores que su valor absoluto. En cuanto a 4), si $a \mid b$ resulta por transitividad que los divisores comunes de a y b son exactamente los divisores de a, el mayor de los cuales es obviamente $|\, a \,|$. La recíproca se obtiene teniendo en cuenta que en cualquier caso $(a : b) \mid b$. La propiedad 5) es consecuencia del hecho de que los divisores comunes de a y b coinciden con los divisores comunes de a y $b - a$, de acuerdo con los ítems 6) y 7) de la proposición 5.1.1. Finalmente, el mismo argumento se aplica para probar 6), escribiendo $b = qa + r_a(b)$. $\quad \Diamond$

El lector podrá advertir que en el sencillo ejemplo de arriba no hemos empleado ningún método sistemático de cálculo, que sólo nos hemos manejado por ensayo y error. Es obvio que ello fue posible debido al reducido tamaño de los números involucrados, pero muy distinta sería la situación si tuviéramos que calcular el máximo común divisor de dos números muy grandes, considerando que no conocemos (en realidad no existe) un algoritmo eficiente que permita listar sus divisores. Veremos sin embargo ahora que esto último no representa una dificultad, ya que mostraremos un método de

cálculo que nos brindará el mcd de dos números sin conocer previamente sus divisores, simplemente a través de restas. Como regalo adicional, el mismo nos revelará una propiedad fundamental del mcd, que también se observa en el ejemplo: no sólo es el mayor de los divisores comunes, *sino que además es múltiplo de todos ellos.*

Supongamos entonces que queremos calcular $(a : b)$, siendo a y b números naturales (basta estudiar este caso). Para ello, definimos inductivamente una sucesión de pares (a_i, b_i) de enteros no negativos a través de las siguientes reglas de recurrencia:

$$(a_1, b_1) = (a, b)$$
$$(a_{k+1}, b_{k+1}) = \begin{cases} (a_k, b_k - a_k) \text{ si } a_k \leq b_k \\ (b_k, a_k - b_k) \text{ si } a_k > b_k \,. \end{cases}$$

Por las mismas razones que en el ítem 5) de la proposición 5.2.1, sigue por inducción que el conjunto de divisores comunes de a y b coincide con el conjunto de divisores comunes de a_k y b_k, y por lo tanto $(a : b) = (a_k : b_k)$ para todo $k \in \mathbb{N}$. Veamos que la secuencia que hemos definido nos conduce directamente al mcd de a y b.

Para ello, designemos por s_n la suma $a_n + b_n$. Si $k \in \mathbb{N}$ y suponiendo por ejemplo que $a_k \leq b_k$, tenemos:

$$s_{k+1} = a_{k+1} + b_{k+1} = a_k + (b_k - a_k) = b_k \leq a_k + b_k = s_k \,,$$

esto es, la sucesión (s_n) es decreciente (el caso $a_k > b_k$ es similar). Puesto que una sucesión infinita de enteros no negativos no puede ser estrictamente decreciente, concluimos que existe un mínimo $t \in \mathbb{N}$ tal que $s_{t+1} = s_t$, en cuyo caso $a_t = 0$ ó $b_t = 0$. Sigue entonces por definición que $a_{t-1} = b_{t-1}$, y por lo tanto

$$(a : b) = a_{t-1} \,.$$

En resumen, siempre se llega a un par cuyas componentes son iguales, y dicha componente común es $(a : b)$.

Ejemplo 5.2.2 Veamos como funciona el algoritmo en el caso de 24 y 132. Aplicando sucesivamente las reglas de recurrencia, tenemos:

$$\begin{aligned} (24 : 132) &= (24 : 108) = (24 : 84) \\ &= (24 : 60) = (24 : 36) \\ &= (24 : 12) = (12 : 12) = 12. \quad \Diamond \end{aligned}$$

Como se puede apreciar, sólo basta restar. Volviendo al caso general, resulta en particular que el conjunto de divisores comunes de a y b es el conjunto de divisores comunes de a_{t-1} y b_{t-1}, que no es otra cosa que el conjunto de divisores de a_{t-1}, pues $a_{t-1} = b_{t-1}$. Se cumple luego en general

la propiedad que conjeturamos más arriba: $(a : b)$ es múltiplo de todos los divisores comunes de a y b, situación que claramente se mantiene en el caso de que a y b sean dos enteros cualesquiera, no necesariamente positivos. Esto permite obtener una definición equivalente de mcd, que enunciamos en la siguiente proposición:

Proposición 5.2.3 El máximo común divisor de dos números enteros a y b es el único número entero no negativo d verificando las condiciones:

$(MCD)_1$ $\quad d \mid a$ y $d \mid b$

$(MCD)_2$ $\quad d$ es múltiplo de todo divisor común de a y b.

Demostracion. El resultado es inmediato si $a = b = 0$, por lo que supondremos que alguno de los números es no nulo. Si $d = (a : b)$, ya sabemos hasta aquí que d satisface $(MCD)_1$ y $(MCD)_2$, por lo que nos ocuparemos de la unicidad. Comencemos observando que $-d$ también verifica ambas condiciones, ya que un número entero tiene los mismos divisores y múltiplos que su inverso aditivo. Finalmente, supongamos en general que t es un entero que verifica las condiciones del enunciado. Puesto que en particular t es un divisor común de a y b, sigue por $(MCD)_2$ que $t \mid d$. Intercambiando los roles de ambos números resulta asimismo que $d \mid t$, y por lo tanto $t = \pm d$.

En conclusión, d y $-d$ son los únicos enteros que verifican $(MCD)_1$ y $(MCD)_2$, siendo d el único no negativo. $\Diamond$

Nota. La definición de mcd se extiende sin dificultad a familias de tres o más enteros. Precisamente, dados enteros $a_1, a_2, \ldots, a_n$, definimos inductivamente su máximo común divisor mediante la recurrencia

$$(a_1 : a_2 : \ldots : a_n) = ((a_1 : a_2 : \ldots : a_{n-1}) : a_n)\,.$$

Por ejemplo,

$$\begin{aligned}(24 : 18 : 42 : 33) &= ((24 : 18 : 42) : 33) = (((24 : 18) : 42) : 33)\\ &= ((6 : 42) : 33) = (6 : 33) = 3.\end{aligned}$$

Mediante sencillos argumentos inductivos se demuestra sin dificultad que esta generalización del mcd también satisface las condiciones $(MCD)_1$ y $(MCD)_2$, debidamente adaptadas. En los ejercicios del final de la sección le encargaremos esta tarea al lector. $\Diamond$

Algoritmo de Euclides.

En el ejemplo del método de cálculo del mcd que hemos mostrado, podemos observar que 24 fue restado 5 veces (la primera vez de 132), hasta alcanzar un número menor que 24. Estas 5 operaciones pueden reemplazarse por una sola, la división entera de 132 por 24. Si efectivamente lo hacemos,

vemos que el cociente es 5 (por eso debimos restar 5 veces) y el resto es 12, que es justamente el número al que arribamos. Siguiendo esta idea en el caso general se obtiene una forma mucho más compacta del algoritmo, que además pondrá en relieve otra propiedad muy importante del máximo común divisor.

Veamos cómo proceder. Partimos como siempre de dos números naturales a y b y comenzamos dividiendo b por a, obteniéndose un cociente q_1 y un resto r_1 de manera que

$$b = q_1 a + r_1 \,.$$

Si $r_1 = 0$ significa que $a \mid b$, y por lo tanto $(a : b) = a$. Si $r_1 > 0$ continuamos, poniendo énfasis en el hecho de que $(a : b) = (a : r_1)$. En el próximo paso dividimos a por r_1 y obtenemos un cociente q_2 y un resto r_2, dando lugar a la igualdad

$$a = q_2 r_1 + r_2.$$

Razonamos entonces de la misma manera que en el paso anterior. Si $r_2 = 0$ sigue que $r_1 \mid a$, luego $(a : b) = (a : r_1) = r_1$ y hemos concluido. Si $r_2 > 0$, notando que $(a : b) = (a : r_1) = (r_1 : r_2)$, continuamos nuestro algoritmo dividiendo r_1 por r_2, y así sucesivamente. Iterando este proceso de división (en cada paso se divide el divisor anterior por el resto anterior), y razonando cada vez como lo hicimos en los dos primeros pasos, vamos obteniendo una sucesión (q_i) de cocientes y una sucesión (r_i) de restos tales que $r_i = q_{i+2} r_{i+1} + r_{i+2}$ y $(a : b) = (r_i : r_{i+1})$ para todo i (para unificar la notación definimos $r_{-1} = b$ y $r_0 = a$).

Puesto que valen las desigualdades $a > r_1 > r_2 > \ldots$ y una sucesión de números naturales no puede ser infinitamente decreciente, necesariamente existe $n \in \mathbb{N}$ tal que $r_n = 0$. Siendo r_n el resto de dividir r_{n-2} por r_{n-1}, resulta entonces que $(r_{n-2} : r_{n-1}) = r_{n-1}$, esto es,

$$(a : b) = r_{n-1} \,.$$

En otras palabras, el máximo común divisor de a y b es *el último resto no nulo*. $\Diamond$

El proceso de divisiones sucesivas para el cálculo del máximo común divisor que acabamos de describir es el llamado *algoritmo de Euclides*, que sistematiza favorablemente nuestro método anterior de diferencias sucesivas.

Ejemplo 5.2.4 Usemos el algoritmo para calcular $(3724 : 13818)$. Efectuando las correspondientes divisiones obtenemos:

$$\begin{aligned} 13818 &= 3 \,.\, 3724 + 2646 \\ 3724 &= 1 \,.\, 2646 + 1078 \\ 2646 &= 2 \,.\, 1078 + 490 \\ 1078 &= 2 \,.\, 490 + \mathbf{98} \\ 490 &= 5 \,.\, 98 = 5 \,.\, 98 + 0 \,. \end{aligned}$$

Por lo tanto $(3724 : 13818) = 98$. $\Diamond$

Combinaciones lineales.

Vamos a ir ahora al encuentro de otra propiedad muy importante del máximo común divisor, que requiere la introducción del siguiente concepto:

> Si a y b son números enteros, cualquier expresión (cualquier número) de la forma $ax + by$ $(x, y \in \mathbb{Z})$ se llamará una *combinación lineal* de a y b.

Así, 28 es combinación lineal de 5 y 12, ya que $28 = 5\,.\,8 + 12\,.(-1)$, mientras que 30 no es combinación lineal de 8 y 20, pues cualquier número de la forma $8x + 20y$ es múltiplo de 4.

Proposición 5.2.5 (Bezout) Sean $a, b \in \mathbb{Z}$ y sea $d = (a : b)$. Entonces d es combinación lineal de a y b.

DEMOSTRACION. El caso $a = b = 0$ es trivial, por lo que podemos suponer que $a \neq 0$. Conservando las notaciones de la descripción del algoritmo de Euclides, probaremos que cada uno de los restos r_k puede expresarse en la forma

$$r_k = au_k + bv_k, \qquad (*)$$

para ciertos enteros u_k y v_k. Resultará entonces en particular que $d = r_{n-1}$ es de la forma indicada.

Suponiendo que nuestra aserción es falsa, consideremos el mínimo índice m tal que r_m no puede expresarse en la forma $(*)$. Puesto que $b = a\,.\,0 + b\,.\,1$ y $a = a\,.\,1 + b\,.\,0$, resulta que $m > 0$. Por la elección de m, sigue entonces que existen enteros u_{m-2}, v_{m-2}, u_{m-1} y v_{m-1} de manera que se satisfacen igualdades del tipo

$$\begin{aligned} r_{m-2} &= au_{m-2} + bv_{m-2} \\ r_{m-1} &= au_{m-1} + bv_{m-1}\,. \end{aligned}$$

Por otro lado, teniendo en cuenta que $r_{m-2} = q_m r_{m-1} + r_m$, obtenemos:

$$\begin{aligned} r_m = r_{m-2} - q_m r_{m-1} &= au_{m-2} + bv_{m-2} - q_m\,(au_{m-1} + bv_{m-1}) = \\ &= a\,(u_{m-2} - q_m u_{m-1}) + b\,(v_{m-2} - q_m v_{m-1})\,, \end{aligned}$$

vale decir, r_m es de la forma $(*)$, lo que es una contradicción. $\Diamond$

NOTA. La demostración anterior nos muestra una forma recurrente de calcular los coeficientes u_i y v_i, a través de las reglas

$$\begin{cases} u_{-1} = 0\,,\ u_0 = 1 & \\ u_m = u_{m-2} - q_m u_{m-1} & \text{si } m \geq 1 \\ v_{-1} = 1\,,\ v_0 = 0 & \\ v_m = v_{m-2} - q_m v_{m-1} & \text{si } m \geq 1. \end{cases}$$

Ejemplo 5.2.6 Listemos en una tabla los valores de las distintas variables del algoritmo de Euclides en el caso $a = 3724\,,\ b = 13818$:

i	r_i	q_i	u_i	v_i
-1	13818	-	0	1
0	3724	-	1	0
1	2646	3	-3	1
2	1078	1	4	-1
3	490	2	-11	3
4	98	2	26	-7

Por lo tanto,

$$98 = (3724 : 13818) = 3724\,.\,26 + 13818\,.\,(-7)\,. \quad \Diamond$$

La propiedad de Bezout permite caracterizar completamente el conjunto de números enteros que pueden expresarse como combinación lineal de dos enteros dados. Precisamente, vale el siguiente resultado:

Proposición 5.2.7 Si $a, b, c \in \mathbb{Z}$, c es combinación lineal de a y b si y sólo si $(a : b) \mid c$. Deducimos de tal hecho que $(a : b)$ es la *mínima* combinación lineal no negativa de a y b.

DEMOSTRACION. Sea $d = (a : b)$ y supongamos primero que $d \mid c$, digamos $c = dk$. Sabemos (proposición 5.2.5) que existen números enteros r y s tales que $d = ar + bs$. Entonces, multiplicando ambos miembros de esta igualdad por k probamos nuestra primera afirmación, ya que entonces

$$c = dk = a(rk) + b(sk)$$

es combinación lineal de a y b. La recíproca es inmediata, ya que siendo d un divisor común de a y de b sigue por propiedades elementales de la divisibilidad (ver 5.1.1) que d es divisor de $ax+by$ cualesquiera sean $x, y \in \mathbb{Z}$.

Finalmente, la afirmación adicional del enunciado es obvia, ya que todo número entero no negativo es el menor de sus múltiplos no negativos. $\Diamond$

La siguiente propiedad del mcd permite en ocasiones agilizar su cálculo.

Lema 5.2.8 Si $t, u, v \in \mathbb{Z}$ y $t > 0$ entonces $(tu : tv) = t(u : v)$.

DEMOSTRACION. Sea $d = (u : v)$. Puesto que $d \mid u$ y $d \mid v$ sigue que td es un divisor común de tu y tv. Luego $td \mid (tu : tv)$, por $(MCD)_2$. Por otro lado, sean $x, y \in \mathbb{Z}$ tales que $d = ux + vy$. Multiplicando por t obtenemos

$td = tux + tvy$, esto es, td es combinación lineal de tu y tv. Deducimos entonces de 5.2.7 que $(tu : tv) \mid td$.

Hemos probado así que los números td y $t(u : v)$ se dividen mutuamente. Siendo ambos no negativos concluimos que son iguales, como queríamos demostrar. $\Diamond$

Por ejemplo,

$$\begin{aligned}(24 : 132) = (2\,.\,12 : 2\,.\,66) = 2\,.\,(12 : 66) = \\ = 2\,.\,(2\,.\,6 : 2\,.\,33) = 4\,.\,(6 : 33) = \\ = 4\,.\,(3\,.\,2 : 3\,.\,11) = \\ = 12\,.\,(2 : 11) = 12\,.\,1 = 12.\end{aligned}$$

Coprimalidad.

Emplearemos terminología especial para referirnos a un caso muy importante.

Dos enteros a y b se dicen *coprimos* si y sólo si $(a : b) = 1$.

Equivalentemente, 1 es el único divisor común positivo de a y b. Además, dado que en general $(a : b)$ es la menor combinación lineal positiva de a y b, resulta que a y b son coprimos si y sólo si 1 es combinación lineal de a y b.

Para abreviar la notación, si u y v son enteros coprimos escribiremos $u \perp v$. Probaremos a continuación varias propiedades referidas al concepto de coprimalidad, que nos darán ocasión de apreciar su relevancia (a, b y c designan números enteros y m y n números naturales).

Proposición 5.2.9 Son válidos los siguientes hechos:

1) Si $d = (a : b)$ entonces $a/d \perp b/d$

2) Si $a \perp b$ entonces $(a : bc) = (a : c)$

3) Si $a \mid bc$ y $a \perp b$ entonces $a \mid c$

4) Si $a \perp b$ entonces $a^m \perp b^n$

5) $(a^m : b^m) = (a : b)^m$

6) $a \mid b$ si y sólo si $a^m \mid b^m$.

Demostracion.

1) Se deduce de 5.2.8, ya que

$$d = (a : b) = (d \times a/d : d \times b/d) = d\,(a/d : b/d)\,,$$

de donde $(a/d : b/d) = 1$. Notemos la razonabilidad del hecho: al dividir dos números por su mayor divisor común suprimimos en ambos todos los factores comunes.

2) Es inmediato probar que $(a : c) \mid (a : bc)$ cualesquiera sean a, b y c. En nuestro caso, escribiendo $1 = ax + by$ y multiplicando por c obtenemos $c = a(cx) + bcy$. Por lo tanto c es combinación lineal de a y bc y entonces $(a : bc) \mid c$. Puesto que obviamente $(a : bc) \mid a$, resulta finalmente que $(a : bc) \mid (a : c)$. En consecuencia $(a : bc) = (a : c)$, pues se dividen mutuamente.

3) Se obtiene como caso particular de 2), ya que

$$a = (a : bc) = (a : c).$$

4) Si $k > 1$, usando 2) tenemos que $(a : b^k) = (a : bb^{k-1}) = (a : b^{k-1})$. Luego, mediante un argumento inductivo deducimos que $a \perp b^n$ cualquiera sea n. Reiterando el razonamiento (b^n ocupa el rol de a y a^m el de b^n), obtenemos $(a^m : b^n) = (b^n : a^m) = 1$, como queríamos probar.

5) Sea $t = (a : b)$. Entonces

$$(a^m : b^m) = (t^m\,(a/t)^m : t^m\,(b/t)^m) = t^m\,((a/t)^m : (b/t)^m) = t^m,$$

por 1) y 4).

6) Sigue de 5), ya que

$$a \mid b \Leftrightarrow (a : b) = a \Leftrightarrow (a^m : b^m) = a^m \Leftrightarrow a^m \mid b^m. \quad \Diamond$$

Todos estos hechos son de interés, pero queremos destacar especialmente la propiedad 3), que como veremos es la llave del teorema de factorización única. Notemos de paso que la condición de que un número divida alguno de los dos factores de un producto es en general suficiente pero no necesaria para que divida al producto, ya que por ejemplo $8 \mid 4 \times 6$ pero $8 \nmid 4$ y $8 \nmid 6$.

MINIMO COMUN MULTIPLO. Existe un concepto dual al de máximo común divisor, que pasamos a definir:

> Dados $a, b \in \mathbb{Z}$, un *mínimo común múltiplo* de a y b es un número entero m satisfaciendo las dos siguientes condiciones:
>
> $(MCM)_1$ $\quad a \mid m$ y $b \mid m$
>
> $(MCM)_2$ $\quad m$ es divisor de todo múltiplo común de a y b.

Independientemente de la existencia de un tal m, observemos la dualidad a la que aludíamos: el mcm de dos números positivos es el menor de sus múltiplos comunes. Vamos a probar ahora que existe un número entero satisfaciendo las condiciones del enunciado.

Si $a = b = 0$, es trivial ver que 0 es el único entero que satisface $(MCM)_1$ y $(MCM)_2$. Si $a \neq 0$ ó $b \neq 0$, sea $m = ab/d$, donde $d = (a : b)$. Es claro que m es un múltiplo común de a y b, ya que

$$m = a(b/d) = b(a/d)\,.$$

Respecto de $(MCM)_2$, supongamos que t es cualquier múltiplo común de a y b, digamos $t = ka = qb$. Dividiendo la última de estas igualdades por d se obtiene $k(a/d) = q(b/d)$, y por lo tanto a/d divide a $q(b/d)$. Siendo a/d y b/d coprimos, sigue por 3) de la proposición 5.2.9 que a/d divide a q. Escribiendo $q = h(a/d)$, tenemos:

$$t = qb = h(ab/d) = hm,$$

y por lo tanto $m \mid t$, como queríamos demostrar.

Como en el caso del mcd, es fácil probar que existen exactamente dos números enteros satisfaciendo las condiciones $(MCM)_1$ y $(MCM)_2$, a saber, m y $-m$. El único no negativo entre ellos será denominado el mínimo común múltiplo de a y b, al que notaremos $[\,a : b\,]$.

Vale la pena remarcar que en cualquier caso vale la fórmula

Fórmula 5.2.10

$$|\,ab\,| = (a : b)\,[\,a : b\,] \qquad \Diamond$$

Ecuación diofántica lineal.

Estudiaremos la cuestión de determinar todos los pares (x, y) de números enteros que satisfacen la ecuación

$$aX + bY = c, \tag{5.10}$$

donde $a, b, c \in \mathbb{Z}$ y $ab \neq 0$. La llamaremos *ecuación diofántica lineal*, por Diofanto de Alejandría, llamándose en general ecuación diofántica a toda aquella de la cual se buscan las soluciones enteras, ó más generalmente, las soluciones racionales. Analizaremos su resolubilidad, y en caso de que admita soluciones mostraremos la forma general de las mismas.

Respecto a la resolubilidad ya podemos dar respuesta al problema, ya que hallar una solución es equivalente a expresar c como combinación lineal de a y b. Luego:

La ecuación 5.10 es resoluble si y sólo si c es múltiplo de $(a : b)$. $\Diamond$

Suponiendo que se verifica la condición de arriba, veamos cómo caracterizar el conjunto de sus soluciones. Para ello, escribamos $a = d\alpha$, $b = d\beta$ y $c = d\gamma$, donde $d = (a : b)$. Si en 5.10 dividimos ambos miembros por d, arribamos a la ecuación

$$\alpha X + \beta Y = \gamma,$$

que claramente tiene las mismas soluciones que la anterior, con la ventaja de que sus coeficientes α y β son coprimos. Fijemos ahora una solución cualquiera (x_0, y_0), que podemos hallar a través del algoritmo de Euclides, y consideremos cualquier solución genérica (x, y). Restando las igualdades

$$\begin{aligned} \alpha x + \beta y &= \gamma \\ \alpha x_0 + \beta y_0 &= \gamma \end{aligned}$$

obtenemos $\alpha(x - x_0) + \beta(y - y_0) = 0$, ó equivalentemente,

$$\beta(y - y_0) = \alpha(x_0 - x). \tag{5.11}$$

Esta última relación nos dice que $\alpha \mid \beta(y - y_0)$, y siendo $\alpha \perp \beta$ resulta que $\alpha \mid y - y_0$, digamos $y - y_0 = \alpha t$.

Razonando en forma totalmente análoga resulta que $\beta \mid x_0 - x$, y por lo tanto existe $k \in \mathbb{Z}$ tal que $x_0 - x = \beta k$. Reemplazando en (5.11) obtenemos $\beta\alpha k = \beta\alpha t$, de donde sigue que $t = k$, pues $\alpha\beta \neq 0$.

Despejando las variables en las igualdades precedentes vemos entonces que toda solución es de la forma $(x_0 - \beta k, y_0 + \alpha k)$, donde k es un cierto entero.

Recíprocamente, estas fórmulas brindan una solución para todo $k \in \mathbb{Z}$, ya que

$$\alpha(x_0 - \beta k) + \beta(y_0 + \alpha k) = \alpha x_0 + \beta y_0 - \alpha\beta k + \beta\alpha k = \gamma\,.$$

Resumiendo, la ecuación (5.10) admite infinitas soluciones (x, y), obtenidas a partir de una solución inicial (x_0, y_0) y a través de un *parámetro* k que varía libremente sobre $\mathbb{Z}$, mediante las fórmulas

Fórmula 5.2.11

$$\begin{cases} x = x_0 - k(b/d) \\ y = y_0 + k(a/d) \end{cases}$$

Ejemplo 5.2.12 Resolvamos completamente la ecuación diofántica

$$3724X + 13818Y = 7448\,.$$

Vimos que $(3724 : 13818) = 98$, y puesto que $7448 = 98\,.\,76$ resulta que la ecuación es resoluble. Sabemos además del ejemplo 5.2.6 que 98 se expresa como combinación lineal de 3724 y 13818 en la forma

$$98 = 3724 \times 26 + 13818 \times (-7).$$

Luego, multiplicando esta última igualdad por 76 obtenemos

$$3724 \times 1976 + 13818 \times (-532) = 7448,$$

vale decir, $(1976, -532)$ es una solución de la ecuación. Por lo tanto, la forma general de las soluciones es

$$\begin{cases} x = 1976 - 141\,k \\ y = -532 + 38\,k\,, \end{cases}$$

con k recorriendo $\mathbb{Z}$. $\quad \Diamond$

5.2.2. Ejercicios

Salvo aclaración explícita, en los siguientes ítems las letras designan números enteros.

1. En cada uno de los siguientes casos calcular $(a : b)$ y expresarlo como combinación lineal de a y b:

   a) $a = 210\,,\ b = 567$
   b) $a = 480\,,\ b = -176$
   c) $a = 15\,.\,36\,,\ b = 15\,.\,22$
   d) $a = -26\,,\ b = 0$
   e) $a = 28\,,\ b = 756$
   f) $a = 3^{18} + 1\,,\ b = 3^{18} - 1\,.$

2. Determinar los posibles valores de:

   a) $(a : a + 1)$
   b) $(a - 1 : a + 1)$
   c) $(4a : 2a + 3)$
   d) $(a^2 + 3 : a^2 - 2)\,.$

3. Si $n \in \mathbb{N}$, probar que $(a^m - 1 : a^n - 1) = a^{(m:n)} - 1\,.$

4. Determinar una progresión artimética de números enteros cuyos términos sean múltiplos de 4 y coprimos con 7. Exhibir otra tal que cada término sea múltiplo de 13 y su desarrollo decimal termine en 1.

5. Calcular $\displaystyle\sum_{a=0}^{99}(a : 15)\,.$

6. Sea d un divisor común de a y b $(d > 0)\,.$ Demostrar las siguientes propiedades:

a) $\left(\frac{a}{d} : \frac{b}{d}\right) = \frac{(a:b)}{d}$.

b) $(a:b) = d$ si y sólo si $\left(\frac{a}{d} : \frac{b}{d}\right) = 1$.

c) $(a:b) = d$ si y sólo si d es combinación lineal de a y b .

7. Demostrar que dos términos consecutivos cualesquiera de la sucesión de Fibonacci son coprimos.

8. Sea $x \in \mathbb{Q}$ tal que $x^n \in \mathbb{Z}$ $(n \in \mathbb{N})$. Probar que $x \in \mathbb{Z}$.

9. Sea $d = (a_1 : a_2 : \ldots : a_n)$.

   a) Probar que $d \mid a_i$ para todo i y que $t \mid d$ para todo divisor común t de los a_i .

   b) Probar que existen $x_1, x_2, \ldots, x_n$ tales que $d = \sum_{i=1}^{n} a_i x_i$.

   c) Demostrar que $\left(\frac{a_1}{d} : \frac{a_2}{d} : \ldots : \frac{a_n}{d}\right) = 1$.

10. Probar que existen en $\mathbb{Z}^2$ infinitos pares $(x, y) \in$ tales que $x + y = 100$ y $(x:y) = 5$.

11. *a*) Sean $a, b, c \in \mathbb{Z}$ tales que $a \mid bc$. Probar que $\frac{a}{(a:b)} \mid c$.

    b) Si $a, b \in \mathbb{N}$, determinar el número de puntos de coordenadas enteras que contiene el segmento del plano de extremos $(0, b)$ y $(a, 0)$.

12. Hallar los pares (x, y) de números naturales coprimos tales que

$$\frac{x+3}{4y} - \frac{2}{x} \in \mathbb{Z} .$$

13. Probar las siguientes afirmaciones $(n \in \mathbb{N})$:

    a) Si $(a:b) = 1$ entonces $(8a + b : 9a + 2b) = 1$ ó 7

    b) Si $(a:b) = 2$ entonces $(3a - 5b : 5a + 4b) = 2$ ó 74

    c) $(4^n + 5^n : 4^n - 5^n) = 1$

    d) $(3^n + 4^{n+1} : 3^{n+1} + 4^n) = 1$.

14. Demostrar las siguientes propiedades $(d, m \in \mathbb{N})$:

    a) $[\, da : db \,] = d\,[\, a : b \,]$

b) $\left[\dfrac{a}{d} : \dfrac{b}{d}\right] = \dfrac{[\,a : b\,]}{d}$ si $d \mid a$ y $d \mid b$

c) $[\,a^m : b^m\,] = [\,a : b\,]^m$.

15. Extender la definición de mínimo común múltiplo al caso de una familia de 3 o más números enteros. Verificar que se conservan, convenientemente adaptadas, las propiedades $(MCM)_1$ y $(MCM)_2$.

16. En cada uno de los siguientes casos determinar los pares $(a, b) \in \mathbb{N}^2$ que satisfacen las condiciones planteadas:

    a) $(a : b) = 25$ y $a + b = 300$

    b) $(a : b) = 6$ y $ab = 648$

    c) $(a : b) = 10$ y $[\,a : b\,] = 1500$.

17. Resolver las siguientes ecuaciones (o sistema de ecuaciones) diofánticas:

    a) $30X + 36Y = 24$

    b) $30X + 36Y = -174$

    c) $41X - 19Y = 8$

    d) $\begin{cases} 16X + 15Y = 1 \\ 6X + 10Y + 15Z = 11 \end{cases}$.

18. Hallar un múltiplo de 1557 cuyo desarrollo decimal termine en 654321.

19. Para adquirir unidades de dos clases de lapiceras, cuyos respectivos precios son \$12 y \$17,50, una escuela dispone de \$1680. ¿ De cuántas maneras puede hacer su compra si debe gastar todo el dinero y comprar al menos una lapicera de cada tipo ?

20. Sea $n \in \mathbb{N}$. Suponiendo que la recta de ecuación $9x + 14y = n$ contiene exactamente un punto de coordenadas enteras positivas, determinar el mayor valor posible de n .

5.3. Factorizacion

5.3.1. Números primos y compuestos

Un número entero a se dice *primo* si y sólo si admite exactamente cuatro divisores, a saber, 1, a, -1 y $-a$. Equivalentemente, a sólo admite las factorizaciones triviales $a = 1\,a$ y $a = (-1)(-a)$. Si $|\,a\,| > 1$ y a no es primo, diremos que a es *compuesto*.

NOTA Puesto que a y $-a$ tienen los mismos divisores, sigue de la definición que a es primo si y sólo si $-a$ lo es. Salvo aclaración explícita, de aquí en adelante emplearemos la palabra primo para referirnos a números primos positivos.

El lector podrá comprobar rápidamente que

$$2, 3, 5, 7, 11, 13, 17, \ldots$$

es la secuencia de los primeros números primos, siendo 2 el único par, ya que por definición los únicos divisores positivos de un primo a son 1 y a.

Observemos finalmente que un entero $a > 1$ es compuesto si y sólo si es posible factorizarlo en la forma $a = bc$ con b y c positivos y distintos de 1. Por ejemplo, $91 = 7\,.\,13$ es compuesto. $\Diamond$

La siguiente propiedad comenzará a revelar el rol de los números primos.

Proposición 5.3.1 Todo número natural $c > 1$ admite un divisor primo.

DEMOSTRACION. Consideremos el conjunto

$$A = \{\, d \in \mathbb{N}_2 \,:\, d \mid c \,\}\,.$$

Siendo $A \neq \emptyset$ ($c \in A$), sigue por el principio de buena ordenación que A admite un primer elemento, digamos q. Puesto que por transitividad todo divisor de q es divisor de c, concluimos que q no tiene divisores positivos propios, con excepción de 1. Por lo tanto q es primo y la propiedad queda demostrada.

Notemos que hemos probado de paso un hecho interesante por sí mismo: el mínimo divisor positivo distinto de 1 de un número natural es siempre un número primo. $\Diamond$

Corolario 5.3.2 (Euclides) Existen infinitos números primos.

DEMOSTRACION. El argumento es muy sencillo y se debe a Euclides, que sin duda con otro lenguaje lo utilizó unos 300 años antes de nuestra era: basta mostrar que dado cualquier conjunto finito $\{\, q_1, q_2, \ldots, q_n \,\}$ de primos existe otro distinto de todos ellos.

Consideremos para ello el número

$$c = 1 + \prod_{i=1}^{n} q_i \,.$$

Puesto que $c > 1$, existe por 5.3.1 un primo q tal que $q \mid c$. Por otro lado, $q_j \nmid c$ cualquiera sea j, ya que en caso contrario resultaría que $1 = c - \prod q_i$ sería divisible por q_j, lo que es absurdo. Luego $q \neq q_j$ para todo j, como queríamos demostrar. $\Diamond$

Corolario 5.3.3 (Eratóstenes) Todo número natural compuesto es divisible por un primo menor o igual que su raíz cuadrada.

DEMOSTRACION. Si c es un tal número, c se factoriza por hipótesis en la forma $c = ab$, donde sin pérdida de generalidad podemos suponer que $1 < a \leq b$. Resulta entonces que $a = c/b \leq c/a$, y por lo tanto $a^2 \leq c$. Bastará luego tomar cualquier divisor primo p de a, ya que entonces $p \mid c$ (por transitividad) y $p \leq a \leq \sqrt{c}$. $\Diamond$

Observemos que el enunciado contrarrecíproco de 5.3.3 establece el siguiente *test de primalidad*: un entero mayor que uno es primo si y sólo si no admite divisores primos menores o iguales que su raíz cuadrada. Por ejemplo 113 es primo, ya que ninguno de los primos 2, 3, 5 y 7 lo divide.

Los números primos tienen características muy especiales que los distinguen del resto de los números enteros. En particular la que exhibiremos en la siguiente proposición, que nos conducirá luego a uno de los teoremas más importantes de la Aritmética elemental.

Proposición 5.3.4 Sea p un número primo y sean $a, b \in \mathbb{Z}$ tales que $p \mid ab$. Entonces $p \mid a$ ó $p \mid b$. Más generalmente, sea $m \in \mathbb{N}$ y sean $a_1, a_2, \ldots, a_m$ enteros tales que

$$p \mid \prod_{i=1}^{m} a_i.$$

Entonces existe un índice k tal que $p \mid a_k$.

DEMOSTRACION. Probaremos solamente la primera parte, ya que la segunda sigue luego fácilmente por inducción en m.

A tal efecto, sea $d = (p : a)$. Puesto que en particular $d \mid p$, resulta que $d = p$ ó $d = 1$, pues p es primo. En el primer caso $p \mid a$ (y la conclusión vale), mientras que en el segundo $p \perp a$. Pero entonces sigue por la propiedad 3) de la proposición 5.2.9 que $p \mid b$, lo que finaliza la prueba. $\Diamond$

Vale la pena remarcar un hecho establecido en el curso de la demostración anterior. Dados un entero cualquiera x y un primo p se verifica alguna de las dos siguientes situaciones: p es un divisor de x ó bien $p \perp x$. Tal hecho

no ocurre en general, ya que por ejemplo 8 no divide a 12 ni es coprimo con 12. Al finalizar esta sección le propondremos demostrar al lector que la propiedad 5.3.4 caracteriza a los números primos, en el sentido de que son los únicos números naturales mayores que 1 que la satisfacen.

5.3.2. Factorización única

Ya estamos en condiciones de mostrar al lector el carácter generador de los números primos:

Proposición 5.3.5 Todo entero $c > 1$ se factoriza de manera única como producto de números primos.

DEMOSTRACION Aplicando el principio de inducción global, mostraremos en primer término la existencia de una tal descomposición en primos. Si $c = 2$ el resultado es inmediato, pues 2 es primo. Suponiendo entonces que $c > 2$ y que el enunciado es válido para enteros x tales que $1 < x < c$, tomemos un divisor primo p de c (proposición 5.3.1) y escribamos $c = p\,c_1$. Si $c_1 = 1$ resulta que c es primo y no hay nada que probar, mientras que si $c_1 > 1$ podemos aplicar la hipótesis inductiva, ya que $c_1 = c/p < c$. Por lo tanto c_1 es producto de números primos, y siendo $c = p\,c_1$ resulta que c también lo es.

Notemos que la prueba por inducción formaliza una idea muy simple: buscamos un divisor primo de c, luego un divisor primo de c_1, etc. Por ejemplo:

$$532 = 2\,.\,266 = 2\,.\,2\,.\,133 = 2\,.\,2\,.\,7\,.\,19.$$

Comentemos que aunque la teoría es clara y ha sido relativamente sencillo probar que todo número se factoriza como producto de primos, en la práctica la situación es más complicada. El problema es que no se conocen algoritmos eficientes de factorización que permitan determinar rápidamente los divisores primos de un número muy grande. Retomaremos más adelante el tema, así como el de analizar la primalidad de un número, ambos muy conectados con las modernas aplicaciones criptográficas de la Aritmética.

Vayamos ahora a la unicidad de la factorización, que como ya veremos también es un hecho de singular importancia. Suponiendo entonces que

$$c = p_1\,p_2\,\ldots\,p_m = q_1\,q_2\,\ldots\,q_n \tag{5.12}$$

son dos posibles factorizaciones de c como producto de primos, probaremos que $m = n$ y que existe una permutación π de $\mathbb{I}_m$ tal que $q_i = p_{\pi(i)}$ para todo índice i. Vale decir, ambas descomposiciones son esencialmente iguales, pudiendo diferir sólo en el orden de los factores.

Suponiendo sin perder generalidad que $m \geq n$, procederemos por inducción en m . El caso $m = 1$ es trivial, ya que entonces $m = n = 1$ y $p_1 = q_1$. Suponiendo $m > 1$, y observando que p_m es un divisor de $q_1 q_2 \ldots q_n$, resulta

por la proposición 5.3.4 que p_m divide alguno de estos factores. Más aún, reordenando los índices si es necesario podemos suponer que $p_m \mid q_n$, de donde sigue que $p_m = q_n$, pues ambos son primos positivos.

Cancelando en (5.12) el factor común p_m arribamos a la igualdad

$$p_1 p_2 \dots p_{m-1} = q_1 q_2 \dots q_{n-1},$$

siendo ahora $m-1$ el máximo número de factores. Sigue entonces por hipótesis inductiva que $n-1 = m-1$ (en cuyo caso $n = m$) y que los factores q_i son un reordenamiento de los p_i $(1 \leq i < m)$. Habiendo probado ya que $q_m = p_m$, queda demostrada la unicidad de la factorización. $\Diamond$

Como vimos en un ejemplo, los factores p_i no son necesariamente distintos. Podemos entonces agrupar los factores repetidos y representar unívocamente a todo número natural c como producto de potencias de números primos distintos, asignando a 1 la representación vacía. Finalmente, si extendemos esta forma de expresión a los enteros negativos insertando un signo delante del producto, obtenemos la versión tradicional del *Teorema Fundamental de la Aritmética* (TFA), que enunciamos a continuación:

Teorema 5.3.6 (Fundamental de la Aritmética) Todo número entero no nulo c se factoriza unívocamente en la forma

$$c = \epsilon \prod_{i=1}^{m} {p_i}^{c_i},$$

donde los p_i son primos positivos distintos, los exponentes c_i son números naturales, $\epsilon = \pm 1$ y $m \in \mathbb{N}_0$. $\Diamond$

El teorema fundamental de la Aritmética es una herramienta muy fuerte, que permite resolver más sencillamente muchas cuestiones, como el análisis de la coprimalidad de dos números, la descripción de los divisores de un número, etc. Mostraremos en la siguiente proposición algunas de sus consecuencias, en las que podremos constatar la importancia de la factorización única. De aquí más, entenderemos por *factorización* de un número a su factorización única como producto de potencias de primos distintos.

Proposición 5.3.7 Se deducen del TFA las siguientes propiedades (las letras $a, b, \dots$ indican números naturales y las letras p números primos):

1) Los primos que aparecen en la factorización de a son exactamente los divisores primos de a.

2) Si p es un divisor primo de a, el exponente de p en la factorización de a es el máximo $m \in \mathbb{N}$ tal que $p^m \mid a$.

3) a y b son coprimos si y sólo si no tienen divisores primos comunes.

4) Sea

$$a = \prod_{i=1}^{r} p_i^{a_i}$$

la factorización de a. Entonces los divisores positivos de a son de la forma

$$d = \prod_{i=1}^{r} p_i^{d_i},$$

donde $0 \leq d_i \leq a_i$ para todo i. Sigue en particular que a admite

$$\prod_{i=1}^{r} (a_i + 1)$$

divisores positivos.

5) Supongamos que $a_1, a_2, \ldots, a_n$ son coprimos dos a dos y que $a_i \mid b$ para todo i. Entonces

$$\prod_{i=1}^{m} a_i \mid b\,.$$

6) Sea

$$a = \prod_{i=1}^{r} p_i^{a_i}$$

la factorización de a. Entonces a es un cuadrado perfecto si y sólo si a_i es par para todo i. En general, a es una potencia $n-$ésima perfecta si y sólo si $n \mid a_i$ para todo i.

Demostracion.

1) Si

$$a = \prod_{i=1}^{r} p_i^{a_i}$$

es la factorización de a, es claro que $p_i \mid a$ para todo i. Tomemos ahora un divisor primo q de a. Puesto que un primo divide a un producto si y sólo si divide a uno de los factores, sigue que $q \mid p_k^{a_k}$ para algún k, y nuevamente por la misma razón resulta que $q \mid p_k$. Siendo ambos primos concluimos que $q = p_k$.

2) Dada la definición de m, podemos escribir $a = p^m t$, donde $p \nmid t$. Entonces, yuxtaponiendo p^m a la factorización de t obtenemos una expresión de a como producto de potencias de números primos en la que p aparece con exponente m, ya que por 1) p no aparece en la factorización de t. Arribamos entonces al resultado usando que la factorización es única.

3) Es evidente que la condición del enunciado es necesaria para la coprimalidad de a y b. Para ver que es suficiente, supongamos que la misma se

cumple y que d es un divisor común de a y b distinto de 1. En tal caso d es divisible por algún primo q, resultando por transitividad que $q \mid a$ y $q \mid b$, lo que es absurdo. Luego $a \perp b$.

4) Supongamos en primer término que d es de la forma indicada. Entonces, podemos escribir

$$a = \prod_{i=1}^{r} p_i^{a_i} = \prod_{i=1}^{r} p_i^{d_i} \prod_{i=1}^{r} p_i^{a_i - d_i},$$

y por lo tanto $d \mid a$.

Recíprocamente, asumamos que $d \mid a$, en cuyo caso todo divisor primo de d es un divisor de a. Sigue entonces por 1) que los $p_i \, (1 \leq i \leq r)$ son los únicos primos que pueden aparecer en la factorización de d. Por lo tanto, d es de la forma

$$\prod_{i=1}^{r} p_i^{d_i},$$

donde los exponentes son enteros no negativos. Insistiendo nuevamente en el hecho de que todo divisor de d es un divisor de a, resulta por la caracterización obtenida en 2) que $d_i \leq a_i$ para todo i, como queríamos probar.

En cuanto a la última afirmación, notemos que para cada i existen $a_i + 1$ formas de elegir el exponente d_i, y que cada elección múltiple corresponde a un divisor distinto de a, por la unicidad de la factorización. Un sencillo cálculo combinatorio nos conduce entonces al resultado deseado.

5) Sea

$$b = \prod_{i=1}^{r} p_i^{b_i}$$

la factorización de b. Puesto que a_j y a_k no tienen divisores primos comunes si $j \neq k$, resulta que la factorización de $a = a_1 \ldots a_m$ se obtiene simplemente yuxtaponiendo las respectivas factorizaciones de los a_i. Puesto que por hipótesis estos números dividen a b, resulta por la propiedad 4) que todo divisor primo de a es algún p_j y aparece en la factorización de a con un exponente menor o igual que b_j. Usando nuevamente 4) concluimos que $a \mid b$.

6) Supongamos primero que a es un cuadrado perfecto, digamos $a = c^2$. Sigue entonces que a y c tienen los mismos factores primos, y por lo tanto c tendrá una factorización de la forma

$$c = \prod_{i=1}^{r} p_i^{c_i}.$$

Luego

$$a = c^2 = \left(\prod_{i=1}^{r} p_i^{c_i} \right)^2 = \prod_{i=1}^{r} p_i^{2c_i}.$$

Por la unicidad de la factorización, concluimos que $a_i = 2c_i$ para todo i.

Para la recíproca basta revertir los pasos. En efecto, supongamos que a_i es par para todo índice i. Entonces

$$a = \prod_{i=1}^{r} p_i^{a_i} = \prod_{i=1}^{r} p_i^{2(a_i/2)} = \left(\prod_{i=1}^{r} p_i^{a_i/2} \right)^2 ,$$

y en consecuencia a es un cuadrado perfecto.

El caso general de caracterización de potencias n-ésimas perfectas a través de la factorización se resuelve en forma idéntica, por lo que dejamos los detalles a cuidado del lector. $\Diamond$

NOTA. Si p y q son primos distintos, es claro que $p^m \perp q^n$ cualesquiera sean $m, n \in \mathbb{N}$, ya que no hay primos comunes en sus factorizaciones. Sigue entonces por TFA que todo número $a > 1$ es producto de números coprimos dos a dos. En consecuencia, la propiedad 5) de la proposición 5.3.7 permite reducir el estudio de la divisibilidad al caso de potencias de números primos. Por ejemplo, un entero es múltiplo de $360 = 2^3 \,.\, 3^2 \,.\, 5$ si y sólo si es múltiplo de 8, 9 y 5. $\Diamond$

Finalizamos la sección con algunos ejemplos de aplicación.

Ejemplo 5.3.8 Veamos que $\sqrt{c}$ es un número irracional si $c \in \mathbb{N}$ y c no es un cuadrado perfecto. En efecto, supongamos por el contrario que $\sqrt{c} = a/b$, donde a y b son números naturales coprimos. Elevando al cuadrado ambos miembros y operando obtenemos entonces la igualdad $cb^2 = a^2$.

Sigue luego que cualquier divisor primo q de c es divisor de a, y por lo tanto no divide a b. Esto significa que q aparece en la factorización de c con el mismo exponente con que aparece en la de a^2, lo que nos muestra que dicho exponente es par. Pero esto contradice la hipótesis de que c no es un cuadrado perfecto, ya que en tal caso alguno de los primos que lo dividen debe aparecer en su factorización con exponente impar. $\Diamond$

Ejemplo 5.3.9 Calculemos en cuántos ceros termina el desarrollo decimal de $15\,!$ Puesto que el desarrollo decimal de un número natural termina en exactamente k ceros si y sólo si 10^k es la máxima potencia de 10 que lo divide, y siendo $10 = 2 \,.\, 5$, debemos calcular los exponentes de 2 y 5 en la factorización de $15\,!$

Será un trabajo sencillo: observemos que cada uno de los 7 números pares entre 1 y 15 aporta un factor 2, cada uno de los tres múltiplos de 4 (4,8,12) aporta dos factores, uno de los cuales ya fue tenido en cuenta, mientras que 8 aporta otro factor adicional. En total, $7+3+1 = 11$ factores 2, esto es, $15\,!$ es divisible por 2^{11} pero no por 2^{12}. En cuanto a 5, aportan un único factor los números 5, 10 y 15. Luego, $15\,!$ es divisible por 10^3 pero no por 10^4, ya que no lo es por 5^4. En definitiva, $15\,!$ termina en exactamente 3 ceros. $\Diamond$

5.3.3. Ejercicios

1. *a*) Probar que todo primo $p > 2$ es de la forma $4k \pm 1$. Demostrar que existen infinitos primos de la forma $4k - 1$.

   b) Probar que todo primo $p > 3$ es de la forma $6k \pm 1$.

2. Hallar el menor primo p tal que $r_{25}(p) = 19$ y $r_{60}(p) = 49$.

3. Sea $a \in \mathbb{N}$ verificando la siguiente propiedad:

   $$a \mid bc \;\Leftrightarrow\; a \mid b \text{ ó } a \mid c\,,$$

   cualesquiera sean los enteros b y c. Probar que $a = 1$ ó a es primo.

4. *a*) Verificar que $a^2 + a + 41$ es primo para todo entero a en el rango $-40 \leq a < 40$.

   b) Demostrar que no existe una función lineal o cuadrática f con coeficientes enteros tal que $f(a)$ es primo para todo $a \in \mathbb{Z}$.

5. Sea p primo y sea $0 < k < p$. Demostrar que $\binom{p}{k}$ es múltiplo de p.

6. Resolver en $\mathbb{N}$ la ecuación $2n^3 + 3n^2 + n = 75174$.

7. Resolver las siguientes ecuaciones diofánticas:

   a) $X^2 - Y^2 = 65$

   b) $X^4 + 12 = Y^2$

   c) $X^2 + 9 = Y^4$

   d) $X^3 + Y^3 = 20$.

8. Dados 29 múltiplos de 12 entre 1 y 500, probar que alguno de ellos es múltiplo de 36.

9. Sea $n \in \mathbb{N}$ ($n > 4$). Probar que n es compuesto si y sólo si $n \mid (n-1)!$

10. ¿Cuántos números naturales menores que 1000 son coprimos con 120?

11. Caracterizar los números naturales no divisibles por ningún primo impar.

12. Caracterizar los $n \in \mathbb{N}$ que pueden descomponerse como suma de 2 o más números naturales consecutivos. Resolver las situaciones $n = 99$ y $n = 180$.

13. Caracterizar los números naturales que pueden expresarse como diferencia de dos cuadrados perfectos.

14. Hallar la suma y el producto de los divisores positivos de $2^{10}3^{20}7^{15}$.

15. Hallar el número de divisores positivos de $6825 \,.\, 8264^2 \,.\, 59049^3$.

16. Dados un primo p y un entero no nulo a, designamos por $\nu_p(a)$ el exponente de p en la factorización de a ($\nu_p(a) = 0$ si $p \nmid a$).

    a) Si $n \in \mathbb{N}$, probar que

    $$\nu_p\,(\,n!\,) \;=\; \sum_{i=1}^{\infty} \left[\frac{n}{p^{\,i}}\right] .$$

    b) Calcular el número de divisores primos de $\binom{50}{22}$.

    c) Determinar en cuántos ceros termina el desarrollo decimal de 80! ¿Cuál es la máxima potencia de 45 que lo divide?

17. *a*) Caracterizar los $a \in \mathbb{Z}$ tales que $\dfrac{(5a+1)(3a+2)}{15} \in \mathbb{Z}$

    b) Probar que $30 \mid a^5 - a$ para todo $a \in \mathbb{Z}$.

18. Determinar los $m, n \in \mathbb{N}$ tales qué $\sqrt[n]{6^m} \in \mathbb{N}$. Probar que $\log_6(12) \notin \mathbb{Q}$.

19. Analizar la resolubilidad de las siguientes ecuaciones ($m, n \in \mathbb{N}$):

    a) $2m^2 = n^2$

    b) $m^3 = 4n^3$

    c) $3^m = 5^n$

    d) $6^m = m^6$.

20. Sean a y b números naturales coprimos tales que ab es un cuadrado perfecto. Probar que a y b son cuadrados perfectos.

21. *a*) Para $n = 936, 6875, 4186$ y 46656, hallar el mayor cuadrado perfecto y el mayor cubo perfecto que dividen a n.

    b) Hallar el menor cubo perfecto divisible por 475 y la menor potencia cuarta divisible por 4056.

22. Hallar el mínimo $n \in \mathbb{N}$ tal que $a^2 \mid n$ para todo $1 \leq a \leq 10$.

23. Probar que existen infinitas ternas de enteros consecutivos cuya suma es un cuadrado perfecto. Si $m > 1$, probar que la suma de 4 enteros consecutivos cualesquiera no es una potencia $m-$ésima perfecta.

24. Sean $a = \prod_{i=1}^{r} p_i^{a_i}$ y $b = \prod_{i=1}^{r} p_i^{b_i}$ números naturales, donde $p_1, p_2, \ldots, p_r$ son primos distintos y los exponentes a_i y b_i son enteros no negativos. Probar la validez de las siguientes fórmulas:

    a) $(a : b) = \prod_{i=1}^{r} p_i^{\min(a_i, b_i)}$

    b) $[\, a : b \,] = \prod_{i=1}^{r} p_i^{\max(a_i, b_i)}$.

25. Sean $a_1, a_2, \ldots, a_r$ y b números enteros tales que $b \perp a_i$ para todo i. Probar que
$$b \perp \prod_{i=1}^{r} a_i \,.$$

26. Si $n \in \mathbb{N}$, calcular $(42^n - 1 : 15867)$.

27. Determinar los $a \in \mathbb{N}$ tales que $[\, a : 90 \,] = 14\,(a : 90)$.

28. En cada uno de los siguientes casos, determinar los $n \in \mathbb{N}$ que satisfacen todas las condiciones planteadas:

    a) $(n : 945) = 63$ y $[\, n : 5950 \,] = 2034900$.

    b) $2 \mid n$, $(n : 495) = 165$ y n tiene exactamente 40 divisores positivos.

29. Un número entero se dice *libre de cuadrados* si 1 es el único cuadrado perfecto que lo divide. Probar que todo $n \in \mathbb{N}$ se expresa unívocamente en la forma $n = a^2 c$, donde $a \in \mathbb{N}$ y c es libre de cuadrados.

30. Si $n > 1$, demostrar que
$$1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \notin \mathbb{Z}.$$

31. Una terna (x, y, z) de números naturales tales que $x^2 + y^2 = z^2$ se dice una *terna pitagórica*. De acuerdo con el ejercicio 18 de la sección 5.1 podemos suponer sin pérdida de generalidad que y es par. Las ternas pitagóricas tales que $(x : y : z) = 1$ se dirán *primitivas*.

    a) Sea (x, y, z) una terna pitagórica y sea $d = (x : y : z)$. Probar que $(x/d, y/d, z/d)$ es una terna pitagórica primitiva.

    b) Sean $u > v$ números naturales coprimos de distinta paridad. Probar que $(u^2 - v^2, 2uv, u^2 + v^2)$ es una terna pitagórica primitiva.

    c) Demostrar que toda terna pitagórica primitiva es del tipo señalado en *b*).

    d) Hallar las ternas pitagóricas primitivas de la forma $(x, 120, z)$.

Nota informativa En el ejercicio precedente se proponen pautas para resolver una ecuación diofántica, a saber, la ecuación cuadrática

$$X^2 + Y^2 = Z^2 .$$

A través de dichas pautas se hallan sus infinitas soluciones no triviales (aquellas en las que las variables son distintas de cero), que pueden expresarse paramétricamente como en el caso de la ecuación diofántica lineal. No hay duda de que hay un salto de dificultad entre ambos tipos de ecuaciones −es necesario emplear otros recursos−, pero todavía podemos encuadrar el método del problema dentro de un nivel razonablemente elemental.

El panorama cambia drásticamente si se encaran situaciones más complejas. Por ejemplo, podríamos preguntarnos si es posible que la suma de los cubos de dos números naturales sea también un cubo perfecto. Esta cuestión, que tiene respuesta negativa y cuya resolución dista de ser elemental, es sólo un caso particular de uno de los más célebres y arduos problemas de la Aritmética y de la Matemática en general, propuesto por el matemático francés Pierre de Fermat a mediados del siglo XVII. La formulación precisa del problema, conocido como la *última conjetura de Fermat*, es la siguiente:

Si $n \in \mathbb{N}$ y $n > 2$, no existen enteros no nulos a, b y c tales que

$$a^n + b^n = c^n.$$

Como se ve, se trata de una ecuación diofántica que generaliza el comentado caso cuadrático, pero según parece de comportamiento muy distinto. Si bien Fermat afirmó tener una demostración de este hecho, nunca la dió a conocer, y se cree que no la tenía o que no era correcta.

Fermat conocía una demostración para el caso de exponente 4, que es relativamente elemental, y con el correr del tiempo la conjetura fue probada para muchos valores particulares del exponente (basta demostrarla para n primo). El problema es que cada caso parecía requerir una técnica particular, y sumamente complicada.

Sin embargo −y aquí reside probablemente el mayor mérito de Fermat−, los esfuerzos desplegados durante más de tres siglos por grandes matemáticos de todo el mundo para convertir la conjetura en un teorema posibilitaron un enorme desarrollo del conocimiento matemático, no sólo dentro de la teoría de números, sino también en otras áreas, como el algebra conmutativa y la geometría algebraica, y como suele ocurrir, las investigaciones realizadas excedieron el marco del problema que les dió origen, permitiendo el establecimiento de teorías matemáticas que hoy ya son clásicas.

Finalmente, en 1994 el matemático inglés Andrew Wiles demostró la conjetura, el ahora sí último teorema de Fermat. Casi no es necesario decir que las técnicas y herramientas que empleó −demostró en realidad una conjetura más general sobre ciertas curvas y que implica el teorema− superan infinitamente el alcance de estas páginas, pero pensamos que es bueno que el lector tenga noticia, aún sea en esta forma imprecisa, de la existencia de problemas de tamaña magnitud, tan simples de contar y tan difíciles de resolver, y que están ciertamente conectados con los temas que estamos tratando.

Capítulo 6

Aritmética Modular

6.1. Congruencia Entera

6.1.1. Definiciones y propiedades

La noción de congruencia, presentada en el capítulo previo, fue explícitamente introducida en 1801 por Gauss en sus *Disquisitiones Arithmeticae*, aunque sin duda la idea no había sido ajena al pensamiento de otros grandes matemáticos anteriores, como Fermat y Euler. Si bien podría creerse que sólo se trata de un lenguaje novedoso, es en realidad mucho más que eso, ya que introduce una forma mucho más adecuada de pensar, simplificando gratamente los algo trabajosos cálculos con restos (básicamente de eso se trata). Recordemos la definición:

> Si $m \in \mathbb{N}$ y $a, b \in \mathbb{Z}$, decimos que a es *congruente con* b *módulo* m si y sólo si $m \mid a - b$. Empleamos en tal caso la notación
>
> $$a \equiv b \,(m)\ ,$$
>
> o alternativamente $a \equiv b \bmod m$.

Para familiarizarse con la definición, sugerimos al lector que verifique por ejemplo la validez de las congruencias $14 \equiv 6\ (2)$, $-3 \equiv 25\ (4)$, $-3 \equiv 25\ (7)$, $54 \equiv 0\ (9)$ y $-6 \equiv -45\ (13)$.

Como hemos establecido en la proposición 5.1.5, la noción de congruencia entre dos enteros (diferir en un múltiplo de m) coincide con la de "tener igual resto" en la división por m. Conviene entonces tener presente que las tres afirmaciones siguientes refieren la misma situación:

$$\begin{array}{ll} (C_1) & a \equiv b\ (m) \\ (C_2) & a = b + km \text{ para algún } k \in \mathbb{Z} \\ (C_3) & r_m(a) = r_m(b). \end{array}$$

Si bien estos tres enunciados equivalentes sólo difieren ligeramente entre sí, ante una situación o problema concreto el empleo de una u otra definición

alternativa de congruencia puede hacernos ver con mayor claridad la cuestión a resolver. Veamos algunas propiedades inmediatas de la congruencia entera.

Proposición 6.1.1 Dado $m \in \mathbb{N}$, valen las siguientes afirmaciones(las letras designan números enteros):

1) Si $a \equiv b\ (m)$ y d un divisor de m entonces $a \equiv b\ (d)$

2) $a \equiv r_m(a)\ (m)$ cualquiera sea $a \in \mathbb{Z}$

3) Sean a, b tales que $|\, a - b \,| < m$. Entonces $a \equiv b\ (m) \Leftrightarrow a = b$

4) $a \equiv 0\ (m) \Leftrightarrow m \mid a\,.$

Demostracion. La afirmación 1) es consecuencia de la transitividad de la relación de divisibilidad, mientras que 2) sigue de la definición de resto y de la caracterización de congruencia dada por (C_2). Por ejemplo, $257 \equiv 5\ (12)$, ya que el resto de dividir 257 por 12 es 5. En cuanto a 3), notemos primero que $a-b$ es múltiplo de m si y sólo si $|\, a - b \,|$ lo es. Luego, $a \equiv b\ (m)$ si y sólo si $|\, a - b \,| = 0$, como queríamos probar. Como corolario de esta propiedad resulta que los restos de dividir m enteros consecutivos por m son todos distintos. Finalmente, 4) sigue directamente de la definición. $\Diamond$

Clases de congruencia.

Es claro que la relación de congruencia módulo m es una relación de equivalencia en $\mathbb{Z}$, ya que ella puede describirse a través de la igualdad de restos. Sus clases de equivalencia se llaman *clases de congruencia* módulo m, y como ocurre en el caso general, las mismas determinan una partición de $\mathbb{Z}$. De acuerdo con las definiciones, la clase de congruencia de un entero a, que notaremos por (a), admite las descripciones:

$$(a) = \{x \in \mathbb{Z} : x \equiv a \bmod m\} = \{x \in \mathbb{Z} : r_m(x) = r_m(a)\} = \\ = \{a + km : k \in \mathbb{Z}\}\,.$$

La idea de la *Aritmética modular* (en este caso módulo m), consiste en identificar entre sí los números pertenecientes a un misma clase de congruencia. Dicho de otro modo, los números se clasifican según sus restos de dividirlos por m, y puesto que hay m restos posibles, existen entonces m clases de congruencia módulo m. Precisamente, $\mathbb{Z}$ admite la partición:

$$\mathbb{Z} = (0) \cup (1) \cup \cdots \cup (m-1)\,.$$

Por ejemplo, hay dos clases de congruencia módulo 2, la de los pares y la de los impares, mientras que la congruencia módulo 1 es trivial (identifica a todos los enteros), pues todos son múltiplos de 1. Volviendo al caso general, observemos que hemos elegido como *representantes* de las clases de congruencia módulo m los números $0, 1, \ldots, m-1$, lo que es completamente natural,

tomando en cuenta que $a \equiv r_m(a)$ para todo a, y que dichos números no son congruentes entre sí módulo m, por 2) de la proposición 6.1.1.

En rigor, tomar como representantes canónicos de las clases de congruencia a los m números que usamos como restos obedece a una simple razón de comodidad, ya que podemos representar a las clases por cualquier conjunto de m números enteros cuyos restos de dividir por m sean todos distintos. Un tal conjunto se llama un *sistema completo de restos módulo m*. Por ejemplo, cualquier conjunto de m enteros consecutivos lo es. Una elección menos natural sería por ejemplo tomar $\{14, -6, -12, 3, 39, 96, -1\}$ como sistema completo de restos módulo 7.

Operaciones modulares La relación de congruencia se comporta en forma muy similar a la igualdad con respecto a las operaciones del anillo $\mathbb{Z}$. En realidad, las propiedades que enunciaremos a continuación traducen en términos de congruencias las propiedades de los restos que hemos establecido en la proposición 5.1.6 del capítulo anterior, por lo que obviaremos algunas demostraciones. En lo que sigue las letras designan números enteros, y salvo mención expresa, todas las congruencias son módulo un cierto número natural m.

Proposición 6.1.2 La relación de congruencia módulo m satisface las siguientes propiedades:

1) Si $a \equiv b$ y $c \equiv d$ entonces $a \pm c \equiv b \pm d$

2) Si $a \equiv b$ y $c \equiv d$ entonces $ac \equiv bd$

3) Más generalmente, supongamos que $a_i \equiv b_i$ para $1 \leq i \leq r$. Entonces

$$\sum_{i=1}^{r} a_i \equiv \sum_{i=1}^{r} b_i \qquad \text{y} \qquad \prod_{i=1}^{r} a_i \equiv \prod_{i=1}^{r} b_i$$

4) Si $a \equiv b$ y $k \in \mathbb{N}$ entonces $a^k \equiv b^k$

5) Sea d un divisor común de a, b y m. Entonces

$$a \equiv b \Leftrightarrow a/d \equiv b/d \ (m/d)$$

6) Si $ac \equiv bc$ y $c \perp m$ entonces $a \equiv b$.

Demostracion. Como señalamos previamente, las propiedades 1) a 4) se deducen de las correspondientes propiedades de los restos respecto a las operaciones. Para demostrar 5), dividiendo por d una relación de la forma $a = b + km$ $(k \in \mathbb{Z})$ obtenemos $a/d = b/d + km/d$, ó equivalentemente $a/d \equiv b/d \ (m/d)$. Para probar la recíproca basta multiplicar por d la relación de la derecha.

Con respecto a 6), sigue por hipótesis que $m \mid (a-b)c$, de lo cual deducimos que $m \mid a-b$, por ser c y m coprimos. Luego $a \equiv b$. Vale la pena hacer notar que la hipótesis de coprimalidad es realmente necesaria, ya que por ejemplo no puede cancelarse el factor 2 en la relación $2\,.\,3 \equiv 2\,.\,5\ (4)$, ya que $3 \not\equiv 5\ (4)$. $\Diamond$

Ejemplo 6.1.3 En este comportamiento tan regular respecto de las operaciones reside gran parte de la eficacia del uso de congruencias, pues hace mucho más sencillos los cálculos y la comprensión de algunos hechos. Como aplicación, calculemos la cifra de las unidades de 7^{7^7}.

Se trata de un número muy grande (tiene 695975 cifras), pero recordando que la última cifra del desarrollo decimal de un número es el resto de dividirlo por 10, podemos hallarla sin mayor dificultad trabajando con congruencias módulo 10. En efecto, calculando mod 10 las primeras potencias de 7 obtenemos

$$7^2 = 49 \equiv -1\ (10),$$

de donde sigue que

$$7^4 = (7^2)^2 \equiv (-1)^2 = 1\ (10).$$

Por otro lado, tomando ahora congruencias módulo 4 tenemos:

$$7^7 \equiv (-1)^7 = -1 \equiv 3\ (4),$$

esto es, el resto de dividir 7^7 por 4 es 3, y por lo tanto 7^7 es de la forma $4k+3$ ($k \in \mathbb{N}$).

Luego,

$$7^{7^7} = 7^{4k+3} = (7^4)^k 7^3 \equiv 7^3 = 7^2 7 \equiv -7 \equiv 3\ (10).$$

Por lo tanto la última cifra de 7^{7^7} es 3. $\Diamond$

Estructura algebraica de las clases de restos.

La compatibilidad que hemos señalado entre la relación de congruencia y las operaciones puede ser descripta en términos más estructurales. En efecto, si designamos por $\mathbb{Z}_m$ el conjunto de las clases de congruencia módulo m, podemos definir naturalmente la suma y el producto de elementos de $\mathbb{Z}_m$, mediante las reglas:

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

y

$$(a)(b) = (ab).$$

Por ejemplo, $(8) + (9) = (5)$ y $(7)(9) = (3)$ en $\mathbb{Z}_{12}$. Las propiedades 1) y 2) de la proposición 6.1.2 aseguran la buena definición de las mismas, en el sentido de que no dependen de los representantes a y b elegidos. Esto confiere a $\mathbb{Z}_m$ una estructura algebraica similar, aunque no idéntica, a la de

$\mathbb{Z}$. Para fijar las ideas, construyamos las tablas de la suma y el producto en $\mathbb{Z}_5$ (clases distintas de (0) en el caso del producto):

$+$	(0)	(1)	(2)	(3)	(4)
(0)	(0)	(1)	(2)	(3)	(4)
(1)	(1)	(2)	(3)	(4)	(0)
(2)	(2)	(3)	(4)	(0)	(1)
(3)	(3)	(4)	(0)	(1)	(2)
(4)	(4)	(0)	(1)	(2)	(3)

$\times$	(1)	(2)	(3)	(4)
(1)	(1)	(2)	(3)	(4)
(2)	(2)	(4)	(1)	(3)
(3)	(3)	(1)	(4)	(2)
(4)	(4)	(3)	(2)	(1)

Por simple inspección de las mismas, podemos advertir algunas características comunes en ambas operaciones. En primer lugar, cada una de ellas admite un elemento neutro, (0) en el caso de la suma y (1) en el del producto, observándose además que cada fila de ambas tablas es una permutación de las clases, esto es, aparecen en ella todas las clases en un cierto orden. Resulta en particular que cada clase tiene un inverso respecto a la suma y cada clase no nula tiene un inverso respecto al producto (esto ya lo distingue del producto de números enteros). Por ejemplo, el inverso aditivo de (4) es (1), mientras que el inverso multiplicativo de (2) es (3). Pasemos ahora a describir la estructura de $\mathbb{Z}_m$ en el caso general.

Proposición 6.1.4 $\mathbb{Z}_m$ es un anillo conmutativo.

DEMOSTRACION Teniendo en cuenta sus definiciones, es claro que el carácter asociativo y conmutativo de ambas operaciones, así como la distributividad del producto respecto de la suma, son consecuencia inmediata de la validez de dichas propiedades para la suma y producto de números enteros. Por idénticas razones resulta también que (0) es el elemento neutro de la suma y (1) es el elemento neutro del producto. Finalmente, para probar que toda clase (a) admite un inverso aditivo observemos que $(a)+(m-a) = (m) = (0)$, y por lo tanto $(m-a)$ es la clase inversa de (a).

Por ejemplo, $-(7) = (3)$ en $\mathbb{Z}_{10}$, mientras que $-(5) = (5)$. $\Diamond$

NOTA Contra lo que podría inducirnos a pensar lo observado en $\mathbb{Z}_5$, donde toda clase no nula admite inverso multiplicativo, el anillo de restos $\mathbb{Z}_m$ no es general un cuerpo. En efecto, observemos por ejemplo que la ecuación $4x \equiv 1\ (6)$ no es resoluble, ya que $4x-1$ es impar y por lo tanto no divisible por 6. Traducido al anillo $\mathbb{Z}_6$, esto significa que no existe una clase (x) tal que $(4)(x) = (1)$, o sea, (4) no es inversible y $\mathbb{Z}_6$ no es un cuerpo.

Otra característica de los anillos $\mathbb{Z}_m$ que los distingue del anillo de números enteros es que no son en general *dominios de integridad*, esto es, el producto de dos elementos no nulos puede ser nulo. Por ejemplo, $(3)(8) = (24) = (0)$ en $\mathbb{Z}_{12}$. Más adelante estudiaremos con mayor detalle estas cuestiones. $\Diamond$

6.1.2. Ejercicios

1. Agrupar los números 0, 3, -1, 2, 18, -9, -75, 28, 30, 1001 y 420 según sus clases de congruencia módulo m, para $m = 4, 7, 15$ y 28.

2. Exhibir un sistema completo de restos módulo 7 cuyos elementos sean potencias de números enteros (con exponente mayor que 1).

3. Sea $a \in \mathbb{N}$ tal que $a \equiv 4\ (8)$. Determinar en función de a la cantidad de números enteros pertenecientes al intervalo $\left[\frac{a-1}{4}\,,\,\frac{a+1}{2}\right]$.

4. Dados 11 números enteros entre 1 y 60, impares y no divisibles por 3, probar que dos de ellos difieren en 30.

5. Probar que todo $x \in \mathbb{Z}$ satisface al menos una de las siguientes congruencias:

$$x \equiv 0\ (2)\ ,\ x \equiv 0\ (3)\ ,\ x \equiv 1\ (4)\ ,\ x \equiv 1\ (6)\ ,\ x \equiv 11\ (12)\,.$$

6. Determinar el último dígito de F_{100} (cf. con el ejercicio 21 de la sección 5.1).

7. Calcular $r_a(b)$ en cada uno de los siguientes casos:

   a) $a = 8^{931}$, $b = 3$
   b) $a = 57^{294}$, $b = 5$
   c) $a = 3\,.\,9^{94} - 2\,.\,41^{23}$, $b = 11$.

8. ¿Qué día de la semana fue el 12 de octubre de 1492?

9. Imaginemos 40 tarjetas numeradas ubicadas correlativamente en una pila (la 1 en el tope y la 40 en el fondo). Se las mezcla de la siguiente forma: la 21 va al tope, luego la 1, luego la 22, luego la 2, etc., quedando en las 2 posiciones inferiores la 40 y y la 20 (en ese orden). Si se repite 10 veces esta operación de mezclado, ¿qué lugar ocupa al final la tarjeta 9?

10. Sea $a \in \mathbb{N}$ tal que $a \equiv 7\ (8)$. Probar que a no es suma de 3 cuadrados perfectos.

11. Sea n un número natural impar y sean $a_1, a_2, \ldots, a_n$ números enteros tales que $a_j \not\equiv a_k\ (n)$ si $j \neq k$. Probar que $\sum a_i$ es múltiplo de n.

12. Sea n primo y sean $a, b \in \mathbb{Z}$. Probar que $a^2 \equiv b^2\ (n)$ si y sólo si $b \equiv a\ (n)$ ó $b \equiv -a\ (n)$ ¿ Vale el mismo resultado si n es compuesto ?

13. *a*) Determinar las "raíces cuadradas" de -3 módulo 31, esto es, los $x \in \mathbb{Z}$ tales que $x^2 \equiv -3\ (31)$

    b) Caracterizar los $x \in \mathbb{Z}$ tales que $7x^2 + 9x + 3 \equiv 0\ (31)$

    c) ¿ Existe $x \in \mathbb{Z}$ tal que $5x^2 + 2x + 8 \equiv 0\ (19)$?

14. Sea p un primo y sean $a, b \in \mathbb{Z}$. Probar que $(a+b)^p \equiv a^p + b^p\ (p)$.

6.2. Ecuaciones modulares

6.2.1. Ecuación lineal de congruencia

Si $a, b \in \mathbb{Z}$ y $m \in \mathbb{N}$, una ecuación del tipo

$$aX \equiv b\ (m) \tag{6.1}$$

se dice una *ecuación lineal de congruencia*.

Estudiaremos en lo que sigue su resolubilidad, y en el caso de que exista alguna solución, exhibiremos un método para hallarlas todas.

Aclaremos en primer lugar qué entenderemos por resolverla. Si consideramos que una solución es un entero x tal que $ax \equiv b\ (m)$, la compatibilidad de la relación de congruencia con las operaciones nos muestra inmediatamente que cualquier entero congruente con x módulo m también es una solución. Debido a esto, lo razonable es pensar que una solución es una clase de congruencia módulo m, es decir, el problema consiste en hallar en $\mathbb{Z}_m$ las soluciones de la ecuación $(a)X = (b)$.

Puesto que $\mathbb{Z}_m$ es finito, para hallarlas podríamos intentar por ejemplo multiplicar la clase (a) por cada elemento de $\mathbb{Z}_m$ y verificar en qué casos se obtiene la clase (b). El método es seguro, pero claramente impracticable si m es grande, con el agravante de que ni siquiera sabemos previamente si existe una solución. Comencemos entonces por determinar una condición necesaria y suficiente para su resolubilidad:

Proposición 6.2.1 La ecuación (6.1) es resoluble si y sólo si $(a : m) \mid b$.

Demostracion Si x representa una solución, la condición $ax \equiv b\ (m)$ es equivalente a la existencia de un entero y tal que $ax + my = b$. Por lo tanto, la ecuación de congruencia (6.1) es resoluble si y solo si esta última ecuación diofántica lo es. Como vimos en el capítulo anterior, ello ocurre si y solo si $(a : m) \mid b$, como queríamos demostrar. Como caso particular importante, notemos que si a y m son coprimos entonces $aX \equiv b\ (m)$ es resoluble cualquiera sea b. $\Diamond$

Forma general de las soluciones.

Establecida la condición de resolubilidad de (6.1), y recordando que podemos hallar una solución de la ecuación diofántica asociada $ax + my = b$ a través del algoritmo de Euclides, veamos cómo hallar todas las soluciones. Para ello, supongamos que x_0 es una solución dada y que x es otra cualquiera. Empleando las notaciones $d = (a : m)$, $\alpha = a/d$ y $\mu = m/d$, y usando propiedades de las congruencias, obtenemos

$$ax \equiv b\ (m) \Leftrightarrow ax \equiv ax_0\ (m) \Leftrightarrow \alpha x \equiv \alpha x_0\ (\mu) \Leftrightarrow x \equiv x_0\ (\mu),$$

ya que podemos cancelar α de la penúltima relación por ser coprimo con μ.

Por lo tanto, la solución es única módulo μ, esto es, los enteros que resuelven la ecuación son todos los de la forma $x_k = x_0 + k\mu$ ($k \in \mathbb{Z}$). Puesto que nos interesan las soluciones módulo m, debemos determinar qué clases de congruencia módulo m recorren exactamente los x_k cuando k recorre $\mathbb{Z}$.

Para ello, tomemos cualquier k y escribámoslo en la forma $k = qd + r$, donde $0 \leq r < d$. Entonces:

$$x_k = x_0 + (qd + r)\mu = x_r + qm \equiv x_r \ (m).$$

Esto significa que la ecuación tiene a lo sumo d soluciones módulo m, a saber $x_0, x_1 \ldots, x_{d-1}$. Para completar nuestra tarea, demostremos que estos números no son congruentres entre sí módulo m, con lo que habremos probado que (6.1) admite exactamente d soluciones. Considerando a tal efecto un par de índices i, j tales que $0 \leq i < j < d$, tenemos que

$$0 < x_j - x_i = (j - i)\mu < d\mu = m,$$

de donde deducimos que $x_j \not\equiv x_i \ (m)$ por la propiedad 3) de la proposición 6.1.1.

En definitiva, y pensando en términos de clases de congruencia, si $d \mid b$ la ecuación $aX \equiv b \ (m)$ admite exactamente d soluciones en $\mathbb{Z}_m$, obtenidas a partir de una solución particular (x_0) en la forma

Fórmula 6.2.2

$$(x_0) \ , \ \left(x_0 + \frac{m}{d}\right) \ , \ \left(x_0 + 2\frac{m}{d}\right) \ , \ \ldots \ , \ \left(x_0 + (d-1)\frac{m}{d}\right) \qquad \Diamond$$

Ejemplo 6.2.3 Resolvamos la ecuación

$$152X \equiv -88 \ (36).$$

Tratándose de congruencias, es claro que podemos reducir los coeficientes módulo 36, por lo que es equivalente resolver la ecuación

$$8X \equiv 20 \ (36).$$

Puesto que $(8 : 36) = 4$ y $4 \mid 20$, resulta que la ecuación es resoluble y tiene 4 soluciones en $\mathbb{Z}_{36}$. Para hallar una, aplicamos el algoritmo que hemos exhibido en el capítulo 5 para escribir $(8 : 36)$ como combinación lineal de 8 y 36, obteniéndose la expresión $4 = 8(-4) + 36 \, . \, 1$, de donde $20 = 8(-20) + 36 \, . \, 5$, o en términos de congruencias: $8(-20) \equiv 20 \ (36)$. Luego (-20) es una solución de la ecuación. De acuerdo con la teoría (en este caso $m/d = 9$), las soluciones en $\mathbb{Z}_{36}$ son (-20), (-11), (-2) y (7). Si nos interesa escribirlas usando los representantes canónicos de las clases, concluimos que el conjunto de soluciones es

$$\{(7), (16), (25), (34)\}.$$

Equivalentemente, un número entero x satisface la relación

$$152x \equiv -88\,(36)$$

si y sólo si el resto de dividir x por 36 es 7, 16, 25 ó 34. $\Diamond$

Inverso modular.

Como caso particular del criterio general de resolubilidad de una ecuación lineal de congruencia, sigue que la ecuación $aX \equiv 1\ (m)$ es resoluble si y sólo si a y m son coprimos, en cuyo caso la solución es única módulo m. Este hecho admite la siguiente traducción en términos de clases:

Proposición 6.2.4 (a) es inversible en $\mathbb{Z}_m$ si y sólo si $a \perp m$.

DEMOSTRACION. Es consecuencia inmediata de la observación anterior, ya que la inversibilidad de (a) es equivalente a la existencia de un entero x tal que $ax \equiv 1\ (m)$. $\Diamond$

NOTA. En las condiciones del enunciado anterior, diremos que a es *inversible* módulo m. Si b es cualquier representante de la clase inversa de (a), esto es, $ab \equiv 1\ (m)$, diremos que b es un *inverso* de a módulo m. Por ejemplo, 13 es inversible módulo 20 y 17 es un inverso modular de 13, como el lector puede verificar. Debe quedar bien claro que los conceptos que acabamos de definir son conceptos modulares. Si a es inversible módulo m también lo es cualquier otro elemento de su clase de congruencia, y cualquier c congruente con b módulo m también es un inverso de a módulo m. Por ejemplo, podemos representar el inverso de 13 módulo 20 por -3.

En general, designaremos por $\mathbb{Z}_m^*$ el conjunto de clases inversibles módulo m, y llamaremos *sistema reducido de restos módulo* m a cualquier conjunto de representantes $\{x_1, x_2, \ldots, x_s\}$ de las clases inversibles módulo m. Equivalentemente,

$$\mathbb{Z}_m^* = \{(x_1), (x_2), \ldots, (x_s)\}\,.$$

Por ejemplo, $\{1, 5, 7, 11\}$ es un sistema reducido de restos módulo 12 y $\{1, 3, 5, -5, -3, -1\}$ es un sistema reducido de restos módulo 14 (verificar). Más genéricamente, $\{1, 2, \ldots, p-1\}$ es un sistema reducido de restos módulo p cualquiera sea el primo p, ya que todo número natural menor que p es coprimo con p. $\Diamond$

La característica que acabamos de señalar (el conjunto de restos no nulos es un sistema reducido de restos) es exclusiva de los números primos. Precisamente, y utilizando el lenguaje de las estructuras algebraicas, vale el siguiente resultado:

Proposición 6.2.5 $\mathbb{Z}_m$ es un cuerpo si y sólo si m es primo.

DEMOSTRACION Ya vimos que toda clase no nula módulo m es inversible si m es primo. Recíprocamente, supongamos que $\mathbb{Z}_m$ es un cuerpo y tomemos cualquier divisor d de $m\,(1 \leq d < m)$. Puesto que (d) es inversible en $\mathbb{Z}_m$, sigue por 6.2.4 que d es coprimo con m, de donde resulta que $d = (d : m) = 1$. Luego m es primo. $\Diamond$

6.2.2. Teorema chino del resto

Para presentar el tema, examinemos el siguiente arreglo matricial de números enteros:

$$\begin{array}{ccccc} 0 & 16 & 12 & 8 & 4 \\ 5 & 1 & 17 & 13 & 9 \\ 10 & 6 & 2 & 18 & 14 \\ 15 & 11 & 7 & 3 & 19 \end{array}$$

Por simple inspección, vemos que sus elementos representan todas las clases de congruencia módulo 20. Si bien en principio podría parecer que están caprichosamente distribuidos, ello no es así: cada fila i consiste de aquellos números entre 0 y 19 que son congruentes con $i-1$ módulo 4 y cada columna j de aquellos que son congruentes con $j-1$ módulo 5. El hecho de que a partir de esta construcción aparezcan todas las clases módulo 20, ó equivalentemente, que no se repita ninguna, nos muestra que la clase de congruencia módulo 20 de un número queda completamente determinada por su clase módulo 4 y por su clase módulo 5. Dicho de otra manera, si conocemos los restos de dividir un número por 4 y por 5, entonces conocemos el resto de dividirlo por 20. Por ejemplo, supongamos que $r_4(x) = 2$ y $r_5(x) = 3$. De acuerdo con las reglas de diseño que fijamos, la clase módulo 20 de x viene dada por el elemento situado en el lugar $(3, 4)$ de la matriz, y por lo tanto $r_{20}(x) = 18$.

Notemos que no ocurre lo mismo en general. Por ejemplo, si un número x es congruente con 2 módulo 4 y con 8 módulo 10, su clase módulo 40 no queda bien determinada, ya que podría ser $x \equiv 18\ (40)$ ó $x \equiv 38\ (40)$.

Obsérvese que en ambas situaciones hemos considerado clases de congruencia de un entero con respecto a dos números, y nos preguntamos si ellas determinan su clase de congruencia respecto al producto de los mismos. La respuesta es afirmativa en el primer caso y negativa en el segundo, y como pronto veremos, la diferencia radica en que los factores sean o no coprimos entre sí. El llamado *teorema chino del resto* (Sun-Tzu, siglo III), que enunciaremos y demostraremos a continuación, generaliza y resuelve esta cuestión.

Teorema 6.2.6 (Chino del resto) Sea $m_1, m_2, \ldots, m_r$ una secuencia de números naturales coprimos dos a dos y sea $a_1, a_2, \ldots, a_r$ una secuencia cualquiera de números enteros. Existe entonces un número entero x verificando

simultáneamente las congruencias

$$\begin{cases} x \equiv a_1 \ (m_1) \\ x \equiv a_2 \ (m_2) \\ \vdots \quad \vdots \qquad \vdots \\ x \equiv a_r \ (m_r) \end{cases}$$

Además, x es único módulo el producto $m_1 m_2 \dots m_r$.

DEMOSTRACION. Para cada índice k el producto

$$M_k = \prod_{i \neq k} m_i$$

es coprimo con m_k, pues cada factor m_i es coprimo con m_k. Por lo tanto M_k es inversible módulo m_k y existe un entero v_k (que podemos calcular mediante el algoritmo de Euclides) tal que $M_k v_k \equiv 1 \ (m_k)$.

Observando que $M_k v_k \equiv 0 \ (m_i)$ para cualquier índice $i \neq k$, ya que en tal caso m_i es un factor de M_k, sea

$$x = \sum_{k=1}^{r} M_k v_k a_k \, .$$

Tomando congruencias módulo m_i para cualquier índice i resulta que

$$x = \sum_{k=1}^{r} M_k v_k a_k \equiv M_i v_i a_i \equiv a_i \ (m_i),$$

y por lo tanto x satisface los requerimientos del enunciado.

Para estudiar la forma general de las soluciones designemos por m el producto $m_1 m_2 \cdots m_r$, y observemos en primer lugar que todo y congruente con x módulo m también es solución, pues en tal caso $y \equiv x \ (m_i)$ para todo i.

Recíprocamente, sea z una solución cualquiera. Sigue entonces por transitividad que $z \equiv x \ (m_i)$ para todo i, ó equivalentemente, $m_i \mid z - x$ para todo i. Puesto que por hipótesis los factores m_i son coprimos dos a dos, concluimos entonces que $m \mid z - x$, esto es, $z \equiv x \ (m)$. Luego la solución es única módulo m, como queríamos probar. $\Diamond$

En los ejercicios del final de la sección, encomendaremos al lector la tarea de estudiar el problema cuando los módulos no necesariamente satisfacen la hipótesis de ser mutuamente coprimos. Veamos ahora un par de ejemplos de aplicación del teorema chino del resto.

Ejemplo 6.2.7 Supongamos que buscamos un entero tal que los restos de dividirlo por 5, 6 y 7 sean 2, 0 y 4, respectivamente. Si partimos de 4 y

vamos sumando 7 (para obtener siempre números congruentes con 4 módulo 7), en el segundo intento llegamos a 18, que verifica también la segunda condición, aunque no la primera. Si ahora vamos sumando 42 (para mantener las últimas dos condiciones), obtenemos sucesivamente 60 y 102, y arribamos así a una solución. Puesto que $5 \,.\, 6 \,.\, 7 = 210$, la segunda parte del teorema chino nos dice que los enteros que satisfacen las condiciones requeridas son los de la forma $210k + 102$, pues la solución es única módulo 210. $\Diamond$

Es innegable que la anterior forma de resolución constituye un método, ya que el teorema nos asegura que así daremos con una solución. Sin embargo, el mismo no resulta eficaz cuando hay muchos factores involucrados o los números son demasiado grandes. Le ofreceremos ahora otro ejemplo, en el que emplearemos para buscar una solución el método descripto en la demostración del teorema.

Ejemplo 6.2.8 Como ilustración del teorema chino del resto, veamos cómo hallar cuatro números naturales consecutivos divisibles por 13, 15, 17 y 19, respectivamente. Designando por x el menor de los cuatro números a determinar, observemos que las condiciones requeridas son equivalentes a la validez simultánea de las congruencias

$$\begin{cases} x \equiv 0 \ (13) \\ x \equiv -1 \ (15) \\ x \equiv -2 \ (17 \\ x \equiv -3 \ (19) \end{cases}$$

Puesto que 13, 15, 17 y 19 son coprimos dos a dos, nos hallamos ante un caso particular del teorema, lo que asegura que el problema tiene solución. Siguiendo los pasos de la demostración, resulta en este caso que $M_1 = 4845$, $M_2 = 4199$, $M_3 = 3705$ y $M_4 = 3315$, como puede verificar el lector.

Nuestra próxima tarea es determinar para cada i el inverso de M_i módulo m_i, para lo cual, reduciendo convenientemente en cada caso, debemos resolver las ecuaciones $9X \equiv 1 \ (13)$, $14X \equiv 1 \ (15)$, $16X \equiv 1 \ (17)$ y $9X \equiv 1 \ (19)$, cuyas soluciones son, como puede comprobarse, $v_1 = 3$, $v_2 = 14$, $v_3 = 16$ y $v_4 = 17$, respectivamente. De acuerdo con la demostración del teorema, obtenemos una solución tomando

$$x = 4845 \,.\, 3 \,.\, 0 + (-1)4199 \,.\, 14 + (-2)3705 \,.\, 16 + (-3)3315 \,.\, 17 = -346411.$$

Siendo $m = 62985$, toda solución es de la forma $-346411 + 62985k$, con $k \in \mathbb{Z}$. Puesto que en nuestro problema ella debe ser positiva, deberemos tomar $k > 0$. Un rápido cálculo nos muestra que $k = 6$ brinda el menor valor positivo posible, que corresponde a la solución $x = 31499$. Por lo tanto, la primera cuaterna de números naturales consecutivos que satisface los requerimentos impuestos es

$$(31499, 31500, 31501, 31502). \quad \Diamond$$

6.2.3. Ejercicios

1. Resolver las siguientes ecuaciones de congruencia:

   a) $21X \equiv 60\ (15)$

   b) $99X \equiv -15\ (13)$

   c) $45X \equiv 96\ (18)$

   d) $85X \equiv 70\ (30)$.

2. Resolver el sistema de ecuaciones

$$\begin{cases} 7X + 3Y \equiv -3\ (23) \\ 2X - 5Y \equiv 8\ (23)\,. \end{cases}$$

3. Hallar los números naturales x de 4 cifras tales que el desarrollo decimal de $18x$ termina en 1256.

4. Calcular cuántos términos de la progresión aritmética $6, 21, \ldots, 2991$ son divisibles por 12 ó por 13.

5. Sea $m \in \mathbb{N}$.

   a) Demostrar que $\mathbb{Z}_m^*$ es cerrado respecto al producto de clases.

   b) Demostrar que $\mathbb{Z}_m^*$ es un grupo respecto al producto de clases.

   c) Si a es un entero inversible módulo m, probar que la multiplicación por a determina una permutación de $\mathbb{Z}_m^*$.

   d) Construir las tablas de multiplicación de $\mathbb{Z}_{36}^*$ y $\mathbb{Z}_{11}^*$.

6. Sea $n \in \mathbb{N}$ $(n > 1)$. Probar que n es primo si y solo si $(n-1)\,! \equiv -1\ (n)$.

7. Si p es primo, determinar los restos de dividir $(p-2)\,!$ y $\binom{2p}{p}$ por p.

8. Probar que $1 + 61\,!$ y $1 + 63\,!$ son múltiplos de 71.

9. *a*) Determinar el mínimo $m \in \mathbb{N}$ que verifica simultáneamente las condiciones $r_4(m) = 5$, $r_5(m) = 2$ y $r_7(m) = 4$.

   b) Como en *a*), verificando $2m \equiv 3\ (5)$, $5m \equiv 2\ (6)$ y $3m \equiv 1\ (7)$.

10. Resolver completamente la ecuación $x^3 \equiv x^2\ (45)$.

11. Hallar un cuadrado perfecto c tal que $c \equiv 7\ (9)$ y $c \equiv 5\ (19)$.

12. Sean $m, n \in \mathbb{N}$ y sean $a, b \in \mathbb{Z}$. Demostrar que el sistema

$$\begin{cases} x \equiv a\ (m) \\ x \equiv b\ (n) \end{cases}$$

es resoluble si y sólo si $a \equiv b \mod (m : n)$. Probar que en tal caso la solución es única módulo $[\, m : n \,]$.

13. Caracterizar los $a \in \mathbb{Z}$ tales que a y $a + 3$ son divisibles por 69 y 120, respectivamente.

14. *a*) Determinar el resto de dividir 65^{60} por 110.
    b) Caracterizar los $a \in \mathbb{Z}$ tales que $63 \mid 2^{250}a + 13^{79}$.
    c) Caracterizar los $a \in \mathbb{N}$ tales que $(8a^{43} + a : 9a) \neq a$.

6.3. Teorema de Fermat

6.3.1. Estructura multiplicativa

En lo que sigue p designará un número primo. Recordando que $\mathbb{Z}_p$ es un cuerpo (proposición 6.2.5), dedicaremos nuestra atención a estudiar la estructura multiplicativa del conjunto $\mathbb{Z}_p^*$ de clases inversibles módulo p, representadas por los elementos del intervalo $\mathbb{I}_{p-1}$. Comenzaremos demostrando un importante resultado debido a Pierre de Fermat (siglo XVII), familiarmente conocido como *pequeño teorema de Fermat*, cuyo enunciado es el siguiente:

Teorema 6.3.1 (Fermat) Sea a un entero no divisible por p. Entonces

$$a^{p-1} \equiv 1\ (p)\,.$$

DEMOSTRACION. Puesto que $a \perp p$ (por ser p primo), la multiplicación por (a) determina una permutación de $\mathbb{Z}_p^*$. Expresado en lenguaje de congruencias, existe un reordenamiento $x_1, x_2, \ldots, x_{p-1}$ de los elementos del intervalo I_{p-1} tal que $ja \equiv x_j\ (p)$ para todo $j \in \mathbb{I}_{p-1}$.

Multiplicando ahora miembro a miembro las $p-1$ congruencias anteriores, obtenemos:

$$a^{p-1}(p-1)\,! \equiv \prod_{t=1}^{p-1} x_t = (p-1)\,!\ (p),$$

ya que ambos miembros de la última igualdad consisten del producto de los mismos factores, posiblemente escritos en diferente orden. Puesto que $(p-1)\,!$ es coprimo con p (ninguno de sus factores es divisible por p), podemos cancelarlo en la relación anterior, resultando finalmente que:

$$a^{p-1} \equiv 1\ (p),$$

como queríamos demostrar. $\lozenge$

NOTA. En la demostración anterior hemos usado reiteradamente la primalidad de p. De hecho el resultado no es válido si el módulo es compuesto, por ejemplo $2^8 \equiv 4\ (9)$. Asimismo, la hipótesis de que $p \nmid a$ es esencial, pues de otro modo $a^{p-1} \equiv 0\ (p)$.

Notemos por último que el teorema muestra que a^{p-2} es el inverso de a módulo p, ya que

$$a\, a^{p-2} = a^{p-1} \equiv 1\ (p)\,. \quad \lozenge$$

El siguiente corolario brinda una versión equivalente de 6.3.1 (algunos llaman teorema de Fermat a este enunciado):

Corolario 6.3.2 $a^p \equiv a\ (p)$ para todo $a \in \mathbb{Z}$.

DEMOSTRACION. Si $p \nmid a$ el resultado sigue multiplicando miembro a miembro por a la congruencia $a^{p-1} \equiv 1\ (p)$, mientras que si a es múltiplo de p tenemos $a^p \equiv 0 \equiv a\ (p)$.

Respecto a la equivalencia que mencionamos arriba, tomemos como válido este enunciado y supongamos que a es un entero no divisible por p. Sigue entonces de las relaciones

$$0 \equiv a^p - a = a(a^{p-1} - 1)\ (p)$$

que $p \mid a^{p-1} - 1$, pues p es primo y $p \nmid a$. Luego $a^{p-1} \equiv 1\ (p)$. $\quad \Diamond$

Ejemplo 6.3.3 Como típica aplicación del teorema de Fermat, calculemos el resto de dividir $x = 2^{979}$ por 55. Si bien el módulo no es primo ($55 = 5 \,.\, 11$), podemos usar Fermat para determinar las clases de x módulo 5 y módulo 11, y utilizar luego el teorema chino del resto. Tenemos

$$2^{979} = 2^{4\,.\,244+3} = \left(2^4\right)^{244} 2^3 \equiv 8 \equiv 3\ (5)$$

y

$$2^{979} = 2^{10\,.\,97+9} = \left(2^{10}\right)^{97} 2^9 \equiv 512 \equiv 6\ (11).$$

Siendo $x \equiv 3\ (5)$ y $x \equiv 6\ (11)$, una sencilla aplicación del teorema chino nos muestra que $x \equiv 28\ (55)$, esto es, $r_{55}(2^{979}) = 28$. $\quad \Diamond$

Orden modular.

El caso primo En conexión con el teorema de Fermat, el lector puede verificar sin dificultad la validez de las relaciones $3^5 \equiv 1\ (11)$ y $10^6 \equiv 1\ (13)$, lo que muestra que dado un número entero a, no divisible por un cierto primo p, es posible que $a^k \equiv 1\ (p)$ para algún $k < p - 1$. Esto conduce naturalmente a considerar el mínimo exponente con dicha propiedad, motivo de la siguiente definición:

> Si p es primo y a es un entero no divisible por p, definimos el *orden* de a módulo p en la forma
>
> $$\operatorname{ord}_p a = \min\left\{k \in \mathbb{N} : a^k \equiv 1\ (p)\right\}.$$

Notemos que el teorema de Fermat garantiza que el conjunto anterior es no vacío, resultando de paso que $\operatorname{ord}_p a \leq p - 1$. Es claro también que la definición anterior sólo depende de la clase de a módulo p, vale decir, $\operatorname{ord}_p b = \operatorname{ord}_p a$ si $b \equiv a\ (p)$.

Tenemos por ejemplo que $\operatorname{ord}_p 1 = 1$, $\operatorname{ord}_{11} 10 = 2$ y $\operatorname{ord}_{13} 2 = 12$. Es claro que este último caso exige comprobar que $2^k \not\equiv 1\ (13)$ para todo $k < 12$, aunque probaremos a continuación (entre otras cosas) que es posible reducir considerablemente el rango de valores a verificar.

Proposición 6.3.4 Sea $a \in \mathbb{Z}$ no divisible por un primo p y sea $m = \operatorname{ord}_p a$. Valen entonces las siguientes propiedades (en todos los casos los exponentes son enteros no negativos):

1) $a^t \equiv 1\ (p) \Leftrightarrow m \mid t$

2) $m \mid p-1$

3) $a^u \equiv a^v\ (p) \Leftrightarrow u \equiv v\ (m)$

4) Dado t existe un único i $(0 \leq i < m)$ tal que $a^t \equiv a^i\ (p)$. Deducimos de ello que los elementos

$$1,\, a,\, a^2, \ldots,\, a^{m-1}$$

representan m clases distintas en $\mathbb{Z}_p^*$.

Demostracion.

1) Supongamos primero que $a^t \equiv 1\ (p)$. Dividiendo t por m, y escribiendo $t = qm + r$ $(0 \leq r < m)$, resulta entonces que

$$1 \equiv a^t = (a^m)^q a^r \equiv a^r\ (p)\,.$$

Luego $r \notin \mathbb{N}$, por la minimalidad del exponente m. Por lo tanto $r = 0$ y $m \mid t$.

La recíproca es inmediata, ya que si $m \mid t$ resulta que

$$a^t = (a^m)^{t/m} \equiv 1\ (p)\,.$$

2) Sigue inmediatamente de 1) ya que $a^{p-1} \equiv 1\ (p)$. Usando esta propiedad en el ejemplo precedente, observemos que basta verificar que $2^k \not\equiv 1\ (13)$ para $k = 1, 2, 3, 4, 6$.

3) Asumamos sin pérdida de generalidad que $u \leq v$. Si $u \equiv v\ (m)$, escribamos $v = u + km$, con $k \in \mathbb{N}_0$. Entonces

$$a^v = a^u\,(a^m)^k \equiv a^u\ (p)\,,$$

como queríamos probar. Inversamente, sea $x = a^u \equiv a^v\ (p)$. Operando, tenemos:

$$x a^{v-u} \equiv a^u a^{v-u} = a^v \equiv x\ (p)\,.$$

Siendo x coprimo con p podemos cancelarlo en la relación anterior, resultando que $a^{v-u} \equiv 1\ (p)$. Sigue luego de 1) que $m \mid v-u$, o sea, $u \equiv v\ (m)$.

4) Es claro que este enunciado se deduce de 3) ya que para todo $t \in \mathbb{Z}$ existe un único i en el rango $0 \leq i < m$ tal que $t \equiv i\ (m)$, a saber, $i = r_m(t)$). Lo hemos incluido para enfatizar el hecho de que existen exactamente m potencias distintas de a módulo p, repitiéndose periódicamente el ciclo

$$1\,,\, a\,,\, a^2\,,\, \ldots\,,\, a^{m-1}.$$

Por ejemplo, las potencias de 5 módulo 13 repiten periódicamente el ciclo 1, 5, 12, 8, pues $\operatorname{ord}_{13} 5 = 4$. En cambio las potencias de 2 constituyen un sistema reducido de restos módulo 13, pues $\operatorname{ord}_{13} 2 = 12$, como ya hemos señalado. O sea, toda clase no nula módulo 13 se realiza como una potencia de 2. $\Diamond$

PRIMOS DE MERSENNE El teorema de Fermat está históricamente ligado a la determinación de la primalidad de ciertos números. Alrededor de 1640, el monje Marin Mersenne comunicó a Fermat y a Descartes que 2^n-1 resultaba ser primo para muchos valores del exponente, por ejemplo 2, 3, 5 y 7. No es casualidad que estos números sean a su vez primos, ya que vale en general el siguiente resultado:

Lema 6.3.5 Si $2^n - 1$ es primo entonces n es primo.

DEMOSTRACION. Si $n = dt$, utilizando un conocido caso de factoreo deducimos que $2^d - 1 \mid 2^n - 1$, ya que:

$$2^n - 1 = \left(2^d\right)^t - 1 = \left(2^d - 1\right)\left(2^{d(t-1)} + \cdots + 2^d + 1\right).$$

Sigue entonces por hipótesis que $2^d - 1 = 1$ ó $2^d - 1 = 2^n - 1$, de donde concluimos que $d = 1$ ó $d = n$. Luego n es primo. $\Diamond$

Leibniz (entre otros) creyó erróneamente que también valía la recíproca, lo que brindaría una forma sistemática de generar números primos. Curiosamente, es muy sencillo hallar un contraejemplo, ya que $2^{11} - 1 = 23 \,.\, 89$ no es primo, a pesar de que 11 lo es. En homenaje al monje francés, un número de la forma $2^p - 1$ (p primo), que designaremos por M_p, se llama un *número de Mersenne*. Como dato ilustrativo, señalemos que los primeros 9 números de Mersenne que resultan primos (primos de Mersenne) son M_2, M_3, M_5, M_7, M_{13}, M_{17}, M_{19}, M_{31} y M_{61}.

Las indagaciones iniciadas por Fermat lo llevaron a formular una primera versión de su teorema, a saber que $2^p - 2$ es múltiplo de p para todo primo p. Su labor fue luego continuada por otros grandes matemáticos, principalmente por Euler y Gauss, que posiblemente en esa dirección desarrollaron la teoría de los residuos cuadráticos.

El problema de estudiar la primalidad de los números de Mersenne sigue plenamente vigente. Una de las conjeturas famosas de la teoría de números es que existen infinitos primos de Mersenne, pero a pesar de los esfuerzos desplegados aún no se ha logrado demostrarla. Se sabe que muchos números de Mersenne no son primos, y se conocen cuarenta y siete que sí lo son, habiéndose hallado en 2009 el cuadragésimo séptimo, a saber $2^{42643801} - 1$, un número cuyo desarrollo decimal supera las 12 millones de cifras. Si bien la cuestión es compleja desde el punto de vista computacional, veamos con un ejemplo cómo el manejo de la teoría ayuda a determinar la primalidad de tales números.

Ejemplo 6.3.6 $M_{19} = 524287$ es primo. Si bien es un número de tamaño ciertamente modesto, nuestro propósito es ilustrar cómo los resultados que hemos establecido permiten reducir notablemente los cálculos. Así, en vez de verificar directamente que 524287 no es divisible por ninguno de los 128 primos menores o iguales que su raíz cuadrada $(724,07\ldots)$, estudiemos la forma de sus posibles divisores primos.

Supongamos entonces que $q \mid 2^{19}-1$, ó equivalentement, $2^{19} \equiv 1\ (q)$. Deducimos entonces que $\mathrm{ord}_q\, 2 = 19$, y por lo tanto $19 \mid q-1$, digamos $q-1 = 19k$. Más aún, k debe ser de la forma $2t$, pues $q-1$ es par. Luego, los posibles divisores primos de M_{19} son de la forma $q = 19k+1 = 38t+1$. Esto reduce muchísimo nuestra tarea, pues ahora sólo debemos buscar posibles divisores entre aquellos primos menores que 724 y de la forma anterior. Un vistazo a una tabla de primos nos muestra que existen 6 de tales números, a saber, 191, 229, 419, 457, 571 y 647. Esto es, hemos descartado sin hacer una sola cuenta a 122 de los 128 candidatos. Como ninguno de ellos lo divide (simple verificación), concluimos que $2^{19}-1$ es primo. $\Diamond$

NUMEROS PERFECTOS Los números de Mersenne están conectados a su vez con los llamados *números perfectos*, que ya aparecen en civilizaciones tan antiguas como la egipcia y la babilónica. Un número natural n se dice perfecto si coincide con la suma de sus divisores propios. Equivalentemente, la suma de todos los divisores de n es $2n$. Por ejemplo, $6 = 1+2+3$ es el menor número perfecto. Desde tiempos remotos estos números llamaron la atención de matemáticos como Euclides, Fermat, Euler, Gauss, etc, y muchos resultados que hoy conocemos tuvieron origen en las investigaciones realizadas en torno a los mismos.

Los primeros números perfectos, son $P_1 = 6$, $P_2 = 28$, $P_3 = 496$ y $P_4 = 8128$. Se observan a simple vista algunos patrones (parecería haber un número perfecto para cada número posible de cifras, terminan alternativamente en 6 y en 8), pero no son características esenciales. Sí lo son otras analogías que estos números presentan, que se descubren en cuanto los factorizamos:

$$\begin{aligned} P_1 &= 2\,.\,3 = 2\,(2^2-1) \\ P_2 &= 2^2\,.\,7 = 2^2\,(2^3-1) \\ P_3 &= 2^2\,.\,31 = 2^4\,(2^5-1) \\ P_4 &= 2^6\,.\,127 = 2^6\,(2^7-1)\,. \end{aligned}$$

Todos ellos son el producto de un primo de Mersenne por una potencia de 2, y los exponentes en cuestión son consecutivos. Es decir, tienen la forma general $2^{p-1}M_p$, con M_p primo. Probaremos a continuación que esta es una condición necesaria y suficiente para que un número par sea perfecto.

Proposición 6.3.7 Un número par a es perfecto si y sólo si $a = 2^{p-1}M_p$, donde p y M_p son primos. En consecuencia, existe una correspondencia biunívoca entre números perfectos pares y primos de Mersenne.

DEMOSTRACION. Si $a = 2^{p-1}M_p$, con p y M_p primos, sigue del TFA que los divisores de a son de dos tipos: los de la forma 2^k y los de la forma $2^k M_p$, donde en ambos casos $0 \leq k \leq p-1$. Si los sumamos, obtenemos:

$$\sum_{k=0}^{p-1} 2^k + \sum_{k=0}^{p-1} 2^k M_p = (M_p + 1) \sum_{k=0}^{p-1} 2^k = 2^p(2^p - 1) = 2a,$$

lo que demuestra que a es perfecto.

Recíprocamente, supongamos que a es un número par perfecto, digamos $a = 2^{m-1}b$, donde $m \geq 2$ y b es impar. Observemos entonces que todo divisor de a es de la forma $2^k d$, donde $d \mid b$ y $0 \leq k \leq m-1$. Designando por s la suma de los divisores de b, tenemos:

$$2^m b = 2a = \sum_k \sum_d 2^k d = \sum_k 2^k \sum_d d = (2^m - 1)s.$$

Teniendo en cuenta que $2^m - 1$ y 2^m son coprimos, deducimos de la igualdad anterior que $2^m - 1 \mid b$. Escribiendo $b = (2^m - 1)c$ resulta entonces que

$$s = 2^m c = (2^m - 1)\, c + c = b + c\,,$$

y siendo s la suma de *todos* los divisores de b, concluimos que b y c son sus únicos divisores. Es claro que esto sólo puede ocurrir si b es primo y $c = 1$, es decir, $b = 2^m - 1$. Luego m es primo y a tiene la forma requerida. $\Diamond$

NOTA Consignemos que la suficiencia de la condición del enunciado fue demostrada por Euclides, mientras que la caracterización completa fue obtenida por Euler unos 2000 años más tarde. Respecto a números perfectos impares, señalemos que no se conoce ninguno y se conjetura que no existen. Hay muchos trabajos sobre el tema, desde Euler hasta nuestros días. Se sabe por ejemplo que un número perfecto impar debiera ser mayor que 10^{300} y divisible por al menos 10 primos distintos.

El caso general Habiendo analizado la estructura multiplicativa de los anillos de restos en el caso primo, nos proponemos ahora estudiar la estructura multiplicativa de $\mathbb{Z}_n^*$ cualquiera sea $n \in \mathbb{N}$. Con la mira puesta en un resultado que extienda el pequeño teorema de Fermat, comencemos con la siguiente definición:

> Llamaremos *indicador de Euler* de un número natural n al cardinal de $\mathbb{Z}_n^*$, que notaremos $\varphi(n)$. Precisamente,
>
> $$\varphi(n) = \#\left(\{k \in \mathbb{I}_m \,:\, k \perp m\}\right).$$

Por ejemplo $\varphi(1) = 1$ y $\varphi(p) = p - 1$ si p es primo, ya que todo natural menor que p es coprimo con p. Más particularmente, no es arduo verificar que $\varphi(11) = 10$, $\varphi(18) = 6$ y $\varphi(40) = 16$, aunque debemos señalar que no existe en general una manera efectiva de calcular $\varphi(n)$, salvo en ciertas situaciones especiales. El siguiente resultado va en esa dirección, ya que nos brindará una fórmula para calcular $\varphi(n)$ en caso de que se conozca la factorización de n.

Proposición 6.3.8 La función de Euler satisface las siguientes propiedades:

1) Si $s \in \mathbb{N}$ y p es primo entonces $\varphi(p^s) = p^{s-1}(p-1)$

2) Si $m, n \in \mathbb{N}$ y $m \perp n$ entonces $\varphi(mn) = \varphi(m)\varphi(n)$

3) Sean $m_1, m_2, \ldots, m_r$ números naturales coprimos dos a dos. Entonces

$$\varphi\left(\prod_{i=1}^{r} m_i\right) = \prod_{i=1}^{r} \varphi(m_i).$$

Demostracion. Para demostrar 1), observemos que un número es coprimo con una potencia de p si y sólo si no es divisible por p. Por lo tanto, para calcular $\varphi(p^s)$ bastará proceder por descarte, contando la cantidad de múltiplos de p menores o iguales que p^s. Puesto que en general el número de múltiplos de a menores o iguales que b es el cociente de dividir b por a, concluimos que hay p^{s-1} múltiplos de p en $\mathbb{I}_{p^s}$. Luego,

$$\varphi(p^s) = p^s - p^{s-1} = p^{s-1}(p-1)\,.$$

Respecto de 2), observemos en primer lugar que el resultado es trivial si $m = 1$ ó $n = 1$, ya que $\varphi(1) = 1$. Supondremos por lo tanto que ambos números son mayores que 1. El hecho de que un entero es coprimo con un producto si y sólo si es coprimo con cada uno de los factores nos permite en este caso definir una función

$$f : \mathbb{Z}_{mn}^* \to \mathbb{Z}_m^* \times \mathbb{Z}_n^*$$

mediante la fórmula $f\left((a)_{mn}\right) = \left((a)_m, (a)_n\right)$, donde en general $(a)_t$ denota la clase de congruencia de a módulo t.

El teorema chino del resto (que podemos aplicar pues m y n son coprimos) dice exactamente que f *es biyectiva*, por lo que en particular ambos conjuntos tienen el mismo cardinal. En consecuencia,

$$\varphi(mn) = \#\left(\mathbb{Z}_{mn}^*\right) = \#\left(\mathbb{Z}_m^* \times \mathbb{Z}_n^*\right) = \#\left(\mathbb{Z}_m^*\right)\#\left(\mathbb{Z}_n^*\right) = \varphi(m)\varphi(n).$$

La última propiedad sigue inmediatamente de 2) por inducción en r. No está de más apuntar que la hipótesis de coprimalidad es necesaria, ya que por ejemplo $\varphi(40) = 16$, mientras que $\varphi(4)\varphi(10) = 8$. $\Diamond$

La propiedad 3) de arriba nos conduce a la fórmula para el cálculo de φ a la que aludimos anteriormente, teniendo en cuenta que todo número natural es producto de potencias de números primos distintos, y por lo tanto mutuamente coprimos. Precisamente, supongamos que

$$n = \prod_{i=1}^{s} p_i^{n_i}$$

es la factorización de un número natural n. Vale entonces la fórmula:

Fórmula 6.3.9

$$\varphi(n) = \prod_{i=1}^{s} p_i^{n_i-1}\,(p_i - 1)$$

Por ejemplo, $\varphi(360) = \varphi(2^3 3^2 5) = 4\,.\,6\,.\,4 = 96$. Si bien la fórmula es sencilla, recalquemos que su utilización supone conocer la factorización de n, que como ya indicamos anteriormente puede ser un problema de difícil solución si n es muy grande. $\Diamond$

El siguiente teorema, demostrado por Leonhard Euler en 1736, generaliza el pequeño teorema de Fermat.

Teorema 6.3.10 (Fermat-Euler) Sea $n \in \mathbb{N}$ y sea a un entero coprimo con n. Entonces

$$a^{\varphi(n)} \equiv 1\ (n)\,.$$

Demostracion. Notemos antes de probar su validez que el enunciado es realmente una extensión del teorema de Fermat, ya que $\varphi(n) = n - 1$ si n es primo y en tal caso las afirmaciones "a es coprimo con n" y "a no es divisible por n" son equivalentes.

En cuanto a la prueba, sea $I = \left\{x_1, x_2, \ldots, x_{\varphi(n)}\right\}$ un sistema reducido de restos módulo n. Puesto que $a \perp n$, la multiplicación por (a) permuta los elementos de $\mathbb{Z}_n^*$. Por lo tanto, para cada índice i existe un índice $j(i)$ tal que

$$ax_i \equiv x_{j(i)}\ (n),$$

siendo $\left\{x_{j(1)}, x_{j(2)}, \ldots, x_{j(\varphi(n))}\right\} = I$. Multiplicando miembro a miembro estas $\varphi(n)$ congruencias, obtenemos:

$$a^{\varphi(n)} \prod_i x_i \equiv \prod_i x_{j(i)} = \prod_i x_i\ (n).$$

Puesto que el producto de los elementos de I es coprimo con n (por serlo sus factores), podemos cancelarlo en la relación anterior. Resulta entonces que $a^{\varphi(n)} \equiv 1\ (n)$, como queríamos probar. $\Diamond$

Por ejemplo, $a^6 \equiv 1\ (18)$ si a es impar y no divisible por 3.

Si a es coprimo con n se define de manera totalmente análoga al caso primo el orden de a módulo n, en la forma:

$$\operatorname{ord}_n a = \min\left\{k \in \mathbb{N} : a^k \equiv 1\ (n)\right\}.$$

Con las modificaciones obvias, todas las propiedades descriptas en la proposición 6.3.4 se extienden sin dificultad al caso general, según el siguiente detalle:

Proposición 6.3.11 Sean $n \in \mathbb{N}$ y $a \in \mathbb{Z}$ coprimos y sea $m = \operatorname{ord}_n a$. Entonces

1) $a^t \equiv 1\ (n) \Leftrightarrow m \mid t$

2) $m \mid \varphi(n)$

3) $a^u \equiv a^v\ (n) \Leftrightarrow u \equiv v\ (m)$

4) Dado t existe un único i $(0 \leq i < m)$ tal que $a^t \equiv a^i\ (n)$. Deducimos de ello que los elementos

$$1,\ a,\ a^2, \ldots,\ a^{m-1}$$

representan m clases distintas en $\mathbb{Z}_n^*$.

Demostracion. Queda a cargo del lector. $\Diamond$

Ejemplo 6.3.12 En el ejemplo 6.1.3 vimos que el desarrollo decimal de $a = 7^{7^7}$ termina en 3. Calculemos ahora sus dos últimas cifras. Puesto que el número determinado por las últimas dos cifras de un número coincide con el resto de dividirlo por 100, bastará calcular las clases de congruencia de a módulo 4 y módulo 25, ya que $100 = 4 \,.\, 25$ y estos números son coprimos entre sí. Emplearemos en ambos casos el teorema de Fermat-Euler, ya que 4 y 25 no son primos.

Con respecto a 4, observando que $\varphi(4) = 2$ y que 7^7 es de la forma $2k+1$, tenemos:

$$a = 7^{\,2k+1} = \left(7^2\right)^k 7 \equiv 7 \equiv 3\ (4).$$

Para calcular la clase de a módulo 25 será conveniente estudiar la clase módulo 20 del exponente 7^7, ya que $\varphi(25) = 20$. A tal efecto, tomamos congruencias módulo 4 y 5, obteniendo:

$$7^7 \equiv (-1)^7 = -1 \equiv 3\ (4)$$

y

$$7^7 \equiv 2^7 = 2^4 2^3 \equiv 8 \equiv 3\ (5).$$

El lector notará que en el último paso usamos el teorema de Fermat. Concluimos entonces que $7^7 \equiv 3\ (20)$, pues el teorema chino del resto asegura

que el sistema de ecuaciones $X \equiv 3\ (4)$ y $X \equiv 3\ (5)$ tiene solución única módulo 20. Luego, escribiendo $7^7 = 20q + 3$, resulta:

$$a = 7^{\,20q+3} = \left(7^{20}\right)^q 7^3 \equiv 7^3 = 7^2 7 \equiv (-1)7 = -7 \equiv 18\ (25).$$

Siendo $a \equiv 18\ (25)$, las posibilidades para las últimas dos cifras de a son 18, 43, 68 y 93, siendo el segundo de estos números el único congruente con 3 módulo 4. Por lo tanto 7^{7^7} termina en 43. $\Diamond$

6.3.2. Ejercicios

1. Probar que $\dfrac{3^{71} - 5}{2}$ es un número entero impar y compuesto.

2. *a*) Caracterizar los $a \in \mathbb{Z}$ tales que $(7a^{111} - 9a^{63} + 8 : 78a) = 26$.
   b) Sea $b \in \mathbb{Z}$ tal que $(5b^{31} + 21 : 132) = 33$. Calcular $r_{66}(b)$.

3. Calcular el resto de dividir $\sum_{k=0}^{100} 3^{k!}$ por 116.

4. Calcular los siguientes órdenes modulares (p denota un primo):
   a) $\mathrm{ord}_{17}(2)$
   b) $\mathrm{ord}_{13}(3)$
   c) $\mathrm{ord}_{23}(-18)$
   d) $\mathrm{ord}_p(p-1)$.

5. Demostrar que toda clase no nula mod 17 es congruente con una potencia de 3. Sabiendo que $14 \equiv 3^9\ (17)$, determinar el inverso de 14 módulo 17.

6. Analizar la primalidad de M_{13}, M_{17} y M_{23}.

7. Probar que el último dígito de un número par perfecto es 6 u 8.

8. Si a es un número entero par, probar que $a^8 + 1$ es divisible por algún primo de la forma $16k + 1$. Concluir que existen infinitos primos de este tipo.

9. Un número compuesto n se dice *pseudoprimo* o *número de Carmichael* si $a^{n-1} \equiv 1\ (n)$ para todo a coprimo con n. Probar que 361 y 1105 son pseudoprimos.

10. Calcular $\varphi(900)$, $\varphi(641)$, $\varphi(1365)$ y $\varphi(10\,!)$.

11. En cada uno de los siguientes casos, hallar los $n \in \mathbb{N}$ que satisfacen la relación planteada:

    a) $\varphi(n) = 14$
    b) $\varphi(n) = 16$
    c) $\varphi(n) = n/2$
    d) $\varphi(n) = n/3$
    e) $\varphi(n) = n - 1$
    f) $\varphi(n) = n - 2\,.$

12. Calcular $r_{7^3}(2^{3538})$ y $r_{119}(13^{73})\,.$

13. Calcular los siguientes órdenes modulares:

    a) $\mathrm{ord}_{63}(8)$
    b) $\mathrm{ord}_{63}(10)$
    c) $\mathrm{ord}_{128}(5)$
    d) $\mathrm{ord}_{128}(-3)\,.$

14. Determinar el primer múltiplo de 49 en la sucesión $1, 11, 111, 1111, \ldots$

Capítulo 7

Complementos de Aritmética Modular

7.1. Raíces primitivas

7.1.1. Ecuaciones no lineales

Una función $f : \mathbb{Z} \to \mathbb{Z}$ de la forma

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \,,$$

donde los a_i son números enteros ($a_n \neq 0$) se llama una *función polinómica* de grado n con coeficientes enteros.

Las funciones polinómicas pueden sumarse y multiplicarse para obtener otras funciones polinómicas, haciendo actuar a la variable x como si fuera un número y aplicando las propiedades estructurales de las operaciones de números reales. Supondremos al lector perfectamente familiarizado con las reglas para hacerlo, por lo que no abundaremos en detalles.

> Conservando la notación de arriba, si $m \in \mathbb{N}$ y $a_n \not\equiv 0 \ (m)$, la ecuación de congruencia
>
> $$f(x) \equiv 0 \ (m)$$
>
> se dice una ecuación de congruencia de grado n.

En el capítulo 6 vimos cómo el algoritmo de Euclides permite resolver con facilidad el caso lineal, esto es, la ecuación de congruencia $aX \equiv b \ (m)$, correspondiente a la función polinómica $f(x) = ax - b$. Por cierto que el método no se extiende a la resolución de ecuaciones de mayor grado, en los que la cuestión es harto más compleja y requiere de otros desarrollos. Ilustraremos la situación a través de un ejemplo, que si bien es particular y de ninguna manera establece un método general, nos ilustrará acerca de la forma de utilizar nuestro conocimiento previo de la estructura multiplicativa de las clases de restos.

Ejemplo 7.1.1 Hallemos las soluciones mod 39 de la ecuación $x^4 \equiv 1$. Se trata de una ecuación de grado 4, y como siempre que resolvemos una ecuación de congruencia pensaremos en términos de clases. Advertimos enseguida que la ecuación es resoluble, ya que obviamente (1) y (38) = (−1) son soluciones, y también que cualquier solución será inversible módulo 39. Veamos cómo hallar todas.

Si $a \in \mathbb{Z}$ representa una solución, resulta que $a^4 \equiv 1\ (3)$ y $a^4 \equiv 1\ (13)$, pues $39 = 3 \,.\, 13$, de donde sigue en particular que a no es divisible ni por 3 ni por 13. Recíprocamente, supongamos que u y v satisfacen las relaciones $u^4 \equiv 1\ (3)$ y $v^4 \equiv 1\ (13)$. Por el teorema chino del resto, sabemos que existe un único entero w módulo 39 tal que $w \equiv u\ (3)$ y $w \equiv v\ (13)$. Por lo tanto, $w^4 \equiv u^4 \equiv 1\ (3)$ y $w^4 \equiv v^4 \equiv 1\ (13)$, de donde resulta que $w^4 \equiv 1\ (39)$ (nuevamente por el teorema chino del resto). En consecuencia, bastará resolver las ecuaciones $x^4 \equiv 1\ (3)$ y $x^4 \equiv 1\ (13)$. La ventaja obtenida reside en que es más sencillo trabajar con módulos primos.

Con respecto a la congruencia módulo 3, toda clase no nula satisface la ecuación, ya que por el teorema de Fermat $a^4 = (a^2)^2 \equiv 1$ para todo $a \in \mathbb{Z}$ no divisible por 3. En el caso de 13, recordemos que todo elemento de $\mathbb{Z}_{13}^*$ es una potencia de 2, pues $\mathrm{ord}_{13}(2) = 12$. Luego, si a es una solución existe un entero s $(0 \leq s < 12)$ tal que $a \equiv 2^i$. Tenemos entonces:

$$a^4 \equiv 1 \Leftrightarrow 2^{4s} \equiv 1 \Leftrightarrow 12 \mid 4s \Leftrightarrow 3 \mid s\,.$$

Por lo tanto, los posibles valores de s son 0, 3, 6 ó 9, siendo entonces a congruente con 1, 8, 12 ó 5 módulo 13, como se verifica inmediatamente. Por el teorema chino del resto, cada uno de estos valores genera dos soluciones módulo 39 (las que corresponden a las dos clases no nulas módulo 3), resultando en definitiva que el problema tiene 8 soluciones en $\mathbb{Z}_{39}^*$, a saber, las clases $(1), (5), (8), (14), (25), (31), (34)$ y (38). $\quad \Diamond$

Repasando los pasos del problema anterior, es evidente que el hecho clave que permitió resolverlo con relativa sencillez es la existencia en $\mathbb{Z}_{13}^*$ de un elemento de orden 12, es decir, un elemento de orden máximo. Nos proponemos mostrar al lector que el del ejemplo 7.1.1 no es un caso especial, y que tal hecho se verifica cualquiera sea el primo p. Dada su importancia, comenzaremos por dar un nombre a dicha situación.

> Si p es un número primo y g es un entero no divisible por p, diremos que g es una *raíz primitiva* módulo p si y sólo si $\mathrm{ord}_p(g) = p - 1$.

Por ejemplo, 2 y 6 son raíces primitivas módulo 13, mientras que 3 no lo es, ya que $\mathrm{ord}_{13}(3) = 3$. En cambio, 3 es una raíz primitiva módulo 17. Le dejamos a su cargo la verificación de estas afirmaciones. Claramente, la

definición anterior sólo depende de la clase de g módulo p, y es equivalente a afirmar que todo entero no divisible por p es congruente módulo p con una potencia de g. Debido a este hecho, también diremos que la clase de g, ó simplemente g, es un *generador* de $\mathbb{Z}_p^*$.

Con el objetivo de probar que para todo primo p existen raíces primitivas módulo p, estableceremos en primer término un importante resultado referente al número de soluciones en $\mathbb{Z}_p$ de una ecuación algebraica de congruencia, conocido como *teorema de Lagrange*. Antes, demostraremos un hecho bien conocido por aquellos lectores familiarizados con el álgebra de polinomios:

Lema 7.1.2 Sea

$$g(x) = a_r x^r + a_{r-1}x^{r-1} + \cdots + a_0$$

una función polinómica de grado $r > 0$ y sea $u \in \mathbb{Z}$. Existen entonces una función polinómica h de grado $r-1$ y un entero a tal que

$$g(x) = (x-u)h(x) + a\,.$$

DEMOSTRACION. Por inducción en r. Si $r = 1$, tenemos

$$g(x) = a_1 x + a_0 = a_1(x-u) + a_1 u + a_0,$$

y el resultado sigue tomando $a = a_1 u + a_0$ y $h(x) = a_1$, que es una función polinómica constante (de grado 0).

Supongamos ahora que r es mayor que 1 y que el lema es válido para funciones polinómicas de grado menor que r. Operando convenientemente, podemos escribir

$$g(x) = x\left(a_r x^{r-1} + a_{r-1}x^{r-2} + \cdots + a_1\right) + a_0\,,$$

y puesto que la expresión entre paréntesis es una función polinómica $t(x)$ de grado $r-1$, ya que $a_r \neq 0$, resulta por hipótesis inductiva que existen una función polinómica $l(x)$ de grado $r-2$ y un entero b tal que

$$t(x) = (x-u)l(x) + b\,.$$

Reemplazando, obtenemos:

$$\begin{aligned} g(x) &= xt(x) + a_0 = x\left((x-u)l(x) + b\right) + a_0 = (x-u)xl(x) + bx + a_0 = \\ &= (x-u)xl(x) + b(x-u) + bu + a_0 = \\ &= (x-u)\left(xl(x) + b\right) + bu + a_0\,. \end{aligned}$$

Arribamos así al resultado que buscábamos, tomando $h(x) = xl(x) + b$ y $a = bu + a_0$, ya que es inmediato verificar que $h(x)$ es una función polinómica de grado $r-1$, por ser $l(x)$ de grado $r-2$.

Es interesante observar que la constante a del enunciado está completamente determinada por g y u, ya que evaluando en $x = u$ la igualdad

$$g(x) = (x - u)h(x) + a \tag{7.1}$$

resulta $a = g(u)$. Deducimos también de (7.1) que el coeficiente de mayor grado de $h(x)$ coincide con el coeficiente de mayor grado de $f(x)$. ♢

Teorema 7.1.3 (Lagrange) Si p es primo, toda ecuación de congruencia de grado n módulo p tiene a lo sumo n soluciones distintas en $\mathbb{Z}_p$.

Demostracion. Aclarando que las congruencias son módulo p, procederemos por inducción en n. Como ya señalamos, el caso $n = 1$ es sencillo y fue tratado en el capítulo 6, donde probamos que la ecuación lineal $ax - b \equiv 0$ tiene exactamente una solución módulo p si a es coprimo con p.

Supongamos ahora $n > 1$ y sea

$$f(x) = \sum_{i=0}^{n} a_i x^i$$

una función polinómica tal que $p \nmid a_n$. Si u es una solución de $f(x) \equiv 0$ (de no haber ninguna no hay nada que probar), aplicando el lema 7.1.2 obtenemos una relación del tipo

$$f(x) = (x - u)h(x) + f(u),$$

donde $h(x)$ es una función polinómica de grado $n - 1$.

Consideremos ahora cualquier otra solución z no congruente con u módulo p y evaluemos ambos miembros de la igualdad anterior en z. Resulta entonces que

$$0 \equiv f(z) = (z - u)h(z) + f(u) \equiv (z - u)h(z),$$

esto es, $p \mid (z-u)h(z)$. Puesto que $p \nmid z-u$, concluimos que z es una solución en $\mathbb{Z}_p$ de la ecuación $h(x) \equiv 0$. Hemos probado así que toda solución distinta de u de la ecuación $f(x) \equiv 0$ es solución de la ecuación $h(x) \equiv 0$. Por un argumento inductivo, esta última admite a lo sumo $n-1$ soluciones distintas en $\mathbb{Z}_p$, de donde concluimos que $f(x) \equiv 0$ admite a lo sumo n soluciones, como queríamos demostrar. ♢

Ejemplo 7.1.4 El teorema de Lagrange brinda una cota para la cantidad de soluciones de una ecuación de congruencia pero no precisa el número exacto de las mismas, que depende del caso particular tratado. De tal modo, la ecuación $4x^5 + x - 5 \equiv 0$ admite dos soluciones en $\mathbb{Z}_7$, mientras que la ecuación $x^5 + 4 \equiv 0$ no admite ninguna en $\mathbb{Z}_{11}$, afirmaciones que el lector podrá comprobar examinando exhaustivamente todos los casos. En el otro extremo, resulta por el pequeño teorema de Fermat que la ecuación $x^p - x \equiv 0$ tiene exactamente p soluciones (el máximo posible) cualquiera sea p. ♢

Continuando con nuestra tarea, recordemos que el orden de cualquier elemento de $\mathbb{Z}_p^*$ es un divisor de $p-1$. Recíprocamente, dado un divisor d de $p-1$, veamos qué podemos decir acerca de la cantidad de elementos de orden d en $\mathbb{Z}_p^*$.

Lema 7.1.5 Sea p un primo y sea d un divisor de $p-1$. Si $\mathbb{Z}_p^*$ admite algún elemento de orden d entonces admite exactamente $\varphi(d)$ elementos de orden d.

DEMOSTRACION. Supongamos que a representa a un elemento de orden d en $\mathbb{Z}_p^*$. Sabemos entonces que a satisface la relación $a^d \equiv 1$ y que los elementos a^i $(0 \leq i < d)$ representan d elementos distintos en $\mathbb{Z}_p^*$. Si j es cualquiera de los exponentes anteriores, resulta que $(a^j)^d = (a^d)^j \equiv 1$, esto es, todas estas d potencias de a satisfacen la ecuación $x^d - 1 = 0$ en $\mathbb{Z}_p^*$. Puesto que por el teorema de Lagrange ésta admite a lo sumo d soluciones distintas, concluimos que los a^i son exactamente sus soluciones.

Sigue luego que todo elemento b de orden d es de la forma a^j $(0 \leq j < d)$, puesto que en particular b satisface la relación $b^d \equiv 1$. Nos resta entonces caracterizar los valores de j en dicho rango para los cuales $\mathrm{ord}_p(a^j) = d$. Probaremos que esto se verifica si y sólo si $j \perp d$.

Supongamos en primer término que $\mathrm{ord}_p(a^j) = d$ y sea $t = (j : d)$. Entonces:

$$(a^j)^{d/t} = (a^d)^{j/t} \equiv 1.$$

Por definición de orden, resulta que $d/t \geq d$, relación que claramente sólo puede verificarse si $t = 1$. Por lo tanto, es necesario que j y d sean coprimos. Recíprocamente, supongamos que $(j : d) = 1$. Para demostrar que $\mathrm{ord}_p(a^j) = d$, deberemos probar que $d \mid m$ para todo $m \in \mathbb{N}$ tal que $(a^j)^m \equiv 1$. En tal caso, tenemos

$$a^{jm} = (a^j)^m \equiv 1,$$

de donde sigue que $d \mid jm$, teniendo en cuenta que $\mathrm{ord}_p(a) = d$. Puesto que d y j son coprimos, concluimos que $d \mid m$, como queríamos probar.

En definitiva, el número de elementos de orden d en $\mathbb{Z}_p^*$ coincide con el número de exponentes j coprimos con d del intervalo $[\,0, d\,)$, que por definición es $\varphi(d)$.

Por ejemplo, no es difícil verificar que 5 (la clase de 5) es un elemento de orden 9 módulo 19. Puesto que $\varphi(9) = 6$, sigue que los elementos de orden 9 de $\mathbb{Z}_{19}^*$ son las 6 potencias de 5 con exponentes coprimos con 9, a saber: 5, $6 \equiv 5^2$, $17 \equiv 5^4$, $9 \equiv 5^5$, $16 \equiv 5^7$ y $4 \equiv 5^8$. $\quad \Diamond$

Dado un divisor d de $p-1$, notemos que el lema 7.1.5 asegura que el número de elementos de orden d en $\mathbb{Z}_p^*$ (que designaremos por $\vartheta(d)$) es 0 ó $\varphi(d)$. El siguiente resultado −interesante en sí mismo−, nos permitirá determinar sin ambigüedad el valor de $\vartheta(d)$ cualquiera sea d.

Proposición 7.1.6 Sea $n \in \mathbb{N}$. Entonces

$$n = \sum_{d|n} \varphi(d),$$

donde el símbolo indica que d recorre los divisores positivos de n.

DEMOSTRACION. Por ejemplo,

$$10 = 1 + 1 + 4 + 4 = \varphi(1) + \varphi(2) + \varphi(5) + \varphi(10)\,.$$

Para demostrarlo en general, clasifiquemos los elementos a del intervalo $\mathbb{I}_n$ según los valores de $(a : n)$, que en cualquier caso será un divisor de n. Precisamente, para cada divisor d de n sea

$$A_d = \{a \in I_n : (a : n) = d\}\,.$$

Puesto que claramente

$$\mathbb{I}_n = \bigcup_{d|n} A_d\,,$$

y la unión es disjunta, tomando cardinales obtenemos la relación

$$n = \sum_{d|n} \#(A_d)\,.$$

Veamos ahora cómo calcular estas cantidades.

Si $d \mid n$ y $a \in A_d$, observemos que a/d y n/d son coprimos y $a/d \leq n/d$. Por otro lado, supongamos que $b \in \mathbb{I}_{n/d}$ es coprimo con n/d. Entonces

$$(db : n) = (db : d\,n/d) = d(b : n/d) = d,$$

y por lo tanto $db \in A_d$. Puesto que la conjunción de estos dos hechos implica que la aplicación $x \mapsto x/d$ establece una correspondencia biunívoca entre A_d y los elementos del intervalo $\mathbb{I}_{n/d}$ coprimos con n/d, deducimos que $\#(A_d) = \varphi(n/d)$. Finalmente, teniendo en cuenta que cuando d recorre todos los divisores de n también lo hace n/d, obtenemos:

$$n = \sum_d \#(A_d) = \sum_d \varphi(n/d) = \sum_d \varphi(d),$$

como queríamos probar. $\Diamond$

Ahora ya estamos en condiciones de demostrar el resultado central de esta sección:

Teorema 7.1.7 Existen raíces primitivas módulo p cualquiera sea p.

DEMOSTRACION. Tomando $n = p-1$ en la proposición 7.1.6, y de acuerdo con la observación derivada del lema 7.1.5, obtenemos

$$p-1 = \sum_{d|p-1} \vartheta(d) \leq \sum_{d|p-1} \varphi(d) = p-1\,.$$

Por consiguiente, las dos sumatorias anteriores son iguales. Más aún, siendo cada término de la primera menor ó igual que el correspondiente término de la segunda, concluimos que todos los términos correspondientes son iguales, ya que la desigualdad de la fórmula anterior sería estricta si para algún d fuera $\vartheta(d) < \varphi(d)$.

Hemos probado entonces que para cada divisor d de $p-1$ existen en $\mathbb{Z}_p^*$ *exactamente* $\varphi(d)$ elementos de orden d. Tomando en particular $d = p-1$, resulta finalmente que existen raíces primitivas módulo p. Más precisamente, existen $\varphi(p-1)$ elementos de orden máximo (generadores) en $\mathbb{Z}_p^*$. $\quad\Diamond$

Ejemplo 7.1.8 Tomando $p = 17$ y operando en $\mathbb{Z}_{17}$ tenemos que

$$3^8 = (3^4)^2 = 81^2 \equiv (-4)^2 = 16 \equiv -1\,,$$

de donde resulta que $3^d \not\equiv 1$ para todo divisor propio d de 16, ya que todos ellos dividen a 8. Por consiguiente, 3 es una raíz primitiva módulo 17.

De acuerdo con lo señalado más arriba, sigue que 17 admite $\varphi(16) = 8$ raíces primitivas, correspondientes a las potencias de 3 con exponente impar (coprimo con 16). Precisamente, ellas son $3 \equiv 3^1$, $10 \equiv 3^3$, $5 \equiv 3^5$, $11 \equiv 3^7$, $14 \equiv 3^9$, $7 \equiv 3^{11}$, $12 \equiv 3^{13}$ y $6 \equiv 3^{11}$. $\quad\Diamond$

El teorema 7.1.7 asegura la existencia de raíces primitivas, pero su demostración no provee ningún método para encontrar una. Si p es relativamente pequeño puede procederse por ensayo y error hasta dar con alguna. Por ejemplo, 2 y 3 no son raíces primitivas módulo 23 ($2^{11} \equiv 3^{11} \equiv 1$), pero 5 sí lo es, como puede verificarse efectuando los cálculos. Habiendo obtenido una podemos fácilmente encontrar todas, ya que ellas son de la forma 5^j, con j variando entre 0 y 21 y coprimo con 22. Siendo $\varphi(22) = 10$, resulta que hay 10 elementos de orden máximo en $\mathbb{Z}_{23}^*$, a saber, las clases de 5, 10, 20, 17, 11, 21, 19, 15, 7 y 14, que se obtienen elevando sucesivamente 5 a los exponentes 1, 3, 5, 7, 9, 13, 15, 17, 19 y 21. Por supuesto que los cálculos se efectúan reduciendo siempre módulo 23, y no requieren en este caso mayor esfuerzo, habida cuenta de que $5^2 \equiv 2$.

Si p es grande la situación es bien distinta, y no se conocen métodos generales efectivos para hallar raíces primitivas módulo p. Existen sin embargo ciertos criterios que en algunos casos permiten aliviar el problema, como el siguiente:

Proposición 7.1.9 Sea p un número primo, sean $q_1, q_2, \ldots, q_m$ los divisores primos de $p-1$ y supongamos que g es un entero no divisible por p tal que $g^{(p-1)/q_i} \not\equiv 1$ para todo i. Entonces g es una raíz primitiva módulo p.

DEMOSTRACION. Sea $d = \mathrm{ord}_p(g)$. Si $d < p-1$, el cociente h de la división entera de $p-1$ por d es un divisor de $p-1$ distinto de 1, por lo que admite algún divisor primo. Puesto que éste también divide a $p-1$, resulta que existe un índice j tal que $q_j \mid h$, digamos $h = q_j\, k$. Luego

$$g^{(p-1)/q_j} = g^{dk} = (g^d)^k \equiv 1,$$

lo que contradice la hipótesis. Luego $d = p-1$ y g es una raíz primitiva módulo p. $\Diamond$

Ejemplo 7.1.10 Apliquemos el criterio para demostrar que 3 es una raíz primitiva módulo 79. En este caso, $79-1 = 78 = 2\,.\,3\,.\,13$, por lo que bastará verificar que ninguno de los números 3^6, 3^{26} y 3^{39} es congruente con 1 módulo 79. Teniendo en cuenta que $3^4 \equiv 2 \ (79)$, obtenemos sucesivamente (las congruencias son módulo 79):

i) $3^6 = 3^4\, 9 \equiv 18$

ii) $3^{26} = (3^{13})^2 = \left((3^4)^3 3\right)^2 \equiv (8\,.\,3)^2 = 24^2 \equiv 23$

iii) $3^{39} = 3^{13} 3^{26} \equiv 24\,.\,23 = 552 \equiv 78\,,$

lo que prueba nuestra afirmación. $\Diamond$

Si bien el criterio anterior reduce sensiblemente la cantidad de pruebas a realizar (recordemos que g es primitiva si y sólo si $g^d \not\equiv 1$ para todo divisor propio d de $p-1$), su aplicación presupone conocer la factorización de $p-1$, que no es sencilla de obtener si p es grande, aún considerando que $p-1$ es par. Debemos tener en claro entonces que se trata de un criterio que no es aplicable a cualquier situación.

Observemos también que hemos ilustrado la situación con un ejemplo sencillo (79 no responde a la denominación de primo "grande"), y hemos realizado las operaciones empleando ciertas destrezas de cálculo. Para trabajar con primos de considerable magnitud es necesario desarrollar técnicas sistemáticas de cálculo. Ahorrando algunos detalles, exhibiremos ahora un algoritmo que permite calcular en general potencias módulo n elevando sucesivamente al cuadrado y reduciendo módulo n, y que por lo tanto sólo trabaja con números cuyo tamaño no supera a n^2.

Algoritmo de cálculo de potencias modulares.

Supongamos que m y n son números naturales y que queremos calcular b^m módulo n, donde b es un entero que podemos suponer en el rango $0 < b < n$. La idea es ir elevando sucesivamente al cuadrado, efectuando en cada paso las correspondientes reducciones. El algoritmo comporta una serie de etapas, cada una de las cuales presenta una disyuntiva. Le mostraremos cómo empieza y cuál es en cada caso la disyuntiva que se presenta, señalando

entonces cuál es el próximo paso a realizar. Tenemos de entrada dos posibilidades, que luego se repetirán en cada paso (todas las operaciones indicadas se realizan módulo n):

1) m es par. Podemos escribir entonces $b^m = (b^2)^{m/2}$. Para continuar, reemplazamos b por b^2 módulo n, el exponente m por $m/2$, y procedemos de igual forma, ahora con respecto al nuevo exponente.

2) m es impar. En tal caso escribimos $b^m = b(b^2)^{(m-1)/2}$. Nuevamente reemplazamos b por b^2 y m por $(m-1)/2$, sólo que ahora aparece un factor b que debemos acumular. Continuamos luego como en el caso 1).

Estas son entonces las sucesivas etapas del algoritmo, que finaliza cuando el exponente alcanza el valor 1. Puesto que en cada paso dividimos por 2 el exponente (más precisamente, tomamos la parte entera de su mitad), el número de etapas coincide con el número de cifras del desarrollo binario de m. En cada una de ellas deberemos elevar un número menor que n al cuadrado, por lo que los números involucrados no superarán a n^2. A su vez, en algunas de las etapas deberemos multiplicar por el valor acumulado (en aquellas en las que el exponente es impar), coincidiendo el número de las mismas con la cantidad de bits no nulos del desarrollo binario de m. Puesto que es relativamente fácil demostrar que el valor esperado de esta cantidad es aproximadamente la mitad del número de cifras binarias, concluimos que el número de multiplicaciones que requerirá el algoritmo será del orden de $1{,}5 \,.\, \log_2 m$, lo que lo hace bastante eficiente.

Para ejecutar el algoritmo en un caso concreto conviene emplear tres variables a, k y c, cuyos roles son los siguientes:

i) La variable a toma el valor inicial b, y se reemplaza en cada paso por a^2 mod n.

ii) La variable k indica el exponente, cuyo valor inicial es m. En cada etapa se reemplaza por la parte entera de su mitad.

iii) La variable c indica el valor acumulado, cuyo valor inicial es 1. En cada etapa, no cambia si el exponente k es par y se la multiplica por a si k es impar (esta operación antecede a calcular a^2 mod n). El valor final de c es el resultado buscado.

Ejemplo 7.1.11 Calculemos 37^{76} mod 113, para lo cual conviene disponer en una tabla como la siguiente los sucesivos valores que van tomando las variables (le encargamos verificar todos los resultados parciales):

Etapa	a	k	c
1	37	76	1
2	13	38	1
3	56	19	56
4	85	9	14
5	106	4	14
6	49	2	14
7	28	1	**53**

Hemos obtenido entonces que $37^{76} \equiv 53 \ (113)$. Obsérvese que c fue multiplicado por a en las etapas 3, 4 y 7, que son las que corresponden a valores impares del exponente.

7.1.2. Ejercicios

1. Hallar las raíces primitivas módulo 17, 31 y 43.

2. Calcular $r_{211}(2^{100})$, $r_{6887}(5^{340})$ y $r_{257}(-3^{121})$.

3. Probar que 2 es una raíz primitiva módulo 13, 5 es una raíz primitiva módulo 23 y 3 es una raíz primitiva módulo 79.

4. En cada uno de los siguientes casos, determinar el menor $k \in \mathbb{N}$ para el cual se satisface la relación planteada:

   a) $2^k \equiv 9 \ (130)$
   b) $5^k \equiv 4 \ (23)$
   c) $3^k \equiv 6 \ (79)$.

5. Resolver las siguientes ecuaciones de congruencia:

   a) $-2x^4 + 5 \equiv 0 \ (13)$
   b) $x^6 \equiv 4 \ (23)$
   c) $5x^{10} \equiv -49 \ (79)$.

6. Resolver las siguientes ecuaciones de congruencia:

   a) $x^5 + x^4 + x^3 + x^2 + x + 1 \equiv 0 \ (31)$
   b) $x^8 \equiv 1 \ (60)$
   c) $x^{10} \equiv 1 \ (29)$.

7. Sea p un primo y sea $m \in \mathbb{N}$ no divisible por $p-1$. Si $x_1, x_2, \ldots, x_p$ son números enteros consecutivos, probar que la suma

$$x_1^m + x_2^m + \ldots + x_p^m$$

es divisible por p.

7.2. Residuos cuadráticos

7.2.1. Definición y propiedades básicas

En lo que sigue la letra p designará un número primo, y salvo mención expresa, el símbolo $\equiv$ denotará congruencia módulo p.

No es difícil demostrar que la conocida fórmula que permite hallar las raíces de una ecuación cuadrática con coeficientes reales puede aplicarse también a la resolución de ecuaciones cuadráticas módulo p. Naturalmente, la condición de resolubilidad es que su discriminante sea un cuadrado en la estructura multiplicativa de $\mathbb{Z}_p$. Motivados por este hecho, introducimos la siguiente definición:

> Si a es un entero, decimos que a es un *residuo cuadrático* módulo p (rc mod p) si y sólo si existe $x \in \mathbb{Z}$ tal que $x^2 \equiv a$. En tal caso diremos informalmente que a es un cuadrado módulo p.

Para ir enfocando la cuestión señalemos algunos hechos, todos ellos de demostración inmediata a partir de la definición.

1) Todo cuadrado perfecto es un rc mod p, aunque no vale la recíproca. Por ejemplo, 2 es un rc mod 23, ya que $2 \equiv 5^2\ (23)$.

2) Es claro que en la definición de residuo cuadrático sólo importa la clase de a módulo p, por lo que podemos limitarnos a considerar valores de a y x pertenecientes al intervalo $[\,0, p-1\,]$. Sigue en particular que todo entero es un residuo cuadrático módulo 2, pues 0 y 1 son cuadrados perfectos. Debido a ello, supondremos de aquí en más que p es impar.

3) Puesto que 0 es un rc mod p, centraremos nuestra atención en el intervalo $\mathbb{I}_{p-1}$, que dicho sea de paso tiene un número par de elementos.

4) Supongamos que a y u pertenen a $\mathbb{I}_{p-1}$ y que $u^2 \equiv a$. Diremos entonces que u es una *raíz cuadrada* de a módulo p. Si v es otra cualquiera, sigue por transitividad que $v^2 \equiv u^2$, de donde deducimos (ejercicio 12 de la sección 6.1) que $v \equiv u$ ó $v \equiv -u$. Por lo tanto a tiene exactamente dos raíces cuadradas módulo p, a saber, u y $p-u$. Por ejemplo, las raíces cuadradas de 2 módulo 23 son 5 y 18. $\Diamond$

Más adelante estableceremos una condición necesaria y suficiente para que 2 sea un residuo cuadrático módulo p. A manera de ilustración, resolvamos ahora esta cuestión en un caso muy especial.

Ejemplo 7.2.1 Sea p un primo de Mersenne mayor que 3. Entonces 2 es un residuo cuadrático módulo p.

Sea $p = 2^q - 1$, donde q es un primo impar, lo que en términos de congruencia significa que $2^q \equiv 1$. Simplemente multiplicando por 2 esta relación, obtenemos:

$$2 \equiv 2^{q+1} = \left(2^{(q+1)/2}\right)^2,$$

esto es, 2 es un rc mod p. Notemos además que hemos determinado una raíz cuadrada de 2 módulo p, a saber, la clase de $2^{(q+1)/2}$. Por ejemplo, 8 es una raíz cuadrada de 2 módulo 31. $\Diamond$

Veamos cuál es el número de residuos cuadráticos módulo p.

Proposición 7.2.2 Exactamente la mitad de los elementos de $\mathbb{I}_{p-1}$ son residuos cuadráticos módulo p. Vale decir, existen $\dfrac{p-1}{2}$ residuos cuadráticos no nulos módulo p.

Brindaremos dos pruebas de este hecho. La primera es una consecuencia casi inmediata de una observación anterior, mientras que la segunda −si bien menos elemental− permite caracterizar completamente los residuos cuadráticos a partir de una raíz primitiva módulo p (de aquí en adelante emplearemos la notación $(p-1)/2 = h$).

PRIMERA DEMOSTRACION. Sigue de la observación 4) que en el caso de que a sea un rc mod p una de sus raíces cuadradas es menor ó igual que h (está en la primera mitad del intervalo), mientras que la otra es mayor, esto es, pertenece a la segunda mitad. Por lo tanto, todos los rc mod p se obtienen elevando al cuadrado los elementos del intervalo $\mathbb{I}_h$. Además, estos h cuadrados son todos distintos módulo p, ya que si $x \in \mathbb{I}_h$, la otra raíz cuadrada mod p de x^2 es $p - x$, que no perteneces a $\mathbb{I}_h$. Luego existen exactamente h residuos cuadráticos no nulos módulo p.

SEGUNDA DEMOSTRACION. Sea g una raíz primitiva módulo p. En ese caso, para todo $a \in \mathbb{I}_{p-1}$ existe un único k $(0 \leq k \leq p-2)$ tal que $a \equiv g^k$. Afirmamos que a es un rc mod p si y sólo si k es par, lo que probaría el resultado, ya que exactamente la mitad de dichos exponentes son pares, por ser $p-2$ impar.

Supongamos en primer término que k es par, digamos $k = 2j$. Luego, $a \equiv g^k = (g^j)^2$ y por lo tanto a es un rc mod p.

Recíprocamente, supongamos que a es un rc mod p, digamos $a \equiv x^2$, y sea i tal que $x \equiv g^i$, de donde resulta que $g^k \equiv a \equiv (g^i)^2 = g^{2i}$. Puesto que $\mathrm{ord}_p(g) = p-1$, sigue que $k \equiv 2i \ (p-1)$, vale decir, $p-1 \mid k-2i$. Siendo $p-1$ par, concluimos inmediatamente que k es par, como queríamos demostrar. $\Diamond$

Ejemplo 7.2.3 Determinemos los residuos cuadráticos módulo 13. Teniendo en cuenta que $h = 6$ en este caso, resulta que ellos son $1^2 \equiv 1$, $2^2 \equiv 4$, $3^2 \equiv 9$, $4^2 \equiv 3$, $5^2 \equiv 12$ y $6^2 \equiv 10$. O sea, $\{1, 3, 4, 9, 10, 12\}$ es el conjunto de residuos cuadráticos no nulos módulo 13. Otra forma de calcularlos es tomar las sucesivas potencias de 2 (que es una raíz primitiva módulo 13) con exponente par variando entre 0 y 10. Las mismas serán entonces de la forma $2^{2i} = 4^i$, donde $0 \leq i \leq 5$, esto es, 4 *genera* los restos cuadráticos módulo 13. De esta manera, y a partir de 1, cada residuo cuadrático se obtiene del

anterior multiplicando por 4 y reduciendo módulo 13. Con este orden de recorrido, obtenemos sucesivamente: 1, 4, 3, 12, 9 y 10. $\Diamond$

Como dijimos al principio, la teoría de residuos cuadráticos está naturalmente ligada a la resolución de ecuaciones cuadráticas de congruencia. Veamos un ejemplo:

Ejemplo 7.2.4 La ecuación $x^2 + 41x + 28 \equiv 0$ tiene 4 soluciones en $\mathbb{Z}_{91}$.

Designaremos por u cualquier solución genérica y procederemos en forma similar a la que se utiliza para resolver una ecuación cuadrática con coeficientes reales. En primer lugar, tratándose de una ecuación de congruencia podemos reemplazar cualquier coeficiente por otro que sea congruente con él módulo 91. Por ejemplo, para completar el cuadrado, sustituimos 41 por -50, resultando que (el lector verificará los cálculos módulo 91)

$$0 \equiv u^2 - 50u + 28 = (u-25)^2 + 28 - 25^2 \equiv (u-25)^2 - 51,$$

ó equivalentemente, $(u-25)^2 \equiv 51$.

Notemos por otro lado que $91 = 7 \,.\, 13$. Puesto que en general $a \equiv b\ (91)$ si y solo si $a \equiv b\ (7)$ y $a \equiv b\ (13)$, sigue que u debe satisfacer las congruencias $(u-4)^2 \equiv 2\ (7)$ y $(u-12)^2 \equiv 12\ (13)$, donde ya hemos reducido módulo 7 y 13, respectivamente. Cada una de estas dos ecuaciones es resoluble en su correspondiente campo, dado que 2 es un rc mod 7 (3 y 4 son sus raíces cuadradas en $\mathbb{Z}_7$) y 12 es un rc mod 13, siendo 5 y 8 sus raíces cuadradas en $\mathbb{Z}_{13}$, como vimos en el ejemplo anterior. Simplemente despejando, vemos que las soluciones módulo 7 son:

$$v_0 \equiv 3 + 4 \equiv 0$$
$$v_1 \equiv 4 + 4 \equiv 1$$

y las soluciones módulo 13:

$$w_0 \equiv 5 + 12 \equiv 4$$
$$w_1 \equiv 8 + 12 \equiv 7\,.$$

Combinando de las 4 maneras posibles estas soluciones obtendremos todas las soluciones módulo 91. En efecto, si (y, z) es cualquiera de los 4 pares, sigue por el teorema chino del resto que existe un único entero x módulo 91 tal que $x \equiv y\ (7)$ y $x \equiv z\ (13)$. Entonces

$$x^2 + 41x + 28 \equiv y^2 + 41y + 28 \equiv 0\ (7)$$

y

$$x^2 + 41x + 28 \equiv z^2 + 41z + 28 \equiv 0\ (13),$$

lo que asegura que x es una solución módulo 91, nuevamente por el teorema chino del resto. Finalmente, calculando en cada caso los valores de x a través del método que brinda la demostración del teorema chino, resulta que las 4 soluciones en $\mathbb{Z}_{91}$ de la ecuación $x^2 + 41x + 28 \equiv 0$ son 7, 43, 56 y 85, como es fácil de verificar. $\Diamond$

Criterio de Euler

No tenemos hasta aquí ninguna forma realmente efectiva de decidir si un entero a no divisible por p es un residuo cuadrático módulo p. Una posibilidad sería confeccionar la lista de los residuos cuadráticos mod p, elevando al cuadrado todos los elementos de la primera mitad del intervalo $\mathbb{I}_{p-1}$ y verificar luego si a está en la lista, pero es obvio que el método es impracticable si p es grande. Alternativamente, podríamos buscar una raíz primitiva g módulo p, expresar a como potencia de g, y verificar finalmente la paridad del exponente. Si bien luce atractiva, esta forma presenta todavía más dificultades que la anterior: en primer lugar ya comentamos que no es fácil hallar una primitiva, y aún en el caso de haber hallado una, no existe en general ningún algoritmo eficiente para calcular el exponente. Sin embargo no debemos desalentarnos, ya que el siguiente resultado −debido a Euler−, permite decidir directamente sobre la cuestión.

Proposición 7.2.5 (Criterio de Euler) Un entero a no divisible por p es un residuo cuadrático módulo p si y sólo si

$$a^h \equiv 1 \ (p\,).$$

Demostracion. Antes de entrar de lleno en la demostración destaquemos un hecho. Dado cualquier entero c no divisible por p, sigue del teorema de Fermat que

$$\left(c^h\right)^2 = c^{2h} = c^{p-1} \equiv 1,$$

de donde resulta que $c^h \equiv 1$ ó $c^h \equiv -1$. Por lo tanto, el criterio de Euler afirma que la mitad de las clases no nulas satisface la primera relación (los residuos cuadráticos), mientras que la otra mitad satisface la segunda (los residuos no cuadráticos). Vamos ahora sí a la demostración.

Supongamos en primer término que a es un rc mod p, y sea b cualquiera de sus raíces cuadradas módulo p. Entonces,

$$a^h \equiv \left(b^2\right)^h = b^{p-1} \equiv 1,$$

lo que prueba la necesidad de la condición.

Habiendo ya probado que los h residuos cuadráticos verifican la ecuación

$$x^h \equiv 1\,,$$

observemos por otro lado que ella admite a lo sumo h soluciones módulo p, por el teorema de Lagrange. Deducimos entonces de ambos hechos que sus soluciones son exactamente los residuos cuadráticos no nulos módulo p, como queríamos demostrar. $\Diamond$

Ejemplo 7.2.6 De acuerdo con el criterio de Euler, los residuos cuadráticos módulo 11 son (además de 1) los elementos de orden 5 de $\mathbb{Z}_{11}^*$. Puesto que 4

es uno de ellos, por ser un cuadrado perfecto, sigue que todos son de la forma 4^i, donde $0 \leq i \leq 4$. Resulta así que los residuos cuadráticos módulo 11 son 1, 4, 5, 9 y 3 (obsérvese que cada uno se obtiene del anterior multiplicando por 4). $\Diamond$

Utilizaremos ahora el criterio de Euler para analizar dos situaciones concretas. Precisamente, nos proponemos caracterizar los números primos p respecto de los cuales -1 y 2 son residuos cuadráticos (más adelante tendremos ocasión de apreciar el interés de estos hechos).

Proposición 7.2.7 Son válidas las siguientes afirmaciones:

1) -1 es un rc mod p si y sólo si $p \equiv 1\ (4)$.

2) 2 es un rc mod p si y sólo si $p \equiv \pm 1\ (8)$.

DEMOSTRACION. Teniendo en cuenta que p es impar, sigue inmediatamente del criterio de Euler que -1 es un rc mod p si y sólo si h es par, condición claramente equivalente a que p sea de la forma $4k+1$, como se postula en 1).

La parte 2) requiere algo más de trabajo. Para demostrarla, designemos por $\mathbb{I}_{p-1}^+$ y $\mathbb{I}_{p-1}^-$ los intervalos $[\,1, h\,]$, y $[\,h+1, p-1\,]$, respectivamente. También, numeremos en la forma $x_1, x_2, \ldots, x_h$ los elementos de $\mathbb{I}_{p-1}^+$.

Si para cada i escribimos $y_i = 2x_i$, es claro que los y_i son h elementos distintos de $\mathbb{I}_{p-1}$. Algunos de ellos pertenecerán a $\mathbb{I}_{p-1}^+$, mientras que los restantes estarán en $\mathbb{I}_{p-1}^-$. Ahora bien, todo elemento y de $\mathbb{I}_{p-1}^-$ es el inverso aditivo módulo p de algún elemento de $\mathbb{I}_{p-1}^+$, ya que $-y \equiv p - y \in \mathbb{I}_{p-1}^+$. Por lo tanto, para cada i se verifica una relación de la forma

$$2x_i \equiv \pm z_i\,, \tag{7.2}$$

donde $z_i \in \mathbb{I}_{p-1}^+$ para todo i (el signo menos aparece cada vez que $y_i > h$). Nuestro próximo paso será mostrar que los z_i recorren (en algún orden) los elementos de $\mathbb{I}_{p-1}^+$. Puesto que son exactamente h, bastará probar *que son todos distintos*, esto es, que $z_i = z_j$ solo si $i = j$.

Si $z_i = z_j$, sigue de las relaciones (7.2) que se verifica alguna de las situaciones $2x_i \equiv 2x_j$, $-2x_i \equiv -2x_j$ ó $2x_i \equiv -2x_j$. Las dos primeras son claramente equivalentes, y puesto que el factor 2 puede cancelarse (es coprimo con p), sólo habrá que analizar los casos $x_i \equiv x_j$ ó $x_i \equiv -x_j$.

En el primero de ellos, siendo x_i y x_j positivos y menores que p, la única posibilidad es que los índices sean iguales. En el otro caso, $x_i + x_j$ resultaría ser un múltiplo positivo de p y menor que p, lo que obviamente no es posible. En consecuencia, el segundo caso no puede ocurrir.

Sabiendo ya que los z_i son todos distintos, multipliquemos ahora miembro a miembro las relaciones (7.2). Obtenemos entonces:

$$2^h \prod_i x_i \equiv (-1)^m \prod_i z_i = (-1)^m \prod_i x_i,$$

donde m es el número de índices i tales que $2x_i \in \mathbb{I}_{p-1}^-$. Observemos que la última igualdad es consecuencia del hecho de que los factores de ambas productorias son los mismos, posiblemente recorridos en diferente orden.

Como el producto de los x_i es coprimo con p (por serlo sus factores), podemos cancelarlo en la relación de congruencia anterior, resultando finalmente que

$$2^h \equiv (-1)^m .$$

Concluimos luego −usando el criterio de Euler−, que 2 es un rc mod p si y sólo si m es par. Sólo resta entonces caracterizar los primos p para los cuales se satisface esta última condición.

Si $x \in \mathbb{I}_{p-1}^+$, es claro que

$$2x \in \mathbb{I}_{p-1}^- \Leftrightarrow h/2 < x \leq h ,$$

por lo que deberemos contar el número de enteros que se hallan en este rango. En general, dados números reales a y b ($a < b$), es un sencillo ejercicio demostrar que existen $[\,b\,] - [\,a\,]$ enteros en el intervalo $(\,a, b\,]$. En nuestra situación, es obvio que los valores de las partes enteras dependen de la clase de congruencia de p módulo 4, por lo que convendrá analizar cada uno de los dos casos posibles:

i) $p = 4k + 1$; resulta entonces que $[h/2] = k$ y $[h] = 2k$, y por lo tanto $m = k$. En consecuencia, 2 es un rc mod p si y sólo si $p \equiv 1\ (8)$.

ii) $p = 4k+3$; En este caso $[h/2] = k$ y $[h] = 2k+1$, de donde obtenemos que $m = k + 1$. Por lo tanto, 2 es un residuo cuadrático si y sólo si k es impar, ó equivalentemente, $p \equiv 7 \equiv -1\ (8)$, lo que prueba nuestra aserción. $\Diamond$

En un lenguaje informal, podríamos decir que tanto -1 como 2 son residuos cuadráticos módulo p respecto del 50 % de los primos impares p, ya que al dividir p por 4 se obtienen dos restos posibles (1 y 3) y al dividirlo por 8 se obtiene alguno de los cuatro restos 1, 3, 5 y 7.

A través de varios ejemplos veremos algunas aplicaciones interesantes de la proposición 7.2.7.

Ejemplo 7.2.8 Existen infinitos primos de la forma $4k + 1$.

Suponiendo que sólo hay un número finito de tales primos y designando por a el producto de todos ellos, consideremos el número

$$c = (2a)^2 + 1,$$

que claramente es impar y no divisible por ningún primo congruente con 1 módulo 4. Tomando cualquier divisor primo q de c (que necesariamente será congruente con 3 módulo 4), resulta que $(2a)^2 + 1 \equiv 0\ (q)$, ó equivalentemente, $(2a)^2 \equiv -1\ (q)$, lo que contradice la primera parte de 7.2.7.

Señalemos que el resultado que acabamos de probar es un caso particular de un célebre teorema de Dirichlet, que afirma que si a y b son coprimos la progresión aritmética

$$a\,,\ a+b\,,\ a+2b\,,\ \ldots\,,\ a+kb\ \ldots$$

contiene infinitos números primos. Claro que las dificultades de su demostración exceden largamente el alcance de estas páginas, mientras que la prueba ofrecida aquí es ciertamente más elemental. $\Diamond$

Ejemplo 7.2.9 Sea p un primo congruente con 3 módulo 4 tal que $q = 2p+1$ también es primo. Entonces $q \mid 2^p - 1$ y por lo tanto M_p es compuesto si $p > 3$.

Casos particulares de la situación del enunciado son $p = 23$ y $p = 83$, que corresponden a $q = 47$ y $q = 167$, respectivamente. Si bien este hecho parece reforzar la conjetura de que existen infinitos números de Mersenne compuestos, aclaremos que permanece abierta la cuestión de la existencia de infinitos pares de primos (p, q) tales que $q = 2p + 1$.

En cuanto a la prueba de nuestra afirmación, sigue fácilmente de las hipótesis que $q \equiv -1\ (8)$, y por lo tanto 2 es un residuo cuadrático módulo q. Entonces

$$2^p = 2^{(q-1)/2} \equiv 1\ (q)\,,$$

esto es, $q \mid 2^p - 1$. Puesto que $q < M_p$ si $p > 3$ (pruebe por inducción que $n + 1 < 2^{n-1}$ si $n > 3$), concluimos que M_p es compuesto.

Volviendo a los ejemplos del principio, señalemos que $M_{23} = 47\,.\,178481$ y que M_{83} es el producto de 167 por un número primo de 23 cifras. $\Diamond$

NUMEROS DE FERMAT Un número natural de la forma $2^{2^m} + 1$ $(m \geq 0)$ se dice un *número de Fermat* y será notado F_m. Al igual que la secuencia de números de Mersenne, esta sucesión atrajo el interés de los estudiosos de la teoría de números desde Fermat hasta nuestros días, pues pareció constituir una posible fuente de obtención de primos (digamos de paso que $2^s + 1$ es primo solo si s es una potencia de 2).

Contribuyó a fomentar tal creencia el hecho de que sus primeros cinco valores son números primos, a saber, 3, 5, 17, 257 y 65537. Curiosamente, ellos son los únicos primos de Fermat conocidos, y actualmente tiende a pensarse que no existen otros. La cuestión presenta grandes dificultades computacionales y no ha sido aún resuelta, habiéndose demostrado hasta aquí que F_m es compuesto para m mayor que 4 y menor que 30. Nosotros examinaremos aquí el caso $m = 5$. Consignemos que Euler fue el primero en demostrar en el siglo XVIII el carácter compuesto de F_5.

Ejemplo 7.2.10 F_5 no es primo.

Antes de pasar al caso particular que nos ocupa, veamos qué podemos decir en general de los posibles divisores primos de F_m $(m \geq 5)$.

Si q es uno de ellos, el hecho de que q divida a $2^{2^m}+1$ implica en términos de congruencias que $2^{2^m} \equiv -1\ (q)$, de donde sigue que $2^{2^{m+1}} \equiv 1\ (q)$ y por lo tanto el orden de 2 módulo q es de la forma 2^j, con $0 \leq j \leq m+1$. Si fuera $j \leq m$, tendríamos

$$2^{2^m} = \left(2^{2^j}\right)^{2^{m-j}} \equiv 1\ (q)\,,$$

lo que no es cierto. En consecuencia $\mathrm{ord}_q(2) = 2^{m+1}$. Puesto que el orden de cualquier elemento de $\mathbb{Z}_q^*$ es un divisor de $q-1$, deducimos que q es de la forma $2^{m+1}k+1$, pero veamos que todavía podemos decir algo más.

En efecto, sigue de la última afirmación que $q \equiv 1\ (8)$, por lo que 2 es un residuo cuadrático módulo q. Con este dato adicional podemos determinar la paridad de k, ya que tomando congruencias módulo q tenemos:

$$1 \equiv 2^{(q-1)/2} = 2^{2^m k} = (2^{2^m})^k \equiv (-1)^k,$$

lo que asegura que k es par. En definitiva, los divisores primos de F_m tienen la forma general $2^{m+2}u+1$. Esto por supuesto no los determina, pero es claro que reduce sensiblemente su búsqueda. Por ejemplo, en el caso particular de $F_5 = 2^{32}+1 = 4294967297$, una búsqueda por ensayo y error de sus divisores primos nos enfrentaría con 6542 primos (los que son menores ó iguales que su raíz cuadrada), pero resulta que sólo 99 de ellos (menos del 2 %) son de la forma $128k+1$, que es la que corresponde al caso $m=5$. Si comenzamos a ensayar con éstos rápidamente nos sonríe la fortuna, ya que el segundo de ellos (641) lo divide. Precisamente, F_5 es producto de dos números primos, a saber, 641 y 6700417. $\Diamond$

7.2.2. Sumas de cuadrados

Sigue del ejercicio 31 de la sección 5.3 que la longitud de la hipotenusa de un triángulo rectángulo de lados enteros es necesariamente una suma de dos cuadrados, por lo que tiene interés caracterizar el conjunto de estos números, que incluye por supuesto los cuadrados perfectos. Por ejemplo, si nos limitamos a los primeros 30 números naturales, un sencillo y algo tedioso examen nos muestra que los números en ese rango que se obtienen como suma de dos cuadrados son 1, 2, 4, 5, 8, 9, 10, 13, 16, 17, 18, 20, 25, 26 y 29. Si bien a simple vista no parece observarse ningún patrón en la lista, pronto veremos que existe uno, relacionado con la factorización de los números.

Comencemos por observar que la propiedad que nos ocupa es *multiplicativa*, vale decir, el producto de dos números enteros que son suma de dos cuadrados también es suma de dos cuadrados, pues vale la igualdad

$$(x^2+y^2)(z^2+w^2) = (xz-yw)^2 + (xw+yz)^2.$$

Ello nos conduce naturalmente a estudiar qué números primos se expresan como suma de dos cuadrados. Puesto que obviamente $2 = 1^2+1^2$ es uno

de ellos, nos ocuparemos en adelante de primos impares. El resultado clave es el siguiente:

Proposición 7.2.11 Un primo impar p es suma de dos cuadrados si y sólo si p es congruente con 1 módulo 4.

DEMOSTRACION Supongamos en primer término que p es suma de dos cuadrados, digamos

$$p = a^2 + b^2.$$

Deducimos entonces que ninguno de los enteros a y b es divisible por p, ya que si alguno de ellos lo fuera también lo sería el otro, por lo que escribiendo $a = p\,a_1$, $b = p\,b_1$ y operando convenientemente resultaría $1 = p\,(a_1 + b_1)$, lo que obviamente es absurdo.

Puesto que en particular b es inversible módulo p, multipliquemos ahora la relación $a^2 \equiv -b^2$ por t^2, donde t designa el inverso de b módulo p. Tenemos entonces:

$$(at)^2 \equiv -(bt)^2 \equiv -1,$$

esto es, -1 es un rc mod p, y por lo tanto $p \equiv 1\ (4)$ (proposición 7.2.7). Como se puede apreciar, la cuestión está íntimamente relacionada con el carácter cuadrático de -1 módulo p, hecho que también se verá reflejado cuando probemos que la condición es suficiente.

Recíprocamente, sea $p \equiv 1\ (4)$ y sea $c \in \mathbb{Z}$ tal que $c^2 \equiv -1$. Designaremos además por m la parte entera de la raíz cuadrada de p. Equivalentemente, m es el único número natural que satisface las desigualdades $m^2 < p < (m+1)^2$.

Motivados por la demostración anterior, comenzaremos asignando a cada par de enteros $(u, v) \in [\,0, m\,] \times [\,0, m\,]$ la clase módulo p de $u + cv$. Puesto que el número de tales pares es $(m+1)^2 > p$, dicha asignación *no puede ser inyectiva*, por lo que existen dos pares *distintos* (u_1, v_1) y (u_2, v_2) tales que

$$u_1 + cv_1 \equiv u_2 + cv_2\,. \tag{7.3}$$

Resulta además que $u_1 \neq u_2$ y $v_1 \neq v_2$. En efecto, supongamos que $v_1 = v_2$. Entonces $u_1 \equiv u_2$, y siendo ambos números no negativos y menores que p (por ser menores ó iguales que m) resulta que deben ser iguales, lo que contradice el hecho de que los pares son distintos. La suposición $u_1 = u_2$ lleva a la misma contradicción, habida cuenta de que c puede cancelarse en la relación $cv_1 \equiv cv_2$, por ser coprimo con p.

Sigue de (7.3) que

$$(u_1 - u_2)^2 \equiv c^2(v_2 - v_1)^2 \equiv -(v_2 - v_1)^2,$$

lo que en términos de divisibilidad significa que $(u_1 - u_2)^2 + (v_2 - v_1)^2$ es múltiplo de p. Examinemos cuidadosamente las diferencias que están entre paréntesis. En primer lugar, ambas son diferentes de 0 y podemos suponerlas positivas, ya que nada cambia si por ejemplo reemplazamos $u_1 - u_2$ por

$u_2 - u_1$. Por otro lado, ninguna de las dos supera a m, que es la medida del intervalo. En conclusión, hemos obtenido un múltiplo de p de la forma $a^2 + b^2$, donde $0 < a \leq m$ y $0 < b \leq m$. Puesto que entonces valen las desigualdades

$$0 < a^2 + b^2 \leq 2m^2 < 2p\,,$$

concluimos que la única posibilidad es $a^2 + b^2 = p$. Luego p es suma de dos cuadrados, como queríamos demostrar. $\Diamond$

NOTA En la prueba de la primera implicación de 7.2.11, luego de establecer que a y b eran coprimos con p, demostramos que -1 era un rc mod p usando solamente que $a^2 + b^2$ era divisible por p. Puesto que tal conclusión no es válida si $p \equiv 3\ (4)$, queda probada la siguiente propiedad, de interés en sí misma:

Corolario 7.2.12 Sea $p \equiv 3\ (4)$ y sean $a, b \in \mathbb{Z}$. Entonces

$$p \mid a^2 + b^2 \Leftrightarrow p \mid a \text{ y } p \mid b\,. \quad \Diamond$$

Ya caracterizados los números primos que se descomponen como suma de dos cuadrados, podemos encarar el problema general de caracterizar los números naturales n que son suma de dos cuadrados. Como anticipamos más arriba, la cuestión depende de la factorización de n. Concretamente, probaremos el siguiente teorema:

Teorema 7.2.13 Un número natural n es suma de dos cuadrados si y sólo todo primo congruente con 3 módulo 4 aparece con exponente par en su factorización.

DEMOSTRACION Por ejemplo, $2754 = 27^2 + 45^2$ es suma de dos cuadrados, correspondiendo al hecho de que $2754 = 2\,.\,3^4\,.\,17$. En cambio 175 no lo es, ya que $175 = 5^2\,.\,7$. De acuerdo con el teorema, observemos además que un número natural no divisible por ningún primo de la forma $4k + 3$ es suma de dos cuadrados, ya que dichos primos aparecen con exponente 0 en su factorización.

Comencemos probando la suficiencia de la condición. Es claro por el teorema fundamental de la Aritmética que todo número natural n se expresa unívocamente en la forma

$$n = q_1 q_2 \ldots q_m u^2,$$

donde los q_i son los primos que aparecen con exponente impar en su factorización, resultando en el caso de que n satisfaga la condición del enunciado que $q_i = 2$ ó $q_i \equiv 1\ (4)$ para todo i. Puesto que u^2 y cada uno de estos primos es suma de dos cuadrados (proposición 7.2.11), concluimos que n también lo es, ya que la propiedad de ser suma de dos cuadrados es multiplicativa.

Para probar la recíproca, supongamos por el absurdo que

$$n = a^2 + b^2 \tag{7.4}$$

y que un cierto primo q congruente con 3 módulo 4 aparece con exponente impar en la factorización de n. Puesto que $q \mid a^2+b^2$, sigue de 7.2.12 que a y b son múltiplos de q ; tomando en cada caso la máxima potencia de q que los divide, podemos escribir $a = q^r c$ y $b = q^s d$, donde $q \nmid c$ y $q \nmid d$. Suponiendo sin pérdida de generalidad que $r \leq s$, reemplazando en (7.4) obtenemos

$$n = q^{2r}\left(c^2 + (q^{s-r}d)^2)\right).$$

Teniendo en cuenta nuestra suposición acerca del exponente de q en la factorización de n, concluimos que q divide al factor $c^2+(q^{s-r}d)^2$. Utilizando nuevamente la propiedad 7.2.12 arribamos a la contradicción deseada, ya que $q \nmid c$. Esto completa la demostración del teorema. $\Diamond$

NOTA Aunque el hecho requiere una demostración bastante más complicada, pueden caracterizarse también los números naturales que son suma de tres cuadrados. En este caso es más práctico brindar una respuesta por la negativa: un número natural no es suma de tres cuadrados si y sólo si es de la forma $4^m(8k + 7)$, donde m y k son enteros no negativos. Por ejemplo, 7 es el menor número natural que no es suma de tres cuadrados.

Señalemos por último que el resultado más importante referido a sumas de cuadrados se debe a Lagrange, que demostró en 1770 que *todo número natural es suma de cuatro cuadrados* . $\Diamond$

7.2.3. Ejercicios

1. Listar los residuos cuadráticos módulo 17, 19 y 37.

2. Calcular 2^{2935} mod 97 y 2^{4150} mod 83.

3. Sea p un primo y sean $a, b \in \mathbb{Z}$. Probar que ab es un residuo cuadrático módulo p si y sólo si a y b lo son ó ninguno de los dos lo es.

4. Probar que -2 es un rc mod p si y solo si $p \equiv 1\ (8)$ ó $p \equiv 3\ (8)$.

5. Resolver las siguientes ecuaciones de congruencia:

   a) $x^2 \equiv -1\ (17)$
   b) $x^2 \equiv -1\ (47)$
   c) $x^2 \equiv 2\ (23)$
   d) $x^4 \equiv 3\ (13)$

e) $2x^2 + 7x - 6 \equiv 2\ (11)$
f) $x(x-1) \equiv 0\ (12)$.

6. Si $n \in \mathbb{N}$ probar que todo divisor impar de n^2+1 es de la forma $4k+1$.

7. Sean p y q primos tales que $p \equiv 1\ (4)$ y $q \equiv 3\ (4)$. Demostrar:
   a) La suma de los residuos cuadráticos módulo p es divisible por p.
   b) $\left(\frac{p-1}{2}\right)!$ es una raíz cuadrada de -1 módulo p.
   c) Si a es un residuo cuadrático módulo q entonces $a^{\frac{q+1}{4}}$ es una raíz cuadrada de a módulo q.

8. Hallar las raíces cuadradas de los rc mod 19 y mod 31.

9. Sean p y q como en el ejemplo 7.2.9. Probar que -2 es una raíz primitiva módulo q.

10. Si $s \in \mathbb{N}$, probar que $2^s + 1$ es primo solo si s es una potencia de 2.

11. Probar que $n + 1 < 2^{n-1}$ para todo $n > 3$.

12. Sea p un primo de Fermat y sea a un residuo no cuadrático módulo p. Probar que a es una raíz primitiva módulo p.

13. Sea p un primo tal que $p \not\equiv 1\ (3)$. Probar que dado $a \in \mathbb{Z}$ existe $x \in \mathbb{Z}$ tal que $x^3 \equiv a\ (p)$.

14. Demostrar que los números de Fermat satisfacen la relación
$$F_m = \prod_{i=0}^{m-1} F_i + 2$$
cualquiera sea $m \in \mathbb{N}$. Deducir que F_r y F_s son coprimos si $r \neq s$ (esto brinda otra demostración de la existencia de infinitos primos).

15. Analizar la primalidad de M_{911}.

16. Verificar que F_4 es primo. ¿Cuántas divisiones son necesarias?

17. Exhibir la lista los números naturales menores ó iguales que 100 que son suma de dos cuadrados y también la de los que son suma de tres cuadrados.

18. Probar que ningún número natural de la forma $4^m(8k+7)$ es suma de tres cuadrados.

Capítulo 8

Números Complejos

8.1. El cuerpo de los números complejos

8.1.1. Introducción

Como recordará el lector, hemos probado en el capítulo 2 que un número real admite raíz cuadrada en $\mathbb{R}$ si y sólo si es no negativo. Nos proponemos ahora ampliar convenientemente el cuerpo de los números reales de manera que la ecuación

$$X^2 = a$$

tenga solución en la nueva estructura cualquiera sea $a \in \mathbb{R}$.

La cuestión se reduce a la búsqueda de una raíz cuadrada de -1. En efecto, supongamos que tenemos (en algún conjunto que contiene a $\mathbb{R}$) una solución t de la ecuación $X^2 = -1$, y sea a un número real negativo. Resulta entonces que

$$\left(t\sqrt{|\,a\,|}\right)^2 = t^2\,|\,a\,| = -\,|\,a\,| = a,$$

y por lo tanto la ecuación $X^2 = a$ tiene solución en el conjunto ampliado.

Por supuesto que la ampliación requerida no se obtiene simplemente agregando un elemento t al cuerpo de los números reales, ya que la misma debe estar dotada de una estructura algebraica de manera que tengan sentido las operaciones efectuadas arriba. A fin de construirla adecuadamente, definiremos dos operaciones en el conjunto de pares ordenados de números reales, a las que denominaremos genéricamente suma y producto.

8.1.2. Definición y estructura de cuerpo.

Designaremos con la letra $\mathbb{C}$ el conjunto $\mathbb{R}^2$. Un elemento de $\mathbb{C}$ será llamado un *número complejo*.

SUMA Y PRODUCTO Dados números complejos $z = (a, b)$ y $w = (c, d)$, de-

finimos su suma y su producto en la forma

$$z + w = (a + c, b + d)$$
$$zw = (ac - bd, ad + bc)\,.$$

Por ejemplo, $(2,4)+(-1,3) = (1,7)$ y $(2,4)(-1,3) = (-14,2)$. Obsérvese que empleamos la simbología habitual para notar estas operaciones entre pares, y que las hemos definido usando, en una forma que será útil a nuestros planes, la suma y el producto de números reales. En el siguiente teorema justificaremos esta última afirmación.

Teorema 8.1.1 Con las operaciones definidas arriba $\mathbb{C}$ es un cuerpo.

DEMOSTRACION. Las propiedades asociativas y conmutativas de la suma y el producto son consecuencia inmediata de las correspondientes propiedades de la suma y el producto en $\mathbb{R}$, y lo mismo sucede con la propiedad distributiva del producto respecto a la suma, por lo que dejamos a cargo del lector la demostración de estas afirmaciones. Sigue también fácilmente de la definición que $(0,0)$ es elemento neutro de la suma y que todo $z = (a,b)$ admite inverso aditivo, a saber $(-a,-b)$, que notaremos $-z$.

Puesto que se comprueba sin inconvenientes que $(1,0)$ es elemento neutro del producto, sólo resta demostrar que todo elemento distinto de $(0,0)$ admite inverso multiplicativo. Sea pues $z = (a,b)$, donde $a \neq 0$ ó $b \neq 0$, y probemos que existe $w = (x,y)$ tal que $zw = (1,0)$.

Usando la definición de producto e igualando componentes, resulta que (x,y) debe satisfacer el sistema de ecuaciones

$$ax - by = 1 \tag{8.1}$$
$$bx + ay = 0\,, \tag{8.2}$$

que podemos resolver en forma sencilla. En efecto, si multiplicamos 8.1 por a y 8.2 por b, y luego sumamos, obtenemos:

$$(a^2 + b^2)x = a,$$

de donde sigue que $x = a/(a^2+b^2)$ (observemos que $a^2+b^2 \neq 0$). De manera similar despejamos y, resultando que $y = -b/(a^2 + b^2)$.

En definitiva z admite inverso multiplicativo, siendo

$$z^{-1} = \left(\frac{a}{a^2+b^2}\,,\ \frac{-b}{a^2+b^2}\right). \quad \Diamond$$

RESTA Y COCIENTE Como en el caso real, la existencia de inverso aditivo y multiplicativo permite definir la resta y el cociente de dos números complejos. Precisamente, si $z = (a,b)$ y $w = (c,d)$ definimos:

$$z - w = z + (-w) = (a - c, b - d)$$

$$\frac{z}{w} = zw^{-1} = \left(\frac{ac+bd}{c^2+d^2}\,,\ \frac{-ad+bc}{c^2+d^2}\right).$$

Naturalmente, en el caso del cociente suponemos $w \neq (0,0)$. $\Diamond$

Cabe ahora preguntarnos si el cuerpo $\mathbb{C}$ de números complejos responde a nuestro objetivo de construir una extensión de $\mathbb{R}$ que contenga una raíz cuadrada de -1. Estrictamente hablando no, ya que $\mathbb{R}$ ni siquiera está contenido en $\mathbb{C}$, pero como veremos a continuación, lograremos nuestro objetivo identificando adecuadamente a $\mathbb{R}$ con un cierto subconjunto de $\mathbb{C}$.

Consideremos para ello la función

$$\theta : \mathbb{R} \to \mathbb{C}$$

definida por $\theta(a) = (a, 0)$. Es trivial probar que θ es inyectiva, lo que establece una biyección entre $\mathbb{R}$ y el conjunto de elementos de $\mathbb{C}$ cuya segunda componente es nula. Si bien hay muchas funciones inyectivas entre ambos conjuntos, por ejemplo las asignaciones $a \mapsto (0, a)$ ó $a \mapsto (a, a)$, y por lo tanto muchas formas de identificar $\mathbb{R}$ con un subconjunto de $\mathbb{C}$, la aplicación θ satisface una serie de propiedades algebraicas que la distingue de cualquier otra inyección. Concretamente, son válidos los siguientes hechos ($x, y \in \mathbb{R}$):

a) $\theta(x) + \theta(y) = \theta(x + y)$.

b) $\theta(x)\theta(y) = \theta(xy)$.

c) $\theta(0) = (0, 0)$.

d) $\theta(1) = (1, 0)$.

e) $\theta(-x) = -\theta(x)$.

f) $\theta(x^{-1}) = (\theta(x))^{-1}$ si $x \neq 0$.

Todas estas propiedades, de verificación inmediata, demuestran que la imagen de θ es un cuerpo (un *subcuerpo* de $\mathbb{C}$). Observando que θ distribuye las operaciones y que aplica elementos neutros en elementos neutros e inversos en inversos, es natural pensar entonces que dicho subcuerpo es una "copia" algebraica del cuerpo de los números reales, por lo que identificaremos ambos conjuntos.

Resumiendo, supondremos a partir de aquí que vale la inclusión $\mathbb{R} \subset \mathbb{C}$ a través de la identificación

$$\mathbb{R} = \{(a, b) \in \mathbb{C} : b = 0\} \ .$$

En consonancia con ello, si x es un número real el elemento $(x, 0)$ será notado simplemente x. Resulta en particular que 0 y 1 son los elementos neutros de la suma y el producto en $\mathbb{C}$, respectivamente. Como en el caso real, el conjunto de números complejos no nulos será notado $\mathbb{C}^*$.

FORMA BINOMICA Designaremos por $\imath$ el número complejo $(0,1)$, y lo llamaremos *unidad imaginaria*. Si $z = (a,b)$ es un número complejo cualquiera, de acuerdo con nuestras convenciones de notación resulta que

$$z = (a,b) = (a,0) + (0,b) = (a,0) + (b,0)(0,1) = a + b\,\imath\,.$$

La expresión $a + b\,\imath$ se llama la *forma binómica* de z, mientras que los números reales a y b se denominan la *parte real* y la *parte imaginaria* de z. Las notaremos $\mathrm{Re}(z)$ y $\mathrm{Im}(z)$, respectivamente. Por ejemplo, si $z = 2 - 3\imath$ tenemos $\mathrm{Re}(z) = 2$ y $\mathrm{Im}(z) = -3$.

Si $z \in \mathbb{C}$, notemos que $\mathrm{Re}(z)$ y $\mathrm{Im}(z)$ son las componentes de z en la presentación original de los números complejos como pares ordenados. Por lo tanto

$$z = w \Leftrightarrow \mathrm{Re}(z) = \mathrm{Re}(w) \text{ y } \mathrm{Im}(z) = \mathrm{Im}(w),$$

cualesquiera sean $z, w \in \mathbb{C}$. Como caso particular resulta que $z = 0$ si y sólo si $\mathrm{Re}(z) = \mathrm{Im}(z) = 0$. Traduzcamos a la nueva notación nuestras definiciones y consideraciones previas:

i) $z \in \mathbb{R} \Leftrightarrow \mathrm{Im}(z) = 0$.

ii) Si $\mathrm{Re}(z) = 0$ diremos que z es *imaginario puro*. Por ejemplo $\imath$ es imaginario puro.

iii) $(a + b\imath) + (c + d\imath) = (a + c) + (b + d)\imath$.

iv) $(a + b\imath)(c + d\imath) = (ac - bd) + (ad + bc)\imath$.

v) $-(a + b\imath) = -a - b\imath$.

vi) $(a + b\imath)^{-1} = a/(a^2 + b^2) - b/(a^2 + b^2)\imath$. $\quad \Diamond$

Volviendo al objetivo principal de nuestra construcción, veamos que efectivamente $\mathbb{C}$ contiene una raíz cuadrada de -1 (en realidad dos).

Proposición 8.1.2 $\imath^2 = -1$.

DEMOSTRACION. Basta usar la fórmula del producto de complejos exhibida arriba en el ítem iv), habida cuenta de que $\mathrm{Re}(\imath) = 0$ y $\mathrm{Im}(\imath) = 1$. Puesto que $-\imath = (-1)\imath$, es claro que $-\imath$ también es una raíz cuadrada de -1. $\quad \Diamond$

NOTA. $\mathbb{C}$ no es un cuerpo ordenado, es decir, de no es posible definir en $\mathbb{C}$ una relación de orden $(\leq)$ que sea compatible con las operaciones, como en el caso real. En efecto, supongamos por el contrario que existe un tal orden. Por las mismas razones que en $\mathbb{R}$ vale entonces la relación $0 < 1$, y por la ley de tricotomía deberá ser $\imath > 0$ ó $\imath < 0$. Suponiendo $\imath > 0$, resulta multiplicando por $\imath$ que $-1 > 0$, o equivalentemente $1 < 0$, lo que es una contradicción. En el segundo caso se llega a la misma situación, ya que $\imath < 0$

si y sólo si $-\imath > 0$, y podemos repetir el argumento anterior con $-\imath$ en vez de $\imath$. $\Diamond$

RAICES CUADRADAS DE UN NUMERO REAL NEGATIVO. Como comentamos en la introducción de este capítulo, la existencia de una raíz cuadrada de -1 asegura que todo número real negativo a admite una raíz cuadrada en $\mathbb{C}$. En realidad, admite exactamente dos, a saber: $\sqrt{|\,a\,|}\,\imath$ y $-\sqrt{|\,a\,|}\,\imath$.

Por ejemplo, las raíces cuadradas complejas de -16 son $4\imath$ y $-4\imath$. Vale la pena detenernos en la afirmación que hemos hecho arriba, acerca del número de raíces cuadradas. Por una parte, es claro que ambos números son soluciones de la ecuación $X^2 = a$ y que son distintos. Para probar que son las únicas soluciones, llamemos u a cualquiera de ellos y consideremos cualquier solución en $\mathbb{C}$ de la ecuación, digamos v. Sigue entonces que

$$0 = v^2 - u^2 = (v-u)(v+u),$$

y por lo tanto $v - u = 0$ ó $v + u = 0$. En definitiva, $v = u$ ó $v = -u$, como queríamos. Demostraremos más adelante un resultado idéntico respecto a la ecuación $X^2 = z$, donde z es cualquier complejo no nulo, pero dejemos en claro que sólo emplearemos el símbolo $\sqrt{x}$ para denotar la única raíz cuadrada real no negativa de un número real x no negativo. $\Diamond$

Puesto que $\mathbb{C}$ es un cuerpo, valen en él todas las propiedades que se derivan de los axiomas de cuerpo. En particular las propiedades de la potencia de exponente entero, que se define en forma totalmente análoga al caso real. Mostraremos a continuación que las potencias de la unidad imaginaria observan un comportamiento muy regular.

Lema 8.1.3 Sea $m \in \mathbb{Z}$. Entonces $\imath^m = \imath^{r_4(m)}$.

DEMOSTRACION. Observemos en primer lugar que $\imath^4 = (\imath^2)^2 = (-1)^2 = 1$. Luego, escribiendo $m = 4t + r_4(m)$, obtenemos

$$\imath^m = \imath^{4t+r_4(m)} = (\imath^4)^t \imath^{r_4(m)} = 1^t \imath^{r_4(m)} = \imath^{r_4(m)},$$

como queríamos probar. Finalmente, siendo $\imath^0 = 1$, $\imath^1 = \imath$, $\imath^2 = -1$ e $\imath^3 = -\imath$, obtenemos la fórmula general:

$$\imath^m = \begin{cases} 1 & \text{si } m \equiv 0\ (4) \\ \imath & \text{si } m \equiv 1\ (4) \\ -1 & \text{si } m \equiv 2\ (4) \\ -\imath & \text{si } m \equiv 3\ (4)\,. \end{cases}$$

Como caso particular resulta que $\imath^{-1} = -i$, pues $-1 \equiv 3\ (4)$. Esto es, el inverso multiplicativo y el inverso aditivo de $\imath$ coinciden. $\Diamond$

8.1.3. Ejercicios

1. Expresar como pares ordenados los números complejos $1+2\imath$, π, $(-\imath)^7$, $2+(-3)$, $\imath-2$ y $4-\imath$.

2. Si $z=2-3\imath$ y $w=1+4\imath$, expresar en forma binómica los siguientes elementos de $\mathbb{C}$:

   a) $z+w$
   b) $w-2z$
   c) $6z$
   d) z^2w
   e) $1+w^{-2}$
   f) $\dfrac{z-2}{w+1}$
   g) $\dfrac{w}{z^3}$
   h) $\mathrm{Im}(w-z)\imath^{154}$
   i) $(\mathrm{Im}(z)+\mathrm{Re}(-w)\imath)\,\imath$.

3. Resolver en $\mathbb{C}$ las siguientes ecuaciones:

   a) $(z+\imath)(z-\imath)=-5/4$
   b) $z^2=3\imath z$
   c) $z^2-2z+5=0$
   d) $z^4=81$.

8.2. El plano complejo

8.2.1. Representación gráfica

Puesto que los números complejos están definidos como pares ordenados de números reales, resulta natural representarlos en el plano, asignando a cada $z = a+b\imath$ el punto de coordenadas cartesianas (a, b). Alternativamente, también pensaremos a z como un segmento dirigido o *vector*, que indicaremos por $\vec{z}$, con origen en el punto $(0, 0)$ y extremo en (a, b). Es claro entonces que a cada número complejo le corresponde un punto del plano y que cada punto del plano representa un número complejo. Este modelo geométrico de $\mathbb{C}$ será llamado el *plano complejo*.

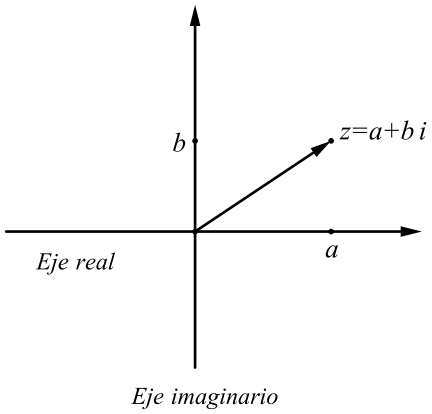


Figura 8.1: El plano complejo

Por ejemplo, los números reales se representan en el eje de las abscisas (eje real) y los imaginarios puros en el eje de las ordenadas (eje imaginario). A través de consideraciones geométricas elementales, es fácil ver que $\overrightarrow{z+w}$ se obtiene aplicando la regla del paralelogramo a los vectores $\vec{z}$ y $\vec{w}$. En particular, y teniendo en cuenta que $z = (z - w) + w$, resulta que $\overrightarrow{z-w}$ es paralelo al segmento que une los puntos z y w, y de igual longitud que éste. Más adelante mostraremos cuál es la interpretación vectorial del producto en $\mathbb{C}$.

Módulo de un numero complejo.

Sea $z = a + b\imath$ en $\mathbb{C}$. Definimos entonces el *módulo* o *valor absoluto* de z en la forma:

$$|\, z \,| = \sqrt{a^2 + b^2}\,.$$

Por ejemplo, $|\, 3 - 4\imath \,| = 5$ y $|\, 2\imath \,| = 2$. Notemos que por definición $|\, z \,|$ es un número real no negativo, y que dicha definición extiende a $\mathbb{C}$ la noción

de módulo definida en $\mathbb{R}$. En efecto, supongamos que $z \in \mathbb{R}$. En tal caso tenemos:

$$|\,z\,| = |\,a + 0\,\imath\,| = \sqrt{a^2 + 0^2} = \sqrt{a^2} = |\,a\,|\,.$$

Más aún, la extensión del módulo a $\mathbb{C}$ tiene idéntico significado geométrico que en $\mathbb{R}$, ya que aplicando el teorema de Pitágoras deducimos inmediatamente que $|\,z\,|$ es la longitud del vector $\vec{z}$, vale decir, representa la distancia de z al origen.

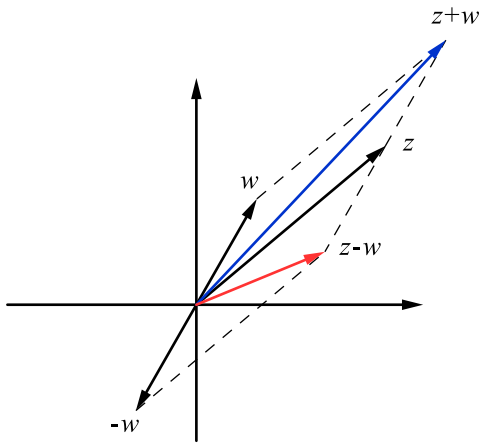


Figura 8.2: Suma y resta en el plano complejo

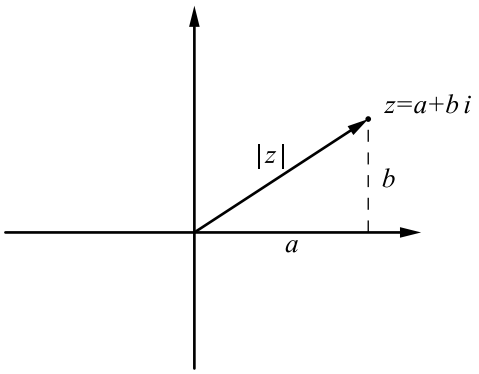


Figura 8.3: Significado geométrico del módulo

En la siguiente proposición exhibiremos las principales propiedades del valor absoluto (las letras designan números complejos).

Proposición 8.2.1 La función módulo satisface las siguientes propiedades:

1) $|\,z\,| \geq 0$ y $|\,z\,| = 0 \iff z = 0$

2) $|\,-z\,| = |\,z\,|$

3) $|\,zw\,| = |\,z\,|\,|\,w\,|$. Deducimos entonces que $\left|\,u^k\,\right| = |\,u\,|^k$ para todo número natural k.

4) $\left|\,w^{-1}\,\right| = |\,w\,|^{-1}$ si $w \neq 0$. Luego $|\,z/w\,| = |\,z\,| \,/\, |\,w\,|$

5) $|\,\mathrm{Re}(z)\,| \leq |\,z\,|$ y $|\,\mathrm{Im}(z)\,| \leq |\,z\,|$. En el primer caso vale la igualdad si y sólo si z es real, y en el segundo caso si y sólo si z es imaginario puro

6) (propiedad triangular) $|\,z + w\,| \leq |\,z\,| + |\,w\,|$.

Demostracion. La propiedad 1) sigue de las propiedades del orden en $\mathbb{R}$, mientras que las propiedades 2) a 4) se obtienen por simple cálculo a partir de las definiciones. Dejamos pues sus demostraciones a cargo del lector. Respecto a 5), si $z = a + b\imath$ tenemos:

$$|\,a\,|^2 = a^2 \leq a^2 + b^2 = |\,z\,|^2\,,$$

de donde resulta que $|\,a\,| \leq |\,z\,|$. Similarmente se prueba que $|\,b\,| \leq |\,z\,|$. Las restantes afirmaciones son evidentes.

Para demostrar 6) escribamos $z = a + b\imath$, $w = c + d\imath$ y sea $u = c - d\imath$. Usando las propiedades 3) y 5) y observando que $|\,u\,| = |\,w\,|$, obtenemos:

$$\begin{aligned} |\,z + w\,|^2 &= (a+c)^2 + (b+d)^2 = a^2 + b^2 + c^2 + d^2 + 2(ac + bd) = \\ &= |\,z\,|^2 + |\,w\,|^2 + 2\mathrm{Re}(zu) \leq |\,z\,|^2 + |\,w\,|^2 + 2\,|\,\mathrm{Re}(zu)\,| \leq \\ &\leq |\,z\,|^2 + |\,w\,|^2 + 2\,|\,zu\,| = |\,z\,|^2 + |\,w\,|^2 + 2\,|\,z\,|\,|\,u\,| = \\ &= |\,z\,|^2 + |\,w\,|^2 + 2\,|\,z\,|\,|\,w\,| = (|\,z\,| + |\,w\,|)^2\,, \end{aligned}$$

lo que prueba nuestra afirmación. $\Diamond$

Raices Cuadradas de un numero complejo. Hemos visto que todo número real no nulo admite dos raíces cuadradas en $\mathbb{C}$. Probaremos ahora que este hecho se extiende a todo número complejo no nulo.

Proposición 8.2.2 Sea $z = a + b\imath$ un número complejo no nulo. Entonces z admite dos raíces cuadradas en $\mathbb{C}$.

Demostracion. Supondremos $b \neq 0$, pues ya hemos resuelto el caso real. Para demostrar el enunciado, estudiemos la forma de las posibles raíces cuadradas de z en $\mathbb{C}$. Si $w = x + y\imath$ es una de ellas (en cuyo caso $-w$ es la otra), igualando partes reales e imaginarias en la ecuación $w^2 = z$ vemos que deben satisfacerse las ecuaciones reales

$$x^2 - y^2 = a \tag{8.3}$$

$$2xy = b\,. \tag{8.4}$$

Ahora bien, puesto que el módulo de un producto es el producto de los módulos, también debe verificarse la condición $|\,w\,|^2 = |\,z\,|$, lo que agrega la ecuación

$$x^2 + y^2 = |\,z\,| \,. \tag{8.5}$$

Sumando ahora (8.3) y (8.5) y dividiendo por 2 obtenemos

$$x^2 = \frac{|\,z\,| + a}{2} \,. \tag{8.6}$$

Similarmente, restando (8.3) de (8.5) y dividiendo por 2 arribamos a la igualdad

$$y^2 = \frac{|\,z\,| - a}{2} \,. \tag{8.7}$$

Observemos que los miembros derechos de las ecuaciones (8.6) y (8.7) son positivos, ya que valen las desigualdades $-|\,z\,| < a < |\,z\,|$, por el inciso 5) de la proposición 8.2.1. Por lo tanto (8.6) y (8.7) son resolubles en $\mathbb{R}$, lo que en principio nos brindaría 4 soluciones, pues hay dos valores posibles para x y dos valores posibles para y. Sin embargo no es así, ya que (8.4) nos indica que los signos de x e y deben coincidir si $b > 0$ y deben ser distintos si $b < 0$.

En definitiva, z admite dos raíces cuadradas en $\mathbb{C}$, a saber:

Fórmula 8.2.3

$$w \;=\; \pm\left(\sqrt{\frac{|\,z\,| + a}{2}} \;+\; \mathrm{sg}(b)\,\sqrt{\frac{|\,z\,| - a}{2}}\;\imath\right) ,$$

donde

$$\mathrm{sg}(b) = \begin{cases} 1 & \text{si } b > 0 \\ -1 & \text{si } b < 0 \,. \end{cases}$$

En rigor sólo hemos demostrado que las raíces cuadradas de z deben ser de la forma anterior, pero es mera rutina verificar que las condiciones halladas también son suficientes, es decir, que los w dados por la fórmula 8.2.3 satisfacen la igualdad $w^2 = z$.

Por ejemplo, aplicando la fórmula resulta que las raíces cuadradas de $3 - 4\imath$ son $\pm(2 - \imath)$, mientras que las de $\imath$ son $\pm\left(\sqrt{2}/2 + \sqrt{2}/2\,\imath\right)$. $\quad \Diamond$

Distancia.

Si $z, w \in \mathbb{C}$, definimos la *distancia* entre z y w en la forma

$$\delta(z, w) = |\,z - w\,| \,.$$

Por ejemplo, $\delta(6+5\imath, 1-7\imath) = 13$, $\delta(1+2\imath, 3-4\imath) = \sqrt{40}$ y $\delta(z, 0) = |\, z\,|$, cualquiera sea $z \in \mathbb{C}$. Observemos que la definición anterior responde a nuestro concepto geométrico de distancia, puesto que como hemos visto, el vector $\overrightarrow{z-w}$ tiene la misma longitud que el segmento que une los puntos z y w. Enunciaremos a continuación las propiedades fundamentales de la distancia. Las mismas se derivan de correspondientes propiedades del valor absoluto, por lo que obviaremos las demostraciones ($z, w, u \in \mathbb{C}$):

Proposición 8.2.4 Valen las siguientes propiedades:

1) $\delta(z, w) \geq 0$ y $\delta(z, w) = 0 \iff z = w$

2) $\delta(z, w) = \delta(w, z)$

3) $\delta(z, u) \leq \delta(z, w) + \delta(w, u)$. $\quad \Diamond$

La última de estas propiedades se deduce de la propiedad triangular del módulo, y justifica esta denominación. En efecto, ella expresa que en un triángulo la longitud de un lado no excede la suma de las longitudes de los otros dos. En realidad la desigualdad es estricta si z, w y u realmente forman un triángulo, es decir, si no están alineados. Le encomendaremos probar esta afirmación en los ejercicios del final de la sección. Ahora veremos un par de ejemplos, en los que ofreceremos descripciones geométricas de ciertos subconjuntos del plano complejo.

Ejemplo 8.2.5 Caractericemos el conjunto:

$$L = \{z \in \mathbb{C} : |\, z + 1 - \imath\,| = |\, z - 1 - 3\imath\,|\}\,.$$

Lo haremos de dos maneras. En la primera, puramente algebraica, escribimos $z = a + b\imath$ y planteamos la condición (los módulos se elevan al cuadrado), resultando que:

$$(a+1)^2 + (b-1)^2 = (a-1)^2 + (b-3)^2\,.$$

Si ahora desarrollamos los cuadrados y simplificamos obtenemos:

$$2a - 2b + 1 = -2a - 6b + 9,$$

ó equivalentemente,

$$a + b = 2,$$

como se comprueba fácilmente. Luego $L = \{z \in \mathbb{C} : \mathrm{Re}(z) + \mathrm{Im}(z) = 2\}$ es una recta del plano complejo.

Alternativamente, podemos proceder de manera geométrica. Para ello, observemos que la condición que define a L puede expresarse en la forma

$$\delta(z, -1 + \imath) = \delta(z, 1 + 3\imath).$$

Por lo tanto (usando las notaciones $u = -1 + \imath$ y $v = 1 + 3\imath$), resulta que L consiste de los puntos del plano que equidistan de u y v. Como sabemos de la geometría elemental, dicho conjunto de puntos es la recta *mediatriz* del segmento que une los puntos u y v, esto es, la recta perpendicular al segmento que pasa por su punto medio. Para hallar su ecuación, notemos que $y = x+2$ es la ecuación de la recta S que pasa por u y v (las componentes de ambos la satisfacen) y que $2\imath$ es el punto medio del segmento, ya que pertenece a S y equidista de u y v. En conclusión, L tiene pendiente -1 y ordenada al origen 2, o sea:

$$L = \{z \in \mathbb{C} : \mathrm{Im}(z) = -\mathrm{Re}(z) + 2\} = \{z \in \mathbb{C} : \mathrm{Re}(z) + \mathrm{Im}(z) = 2\}.$$

Naturalmente, arribamos al mismo resultado. Ambas formas de caracterizar L son válidas, aunque con la segunda descubrimos rápidamente qué tipo de conjunto del plano es L.

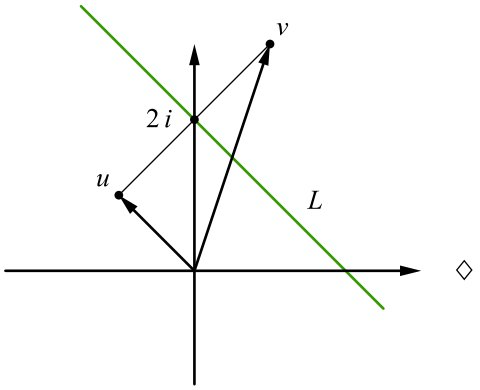


Figura 8.4: Recta mediatriz L

Ejemplo 8.2.6 Consideremos el subconjunto de $\mathbb{C}$

$$A = \{z \in \mathbb{C} : 1 < |\, z - 2\imath \,| < 3\}.$$

Puesto que la distancia entre dos puntos es el módulo de su diferencia, resulta que A consiste de los puntos cuya distancia a $2\imath$ es mayor que 1 y menor que 3. Por lo tanto, A es la corona circular determinada por las circunferencias de centro $2\imath$ y radios 1 y 3, respectivamente.

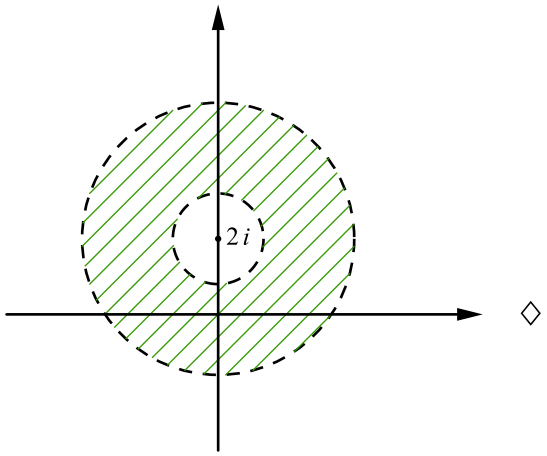


Figura 8.5: Corona circular

CONJUGACIÓN. Si $z = a + b\imath \in \mathbb{C}$, definimos el *conjugado* de z en la forma

$$\overline{z} = a - b\imath\,.$$

Por ejemplo, los conjugados de $1 + 3\imath$, 2 y $-4\imath$ son $1 - 3\imath$, 2 y $4\imath$, respectivamente. Notemos que la función conjugación $\sigma(z) = \bar{z}$ determina una biyección de $\mathbb{C}$, ya que

$$\sigma(\sigma(z)) = \sigma(a - b\imath) = a + b\imath = z,$$

esto es, σ es inversible y coincide con su inversa. En la figura de abajo ilustramos la posición relativa de un vector (de un número) respecto de su conjugado:

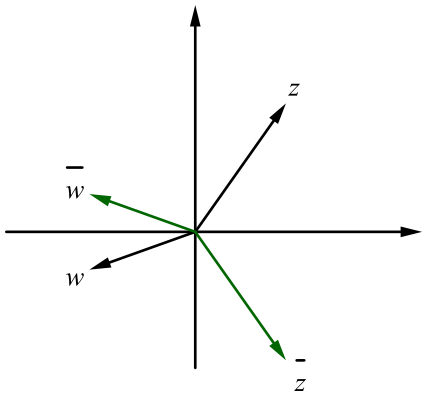


Figura 8.6: Conjugación

En la siguiente proposición apreciaremos el comportamiento regular de la conjugación con respecto a las operaciones, así como su relación con el módulo.

Proposición 8.2.7 La conjugación satisface las siguientes propiedades (las letras designan números complejos):

1) $\overline{z+w} = \overline{z} + \overline{w}$

2) $\overline{-u} = -\overline{u}$. Luego $\overline{z-w} = \overline{z} - \overline{w}$

3) $\overline{zw} = \overline{z}\ \overline{w}$. Usando un argumento inductivo sigue entonces que $\overline{u^k} = \overline{u}^{\,k}$ cualquiera sea $k \in \mathbb{N}$

4) Si $u \neq 0$ entonces $\overline{u^{-1}} = \overline{u}^{\,-1}$. Luego $\overline{z/w} = \overline{z}/\overline{w}$ si $w \neq 0$ y $\overline{u^k} = \overline{u}^{\,k}$ para todo $k \in \mathbb{Z}$

5) $z + \overline{z} = 2\mathrm{Re}(z)$ y $z - \overline{z} = 2\imath\,\mathrm{Im}(z)$. Se deduce luego que z es real si y sólo si $\overline{z} = z$ y que z es imaginario puro si y sólo si $\overline{z} = -z$

6) $|\,\overline{z}\,| = |\,z\,|$

7) $z\overline{z} = |\,z\,|^2$. Resulta entonces que

$$z^{-1} = \frac{\overline{z}}{|\,z\,|^2}$$

si $z \neq 0$.

DEMOSTRACION. Todas las fórmulas se obtienen y demuestran por simple aplicación de las correspondientes definiciones, por lo que dejamos los detalles a cargo del lector. ♢

8.2.2. Forma trigonométrica

Todo punto del plano complejo está unívocamente determinado por sus coordenadas cartesianas, que corresponden a la parte real e imaginaria del número complejo que representa. Hay sin embargo otros sistemas de coordenadas, esto es, otras formas posibles de situar un punto a través de la asignación de un par de datos numéricos. Por ejemplo, es geométricamente intuitivo que un punto P, distinto del origen O, queda completamente determinado si conocemos la longitud r del vector $\overrightarrow{OP}$ y la medida en radianes φ del ángulo que éste forma con la dirección positiva del eje horizontal. De esta manera se asocia a cada $P \neq O$ un par de números reales (r, φ), donde $r > 0$ y $0 \leq \varphi < 2\pi$, que llamamos las *coordenadas polares* de P.

Por supuesto que todo esto es algo impreciso, ya que debemos mostrar cómo calcular efectivamente r y φ y exhibir fórmulas de conversión entre coordenadas cartesianas y polares. Lo haremos trabajando directamente sobre números complejos, lo que nos dará una forma alternativa de expresarlos

que resultará muy útil a la hora de multiplicar y dividir en $\mathbb{C}$. Supondremos en lo que sigue que el lector está familiarizado con las propiedades básicas de las funciones trigonométricas.

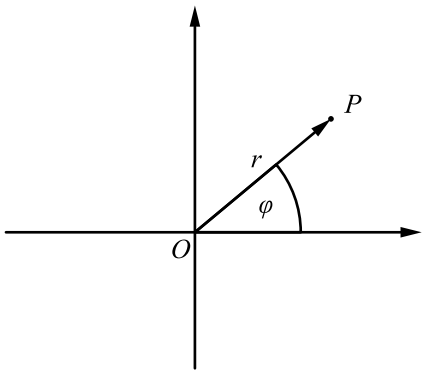


Figura 8.7: Coordenadas polares (r, φ)

ARGUMENTO DE UN NUMERO COMPLEJO Sea $z = a+b\imath$ un número complejo no nulo y designemos por r su módulo. Utilizando el ítem 4) de la proposición 8.2.1 sigue entonces que

$$\left| (a/r) + (b/r)\,\imath \right| = \left| z/r \right| = |z|\,/r = 1,$$

esto es, $(a/r) + (b/r)\,\imath$ pertenece a la circunferencia de radio 1 y centro en el origen. Puesto que todo punto de la misma es de la forma $(\cos t, \text{sen } t)$, con t variando entre 0 y 2π, concluimos que existe un número real $\theta\,, 0 \leq \theta < 2\pi$, tal que $a/r = \cos\theta$ y $b/r = \text{sen } \theta$. Despejando a y b en estas fórmulas y operando, llegamos entonces a la siguiente expresión de z:

$$z = r\,(\cos\theta + \imath \text{ sen } \theta)\,. \tag{8.8}$$

> El número θ se llamará *argumento principal* de z y lo notaremos $\arg(z)$, mientras que la expresión (8.8) se denominará la forma *trigonométrica* ó *polar* de z. Diremos asimismo que $(|\,z\,|\,, \arg(z))$ son sus coordenadas polares.

Resumiendo, si $z \in \mathbb{C}^*$ sus coordenadas cartesianas (a, b) y sus coordenadas polares (r, θ) se ligan por las fórmulas:

$$\begin{cases} r = \sqrt{a^2 + b^2} \\ \cos\theta = a/r \\ \text{sen } \theta = b/r\,. \end{cases} \qquad \Diamond$$

Ejemplos 8.2.8 Sea $z = -1 + \sqrt{3}\,\imath$. Entonces $|\,z\,| = 2$ y $\cos\theta = -1/2$, lo que nos brinda dos alternativas: $\theta = 2/3\,\pi$ ó $\theta = 2\pi - 2/3\,\pi = 4/3\,\pi$. Puesto que además $\operatorname{sen}\theta = \sqrt{3}/2 > 0$, resulta que $\theta < \pi$ y por lo tanto $\arg(z) = 2/3\,\pi$. Luego, la forma trigonométrica de z es

$$z = 2\,(\cos 2/3\,\pi + \imath \operatorname{sen} 2/3\,\pi)\,.$$

Si z es real o imaginario puro es muy sencillo determinar su argumento. En efecto, si $z \in \mathbb{R}$ resulta que $\cos\theta = \pm 1$, según que z sea positivo o negativo, y $\operatorname{sen}\theta = 0$. Luego $\arg(z) = 0$ si $z > 0$ y $\arg(z) = \pi$ si $z < 0$.

Si $z = b\imath$ es imaginario puro, sigue que $\cos\theta = 0$ y $\operatorname{sen}\theta = \operatorname{sg}(b)$. Por lo tanto, $\arg(z) = \pi/2$ si $b > 0$ y $\arg(z) = 3/2\,\pi$ si $b < 0$. $\quad\Diamond$

NOTA. Recordando que $\arg(z)$ sólo está definido para $z \neq 0$, hagamos notar que el argumento de z posee el sentido geométrico que queríamos darle, vale decir, es la medida del ángulo que forma el vector $\overrightarrow{z}$ con la dirección positiva del eje de abscisas, hecho que por ejemplo se manifiesta claramente si z es real o imaginario puro. Si bien no haremos una demostración general de esta afirmación, la ilustraremos con uno de los casos posibles.

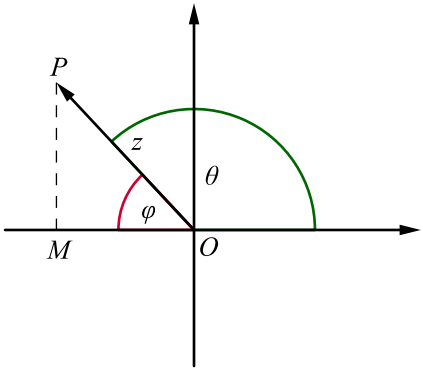


Figura 8.8: Forma trigonométrica

Supongamos que $z = a + b\imath$ es del segundo cuadrante ($a < 0$ y $b > 0$) y sean (r, θ) las coordenadas polares de z. Si consideramos en la figura 8.8 el triángulo rectángulo OPM, resulta que

$$\cos\varphi = OM/OP = -a/r = -\cos\theta = \cos(\pi - \theta)$$

y

$$\operatorname{sen}\varphi = PM/OP = b/r = \operatorname{sen}\theta = \operatorname{sen}(\pi - \theta).$$

Por lo tanto $\varphi = \pi - \theta$. Luego $\theta = \pi - \varphi$, como asegurábamos. $\quad\Diamond$

En adelante, por practicidad en la escritura emplearemos la notación

$$r\,(\cos x + \imath \operatorname{sen} x) = re^{\imath\, x}.$$

Por ejemplo, $e^{\imath\,\pi} = -1$. Esta manera alternativa de expresar los elementos de $\mathbb{C}^*$ debe interpretarse formalmente, pero ya veremos más adelante que es natural adoptarla.

Observemos que un número complejo no nulo admite en realidad infinitas expresiones de tipo (8.8). En efecto, siendo el seno y el coseno funciones reales de período 2π, dado $z = re^{\imath\,\theta}$ resulta que

$$z = re^{\imath(\theta+2\pi)} = re^{\imath(\theta+4\pi)} = \dots,$$

y en general

$$z = re^{\imath(\theta+2k\pi)} \tag{8.9}$$

cualquiera sea $k \in \mathbb{Z}$. Para clarificar la cuestión, comencemos por introducir la siguiente definición:

> Si $x, y \in \mathbb{R}$, diremos que x es *congruente* con y módulo 2π si y sólo si existe $m \in \mathbb{Z}$ tal que $x - y = 2m\pi$. Emplearemos en tal caso la notación $x \equiv y\ (2\pi)$.

Es inmediato probar que la congruencia módulo 2π es una relación de equivalencia en $\mathbb{R}$ y que toda clase de equivalencia tiene un único representante en el intervalo $[0, 2\pi)$. En el siguiente lema caracterizaremos las posibles expresiones trigonométricas de un número complejo.

Lema 8.2.9 Sea $w \in \mathbb{C}$ y sean $t, x \in \mathbb{R}\,(t > 0)$. Entonces $w = te^{\imath x}$ si y sólo si $t = |\,w\,|$ y $x \equiv \arg(w)\ (2\pi)$.

Demostracion. Sigue de (8.9) que las condiciones son suficientes. Recíprocamente, supongamos que $w = te^{\imath x}$ y sea $re^{\imath\,\theta}$ la forma trigonométrica de w. Tenemos en primer lugar que

$$r^2 = |\,w\,|^2 = |\,t\,|^2\,|\cos x + \imath\,\operatorname{sen} x|^2 = t^2\left(\cos^2 x + \operatorname{sen}^2 x\right) = t^2.$$

Siendo r y t positivos resulta que $t = r = |\,w\,|$.

Si igualamos ahora partes reales y partes imaginarias en las dos expresiones de w concluimos que $\cos x = \cos\theta$ y $\operatorname{sen} x = \operatorname{sen}\,\theta$. Por lo tanto:

$$\cos(x - \theta) = \cos x \cos\theta + \operatorname{sen} x\ \operatorname{sen}\theta = \cos^2 x + \operatorname{sen}^2 x = 1.$$

En consecuencia, existe $q \in \mathbb{Z}$ tal que $x - \theta = 2q\pi$, esto es,

$$x \equiv \arg(w)\ (2\pi)\,. \quad \Diamond$$

Resaltemos el significado de lo que hemos probado. Un complejo no nulo w admite infinitos argumentos, todos congruentes entre sí módulo 2π, siendo

el que designamos por $arg(w)$ y llamamos argumento principal el único de ellos que se encuentra en el rango $0 \leq \arg(w) < 2\pi$.

Calculemos por ejemplo el argumento de $w = \cos(\pi/5) - \imath \operatorname{sen}(\pi/5)$. Usando propiedades del seno y el coseno tenemos:

$$w = \cos(-\pi/5) + \imath \operatorname{sen}(-\pi/5) = \cos(9/5\,\pi) + \imath \operatorname{sen}(9/5\,\pi),$$

pues $-\pi/5 \equiv 9/5\,\pi\ (2\pi)$. Luego $\arg(w) = 9/5\,\pi$.

A partir de la siguiente proposición podremos apreciar la ventaja de emplear la forma polar para describir el significado geométrico de productos y cocientes y para el cálculo de potencias. Usaremos sin referencia algunas propiedades de la congruencia módulo 2π atinentes a su comportamiento respecto a las operaciones, ya que éste es idéntico al de la congruencia entera (las letras denotan números complejos no nulos y el símbolo $\equiv$ congruencia módulo 2π).

Proposición 8.2.10 Son válidas las siguientes propiedades:

1) $\arg(zw) \equiv \arg(z) + \arg(w)$

2) $\arg(z/w) \equiv \arg(z) - \arg(w)$

3) $\arg(u^{-1}) = \arg(\overline{u}) \equiv -\arg(u)$.

Demostracion.

1) Si $z = re^{\imath\alpha}$ y $w = se^{\imath\beta}$ son las respectivas formas polares, tenemos:

$$\begin{aligned} zw &= rs\,(\cos\alpha + \imath \operatorname{sen}\,\alpha)\,(\cos\beta + \imath \operatorname{sen}\,\beta) = \\ &= rs\,((\cos\alpha\cos\beta - \operatorname{sen}\,\alpha \operatorname{sen}\,\beta) + \imath\,(\cos\alpha \operatorname{sen}\,\beta + \operatorname{sen}\,\alpha\cos\beta)) \\ &= rs\,(\cos(\alpha+\beta) + \imath \operatorname{sen}\,(\alpha+\beta))\,. \end{aligned}$$

Sigue entonces del lema 8.2.9 que

$$\arg(zw) \equiv \alpha + \beta = \arg(z) + \arg(w),$$

como queríamos demostrar.

2) Escribiendo $z = (z/w)w$ y usando 1), tenemos:

$$\arg(z) \equiv \arg(z/w) + \arg(w),$$

de donde

$$\arg(z/w) \equiv \arg(z) - \arg(w).$$

3) Puesto que $u^{-1} = \overline{u}/\,|\,u\,|^2$, sigue aplicando 2) que

$$\arg(u^{-1}) \equiv \arg(\overline{u}) - \arg\left(|\,u\,|^2\right) = \arg(\overline{u}),$$

ya que $|\,u\,|^2$ es un número real positivo. Finalmente, usando nuevamente 2) resulta que

$$\arg(u^{-1}) = \arg(1/u) \equiv \arg(1) - \arg(u) = -\arg(u)\,,$$

como queríamos probar. $\Diamond$

INTERPRETACION GEOMETRICA Ahora podemos interpretar gráficamente el producto de números complejos. En efecto, si $z, w \in \mathbb{C}^*$ y $\arg(w)$ es la medida en radianes de un ángulo de β grados sexagesimales, la fórmula 1) de 8.2.10 nos muestra que la dirección del vector zw se obtiene rotando β grados (en sentido antihorario) el vector z. Conocida su dirección, su longitud $|\,z\,|\,|\,w\,|$ lo determina completamente. Por ejemplo, la multiplicación por $\imath$ rota cada vector 90° sin cambiar su longitud.

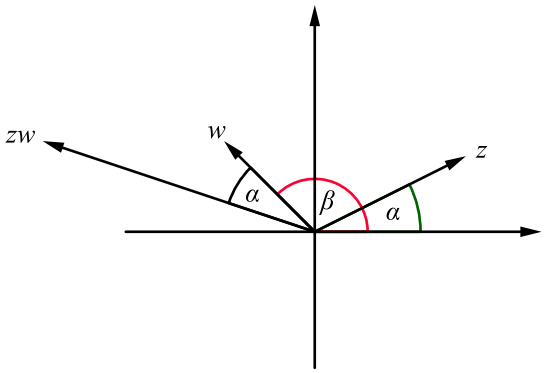


Figura 8.9: Representación gráfica del producto

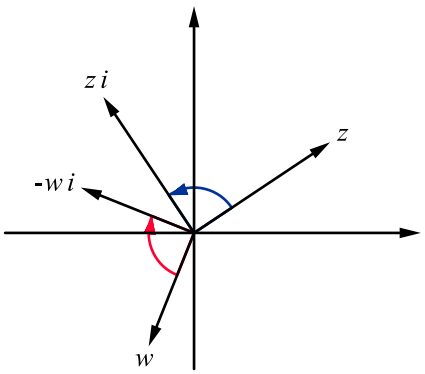


Figura 8.10: Rotaciones de 90°

Una interpretación similar a la del producto cabe para el cociente, sólo que la rotación debe hacerse en el sentido de las agujas del reloj.

Corolario 8.2.11 (De Moivre) Sea $z \in \mathbb{C}^*$ y sea $n \in \mathbb{Z}$. Entonces

$$\arg(z^n) \equiv n \arg(z)\,. \tag{8.10}$$

DEMOSTRACION. Para $n \geq 0$ la fórmula sigue fácilmente por inducción, ya que $\arg(z^0) = \arg(1) = 0$ y

$$\arg(z^{n+1}) = \arg(z^n z) \equiv \arg(z^n) + \arg(z) = (n+1)\arg(z)\,.$$

Usando este hecho y la parte 3) de la proposición 8.2.10 la fórmula también resulta válida para $n < 0$, pues

$$\arg(z^n) = \arg\left((z^{-n})^{-1}\right) \equiv -\arg(z^{-n}) \equiv -(-n)\arg(z) = n\arg(z)\,. \quad \Diamond$$

Ejemplo 8.2.12 Si $z = -1 + \sqrt{3}\,\imath$, hallemos la forma binómica de $w = z^{50}$. Puesto que ya vimos en 8.2.8 que z tiene módulo 2 y argumento $2/3\,\pi$, usando la fórmula de De Moivre resulta entonces que

$$\arg(w) \equiv 100/3\,\pi = 32\pi + 4/3\,\pi \equiv 4/3\,\pi\ (2\pi)\,,$$

y por lo tanto $\arg(w) = 4/3\,\pi$. Puesto que $|\,w\,| = 2^{50}$, tenemos:

$$w = 2^{50} e^{\imath(4/3\,\pi)} = 2^{50}\left(-1/2 - \imath\,\sqrt{3}/2\right) = -2^{49} - 2^{49}\sqrt{3}\,\imath. \quad \Diamond$$

Ejemplo 8.2.13 Determinemos los números complejos z tales que

$$\pi/2 < \arg(z^3) < \pi\,.$$

Si $\arg(z) = x$, sigue por De Moivre que $\arg(z^3) = 3x - 2k\pi$, para algún entero $k \geq 0$. Siendo $3x < 6\pi$ resulta que $k < 3$, lo que nos deja tres casos a considerar:

i) $x < 2/3\,\pi$, en cuyo caso $3x < 2\pi$ y y por lo tanto $k = 0$. Deben verificarse entonces las desigualdades $\pi/2 < 3x < \pi$, lo que nos muestra que en el primer caso la solución es

$$\pi/6 < x < \pi/3\,.$$

ii) $2/3\,\pi \leq x < 4/3\,\pi$ y por lo tanto $2\pi \leq 3x < 4\pi$. Luego $k = 1$ y las condiciones son ahora $\pi/2 < 3x - 2\pi < \pi$. Operando, obtenemos la solución

$$5/6\,\pi < x < \pi\,.$$

iii) Finalmente, consideremos el caso $4/3\,\pi \leq x < 2\pi$. Similarmente a los casos anteriores, arribamos a las desigualdades $\pi/2 < 3x - 4\pi < \pi$, equivalentes a la solución

$$3/2\,\pi < x < 5/3\,\pi\,.$$

Resumiendo, los números complejos que satisfacen las condiciones dadas se distribuyen en 3 regiones del plano, delimitada cada una de ellas por dos semirrectas, cuyas inclinaciones son de 30° y 60° en la primera región, de 150° y 180° en la segunda y de 270° y 300° en la última. Visualmente, forman una figura que semeja un molino de 3 aspas.

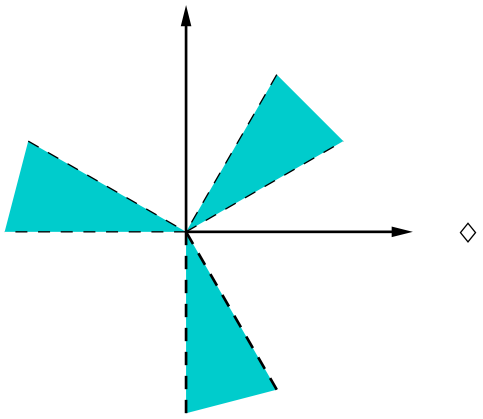


Figura 8.11: Vista gráfica de las soluciones

8.2.3. Ejercicios

1. Sean u y v vectores no colineales del primer cuadrante del plano, representando números complejos z y w. Respresentar gráficamente los vectores correspondientes a:

   a) $tz\,(t \in \mathbb{R})$
   b) $\overline{w - z}$
   c) z^2
   d) $|\,z\,|$
   e) z/w
   f) $|\,z - w\,|$.

2. En cada uno de los siguientes casos, determinar los $z \in \mathbb{C}$ que satisfacen la condición dada:

   a) $z = -\overline{z}$
   b) $\overline{z} = \imath z$
   c) $z = \imath\overline{z}$
   d) $z^2 \in \mathbb{R}$
   e) $z^2 = \overline{z}$
   f) $|\,z\,| = z$
   g) $|\,z\,| = \imath z$
   h) $|\,z + 1\,| = |\,z\,| + 1$
   i) $z^{-1} = \overline{z}$.

3. Sean $z, w \in \mathbb{C}$.

   a) Probar que $|z+w|^2 + |z-w|^2 = 2\left(|z|^2 + |w|^2\right)$. Interpretar geométricamente.

   b) Si z y w son no nulos, probar que $|z+w| = |z| + |w|$ si y sólo si existe un número real positivo λ tal que $w = \lambda z$.

4. *a*) Hallar las raíces cuadradas de $-25\imath$ y de $7 + 24\imath$

   b) Hallar las raíces cuartas de $(3+4\imath)^2$

   c) Resolver las ecuaciones

   i) $x^2 - (1+2\imath)x - (1-\imath) = 0$

   ii) $x^3 - 5x^2 + 17x = 13$.

5. Sean $z, w \in \mathbb{C}^*$ tales que $\arg(z) = \alpha$ y $\arg(w) = \beta$. Probar que

$$\arg(zw) = \begin{cases} \alpha + \beta & \text{si } \alpha + \beta < 2\pi \\ \alpha + \beta - 2\pi & \text{si } \alpha + \beta \geq 2\pi\,. \end{cases}$$

   Determinar fórmulas análogas para $\arg(z/w)$, $\arg(\overline{z})$ y $\arg(z^{-1})$.

6. Hallar módulo y argumento principal de los siguientes números complejos:

   a) $3/2 + \sqrt{27}/2\,\imath$

   b) $3/2 - \sqrt{3}/2\,\imath$

   c) $\cos 11/5\,\pi - \imath \operatorname{sen} 19/5\,\pi$

   d) $-\cos 8/3\,\pi + \imath \operatorname{sen} 8/3\,\pi$

   e) $\operatorname{sen} 3/4\,\pi + \imath \cos 3/4\,\pi$

   f) $\cos 55/3\,\pi - \operatorname{sen} 56/3\,\pi$.

7. Representar gráficamente los siguientes subconjuntos de $\mathbb{C}$:

   a) $\{z \in \mathbb{C} : |z-1| < 5 \text{ y } \operatorname{Re}(z) < \operatorname{Im}(z)\}$

   b) $\{z \in \mathbb{C} : z + \overline{z} = z\overline{z}\}$

   c) $\{z \in \mathbb{C} : |\overline{z} + 1 - \imath| > |z|\}$

   d) $\left\{z \in \mathbb{C} : |z| \leq 2 \text{ y } \arg\left(\imath z^4\right) < \pi/2\right\}$.

8. Hallar la forma binómica de z en los siguientes casos:

$$z = \left(-\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}\,\imath\right)^{56} \;;\; z = \left(\frac{3\sqrt{3} + 9\imath}{4 + 4\imath}\right)^{40} \;;\; z = \left(\frac{1}{2} - \frac{\sqrt{3}}{2}\,\imath\right)^{25}.$$

8.3. Radicación compleja

8.3.1. Raíces enésimas

En lo que sigue n denotará un número natural cualquiera. Habiendo ya establecido que la ecuación $X^2 = z$ admite dos soluciones complejas cualquiera sea $z \in \mathbb{C}^*$, demostraremos ahora que algo similar ocurre con la ecuación $X^n = z$. Precisamente, probaremos que todo complejo z no nulo admite n raíces n-ésimas en $\mathbb{C}$ (observemos que 0 es la única raíz n-ésima de 0, ya que $w^n = 0$ si y sólo si $w = 0$).

Antes de encarar esta tarea, recordemos que dado un número real $c \geq 0$ el símbolo $\sqrt[n]{c}$ designa el único número real $b \geq 0$ tal que $b^n = c$. Como en el caso de la raíz cuadrada, vale la pena remarcar que sólo usaremos dicho símbolo en la situación que acabamos de referir.

Teorema 8.3.1 Si $z \in \mathbb{C}^*$, z admite exactamente n raíces n-ésimas en $\mathbb{C}$.

DEMOSTRACION. Supongamos que $w^n = z$ y sean

$$\begin{aligned} z &= |\, z \,|\, (\cos\alpha + \imath \text{ sen } \alpha) \qquad \text{y} \\ w &= |\, w \,|\, (\cos\beta + \imath \text{ sen } \beta) \end{aligned}$$

las formas polares de z y de w. Igualando primero módulos en la igualdad $w^n = z$, deducimos de la multiplicatividad del módulo que $|\, w \,|^n = |\, z \,|$ y por lo tanto $|\, w \,| = \sqrt[n]{|\, z \,|}$.

Considerando ahora los argumentos, sigue de la fórmula de De Moivre que

$$\alpha = \arg(z) = \arg(w^n) \equiv n \arg(w) = n\beta,$$

donde $\equiv$ indica congruencia módulo 2π. Por lo tanto, existe $s \in \mathbb{Z}$ tal que $\alpha = n\beta - 2s\pi$, ó equivalentemente,

$$\beta = \frac{\alpha + 2s\pi}{n}\,.$$

Recíprocamente, es fácil probar que $\sqrt[n]{|\, z \,|}\, e^{\imath\left(\frac{\alpha+2s\pi}{n}\right)}$ es una raíz n-ésima de z *cualquiera sea* $s \in \mathbb{Z}$, pues $\left(\sqrt[n]{|\, z \,|}\right)^n = |\, z \,|$ y $n\left(\frac{\alpha+2s\pi}{n}\right) = \alpha + 2s\pi \equiv \alpha$. En consecuencia, w es una raíz n-ésima de z si y sólo si w es de la forma

$$w = \sqrt[n]{|\, z \,|}\left(\cos\left(\frac{\alpha + 2k\pi}{n}\right) + \imath \text{ sen}\left(\frac{\alpha + 2k\pi}{n}\right)\right), \tag{8.11}$$

donde k es un entero cualquiera (designaremos por w_k el número determinado por la expresión de arriba).

La fórmula 8.11 estaría indicando que z admite infinitas raíces n-ésimas, pues k varía libremente sobre $\mathbb{Z}$, pero veremos que no es así, ya que la correspondencia $k \mapsto w_k$ no es inyectiva.

Para ver esto último, dado $k \in \mathbb{Z}$ designemos por q y r el cociente y el resto de dividir k por n, respectivamente. Resulta entonces que

$$\frac{\alpha + 2k\pi}{n} = \frac{\alpha + 2(qn + r)\pi}{n} = \frac{\alpha + 2r\pi}{n} + 2q\pi \equiv \frac{\alpha + 2r\pi}{n}\,,$$

y por lo tanto $w_k = w_r$. Puesto que $0 \leq r < n$, deducimos que z admite a lo sumo n raíces n-ésimas. Por otro lado, dados i, j tales que $0 \leq i < j < n$, tenemos que

$$0 < \frac{\alpha + 2j\pi}{n} - \frac{\alpha + 2i\pi}{n} = \frac{2(j-i)\pi}{n} < 2\pi\,,$$

y por lo tanto $\dfrac{\alpha + 2i\pi}{n} \not\equiv \dfrac{\alpha + 2j\pi}{n}\ (2\pi)$. Esto implica que $w_i \neq w_j$.

En conclusión, z admite exactamente n raíces n-ésimas $w_0, w_1, \ldots, w_{n-1}$ en $\mathbb{C}$, dadas por la fórmula:

Fórmula 8.3.2

$$w_k = \sqrt[n]{|\,z\,|}\left(\cos\left(\frac{\alpha + 2k\pi}{n}\right) + \imath\ \mathrm{sen}\left(\frac{\alpha + 2k\pi}{n}\right)\right) \quad (0 \leq k < n)$$

Ejemplo 8.3.3 Hallemos la forma binómica de las raíces cúbicas de $z = 8\imath$. Siendo en este caso $|\,z\,| = 8$ y $\alpha = \arg(z) = \pi/2$, aplicando la fórmula 8.3.2 resulta que las raíces cúbicas de $8\imath$ son de la forma

$$w_k = 2\,e^{\imath\left(\frac{\pi}{6} + \frac{2k\pi}{3}\right)},$$

con $0 \leq k \leq 2$. Resolviendo, obtenemos:

$$\begin{aligned} w_0 &= 2\,(\cos(\pi/6) + \imath\ \mathrm{sen}(\pi/6)) = \sqrt{3} + \imath \\ w_1 &= 2\,(\cos(5/6\,\pi) + \imath\ \mathrm{sen}(5/6\,\pi)) = -\sqrt{3} + \imath \\ w_2 &= 2\,(\cos(3/2\,\pi) + \imath\ \mathrm{sen}(3/2\,\pi)) = -2\imath\,. \quad \Diamond \end{aligned}$$

NOTA. Las raíces n-ésimas de un número complejo no nulo z se disponen regularmente en el plano complejo. En efecto, en primer lugar todas tienen el mismo módulo, y por otro lado, por simple inspección de la fórmula 8.3.2 observamos que

$$\arg(w_{k+1}) - \arg(w_k) = \frac{\alpha + 2(k+1)\pi}{n} - \frac{\alpha + 2k\pi}{n} = \frac{2\pi}{n}\,,$$

esto es, la medida del ángulo formado por los vectores w_k y w_{k+1} es constante.

A través de consideraciones geométricas elementales deducimos entonces que las raíces n-ésimas de z *son vértices de un polígono regular* de n lados inscripto en la circunferencia de radio $\sqrt[n]{|\,z\,|}$ y centro en el origen. Por ejemplo, las raíces cúbicas de $8\imath$ son los vértices de un triángulo equilátero inscripto en la circunferencia de radio 2 y centro en el origen.

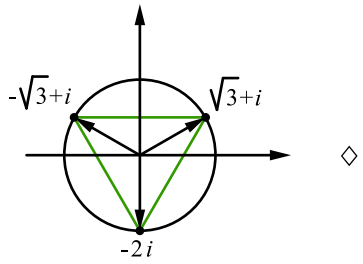


Figura 8.12: Raíces cúbicas de $8\imath$

Raíces de la unidad.

Dado $n \in \mathbb{N}$ consideraremos especialmente el conjunto $\{u_0, u_1, \ldots, u_{n-1}\}$ de raíces n-ésimas de 1, que notaremos G_n. Puesto que $|\,1\,| = 1$ y $\arg(1) = 0$, aplicando la fórmula general resulta en este caso que

$$u_k = \cos\frac{2k\pi}{n} + \imath \text{ sen } \frac{2k\pi}{n} = e^{\frac{2k\pi\imath}{n}} \quad (0 \leq k < n)\,. \tag{8.12}$$

Notemos que $u_0 = 1$ y que toda raíz de la unidad tiene módulo 1.
Por ejemplo, es inmediato verificar que para $n \leq 4$ se tiene

$$\begin{aligned} G_1 &= \{\,1\} \\ G_2 &= \{\,1\,,\,-1\} \\ G_3 &= \left\{\,1\,,\,-1/2 + \sqrt{3}/2\,\imath\,,\,-1/2 - \sqrt{3}/2\,\imath\right\} \\ G_4 &= \{\,1\,,\,\imath\,,\,-1\,,\,-\imath\}\,. \end{aligned}$$

Como casos particulares de una situación general ya mencionada, observemos que las raíces cúbicas (cuartas) de 1 son los vértices de un triángulo equilátero (de un cuadrado) inscripto en la circunferencia de radio 1 y centro en el origen.

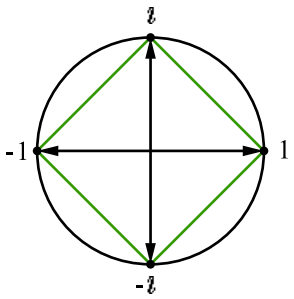


Figura 8.13: Raíces cuartas de la unidad

El siguiente lema revelará un rol importante de las raíces de la unidad.

Lema 8.3.4 Sea $z \in \mathbb{C}^*$ y sea $w \in \mathbb{C}$ tal que $w^n = z$. Entonces

$$w\,G_n = \{w\,u_0\,,\, w\,u_1\,\ldots\,,\, w\,u_{n-1}\}$$

es el conjunto de raíces n-ésimas de z.

DEMOSTRACION. Notemos el significado del enunciado: si conocemos una raíz n-ésima cualquiera de z, multiplicándola por los elementos de G_n obtenemos todas las raíces n-ésimas de z. Por ejemplo, las raíces cuartas de 16 son 2, $2\imath$, -2 y $-2\imath$, pues 2 es una raíz cuarta de 16.

En cuanto a la demostración, sigue fácilmente que $w\,G_n$ está contenido en el conjunto A de raíces n-ésimas de z, ya que $(w\,u_i)^n = w^n u_i^n = w^n = z$ para todo i. Puesto que por otra parte ambos conjuntos tienen n elementos, pues $w\,u_i = w\,u_j$ si y solo si $u_i = u_j$ (por ser $w \neq 0$), concluimos que $wG_n = A$, como queríamos demostrar. $\Diamond$

Destacaremos a continuación algunas propiedades básicas de las raíces de la unidad (z y w designan números complejos y m y n números naturales).

Proposición 8.3.5 Son válidas las siguientes propiedades:

1) Si $z, w \in G_n$ entonces $zw \in G_n$

2) Si $z \in G_n$ entonces $z^{-1} \in G_n$

3) $G_m \cap G_n = G_{(m:n)}$

4) $G_m \subseteq G_n$ si y sólo si $m \mid n\,.$

DEMOSTRACION. Las propiedades 1) y 2) siguen directamente de la definición de G_n, ya que $(zw)^n = z^n w^n = 1\,.\,1 = 1$ y $(z^{-1})^n = (z^n)^{-1} = 1^{-1} = 1\,.$

Para probar 3), sea $d = (m : n)$ y sea $u \in G_d$. Escribiendo $m = dr$ y $n = ds$ resulta que $u^m = \left(u^d\right)^r = 1^r = 1$ y $u^n = \left(u^d\right)^s = 1^s = 1$. Luego $u \in G_m \cap G_n$.

Recíprocamente, sea $v \in G_m \cap G_n$ y sean $h, k \in \mathbb{Z}$ tales que $d = mh + nk$. Entonces

$$v^d = v^{mh+nk} = (v^m)^h\,(v^n)^k = 1\,.\,1 = 1\,,$$

y por lo tanto $v \in G_d$. En consecuencia $G_d = G_m \cap G_n\,.$

El ítem 4) se deduce inmediatamente de 3). $\Diamond$

NOTA. La propiedad 1) afirma que la restricción a G_n del producto de números complejos define una operación binaria en G_n, que obviamente es asociativa y conmutativa. Puesto que $1 \in G_n$, dicha operación admite un elemento neutro, y además, todo elemento tiene inverso, por 2). Como sabemos, un conjunto G dotado de una operación binaria asociativa, con elemento neutro, y en el que todo elemento tiene inverso respecto a dicha operación se dice un *grupo*. Si por añadidura la operación es conmutativa se dice que G

es un *grupo conmutativo o abeliano*. Por lo tanto, G_n es un grupo abeliano de n elementos.

Notemos también que $\overline{z} \in G_n$ si $z \in G_n$, ya que $|\, z\,| = 1$ y en tal caso $\overline{z} = z^{-1}$.

Por último, deducimos de 4) que cualquier raíz de la unidad pertenece a G_k para infinitos valores de k. Por ejemplo, $-1 \in G_k$ si y sólo si k es par. Análogamente, $\imath \in G_k$ si y sólo si k es múltiplo de 4. $\Diamond$

RAICES PRIMITIVAS Dado cualquiera de los elementos u_k de G_n, observemos que

$$\arg(u_k) = 2k\pi/n = k(2\pi/n) = k\arg(u_1),$$

de donde concluimos que $u_k = u_1^k$. Vale decir, toda raíz n-ésima de 1 es una potencia de u_1. Este último hecho tiene una clara interpretación gráfica, ya que calcular las sucesivas potencias de u_1 corresponde geométricamente a rotar $n-1$ veces el vector u_1, siempre con un ángulo de giro de $2\pi/n$ radianes. Puesto que éste es el ángulo determinado por dos raices n-ésimas consecutivas de 1, resulta que al tomar las sucesivas potencias de u_1 barremos todos los elementos de G_n.

Lo anterior sugiere la siguiente definición:

> Diremos que $w \in G_n$ es una raíz n-ésima *primitiva* de 1 si y sólo si para cada $z \in G_n$ existe un entero j $(0 \le j < n)$ tal que $z = w^j$. Alternativamente, diremos que w es una raíz primitiva de *orden* n ó que w es un *generador* de G_n.

Así, $e^{2\pi\imath/n}$ es una raíz n-ésima primitiva de 1 cualquiera sea $n \in \mathbb{N}$. Como casos particulares resulta que -1 es una raíz primitiva de orden 2 y que la unidad imaginaria es una raíz primitiva de orden 4. Notemos que $\imath \in G_8$ pero no es una raíz octava primitiva de 1, ya que obviamente sus potencias no recorren todos los elementos de G_8. El siguiente resultado explicará este hecho y nos brindará otras formas equivalentes de la definición anterior.

Proposición 8.3.6 Si $w \in G_n$, las siguientes afirmaciones son equivalentes:

a) w es una raíz n-ésima primitiva de 1

b) $n = \min\left\{j \in \mathbb{N} \,:\, w^j = 1\right\}$

c) Si $j \in \mathbb{Z}$, $w^j = 1 \Leftrightarrow n \mid j$

d) $w = u_k$ con k y n coprimos.

DEMOSTRACION.

a) $\Rightarrow$ *b*) Por hipótesis, la aplicación

$$f : \{\,0,\, 1, \ldots,\, n-1\} \to G_n$$

definida por $f(i) = w^i$ es suryectiva. Puesto que ambos conjuntos tienen n elementos, resulta que la misma también es inyectiva. En particular $w^i \neq 1$ si $0 < i < n$, ya que $1 = w^0 = f(0)$. Esto claramente prueba b).

$b) \Rightarrow c)$ Si $n \mid j$ sigue por la proposición 8.3.5 que $G_n \subseteq G_j$, y por lo tanto $w^j = 1$. Recíprocamente, supongamos que $w^j = 1$ y sea $r = r_n(j)$. Escribiendo $j = qn + r$ resulta entonces que

$$1 = w^j = w^{qn+r} = (w^n)^q\, w^r = w^r,$$

y siendo $r < n$ la única posibilidad es $r = 0$, por b). Luego $n \mid j$.

$c) \Rightarrow d)$ Si $d = (k : n)$ tenemos

$$\arg(w^{n/d}) = \arg(u_k^{n/d}) \equiv (n/d)(2k\pi/n) = 2(k/d)\pi \equiv 0\ (2\pi),$$

esto es, $w^{n/d} = 1$. Sigue entonces por hipótesis que n/d es múltiplo de n, resultando en particular que $n \leq n/d$. Esto obviamente implica que $d = 1$.

$d) \Rightarrow a)$ Bastará probar que las n potencias $u_k^0, u_k^1, \ldots, u_k^{n-1}$ son distintas, ya que en tal caso la función f es inyectiva y por lo tanto suryectiva. Razonando por el absurdo, supongamos que existen exponentes i, j $(0 \leq i < j < n)$ tales que $u_k^i = u_k^j$. Examinando los argumentos, resulta que

$$i(2k\pi/n) \equiv j(2k\pi/n)\ (2\pi),$$

de donde sigue que existe $s \in \mathbb{Z}$ tal que

$$i(2k\pi/n) = j(2k\pi/n) + 2s\pi.$$

Operando convenientemente obtenemos $ki = kj + sn$, ó equivalentemente, $ki \equiv kj\ (n)$. Cancelando k (recordemos que k y n son coprimos) arribamos a la contradicción requerida, ya que la relación $i \equiv j\ (n)$ no es posible en el rango de variación de i y j. $\quad \Diamond$

NOTA. Observemos que la caracterización que brinda d) muestra que existen exactamente $\varphi(n)$ raíces $n-$ésimas primitivas de la unidad.

Por otra parte, toda raíz de la unidad es una raíz primitiva de algún orden conveniente. Precisamente, conservando las notaciones de 8.3.6 la condición d) se generaliza de la siguiente forma: $w = u_k$ es una raíz primitiva de orden n/d, donde $d = (k : n)$. En efecto, sea $m = n/d$ y sea $q = k/d$. Entonces

$$\arg(w) = 2k\pi/n = 2qd\pi/md = 2q\pi/m\,,$$

lo que prueba que $w \in G_m$. Puesto que además q y m son coprimos sigue de 8.3.6 que w es una raíz m-ésima primitiva de 1. $\quad \Diamond$

Ejemplo 8.3.7 Tomemos $n = 12$. De acuerdo con los resultados de arriba, podemos clasificar de la siguiente manera los elementos de G_{12}:

- $u_0 = 1$ es primitiva de orden 1
- $u_6 = -1$ es primitiva de orden 2
- u_4 y u_8 son primitivas de orden 3
- $u_3 = \imath$ y $u_9 = -\imath$ son primitivas de orden 4
- u_2 y u_{10} son primitivas de orden 6
- u_1, u_5, u_7 y u_{11} son las raíces duodécimas primitivas de 1.

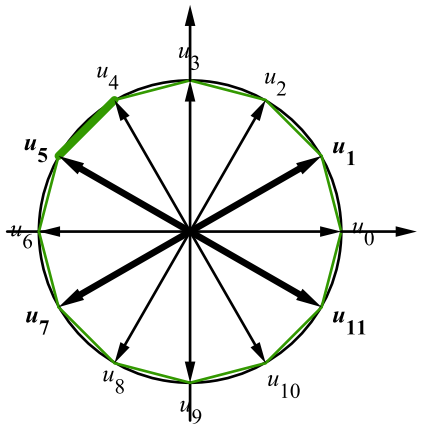


Figura 8.14: Vista gráfica de G_{12} ◇

Ecuaciones en el cuerpo de números complejos.

Cerraremos el capítulo resolviendo algunas ecuaciones en $\mathbb{C}$. No es nuestro propósito hacer un estudio sistemático de algún determinado tipo de ecuación ni brindar fórmulas generales para su resolución, sino servirnos de algunos casos particulares para ilustrar y aplicar los temas desarrollados. Comencemos pues con nuestra serie de ejemplos.

I Resolvamos la ecuación cuadrática

$$z^2 - 3z + 3 + \imath = 0\,.$$

Si bien la tarea es sencilla, queremos establecer claramente que puede usarse para resolverla la misma fórmula que desarrollamos en el caso real, y que se deduce de idéntica manera en el caso complejo. La diferencia es que toda ecuación cuadrática es resoluble en el campo complejo, pues todo número complejo admite raíces cuadradas en $\mathbb{C}$.

En este caso el discriminante es $9 - 4(3 + \imath) = -3 - 4\imath$, cuyas raíces cuadradas son $1 - 2\imath$ y $-1 + 2\imath$. Aplicando la fórmula y efectuando las

operaciones resulta que las soluciones son $z_1 = 2 - \imath$ y $z_2 = 1 + \imath$, como el lector puede verificar. □

II Consideremos la ecuación *bicuadrática*

$$z^4 + (9 - 2\imath)z^2 - 18\imath = 0\,,$$

una ecuación de grado 4 cuyos coeficientes de grado impar son nulos. Realizando el cambio de variable $w = z^2$ deberemos resolver entonces la ecuación cuadrática $w^2 + (9 - 2\imath)w - 18\imath = 0$.

Procediendo como en I, vemos que las soluciones de la misma son $w_1 = 2\imath$ y $w_2 = -9$. Puesto que $z^2 = w$, calculamos ahora las raíces cuadradas de w_1 y w_2, que nos darán las cuatro soluciones que buscamos. Aplicando la fórmula 8.2.3 resulta que ellas son $z_1 = 1 + \imath$, $z_2 = -1 - \imath$, $z_3 = 3\imath$ y $z_4 = -3\imath$. □

III Vayamos ahora a la ecuación

$$\left(\frac{4z - 1}{4}\right)^6 = z^3.$$

Se trata de una ecuación de grado 6, que en principio parece complicada de resolver. Sin embargo, su peculiar forma nos facilita la cuestión. En efecto, dividiendo ambos miembros por z^3 (previa verificación de que $z = 0$ no es solución) y operando obtenemos la ecuación equivalente

$$\left(\frac{(4z - 1)^2}{16z}\right)^3 = 1,$$

esto es, z es solución si y sólo si $(4z-1)^2/16z \in G_3$. Luego, deberemos hallar, para cada $u \in G_3$, las soluciones de la ecuación $(4z-1)^2 = 16zu$, que desarrollada resulta ser la ecuación cuadrática

$$16z^2 - 8(1 + 2u)z + 1 = 0, \tag{8.13}$$

cuyo discriminante es

$$\Delta = 64(1 + 2u)^2 - 64 = 256(u + u^2).$$

Supongamos en primer término que $u \neq 1$. En tal caso, aplicando la fórmula de la suma de los términos de una progresión geométrica tenemos:

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

Por lo tanto $\Delta = -256$ y sus raíces cuadradas son $16\imath$ y $-16\imath$. Aplicando la fórmula de la ecuación cuadrática y teniendo en cuenta que

$2u = -1 + \sqrt{3}\,\imath$ ó $2u = -1 - \sqrt{3}\,\imath$ resulta que las cuatro soluciones correspondientes a los dos valores de u distintos de 1 son

$$\pm\left(\frac{2+\sqrt{3}}{4}\right)\imath \quad \text{y} \quad \pm\left(\frac{2-\sqrt{3}}{4}\right)\imath\,.$$

Si $u = 1$ tenemos $\Delta = 512$, resultando que las dos soluciones de (8.13) correspondientes a este caso son reales, a saber:

$$\frac{3+2\sqrt{2}}{4} \quad \text{y} \quad \frac{3-2\sqrt{2}}{4}\,.$$

Encargamos al lector la tarea de verificar todos los cálculos. □

IV Veamos cómo resolver la ecuación

$$1 + z^3 + z^6 + z^9 = 0. \tag{8.14}$$

Si $x \in \mathbb{C}$, es inmediato demostrar, a partir de las propiedades de la suma y el producto de números complejos, la validez de la relación

$$x^4 - 1 = (x-1)(1 + x + x^2 + x^3).$$

Entonces, tomando $x = z^3$ resulta que todo número complejo z satisface la igualdad

$$z^{12} - 1 = (z^3 - 1)(1 + z^3 + z^6 + z^9). \tag{8.15}$$

Sigue de ello que toda solución z de nuestra ecuación es un elemento de G_{12}, ya que en tal caso $z^{12} - 1 = 0$. Notemos de paso que ningún elemento z de G_3 es solución de (8.14), ya que $1 + z^3 + z^6 + z^9 = 4$ si $z^3 = 1$.

Recíprocamente, supongamos que $z \in G_{12}$ y $z \notin G_3$. Tenemos entonces que el miembro de la izquierda de (8.15) es cero, por lo que alguno de los dos factores del miembro de la derecha debe ser nulo. Puesto que el primero no lo es, concluimos que z satisface (8.14).

En definitiva, las soluciones de $1 + z^3 + z^6 + z^9 = 0$ son exactamente los elementos del conjunto $G_{12} - G_3$. □

V Consideremos finalmente la ecuación

$$z^{10} = \imath\overline{z}^{\,2}. \tag{8.16}$$

Es claro que $z = 0$ es solución, por lo que supondremos en adelante que z es una solución no nula. Tomando módulos tenemos:

$$|\,z\,|^{10} = \left|\,z^{10}\,\right| = \left|\,\imath\,\overline{z}^{\,2}\right| = |\,\imath\,|\,|\,\overline{z}\,|^2 = |\,z\,|^2\,.$$

Deducimos entonces que $|z| = 1$, pues $|z|$ es un número real positivo y $|z|^8 = 1$. Luego, multiplicando ambos miembros de (8.16) por z^2 obtenemos:

$$z^{12} = \imath\,(\overline{z}z)^2 = \imath\,|z|^4 = \imath\,,$$

vale decir, z es una raíz duodécima de $\imath$. Recíprocamente, supongamos que $z^{12} = \imath$. Entonces

$$z^{10} = z^{12}(z^{-1})^2 = \imath\,\overline{z}^{\,2}\,,$$

pues $w^{-1} = \overline{w}$ para todo $w \in \mathbb{C}$ de módulo 1.

En consecuencia, z es solución de (8.16) si y sólo si $z = 0$ ó z es una raíz duodécima de $\imath$ (13 soluciones en total). □

8.3.2. Ejercicios

1. En cada uno de los siguientes casos hallar la forma binómica de las raíces $m-$ésimas de z:

   a) $z = 27\,,\ m = 6$

   b) $z = -5\,,\ m = 4$

   c) $z = -8\imath\,,\ m = 6$

   d) $z = (3 + 4\imath)^2\,,\ m = 8\,.$

2. Determinar los $z \in \mathbb{C}$ tales que

   a) $z^3 = \overline{z}^{\,5}$

   b) $\overline{z}^{\,4} = -\imath z^7$

   c) $z^9 = \overline{z}^{\,9}$

   d) $z^4 = (3 + 3\imath)^6$

   e) $(z + 1)^4 = (\overline{z} - \imath)^4$

   f) $z^{10} - 4z^5 + 13 = 0\,.$

3. Si $n \in \mathbb{N}$, determinar las regla de multiplicación de los elementos u_i de G_n (por ejemplo, $u_8u_{10} = u_6$ en G_{12}). Exhibir la tabla completa de multiplicación de $G_6\,.$

4. Sea $n \in \mathbb{N}$. Calcular:

   a) $\displaystyle\sum_{w \in G_n} w$

   b) $\displaystyle\prod_{w \in G_n} w$

c) $\sum_{i=0}^{n-1} z^i \ (z \in G_n)$.

5. *a*) Si $n \in \mathbb{N}$ y $z \in G_n$, probar que $(1+z)^n \in \mathbb{R}$
   b) Calcular $w^{22} + 2w^{15} + 3w^9 + 3\bar{w}^{\,4} + 3w^3 + 2w^{-3} + 6$ para cada $w \in G_5$
   c) Probar que $\operatorname{Re}\big((z^{44}-1)(z^{19}+1)\big) = 0$ para todo $z \in G_7$.

6. Sea $n \in \mathbb{N}$ y sea w una raíz primitiva de orden n. Probar:
   a) w^k es raíz primitiva de orden n si y sólo si $(k:n) = 1 \, (k \in \mathbb{Z})$
   b) w^{-1} es raíz primitiva de orden n
   c) $-w$ es raíz primitiva de orden $2n$ si n es impar.

7. Sea u una raíz primitiva de orden 15.
   a) Caracterizar los $m \in \mathbb{N}$ tales que $u^{6m} = \overline{u}^{\,51}$
   b) Demostrar que $1 + u^5 + u^{10} = 1 + u^3 + u^6 + u^9 + u^{12} = 0$
   c) Calcular la suma de las raíces primitivas de orden 15.

Capítulo 9

Polinomios

9.1. El anillo de polinomios

9.1.1. Introducción

Si $n \in \mathbb{N}$, se denomina *ecuación algebraica* de grado n a toda ecuación del tipo

$$a_n X^n + a_{n-1} X^{n-1} + \cdots + a_1 X + a_0 = 0, \tag{9.1}$$

donde $a_0, a_1, \ldots, a_n$ son números complejos y $a_n \neq 0$.

En su máxima generalidad, se trata de hallar todos los $z \in \mathbb{C}$ tales que

$$\sum_{i=0}^{n} a_i z^i = 0\,,$$

aunque en ocasiones, dependiendo de la naturaleza del problema que origina la ecuación, es posible que sólo nos interese un cierto tipo específico de soluciones. Por ejemplo, si los coeficientes a_i fueran números reales podríamos concentrarnos sólo en la búsqueda de las soluciones reales, o en las soluciones racionales en el caso de que los coeficientes fueran números enteros.

Sin duda el lector está familiarizado con dos casos sencillos del problema general de resolución de ecuaciones algebraicas de grado arbitrario, como lo son la ecuación de primer grado $aX + b = 0$, cuya única solución se obtiene "despejando" directamente X, y la ecuación cuadrática $aX^2 + bX + c$, cuyas soluciones pueden hallarse con facilidad a través de una fórmula que hemos exhibido en estas páginas.

Sin entrar en detalles que están fuera del alcance de este libro, señalemos que la complejidad de la cuestión aumenta notablemente a medida que crece n. Ya en el siglo XVI fueron establecidas fórmulas similares a la del caso cuadrático para resolver las ecuaciones de grado 3 y 4 (en el sentido de que involucran operaciones racionales entre los coeficientes y el cálculo de algunos radicales), aunque bastante más complicadas, y luego de siglos de esfuerzos dirigidos a hallar fórmulas generales, se probó a principios del siglo XIX que

no existen fórmulas de tal tipo que permitan resolver la ecuación general de grado n si $n > 4$. Este impactante resultado, debido a Ruffini y Abel, no debe interpretarse como una imposibilidad práctica de hallar las soluciones de una ecuación algebraica, ya que existen métodos analíticos sencillos que permiten aproximarlas con el grado de precisión que se requiera.

Motivados por lo anterior, introduciremos en este capítulo la noción de polinomio, objeto que podríamos describir informalmente como cualquiera de las expresiones que aparecen en el miembro de la izquierda de una ecuación algebraica, a los que en principio apartaremos de su contexto original, pensándolos como entes matemáticos abstractos. Dotaremos entonces de estructura algebraica al conjunto de tales objetos, lo que nos permitirá desarrollar una "aritmética" de polinomios, similar a la desarrollada en $\mathbb{Z}$, para volver luego con mejores herramientas al problema inicial de resolución de ecuaciones algebraicas. Aclaremos de todos modos que ésta no es la única motivación posible para introducir la teoría de polinomios, ya que éstos son muy importantes en otros campos de la Matemática, como por ejemplo el análisis y el álgebra lineal, lo que hace necesario que cualquier curso básico de álgebra contenga una exposición detallada de la misma.

9.1.2. Definiciones

Salvo mención expresa en otro sentido, en lo que sigue la letra A designará el anillo $\mathbb{Z}$ de números enteros o bien cualquiera de los cuerpos numéricos $\mathbb{Q}$, $\mathbb{R}$ ó $\mathbb{C}$.

> Llamaremos *polinomio con coeficientes en* A a toda expresión de la forma
>
> $$f = a_0 + a_1 X + \ldots + a_{n-1} X^{n-1} + a_n X^n = \sum_{i=0}^{n} a_i X^i,$$
>
> donde $n \in \mathbb{N}_0$, X es un símbolo llamado *indeterminada* y los a_i son elementos de A llamados los *coeficientes* de f. Los términos $a_i X^i$ se denominan *monomios*, y a_i se dice el coeficiente de *grado* i de f.

En general, notaremos los polinomios con letras del tipo $f, g, h, \ldots$, y designaremos por $A[X]$ el conjunto de polinomios con coeficientes en A. Es claro entonces que valen las inclusiones

$$\mathbb{Z}[X] \subset \mathbb{Q}[X] \subset \mathbb{R}[X] \subset \mathbb{C}[X]\,.$$

Para fijar las ideas, y convenir de paso cierta flexibilidad en la notación, veamos algunos ejemplos:

Ejemplos 9.1.1 Consideremos los polinomios

$$\begin{aligned} f_1 &= 1 + 5X + 2X^2 + 4X^3 \\ f_2 &= -1 + 2/3X^3 - 2X^4 + (1+\imath)X^6 + X^7 \\ f_3 &= \sqrt{5}X^3 - 4X^2 + 7X \\ f_4 &= 8\pi \\ f_5 &= -12/7X^5\,. \end{aligned}$$

Usando la notación de la definición, observemos que los valores de n son 3, 7, 3, 0 y 5, respectivamente. Es claro que $f_j \in \mathbb{C}[X]$ en todos los casos, $f_j \in \mathbb{R}[X]$ si $j \neq 2$, $f_j \in \mathbb{Q}[X]$ en los casos $j = 1$ y $j = 5$ y f_1 es el único con coeficientes enteros.

Como puede apreciarse, por razones de comodidad se emplean ciertas convenciones de notación: se obvia la escritura de un monomio cuyo coeficiente es nulo, no se escribe un coeficiente si su valor es ± 1 (a menos que sea de grado 0), y la expresión $+(-a_i)X^i$ se escribe simplemente $-a_iX^i$. Por ejemplo, el polinomio $g = 0 + 3X + (-1)X^2 + 0X^3 + 2X^4 + 1X^5$ se escribe más sencillamente $g = 3X - X^2 + 2X^4 + X^5$.

Notemos también (es el caso del polinomio f_4), que todo $a \in A$ determina un elemento de $A[X]$, a saber $f = aX^0$. A un tal polinomio se lo llama *constante*, y suele designarse simplemente por a. Como último comentario acerca de la notación, observemos que en ocasiones expresaremos un polinomio en potencias decrecientes de la indeterminada (como por ejemplo el polinomio f_3). $\Diamond$

IGUALDAD DE POLINOMIOS Sean

$$f = \sum_{i=0}^{m} a_iX^i \quad \text{y} \quad g = \sum_{i=0}^{n} b_iX^i$$

polinomios con coeficientes en A, y supongamos sin pérdida de generalidad que $m \leq n$. Diremos que $f = g$ si y sólo si $a_i = b_i$ para $0 \leq i \leq m$ y $b_i = 0$ para todo $i > m$.

Observemos el significado de la definición anterior: si a la expresión que define un polinomio se le agrega un número arbitrario de coeficientes nulos se obtiene el mismo polinomio. Por ejemplo, sea $f = 1 - X + 2X^3$. Entonces

$$f = 1 - X + 2X^3 + 0X^4 = 1 - X + 2X^3 + 0X^4 + 0X^5 = \ldots$$

NOTA Si bien la definición que hemos brindado al principio del capítulo parece indicar que un polinomio está determinado por la secuencia finita de sus coeficientes, la noción de igualdad que acabamos de dar nos muestra que un polinomio admite en realidad infinitas secuencias de coeficientes. Así, el polinomio del ejemplo anterior admite como "vector" de coeficientes a

cualquiera de las secuencias $(1,-1,0,2)$, $(1,-1,0,2,0)$, $(1,-1,0,2,0,0,0,0)$, etc.

Sin embargo, existe una manera sencilla de salvar esta aparente anomalía de que un polinomio admita diferentes sucesiones de coeficientes, y es la de asignarle a cada polinomio una secuencia infinita $(a_i)_{i\in\mathbb{N}_0}$ de coeficientes, donde $a_i \in A$ y $a_i = 0$ para todo i mayor o igual que un cierto índice r. Por ejemplo, al polinomio anterior le corresponde la secuencia infinita definida por $a_0 = 1$, $a_1 = -1$, $a_2 = 0$, $a_3 = 2$ y $a_i = 0$ para todo $i \geq 4$. Recíprocamente, cualquier secuencia de este tipo corresponde a un elemento de $A[X]$. Por ejemplo, la sucesión

$$(2, 3, 0, 1/2, -1, 0, 0, \ldots, 0, \ldots)$$

es la secuencia de coeficientes del polinomio $2+3X+1/2X^3-X^4$. Como caso particular importante, consideremos el elemento de $A[X]$ correspondiente a la secuencia

$$(0, 0, 0, \ldots, 0, \ldots),$$

esto es, $a_i = 0$ para todo i. Lo llamaremos polinomio *nulo* y lo notaremos 0.

Como último apunte, señalemos que esta definición de polinomio como una secuencia infinita de coeficientes, con "colas" de ceros simplifica por ejemplo la definición de igualdad, ya que conservando las notaciones de arriba sigue que $f = g$ si y sólo si $a_i = b_i$ para todo $i \in \mathbb{N}_0$. En particular, $f \neq 0$ si y sólo existe un índice j tal que $a_j \neq 0$. Aclaremos de todos modos que seguiremos empleando la tradicional representación finita de un polinomio, en la que se acostumbra a escribir hasta el último coeficiente no nulo, digamos a_n, quedando sobrentendido que todos los coeficientes de grado mayor que n son nulos. $\Diamond$

GRADO DE UN POLINOMIO Sea f un polinomio no nulo y sea $(a_i)_{i\in\mathbb{N}_0}$ su secuencia de coeficientes. Definimos entonces el *grado* de f en la forma

$$gr(f) = \max\{i \in \mathbb{N}_0 \,:\, a_i \neq 0\}.$$

Observemos que el grado está bien definido, ya que el conjunto de índices correspondientes a coeficientes no nulos de f es finito (por definición de polinomio) y no vacío, pues $f \neq 0$. Es claro que el grado es un entero no negativo y que admite la siguiente caracterización:

$$gr(f) = m \text{ si y solo si } a_m \neq 0 \text{ y } a_i = 0 \text{ para todo } i > m.$$

Equivalentemente, f puede expresarse en la forma

$$f = a_0 + a_1X + \cdots + a_{m-1}X^{m-1} + a_mX^m,$$

donde a_m (llamado el coeficiente *principal* de f) es distinto de cero. Si $a_m = 1$ diremos que f es *mónico*.

Tenemos por ejemplo:

$$gr(2 + X^2 - 8X^3 + 7X^5) = 5$$
$$gr(1 - X + 7X^4 - \sqrt{8}X^6 + 0X^7) = 6$$
$$gr(-4X^{25}) = 25$$
$$gr(5/2) = 0\,.$$

El último ejemplo es un caso particular de una situación general: los polinomios de grado cero son exactamente los polinomios constantes no nulos (recordemos que no está definido el grado en el caso del polinomio nulo).

Estructura de anillo.

La estructura de anillo de A nos permitirá definir a continuación dos operaciones binarias en $A[X]$ que dotarán a éste de una estructura de anillo conmutativo.

SUMA DE POLINOMIOS Sean $f = \sum_{i=0}^{n} a_i X^i$ y $g = \sum_{i=0}^{n} b_i X^i$ polinomios con coeficientes en A (obsérvese que no hay inconveniente en unificar así las representaciones de ambos). Definimos entonces la *suma* de f y g en la forma

$$f + g = \sum_{i=0}^{n} (a_i + b_i) X^i,$$

esto es, el coeficiente de grado i de $f + g$ es la suma (en A) de los correspondientes coeficientes de grado i.

Por ejemplo, si $f = -2 + 3X - X^2 + 2X^3$, $g = 1 - 2X + 2X^2 + 5X^4 + X^5$ y $h = 1 + X - 2X^2 + X^3 - X^5$, aplicando la definición de arriba resulta que valen las igualdades

$$f + g = -1 + X + X^2 + 2X^3 + 5X^4 + X^5$$
$$g + h = 2 - X + X^3 + 5X^4.$$

NOTA Supongamos que

$$f = a_0 + a_1 X + \cdots + a_n X^n \in A[X]$$

y consideremos los polinomios (monomios) $f_0, f_1, \ldots, f_n$, donde $f_i = a_i X^i$. Sigue inmediatamente de la definición de suma en $A[X]$ que vale la igualdad

$$f = f_0 + f_1 + \cdots + f_n,$$

vale decir, todo polinomio es suma (en el sentido de la operación recién definida) de los monomios que definen su expresión. Como consecuencia de ello los símbolos "+" que aparecen en la representación de f dejan de ser meramente formales, lo que entre otras cosas legitima el uso del símbolo de sumatoria en la expresión de un polinomio. $\Diamond$

PRODUCTO DE POLINOMIOS En vez de definir directamente el producto de dos polinomios cualesquiera, definiremos en primer término el producto de monomios, teniendo en cuenta que todo polinomio es suma de monomios y que el producto debe ser distributivo con respecto a la suma. Así, dados $a, b \in A$ definimos

$$(aX^r)(bX^s) = abX^{r+s},$$

esto es, el producto de dos monomios es otro monomio cuyo grado (si existe) es la suma de los correspondientes grados y su coeficiente es el producto (en el anillo A) de los respectivos coeficientes. Vale la pena hacer notar que las potencias de la indeterminada, que en principio son expresiones formales, multiplican entre sí mediante una regla que vale en cualquier anillo conmutativo.

De acuerdo con el propósito formulado arriba, vayamos ahora al caso general. Si $f = \sum_{i=0}^{m} a_i X^i$ y $g = \sum_{i=0}^{n} b_i X^i$ son elementos de $A[X]$ definimos el *producto* de f y g en la forma

$$fg = \sum_{i,j} (a_i X^i)(b_j X^j) = \sum_{i,j} a_i b_j X^{i+j}\,,$$

esto es, multiplicamos cada monomio de f por cada monomio de g y luego sumamos. Naturalmente, podemos (y debemos) detallar mejor la expresión de arriba. Observemos primero que en el miembro de la derecha aparecerán monomios de igual grado correspondientes a distintos pares de índices, que de acuerdo con la definición de suma deberán ser agrupados. Por ejemplo, $a_0b_2 + a_1b_1 + a_2b_0$ será el coeficiente de grado 2 del producto, y en cuanto a los límites de la suma, es claro que el grado $i + j$ varía entre 0 y $m + n$, situaciones que como es inmediato verificar sólo se alcanzan en los casos $(i, j) = (0, 0)$ y $(i, j) = (m, n)$, respectivamente. En atención a estas consideraciones, tenemos:

$$fg = a_0b_0 + (a_0b_1 + a_1b_0)\,X + (a_0b_2 + a_1b_1 + a_2b_0)\,X^2 + \cdots + a_mb_nX^{m+n},$$

o en forma más compacta,

$$fg = \sum_{k=0}^{m+n} \left(\sum_{i+j=k} a_i b_j \right) X^k = \sum_{k=0}^{m+n} \left(\sum_{i=0}^{k} a_i b_{k-i} \right) X^k.$$

Por ejemplo, si $f = -2 + 3X^2 - X^3$, $g = 1 - 2X + X^2 + 3X^4 + X^5$ y $h = 1 + X - 2X^2 + X^3 - X^5$ tenemos

$$fg = -2 - 4X + X^2 - 6X^3 - 3X^4 - 3X^5 + 9X^6 - X^8$$
$$h^2 = 1 + 2X - 3X^2 - 2X^3 + 6X^4 - 6X^5 - X^6 + 4X^7 - 2X^8 + X^{10}.$$

Encargamos al lector la verificación de estos resultados. Como ejemplo de carácter general, destaquemos la forma sencilla que adopta el producto en

$A[X]$ de un polinomio cualquiera $f = \sum_i a_i X^i$ por un polinomio constante c. En efecto, aplicando la definición de producto sigue inmediatamente que

$$cf = \sum_i (ca_i)X^i,$$

esto es, basta multiplicar por c cada coeficiente de f. Resulta en particular que $0f = 0$ y $1f = f$ cualquiera sea $f \in A[X]$.

En el lema que sigue estableceremos resultados acerca del grado de la suma y el producto.

Lema 9.1.2 Si f y g son polinomios no nulos en $A[X]$ valen las siguientes propiedades:

1) $f + g = 0$ ó $gr(f + g) \leq \max\{gr(f), gr(g)\}$

2) Si $gr(f) \neq gr(g)$ entonces $gr(f + g) = \max\{gr(f), gr(g)\}$

3) $fg \neq 0$ y $gr(fg) = gr(f) + gr(g)$.

DEMOSTRACION. Si $f + g \neq 0$, supongamos sin pérdida de generalidad que $m = gr(f) \leq gr(g) = n$. Si designamos por $(a_i)_{i\in\mathbb{N}_0}$ y $(b_i)_{i\in\mathbb{N}_0}$ las secuencias de coeficientes de f y g, respectivamente, y k es cualquier índice mayor que n, tenemos que $a_k + b_k = 0 + 0 = 0$, pues $k > n \geq m$. Luego el coeficiente de grado k de $f + g$ es nulo, lo que asegura que el grado de $f + g$ es a lo sumo n, como queríamos probar. Si asumimos además que $m < n$, resulta que $a_n + b_n = b_n \neq 0$, de donde $gr(f + g) = n$. Esto completa la demostración de 1) y 2).

Respecto a 3), conservando las notaciones de la primera parte y designando por (c_k) la secuencia de coeficientes de fg, notamos en primer término que $c_{m+n} \neq 0$, ya que $c_{m+n} = a_m b_n$ es producto de dos elementos no nulos.

En particular, esta última afirmación prueba que $fg \neq 0$, resultando entonces que $gr(fg) = m + n$, pues ya habíamos observado en la definición de producto que $c_k = 0$ para todo $k > m + n$. Recomendamos al lector repasar los ejemplos anteriores, en los que se ilustran diversas situaciones que pueden registrarse respecto al grado de una suma. $\Diamond$

Proposición 9.1.3 $A[X]$ es un anillo conmutativo.

DEMOSTRACION. Es sencillo verificar, usando la validez de las correspondientes propiedades en el anillo de coeficientes A, que la suma y el producto de polinomios son operaciones asociativas y conmutativas, y que el producto es distributivo respecto de la suma. A manera de ilustración demostraremos este último hecho, dejando las restantes demostraciones a cargo del lector.

Sean pues $f = \sum_{i=0}^{m} a_i X^i$, $g = \sum_{i=0}^{n} b_i X^i$ y $h = \sum_{i=0}^{n} c_i X^i$ polinomios con coeficientes en A y designemos con las letras p y q los coeficientes de fg y fh, respectivamente. Tenemos entonces:

$$\begin{aligned} f\,(g+h) &= \sum_{k=0}^{m+n} \left(\sum_{i+j=k} a_i(b_j + c_j) \right) X^k = \sum_{k=0}^{m+n} \left(\sum_{i+j=k} (a_i b_j + a_i c_j) \right) X^k = \\ &= \sum_{k=0}^{m+n} \left(\sum_{i+j=k} a_i b_j + \sum_{i+j=k} a_i c_j \right) X^k = \sum_{k=0}^{m+n} (p_k + q_k)\, X^k = \\ &= \sum_{k=0}^{m+n} p_k X^k + \sum_{k=0}^{m+n} q_k X^k = fg + fh. \end{aligned}$$

Adviértase cómo hemos empleado las propiedades de las operaciones en A. Por otro lado, es claro que ambas operaciones admiten elemento neutro, el polinomio 0 en el caso de la suma y el polinomio 1 en el del producto, y que todo polinomio $f = \sum_i a_i X^i$ admite un inverso aditivo, a saber, el polinomio $\sum_i (-a_i) X^i$, que notaremos $-f$.

Se verifican en definitiva todas las propiedades que definen una estructura de anillo. Debido a ello, de aquí en adelante $A[X]$ será designado el *anillo de polinomios* en la indeterminada X con coeficientes en A. $\Diamond$

DIFERENCIA DE POLINOMIOS Como ya vimos en otras situaciones, la existencia de inverso aditivo permite definir la *diferencia* o *resta* de polinomios. En efecto, si $f = \sum_{i=0}^{n} a_i X^i$ y $g = \sum_{i=0}^{n} b_i X^i$ son polinomios con coeficientes en A, definimos la diferencia de f y g en la forma

$$f - g = f + (-g) = \sum_{i=0}^{n} (a_i + (-b_i)) X^i = \sum_{i=0}^{n} (a_i - b_i) X^i,$$

esto es, el coeficiente de grado i de $f - g$ es la diferencia (en A) de los correspondientes coeficientes de grado i.

De tal forma, si f y g son los polinomios que usamos para ilustrar la definición de suma tenemos que

$$f - g = -3 + 5X - 3X^2 + 2X^3 - 5X^4 - X^5.$$

Puesto que $gr(-g) = gr(g)$ si $g \neq 0$, es claro que el grado de la diferencia se comporta en forma idéntica al grado de la suma, esto es,

$$gr(f - g) \leq \max \{gr(f), gr(g)\}$$

si f y g son polinomios no nulos distintos.

POLINOMIOS INVERSIBLES Si $f \in A[X]$, diremos que f es *inversible* si y sólo si existe $g \in A[X]$ tal que $fg = 1$, vale decir, f admite inverso respecto al producto de polinomios.

La condición de inversibilidad es sumamente restrictiva, como apreciaremos a través de la caracterización que brinda el siguiente lema:

Lema 9.1.4 Sea A un cuerpo numérico y sea $f \in A[X]$. Entonces f es inversible si y sólo si $gr(f) = 0$.

DEMOSTRACION. Supongamos que f es inversible y sea $g \in A[X]$ tal que $fg = 1$. Tomando grados, obtenemos:

$$0 = gr(1) = gr(fg) = gr(f) + gr(g).$$

Siendo el grado un entero no negativo, es claro que la única posibilidad es $gr(f) = gr(g) = 0$.

Recíprocamente, supongamos que $gr(f) = 0$, digamos $f = aX^0$, donde a es un elemento no nulo de A. Sigue entonces por hipótesis que a admite un inverso b en A, y por lo tanto, designando por g al polinomio bX^0, tenemos:

$$fg = \left(aX^0\right)\left(bX^0\right) = abX^0 = 1\,X^0 = 1,$$

vale decir, f es inversible.

La situación varía en el caso de polinomios con coeficientes enteros, ya que si bien sigue siendo cierto que un elemento inversible f debe ser una constante no nula (basta repetir la primera parte de la demostración anterior), la misma debe por su parte ser inversible en $\mathbb{Z}$, resultando por lo tanto que f es inversible en $\mathbb{Z}[X]$ si y sólo si $f = 1$ ó $f = -1$. $\quad \Diamond$

POLINOMIOS SOBRE UN ANILLO ARBITRARIO La estructura de anillo de $A[X]$ sólo requiere disponer de operaciones de suma y producto en el conjunto A de coeficientes, satisfaciendo las propiedades asociativas, conmutativas, etc. Por lo tanto, dado cualquier anillo conmutativo B podemos construir, en idéntica forma a los casos $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ ó $\mathbb{C}$ el anillo de polinomios con coeficientes en B, que por supuesto notaremos $B[X]$.

Por ejemplo, consideremos el anillo $\mathbb{Z}_{12}[X]$ de polinomios con coeficientes en el anillo de clases de restos módulo 12. Aplicando las definiciones de suma y producto a los polinomios $f = 7 + 2X + 10X^2$ y $g = 1 - 5X + 4X^2 + 10X^3$ (en cada caso el coeficiente a representa su clase módulo 12), obtenemos:

$$f + g = 8 + 9X + 2X^2 + 10X^3 \quad \text{y} \quad fg = 7 + 3X + 4X^2 + 4X^3 + 4X^5.$$

El lector podrá demostrar sin problemas que, cualquiera sea B, es válida en $B[X]$ la relación que hemos establecido en el lema 9.1.2 para el grado de una suma, aunque no ocurre lo mismo respecto del grado de un producto, ya que por ejemplo tenemos en $\mathbb{Z}_{12}[X]$ que

$$(7 + 3X - 6X^2 + 3X^4)(2 + 9X + 3X^2 + 4X^3) = 2 + 9X + 7X^3 + 3X^5 + 9X^6,$$

esto es, el grado de un producto puede ser menor que la suma de los grados.

Para mostrar otras situaciones "extrañas" que pueden ocurrir, consideremos en $\mathbb{Z}_8[X]$ los polinomios $u = 2 + 2X$, $v = 4 + 4X^2$ y $w = 1 + 4X$. Un sencillo cálculo nos muestra entonces que $uv = 0$ y $w^2 = 1$, vale decir,

el producto de dos polinomios no nulos puede ser nulo, y un polinomio inversible puede tener grado mayor que 0. Naturalmente, estas anomalías dependen del anillo B de coeficientes, y motivados por ellas introducimos la siguiente definición:

Un anillo B se dice *íntegro* si y sólo si

$$xy = 0 \Rightarrow x = 0 \text{ ó } y = 0\,,$$

cualesquiera sean los elementos $x, y \in B$.

Equivalentemente, el producto de dos elementos no nulos de B es distinto de 0. Si además B es conmutativo, se lo llama un *dominio de integridad*.

Con la misma demostración que brindamos en el capítulo 2 para el caso de $\mathbb{R}$ es inmediato probar que los cuerpos son dominios de integridad, resultando en particular que los anillos numéricos lo son, ya que $\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$ son cuerpos y el producto en $\mathbb{Z}$ es la restricción del producto en $\mathbb{Q}$. En cambio $\mathbb{Z}_{12}$ no es un dominio de integridad, pues por ejemplo $3\,.\,4 = 0$ en este anillo.

En los anillos de polinomios con coeficientes en un dominio de integridad valen propiedades muy similares a las ya establecidas para polinomios sobre anillos numéricos. A continuación enunciamos dos de ellas, dejando las demostraciones a cargo del lector, ya que las mismas son idénticas a las efectuadas anteriormente (B designa un dominio de integridad y f y g polinomios con coeficientes en B).

Lema 9.1.5 Son válidas las siguientes propiedades:

1) Si f y g son no nulos entonces $fg \neq 0$, esto es, $B[X]$ es un dominio de integridad. Además, $gr(fg) = gr(f) + gr(g)$.

2) Los elementos inversibles de $B[X]$ son exactamente los polinomios a de grado 0 tales que a es inversible en B (respecto al producto). $\Diamond$

Derivación.

Sea A cualquiera de los anillos numéricos y sea

$$f = \sum_{i=0}^{n} a_i X^i$$

un polinomio con coeficientes en A. Definimos entonces el *polinomio derivado* de f en la forma

$$f' = \sum_{i=1}^{n} i a_i X^{i-1}.$$

Por ejemplo, si $f = 2 - 3X + 1/6X^3 + X^4$ entonces $f' = -3 + 1/2X^2 + 4X^3$.

NOTA Destaquemos algunos hechos relacionados con la definición anterior:

i) Si bien un polinomio f admite infinitas representaciones (recordemos que pueden agregarse coeficientes nulos), es claro que la definición de f' no depende de la forma particular elegida, ya que $ia_i = 0$ si $a_i = 0$. Notemos de paso que si

$$(a_0, a_1, a_2, a_3, \ldots, a_n, 0, 0, \ldots)$$

es la secuencia infinita de coeficientes de f entonces

$$(a_1, 2a_2, 3a_3, 4a_4, \ldots, na_n, 0, 0, \ldots)$$

es la secuencia de coeficientes de f'. Sigue en particular que $f' = 0$ si y sólo si $f = 0$ ó $gr(f) = 0$, ya que $ka_k \neq 0$ si $a_k \neq 0$ y $k > 0$.

ii) Al lector familiarizado con las nociones básicas del cálculo diferencial le resultará sin duda conocida la expresión del polinomio derivado, ya que el mismo se obtiene aplicando formalmente las reglas de derivación a la expresión de f. Más adelante quedará clara la conexión existente entre ambos conceptos, pues asociaremos a cada polinomio f con coeficientes reales una función de $\mathbb{R}$ en $\mathbb{R}$ cuya función derivada (en el sentido del análisis matemático) coincide con la función asociada al polinomio derivado f'.

iii) La definición de polinomio derivado determina una aplicación (operador)

$$\partial : A[X] \to A[X],$$

a través de la fórmula $\partial(f) = f'$, ya que claramente $i\, a_i \in A$ si $a_i \in A$. Lo llamaremos el *operador de derivación* en $A[X]$.

iv) Si f es un polinomio, aplicando el operador de derivación a f' obtenemos un nuevo polinomio, que llamaremos polinomio derivado segundo de f y notaremos $f^{(2)}$. Iterando esta operación obtenemos el derivado tercero $f^{(3)}$, el cuarto $f^{(4)}$, y así sucesivamente. En general, dado un entero no negativo k, notaremos por $f^{(k)}$ el polinomio derivado k-ésimo de f, obtenido aplicando k veces el operador ∂ a f. Rigurosamente hablando, definimos inductivamente los derivados sucesivos de f en la forma:

$$f^{(k)} = \begin{cases} f & \text{si } k = 0 \\ \partial(f^{(k-1)}) & \text{si } k > 0\,. \end{cases}$$

Por ejemplo, sea $f = 4 + 2X^2 + 5X^3$. Entonces $f' = 4X + 15X^2$, $f^{(2)} = 4 + 30X$, $f^{(3)} = 30$ y $f^{(k)} = 0$ si $k \geq 4$. $\quad \Diamond$

Proposición 9.1.6 La derivación satisface las siguientes propiedades (las letras f y g designan polinomios y las letras c y d elementos de A):

1) Si $gr(f) > 0$ entonces $gr(f') = gr(f) - 1$

2) $\partial(cf + dg) = c\,\partial(f) + d\,\partial(g)$. Más generalmente,

$$\partial\left(\sum_{i=1}^{m} c_i f_i\right) = \sum_{i=1}^{m} c_i\,\partial(f_i)$$

3) $\partial(fg) = \partial(f)g + f\partial(g)$. En general,

$$\partial\left(\prod_{i=1}^{m} f_i\right) = \sum_{i=1}^{m}\left(\prod_{j\neq i} f_j\right)\partial(f_i)$$

4) $\partial(f^m) = mf^{m-1}\partial(f)$ para todo $m \in \mathbb{N}$.

DEMOSTRACION. La propiedad 1) sigue inmediatamente de la definición de f'. Para probar 2), designemos por a_i y b_i los coeficientes de f y g, respectivamente. Entonces:

$$\begin{aligned}\partial(cf + dg) &= \partial\left(\sum_i (ca_i + db_i)X^i\right) = \sum_i i(ca_i + db_i)X^{i-1} = \\ &= \sum_i ica_i X^{i-1} + \sum_i idb_i X^{i-1} = c\sum_i ia_i X^{i-1} + d\sum_i ib_i X^{i-1} = \\ &= c\,\partial(f) + d\,\partial(g).\end{aligned}$$

El enunciado 2) se denomina propiedad de *linealidad* de la derivación, y su generalización sigue trivialmente por inducción en m. Respecto al ítem 3), puesto que un producto es una suma de productos de monomios y ya hemos probado que el derivado de una suma es la suma de los derivados, bastará demostrarlo en el caso de monomios, digamos $f = aX^r$ y $g = bX^t$. Tenemos entonces:

$$\begin{aligned}\partial(fg) = \partial(abX^{r+t}) &= (r+t)abX^{r+t-1} = rabX^{r+t-1} + tabX^{r+t-1} = \\ &= (raX^{r-1})(bX^t) + (aX^r)(tbX^{t-1}) = \partial(f)g + f\partial(g).\end{aligned}$$

Finalmente, la segunda parte de 3) y la propiedad 4) se deducen mediante sencillos argumentos inductivos a partir de la fórmula hallada para el derivado de un producto . ♢

9.1.3. Ejercicios

1. Determinar la secuencia (infinita) de coeficientes de los siguientes elementos de $\mathbb{C}[X]$:

   a) $2X^7 - X^5 + 3X^4 + X$

   b) $8(X+1)^3 - 4(X-1)^4$

c) $\displaystyle\sum_{k=0}^{10} X^{2k}$

d) $\displaystyle\sum_{k=0}^{6} \imath^k \binom{5}{k} X^{20-3k}$.

2. *a)* Sea f un polinomio con coeficientes en un anillo conmutativo. Sabiendo que los grados de f, $f + X^4$ y Xf son tres enteros consecutivos, determinar los posibles valores de $gr(f)$

   b) Hallar $f \in \mathbb{Z}_{12}[X]$, de grado 3, tal que $f^2 = 0$. Determinar todos los polinomios de tal tipo en $\mathbb{Z}_{12}[X]$

   c) Sean $f, g \in \mathbb{C}[X]$ tales que $gr(f + Xg) = 5$ y $gr(fg^2) = 16$. Calcular $gr(f)$ y $gr(g)$.

3. Sea $g = X^3 + aX^2 + bX + c$, donde las letras designan números complejos. En cada uno de los siguientes casos, determinar a, b y c de manera que g satisfaga la relación planteada:

   a) $g = (1/3)Xg' - X^2 - 1$

   b) $g^2 = X^6 - 2\imath X^4 + 2\imath X^3 - X^2 + 2X - 1$

   c) $(X + 1)g = X^4 - 2g$.

4. En cada uno de los siguientes casos, determinar los $f \in \mathbb{R}[X]$ que satisfacen la relación planteada:

   a) $f^2 + 2X = (X + 6)f - 8$

   b) $X^2 f + 2f = X^4 - 4$

   c) $f^2 = X^3 - 3X^2 + (3X - 1)f + X - 2$

   d) $(f')^3 = 8Xf$.

5. Consideremos en $\mathbb{C}[X]$ los polinomios $f = X^3 + \imath X^2 - 1$, $g = X^2 + X + 4$ y $h = X^4 + (2 + \imath)X^3 + 2\imath X^2 - X - 2$.

   a) Calcular el coeficiente principal y los coeficientes de grado 0, 3, 6 y 12 de los polinomios

      i) $f^2 g + h^3$

      ii) $(1 + g - h)^3$

      iii) $f^2 + g^2 + h^2$

      iv) fgh

      v) $Xf + X^2 - 2h$

      vi) $g^4 - h^2$.

b) Analizar la existencia de $u \in \mathbb{C}[X]$ tal que $f = gu$. Lo mismo respecto a las situaciones $h = fu$, $h = gu$ y $g = fu$.

6. Si A es un anillo conmutativo, demostrar que las fórmulas de Newton y Leibniz son válidas en $A[X]$.

7. Demostrar que $f = 1 + 2X + 4X^3$ es inversible en $\mathbb{Z}_8[X]$ (sug: hallar en $\mathbb{Z}_8[X]$ un polinomio g de grado 3 tal que $fg = 1$).

8. Sea $P = \{f = \sum a_i X^i \in \mathbb{C}[X] : a_i = 0 \text{ para todo } i \text{ impar}\}$.

   a) Probar que la suma y el producto de polinomios son operaciones cerradas en P

   b) Dado $g \in \mathbb{C}[X]$ probar que existe un único par g_0, g_1 de elementos de P tales que $g = g_0 + Xg_1$

   c) Hallar g_0 y g_1 en los casos

   i) $g = X^7 - 2X^6 + 3X^4 - X^3 + X - 1$

   ii) $g = 3$

   iii) $g = -2X$

   iv) $g = \sum_{i=0}^{9} X^{2i+1}$.

9. Si $n \in \mathbb{N}$, exhibir un polinomio $f \in \mathbb{Z}_n[X]$ de grado positivo tal que $f' = 0$.

10. Si $n \in \mathbb{N}$ y $h = \prod_{k=1}^{n}(X - k)$, calcular $h'(0)$ y $h'(n)$.

9.2. Raíces

9.2.1. Especialización y raíces

Todo polinomio

$$f = \sum_{i=0}^{m} a_i X^i$$

con coeficientes en A determina una función de A en A (que también designaremos por f), a través de la fórmula

$$f(c) = \sum_{i=0}^{m} a_i c^i \ \ (c \in A)\,.$$

En otras palabras, $f(c)$ es el valor numérico obtenido efectuando las operaciones indicadas en la expresión de f luego de reemplazar X por c. Lo llamaremos *especialización* o *evaluación* de f en c.

Ejemplos 9.2.1 Tomemos por caso $f = X^4 - X^3 + 3X^2 - 5X - 10$. Puesto que un polinomio con coeficientes en un anillo numérico es en definitiva un polinomio con coeficientes complejos, podemos especializar f en cualquier elemento de $\mathbb{C}$. Por ejemplo, un sencillo cálculo nos muestra que $f(-1) = 1$, $f(\sqrt{3}) = 8 - 8\sqrt{3}$ y $f(-\imath) = -12 + 4\imath$. Notemos además que $f(0) = -10$, y que, en general, $f(0)$ es el coeficiente de grado 0 de f.

Ciertas funciones, bien conocidas por el lector, son casos particulares de especialización. Así, la función lineal $f(x) = ax + b$ es la función de especialización de un polinomio de grado 1, mientras que la función cuadrática $f(x) = ax^2 + bx + c$ está determinada por un polinomio de grado 2. Finalmente, las funciones constantes están asociadas a polinomios de grado 0 ó al polinomio nulo, lo que justifica el nombre que les hemos dado a estos polinomios. $\Diamond$

RAICES

Conservando las notaciones de la definición anterior, diremos que c es una *raíz* de f si y sólo si $f(c) = 0$, esto es, c es solución de la ecuación algebraica

$$a_m X^m + a_{m-1} X^{m-1} + \cdots + a_1 X + a_0 = 0\,.$$

Alternativamente, diremos que c es un *cero* de f, o también que c anula a f.

Ejemplos 9.2.2 Comencemos con un caso trivial, el de los polinomios constantes. Es claro por definición que los polinomios de grado 0 no admiten

ninguna raíz. Por el contrario, todo $c \in \mathbb{C}$ es raíz del polinomio nulo, siendo éste el único polinomio que tiene infinitos ceros, como veremos más adelante. Observemos también que 0 es raíz de f si y sólo si $a_0 = 0$ y que 1 es raíz de f si y sólo si la suma de los coeficientes a_i es cero.

Como ejemplo más concreto, el lector podrá verificar inmediatamente que 2 es una raíz (racional) del polinomio f del ejemplo 9.2.1, que tiene coeficientes en $\mathbb{Q}$. En general, un polinomio con coeficientes en A no necesariamente admite una raíz en A. Si bien conocemos un ejemplo paradigmático de tal hecho (el polinomio X^2+1 no tiene raíces en $\mathbb{R}$), veremos ahora un caso bastante menos obvio, como es el del polinomio con coeficientes racionales

$$h = X^4 - 2X^3 - X^2 + 6X - 6\,.$$

A pesar de que no tenemos una fórmula para el cálculo de sus ceros, podemos hallarlos factorizando convenientemente. En efecto, si $z \in \mathbb{C}$ tenemos:

$$\begin{aligned} h(z) &= z^4 - 2z^3 - z^2 + 6z - 6 = -2z(z^2-3) + z^4 - z^2 - 6 = \\ &= -2z(z^2-3) + (z^2-3)^2 + 5z^2 - 15 = \\ &= -2z(z^2-3) + (z^2-3)^2 + 5(z^2-3) = \\ &= (z^2-3)(-2z + z^2 - 3 + 5) = (z^2-3)(z^2-2z+2). \end{aligned}$$

Teniendo en cuenta que un producto es nulo si y sólo si alguno de los factores es 0, concluimos que z es raíz de h si y sólo si z es raíz de X^2-3 ó z es raíz de X^2-2X+2. Puesto que éstos polinomios son de grado 2 podemos hallar sus ceros, resultando que ellos son $\pm\sqrt{3}$ y $1 \pm \imath$. En resumen, h tiene 4 raíces complejas (dos de las cuales son reales) y no admite raíces racionales.

Pronto veremos que algunos de estos hechos se enmarcan en otros más generales, cuyo conocimiento nos hubiera permitido anticiparlos. Aclaremos de todos modos que lo anterior no constituye un método eficaz de cálculo, ya que en general no es fácil factorizar una expresión polinomial. $\lozenge$

En la siguiente proposición veremos cómo se comporta la especialización respecto de las operaciones en $A[X]$.

Proposición 9.2.3 Si $f, g \in A[X]$ y $z \in A$, son válidas las siguientes afirmaciones:

1) $(f+g)(z) = f(z) + g(z)$

2) $(fg)(z) = f(z)g(z)$.

DEMOSTRACION. Básicamente, estas propiedades son consecuencia del hecho de que en las operaciones de polinomios la indeterminada X satisface

reglas que valen en cualquier anillo conmutativo. De todas maneras brindaremos una demostración formal. Para probar 1), supongamos sin pérdida de generalidad que $f = \sum_{i=0}^{m} a_i X^i$ y $g = \sum_{i=0}^{m} b_i X^i$. Entonces:

$$(f+g)(z) = \sum_{i=0}^{m} (a_i + b_i) z^i = \sum_{i=0}^{m} a_i z^i + b_i z^i = \\ = \sum_{i=0}^{m} a_i z^i + \sum_{i=0}^{m} b_i z^i = f(z) + g(z).$$

Habiendo demostrado 1), razonando como en una oportunidad anterior bastará probar 2) para el caso en que f y g son monomios, digamos $f = aX^m$ y $g = bX^n$. Tenemos entonces:

$$(fg)(z) = abz^{m+n} = abz^m z^n = (az^m)(bz^n) = f(z)g(z). \quad \Diamond$$

Las fórmulas precedentes se generalizan sin dificultad a los casos de sumas y productos de cualquier familia finita de polinomios (r denota un número natural):

Corolario 9.2.4 Sean $h_1, h_2, \ldots, h_r$ polinomios con coeficientes en A y sea $z \in A$. Entonces

1) $\left(\sum_{j=1}^{r} h_j\right)(z) = \sum_{j=1}^{r} h_j(z)$

2) $\left(\prod_{j=1}^{r} h_j\right)(z) = \prod_{j=1}^{r} h_j(z)$

3) z es raíz de $h = \prod_{j=1}^{r} h_j$ si y sólo si z es raíz de h_k para algún índice k.

DEMOSTRACION. Las propiedades 1) y 2) se deducen de 9.2.3 mediante argumentos inductivos, mientras que 3) es consecuencia de 2), teniendo en cuenta que todo anillo numérico es un dominio de integridad.

Por ejemplo, las raíces del polinomio

$$X(X+1)(X-2)(X^2 - X - 12)(X^2 + 4X - 5)$$

son $0, -1, 2, -3, 4, 1$ y -5, ya que éstas son las raíces de los factores, como se comprueba fácilmente. Asimismo, se deduce de 3) que la búsqueda de raíces siempre puede circunscribirse a polinomios mónicos, ya que todo polinomio no nulo puede expresarse en la forma $f = ag$, donde a es el coeficiente

prinicipal de f y g es mónico, resultando por lo tanto que z es raíz de f si y sólo si z es raíz de g. $\Diamond$

La conexión entre factorización y raíces que establece la propiedad 3) del corolario anterior admite una especie de recíproca, en el sentido de que toda raíz de un polinomio determina una factorización del mismo. Este hecho, que demostraremos muy pronto, nos permitirá obtener una cota para el número de raíces de un polinomio no nulo, además de otros resultados importantes. Probaremos en primer término un resultado auxiliar.

Lema 9.2.5 Sea $z \in A$ y sea $m \in \mathbb{N}$. Existe entonces un polinomio g_m en $A[X]$ tal que

$$X^m - z^m = (X - z)g_m\,.$$

DEMOSTRACION. Por inducción en m. El resultado es trivial si $m = 1$ (tomando $g_1 = 1$). Para el paso inductivo, supongamos que $X^k - z^k = (X - z)g_k$ con $g_k \in A[X]$. Resulta entonces que

$$\begin{aligned} X^{k+1} - z^{k+1} &= X(X^k - z^k) + z^k X - z^{k+1} = \\ &= X(X - z)g_k + z^k(X - z) = (X - z)(Xg_k + z^k). \end{aligned}$$

Luego el resultado es válido para $k + 1$ tomando $g_{k+1} = Xg_k + z^k$. $\Diamond$

Si $z \in A$ es claro que z es raíz de $X - z$, de donde deducimos por el ítem 3) de 9.2.4 que z es raíz de todo polinomio que admita a $X - z$ como factor. Probaremos a continuación que la recíproca de este hecho también es válida.

Proposición 9.2.6 Sea

$$f = \sum_{i=0}^{n} a_i X^i \in A[X]$$

y sea z una raíz de f en A. Existe entonces $h \in A[X]$ tal que $f = (X - z)h$.

DEMOSTRACION. El resultado deseado es consecuencia del lema previo, ya que operando convenientemente tenemos:

$$\begin{aligned} f = f - f(z) &= \sum_{i=0}^{n} a_i X^i - \sum_{i=0}^{n} a_i z^i = \sum_{i=1}^{n} a_i (X^i - z^i) = \\ &= \sum_{i=1}^{n} a_i (X - z) g_i = (X - z) \sum_{i=1}^{n} a_i g_i\,. \end{aligned}$$

Luego basta tomar $h = \sum_{i=1}^{n} a_i g_i$. $\Diamond$

Corolario 9.2.7 Sea $n \in \mathbb{N}_0$ y sea $f \in A[X]$ un polinomio de grado n. Entonces f tiene a lo sumo n raíces en A.

DEMOSTRACION. Haremos inducción en n. El enunciado es ciertamente válido si $n = 0$ ó f no admite raíces en A. Supongamos entonces que $n > 0$ y consideremos cualquier conjunto $\{z_1, z_2, \ldots, z_r\}$ de raíces de f en A. Bastará probar que $r \leq n$.

Por proposición 9.2.6, podemos escribir $f = (X - z_1)h$, donde $h \in A[X]$ y $gr(h) = n-1$, resultando por 3) del corolario 9.2.4 que z_i es raíz de h para todo $i > 1$. Puesto que por un argumento inductivo h tiene a lo sumo $n-1$ raíces en A, sigue que $r - 1 \leq n - 1$, y por lo tanto $r \leq n$. $\quad \Diamond$

Corolario 9.2.8 Sean $g, h \in A[X]$ tales $g(z) = h(z)$ para todo $z \in A$. Entonces $g = h$.

DEMOSTRACION. Observemos el significado de este hecho: la función asociada a un polinomio f determina completamente los coeficientes de f. Para probarlo, consideremos la diferencia $t = g - h$. Si $z \in A$, sigue por las propiedades de la especialización que $t(z) = (g - h)(z) = g(z) - h(z) = 0$, por hipótesis. Esto es, t se anula sobre todo elemento de A, de donde sigue por corolario 9.2.7 que $t = 0$, ya que en otro caso tendría a lo sumo tantas raíces como su grado. Luego $g = h$. $\Diamond$

Deducimos de la demostración anterior que hubiéramos arribado a la misma conclusión suponiendo que $g(z)$ y $h(z)$ coinciden sobre *infinitos* valores de z, ó más generalmente, sobre un conjunto "suficientemente grande" de valores de z. En el siguiente (e interesante) resultado quedará claro el sentido de esta última afirmación.

Proposición 9.2.9 (Fórmula interpoladora de Lagrange) Sea $n \in \mathbb{N}$ y sea K un cuerpo numérico. Si $a_0, a_1, \ldots, a_n$ son $n+1$ elementos distintos de K y $b_0, b_1, \ldots, b_n$ son $n+1$ elementos cualesquiera de K, existe un único $f \in K[X]$ de grado menor o igual que n (o nulo) tal que $f(a_i) = b_i$ para $i = 0, 1, \ldots, n$.

DEMOSTRACION. Sea

$$f = \sum_{k=0}^{n} b_k \frac{(X - a_0)\ldots(X - a_{k-1})(X - a_{k+1})\ldots(X - a_n)}{(a_k - a_0)\ldots(a_k - a_{k-1})(a_k - a_{k+1})\ldots(a_k - a_n)} . \qquad (9.2)$$

Observemos que cada uno de los $n+1$ sumandos de (9.2) es de la forma $b_k(f_k/c_k)$, donde $f_k = \prod_{j \neq k}(X - a_j)$ es un polinomio de grado n y la constante $c_k = \prod_{j \neq k}(a_k - a_j)$ es no nula. Podemos afirmar entonces en primer lugar que f es nulo o tiene grado menor o igual que n, por ser suma de polinomios nulos o de grado n. Por otro lado, es claro a partir de las definiciones que

$f_k(a_j) = 0$ para todo $j \neq k$. Luego, si i es cualquier índice entre 0 y n, tenemos:

$$f(a_i) = \sum_{k=0}^{n} (b_k/c_k)\, f_k(a_i) = (b_i/c_i)\, f_i(a_i) = (b_i/c_i)\, c_i = b_i\,,$$

esto es, f cumple los requerimientos del enunciado. Para probar la unicidad, supongamos que $g \in K[X]$ también los satisface. Resulta entonces que $f - g$ admite por lo menos $n + 1$ raíces en K, ya que

$$(f - g)(a_i) = f(a_i) - g(a_i) = b_i - b_i = 0$$

para todo índice $i\,(0 \leq i \leq n)$, lo que nos permite concluir que $f - g$ es nulo, pues en otro caso resultaría ser un polinomio de grado menor o igual que n (por serlo f y g) admitiendo más de n raíces en K. Luego $g = f$ $\quad \Diamond$

La expresión (9.2) se llama *fórmula de interpolación* de Lagrange, y f el *polinomio interpolador*. A manera de interpretación geométrica, notemos que la foŕmula de Lagrange permite definir, digamos en el caso real, una función polinómica de grado a lo sumo n cuyo gráfico pase por los $n + 1$ puntos (a_i, b_i) del plano. Veamos algunos ejemplos.

Ejemplos 9.2.10 Construyamos primero una cúbica que pase por los puntos $(-1, -3/2)$, $(0, 1)$, $(1, 7/2)$ y $(2, 12)$. Aplicando (9.2) obtenemos (el lector verificará los cálculos):

$$\begin{aligned} f &= (1/4)(X^3 - 3X^2 + 2X) + (1/2)(X^3 - 2X^2 - X + 2) - \\ &\quad - (7/4)(X^3 - X^2 - 2X) + 2(X^3 - X) = \\ &= X^3 + (3/2)X + 1\,. \end{aligned}$$

Tomemos ahora $n = 2$ y construyamos el polinomio interpolador de Lagrange en el caso $a_0 = -2$, $a_1 = 1$, $a_2 = 3$, $b_0 = -10$, $b_1 = 2$ y $b_2 = 10$. Tenemos entonces:

$$\begin{aligned} f &= (-2/3)(X^2 - 4X + 3) - (1/3)(X^2 - X - 6) + (X^2 + X - 2) = \\ &= 4X - 2\,. \end{aligned}$$

Puesto que por 9.2.9 f es el único polinomio de grado menor o igual que 2 que satisface las condiciones planteadas, deducimos que no existe ninguna parábola (gráfico de una función cuadrática) que pase por los puntos $(-2, -10)$, $(1, 2)$ y $(3, 10)$. La razón es muy sencilla desde el punto de vista geométrico: dichos puntos están alineados. $\quad \Diamond$

NOTA. Las nociones de especialización y raíces se extienden sin dificultad a polinomios con coeficientes en cualquier anillo conmutativo B. Así, todo f en $B[X]$ determina una aplicación (función de especialización) de B en

B, siendo sus raíces el conjunto de elementos $b \in B$ tales que $f(b) = 0$. Es inmediato verificar que dicha especialización conserva las propiedades enunciadas en la proposición 9.2.3, aunque dependiendo del anillo B existen diferencias con respecto a los anillos numéricos que vale la pena remarcar.

Tomemos en primer lugar un primo p y consideremos en $\mathbb{Z}_p[X]$ el polinomio $f = X^p - X$. Si lo evaluamos en un elemento genérico de $\mathbb{Z}_p$, digamos la clase (a) de cualquier $a \in \mathbb{Z}$, obtenemos $f((a)) = (a^p - a) = (0)$, pues $a^p \equiv a\ (p)$ por el teorema de Fermat. En consecuencia, f es un ejemplo de un polinomio no nulo cuya función de especialización es idénticamente nula, lo que muestra que el corolario 9.2.8 no es válido en $\mathbb{Z}_p[X]$.

Consideremos ahora en $\mathbb{Z}_6[X]$ el polinomio $g = X^4 + X^3 + 2X^2 + 2X$, que podemos descomponer como producto de factores de menor grado como sigue:

$$g = X^2(X^2+2) + X(X^2+2) = (X^2+2)(X^2+X) = X(X+1)(X^2+2).$$

Si a es un entero, es claro que $a(a+1)$ es par, y sigue además del teorema de Fermat que $a(a^2+2)$ es múltiplo de 3. Por lo tanto $a(a+1)(a^2+2)$ es divisible por 6, lo que significa que $g(t) = 0$ para todo $t \in \mathbb{Z}_6$. Luego, hemos hallado un ejemplo de un polinomio que admite un número de raíces mayor que su grado. Veremos más adelante que esta situación no puede ocurrir si el anillo de coeficientes es un cuerpo.

Notemos finalmente que (1), que es una de las raíces de g en $\mathbb{Z}_6$, no es raíz de ninguno de los factores X, $X+1$ y X^2+2. Naturalmente, esto se debe a que $\mathbb{Z}_6$ no es un dominio de integridad. $\Diamond$

Números algebraicos

> Si $z \in \mathbb{C}$, diremos que z es un *número algebraico* si y sólo si z es raíz de un polinomio no nulo con coeficientes racionales. Si z no es algebraico, diremos que es *trascendente*.

Antes de la ilustrar con ejemplos estas definiciones resaltaremos algunos hechos. Lo haremos principalmente a título informativo, ya que las dificultades que presentan sus demostraciones exceden largamente el alcance de estas páginas. Digamos por ejemplo que la suma, resta, producto y cociente de números algebraicos es algebraico, por lo que el conjunto de tales números es un subcuerpo del cuerpo de números complejos. El mismo es *numerable* (coordinable con el conjunto de números naturales), mientras que el conjunto de números trascendentes (su complemento) es coordinable con el conjunto de números reales. Expresado en lenguaje coloquial, esto significa que hay una abrumadora "mayoría" de números trascendentes respecto de los algebraicos, no obstante lo cual no es tarea sencilla exhibir números trascendentes. Señalemos en cuanto a esto que Liouville (1844) fue el primero en demostrar rigurosamente la existencia de tales números, y que Hermite (1873) y Lindemann (1882) probaron, con recursos algebraicos y analíticos,

la trascendencia de e (base de los logaritmos naturales) y π, respectivamente. En general es un problema muy difícil decidir si un número es algebraico o trascendente, y si bien existen teoremas generales que engloban muchas situaciones, quedan aún muchas cuestiones pendientes, como por caso la naturaleza de $e + \pi$.

Ejemplos 9.2.11 Veamos algunos ejemplos de números algebraicos y trascendentes. En el caso de estos últimos sólo haremos referencia a ellos, sin brindar las correspondientes justificaciones.

Como caso trivial comencemos observando que todo número racional r es algebraico, pues es raíz del polinomio $X - r$, y como primeros ejemplos sencillos de números algebraicos no racionales citemos a $\sqrt{2}$ y $\sqrt[3]{4}$, que anulan a los polinomios X^2-2 y X^3-4, respectivamente. Más generalmente, dados $n \in \mathbb{N}$ y $a \in \mathbb{Q}$, todas las raíces $n-$ésimas de a en $\mathbb{C}$ son números algebraicos, pues son las raíces del polinomio $X^n - a$.

Un ejemplo célebre de irracional algebraico es

$$\phi = \frac{1+\sqrt{5}}{2}\,,$$

llamado el *número de oro* y raíz del polinomio X^2-X-1. Más generalmente, todo número real de la forma $a + b\sqrt{5}$ $(a, b \in \mathbb{Q})$ es algebraico, pues es raíz del polinomio con coeficientes racionales $X^2-2aX+a^2-5b^2$, como se verifica fácilmente.

Como un último ejemplo de número algebraico veamos que $\omega = \sqrt{2} + \imath$ lo es. En efecto, elevando al cuadrado tenemos

$$\omega^2 = \left(\sqrt{2} + \imath\right)^2 = 2 - 1 + 2\imath\sqrt{2} = 1 + 2\imath(\omega - \imath) = 3 + 2\imath\omega,$$

de donde resulta que $\omega^2 - 3 = 2\imath\omega$. Elevando nuevamente al cuadrado obtenemos

$$\omega^4 - 6\omega^2 + 9 = -4\omega^2,$$

de donde $\omega^4 - 2\omega^2 + 9 = 0$. Hemos hallado así una relación algebraica con coeficientes racionales entre potencias de ω, lo que indica que ω es un número algebraico. Precisamente, ω es raíz del polinomio $X^4 - 2X^2 + 9$.

Respecto a números no algebraicos, consignemos que Liouville probó que la serie

$$\sum_{k=1}^{\infty} 10^{-k!}$$

converge a un número trascendente, llamado por ello la constante de Liouville. Observemos que su desarrollo decimal es de la forma

$$0{,}11000100000000000000000100000\ldots,$$

donde los unos aparecen en las posiciones 1,2,6,24, etc., correspondientes a números factoriales. Como ejemplo algo curioso, señalemos que también es trascendente el número

$$0{,}12345678910111213\ldots,$$

cuyo desarrollo decimal consiste de la sucesión infinita de números naturales, hecho demostrado por K. Malher en 1937.

Para terminar mencionemos a $2^{\sqrt{2}}$ y e^{π}. Ambos son trascendentes y constituyen casos particulares de una situación más general, resuelta por Guelfond en 1934, en conexión con el famoso séptimo problema de Hilbert acerca de la trascendencia de números de la forma α^{β}, donde α es algebraico $(\alpha \neq 0, 1)$ y β no es racional. $\Diamond$

Determinación de raíces.

Hemos señalado al iniciar este capítulo que no existen fórmulas generales que permitan hallar las raíces de cualquier polinomio de grado mayor que 4. Sin embargo, en muchos casos es posible establecer interesantes resultados teóricos acerca de las mismas, algunos de los cuales vamos a destacar a lo largo de este apartado. Previamente daremos un informe acerca de la existencia de fórmulas de cálculo en los grados bajos, puntualizando algunos casos ya conocidos e incluyendo alguno nuevo. Comenzaremos con el más simple.

POLINOMIOS DE GRADO UNO. Sea

$$f = aX + b$$

un polinomio de grado 1 con coeficientes en A. Es claro entonces que f tiene una única raíz en $\mathbb{C}$, a saber $z = -b/a$. Notemos que $z \in A$ si A es un cuerpo, mientras que sólo podemos asegurar que $z \in \mathbb{Q}$ si $f \in \mathbb{Z}[X]$.

POLINOMIOS CUADRATICOS Y BICUADRATICOS. Se trata de polinomios de grado 2, digamos

$$f = aX^2 + bX + c,$$

o de grado 4 de la forma

$$g = aX^4 + bX^2 + c,$$

respectivamente. En el primer caso vimos que z es una raíz de f en $\mathbb{C}$ si y sólo si

$$(2az + b)^2 = b^2 - 4ac = \Delta, \tag{9.3}$$

obteniéndose la fórmula

$$z = \frac{-b \pm \omega}{2a},$$

donde ω es cualquiera de las raíces cuadradas complejas de Δ. En la situación particular de que los coeficientes de f sean reales (resp. racionales), deducimos de la fórmula anterior que f admite raíces reales (resp. racionales) si y sólo si Δ es un cuadrado en $\mathbb{R}$ (resp. en $\mathbb{Q}$), lo que equivale en el caso real a que Δ sea mayor o igual que cero, mientras que en el caso racional dicha condición es necesaria pero no suficiente. Por ejemplo, $X^2 - 3X + 1$ tiene raíces irracionales, pues $\Delta = 5$ es un cuadrado en $\mathbb{R}$ pero no en $\mathbb{Q}$.

El caso bicuadrático sigue del anterior mediante la sustitución $Y = X^2$, resultando entonces que g admite cuatro raíces en $\mathbb{C}$, a saber, las raíces cuadradas complejas de cada una de las dos raíces de f. De acuerdo con esto, podemos analizar en forma similar a la anterior la pertenencia a $\mathbb{R}$ ó $\mathbb{Q}$ de las raíces de polinomios bicuadráticos con coeficientes reales o racionales. Por ejemplo, $X^4 - 3X^2 - 4$ tiene exactamente dos raíces reales (racionales), ya que $X^2 - 3X - 4$ tiene una raíz negativa (-1) y otra positiva (4) que es un cuadrado en $\mathbb{Q}$.

NOTA. La fórmula para hallar los ceros de un polinomio de grado 2 es válida para polinomios con coeficientes en cualquier cuerpo K de característica distinta de 2 (lo que significa que $1 + 1 \neq 0$ en K), resultando al igual que en los cuerpos numéricos que $f = aX^2 + bX + c$ admite raíces en K si y sólo si $\Delta = b^2 - 4ac$ es un cuadrado en K.

Por ejemplo, consideremos en $\mathbb{Z}_7[X]$ los polinomios $f_1 = X^2 - X - 2$, $f_2 = 2X^2 + 3X + 2$ y $f_3 = 3X^2 + 2X - 4$, cuyos respectivos discriminantes son $\Delta_1 = 9 \equiv 2$, $\Delta_2 = -7 \equiv 0$ y $\Delta_3 = 52 \equiv 3$, donde naturalmente $\equiv$ indica congruencia módulo 7. A partir de una simple inspección de casos, vemos fácilmente que 2 es un cuadrado en $\mathbb{Z}_7$, siendo 3 y 4 sus raíces cuadradas, y que la ecuación $x^2 \equiv 3$ no admite solución módulo 7. En consecuencia f_3 no tiene raíces en $\mathbb{Z}_7$, mientras que aplicando la fórmula obtenemos que f_1 admite dos raíces en K (2 y 6) y f_2 una sola (2). Como vemos, se presentan las mismas alternativas que en el caso real.

LA ECUACION CUBICA El hallazgo de fórmulas para resolver ecuaciones algebraicas de grados 3 y 4 se remonta al siglo XVI, a través de los trabajos de los algebristas italianos Scipione Del Ferro, Niccolo Tartaglia, Girolamo Cardano y Ludovico Ferrari. Deduciremos a continuación la fórmula para resolver la ecuación cúbica, atribuida a Tartaglia y publicada por Cardano en su obra *Ars magna*, razón por la cual se la conoce bajo la denominación de *fórmula de Cardano-Tartaglia*.

Consideremos para ello cualquier polinomio de grado 3 con coeficientes en $\mathbb{C}$, que sin pérdida de generalidad podemos suponer mónico, digamos

$$f = X^3 + aX^2 + bX + c.$$

Completando adecuadamente el cubo de un binomio logramos "des-

hacernos" del término cuadrático, ya que podemos escribir:

$$\begin{aligned} f &= (X + a/3)^3 + \left(b - a^2/3\right) X + c - a^3/27 = \\ &= (X + a/3)^3 + \left(b - a^2/3\right)(X + a/3) + 2a^3/27 - ab/3 + c = \\ &= (X + a/3)^3 + p\,(X + a/3) + q, \end{aligned}$$

donde $p = b - a^2/3$ y $q = 2a^3/27 - ab/3 + c = f\,(-a/3)$.

Designando por g el polinomio $X^3 + pX + q$, sigue entonces que u es raíz de f si y sólo si $u + a/3$ es raíz de g, ó equivalentemente, v es raíz de g si y sólo si $v - a/3$ es raíz de f. Bastará por lo tanto resolver la ecuación

$$X^3 + pX + q = 0\,, \tag{9.4}$$

en la que podemos suponer $p \neq 0$ y $q \neq 0$, ya que en el caso $p = 0$ sus soluciones son las raíces cúbicas de $-q$, mientras que si $q = 0$ es claro que ellas son 0 y las raíces cuadradas de $-p$.

A tal efecto, realizamos la sustitución

$$X = Z - \frac{p}{3Z} = h(Z)\,. \tag{9.5}$$

Reemplazando en (9.4) y operando convenientemente, es inmediato verificar entonces que se obtiene la ecuación en Z

$$Z^3 - \frac{p^3}{27Z^3} + q = 0\,, \tag{9.6}$$

lo que nos conduce, multiplicando por Z^3, a la ecuación

$$Z^6 + qZ^3 - p^3/27 = 0\,. \tag{9.7}$$

Observemos que $Z = 0$ no es solución de esta última ecuación, pues hemos supuesto $p \neq 0$, y puesto que (9.6) se obtiene de (9.7) dividiendo por Z^3, concluimos que ambas tienen las mismas soluciones.

Finalmente, mediante el cambio de variable $T = Z^3$ (9.7) se convierte en la ecuación cuadrática

$$T^2 + qT - p^3/27 = 0\,, \tag{9.8}$$

cuyas soluciones, como puede comprobar el lector, son

$$t_1 = -q/2 + \alpha \;\text{ y }\; t_2 = -q/2 - \alpha\,,$$

donde α es una raíz cuadrada fija de $(q/2)^2 + (p/3)^3$. Concluimos luego, dada la relación existente entre las variables Z y T, que la ecuación (9.7) admite 6 soluciones, a saber, las 3 raíces cúbicas de t_1 y las 3 raíces cúbicas de t_2.

Sin embargo, será preciso realizar un análisis de las mismas, ya que nuestra ecuación original (9.4) no puede tener más de 3 soluciones. Consideremos

para ello cualquier raíz cúbica γ de t_2 y elijamos una raíz cúbica β de t_1. Operando, tenemos:

$$(\beta\gamma)^3 = (-q/2+\alpha)(-q/2-\alpha) = q^2/4 - \alpha^2 = (-p/3)^3,$$

de donde sigue que $\beta\gamma = (-p/3)\zeta$, para algún $\zeta \in G_3$. Por lo tanto:

$$h(\gamma) = \gamma - p/3\gamma = -p/3\beta\zeta^{-1} + \beta\zeta^{-1} = h(\beta\zeta^{-1}),$$

vale decir, las soluciones γ y $\beta\zeta^{-1}$ de (9.7) determinan a través de la sustitución (9.5) la misma solución de (9.4). Esto prueba que sólo necesitamos las raíces cúbicas de t_1 para generar todas las soluciones de (9.4), resultando que estas son $h(\beta)$, $h(\beta\omega)$ y $h(\beta\omega^2)$, donde ω es una raíz cúbica primitiva de la unidad.

Con la intención de obtener una forma más explícita de las soluciones, que destaque por ejemplo los coeficientes de la ecuación, emplearemos por esta vez la notación

$$\beta = \sqrt[3]{-q/2 + \sqrt{q^2/4 + p^3/27}},$$

donde los radicales designan alguna raíz cuadrada o cúbica fija del radicando correspondiente. De acuerdo con nuestro desarrollo anterior, resulta entonces que las soluciones x en $\mathbb{C}$ de la ecuación $X^3 + pX + q = 0$ responden a la siguiente fórmula:

Fórmula 9.2.12 (Fórmula de Cardano-Tartaglia)

$$x = \omega^k \sqrt[3]{-q/2 + \sqrt{q^2/4 + p^3/27}} - \omega^{-k} p \left(3\sqrt[3]{-q/2 + \sqrt{q^2/4 + p^3/27}}\right)^{-1},$$

donde ω es una raíz cúbica primitiva de la unidad y $0 \leq k \leq 2$. $\Diamond$

Ejemplo 9.2.13 Calculemos las raíces de $f = X^3 - 6X^2 + 10X - 8$. Completando el cubo resulta que $f = (X-2)^3 - 2(X-2) - 4$, y en consecuencia las raíces de f son de la forma $z+2$, donde z es raíz de $g = X^3 - 2X - 4$. Empleando la notación del caso general, tenemos en este caso $-q/2 = 2$ y $q^2/4 + p^3/27 = 100/27$, de donde sigue que

$$\beta = \sqrt[3]{2 + \sqrt{\frac{100}{27}}} = \sqrt[3]{2 + \frac{10}{9}\sqrt{3}} = \frac{3+\sqrt{3}}{3},$$

como puede verificarse. Luego, tomando $\omega = -1/2 + (\sqrt{3}/2)\imath$, y aplicando

la fórmula de Cardano-Tartaglia obtenemos que las raíces de g son:

$$z_1 = \frac{3+\sqrt{3}}{3} + \frac{2}{3+\sqrt{3}} = \frac{3+\sqrt{3}}{3} + \frac{3-\sqrt{3}}{3} = 2$$

$$z_2 = \left(\frac{-1+\sqrt{3}\,\imath}{2}\right)\left(\frac{3+\sqrt{3}}{3}\right) - \left(\frac{1+\sqrt{3}\,\imath}{2}\right)\left(\frac{3-\sqrt{3}}{3}\right) = -1+\imath$$

$$z_3 = \left(\frac{-1-\sqrt{3}\,\imath}{2}\right)\left(\frac{3+\sqrt{3}}{3}\right) + \left(\frac{-1+\sqrt{3}\,\imath}{2}\right)\left(\frac{3-\sqrt{3}}{3}\right) = -1-\imath\,.$$

Dada la relación existente entre las raíces de f y g que hemos señalado arriba, resulta finalmente que las raíces de f son 4, $1+\imath$ y $1-\imath$. $\Diamond$

RAICES RACIONALES. En los ejemplos 9.2.2 de la presente sección vimos que un polinomio f con coeficientes racionales puede admitir o no raíces racionales. Mostraremos en este apartado que es posible determinar *a priori* un número finito de valores entre los cuales deberán buscarse las posibles raíces racionales de f. Técnicamente hablando, brindaremos una condición necesaria −llamada *criterio de Gauss*− para que un número racional sea raíz de un polinomio con coeficientes racionales. En realidad, podemos siempre suponer que f tiene coeficientes enteros, ya que si m es el mínimo común múltiplo de los denominadores de sus coeficientes resulta que $mf \in \mathbb{Z}[X]$ y tiene las mismas raíces que f. Hecha esta reducción, enunciemos y demostremos el mencionado criterio.

Proposición 9.2.14 (Criterio de Gauss) Sea

$$f = \sum_{i=0}^{n} a_i X^i \in \mathbb{Z}[X]$$

y sea r/s una raíz racional de f, con r y s enteros coprimos. Entonces $r \mid a_0$ y $s \mid a_n$.

DEMOSTRACION. Multiplicando por s^n la igualdad $f(r/s) = 0$, obtenemos

$$0 = \sum_{i=0}^{n} a_i r^i s^{n-i} = \sum_{i=0}^{n} a_{n-i} r^{n-i} s^i = a_n r^n + s\left(\sum_{i=1}^{n} a_{n-i} r^{n-i} s^{i-1}\right),$$

de donde deducimos que $s \mid a_n$, pues $s \mid a_n r^n$ y es coprimo con r^n. Similarmente, reescribiendo la primera de las igualdades de arriba tenemos:

$$0 = \sum_{i=0}^{n} a_i r^i s^{n-i} = a_0 s^n + \sum_{i=1}^{n} a_i r^i s^{n-i} = a_0 s^n + r\left(\sum_{i=1}^{n} a_i r^{i-1} s^{n-i}\right),$$

esto es, $r \mid a_0 s^n$. Puesto que r es coprimo con s^n concluimos que $r \mid a_0$, como queríamos probar. $\Diamond$

Corolario 9.2.15 Sea f un polinomio mónico con coeficientes en $\mathbb{Z}$. Entonces toda raíz racional de f es entera.

Demostracion. Inmediata. $\Diamond$

Ejemplos 9.2.16 Apliquemos lo anterior para hallar las posibles raíces en $\mathbb{Q}$ del polinomio $f = X^5 + 7/6X^4 - 5/2X^3 - 4/3X^2 + 2/3X$, una de las cuales obviamente es 0. Multiplicando por 6 resulta

$$6f = 6X^5 + 7X^4 - 15X^3 - 8X^2 + 4X = X(6X^4 + 7X^3 - 15X^2 - 8X + 4),$$

por lo que bastará hallar las raíces racionales de

$$g = 6X^4 + 7X^3 - 15X^2 - 8X + 4,$$

que tiene coeficientes enteros.

De acuerdo con el criterio de Gauss, y conservando la notación de 9.2.14, los valores admisibles de r son los divisores de 4, esto es, ± 1, ± 2 y ± 4, y los de s los divisores positivos de 6, a saber 1, 2, 3 y 6. Por lo tanto, y recordando la hipótesis de coprimalidad, las posibles raíces racionales de g pertenecen al conjunto

$$\{\pm 1, \pm 2, \pm 4, \pm 1/2, \pm 1/3, \pm 2/3, \pm 4/3, \pm 1/6\}.$$

Una simple verificación nos muestra que de estos 16 valores solo -2 y $1/3$ son raíces de g. Luego, f admite 3 raíces racionales, a saber, 0, -2 y $1/3$.

Acotemos que un uso adecuado de la teoría permite aliviar los cálculos. Por ejemplo (ver proposición 9.2.6), el hecho de que -2 sea raíz de g determina la factorización

$$g = (X + 2)(6X^3 - 5X^2 - 5X + 2).$$

Designando por h el segundo factor del miembro de la derecha en la igualdad anterior, sigue que las restantes raíces racionales de g (si existen) son raíces de h. Aplicando nuevamente Gauss vemos que los posibles valores de estas son $\pm 1, \pm 2, \pm 1/2, \pm 1/3, \pm 2/3$ y $\pm 1/6$, vale decir, hemos descartado algunos casos. Por supuesto, en cuanto hayamos determinado que $1/3$ es raíz de h podemos repetir el procedimiento. En tal caso, obtendríamos la descomposición

$$g = (X + 2)(X - 1/3)(6X^2 - 3X - 6).$$

Puesto que el último factor de arriba es de grado 2 es más sencillo calcular directamente sus ceros a través de la fórmula, resultando que ninguna de ellas es racional, lo que corrobora que -2 y $1/3$ son las únicas raíces de g en $\mathbb{Q}$.

Como ejemplo más general, consideremos el polinomio $f = X^3 - X + p$, donde p es un número primo. De acuerdo con el criterio de Gauss sus posibles raíces racionales (en realidad enteras) son ± 1 y $\pm p$, pero un rápido cálculo nos muestra que ninguno de estos números anula a f. En efecto, es claro que 1, -1 y p no son raíces de f, pues $f(1) = f(-1) = p$ y $f(p) = p^3$. En cuanto al último valor, tenemos $f(-p) = -p^3 + 2p = p\,(2 - p^2) \neq 0$, ya que obviamente $p^2 \neq 2$. Luego f no tiene raíces racionales. $\Diamond$

RAICES DE POLINOMIOS CON COEFICIENTES REALES. Los polinomios con coeficientes reales admiten cierto tratamiento especial en conexión con sus raíces, debido a que las funciones polinómicas asociadas con ellos por especialización son funciones reales *continuas*, y podemos estudiarlas entonces aplicando recursos del análisis matemático. A manera de ejemplo, mencionemos el hecho de que toda ecuación cuadrática con coeficientes reales y discriminante positivo admite soluciones reales, resultado que se deriva del axioma de completitud a través de la existencia de raíces cuadradas, y que no es válido para ecuaciones cuadráticas con coeficientes racionales. Nos dedicaremos ahora a mostrar otra importante situación en la que puede asegurarse la existencia de soluciones reales, y que es consecuencia de un teorema básico del cálculo.

Teorema 9.2.17 Todo polinomio f de grado impar con coeficientes reales admite al menos una raíz real.

DEMOSTRACION. Supondremos sin pérdida de generalidad que f es mónico, digamos

$$f = X^n + \sum_{i=1}^{n} a_{n-i} X^{n-i},$$

con n impar. Para demostrar el enunciado, bastará hallar $u, v \in \mathbb{R}\,(u < v)$ tales que $f(u) < 0$ y $f(v) > 0$, ya que entonces el *teorema de Bolzano* aplicado a la función continua f nos asegurará la existencia de un número real $\theta \in (u, v)$ tal que $f(\theta) = 0$.

A tal efecto, sea $c = \text{máx}\,\{\,|\,a_i\,|\,:\,0 \leq i < n\}$ y sea x cualquier número real tal que $|x| > \text{máx}\,\{1, cn\}$. Operando, obtenemos:

$$f(x) = x^n + \sum_{i=1}^{n} a_{n-i} x^{n-i} = x^n \left(1 + \sum_{i=1}^{n} \frac{a_{n-i}}{x^i}\right),$$

ó equivalentemente:

$$\frac{f(x)}{x^n} = 1 + \sum_{i=1}^{n} \frac{a_{n-i}}{x^i}\,. \tag{9.9}$$

Por otro lado, teniendo en cuenta la elección de x, resulta:

$$\left| \sum_{i=1}^{n} \frac{a_{n-i}}{x^i} \right| \leq \sum_{i=1}^{n} \frac{|\, a_{n-i} \,|}{|\, x \,|^i} \leq \sum_{i=1}^{n} \frac{|\, a_{n-i} \,|}{|\, x \,|} \leq \frac{cn}{|\, x \,|} < 1\,,$$

de donde sigue que

$$1 + \sum_{i=1}^{n} \frac{a_{n-i}}{x^i} > 0\,. \tag{9.10}$$

Deducimos entonces de (9.9) y (9.10) que $f(x)$ y x^n tienen el mismo signo para tales valores de x, en cuyo caso también coinciden los signos de $f(x)$ y x, pues n es impar. Luego, para obtener las condiciones requeridas bastará tomar por ejemplo $v = \text{máx}\,\{1, cn\} + 1$ y $u = -v$. $\Diamond$

En el próximo capítulo estaremos en condiciones de establecer un resultado más fuerte que el anterior, que afirma que todo polinomio de grado impar con coeficientes en $\mathbb{R}$ tiene un número impar de raíces reales. Probaremos por ahora otro hecho −conectado con el que acabamos de mencionar− acerca de las raíces complejas de un polinomio con coeficientes reales.

Proposición 9.2.18 Sea $f \in \mathbb{R}[X]$ y sea z una raíz de f en $\mathbb{C}$. Entonces $\overline{z}$ también es raíz de f.

Demostracion. Obviamente, el resultado tiene interés si z no es real. Notemos además que no vale la misma afirmación si f no tiene coeficientes reales, ya que por ejemplo $\imath$ es raíz de $X - \imath$ pero $\overline{\imath} = -\imath$ no lo es. Para demostrar la proposición escribamos

$$f = \sum_k c_k X^k,$$

donde $c_k \in \mathbb{R}$ para todo k. Empleando la notación $\overline{w} = \sigma(w)$ y tomando conjugados en la igualdad $0 = f(z)$ obtenemos, usando las propiedades de la conjugación:

$$\begin{aligned} 0 = \sigma(0) = \sigma(\sum_k c_k\, z^k) = \sum_k \sigma(c_k\, z^k) = \\ = \sum_k \sigma(c_k)\; \sigma(z)^k = \sum_k c_k \sigma(z)^k = f\,(\sigma(z))\,, \end{aligned}$$

como queríamos demostrar. $\Diamond$

Ejemplo 9.2.19 Calculemos todas las raíces en $\mathbb{C}$ del polinomio

$$f = X^5 + 5X^4 - 12X^3 + 20X^2 + 27X - 105,$$

sabiendo que una de ellas es $1-2\imath$. De acuerdo con 9.2.18 sigue que $1+2\imath$ es otra de las raíces de f, lo que nos permite obtener la factorización

$$f = (X-(1+2\imath))\,(X-(1-2\imath))\left(X^3+7X^2-3X-21\right),$$

como el lector se encargará de comprobar. Puesto que el factor

$$g = X^3+7X^2-3X-21$$

tiene coeficientes en $\mathbb{Z}$ podemos emplear el criterio de Gauss para determinar si admite alguna raíz racional, que debiera ser en este caso un divisor entero de 21. Una rápida verificación nos muestra que -7 es su única raíz en $\mathbb{Q}$, lo que conduce a la descomposición $g = (X+7)\left(X^2-3\right)$. Luego, las 5 raíces complejas de f son -7, $\sqrt{3}$, $-\sqrt{3}$, $1+2\imath$ y $1-2\imath$. $\quad\Diamond$

9.2.2. Ejercicios

1. *a*) Sea $f \in \mathbb{Z}[X]$ y sea m un número natural que no es un cuadrado perfecto. Si $f(\sqrt{m}) \in \mathbb{Q}$ probar que $f(\sqrt{m}) = f(-\sqrt{m})$
   b) Hallar en un polinomio mónico con coeficientes enteros g, de grado 3, tal que $g(\sqrt{2}) = 3$ y $g(-1) = 7$
   c) Hallar en $\mathbb{Z}[X]$ un polinomio h con coeficiente principal 2 tal que $h\left(\sqrt{3}\right) = h\,(1-\imath) = 0$.

2. Sea $z \in \mathbb{C}$ y sea $n \in \mathbb{N}$. Probar las siguientes igualdades polinomiales:

   a) $X^n - z^n = (X-z)\displaystyle\sum_{i=0}^{n-1} z^i X^{n-1-i}$

   b) $X^n + z^n = (X+z)\displaystyle\sum_{i=0}^{n-1} (-z)^i X^{n-1-i}$ si n es impar

   c) $X^n - z^n = (X+z)\displaystyle\sum_{i=0}^{n-1} (-z)^i X^{n-1-i}$ si n es par.

3. Hallar las raíces complejas de los siguientes polinomios:

   a) X^3+2X^2+4X+8
   b) $X^4-3X^3+9X^2-27X+81$
   c) X^3-2X^2+4X-8.

4. Sea $f \in \mathbb{Z}[X]$ tal que $f(a)$ es primo para todo entero a. Probar que f es constante.

5. *a*) Hallar $f \in \mathbb{Q}[X]$, de grado 3, tal que $f(1) = 2$, $f(-1) = 0$, $f(2) = 1$ y $f(3) = 1$
   b) ¿Existe $f \in \mathbb{Z}[X]$, de grado 3, tal que $f(\sqrt{2}) = 6\sqrt{2}$, $f(\imath) = 21$ y $f(-1) = 3$?
   c) ¿Existe $f \in \mathbb{R}[X]$, de grado 2, tal que $f(1) = -1$, $f(2) = -2$ y $f(\imath) = -7$?
   d) Hallar $f \in \mathbb{R}[X]$, de grado mínimo, tal que $f(1) = 0$, $f(3) = 18$, $f(5) = 52$ y $f(7) = 102$.

6. Sea $f \in \mathbb{R}[X]$ de grado n y sean $x_1, \ldots, x_{n+1}$ números racionales distintos tales que $f(x_i) \in \mathbb{Q}$ para todo i. Probar que $f \in \mathbb{Q}[X]$. Demostrar que vale un resultado análogo reemplazando $\mathbb{R}$ por $\mathbb{C}$ y $\mathbb{Q}$ por $\mathbb{R}$.

7. Demostrar que los números $\sqrt{2}-\sqrt{7}$, $3+\sqrt{2}\,\imath$, $\sqrt[3]{2}-\sqrt{2}$ y $\sqrt{3\sqrt{5}-\sqrt{8}}$ son algebraicos.

8. Sea ρ un número real trascendente. Demostrar que $\sqrt{|\,\rho\,|}$ es trascendente y que ρ^m es trascendente para todo entero no nulo m.

9. Hallar las raíces complejas de los siguientes polinomios:
   a) $X^4 - 4X^2 + 16$
   b) $X^4 - X^3 - X^2 - X - 2$
   c) $X^3 - 2(1+\sqrt{3})X^2 + 2(1+2\sqrt{3})X - 4\sqrt{3}$
   d) $X^6 - 4X^4 - 41X^2 - 36$
   e) $X^5 + X^4 + 2X^3 + 2X^2 - 8X - 8$.

10. Hallar las raíces racionales de los siguientes polinomios:
    a) $X^5 - X^4 - 8X^3 + 2X^2 + 12X$
    b) $(X+1)(X-5)(X^3 + 2X^2 + 2X + 1)$
    c) $X^5 + (23/6) + (1/6)X^3 + (19/6)X^2 - (5/6)X - 2/3$
    d) $1 + X + \cdots + X^n$ $(n \in \mathbb{N})$.

11. Analizar en qué casos el polinomio $X^3 + pX - q$ (p y q primos) admite una raíz racional.

12. Sea $f \in \mathbb{Z}[X]$ y sea a/b una raíz racional de f, con a y b enteros coprimos. Probar que $a - b\,n \mid f(n)$ para todo $n \in \mathbb{Z}$.

13. Sea $f = 3X^5 + X^4 - 30X^3 - 19X^2 + 51X + 18$.

a) De acuerdo con el criterio de Gauss, determinar las posibles raíces racionales de f

b) Determinar cuáles de los valores hallados en *a*) satisfacen la condición demostrada en el ejercicio precedente para $n = 1, 2$ y -1

c) Hallar las raíces racionales de f.

Capítulo 10

Divisibilidad en anillos de polinomios

10.1. Divisibilidad

10.1.1. Divisores y múltiplos

La teoría elemental de divisibilidad de números enteros que hemos estudiado en los capítulos 5 y 6 puede desarrollarse de manera muy similar en los anillos de polinomios con coeficientes en un cuerpo arbitrario K, obteniéndose resultados esencialmente idénticos a los obtenidos en el caso de $\mathbb{Z}$. Comenzaremos definiendo la noción de divisibilidad en $K[X]$:

> Si f y g son polinomios con coeficientes en K decimos que f *divide* a g si y sólo si existe $h \in K[X]$ tal que $g = fh$.

Como en el caso numérico emplearemos las notación $f \mid g$ para indicar que f divide a g, y también nos referiremos a la situación diciendo que f es *divisor* de g o que g es *múltiplo* de f. Si f no divide a g escribiremos $f \nmid g$.

Ejemplos 10.1.1 $X^2 + (\sqrt{2}-1)X - \sqrt{2}$ divide a $X^4 + \sqrt{2}X^3 - X - \sqrt{2}$ en $\mathbb{R}[X]$, ya que vale la igualdad

$$X^4 + \sqrt{2}X^3 - X - \sqrt{2} = \left(X^2 + (\sqrt{2}-1)X - \sqrt{2}\right)\left(X^2 + X + 1\right),$$

como es fácil de verificar. Análogamente, $2X^4 + 2X^3 - 11X^2 + X - 6$ es múltiplo de $2X^2 - 2X - 12$ en $\mathbb{Q}[X]$, ya que un simple cálculo nos muestra que

$$2X^4 + 2X^3 - 11X^2 + X - 6 = (2X^2 + 2X - 12)(X^2 + 1/2). \quad \Diamond$$

NOTA. El factor h de la definición de arriba es único si $f \neq 0$. En efecto, supongamos que $g = fh = fh_1$. Tenemos entonces que $f(h - h_1) = 0$, y

siendo $K[X]$ un dominio de integridad, resulta que $h - h_1 = 0$, esto es, $h_1 = h$. Observemos de paso que obtendríamos la misma conclusión si sólo exigiéramos que el anillo de coeficientes fuese un dominio de integridad.

En realidad, en la definición de divisibilidad no es esencial que K sea un cuerpo, y de idéntica manera puede definirse dicha noción para polinomios con coeficientes en cualquier anillo conmutativo. Sin embargo, la teoría es bastante más complicada cuando el anillo de coeficientes no es un cuerpo, y este es el motivo por el cual hemos optado por trabajar primordialmente en la situación referida.

Señalemos por último que en el segundo de los ejemplos de 10.1.1, en el que ambos polinomios tienen coeficientes enteros, la relación de divisibilidad que mostramos es válida en $\mathbb{Q}[X]$ pero no en $\mathbb{Z}[X]$, ya que el factor de divisibilidad tiene coeficientes racionales pero no enteros. $\Diamond$

Enumeraremos a continuación las propiedades básicas de la divisibilidad de polinomios. No ofreceremos todas las demostraciones, ya que casi todas ellas se obtienen usando propiedades válidas en cualquier anillo conmutativo, y son prácticamente idénticas a las que ya brindamos en el caso de $\mathbb{Z}$. De todos modos, nos ocuparemos en señalar algunas diferencias de forma existentes entre ambas nociones de divisibilidad. Salvo aclaración expresa, las letras designan polinomios.

Proposición 10.1.2 Valen en $K[X]$ las siguientes propiedades:

1) La relación de divisibilidad en $K[X]$ es reflexiva y transitiva

2) $f \mid 0$ para todo f y $0 \mid g \Leftrightarrow g = 0$

3) $a \mid f$ para todo $a \in K^*$

4) $f \mid g \Leftrightarrow af \mid bg$ cualesquiera sean $a, b \in K^*$. Resulta en particular que $cf \mid f$ para todo $c \in K^*$

5) Si $f \mid g$ y $g \neq 0$ entonces $gr(f) \leq gr(g)$

6) $f \mid g$ y $g \mid f$ si y sólo si existe $c \in K^*$ tal que $g = cf$

7) Sea $n \in \mathbb{N}$ y supongamos que $f \mid g_i$ para $i = 1, 2, \ldots, n$. Entonces

$$f \mid \sum_{i=1}^{n} h_i g_i$$

cualesquiera sean $h_1, h_2, \ldots, h_n$

8) Conservando la notación de 7), supongamos que

$$f \mid \sum_{i=1}^{n} g_i$$

y $f \mid g_i$ para todo $i < n$. Entonces $f \mid g_n$.

DEMOSTRACION. Las propiedades 1), 2), 7) y 8) se demuestran como en el caso de $\mathbb{Z}$, por lo que nos dedicaremos a las restantes. Comenzando por 3), notemos simplemente que $f = a(a^{-1}f)$. Respecto de 4), escribiendo $g = fh$ obtenemos $bg = bfh = af(ba^{-1}h)$, de donde sigue que $af \mid bg$, mientras que la recíproca es un caso particular de lo que acabamos de probar, ya que $f = a^{-1}(af)$ y $g = b^{-1}(bg)$. La segunda afirmación resulta de tomar $g = f$, $a = c$ y $b = 1$. Para probar 5), volvamos a escribir $g = fh$ y observemos que por hipótesis todos los elementos involucrados son no nulos. Luego, tomando grados resulta:

$$gr(g) = gr(fh) = gr(f) + gr(h) \geq gr(f).$$

En cuanto a 6), supongamos primero que f y g se dividen mutuamente. Sigue entonces de 2) que $f = 0 \Leftrightarrow g = 0$, en cuyo caso basta tomar $c = 1$. Si ambos son no nulos, usando la propiedad 5) concluimos que $gr(f) = gr(g)$, y por lo tanto, en una relación del tipo $g = fh$ el factor h debe ser de grado 0, como queríamos demostrar. Finalmente, notemos que $g = cf$ si y sólo si $f = c^{-1}g$, por lo que también vale la recíproca. $\Diamond$

NOTA Repasando las propiedades anteriores, y suponiendo que el cuerpo K es infinito, descubrimos algunas diferencias entre la divisibilidad en $\mathbb{Z}$ y la divisibilidad en $K[X]$. Por ejemplo, no es válido en $K[X]$ que todo elemento no nulo admite sólo un número finito de divisores, ya que todo polinomio de grado 0 divide a cualquier polinomio, propiedad que sólo satisfacen en $\mathbb{Z}$ los números 1 y -1. Más aún, a la lista de divisores de cualquier polinomio f debemos agregarle el conjunto −también infinito−, de todos los polinomios de la forma af, donde $a \in K^*$. Para disponer de un lenguaje cómodo, diremos que af es un *asociado* de f. Observemos que todo polinomio no nulo tiene infinitos asociados, y que de acuerdo con 6) dos polinomios son asociados si y sólo si se dividen mutuamente.

Sin embargo, estas diferencias son meramente cuantitativas, ya que los hechos obedecen a una misma razón. En efecto, las constantes no nulas juegan el mismo rol que 1 y -1 en $\mathbb{Z}$ debido a que ellas son las *unidades* del anillo $K[X]$ (elementos inversibles con respecto al producto), mientras que 1 y -1 son las unidades del anillo $\mathbb{Z}$. La siguiente descripción permite apreciar aún más la analogía: todo elemento f de $K[X]$ admite entre sus divisores a las unidades (constantes no nulas) y a sus asociados (productos de f por unidades), de la misma forma que todo entero m es divisible por las unidades (1 y -1) y por sus asociados (m y $-m$).

Notemos también que con respecto a la divisibilidad el grado cumple en $K[X]$ un rol similar al del valor absoluto en $\mathbb{Z}$, y que por la propiedad 4) es suficiente estudiar relaciones de divisibilidad entre polinomios mónicos, ya que todo polinomio no nulo f puede escribirse en la forma $f = ag$, donde a es el coeficiente principal de f y g es mónico.

Por ejemplo, determinemos los divisores con coeficientes racionales del polinomio $f = 2X^4 - 2X^2 - 12 = 2(X^4 - X^2 - 6)$, que de acuerdo con lo dicho coinciden los divisores con coeficientes racionales de $g = X^4 - X^2 - 6$.

Si $g = ht$ es una factorización no trivial de g (ninguno de los factores es constante), sigue por razones de grado que ambos son de grado 2 o bien uno de ellos tiene grado 1 y el otro grado 3. Ahora bien, una sencilla aplicación del criterio de Gauss nos muestra que g no tiene raíces racionales, por lo que g no puede tener un factor de grado 1 con coeficientes racionales, ya que en tal caso este tendría una raíz en $\mathbb{Q}$, que también sería raíz de g. Luego, suponiendo h y t mónicos, la factorización debe ser de la forma:

$$g = (X^2 + aX + b)(X^2 + cX + d),$$

donde $a, b, c, d \in \mathbb{Q}$. Expandiendo el segundo miembro e igualando coeficientes, arribamos al sistema de ecuaciones

$$\begin{cases} a + c = 0 \\ b + ac + d = -1 \\ ad + bc = 0 \\ bd = -6\,. \end{cases}$$

Sigue de la primera ecuación que $c = -a$, y reemplazando en la tercera obtenemos $a(d - b) = 0$. Puesto que $d = b$ es imposible, ya que en tal caso tendríamos $b^2 = -6$, concluimos que $a = c = 0$. Multiplicando ahora ambos miembros de la segunda ecuación por b tenemos que $b^2 - 6 = -b$, ó equivalentemente, $b^2 + b - 6 = 0$. Resolviendo esta ecuación cuadrática hallamos las soluciones $b = 2$ y $b = -3$, que corresponden a $c = -3$ y $c = 2$. Puesto que no importa el orden de los factores, resulta que la única factorización hallada es $g = (X^2 + 2)(X^2 - 3)$.

En definitiva, los divisores de f en $\mathbb{Q}[X]$ son sus asociados, los asociados de $X^2 + 2$, los asociados de $X^2 - 3$ y las constantes no nulas.

10.1.2. Algoritmo de división

Al igual que en $\mathbb{Z}$, también existe un *algoritmo de división* en $K[X]$, que entre otras cosas permite decidir si un polinomio es múltiplo de otro. Precisaremos los detalles a través del siguiente enunciado, en el que se podrá apreciar la similitud de la situación respecto del caso entero, con el grado cumpliendo un rol equivalente al del valor absoluto:

Teorema 10.1.3 (Algoritmo de división de polinomios) Sean f y g polinomios con coeficientes en K ($f \neq 0$). Existe entonces un único par de polinomios h y r en $K[X]$, llamados respectivamente el *cociente* y el *resto* de la división de g por f, satisfaciendo las siguientes condiciones:

$$\begin{aligned} &DP_1) \quad g = hf + r \\ &DP_2) \quad r = 0 \text{ ó } gr(r) < gr(f)\,. \end{aligned}$$

DEMOSTRACION. Probaremos en primer término la existencia de elementos h y r en $K[X]$ verificando las condiciones DP_1) y DP_2). Si $gr(g) < gr(f)$ ó $g = 0$ el resultado es inmediato, ya que en tales casos basta tomar $h = 0$ y $r = g$. Supongamos pues que $m = gr(f) \leq gr(g) = n$ y procedamos por inducción fuerte en n. El caso $m = n = 0$ es trivial, ya que si f y g son constantes no nulas, digamos $f = b$ y $g = a$, basta tomar $h = a/b$ y $r = 0$ (observemos que $a/b \in K$ por ser K un cuerpo).

Asumamos ahora que $n > 0$ y utilicemos las notaciones

$$f = \sum_{i=0}^{m} b_i X^i \quad \text{y} \quad g = \sum_{i=0}^{n} a_i X^i.$$

Podemos escribir entonces:

$$g = (a_n/b_m)\, X^{n-m} f + g_1\,, \tag{10.1}$$

donde $g_1 = g - (a_n/b_m)\, X^{n-m} f$. Notemos además que $g_1 = 0$ ó $gr(g_1) < n$, pues $(a_n/b_m)\, X^{n-m} f$ es de grado n y tiene coeficiente principal a_n, al igual que g.

Razonamos ahora de la siguiente manera. Si $g_1 = 0$ ó $gr(g_1) < m$, obtenemos el resultado tomando $h = (a_n/b_m)\, X^{n-m}$ y $r = g_1$ mientras que si $m \leq gr(g_1) \leq n-1$ sigue por un argumento inductivo (con g_1 en lugar de g) que existen polinomios t y s en $K[X]$ ($s = 0$ ó $gr(s) < gr(f)$) tales que $g_1 = tf + s$. Luego, reemplazando en 10.1 obtenemos:

$$\begin{aligned} g &= (a_n/b_m)\, X^{n-m} f + g_1 = (a_n/b_m)\, X^{n-m} f + tf + s = \\ &= \big((a_n/b_m)\, X^{n-m} + t\big)\, f + s\,. \end{aligned}$$

Tomando $h = (a_n/b_m)\, X^{n-m} + t$ y $r = s$ llegamos entonces al resultado deseado.

Para probar la unicidad, consideremos cualquier otro par (h_1, r_1) de elementos de $K[X]$ satisfaciendo las condiciones DP_1) y DP_2) del enunciado, y supongamos que $r_1 \neq r$. Restando las correspondientes igualdades, tenemos:

$$0 = g - g = (h_1 f + r_1) - (hf + r) = (h_1 - h)f + (r_1 - r)\,,$$

de donde

$$r - r_1 = (h_1 - h)f\,,$$

esto es, f divide a $r - r_1$. Siendo $r - r_1 \neq 0$, obtenemos por 5) de la proposición 10.1.2 las desigualdades:

$$gr(f) \leq gr(r - r_1) \leq \max\{gr(r), gr(r_1)\} < gr(f),$$

lo que es absurdo. Luego $r_1 = r$ y por lo tanto $h_1 = h$, ya que $f \neq 0$ y $K[X]$ es un dominio de integridad.

Por razones de simplicidad hemos asumido en la demostración anterior que ambos términos de la diferencia $r - r_1$ son distintos de cero. El lector no tendrá dificultad en probar que la desigualdad $gr(r - r_1) < gr(f)$ también es válida en el caso de que alguno de ellos sea nulo. $\Diamond$

Para demostrar formalmente la existencia de cociente y resto hemos procedido por inducción, pero es importante que el lector no pierda de vista el proceso iterativo que conduce a ellos, que seguramente aprendió en sus estudios secundarios. En el primer paso se obtiene un monomio M_1 de grado $n - m$ cuyo coeficiente principal es el cociente de los coeficientes principales de g y f. Restamos entonces $M_1 f$ de g, repetimos el proceso anterior con $g_1 = g - M_1 f$ en lugar de g, y así siguiendo. Vamos determinando de tal modo una secuencia $M_1, M_2, \ldots$ de monomios y una secuencia $g_0, g_1, g_2, \ldots$ de polinomios ($g_0 = g$), definidos por la relaciones

$$M_{i+1} = (c_i/b_m)\, X^{gr(g_i)-gr(f)} \quad \text{y} \quad g_{i+1} = g_i - M_{i+1} f$$

para todo $i \geq 0$, donde c_j denota el coeficiente principal de g_j. Puesto que la secuencia de los grados de los polinomios g_i es estrictamente decreciente, es claro que existe un mínimo entero $k > 0$ tal que $g_k = 0$ ó $gr(g_k) < gr(f)$. Resulta entonces que g_k es el resto y $M_1 + M_2 + \ldots + M_k$ es el cociente de la división de g por f.

Ejemplo 10.1.4 Para ilustrar el proceso de división, tomemos en $\mathbb{Q}[X]$ $g = 4X^7 - X^6 + x^5 + 3X^4 + X^3 + 3X - 5$ y $f = 2X^3 - 2X^2 + 2X + 1$. Siguiendo los pasos del algoritmo, obtenemos sucesivamente:

$$\begin{aligned}
(M_1, g_1) &= \left(2X^4\,,\, 3X^6 - 3X^5 + X^4 + X^3 + 3X - 5\right) \\
(M_2, g_2) &= \left(3/2X^3\,,\, -2X^4 - 1/2X^3 + 3X - 5\right) \\
(M_3, g_3) &= \left(-X\,,\, -5/2X^3 + 2X^2 + 4X - 5\right) \\
(M_4, g_4) &= \left(-5/4\,,\, -1/2X^2 + 13/2X - 15/4\right)\,.
\end{aligned}$$

Luego el cociente de la división de g por f es $2X^4 + 3/2X^3 - X - 5/4$ y el resto es $-1/2X^2 + 13/2X - 15/4\,.$ $\Diamond$

NOTA. El ejemplo que acabamos de desarrollar nos muestra que la hipótesis de que K sea un cuerpo es esencial para la existencia de algoritmo de división en $K[X]$, y que éste no se extiende a polinomios con coeficientes en un anillo A cualquiera. Así, en el caso anterior el cociente y el resto no pertenecen a $\mathbb{Z}[X]$, a pesar de que f y g tienen coeficientes enteros. Sin embargo, podemos asegurar en $\mathbb{Z}[X]$ la existencia de un *algoritmo de división por polinomios mónicos*, habida cuenta de que en las sucesivas etapas del algoritmo se divide siempre por el coeficiente principal de f. Por ejemplo, el lector puede verificar que el cociente y el resto de dividir $3X^5 - X^2 + X - 1$ por $X^2 + X - 4$ son $3X^3 - 3X^2 + 15X - 28$ y $89X - 113$, respectivamente.

Volviendo a la situación general, veamos que la división de g por f en $K[X]$ puede siempre obtenerse dividiendo por un polinomio mónico. En efecto, escribamos $f = bf^*$, donde b es el coeficiente principal de f y f^* es mónico. Entonces, si h_1 y r_1 son el cociente y el resto de dividir g por f^*, tenemos:

$$g = h_1 f^* + r_1 = (b^{-1}h_1)(bf^*) + r_1 = (b^{-1}h_1)f + r_1.$$

Puesto que $b^{-1}h_1$ y r_1 satisfacen las condiciones DP_1) y DP_2), concluimos que ellos son el cociente y el resto de dividir g por f. $\Diamond$

Ejemplo 10.1.5 Para ilustrar un caso importante, dividamos en $\mathbb{C}[X]$ el polinomio $g = X^4 + (1-2\imath)X^3 - (3+2\imath)X^2 - (1-4\imath)X + 2$ por el polinomio $f = 2X^2 + 2X - 4 = 2(X^2 + X - 2)$.

Aplicando el recurso anterior dividimos g por $f^* = X^2 + X - 2$, y obtenemos sucesivamente:

$$\begin{aligned}(M_1, g_1) &= \left(X^2\,,\, -2\imath X^3 - (1+2\imath)X^2 - (1-4\imath)X + 2\right)\\ (M_2, g_2) &= \left(-2\imath X\,,\, -X^2 - X + 2\right)\\ (M_3, g_3) &= (-1\,,\, 0)\ .\end{aligned}$$

Por lo tanto el resto de la división de g por f es nulo, mientras que el cociente es

$$1/2\,(X^2 - 2\imath X - 1) = 1/2X^2 - \imath X - 1/2\,. \qquad \Diamond$$

Como en el caso de $\mathbb{Z}$, el algoritmo de división en $K[X]$ brinda un mecanismo efectivo para decidir si un cierto polinomio es múltiplo de otro. Precisamente, tenemos:

Corolario 10.1.6 Sean $f, g \in K[X]$ $(f \neq 0)$. Entonces $f \mid g$ si y sólo si el resto de dividir g por f es cero.

Demostracion. Es idéntica a la del caso entero, por lo que dejamos los detalles a cargo del lector. $\Diamond$

Por ejemplo, $g = X^4 + (1-2\imath)X^3 - (3+2\imath)X^2 - (1-4\imath)X + 2$ es múltiplo de $f = 2X^2 + 2X - 4$, pues vimos arriba que el resto de la división de g por f es cero, siendo el cociente de la división el correspondiente factor de divisibilidad. Esto es:

$$g = \left(1/2X^2 - \imath X - 1/2\right) f\,.$$

Nota La analogía que venimos señalando entre $\mathbb{Z}$ y $K[X]$ se extiende a la noción de congruencia, que también podemos definir en el anillo de polinomios. Así, dados f, g y h en $K[X]$, decimos que g y h son *congruentes*

módulo f si y solo si g y h difieren en un múltiplo de f. Empleando las mismas notaciones que en el caso entero, tenemos entonces que

$$g \equiv h\ (f) \Leftrightarrow f \mid h - g\,.$$

Es inmediato probar que $\equiv$ es una relacion de equivalencia en $K[X]$, que al igual que en $\mathbb{Z}$ puede describirse también en términos de restos. Precisamente,

$$g \equiv h\ (f) \Leftrightarrow r_f(g) = r_f(h),$$

donde, en general, $r_f(l)$ designa el resto de dividir un polinomio l por f.

Dejaremos a cargo del lector la tarea de demostrar que la congruencia de polinomios satisface las mismas propiedades que la congruencia entera, en particular aquellas referidas a su comportamiento respecto a las operaciones, propiedades que usaremos de aquí en adelante cuando sea necesario. $\Diamond$

10.1.3. Máximo común divisor

Prosiguiendo con el desarrollo de la teoría de la divisibilidad en el anillo de polinomios con coeficientes en un cuerpo K, nos ocuparemos ahora del concepto de máximo común divisor en $K[X]$. La idea es enteramente similar a la del caso entero, salvo que aquí la introduciremos axiomáticamente:

> Dados polinomios $f, g \in K[X]$, diremos que $h \in K[X]$ es un *máximo común divisor* de f y g si y sólo si h verifica las dos siguientes propiedades:
>
> $(MCDP)_1$ $\quad h \mid f$ y $h \mid g$
>
> $(MCDP)_2$ $\quad t \mid h$ para todo $t \in K[X]$ tal que $t \mid f$ y $t \mid g$.

Como se ve, requerimos que h cumpla las mismas condiciones que el máximo común divisor de dos números enteros. Debido a ello, y a pesar de que aún no hemos probado la existencia de un tal h cualesquiera sean f y g, es razonable pensar que valen en $K[X]$ propiedades similares a las del caso entero, como veremos a continuación.

Proposición 10.1.7 Son válidas las siguientes propiedades (las letras designan polinomios):

1) $f \mid g$ si y sólo si f es un mcd de f y g

2) 0 es un mcd de f y g si y sólo si $f = g = 0$

3) Supongamos que h es un mcd de f y g y sea $l \in K[X]$. Entonces l es un mcd de f y g si y sólo si l es un asociado de h

4) Sea $f \neq 0$ y sea $r = r_f(g)$. Entonces h es un mcd de f y g si y sólo si h es un mcd de f y r.

DEMOSTRACION. Todas las demostraciones son consecuencia inmediata de la definición y las propiedades elementales de la divisibilidad, por lo que quedan a cargo del lector. Señalemos eso sí un hecho interesante, que se registra por ejemplo sobre los cuerpos numéricos: si dos polinomios f y g no simultámeante nulos admiten un máximo común divisor h en $K[X]$ entonces admiten infinitos, ya que h es distinto de 0 y tiene por lo tanto infinitos asociados en $K[X]$. Adicionalmente, resulta que f y g admiten un único máximo común divisor mónico, a saber $c^{-1}h$, donde c es el coeficiente principal de h. $\Diamond$

La propiedad 4) de 10.1.7 es clave para probar la existencia de máximo común divisor de dos polinomios cualesquiera. Procederemos por inducción, adelantando que en la demostración subyace un método de cálculo del mcd enteramente análogo al que empleamos en el anillo de números enteros.

Proposición 10.1.8 Dados f y g en $K[X]$, existe en $k[X]$ un máximo común divisor h de f y g. Además, éste puede expresarse como *combinación lineal* de f y g, es decir, existen l y t en $K[X]$ tales que

$$h = lf + tg\,.$$

DEMOSTRACION. Ambas afirmaciones son trivialmente ciertas si alguno de los polinomios divide al otro. Puesto que esta situación incluye los casos $f = 0$ ó $g = 0$, supondremos que f y g son no nulos y haremos inducción en el mínimo m entre sus grados, asumiendo sin pérdida de generalidad que $m = gr(f)$.

Si $m = 0$ estamos nuevamente en el caso anterior, ya que las constantes no nulas dividen a cualquier polinomio. Supongamos entonces que $m > 0$ y que el enunciado es válido para cualquier par de polinomios no nulos u y v tales que $gr(u) < m$ ó $gr(v) < m$. Dividiendo g por f, y pudiendo suponer en virtud de nuestras consideraciones anteriores que g no es múltiplo de f, obtenemos una relación del tipo

$$g = qf + r, \tag{10.2}$$

donde $gr(r) < m$. Resulta luego por hipótesis inductiva que f y r admiten un mcd h en $K[X]$, y puesto que r es el resto de la división de g por f, concluimos por la propiedad 4) de 10.1.7 que h es un mcd de f y g.

Asimismo, el argumento inductivo nos asegura la existencia de polinomios l_1 y t_1 en $K[X]$ tales que

$$h = l_1 r + t_1 f,$$

de donde sigue usando (10.2) que

$$h = l_1\,(g - qf) + t_1 f = (t_1 - l_1 q)\,f + l_1 g,$$

esto es, h es combinación lineal de f y g, siendo $l = t_1 - l_1 q$ y $t = l_1$. $\Diamond$

ALGORITMO DE EUCLIDES. Como dijimos más arriba, la demostración anterior provee un método iterativo de cálculo del máximo común divisor en $K[X]$, llamado como en el caso de números enteros *algoritmo de Euclides*. Así, si f y g son polinomios no nulos se comienza dividiendo g por f, obteniéndose un cociente q_1 y un resto r_1. Si $r_1 = 0$ significa que $f \mid g$, y por lo tanto f es un mcd de f y g. Si $r_1 \neq 0$ debemos continuar, recordando que cualquier mcd de f y r_1 es también un mcd de f y g. En el próximo paso dividimos f por r_1 y así sucesivamente.

Vamos obteniendo de tal forma dos secuencias (q_i) y (r_i) de polinomios, donde q_{i+1} y r_{i+1} son el cociente y el resto de dividir r_{i-1} por r_i, respectivamente, verificándose que cualquier mcd de r_{i-1} y r_i $(i \geq 0)$ es un mcd de f y g (unificamos la notación escribiendo $r_{-1} = g$ y $r_0 = f$).

Puesto que la sucesión de los grados de los restos r_j no puede ser infinitamente decreciente, es claro que existe $n \in \mathbb{N}$ tal que $r_n = 0$. Siendo r_n el resto de dividir r_{n-2} por r_{n-1}, resulta entonces que r_{n-1} es un mcd de r_{n-2} y r_{n-1}, y por lo tanto de f y g. En definitiva, en este proceso de divisiones sucesivas *el último resto no nulo* es un máximo común divisor de f y g.

> Trasladando a $K[X]$ las convenciones empleadas en $\mathbb{Z}$, llamaremos máximo común divisor de dos polinomios no simultáneamente nulos f y g a su único *máximo común divisor mónico*, al que notaremos $(f : g)$.

Para fijar el concepto, supongamos que $h = 4X^3 + 6X - 8$ es un máximo común divisor de dos polinomios u y v. Entonces

$$(u : v) = X^3 + 3/2X - 2,$$

ya que éste es el único polinomio mónico asociado a h. Similarmente, observemos que $(-3X + 2 : 0) = X - 2/3$.

Desarrollemos en detalle un ejemplo de aplicación del algoritmo de Euclides:

Ejemplo 10.1.9 Calculemos $(f : g)$ siendo $f = 4X^6 - 13X^4 + 7X^2 - 3X + 2$ y $g = 4X^7 - 4X^6 - 9X^5 + 5X^4 + 6X^3 - 5X + 2$. Para ello, disponenemos en el siguiente esquema, ya utilizado para números enteros, los cómputos obtenidos en las sucesivas etapas del algoritmo. La segunda columna corresponde a los restos (comenzando por g y f), la tercera a los cocientes, y las rotuladas por l_i y t_i a los coeficientes polinomiales que permiten expresar r_i como combinación lineal de f y g, respectivamente, recordando que satisfacen las reglas de recurrencia $l_{i+1} = l_{i-1} - q_i l_i$ y $t_{i+1} = t_{i-1} - q_i t_i$.

i	r_i	q_{i+1}	l_i	t_i
-1	g	–	0	1
0	f	X–1	1	0
1	$4X^5$–$8X^4$–X^3+$10X^2$–$10X$+4	X+2	$-X$+1	1
2	$4X^4-8X^3-3X^2+13X-6$	X	$-X^3-X^2+1$	X^2+2X+1
3	$2X^3-3X^2-4X+4$	$2X-1$	$-X^3$–X^2+1	X^2+2X+1
4	$2X^2+X-2$	$X-2$	$2X^4+X^3-X$	$-2X^3-3X^2-X-1$
5	0	–	–	–

Resulta entonces que $r_4 = 2X^2 + X - 2$ es un mcd de f y g, por ser el último resto no nulo, verificándose además la relación

$$r_4 = (2X^4+X^3-X)f - (2X^3+3X^2+X+1)g\,. \tag{10.3}$$

Dividiendo r_4 por su coeficiente principal hallamos el máximo común divisor mónico de f y g, a saber:

$$(f:g) = X^2 + \frac{1}{2}\,X - 1\,,$$

mientras que para expresar $(f:g)$ como combinación lineal de f y g dividimos también por 2 los dos miembros de (10.3), en cuyo caso obtenemos

$$(f:g) = \left(X^4 + \frac{1}{2}\,X^3 - \frac{1}{2}\,X\right) f - \left(X^3 + \frac{3}{2}\,X^2 + \frac{1}{2}\,X + \frac{1}{2}\right) g\,. \quad \Diamond$$

Veamos una consecuencia importante de la existencia de máximo común divisor.

Corolario 10.1.10 Si $f, g \in K[X]$, las raíces comunes de f y g en K son las raíces de $(f:g)$ en K.

Demostracion. Es claro que toda raíz de $h = (f:g)$ es raíz de f y de g, pues h es divisor de ambos. Inversamente, sea $\alpha \in K$ tal que $f(\alpha) = g(\alpha) = 0$ y expresemos h como combinación lineal de f y g, digamos $h = pf + qg$. Arribamos entonces al resultado especializando en α, ya que

$$h(\alpha) = p(\alpha)f(\alpha) + q(\alpha)g(\alpha) = 0. \quad \Diamond$$

Ejemplo 10.1.11 Hallemos las raíces $u = X^4 - 4X^3 + 2X^2 + 12X - 15$ sabiendo que tiene una raíz común con $v = X^3 - 2X^2 - 3X + 10$.

De acuerdo con el corolario anterior lo aconsejable es calcular $(u:v)$. Aplicando el algoritmo de Euclides dividimos en primer termino u por v,

obteniendo cociente $q_1 = X - 2$ y resto $r_1 = X^2 - 4X + 5$, mientras que el cociente q_2 y el resto r_2 de dividir v por r_1 son $X + 2$ y 0, respectivamente, como el lector puede verificar. Concluimos luego que $(u : v) = r_1$, resultando además:

$$u = q_1 v + r_1 = q_1 q_2 r_1 + r_1 = (q_1 q_2 + 1)\, r_1 = \left(X^2 - 3\right)\left(X^2 - 4X + 5\right).$$

La descomposición de u obtenida nos permite calcular fácilmente sus raíces, siendo las mismas los ceros de los dos factores cuadráticos de arriba, a saber, $\sqrt{3}$, $-\sqrt{3}$, $2 + \imath$ y $2 - \imath$. $\Diamond$

COPRIMALIDAD. Dos polinomios f y g con coeficientes en K se dirán *coprimos* si y solo si $(f : g) = 1$. Simbolizaremos la situación en la forma $f \perp g$.

Es claro que la definición anterior es equivalente a afirmar que los únicos divisores comunes de f y g son las constantes no nulas, siendo entonces cualquiera de ellas un máximo común divisor de f y g.

Ejemplo 10.1.12 Dos polinomios f y g con coeficientes complejos, de grado a lo sumo 3 y sin raíces comunes en $\mathbb{C}$, son coprimos. En efecto, supongamos por el contrario que $h = (f : g)$ tiene grado positivo. Siendo $gr(h) \leq 3$ podemos asegurar entonces que h tiene una raíz z en $\mathbb{C}$, lo que es contradictorio, ya que en tal caso z sería raíz de f y de g, por corolario 10.1.10. Luego $f \perp g$. $\Diamond$

El concepto que acabamos de definir en $K[X]$ es formalmente idéntico al de coprimalidad en $\mathbb{Z}$, y goza por lo tanto de propiedades muy similares. Enunciaremos a continuación algunas de las más relevantes.

Proposición 10.1.13 Son válidas en $K[X]$ las siguientes propiedades (todas las letras designan polinomios):

1) $f \perp g$ si y solo si existen p y q tales que $1 = pf + qg$

2) Sean f y g polinomios tales que $f = (f : g)f_1$ y $g = (f : g)g_1$. Entonces f_1 y g_1 son coprimos

3) Recíprocamente, supongamos que $f = hu$ y $g = hv$ con $u \perp v$. Entonces h es un máximo común divisor de f y g

4) Sean f, g y t polinomios tales que $f \mid gt$ y $f \perp g$. Entonces $f \mid t$

5) Si $f \perp g$ entonces f y g no tienen raíces comunes en K.

DEMOSTRACION. Encargamos al lector la tarea de probar los ítems 1) a 4), simplemente adaptando las demostraciones que efectuamos en el caso de números enteros. En cuanto a 5), es consecuencia inmediata del corolario 10.1.10. $\Diamond$

Ejemplo 10.1.14 Consideremos en $\mathbb{Q}[X]$ los polinomios

$$f = 2X^4 - 4X^3 + 8X^2 - 12X + 6 = 2(X^2+3)(X-1)^2 \text{ y}$$
$$g = 2X^5 + 2X^4 - 6X^3 + 6X^2 - 36X = 2(X^2+3)X(X-2)(X+3).$$

Resulta por 10.1.12 que los polinomios $(X-1)^2$ y $X(X-2)(X+3)$ son coprimos, pues son de grado menor que 4 y no tienen raíces comunes en $\mathbb{C}$. Luego $(f:g) = X^2+3$, por la propiedad 3) de 10.1.13. $\Diamond$

MINIMO COMUN MULTIPLO. Al igual que en números enteros, el concepto de máximo común divisor en $K[X]$ admite la siguiente noción dual: dados f, g y h en $K[X]$, diremos que h es un *mínimo común múltiplo* de f y g si y solo si se satisfacen las dos siguientes condiciones:

$(MCMP)_1$ $\quad f \mid h$ y $g \mid h$

$(MCMP)_2$ $\quad h \mid t$ para todo $t \in K[X]$ tal que $f \mid t$ y $g \mid t$.

Observemos que de acuerdo con la definición anterior, un mínimo común múltiplo de f y g tendrá grado mínimo entre los múltiplos no nulos de ambos. Veamos cómo asegurar la existencia de un tal polinomio y cómo se caracteriza el conjunto de todos los polinomios que verifican las condiciones de la definición.

Proposición 10.1.15 Todo par de polinomios f y g con coeficientes en K admite un mínimo común múltiplo en $K[X]$. Designando por h cualquiera de ellos, resulta además que un elemento l de $K[X]$ también lo es si y sólo si es un asociado de h.

DEMOSTRACION. La última afirmación sigue fácilmente de las propiedades generales de la divisibilidad, por lo que solo demostraremos en detalle la primera.

Si $f = 0$ ó $g = 0$ sigue inmediatamente que $h = 0$ es el único mínimo común múltiplo de f y g, por lo que supondremos que ambos polinomios son no nulos. Si $d = (f:g)$, probaremos que $h = dpq$ es un mínimo común múltiplo de f y g, donde p y q son polinomios tales que $f = dp$ y $g = dq$.

Es claro que h es múltiplo de f y g, ya que valen las relaciones

$$h = fq = gp\,.$$

Respecto de $(MCMP)_2$, supongamos que t es múltiplo de f y g, digamos

$$t = ff_1 = gg_1.$$

Reemplazando f por dp y g por dq en la última igualdad, y cancelando d, obtenemos

$$pf_1 = qg_1,$$

y por lo tanto $q \mid pf_1$. Siendo p y q coprimos sigue que $q \mid f_1$ (propiedades 3) y 4) de 10.1.13). Luego, escribiendo $f_1 = qk$ ($k \in K[X]$) tenemos:

$$t = ff_1 = dpf_1 = dpqk = hk,$$

y en consecuencia $h \mid t$, como queríamos demostrar. $\Diamond$

Como en el caso del máximo común divisor, dados polinomios no nulos f y g designaremos por $[\, f : g \,]$ el único mínimo común múltiplo mónico de f y g. Observemos que si estos son mónicos (siempre podemos situarnos en ese caso) vale la fórmula

$$fg = (f : g)\,[\, f : g \,]\,. \qquad \Diamond$$

10.1.4. Ejercicios

1. En cada uno de los siguientes casos determinar todas las descomposiciones de f como producto de dos polinomios mónicos con coeficientes complejos:
   a) $f = X^4 - 2$
   b) $f = X^3 - (4+\sqrt{2})X^2 + (5+4\sqrt{2})X - 5\sqrt{2}$
   c) $f = X^4 + 1$
   d) $f = X^4 - X^3 - 5X^2 - X - 6$
   e) $f = X^6 + X^3$
   f) $f = X^4 + 2X^2 - 8\,.$

2. Demostrar que dos polinomios mónicos son asociados si y solo si son iguales.

3. *a)* Mostrar que $X^2 - 2X + 2$ admite en $\mathbb{Q}[X]$ un múltiplo de la forma $X^m + c$
   b) Analizar si es posible afirmar lo mismo con respecto al polinomio $X^3 - 3X^2 + 4X - 2$.

4. Hallar $n > 4$ tal que $X^3 - X - 1 \mid X^n - X^4 - 1$ en $\mathbb{Q}[X]\,.$

5. Si p es un número primo, demostrar que $X^p - X$ es divisible en $\mathbb{Z}_p[X]$ por todos los polinomios de grado 1.

6. En cada uno de los siguientes casos dividir:
   a) $2X^5 - (2+4\imath)X^4 + 6\imath X^3 + 10X^2 - 5X - 5 + 6\imath$ por $2X^3 - 4\imath X^2 + 6$
   b) $X^8 - 2X^6 + 2X^5 + 3X^4 - 2X^3 - X^2 + 2X - 8$ por $X^4 - X^2 + X + 1$

c) $-8X^3 + 6X^2 - 4X + 1$ por $X^5 - X^3 + 5X^2 + 6X$

d) $1 + X + \ldots + X^{12}$ por $X^3 + X^2 + 1$

e) $3X^{20} + 2X^{15} - 2X^{13} - 6X^9 + 6X^8 + 9X^3 - 8X - 8$ por X^5 .

7. Sea $g \in \mathbb{R}[X]$ y sea $f = 2X^3 - 5X + 4$. Si h y r son el cociente y el resto de dividir g por f, determinar el cociente y el resto de dividir:

a) g por $f/2$

b) $3g + X^2 - 1$ por f

c) $g - 3r$ por $-3f$

d) $g + Xh$ por $f + X$.

8. Sean $m, n \in \mathbb{N}$ y sea $r = r_m(n)$.

a) Probar que $X^r - 1$ es el resto de dividir $X^n - 1$ por $X^m - 1$

b) Probar que $X^m - 1 \mid X^n - 1 \Leftrightarrow m \mid n$.

9. Sea $f = X^3 - aX^2 + 2X - 1$ ($a \in \mathbb{C}$). Hallar los valores de a tales que:

a) $f \mid X^6 + X^5 + X^4 - 2X^3 - 2X^2 + X$

b) f es múltiplo de $X^2 + X - 1$

c) $f \equiv X - 6\ (X^2 + 1)$

d) $X^4 - 1 \equiv -X^2 - X\ (f)$.

10. *a*) Sea K un cuerpo numérico y sea $f \in K[X]$ un polinomio de grado $n > 0$. Probar que para cada $g \in K[X]$ existe una única secuencia $a_0, a_1, \ldots, a_{n-1}$ de elementos de K tales que

$$g \equiv \sum_{i=0}^{n-1} a_i X^i \ (f)$$

b) En relación con la parte *a*), sean $g, h \in \mathbb{R}[X]$ y sean a, b, c y d en $\mathbb{R}$ tales que $g \equiv a + bX\ (X^2 + 1)$ y $h \equiv c + dX\ (X^2 + 1)$. Probar las relaciones

(i) $g + h \equiv (a + c) + (b + d)X\ (X^2 + 1)$

(ii) $gh \equiv (ac - bd) + (ad + bc)X\ (X^2 + 1)$.

11. En cada uno de los siguientes casos calcular $(f : g)$ y expresarlo como combinación lineal de f y g:

a) $f = X^4 - 5X^3 + X - 5$, $g = 2X^6 - X^5 + X^4 + X^2 - X$

b) $f = X^5 + X^4 + 3X^3 - X^2 - 4$, $g = X^3 + X - 2$

c) $f = 3(X+3)^7(X^2 - X - 6)$, $g = (X+3)^6(2X^2 - 5X - 3)$

d) $f = X^5 - X^4 - X^2 + X - 2$, $g = 4 - (X+5)^{20} f$.

12. Si $m, n \in \mathbb{N}$, probar que $(X^m - 1 : X^n - 1) = X^{(m:n)} - 1$.

13. En cada uno de los siguientes casos hallar las raíces complejas comunes de f y g:

a) $f = X^4 + \imath X^3 - X^2 + (2 - \imath)X + 2\imath$; $g = X^4 + X^3 + 2X^2 + X + 1$

b) $2X^3 - X^2 + 4$; $g = -2X^3 - X^2 + X - 1$

c) $f = \sum_{k=0}^{20} X^k$; $g = \sum_{k=0}^{10} X^k$.

14. En cada uno de los ítems del ejercicio 11 calcular el mínimo común múltiplo de f y g .

15. Sea K un cuerpo numérico y sean f, f_1, g y g_1 en $K[X]$ tales que f_1 es un asociado de f y g_1 es un asociado de g. Probar que $(f_1 : g_1) = (f : g)$ y $[\, f_1 : g_1 \,] = [\, f : g \,]$.

16. Sea $g = X^5 + 3X^4 + 3X^3 + X^2$. Determinar $f \in \mathbb{Q}[X]$ de grado 4 tal que $(f : g) = X^3 + 2X^2 + X$ y $f(1) = -20$.

17. Hallar las raíces complejas de $X^4 - X^3 - X - 1$ sabiendo que tiene una raíz común con $X^4 + X^3 + 4X^2 + 3$.

10.2. Divisibilidad y raíces

10.2.1. Factores de grado uno

De acuerdo con la teoría, el resto de la división en $K[X]$ de cualquier polinomio por otro de grado uno, que sin pérdida de generalidad podemos suponer mónico, es un polinomio de grado cero o el polinomio nulo. En cualquier caso será un elemento de K, que como ahora veremos puede calcularse por especialización:

Proposición 10.2.1 (Teorema del resto) Si $g \in K[X]$ y $c \in K$, el resto de la división de g por $X - c$ es $g(c)$.

Demostracion. Sea h el cociente y sea $r \in K$ el resto de dividir g por $X - c$. Evaluando en c cada miembro de la igualdad $g = (X - c)h + r$, y aplicando las propiedades de la especialización obtenemos:

$$g(c) = (c - c)h(c) + r(c) = (c - c)h(c) + r = r\,,$$

como queríamos demostrar.

Por ejemplo, si $g = \sum_i a_i X^i$ resulta que el resto de dividir g por X es a_0, mientras que los restos de dividirlo por $X-1$ y $X+1$ son $\sum_i a_i$ y $\sum_i (-1)^i a_i$, respectivamente.

Teniendo en cuenta un apunte que hicimos en la sección anterior, acerca de que cualquier división puede reducirse a la división por un polinomio mónico, el lector probará sin inconvenientes la siguiente generalización del enunciado anterior: el resto de la división de g por $aX + b$ es $g(-b/a)$. $\Diamond$

Como corolario del teorema del resto reaparece un hecho ya demostrado (algo trabajosamente) en el capítulo 9, que muestra que las raíces de un polinomio se corresponden con sus divisores de grado 1:

Corolario 10.2.2 Sea $g \in K[X]$ y sea $c \in K$. Entonces c es raíz de g si y sólo si $X - c$ divide a g. Equivalentemente, c es raíz de g si y sólo si g se factoriza en la forma

$$g = (X - c)h,$$

con $h \in K[X]$.

Demostracion. Queda a cargo del lector. $\Diamond$

Regla de Ruffini.

El proceso de hallar el cociente h y el resto r de la división de un polinomio $g \in K[X]$ por otro de la forma $X - c$ es particularmente simple. En efecto, supongamos que g es de grado n ($n > 0$), y designemos por a_i y b_i los

coeficientes de g y h, respectivamente. Puesto que claramente $gr(h) = n-1$, deducimos de la igualdad

$$\sum_{i=0}^{n} a_i X^i = (X - c) \sum_{i=0}^{n-1} b_i X^i + r$$

la validez de las relaciones $a_n = b_{n-1}$ y $a_i = b_{i-1} - cb_i$ para todo $0 \leq i < n$, definiendo $b_{-1} = r$.

Despejando, obtenemos entonces las fórmulas

$$b_{i-1} = a_i + cb_i\,, \qquad (0 \leq i \leq n-1) \tag{10.4}$$

que constituyen una regla práctica para determinar iterativamente (en orden decreciente y comenzando por $b_{n-1} = a_n$) los coeficientes b_j de h, recordando que el último es el resto de la división. Naturalmente, las fórmulas descriptas se aplican también al caso de la división por un polinomio de la forma $X+c$, simplemente reemplazando c por $-c$ en (10.4).

La fórmula de cálculo dada por las relaciones (10.4) es la llamada *regla de Ruffini*. Si bien es válida cualquiera sea c, resulta particularmente útil en el caso de que c sea una raíz de g, ya que nos brinda una manera ágil de hallar el cociente h y encarar entonces el cálculo de otras raíces de g.

Mostremos a través de un ejemplo cómo funciona la regla.

Ejemplo 10.2.3 Efectuemos la división de $g = 3X^6 - 15X^4 + 6X^3 + X^2 - 5$ por $X - 2$ usando la regla de Ruffini. La siguiente tabla, de uso corriente, ilustra los pasos del algoritmo:

	3	0	-15	6	1	0	-5
2		6	12	-6	0	2	4
	3	6	-3	0	1	2	-1

Se comienza escribiendo en la fila superior los coeficientes de g, en orden decreciente de grado, mientras que en la inferior se ubica los sucesivos b_i obtenidos. Las celdas de la fila central se llenan con los resultados de multiplicar c (situado en el extremo izquierdo de la misma) por el anterior b_j hallado. El lector podrá observar que la tercera fila comienza con el coeficiente principal de g, y que cada uno de los restantes elementos (el último de los cuales es el resto de la división) se obtienen sumando los otros dos elementos de la columna, en acuerdo con las fórmulas (10.4). En este caso, resulta entonces $h = 3X^5 + 6X^4 - 3X^3 + X + 2$ y $r = -1$. $\Diamond$

10.2.2. Multiplicidad

Sabemos que un polinomio $f = aX^2 + bX + c$ de grado 2 cuyo discriminante Δ es nulo admite una única raíz, a saber, $z = -b/2a$, que se dice entonces una raíz doble de f. Examinaremos un poco más esta situación, como paso previo a definir en general el concepto de multiplicidad de una raíz. Informalmente hablando, dicha raíz será "contada" tantas veces como su multiplicidad en la lista de raíces del polinomio, aunque obviamente debemos precisar el significado de esta frase.

En el caso cuadrático mencionado, completando cuadrados obtenemos

$$f = a\left(X + \frac{b}{2a}\right)^2 + c - \frac{b^2}{4a} = a\left(X + \frac{b}{2a}\right)^2 - \frac{\Delta}{4a} = a\,(X - z)^2\,, \quad (10.5)$$

observándose entonces que f no solamente es divisible por $X - z$ sino también por $(X - z)^2$, lo que motiva el nombre de raíz doble asignado a z. Inspirados por esta situación, definimos el concepto de multiplicidad de una raíz en la siguiente forma:

> Sea f un polinomio no nulo con coeficientes en un cuerpo K y sea z una raíz de f en K. Diremos que z es una raíz de *multiplicidad* m de f ($m \in \mathbb{N}$) si y sólo si $(X - z)^m \mid f$ y $(X - z)^{m+1} \nmid f$.

Diremos también que z es una raíz de *orden* m de f, y emplearemos la notación $\mathrm{mult}(z, f) = m$. Para valores bajos de m se emplean algunos nombres especiales, como los de raíces *dobles* y *triples* para designar los casos $m = 2$ y $m = 3$, mientras que una raíz de multiplicidad 1 se dice *simple*. Además, es costumbre llamar raíz *múltiple* a toda raíz de multiplicidad mayor que 1.

NOTA. Respecto a la buena definición de la multiplicidad, notemos que el conjunto de exponentes j tales que $(X - z)^j \mid f$ es no vacío (por ser z una raíz de f) y está acotado por el grado de f, hechos que garantizan la existencia de un tal m, resultando en particular que $\mathrm{mult}(z, f) \leq gr(f)$. Concomitantemente, es claro que no tiene sentido referirse a la multiplicidad de z respecto del polinomio nulo, pues éste es divisible por cualquier potencia de $X - z$.

Consignemos también que la noción de multiplicidad respecto de un polinomio f puede extenderse a todo elemento z de K, definiendo $\mathrm{mult}(z, f) = 0$ si z no es raíz de f. Esto permite brindar entonces la siguiente definición (más general) de multiplicidad:

$$\mathrm{mult}(z, f) = \max\left\{j \in \mathbb{N}_0 \,:\, (X - z)^j \mid f\right\},$$

para todo polinomio no nulo f y para todo $z \in K$, resultando que z es raíz de f si y sólo si $\mathrm{mult}(z, f) \geq 1$. $\Diamond$

Ejemplos 10.2.4 Consideremos el polinomio

$$f = X^7 - 6X^6 + 13X^5 - 14X^4 + 12X^3 - 8X^2.$$

Es claro que 0 es raíz de f y que f es divisible por X^2, a través de la factorización $f = X^2(X^5 - 6X^4 + 13X^3 - 14X^2 + 12X - 8)$. Designando por g este último factor resulta que f no es divisible por X^3, ya que si fuera $f = X^3 g_1$ tendríamos $X^2 g = X^3 g_1$, de donde $g = X g_1$ por ser $\mathbb{Q}[X]$ un dominio de integridad. Puesto que esto no es posible (0 no es raíz de g), sigue que 0 es raíz doble de f.

Para hallar otras raíces de f bastará buscar las de g, resultando por el criterio de Gauss que 2 es raíz del mismo. Para hallar la máxima potencia de $X - 2$ que divide a g, y por lo tanto a f, comenzamos dividiendo g por $X - 2$, luego dividimos el cociente por $X - 2$, y así sucesivamente, hasta encontrarnos con un cociente que no sea divisible por $X - 2$. Obtenemos de tal manera las factorizaciones:

$$\begin{aligned} g &= (X-2)(X^4 - 4X^3 + 5X^2 - 4X + 4) = \\ &= (X-2)^2(X^3 - 2X^2 + X - 2) = (X-2)^3(X^2 + 1) \end{aligned}$$

Puesto que $X - 2 \nmid X^2 + 1$, pues las raíces de este son $\imath$ y $-\imath$, razonando como antes resulta que 2 es una raíz de multiplicidad 3 de f.

Finalmente, f se descompone en la forma:

$$f = X^2(X-2)^3(X-\imath)(X+\imath),$$

lo que nos permite concluir que f admite 7 raíces en $\mathbb{C}$ (cada una contada tantas veces como su multiplicidad): 0 (doble), 2 (triple) y las raíces simples $\imath$ y $-\imath$.

Como ejemplo más general señalemos que z es raíz de multiplicidad r del polinomio $h = a(X - z)^r$, donde $z \in K$, $a \in K^*$ y r un entero no negativo, ya que por razones de grado es obvio que $(X - z)^{r+1} \nmid h$. $\quad \Diamond$

Estudiaremos ahora algunas propiedades de la multiplicidad, la primera de las cuales brinda una caracterización muy útil para su cálculo. Asimismo, la propiedad 5) afina un resultado que probamos en el capítulo 9 acerca del número de raíces de un polinomio.

Proposición 10.2.5 Son válidas las siguientes propiedades (los símbolos z y z_i designan elementos de K y las otras letras polinomios con coeficientes en K):

1) $\text{mult}(z, f) = m$ si y sólo si f se factoriza en la forma

$$f = (X - z)^m f_1$$

con $f_1(z) \neq 0$

2) Si $f = gh$ entonces $\text{mult}(z, f) = \text{mult}(z, g) + \text{mult}(z, h)$

3) Si $g \mid f$ entonces $\text{mult}(z, g) \leq \text{mult}(z, f)$

4) Sea $f = (X - z)^r g$. Entonces $\text{mult}(z, f) \geq r$

5) Sean $z_1, \ldots, z_s$ raíces distintas de f en K, con $\text{mult}(z_i, f) = m_i$. Entonces
$$\prod_{i=1}^{s} (X - z_i)^{m_i} \mid f\,.$$

Resulta en particular que un polinomio de grado n con coeficientes en K tiene a lo sumo n raíces en K, *aún contando cada una de ellas tantas veces como su multiplicidad*.

DEMOSTRACION. Para probar 1), supongamos primero que $\text{mult}(z, f) = m$. Sigue entonces por definición que $f = (X - z)^m f_1$. Si $f_1(z) = 0$ resulta por 10.2.2 que f_1 es divisible por $X - z$, digamos $f_1 = (X - z)f_2$, y por lo tanto $f = (X - z)^m (X - z) f_2 = (X - z)^{m+1} f_2$, lo que absurdo. Luego $f_1(z) \neq 0$.

Recíprocamente, supongamos que f admite una descomposición como la del enunciado. Entonces $(X-z)^{m+1} \nmid f$, pues en caso contrario obtendríamos una igualdad del tipo $(X - z)^m f_1 = f = (X - z)^{m+1} t$, resultando por cancelación que $f_1 = (X - z)t$, lo que implicaría que z es raíz de f_1. Puesto que obviamente $(X - z)^m \mid f$, concluimos que $\text{mult}(z, f) = m$.

Respecto a 2), supongamos que las multiplicidades de z con respecto a g y h son i y j, respectivamente. De acuerdo con 1) podemos entonces escribir $g = (X-z)^i g_1$ y $h = (X-z)^j h_1$, donde z no es raíz ni de g_1 ni de h_1. Luego
$$f = gh = (X - z)^i (X - z)^j g_1 h_1 = (X - z)^{i+j} g_1 h_1,$$
y sigue de 1) que $\text{mult}(z, f) = i + j$, ya que $g_1 h_1(z) = g_1(z) h_1(z) \neq 0$.

La propiedad 3) es inmediata a partir de 2), pues la multiplicidad es un entero no negativo, mientras que 4) es un caso particular de 3), teniendo en cuenta que $\text{mult}\,(z, (X - z)^r) = r$.

Vayamos finalmente a 5), observando que el resultado sigue directamente de la definición si $s = 1$. Si $s > 1$, escribamos $f = (X - z_1)^{m_1} p$, y consideremos cualquier índice $j > 1$. Tenemos entonces aplicando 2) que
$$\text{mult}(z_j, p) = \text{mult}(z_j, f) - \text{mult}(z_j, (X - z_1)^{m_1}) = m_j,$$
pues z_j no es raíz de $(X - z_1)^{m_1}$. Sigue luego por un argumento inductivo que existe q tal que
$$p = \left(\prod_{i=2}^{s} (X - z_i)^{m_i} \right) q\,,$$
de donde reemplazando obtenemos:
$$f = (X - z_1)^{m_1} \left(\prod_{i=2}^{s} (X - z_i)^{m_i} \right) q = \left(\prod_{i=1}^{s} (X - z_i)^{m_i} \right) q\,,$$
como queríamos.

Una simple consideración de grados nos permite probar la última afirmación del enunciado, ya que

$$n = gr(f) \geq gr\left(\prod_{i=1}^{s}(X - z_i)^{m_i}\right) = \sum_{i=1}^{s} m_i\,. \quad \Diamond$$

Corolario 10.2.6 Sea $f \in K[X]$ un polinomio de grado n admitiendo n raíces $z_1, z_2, \ldots, z_n$ (no necesariamente distintas) en K. Entonces

$$f = c\prod_{i=1}^{n}(X - z_i)$$

para algún $c \in K^*$.

DEMOSTRACION. Por la propiedad 5) de 10.2.5 el polinomio $g = \prod_{i=1}^{n}(X - z_i)$ divide a f. Siendo ambos de grado n el factor de divisibilidad debe ser de grado 0, lo que prueba nuestra afirmación. Notemos de paso que c es el coeficiente principal de f, ya que g es mónico. $\Diamond$

Multiplicidad y derivación.

Hasta aquí, para calcular la multiplicidad de un elemento z respecto a un polinomio hemos procedido dividiendo reiteradamente por $X - z$. Existe sin embargo otra forma de hacerlo, a través de los polinomios derivados. Para tantear la situación, consideremos nuevamente el caso de un polinomio $f = aX^2 + bX + c$ de grado 2. Puesto que $-b/2a$ es la única raíz de su derivado $f' = 2aX + b$, deducimos de (10.5) que f admite una raíz doble si y sólo si tiene una raíz en común con f'. En la siguiente proposición demostraremos la validez de este hecho en general, y mostraremos cómo usar los derivados de un polinomio para el cálculo de multiplicidades.

Proposición 10.2.7 Sea $f \in K[X]$ y sea z una raíz de f en K. Entonces:

a) z es raíz múltiple de f si y sólo si z es raíz de f'. Precisamente, vale la fórmula

$$\text{mult}(z, f') = \text{mult}(z, f) - 1\,. \tag{10.6}$$

b) $\text{mult}(z, f) = m$ $(m \in \mathbb{N})$ si y sólo si

$$f(z) = f^{(\prime)}(z) = f^{(2)}(z) = \cdots = f^{(m-1)}(z) = 0$$

y $f^{(m)}(z) \neq 0\,.$

DEMOSTRACION. Probaremos directamente la fórmula (10.6), ya que la primera afirmación de a) se deduce de ella. Sea entonces $\text{mult}(z, f) = r$ y escribamos $f = (X - z)^r f_1$, con $f_1(z) \neq 0$. Usando las reglas de derivación tenemos:

$$f' = r(X-z)^{r-1} f_1 + (X-z)^r f_1' = (X-z)^{r-1}(r f_1 + (X-z) f_1') = (X-z)^{r-1} p\,,$$

y puesto que $p(z) = r f_1(z) \neq 0$ (notemos que $r \geq 1$), resulta por la propiedad 1) de 10.2.5 que $\text{mult}(z, f') = r - 1$, como queríamos probar.

Para probar la parte b) procederemos por inducción en m, empleando la notación $f' = g$.

Aplicando a), sigue que $\text{mult}(z, f) = 1$ si y solo si $\text{mult}(z, f) = 0$, lo que prueba el caso $m = 1$. Supongamos ahora que $m > 1$ y que el resultado es válido para raíces de multiplicidad menor que m.

Si $\text{mult}(z, f) = m$, resulta que $\text{mult}(z, g) = m - 1$, de donde sigue por hipótesis inductiva que $g^{(i)}(z) = 0$ para $i = 0, 1, \ldots, m-2$, mientras que $g^{(m-1)}(z) \neq 0$. Puesto que $g^{(i)} = f^{(i+1)}$ y $f(z) = 0$, las igualdades anteriores indican que $f^{(i)}(z) = 0$ para $0 \leq i \leq m-1$ y $f^{(m)}(z) \neq 0$.

Recíprocamente, las condiciones del enunciado aseguran por hipótesis inductiva que $\text{mult}(z, f') = m - 1$, por lo que $\text{mult}(z, f) = m$. $\quad \Diamond$

Como ilustración del criterio que acabamos de demostrar consideremos nuevamente el polinomio $f = X^7 - 6X^6 + 13X^5 - 14X^4 + 12X^3 - 8X^2$, cuyos primeros derivados son

$$\begin{aligned} f' &= 7X^6 - 36X^5 + 65X^4 - 56X^3 + 36X^2 - 16X \\ f'' &= 42X^5 - 180X^4 + 260X^3 - 168X^2 + 72X - 16 \\ f''' &= 210X^4 - 720X^3 + 780X^2 - 336X + 72\,. \end{aligned}$$

Es obvio que 0 es raíz de f y f' pero no de f'', por lo que es una raíz doble de f. Por otra parte 2 es raíz triple de f, ya que $f(2) = f'(2) = f''(2) = 0$ y $f'''(2) = 120$. Finalmente, $\imath$ y $-\imath$ son raíces simples, ya que ambas son raíces de f y no de f'.

Resaltemos que para aplicar la fórmula (10.6) y el criterio de los derivados es esencial que z *sea raiz* de f. Por ejemplo, es inmediato verificar que 1 anula al primero y al segundo derivado del polinomio $f = 2X^3 - 6X^2 + 6X - 1$, no obstante lo cual $\text{mult}(1, f) = 0$, por ser $f(1) \neq 0$.

10.2.3. Relaciones entre coeficientes y raices

Podemos determinar completamente los valores de dos números reales u y v (digamos $u \leq v$) si conocemos su suma y su producto. Por ejemplo, sabiendo que $u + v = 7$ y $uv = 12$ resulta $u = 3$ y $v = 4$, mientras que $u = -1/2$ y $v = 2$ si $u + v = 3/2$ y $uv = -1$. Para justificar nuestra afirmación, observemos que existe un polinomio mónico de grado 2 cuyos

coeficientes son los datos $u+v$ y uv del problema y cuyas raíces (que sabemos entonces cómo hallar) son nuestras incógnitas u y v, a saber:

$$p = (X-u)(X-v) = X^2 - (u+v)X + uv\,.$$

La igualdad polinomial de arriba muestra que los coeficientes a_i de un polinomio mónico de grado 2 quedan unívocamente determinados por sus raíces, a través de las relaciones $a_1 = -(u+v)$ y $a_0 = uv$. Como pronto veremos, es posible establecer relaciones similares para polinomios de grado arbitario (no necesariamente mónicos) con coeficientes en cualquier cuerpo K.

Por ejemplo, supongamos que queremos determinar $g \in K[X]$, de grado 3 y con coeficiente principal c, cuyas raíces sean tres elementos dados u, v y w de K. Por un lado, g será de la forma $g = cX^3 + a_2X^2 + a_1X + a_0$, mientras que por otra parte (corolario 10.2.6) deberá ser

$$\begin{aligned} g &= c(X-u)(X-v)(X-w) = \\ &= cX^3 - c(u+v+w)X^2 + c(uv+uw+vw)X - cuvw\,. \end{aligned}$$

Igualando coeficientes, obtenemos entonces las igualdades

$$\begin{aligned} a_2 &= -c\,(u+v+w) \\ a_1 &= c\,(uv+uw+vw) \\ a_0 &= -c\,(uvw)\,, \end{aligned}$$

que permiten expresar los coeficientes de un polinomio de grado 3 en términos de sus raíces.

Salvo signo y el factor constante c (que podemos elegir arbitrariamente), observemos que a_2 es la suma de las raíces, a_1 es la suma de todos los posibles productos de 2 raíces y a_0 es el producto de las 3 raíces. Para apreciar mejor la existencia de un patrón similar al del caso cuadrático, digamos que los coeficientes se obtienen, en orden decreciente de grado, como sumas de productos de raíces, tomadas éstas de a una, de a dos ó de a tres, respectivamente.

Habiendo estudiado estos casos particulares demostraremos ahora en general las fórmulas de *Vieta*, que establecen las relaciones existentes entre coeficientes y raíces de un polinomio arbitrario de grado n con coeficientes en un cuerpo K.

Proposición 10.2.8 (Fórmulas de Vieta) Sea

$$f = \sum_{i=0}^{n} a_i X^i$$

un polinomio de grado n con coeficientes en un cuerpo K admitiendo n raíces $x_1, x_2, \ldots, x_n$ (no necesariamente distintas) en K. Entonces

$$a_{n-k} = (-1)^k a_n \left(\sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} x_{i_1} x_{i_2} \cdots x_{i_k} \right) \tag{10.7}$$

para $1 \leq k \leq n$.

DEMOSTRACION. Observemos que las fórmulas (10.7) generalizan las obtenidas en los casos $n = 2$ y $n = 3$, ya que la expresión entre paréntesis indica que la suma se efectúa sobre *todos los posibles productos de k raíces* de f, escribiéndose los subíndices en forma creciente para destacar el hecho de que cada sumando corresponde a una elección de k índices distintos entre n, sin que importe el orden de elección. Resulta en particular que dicha suma, que notaremos $s_k(x_1, x_2, \ldots, x_n)$, tiene $\binom{n}{k}$ términos.

Podemos demostrarlas en forma combinatoria, usando la descomposición

$$f = a_n\,(X - x_1)(X - x_2)\cdots(X - x_{n-1})(X - x_n) \tag{10.8}$$

dada por el corolario 10.2.6. En efecto, siendo el miembro de la derecha de (10.8) un producto de n binomios de primer grado (aparte del factor a_n), es claro que un término genérico de grado $n - k$ de su expansión se obtiene como producto de k monomios de grado 0 y $n - k$ monomios de grado 1, resultando entonces de la forma

$$(-1)^k a_n x_{i_1} \cdots x_{i_k} X^{n-k},$$

para una cierta elección de índices $i_1 < i_2 < \cdots < i_k$. Por lo tanto, el coeficiente a_{n-k} de dicho producto se obtiene sumando las expresiones $(-1)^k a_n x_{i_1} \cdots x_{i_k}$ sobre todas las formas posibles de elegir dichos índices. Vale decir, sobre todas las formas posibles de elegir k raíces de f, como queríamos demostrar. $\Diamond$

NOTA Las sumas $s_i(x_1, x_2, \ldots, x_n)$, que de no existir riesgo de confusión notaremos más concisamente s_i, se llaman *funciones simétricas elementales* de los elementos $x_1, x_2, \ldots, x_n$. Por ejemplo, para $n = 4$ tenemos

$$\begin{aligned}
s_1 &= x_1 + x_2 + x_3 + x_4 \\
s_2 &= x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4 \\
s_3 &= x_1x_2x_3 + x_1x_2x_4 + x_1x_3x_4 + x_2x_3x_4 \\
s_4 &= x_1x_2x_3x_4\,.
\end{aligned}$$

Puesto que en un cuerpo la suma y el producto son operaciones conmutativas, es inmediato probar que $s_k(x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(n)}) = s_k(x_1, x_2, \ldots, x_n)$ para toda permutación π de los índices, lo cual explica su nombre de simétricas. En cuanto a su designación de elementales, se debe al hecho de que toda función polinómica simétrica en las variables $x_1, \ldots, x_n$ puede expresarse como función polinómica en las variables s_i. No ofreceremos aquí una demostración de este teorema, dadas las dificultades que la misma presenta, pero ilustraremos la situación para $n = 3$ considerando la función

$g(x_1, x_2, x_3) = x_1^3 + x_2^3 + x_3^3$, que es obviamente simétrica. En este caso, aplicando las definiciones y la fórmula de Leibniz, obtenemos:

$$\begin{aligned} g(x_1, x_2, x_3) &= s_1^3 - 3\left(x_1^2x_2 + x_1^2x_3 + x_1x_2^2 + x_2^2x_3 + x_1x_3^2 + x_2x_3^2\right) - 6s_3 = \\ &= s_1^3 - 3s_1(x_1x_2 + x_1x_3 + x_2x_3) + 9x_1x_2x_3 - 6s_3 = \\ &= s_1^3 - 3s_1s_2 + 3s_3, \end{aligned}$$

como el lector puede constatar. ◊

Ejemplos 10.2.9 Construyamos en $\mathbb{R}[X]$ un polinomio mónico f de grado 3 cuyas raíces sean -1, 2 y 4. De acuerdo con las definiciones, tenemos en este caso que $s_1 = 5$, $s_2 = 2$ y $s_3 = -8$, por lo que aplicando (10.7) obtenemos

$$f = X^3 - 5X^2 + 2X + 8,$$

siendo inmediato verificar que se satisfacen las condiciones requeridas.

Hallemos ahora un polinomio g de grado 4 con coeficiente principal 6 y cuyas raíces sean $\sqrt{3}$, $-\sqrt{3}$ y $1/2$, esta última con multiplicidad 2 (lo que significa que debe ser contada dos veces en la lista de raíces de g). Efectuando las correspondientes operaciones resulta entonces que $s_1 = 1$, $s_2 = -11/4$, $s_3 = -3$ y $s_4 = -3/4$, y por lo tanto

$$g = 6(X^4 - X^3 - 11/4X^2 + 3X - 3/4) = 6X^4 - 6X^3 - 33/2X^2 + 18X - 9/2\,.$$

Dejamos las verificaciones a cargo del lector. ◊

Ejemplo 10.2.10 En la búsqueda de un caso más interesante consideremos el polinomio $h = X^3 - 6X + 3$. Empleando el teorema de Bolzano es fácil ver que h tiene 3 raíces reales, situadas en los intervalos $(-3, -2)$, $(0, 1)$ y $(2, 3)$, respectivamente. Deducimos en particular que las mismas son números irracionales, ya que las posibles raíces racionales de un polinomio mónico con coeficientes enteros deben ser enteras, por el criterio de Gauss.

Nos proponemos caracterizar los polinomios de grado 3 cuyas raíces sean los cuadrados de las raíces u, v y w de h, aún sin conocer éstas explícitamente. Lo haremos calculando formalmente las funciones simétricas elementales s_i de u^2, v^2 y w^2, usando las relaciones $u + v + w = 0$, $uv + uw + vw = -6$ y $uvw = -3$. Operando, obtenemos:

$$\begin{aligned} s_1 &= u^2 + v^2 + w^2 = (u + v + w)^2 - 2(uv + uw + vw) = -2(-6) = 12 \\ s_2 &= u^2v^2 + u^2w^2 + v^2w^2 = (uv + uw + vw)^2 - 2uvw(u + v + w) = 36 \\ s_3 &= u^2v^2w^2 = (uvw)^2 = 9\,. \end{aligned}$$

Por lo tanto,

$$h_1 = X^3 - 12X^2 + 36X - 9$$

es un polinomio como el que buscamos. Más generalmente, los polinomios de grado 3 con coeficientes complejos que admiten a u^2, v^2 y w^2 por raíces son los de la forma αh_1, donde α es cualquier número complejo no nulo. ◊

10.2.4. Ejercicios

1. Calcular en $\mathbb{Q}[X]$:

   a) El resto de dividir X^n por X^2-2 (analizar los casos n par y n impar)

   b) El resto de dividir $X^{50}-2X^{49}+X^6-1$ por X^3-2X^2-X+2

   c) El resto de dividir $X^{160}+4X^{81}+1$ por $X^{80}-X^2-3$

   d) El resto de dividir $\sum_{k=0}^{n} X^k$ por X^3-X $(n\in\mathbb{N})$.

2. Sea $f\in\mathbb{Q}[X]$ tal que $f(1+\sqrt{3})=0$. Probar que f es divisible por X^2-2X-2.

3. Sea $f\in\mathbb{C}[X]$ de grado $n\leq 3$. Probar que f admite n raíces (no necesariamente distintas) en $\mathbb{C}$.

4. Calcular $\text{mult}(z,f)$ en cada uno de los siguientes casos:

   a) $f=X^5-5X^4+9X^3-9X^2+8X-4$; $z=-1$, $z=2$, $z=\imath$

   b) $f=X^4+8\imath X^3-24X^2-32\imath X+16$; $z=2\imath$, $z=-2\imath$

   c) $f=(X+1)\sum_{k=0}^{n}(-1)^k X^k$ $(n\in\mathbb{N})$; $z=1$, $z=-1$

   d) $f=(X-z)^{100}+(X-z)^{15}(X^2-X+1+z-z^2)$; $z\in\mathbb{C}$

5. Si $f=\sum_i a_i X^i\in\mathbb{C}[X]$ se define el *polinomio conjugado* de f en la forma
$$\overline{f}=\sum_i \overline{a_i}\,X^i\,.$$
   Probar las siguientes propiedades ($f,g\in\mathbb{C}[X]$ y $z\in\mathbb{C}$):

   a) $\overline{f}=f$ si y solo si $f\in\mathbb{R}[X]$

   b) $\overline{f+g}=\overline{f}+\overline{g}$ y $\overline{fg}=\overline{f}\,\overline{g}$

   c) $\overline{f^s}=\overline{f}^{\,s}$ para todo $s\in\mathbb{N}$

   d) z es raíz de f si y solo si $\overline{z}$ es raíz de $\overline{f}$

   e) $\text{mult}(z,f)=\text{mult}(\overline{z},f)$.

6. En cada uno de los siguientes incisos hallar $z\in\mathbb{C}$ de manera que se satisfaga la condición planteada:

   a) $(X+1)^3 \mid X^6+(z+6)X^5+7X^4+(4-2z)X^3+7X^2+(z+6)X+1$

b) $X^3 + 4X^2 - 3X + 2z$ admite una raíz múltiple en $\mathbb{C}$

c) $X^4 - X^3 - 3X^2 - 5zX - 2$ admite una raíz triple en $\mathbb{C}$.

7. Sea $f \in \mathbb{C}[X]$ tal que $(f : f') = X^4 - 6X^3 + 10X^2 - 6X + 9$. Determinar el mínimo valor posible de $gr(f)$.

8. Si $n \in \mathbb{N}$, probar que las raíces complejas de $\sum_{k=0}^{n} (k!)^{-1} X^k$ y $\sum_{k=0}^{n} X^k$ son simples.

9. Calcular los coeficientes de grado 0 y 59 de

$$(X-1)^{30}(X+2)^{20}(X^2+X-2)^5.$$

10. En cada uno de los siguientes casos, hallar los coeficientes del único polinomio $f \in \mathbb{Q}[X]$ que satisface las condiciones dadas:

a) f es mónico de grado 4 y sus raíces son 1, -3 y 8, esta última doble.

b) f es de grado 4 con coeficiente principal 5, admite a 1, 2 y 3 como raíces simples y $f(0) = 8$.

c) f es mónico de grado 5, $f(2) = -3$ y $1 + \imath$ es raíz doble de f.

11. Hallar las raíces de $X^3 - 9\sqrt{2}X^2 + 55X - 57\sqrt{2}$ sabiendo que las mismas determinan una progresión aritmética.

12. El volumen de un paralelepípedo es 36 cm^3, su superficie lateral es 66 cm^2 y la suma de las longitudes de todas sus aristas es 40 cm. Hallar la longitud de cada una de sus aristas.

13. Probar que el polinomio $((x_2 - x_1)(x_3 - x_1)(x_3 - x_2))^2$ es simétrico y expresarlo en términos de las funciones simétricas elementales s_1, s_2 y s_3.

14. Sean α, β y γ las raíces complejas de $X^3 + 3X - 1$.

a) Calcular

i) $\alpha^2 + \beta^2 + \gamma^2$

ii) $\alpha^{-1} + \beta^{-1} + \gamma^{-1}$

iii) $\alpha\beta/\gamma + \alpha\gamma/\beta + \beta\gamma/\alpha$

iv) $\alpha^3 + \beta^3 + \gamma^3$.

b) Determinar un polinomio mónico de grado 3 cuyas raíces sean:

i) $\alpha^2,\ \beta^2$ y γ^2
ii) $\alpha+1,\ \beta+1$ y $\gamma+1$
iii) $\alpha\beta,\ \alpha\gamma$ y $\beta\gamma$
iv) $\alpha+\beta,\ \alpha+\gamma$ y $\beta+\gamma$.

10.3. Irreducibilidad y factorización única

10.3.1. Polinomios irreducibles

Existe en el anillo de polinomios con coeficientes en un cuerpo K una noción equivalente a la de número primo en $\mathbb{Z}$, en el sentido de elementos que sólo pueden factorizarse en forma trivial. Precisamente:

> Diremos que un polinomio f de grado positivo con coeficientes en K es *irreducible* en $K[X]$ si y solo si los únicos divisores de f en $K[X]$ son sus asociados y las constantes no nulas. Equivalentemente, no es posible descomponer f en $K[X]$ como producto de dos polinomios de grado positivo. En caso contrario, diremos que f es *reducible* en $K[X]$. Emplearemos también las expresiones f es irreducible o reducible sobre K, según el caso.

NOTA. Obsérvese que no consideramos irreducibles a los polinomios de grado 0, a pesar de que ellos claramente satisfacen la definición anterior. Se los excluye debido al rol trivial que juegan respecto de la divisibilidad en $K[X]$, el mismo que cumplen 1 y -1 en $\mathbb{Z}$.

Notemos además que cualquier asociado de un polinomio irreducible también es irreducible, ya que ambos tienen los mismos divisores. Deducimos de ello que todo polinomio irreducible es asociado de un único polinomio irreducible mónico.

Ejemplos 10.3.1 Como ejemplo genérico, señalemos que todo polinomio de grado 1 es irreducible, ya que el producto de dos polinomios de grado positivo tiene grado mayor o igual que 2.

Más particularmente, probemos que el polinomio $f = X^4 + 1$ es irreducible en $\mathbb{Q}[X]$. Para ello, y siendo $gr(f) = 4$, observemos que habría en principio dos formas posibles de factorizar no trivialmente a f: como producto de un polinomio de grado 1 por otro de grado 3, ó como producto de dos polinomios de grado 2. Puesto que f no admite factores de grado 1 en $\mathbb{Q}[X]$ por no tener raíces racionales (en realidad tampoco reales), solo debemos considerar la posibilidad de que f se descomponga como producto de dos factores (que podemos suponer mónicos) de grado 2, digamos

$$f = (X^2 + aX + b)(X^2 + cX + d),$$

donde $a, b, c, d \in \mathbb{Q}$.

Desarrollando el segundo miembro de la igualdad anterior e igualando coeficientes, resulta que deben satisfacerse las ecuaciones

$$a + c = 0 \tag{i}$$

$$b + ac + d = 0 \tag{ii}$$

$$ad + bc = 0 \tag{iii}$$

$$bd = 1, \tag{iv}$$

que podemos resolver con facilidad. En efecto, sigue de (i) y (iii) que $c = -a$ y $a(d-b) = 0$, de donde $a = 0$ ó $d = b$. En el primer caso resulta por (ii) y (iv) que $d = -b$ y $b^2 = -1$, lo que es absurdo pues por hipótesis b es real. Deberá ser entonces $d = b$ y por lo tanto $b^2 = 1$.

Sin embargo, esto último también nos conduce a una contradicción, ya que multiplicando los dos miembros de (ii) por b resulta $2 - a^2 = 0$, esto es, $a = \sqrt{2}$ ó $a = -\sqrt{2}$. Luego el sistema de ecuaciones de arriba no admite soluciones racionales, y en consecuencia $X^4 + 1$ es irreducible sobre $\mathbb{Q}$.

Los cálculos precedentes parecen indicar que las ecuaciones (i) a (iv) son resolubles en $\mathbb{R}$, hecho que permitiría factorizar $X^4 + 1$ en $\mathbb{R}[X]$ de manera no trivial. Para mostrar que efectivamente es así podemos factorizar directamente f, completando el cuadrado y aplicando un conocido caso de factoreo. Concretamente, tenemos:

$$X^4 + 1 = \left(X^2 + 1\right)^2 - 2X^2 = \left(X^2 + 1\right)^2 - \left(\sqrt{2}X\right)^2 =$$
$$= \left(X^2 + \sqrt{2}X + 1\right)\left(X^2 - \sqrt{2}X + 1\right).$$

Vale decir, $X^4 + 1$ se factoriza en $\mathbb{R}[X]$ (luego también en $\mathbb{C}[X]$) como producto de dos polinomios de grado 2, y por lo tanto es *reducible* sobre $\mathbb{R}$ y sobre $\mathbb{C}$. Esto nos muestra que la noción de irreducibilidad de un polinomio es relativa al cuerpo al que pertenecen sus coeficientes. $\Diamond$

Tanto el concepto de raíz de un polinomio como el de irreducibilidad del mismo están ligados a la idea de factorización, por lo que es razonable esperar que haya algún tipo de conexión entre ambos. Precisaremos la cuestión en la siguiente proposición, en la que además nos ocuparemos de otras propiedades de los polinomios irreducibles, equiparables a las que satisfacen los números primos en $\mathbb{Z}$.

Proposición 10.3.2 Son válidas en $K[X]$ las siguientes propiedades:

1) Si f es irreducible en $K[X]$ y $gr(f) > 1$ entonces f no tiene raíces en K

2) Sea $f \in K[X]$ un polinomio de grado 2 ó 3, sin raíces en K. Entonces f es irreducible sobre K

3) Sea f irreducible en $K[X]$ y sea $g \in K[X]$. Entonces $f \mid g$ ó f y g son coprimos

4) Sea f irreducible en $K[X]$ y sean $g, h \in K[X]$ tales que $f \mid gh$. Entonces $f \mid g$ ó $f \mid h$

5) Más generalmente, sea f irreducible en $K[X]$ y sean $g_1, g_2, \ldots, g_m$ en $K[X]$ tales que
$$f \mid \prod_{i=1}^{m} g_i.$$

Existe entonces un índice k tal que $f \mid g_k$.

DEMOSTRACION. Las propiedades 3), 4) y 5) se demuestran como en el caso de números enteros, por lo que dejamos los detalles a cargo del lector. En cuanto a las dos primeras, notemos antes de demostrarlas que 2) es recíproca de 1), pero sólo para polinomios de grado 2 ó 3, ya que en general *dicha recíproca es falsa.* Por ejemplo, el polinomio $(X^2+1)(X^2+3)$ es (obviamente) reducible sobre $\mathbb{R}$, aún cuando no tiene raíces reales.

Para probar 1), supongamos por el absurdo que f admite una raíz c en K. Como sabemos, f se factoriza en tal caso en la forma $f = (X-c)t$, con $t \in K[X]$, resultando entonces por hipótesis que

$$gr(t) = gr(f) - 1 > 1 - 1 = 0,$$

lo que contradice la irreducibilidad de f.

En cuanto a 2), es claro por la hipótesis sobre su grado que cualquier posible factorización no trivial de f en $K[X]$ debiera contener un factor de grado 1. Puesto que esto es imposible, por no tener f raíces en K, concluimos que es irreducible sobre K. $\Diamond$

Factorización única.

La noción de irreducibilidad y sus propiedades conducen a un teorema de factorización única en el anillo de polinomios con coeficientes en K, similar al teorema fundamental de la Aritmética en $\mathbb{Z}$. Su enunciado preciso es el siguiente (todos los polinomios tienen coeficientes en K y la irreducibilidad mencionada es sobre K):

Teorema 10.3.3 Todo polinomio no nulo f se factoriza unívocamente en la forma

$$f = a \prod_{i=1}^{r} f_i^{m_i}, \tag{10.9}$$

donde los f_i son polinomios irreducibles mónicos distintos, los exponentes m_i son números naturales, $a \in K^*$ y $r \in \mathbb{N}_0$.

DEMOSTRACION. Resulta sencillo probar por inducción en el grado de f la existencia de una factorización como la del enunciado, mientras que la unicidad de la misma es consecuencia de la propiedad 5) de 10.3.2. La situación es enteramente análoga a la del anillo $\mathbb{Z}$, por lo que encomendamos al lector la tarea de completar los detalles de la demostración. $\Diamond$

NOTA. El lector advertirá la fuerte analogía existente entre la situación que acabamos de describir y la ya estudiada en el caso de $\mathbb{Z}$. De todos modos, vale la pena señalar algunas peculiaridades que son propias del anillo de polinomios:

(i) Los polinomios irreducibles mónicos desempeñan en $K[X]$ el mismo rol que los primos positivos en $\mathbb{Z}$. Así, los divisores irreducibles de f en $K[X]$ son exactamente los polinomios irreducibles asociados a los f_i. También como en $\mathbb{Z}$, $f_i^{m_i}$ es la mayor potencia de f_i que divide a f.

(ii) Puesto que los f_i son mónicos es claro que la constante a es el coeficiente principal de f. Además, $r = 0$ si y solo si $gr(f) = 0$.

(iii) Considerando que c es raíz de f si y solo si $X - c \mid f$, sigue que el número de raíces distintas de f en K coincide con el número de factores f_i de grado 1 en la descomposición (10.9). Más aún, si $f_i = X - c_i$ resulta que $m_i = \text{mult}(c_i, f)$, por definición de multiplicidad.

(iv) Conocidas las factorizaciones de dos polinomios f y g, es tarea sencilla calcular el máximo común divisor y el mínimo común múltiplo de f y g, a través de fórmulas idénticas a las que exhibimos en el caso de números enteros. Se obtiene en particular que f y g son coprimos si y solo si sus factorizaciones no tienen ningún factor irreducible mónico en común. $\Diamond$

Ejemplo 10.3.4 Consideremos en $\mathbb{Q}[X]$ el polinomio

$$f = 2X^7 - 4X^6 - 6X^5 + 4X^4 + 6X^3 + 12X^2 + 14X + 4\,.$$

Utilizando el criterio de Gauss, resulta que las únicas raíces racionales de f son 2 (simple) y -1 (doble), lo que nos permite hallar una primera descomposición de f, a saber:

$$f = 2(X-2)(X+1)^2(X^4 - 2X^3 - 2X - 1).$$

Si bien el polinomio $h = X^4 - 2X^3 - 2X - 1$ no tiene raíces racionales, podemos factorizarlo agrupando adecuadamente sus términos, en la forma

$$h = (X^2-1)(X^2+1) - 2X(X^2+1) = (X^2+1)(X^2 - 2X - 1).$$

Puesto que los dos últimos factores de h son irreducibles sobre $\mathbb{Q}$, por ser de grado 2 y sin raíces racionales (proposición 10.3.2), resulta que

$$f = 2(X - 1/2)(X+1)^2(X^2 - 2X - 1)(X^2+1).$$

es la factorización completa de f en $\mathbb{Q}[X]$.

Si miramos a f como polinomio con coeficientes reales ó complejos podemos refinar su factorización, ya que $X^2 - 2X - 1$ admite las raíces reales $1 + \sqrt{2}$ y $1 - \sqrt{2}$ y $X^2 + 1$ se anula en $\imath$ y $-\imath$. Reuniendo esta información concluimos que

$$f = 2(X-2)(X+1)^2(X - (1+\sqrt{2}))(X - (1-\sqrt{2}))(X^2+1)$$

es la factorización de f en $\mathbb{R}[X]$ y

$$f = 2(X-2)(X+1)^2(X - (1+\sqrt{2}))(X - (1-\sqrt{2}))(X-\imath)(X+\imath)$$

es su factorización en $\mathbb{C}[X]$. $\Diamond$

Ejemplo 10.3.5 Factoricemos sobre $\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$ el polinomio $f = X^{12} - 1$. El caso más sencillo es sin duda el caso complejo, ya que f tiene 12 raíces en $\mathbb{C}$ (las raíces duodécimas de la unidad) y por lo tanto se descompone en $\mathbb{C}[X]$ en la forma

$$f = \prod_{w \in G_{12}} (X - w)\,.$$

Para resolver los otros casos, podemos factorizar f empleando reiteradamente ciertas identidades válidas en cualquier anillo conmutativo. Precisamente, tenemos:

$$\begin{aligned} f &= (X^6 - 1)(X^6 + 1) = (X^3 - 1)(X^3 + 1)(X^6 + 1) = \\ &= (X - 1)(X^2 + X + 1)(X + 1)(X^2 - X + 1)(X^2 + 1)(X^4 - X^2 + 1). \end{aligned}$$

Es inmediato verificar que ninguno de los factores de grado 2 del miembro de la derecha admite raíces reales (sus discriminantes son negativos), resultando entonces que ellos son irreducibles sobre $\mathbb{R}$, y por lo tanto también sobre $\mathbb{Q}$. Por otro lado, completando cuadrados podemos descomponer $g = X^4 - X^2 + 1$ en la forma:

$$\begin{aligned} g &= \left(X^2 + 1\right)^2 - 3X^2 = \left(X^2 + 1\right)^2 - \left(\sqrt{3}X\right)^2 = \\ &= \left(X^2 + \sqrt{3}X + 1\right)\left(X^2 - \sqrt{3}X + 1\right), \end{aligned}$$

esto es, g se factoriza en $\mathbb{R}[X]$ como producto de dos polinomios u y v de grado 2. Puesto que los dos tienen discriminante -1 ambos son irreducibles sobre $\mathbb{R}$.

Veamos finalmente que g es irreducible sobre $\mathbb{Q}$. Supongamos por el contrario que g admite una factorización no trivial en $\mathbb{Q}[X]$, digamos $g = ht$, con h y t polinomios mónicos de grado positivo. Puesto que g no tiene raíces racionales (ni reales), es claro que ambos factores deben ser de grado 2.

Ahora bien, de las igualdades $g = uv = ht$ sigue que u divide a ht, y siendo u irreducible en $\mathbb{R}[X]$ resulta por propiedad 4) de 10.3.2 que divide a uno de los factores, digamos $u \mid h$. Puesto que ambos polinomios son mónicos de grado 2 concluimos que $u = h$, lo que es absurdo, pues $u \notin \mathbb{Q}[X]$. Luego g es irreducible sobre $\mathbb{Q}$.

En conclusión, $f = X^{12} - 1$ se factoriza en $\mathbb{Q}[X]$ en la forma

$$f = (X - 1)(X + 1)(X^2 + X + 1)(X^2 - X + 1)(X^2 + 1)(X^4 - X^2 + 1),$$

mientras que su descomposición en $\mathbb{R}[X]$ es

$$f = (X-1)(X+1)(X^2+X+1)(X^2-X+1)(X^2+1)(X^2+\sqrt{3}X+1)(X^2-\sqrt{3}X+1).$$

Es interesante señalar que en la factorización de f sobre $\mathbb{Q}$ los factores se corresponden con los divisores positivos de 12, hecho que tiene conexión con un ejemplo que exhibimos en el capitulo 8, respecto a la clasificación de

los elementos de G_{12} según su orden. Precisamente, los ceros de los factores cuadráticos de f son las raíces duodécimas de la unidad de orden 3, 6, y 4, respectivamente, las raíces de $X^4 - X^2 + 1$ son las raíces primitivas de orden 12, y, obviamente, las de $X - 1$ y $X + 1$ son las primitivas de orden 1 y 2, respectivamente. Subrayemos que esta distribución de las raíces de la unidad como ceros de los distintos factores irreducibles de $X^n - 1$ en $\mathbb{Q}[X]$ es válida cualquiera sea $n \in \mathbb{N}$, aunque la demostración general de este hecho es bastante complicada y supera el nivel que nos hemos propuesto alcanzar en estas páginas $\quad \Diamond$

10.3.2. Ejercicios

En los ejercicios que siguen la letra K designa cualquiera de los cuerpos numéricos $\mathbb{Q}$, $\mathbb{R}$ ó $\mathbb{C}$.

1. Sean K y L cuerpos numéricos tal que $K \subset L$. Si $f \in K[X]$ es irreducible sobre L probar que f es irreducible sobre K.

2. En cada uno de los siguientes ítems factorizar completamente el polinomio dado en $\mathbb{Q}[X]$, $\mathbb{R}[X]$ y $\mathbb{C}[X]$. Entendemos por factorización completa su descomposición en $K[X]$ como producto de polinomios irreducibles mónicos (salvo factor constante):

   a) $2X^5 - 14X^4 + 28X^3 - 6X^2 - 18X$

   b) $2X^4 - 4X^2 - 30$

   c) $X^6 - X^4 - 14X^2 + 24$

   d) $(4X^4 + 6X^3 + 8X^2 + 2X - 4)^3(X^2 - 2X - 3)^2$

   e) $X^8 - 16$.

3. Supongamos que un polinomio de grado 8 con coeficientes reales tiene una raíz real doble, dos raíces reales simples y cuatro raíces complejas no reales. ¿Cuántos factores irreducibles distintos puede contener su factorización en $\mathbb{R}[X]$? Ejemplificar.

4. Utilizando la información suministrada, factorizar f sobre $\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$ en cada uno de los siguientes casos:

   a) $f = X^4 - X^3 + 13X^2 - 9X + 36$, sabiendo que tiene una raíz imaginaria pura.

   b) $f = 4X^4 - 31X^2 + 5X + 44$, sabiendo que tiene una raíz de la forma $a + \sqrt{3}$ $(a \in \mathbb{Q})$.

c) $f = 7X^6 + 6X^5 + 18X^4 + 48X^3 - 192X^2 + 96X - 608$, sabiendo que tiene una raíz múltiple.

d) $f = X^6 - X^5 + 4X^4 - 3X^3 + 5X^2 - 2X + 2$, sabiendo que tiene entre sus ceros una raíz sexta primitiva de la unidad.

5. *a*) Sea f un polinomio irreducible sobre K. Demostrar que todas sus raíces complejas son simples.

   b) Sea f un polinomio con coeficientes en K que se factoriza en $K[X]$ como producto de polinomios irreducibles distintos. Demostrar que todas las raíces complejas de f son simples.

6. Sean $f, g \in K[X]$ polinomios irreducibles mónicos distintos. Probar que f y g no tienen raíces comunes en $\mathbb{C}$.

Capítulo 11

Polinomios sobre cuerpos numéricos

11.1. Irreducibilidad y factorización

11.1.1. Teorema Fundamental del Algebra

En el capítulo anterior desarrollamos la teoría de divisibilidad de polinomios con coeficientes en cualquier cuerpo K, hasta arribar al teorema de factorización única en polinomios irreducibles. Vamos ahora a particularizar algunos de los resultados obtenidos, concentrando nuestra atención en los anillos de polinomios con coeficientes en los cuerpos numéricos $\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$. Comenzando con el caso complejo, recordemos que en la construcción de $\mathbb{C}$ nos movía el propósito de obtener una extensión de $\mathbb{R}$ en la cual todo número real negativo admitiera una raíz cuadrada. Vimos luego que obteníamos bastante más, al probar en el capítulo 8 que todo número complejo admite raíces $n-$ésimas en $\mathbb{C}$ cualquiera sea $n \in \mathbb{N}$. A través de la formulación del siguiente resultado, conocido como *Teorema Fundamental del Algebra*, apreciaremos a continuación que el logro es aún más importante: *toda ecuación algebraica de grado positivo con coeficientes complejos es resoluble en* $\mathbb{C}$. Su enunciado preciso es el siguiente:

Teorema 11.1.1 Todo polinomio no constante con coeficientes en $\mathbb{C}$ admite al menos una raíz en $\mathbb{C}$. $\Diamond$

Debido a este hecho se dice que $\mathbb{C}$ es un cuerpo *algebraicamente cerrado*. El teorema fue probado por D'Alambert (1746) y por Gauss en su tesis doctoral (1799), pero ambas demostraciones, sin ser incorrectas, utilizaban resultados que no fueron probados hasta mucho después. Más tarde, en 1814, fue demostrado correctamente por Jean Robert Argand mediante métodos geométricos. Se conocen actualmente muchas pruebas distintas del teorema, y todas emplean en mayor o menor medida recursos del análisis matemático.

Por razones de complejidad no brindaremos aquí ninguna de ellas, pero exhibiremos a continuación varias de sus muchas e importantes consecuencias.

Corolario 11.1.2 Un polinomio con coeficientes complejos es irreducible en $\mathbb{C}[X]$ si y solo si es de grado uno.

DEMOSTRACION. Sabemos que los polinomios de grado uno son irreducibles sobre cualquier cuerpo K y que los polinomios irreducibles de grado mayor que uno con coeficientes en K no tienen raíces en K (proposición 10.3.2). Puesto que cualquier polinomio de grado positivo con coeficientes en $\mathbb{C}$ tiene alguna raíz en $\mathbb{C}$, concluimos que los únicos elementos irreducibles del anillo $\mathbb{C}[X]$ son los polinomios de grado uno. $\Diamond$

Factorización en C[X] .

Deducimos del corolario precedente que todo polinomio no nulo se *descompone linealmente* en $\mathbb{C}[X]$. Precisamente:

Proposición 11.1.3 Todo polinomio no nulo $f \in \mathbb{C}[X]$ se factoriza unívocamente en la forma

$$f = w \prod_{i=1}^{r} (X - z_i)^{m_i}\,, \tag{11.1}$$

donde $r \geq 0$, $w \in \mathbb{C}^*$, los z_i son números complejos distintos y los exponentes m_i son números naturales.

DEMOSTRACION. Obtenemos una descomposición de tipo (11.1) aplicando al caso $K = \mathbb{C}$ el teorema 10.3.3 de factorización única, habida cuenta de que por el corolario anterior los elementos irreducibles mónicos de $\mathbb{C}[X]$ son de la forma $X - z$. $\Diamond$

El siguiente resultado afina el enunciado del teorema fundamental del Algebra y justifica la denominación de cuerpo algebraicamente cerrado dada a $\mathbb{C}$.

Corolario 11.1.4 Sea $n \in \mathbb{N}$ y sea $f \in \mathbb{C}[X]$ de grado n. Entonces f tiene exactamente n raíces en $\mathbb{C}$ (cada una de ellas contada según su multiplicidad).

DEMOSTRACION. Escribiendo como antes

$$f = w \prod_{i=1}^{r} (X - z_i)^{m_i}\,,$$

resulta claro que $\{z_1, \cdots, z_r\}$ es el conjunto de raíces distintas de f en $\mathbb{C}$ y que $\text{mult}(z_i, f) = m_i$ para todo i. Obtenemos entonces:

$$n = gr(f) = \sum_{i=1}^{r} gr\left((X - z_i)^{m_i}\right) = \sum_{i=1}^{r} m_i,$$

como queríamos demostrar. $\Diamond$

El hecho de que todo polinomio con coeficientes complejos admita un juego completo de raíces en $\mathbb{C}$ permite describir las relaciones de divisibilidad en $\mathbb{C}[X]$ en términos de raíces. Veamos algunas muestras de esta afirmación.

Proposición 11.1.5 Si $f, g \in \mathbb{C}[X]$, son válidas las siguientes propiedades:

1) f y g son coprimos si y solo si f y g no tienen raíces comunes en $\mathbb{C}$

2) $f \mid g$ si y solo si $\text{mult}(z, f) \leq \text{mult}(z, g)$ para toda raíz z de f

3) f y g son asociados si y solo si $\text{mult}(z, f) = \text{mult}(z, g)$ para todo $z \in \mathbb{C}$.

Demostracion. Para probar 1), recordemos (corolario 10.1.10) que las raíces comunes de f y g son exactamente las raíces de $(f : g)$. Luego f y g no tienen raíces comunes en $\mathbb{C}$ si y solo si $(f : g)$ no tiene raíces en $\mathbb{C}$, lo cual es equivalente a la condición $(f : g) = 1$, por teorema fundamental del álgebra.

Ya hemos demostrado (proposición 10.2.5) que la condición del enunciado de 2) es necesaria cualquiera sea el cuerpo K de coeficientes. En cuanto a la suficiencia de la misma en $\mathbb{C}[X]$, si

$$f = v \prod_{i=1}^{s} (X - u_i)^{s_i}$$

es la factorización única de f sobre $\mathbb{C}$ observemos que la hipótesis asegura que cada raíz u_i de f es también raíz de g y que el factor lineal $X - u_i$ aparece en la descomposición de g con un exponente $t_i \geq s_i$, lo que obviamente implica que f es un factor de g, como se quería probar.

La parte 3) es consecuencia directa de 2), ya que dos polinomios son asociados si y solo si se dividen mutuamente. $\Diamond$

Ejemplo 11.1.6 Sean m y n números naturales y sea $d = (m : n)$. Entonces

$$(X^m - 1 : X^n - 1) = X^d - 1\,.$$

Si bien vale sobre cualquier cuerpo K, demostraremos la igualdad anterior en $\mathbb{C}[X]$. Puesto que para cualquier $k \in \mathbb{N}$ el conjunto de raíces de $f_k = X^k - 1$ es G_k, resulta que el conjunto de raíces comunes de f_m y f_n es $G_m \cap G_n = G_d$. Por lo tanto, $(f_m : f_n)$ y f_d tienen las mismas raíces. Teniendo en cuenta que en ambos casos ellas son simples, sigue de la proposición anterior que $(f_m : f_n)$ y f_d son asociados. Al ser mónicos, concluimos que son iguales. $\Diamond$

POLINOMIOS CICLOTOMICOS Si $n \in \mathbb{N}$, designaremos por H_n el conjunto de raíces $n-$ésimas primitivas de la unidad, y llamaremos *polinomio ciclotómico* de orden n al polinomio

$$\Phi_n = \prod_{w \in H_n} (X - w),$$

vale decir, Φ_n es mónico y sus ceros son las raíces $n-$ésimas primitivas de la unidad. Sigue en particular que $gr(\Phi_n) = \varphi(n)$.

Es fácil verificar por ejemplo que $\Phi_1 = X - 1$, $\Phi_2 = X + 1$ y $\Phi_4 = X^2 + 1$. Asimismo, $\Phi_p = X^{p-1} + X^{p-2} + \ldots + X + 1$ si p es primo, teniendo en cuenta la factorización

$$X^p - 1 = (X - 1)(X^{p-1} + X^{p-2} + \ldots + X + 1)$$

y el hecho de que toda raíz $p-$ésima de la unidad distinta de 1 es primitiva.

Como se puede apreciar, los polinomios ciclotómicos de los ejemplos anteriores tienen coeficientes enteros. Veremos a continuación que tal hecho no es casual.

Proposición 11.1.7 $\Phi_n \in \mathbb{Z}[X]$ para todo $n \in \mathbb{N}$.

DEMOSTRACION Procederemos por inducción. Puesto que el resultado vale para $n = 1$, probémoslo para $n > 1$ suponiendo que $\Phi_k \in \mathbb{Z}[X]$ para todo k menor que n.

Si d es un divisor positivo de n valen las inclusiones $H_d \subseteq G_d \subseteq G_n$, mientras que por otro lado vimos en el capítulo 8 que todo elemento de G_n es una raíz $m-$ésima primitiva de 1 para algún divisor m de n. La conjunción de ambos hechos brinda entonces la igualdad

$$G_n = \bigcup_{d|n} H_d \,.$$

Puesto que G_n y H_d son los conjuntos de raíces de $X^n - 1$ y Φ_d, respectivamente, y $H_r \cap H_s = \emptyset$ si $r \neq s$, dicha igualdad de conjuntos se traduce en la igualdad de polinomios

$$X^n - 1 = \prod_{d|n} \Phi_d = \Phi_n \prod_{\substack{d|n \\ d<n}} \Phi_d \,.$$

Designando por h el polinomio de la productoria del segundo miembro, resulta por hipótesis inductiva que h tiene coeficientes enteros. En consecuencia, Φ_n es el cociente de la división de dos elementos de $\mathbb{Z}[X]$, a saber, $X^n - 1$ y h. Como éste además es mónico, concluimos que $\Phi_n \in \mathbb{Z}[X]$. $\Diamond$

Ejemplos 11.1.8 Si bien existen fórmulas que en teoría permitirían determinar la expansión de Φ_n para cualquier n, no siempre es posible utilizarlas,

ya que su aplicación requiere conocer todos los divisores de n. De todos modos vamos a calcular Φ_n para ciertos valores bajos de n, mostrando qué tipo de recursos pueden usarse. Comencemos por Φ_9.

Si z es una raíz novena primitiva de la unidad sigue inmediatamente que z^3 es una raíz cúbica primitiva de 1, y por lo tanto es un cero de $\Phi_3 = X^2 + X + 1$. Especializando, resulta entonces que

$$0 = \Phi_3(z^3) = (z^3)^2 + z^3 + 1 = z^6 + z^3 + 1.$$

Hemos probado así que toda raíz de Φ_9 es raíz del polinomio X^6+X^3+1, y puesto que ambos polinomios son mónicos y de igual grado ($\varphi(9) = 6$), concluimos que $\Phi_9 = X^6 + X^3 + 1$.

Determinemos ahora los coeficientes de Φ_{18}, cuyo grado coincide con el de Φ_9. Observemos para ello que $z^9 = -1$ para todo $z \in H_{18}$, ya que $(z^9)^2 = 1$ y $z^9 \neq 1$. Resulta entonces que $-z$ es una raíz primitiva de orden 9, ya que $(-z)^9 = -z^9 = 1$, y si fuera $(-z)^3 = 1$ tendríamos $z^6 = (z^3)^2 = (-1)^2 = 1$, contradiciendo el hecho de que z tiene orden 18.

Luego, usando que $-z$ es raíz de $\Phi_9 = X^6 + X^3 + 1$, obtenemos:

$$0 = \Phi_9(-z) = (-z)^6 + (-z)^3 + 1 = z^6 - z^3 + 1.$$

Resulta por lo tanto que $\Phi_{18} = X^6 - X^3 + 1$, ya que ambos son mónicos de grado 6 y cada una de las 6 raíces simples del primero también es raíz del segundo. Más generalmente, el lector puede verificar que de manera completamente análoga se obtiene la expansión de Φ_{2m} a partir de la de Φ_m para todo m impar.

Calculemos por último Φ_{15}. Procediendo como en los casos anteriores, tomemos z en H_{15}. En tal caso z^5 es una raíz primitiva de orden 3, lo que nos permite deducir que z es raíz del polinomio $f = X^{10} + X^5 + 1$. Por otra parte, si w es una raíz cúbica primitiva de la unidad tenemos que

$$w^{10} + w^5 + 1 = w + w^2 + 1 = \Phi_3(w) = 0,$$

esto es, w es raíz de f. Por lo tanto, los ceros de f son los 8 elementos de H_{15} (notar que $\varphi(15) = 8$) y las dos raíces de $X^2 + X + 1$, de donde sigue en particular que este último polinomio es un divisor de f. Efectuando la división, se obtiene la factorización

$$f = X^{10} + X^5 + 1 = (X^2 + X + 1)(X^8 - X^7 + X^5 - X^4 + X^3 - X + 1),$$

como el lector puede constatar. Iterando argumentos ya usados en varias oportunidades, concluimos que

$$\Phi_{15} = X^8 - X^7 + X^5 - X^4 + X^3 - X + 1\,. \quad \Diamond$$

Factorización en R[X].

A partir del teorema fundamental del álgebra también es posible caracterizar los elementos irreducibles de $\mathbb{R}[X]$. Precisamente:

Proposición 11.1.9 Todo polinomio f con coeficientes reales de grado mayor que 2 es reducible sobre $\mathbb{R}$. Equivalentemente, los únicos polinomios irreducibles de $\mathbb{R}[X]$ son los de grado 1 y los de grado 2 con discriminante negativo.

DEMOSTRACION. La equivalencia de ambas afirmaciones es clara, ya que un polinomio de grado 2 con coeficientes en cualquier cuerpo K es irreducible (sobre K) si y solo si no tiene raíces en K. Bastará probar luego la primera afirmacion.

Consideremos para ello una raíz z de f en $\mathbb{C}$. Si $z \in \mathbb{R}$ es claro que f es reducible en $\mathbb{R}[X]$. Si $z \notin \mathbb{R}$, resulta que z y $\overline{z}$ son raíces distintas de f en $\mathbb{C}$, y por lo tanto el polinomio $g = (X - z)(X - \overline{z})$ es un divisor de f en $\mathbb{C}[X]$, digamos $f = gh$.

Puesto que $g \in \mathbb{R}[X]$ y h es el cociente de la división de f por g sigue que h también pertenece a $\mathbb{R}[X]$. Teniendo en cuenta por último que

$$gr(h) = gr(f) - 2 > 0\,,$$

concluimos que f es reducible sobre $\mathbb{R}$, como queríamos probar. $\Diamond$

De acuerdo con esta caracterización de irreducibles, resulta que todo polinomio no nulo con coeficientes reales se descompone como producto de polinomios de grado 1 ó 2. Precisamente, el teorema de factorización única adopta en $\mathbb{R}[X]$ la siguiente forma (los límites de las productorias y los exponentes son como en el caso complejo, y todos los coeficientes son números reales):

Proposición 11.1.10 Todo polinomio no nulo $f \in \mathbb{R}[X]$ se factoriza unívocamente en la forma

$$f = a \prod_{i=1}^{r} (X - b_i)^{m_i} \prod_{i=1}^{s} (X^2 + c_i X + d_i)^{n_i}\,, \tag{11.2}$$

donde $a \in \mathbb{R}^*$ y $c_i^2 - 4d_i < 0$ para todo i. Como siempre, suponemos que todos los factores irreducibles son distintos.

DEMOSTRACION. Basta emplear el teorema de factorización única en $\mathbb{R}[X]$ y la caracterización de irreducibles dada por la proposición 11.1.9 .

Observemos que $b_1, \ldots, b_r$ son las raíces distintas de f en $\mathbb{R}$, mientras que las raíces distintas de f en $\mathbb{C}$ son $z_1, \overline{z}_1, \ldots, z_s, \overline{z}_s$, donde z_i es raíz de $X^2 + c_i X + d_i$.

Por otra parte, igualando grados en (11.2) obtenemos

$$gr(f) = \sum_{i=1}^{r} m_i + 2\sum_{i=1}^{s} n_i\,,$$

de donde resulta que el número de raíces reales de un polinomio con coeficientes reales tiene la misma paridad que su grado. En particular, esto brinda otra demostración de un hecho ya establecido en el capitulo 9, a saber, que un polinomio con coeficientes reales de grado impar admite al menos una raíz real. $\Diamond$

Ejemplo 11.1.11 Factorizemos en $\mathbb{R}[X]$ el polinomio

$$f = X^8 - 5X^7 + 5X^6 - 3X^5 + 5X^4 + 9X^3 - X^2 + 7X - 2\,.$$

Empleando el criterio de Gauss resulta que f admite dos raíces racionales simples (-1 y 2), lo que determina la factorización

$$f = (X+1)(X-2)(X^6 - 4X^5 + 3X^4 - 8X^3 + 3X^2 - 4X + 1),$$

como el lector puede comprobar. Designando por g el último factor de arriba y por a_i sus coeficientes, observemos que $a_i = a_{6-i}$ para todo i. Esta peculiaridad permite calcular con cierta facilidad sus raíces (que obviamente son distintas de cero), ya que si z es cualquiera de ellas, dividiendo por z^3 la igualdad $0 = g(z)$ y operando convenientemente, obtenemos:

$$\begin{aligned} 0 &= z^3 - 4z^2 + 3z - 8 + 3z^{-1} - 4z^{-2} + z^{-3} = \\ &= z^3 + 3z + 3z^{-1} + z^{-3} - 4z^2 - 8 - 4z^{-2} = \\ &= \left(z + z^{-1}\right)^3 - 4\left(z + z^{-1}\right)^2 = \left(z + z^{-1}\right)^2\left(z + z^{-1} - 4\right). \end{aligned}$$

Por lo tanto, $z + z^{-1} = 0$ ó $z + z^{-1} = 4$. Multiplicando en cada caso por z, arribamos a las condiciones equivalentes $z^2 + 1 = 0$ ó $z^2 - 4z + 1 = 0$, lo que nos muestra que los ceros de g son $\imath$ y $-\imath$ (ambos con multiplicidad 2) y las raíces $2 + \sqrt{3}$ y $2 - \sqrt{3}$ del polinomio $X^2 - 4X + 1$. En conclusión,

$$f = (X+1)\,(X-2)\left(X - (2+\sqrt{3})\right)\left(X - (2-\sqrt{3})\right)\left(X^2+1\right)^2$$

es la factorizacion de f en $\mathbb{R}[X]$, mientras que su descomposicion en $\mathbb{C}[X]$ es

$$f = (X+1)\,(X-2)\left(X - (2+\sqrt{3})\right)\left(X - (2-\sqrt{3})\right)(X-\imath)^2\,(X+\imath)^2\,. \quad \Diamond$$

Factorización en Q[X].

La situación en $\mathbb{Q}[X]$ es esencialmente distinta a la de $\mathbb{R}[X]$ y $\mathbb{C}[X]$, ya que existen polinomios irreducibles de cualquier grado sobre $\mathbb{Q}$. Justificaremos esta afirmación exhibiendo el siguiente ejemplo general.

Ejemplo 11.1.12 $X^n - 2$ es irreducible en $\mathbb{Q}[X]$ para todo $n \in \mathbb{N}$.

Fijado n, consideremos cualquier factor mónico de grado positivo de $X^n - 2$ en $\mathbb{Q}[X]$, digamos

$$f = X^s + \sum_{i=0}^{s-1} c_i X^i.$$

Puesto que toda raíz de f en $\mathbb{C}$ es entonces una raíz $n-$ésima de 2, resulta que f se factoriza en $\mathbb{C}[X]$ en la forma

$$f = \prod_{i=1}^{s} (X - \sqrt[n]{2}\, w_i),$$

para ciertos elementos $w_1, \ldots, w_s$ de G_n. Igualando los coeficientes de grado 0 en la expresión de arriba y tomando módulos, obtenemos

$$|\, c_0 \,| = \left| (-1)^s \prod_{i=1}^{s} \sqrt[n]{2}\, w_i \right| = \left(\sqrt[n]{2} \right)^s \prod_{i=1}^{s} |\, w_i \,| = \sqrt[n]{2^s}\,,$$

y por lo tanto $\sqrt[n]{2^s}$ es un número racional, digamos $\sqrt[n]{2^s} = a/b$ $(a, b \in \mathbb{N})$.

Utilizando el teorema fundamental de la aritmética, deducimos de la igualdad $2^s b^n = a^n$ que 2^s es una potencia $n-$ésima en $\mathbb{Z}$, y por lo tanto s es un múltiplo de n. Puesto que además $s \leq n$ (por ser f un divisor de $X^n - 2$), concluimos que $s = n$. Luego, $X^n - 2$ sólo admite factores de grado 0 ó n en $\mathbb{Q}[X]$ y en consecuencia es irreducible sobre $\mathbb{Q}$. $\quad \Diamond$

Si bien en general es difícil decidir acerca de la reducibilidad o irreducibilidad de un elemento de $\mathbb{Q}[X]$, es posible emplear en ocasiones algunos criterios útiles. A manera de ilustración, y luego de establecer unos pocos hechos necesarios, exhibiremos el criterio de irreducibilidad de *Eisenstein*, aplicable a polinomios con coeficientes enteros.

Como ya hemos comentado a principios del capítulo 10, tiene perfecto sentido aplicar el concepto de divisibilidad a polinomios con coeficientes en un anillo conmutativo cualquiera, aunque éste no sea un cuerpo. Así, y como caso particular de dicha situación, dados $m \in \mathbb{Z}$ y $g \in \mathbb{Z}[X]$ decimos que m *divide* a g si y solo si existe $f \in \mathbb{Z}[X]$ tal que $g = mf$. Notaremos este hecho, equivalente a afirmar que todos los coeficientes de g son múltiplos de m, en la forma usual $m \mid g$. Respecto de esta cuestión, veamos que los números primos conservan en $\mathbb{Z}[X]$ la propiedad fundamental que los caracteriza en $\mathbb{Z}$.

Proposición 11.1.13 Sea p un primo y sean $g, h \in \mathbb{Z}[X]$ tales que $p \mid gh$. Entonces $p \mid g$ ó $p \mid h$.

DEMOSTRACION. Designando por a_i, b_i y c_i los coeficientes de g, h y gh, respectivamente, supongamos que ni g ni h son divisibles por p. Usando el principio de buena ordenación, consideremos en tal caso el menor índice k tal que $p \nmid a_k$ y el menor índice l tal que $p \nmid b_l$. Calculando el coeficiente de grado $k+l$ de gh, y tomando congruencias módulo p, resulta entonces que

$$0 \equiv c_{k+l} = \sum_i a_i b_{k+l-i} = a_k b_l + \sum_{i \neq k} a_i b_{k+l-i} \equiv a_k b_l \,,$$

pues $i < k$ ó $k+l-i < l$ en cada término de la última sumatoria. Arribamos de esta forma a la deseada contradicción, ya que p es primo y no divide a ninguno de los dos factores del producto $a_k b_l$. ♢

Lema 11.1.14 Sea $m \in \mathbb{N}$ y sean $f, g, h \in \mathbb{Z}[X]$ tales que $mf = gh$. Existen entonces polinomios u y v con coeficientes enteros tales que $gr(u) = gr(g)$, $gr(v) = gr(h)$ y $f = uv$.

DEMOSTRACION. La afirmación es trivial si $m = 1$. Suponiendo $m > 1$ y empleando un argumento inductivo, bastará probar el resultado para m primo (dejamos los detalles a cargo del lector). En tal caso, usando la proposición 11.1.13 podemos suponer sin pérdida de generalidad que m divide a g, digamos $g = mg_1$, donde g_1 es un polinomio con coeficientes enteros del mismo grado que g. Cancelando el factor m en ambos miembros de la igualdad $mf = (mg_1)h$ obtenemos $f = g_1 h$, y el resultado sigue tomando $u = g_1$ y $v = h$. ♢

Ya podemos enunciar y demostrar el criterio de irreducibilidad que mencionamos anteriormente:

Proposición 11.1.15 (Criterio de Eisenstein) Sea p un primo y sea

$$f = \sum_{i=0}^{n} a_i X^i$$

un polinomio con coeficientes enteros tal que $p \mid a_i$ para todo $i < n$, $p \nmid a_n$ y $p^2 \nmid a_0$. Entonces f es irreducible en $\mathbb{Q}[X]$.

DEMOSTRACION. Un polinomio que satisface las condiciones del enunciado en relación con algún primo p se dirá un polinomio de Eisenstein. Anticipándonos a la demostración, resulta por ejemplo que $2X^4 + 9X^3 - 6X + 15$ es irreducible sobre $\mathbb{Q}$, pues es un polinomio de Eisenstein respecto de $p = 3$. Observemos también que $X^n - 2$ −cuya irreducibilidad sobre $\mathbb{Q}$ demostramos en el ejemplo 11.1.12− es un polinomio de Eisenstein respecto de $p = 2$.

Iniciamos la demostración probando en primer término que f sólo admite factores de grado 0 ó n en $\mathbb{Z}[X]$. En efecto, supongamos que $f = gh$ con g y h de coeficientes enteros b_i y c_i respectivamente. En tal caso, puesto que

$a_0 = b_0c_0$ y $p^2 \nmid a_0$, podemos asumir sin pérdida de generalidad que $p \mid b_0$ y $p \nmid c_0$. Por otro lado $p \nmid g$, ya que $p \nmid gh$, lo que garantiza la existencia de un índice k tal que $p \nmid b_k$ y $p \mid b_i$ para todo $i < k$. Deducimos entonces que

$$a_k = b_k c_0 + \sum_{i=0}^{k-1} b_i c_{k-i}$$

no es divisible por p, lo que implica por hipótesis que $k = n$. Puesto que valen las desigualdades $k \leq gr(g) \leq n$, concluimos que $gr(g) = n$ y $gr(h) = 0$, como queríamos probar.

Veamos finalmente que f no puede descomponerse en $\mathbb{Q}[X]$ como producto de dos polinomios de grado positivo. Supongamos para ello que $f = f_1 f_2$ en $\mathbb{Q}[X]$, y tomemos números naturales m_1 y m_2 tales que $m_i f_i \in \mathbb{Z}[X]$ (podemos por ejemplo elegir como m_i el mínimo común múltiplo de los denominadores de los coeficientes de f_i). Designando por m el producto $m_1 m_2$, obtenemos en $\mathbb{Z}[X]$ la igualdad $mf = (m_1 f_1)(m_2 f_2)$, de la que deducimos, aplicando el lema 11.1.14, la existencia de polinomios g_1 y g_2 con coeficientes enteros tales que $f = g_1 g_2$ y $gr(g_i) = gr(m_i f_i) = gr(f_i)$ para $i = 1, 2$.

De acuerdo con lo demostrado más arriba, sigue entonces que $gr(g_1) = 0$ ó $gr(g_2) = 0$ y por lo tanto alguno de los polinomios f_i es constante, lo que prueba la irreducibilidad de f sobre $\mathbb{Q}$. $\Diamond$

Como aplicación del criterio de Eisenstein, probaremos la irreducibilidad sobre $\mathbb{Q}$ de los polinomios ciclotómicos de orden primo p.

La aplicación no es automática, ya que Φ_p no es un polinomio de Eisenstein, pero ya veremos que mediante un "cambio de variable" adecuado podremos colocarnos en ese caso. Para darle rigor a este lenguaje bastante impreciso, comencemos por ampliar el concepto de especialización.

Así como hemos definido la especialización de un polinomio f en los elementos del anillo A de coeficientes, también tiene pleno sentido definir la especialización de f en un polinomio P, a través de la sustitución $X \mapsto P$. Precisamente, si

$$f = \sum_i a_i X^i$$

es un polinomio con coeficientes en A y $P \in A[X]$, definimos la especialización de f en P en la forma

$$f(P) = \sum_i a_i P^i,$$

que claramente es un polinomio con coeficientes en A. Es inmediato verificar que este tipo de especialización también satisface las propiedades enunciadas en la proposición 9.2.3, ya que la validez de las mismas es consecuencia de propiedades válidas en cualquier anillo conmutativo. De tal modo, si f, g y P son polinomios con coeficientes en A, valen las igualdades

$$(f+g)(P) = f(P) + g(P) \quad \text{y} \quad (fg)(P) = f(P)g(P).$$

Asimismo, el lector podrá probar sin dificultad la fórmula

$$gr\,(f(P)) = gr(f)gr(P),$$

válida para polinomios no nulos con coeficientes en cualquier anillo numérico.

Con estas nuevas herramientas, ya podemos resolver el asunto que dejamos pendiente:

Ejemplo 11.1.16 Si p es primo, el polinomio ciclotómico

$$\Phi_p = X^{p-1} + X^{p-2} + \ldots + X + 1$$

es irreducible sobre $\mathbb{Q}$.

Especializando la igualdad

$$(X-1)\Phi_p = X^p - 1$$

en el polinomio $X+1$, y usando las propiedades de la especialización, tenemos

$$X\Phi_p\,(X+1) = (X+1)^p - 1 = \sum_{i=0}^{p} \binom{p}{i} X^i - 1 = \sum_{i=1}^{p} \binom{p}{i} X^i =$$

$$= X \sum_{i=1}^{p} \binom{p}{i} X^{i-1} = X \sum_{i=0}^{p-1} \binom{p}{i+1} X^i,$$

de donde obtenemos, cancelando el factor X en ambos términos:

$$\Phi_p\,(X+1) = \sum_{i=0}^{p-1} \binom{p}{i+1} X^i = p + \sum_{i=1}^{p-2} \binom{p}{i+1} X^i + X^{p-1}.$$

Teniendo en cuenta que todos los números combinatorios de la última suma son divisible por p, resulta que $\Phi_p\,(X+1)$ es un polinomio de Eisenstein respecto de p y por lo tanto es irreducible sobre $\mathbb{Q}$.

Finalmente, para demostrar que Φ_p también es irreducible, supongamos que $\Phi_p = fg$ con f y g en $\mathbb{Q}[X]$. Sigue entonces especializando que

$$\Phi_p\,(X+1) = f(X+1)g(X+1),$$

lo que implica que alguno de los factores del segundo miembro es constante, por la irreducibilidad de $\Phi_p\,(X+1)$. Puesto que

$$gr\,(h(X+1)) = gr(h)gr(X+1) = gr(h)$$

cualquiera sea el polinomio no nulo h, concluimos que $gr(f) = 0$ ó $gr(g) = 0$, lo que prueba que Φ_p es irreducible en $\mathbb{Q}[X]$.

No está de más señalar que $X^{m-1} + \ldots + X + 1$ no es irreducible sobre $\mathbb{Q}$ si m no es primo. Por ejemplo, para $m = 6$ se tiene la factorización no trivial

$$X^5 + X^4 + X^3 + X^2 + X + 1 = (X^2 + X + 1)(X^3 + 1). \quad \Diamond$$

11.1.2. Ejercicios

1. En cada uno de los siguientes casos determinar todas las formas posibles de descomponer f como producto de dos polinomios mónicos con coeficientes reales:

   a) $f = (X+3)^4(X^2-5)$
   b) $f = (X^2+X+1)(X^4+X^2-2)(X-1)^2$
   c) $f = (X^3 - 1/2\,X^2 - 2X - 1)^3(X+4\imath)(X-4\imath)$.

2. Demostrar que los siguientes polinomios son irreducibles sobre $\mathbb{Q}$:

   a) $X^5 - 6X^4 + 9X^3 - 3X^2 + 30X - 5$
   b) $X^4 + 102X^3 - 68X + 136$
   c) $1/2\,X^4 - 4X^3 + 2X^2 + X - 3$
   d) $X^4 + 1$
   e) $X^n - q$ (q primo)
   f) $X^3 - aX^2 + bX + 1$ (a y b enteros impares) .

3. Hallar el desarrollo de los polinomios ciclotómicos Φ_8, Φ_{20} y Φ_{24} .

4. Factorizar en $\mathbb{Q}[X]$, $\mathbb{R}[X]$ y $\mathbb{C}[X]$ los polinomios X^6+1, X^8-4, X^8+4, Φ_8 y Φ_{24} .

5. Si z es una raíz quinta primitiva de la unidad, probar que $z + z^{-1}$ es raíz de $X^2 + X - 1$.

6. Factorizar Φ_5 en $\mathbb{R}[X]$ y $\mathbb{Q}[X]$.

7. Sea $n \in \mathbb{N}$. Determinar (según sea n par o impar) el número de factores irreducibles mónicos de $X^n + X^{n-1} + \ldots + X + 1$ en $\mathbb{R}[X]$.

8. Sea $f = 2X^5 + 5X^4 + X^3 - 4X^2 - X + 1$. En cada uno de los siguientes casos determinar los polinomios mónicos g con coeficientes racionales que satisfacen la condición planteada:

   a) $g \mid f$
   b) $g \mid f$ y g no tiene raíces racionales
   c) $g \mid f$ y g admite exactamente dos raíces racionales
   d) $g \mid f$ y $g(1) = 1$
   e) $gr(g) = 3$ y $gr\,((f : g)) = 2$.

9. En los siguientes incisos las letras f, g y P designan polinomios con coeficientes complejos.

   a) Si f y g son coprimos probar que $f(P)$ y $g(P)$ son coprimos.

   b) Si $(f : g) = h$ probar que $(f(P) : g(P)) = h(P)$ (suponiendo P mónico).

11.2. Cuerpo de fracciones racionales

11.2.1. Construcción

Debido a su profusa aparición en cursos de Cálculo, el lector está familiarizado con las llamadas funciones racionales, definidas por expresiones de la forma

$$f(x) = \frac{g(x)}{h(x)}\,,$$

donde g y h son polinomios con coeficientes reales. Dejando un poco de lado su significado funcional, daremos un tratamiento abstracto a dichas expresiones fraccionarias, que podemos describir informalmente como "cocientes de polinomios", y mostraremos que el conjunto de las mismas admite una estructura natural de cuerpo, derivada de la estructura algebraica del anillo de polinomios. Como aplicación útil, demostraremos luego que toda fracción se descompone en forma única como suma de fracciones simples, hecho del cual probablemente el lector tenga noticia y que brinda un método de integración de funciones racionales.

Sin duda es necesario darle sentido a la frase que hemos escrito entre comillas en el párrafo anterior, ya que no está claro en absoluto qué clase de objeto es g/h si g y h son polinomios. Como en el caso de la construcción de los números complejos, el asunto requerirá la creación de un conjunto adecuado y la definición en él de ciertas operaciones, que muy probablemente le recordarán al lector las conocidas operaciones entre fracciones numéricas.

Nos proponemos tratar el tema con la mayor generalidad posible, por lo que la letra K designará en lo que sigue un cuerpo arbitrario. Comencemos definiendo en el conjunto

$$\mathcal{F} = \{(f, g) \in K[X] \times K[X] \,:\, g \neq 0\}$$

la relación

$$(f_1, g_1) \sim (f_2, g_2) \Leftrightarrow f_1 g_2 = f_2 g_1\,,$$

que resulta ser una relación de equivalencia.

Dejando a cargo del lector la demostración de la reflexividad y la simetría de la misma, probemos que es transitiva. Supongamos para ello que z_1, z_2 y z_3 son elementos de $\mathcal{F}$ tales que $z_1 \sim z_2$ y $z_2 \sim z_3$. Escribiendo $z_i = (f_i, g_i)$ y usando las propiedades de las operaciones en $K[X]$ resulta entonces que

$$g_2(f_1 g_3 - f_3 g_1) = f_1 g_2 g_3 - f_3 g_2 g_1 = f_2 g_1 g_3 - f_2 g_3 g_1 = 0\,,$$

de donde sigue que $f_1 g_3 - f_3 g_1 = 0$, ya que $g_2 \neq 0$ y $K[X]$ es un dominio de integridad. En consecuencia $f_1 g_3 = f_3 g_1$ y $z_1 \sim z_3$.

Habiendo demostrado que la relación definida es de equivalencia, emplearemos cualquiera de los símbolos

$$\frac{f}{g} \;\text{ ó }\; f/g$$

para designar la clase de equivalencia de un elemento genérico (f, g) de $\mathcal{F}$, mientras que el conjunto cociente $\mathcal{F}/\sim$, cuyos elementos llamaremos *fracciones racionales* con coeficientes en K, será notado $K(X)$.

Resumiendo:

$$K(X) = \{f/g \,:\, f, g \in K[X]\,, g \neq 0\}\,,$$

con la igualdad de fracciones dada por la condición (sin duda familiar para el lector)

$$f/g \,=\, h/l \,\Leftrightarrow\, fl = gh\,.$$

Representacion irreducible.

Una fracción p/q se dice *irreducible* si y solo si p y q son coprimos. Si $z = f/g$ es cualquier fracción y $d = (f : g)$, escribiendo $f = du$ y $g = dv$ resulta que

$$f/g = du/dv = u/v,$$

lo que muestra que u/v es una representación irreducible de z. Multiplicando numerador y denominador por una constante adecuada podemos suponer además que v es mónico, siendo entonces u/v la única representación de este tipo que admite z. En efecto, si $z = u_1/v_1$ con $(u_1 : v_1) = 1$ y v_1 mónico, deducimos de la igualdad $uv_1 = vu_1$ que v y v_1 se dividen mutuamente y por lo tanto son iguales, por ser mónicos. Luego $u_1 = u$ y nuestra afirmación queda probada. La fracción u/v puede elegirse como representante canónico de la clase de z, y se llama la *representación irreducible* de z.

Por ejemplo, sean $f = 6X^2 + 4X - 2$ y $g = 2X^3 + 2X + 4$ y sea $z = f/g$ en $\mathbb{Q}(X)$. Un sencillo cálculo nos muestra que $(f : g) = X + 1$, resultando que $f = 2(X+1)(3X-1)$ y $g = 2(X+1)(X^2 - X + 2)$. Puesto que entonces $u = 3X - 1$ y $v = X^2 - X + 2$ son coprimos, concluimos que

$$\frac{u}{v} = \frac{3X-1}{X^2 - X + 2} = \frac{f}{g}$$

es la representación irreducible de z.

11.2.2. Estructura algebraica

Para obtener la prometida estructura de cuerpo, definimos en $K(X)$ las dos siguientes operaciones, que como de costumbre denominaremos genéricamente suma y producto:

$$\frac{f}{g} + \frac{h}{t} = \frac{ft + gh}{gt} \tag{11.3}$$

$$\frac{f}{g} \,.\, \frac{h}{t} = \frac{fh}{gt}\ . \tag{11.4}$$

Obviamente, las operaciones indicadas en los numeradores y denominadores corresponden a la suma y el producto usuales de polinomios. Observemos que la suma de fracciones de igual denominador adopta una forma particularmente sencilla, ya que

$$\frac{f}{g} + \frac{h}{g} = \frac{fg + gh}{gg} = \frac{g(f+h)}{gg} = \frac{f+h}{g} \, .$$

Antes de estudiar las propiedades de la suma y el producto de fracciones debemos verificar que las mismas están bien definidas, habida cuenta de que la representación de una fracción racional como cociente de polinomios no es única. Supongamos entonces que $f/g = f_1/g_1$ y $h/t = h_1/t_1$, ó equivalentemente, que $fg_1 = gf_1$ y $ht_1 = th_1$. Operando, tenemos:

$$(ft + gh)g_1t_1 = fg_1tt_1 + ht_1gg_1 = gf_1tt_1 + th_1gg_1 = gt(f_1t_1 + g_1h_1)$$

y

$$(fh)(g_1t_1) = fg_1ht_1 = gf_1th_1 = (gt)(f_1h_1),$$

lo que prueba que

$$f/g + h/t = f_1/g_1 + h_1/t_1 \quad \text{y} \quad (f/g)(h/t) = (f_1/g_1)(h_1/t_1) \, ,$$

esto es, la suma y el producto de fracciones no dependen de las representaciones elegidas para éstas. Veamos ahora que se satisfacen todos los axiomas de cuerpo.

Proposición 11.2.1 El conjunto de fracciones racionales con coeficientes en K es un cuerpo respecto a la operaciones definidas en (11.3) y (11.4).

Demostracion. La validez de las propiedades asociativas y conmutativas de la suma y el producto de fracciones, y de la propiedad distributiva del producto respecto a la suma son consecuencia inmediata de las propiedades de la suma y el producto de polinomios, por lo que dejamos los detalles de sus demostraciones a cargo del lector. Admitidos estos hechos, resta probar entonces que ambas operaciones admiten elemento neutro, que toda fracción admite inverso aditivo y que toda fracción no nula (esto es, distinta del elemento neutro de la suma) admite inverso multiplicativo.

Con respecto al primer punto, es inmediato verificar que $0/1$ es el elemento neutro de la suma de fracciones y que $1/1$ es el elemento neutro del producto. Vale la pena remarcar que $0/1$ es la representación irreducible de todas las fracciones de la forma $0/g$, mientras que $1/1$ es la representación irreducible de las fracciones de la forma g/g. Por simplicidad, las notaremos 0 y 1, respectivamente.

Si $z = f/g$ es una fracción cualquiera, sigue inmediatamente que la fracción $(-f)/g$, que notaremos $-z$, es su inverso aditivo, ya que

$$\frac{f}{g} + \frac{(-f)}{g} = \frac{f + (-f)}{g} = \frac{0}{g} = 0 \, ,$$

mientras que g/f es su inverso multiplicativo si $z \neq 0$, ya que f es no nulo en tal caso. Como es de imaginar, la notaremos z^{-1}. $\Diamond$

NOTA La aplicación $\vartheta : K[X] \to K(X)$ definida por $\vartheta(f) = f/1$ es inyectiva, pues

$$f/1 = g/1 \Leftrightarrow g = f,$$

por definición de igualdad de fracciones. Luego, y a similitud del caso complejo, identificando cada polinomio h con la fracción $h/1$ podemos asumir que vale la inclusión $K[X] \subset K(X)$. Más aún, la inmersión anterior es compatible con las correspondientes estructuras algebraicas, ya que si h y t son polinomios valen las igualdades

$$\vartheta(h) + \vartheta(t) = \frac{h}{1} + \frac{t}{1} = \frac{h+t}{1} = \vartheta(h+t)$$

y

$$\vartheta(h)\,.\,\vartheta(t) = (h/1)(t/1) = ht/1 = \vartheta(ht),$$

lo que significa que las operaciones definidas en $K(X)$ coinciden con las operaciones usuales de polinomios cuando se las restringe a $K[X]$. Asimismo, dada una fracción z, representada en la forma f/g, resulta que

$$z \in K[X] \Leftrightarrow \text{ existe } h \in K[X] \text{ tal que } f/g = h/1 \Leftrightarrow f = gh,$$

esto es, una fracción racional es un polinomio si y solo si el numerador es múltiplo del denominador. El lector deberá probar que esta condición es independiente de la representación elegida para z.

Conservando la notación precedente, y teniendo en cuenta que el inverso multiplicativo en $K(X)$ de un polinomio no nulo cualquiera l es $1/l$, notemos por último que podemos escribir z en la forma

$$z = f/g = (f/1)(1/g) = fg^{-1},$$

igualdad que da sentido a la descripción informal de las fracciones racionales como cocientes de polinomios. $\Diamond$

11.2.3. Fracciones simples

> Una fracción racional z se dice una *fracción simple* si y solo si su representación canónica es de la forma $z = u/v^m$, donde v es un polinomio mónico irreducible, $m \in \mathbb{N}$ y $gr(u) < gr(v)$.

Por ejemplo,

$$\frac{1}{(X-3)^2} \quad \text{y} \quad \frac{2X-1}{X^2+X+1}$$

son fracciones simples en $\mathbb{Q}(X)$.

Asimismo, una fracción $z = f/g$ con $gr(f) < gr(g)$ se dice una *fracción propia*. Es fácil ver que la propiedad es independiente de la representación elegida para z.

En lo que resta de la sección, dedicaremos nuestros esfuerzos a probar que toda fracción propia se descompone en forma única como suma de fracciones simples. Por ejemplo, el lector puede constatar que vale en $\mathbb{Q}(X)$ la descomposición

$$\frac{X^2 - X + 3}{X^4 + X^2} = \frac{3}{X^2} - \frac{1}{X} + \frac{X - 2}{X^2 + 1}.$$

Comenzaremos la tarea demostrando dos lemas preparatorios:

Lema 11.2.2 Sea $f \in K[X]$ un polinomio de grado positivo s y sea h un polinomio no nulo con coeficientes en K. Entonces h se expresa unívocamente en la forma

$$h = \sum_{i=0}^{m} h_i f^i, \tag{11.5}$$

donde $h_i = 0$ ó $gr(h_i) < s$ para todo i, $h_m \neq 0$ y $m \in \mathbb{N}_0$. Más precisamente, m es la parte entera de $gr(h)/s$.

DEMOSTRACION. Por razones prácticas, escribiremos abreviadamente las condiciones $t = 0$ ó $gr(t) < s$ en la forma $gr(t) \preccurlyeq s$.

Para probar la existencia de un desarrollo de la forma (11.5), observemos en primer término que el caso $gr(h) < s$ es trivial, ya que tomando $m = 0$ y $h_0 = h$ se satisfacen las condiciones requeridas.

Si $gr(h) \geq s$, dividiendo h por f podemos escribir

$$h = qf + l,$$

con $gr(l) \preccurlyeq s$ y $q \neq 0$.

Puesto que

$$gr(q) = gr(qf) - gr(f) = gr(qf + l) - s = gr(h) - s < gr(h),$$

sigue por un argumento inductivo que q admite un desarrollo de la forma

$$q = \sum_{i=0}^{k} q_i f^i,$$

con $q_k \neq 0$ y $gr(q_i) \preccurlyeq s$ para todo i, verificándose además que $k = [\, gr(q)/s \,]$. Luego:

$$h = qf + l = \left(\sum_{i=0}^{k} q_i f^i \right) f + l = \sum_{i=0}^{k} q_i f^{i+1} + l = \sum_{i=1}^{k+1} q_{i-1} f^i + l,$$

y tomamos entonces $m = k+1$, $h_0 = l$ y $h_i = q_{i-1}$ $(1 \leq i \leq m)$. Completamos la prueba notando que

$$\left[\frac{gr(h)}{s}\right] = \left[\frac{gr(q)+s}{s}\right] = \left[\frac{gr(q)}{s}+1\right] = \left[\frac{gr(q)}{s}\right]+1 = k+1 = m.$$

Con respecto a la unicidad, supongamos que

$$h = \sum_{i=0}^{m} g_i f^i$$

es cualquier otro desarrollo de tipo (11.5) de h (ya demostramos que el número de términos depende sólo de h), y probemos que $g_i = h_i$ para todo i.

La conclusión es obvia si $m = 0$. Supongamos entonces que $m > 0$ y que el resultado vale para desarrollos de longitud menor que m. Puesto que $gr(h_0) \preccurlyeq s$ y $gr(g_0) \preccurlyeq s$, escribiendo

$$h = f\left(\sum_{i=0}^{m-1} h_{i+1} f^i\right) + h_0 = f\left(\sum_{i=0}^{m-1} g_{i+1} f^i\right) + g_0$$

sigue por unicidad de cociente y resto que $g_0 = h_0$ y

$$\sum_{i=0}^{m-1} h_{i+1} f^i = \sum_{i=0}^{m-1} g_{i+1} f^i,$$

de donde resulta por hipótesis inductiva que $g_{i+1} = h_{i+1}$ para todo i, o equivalentemente, $g_i = h_i$ para $i \leq i \leq m$, como queríamos. $\Diamond$

Conservando las notaciones del lema y usando el mismo lenguaje que en $\mathbb{Z}$, diremos que

$$h = \sum_{i=0}^{m} h_i f^i$$

es el desarrollo de h en *base* f.

Por ejemplo, tomando $h = X^5 + 4X^3 - X^2 + 3X + 2$ en $\mathbb{Q}[X]$, resulta que

$$h = X(X^2+1)^2 + (2X-1)(X^2+1) + 3$$

es el desarrollo de h en base X^2+1, como el lector puede verificar.

Lema 11.2.3 Sea $z = f/g$ una fracción propia irreducible donde g es producto de dos polinomios de grado positivo g_1 y g_2 coprimos entre sí. Entonces z puede descomponerse en la forma

$$z = \frac{f_1}{g_1} + \frac{f_2}{g_2}\,,$$

donde f_1/g_1 y f_2/g_2 también son fracciones propias irreducibles.

Más generalmente, supongamos que g es producto de m polinomios g_i de grado positivo y coprimos dos a dos. Entonces z puede expresarse en la forma

$$z = \sum_{i=1}^{m} \frac{f_i}{g_i}\,,$$

donde f_i/g_i es una fracción propia irreducible para todo i.

DEMOSTRACION. Probaremos solo la primera parte del lema, ya que habiendo probado el caso $m = 2$ es fácil obtener la demostración del caso general mediante un argumento inductivo, teniendo en cuenta que g_1 y $g_2 g_3 \dots g_m$ son coprimos (dejamos los detalles a cargo del lector).

Respecto a la primera afirmación, por hipótesis podemos expresar f como combinación lineal de g_1 y g_2, digamos

$$f = ug_1 + vg_2 \quad (*),$$

donde u y v son polinomios no nulos. Dividiendo v por g_1 y escribiendo

$$v = tg_1 + r \ (gr(r) < gr(g_1))\,,$$

resulta reemplazando que

$$f = ug_1 + (tg_1 + r)g_2 = (u + tg_2)g_1 + rg_2,$$

por lo que podemos suponer sin pérdida de generalidad que $gr(v) < gr(g_1)$ (observemos que $r \neq 0$, ya que $g_1 \nmid f$).

Deducimos a la vez que en tal caso $gr(u) < gr(g_2)$, ya que

$$gr(u) = gr(ug_1) - gr(g_1) = gr(f - vg_2) - gr(g_1) < gr(g) - gr(g_1) = gr(g_2)\,,$$

por ser $gr(f) < gr(g)$ y $gr(vg_2) < gr(g_1) + gr(g_2) = gr(g)$.

Operando, obtenemos

$$z = \frac{f}{g} = \frac{ug_1 + vg_2}{g_1 g_2} = \frac{vg_2}{g_1 g_2} + \frac{ug_1}{g_1 g_2} = \frac{v}{g_1} + \frac{u}{g_2}\,,$$

y el resultado sigue tomando $f_1 = v$ y $f_2 = u$. Advierta el lector que efectivamente v y g_1 son coprimos, ya que por hipótesis $(f : g_1) = 1$ y cualquier factor común de v y g_1 es un factor de f, por $(*)$. Obviamente el mismo argumento se aplica a u y g_2. $\Diamond$

Descomposicion en fracciones simples.

Ya estamos en condiciones de probar el resultado mencionado en el inicio de este apartado.

Teorema 11.2.4 Toda fracción racional propia $z = f/g$ se descompone como suma de fracciones simples. Precisamente, existe una familia $(g_i)_{1\leq i\leq m}$ de polinomios irreducibles mónicos distintos y una familia $(a_i)_{1\leq i\leq m}$ de números naturales tales que

$$z = \sum_{i=1}^{m}\sum_{j=1}^{a_i} \frac{f_{ij}}{g_i^j} ,$$

con $f_{ia_i} \neq 0$ y $gr(f_{ij}) \preccurlyeq gr(g_i)$ para todo par de índices (i, j). Además, los parámetros m, g_i, a_i y f_{ij} están unívocamente determinados por z.

Demostracion. Suponiendo que f/g es la representación irreducible de z, designemos por $g = \prod_{i=1}^{m} g_i^{a_i}$ la factorización canónica de g en $K[X]$ como producto de irreducibles mónicos. Puesto que $(g_i^{a_i} : g_j^{a_j}) = 1$ si $i \neq j$, sigue por el lema 12.2.3 que existen polinomios no nulos $f_1, f_2, \ldots, f_m$ tales que

$$z = \frac{f_1}{g_1^{a_1}} + \frac{f_2}{g_2^{a_2}} + \cdots + \frac{f_m}{g_m^{a_m}} , \tag{11.6}$$

con $gr(f_i) < gr(g_i^{a_i})$ y $(f_i : g_i) = 1$ para todo i.

Fijemos ahora cualquier índice $i\,(1 \leq i \leq m)$ y sea $s_i = gr(g_i)$. Puesto que

$$\frac{gr(f_i)}{s_i} < \frac{gr(g_i^{a_i})}{s_i} = a_i ,$$

sigue del lema 11.2.3 que el desarrollo de f_i en base g_i tiene a lo sumo a_i términos, y por lo tanto podemos escribirlo en la forma

$$f_i = \sum_{j=0}^{a_i-1} t_j g_i^j,$$

con $gr(t_j) \preccurlyeq s_i$ para todo j. Operando en $K(X)$ sigue entonces que

$$\frac{f_i}{g_i^{a_i}} = \sum_{j=0}^{a_i-1} \frac{t_j\, g_i^j}{g_i^{a_i}} = \sum_{j=0}^{a_i-1} \frac{t_j}{g_i^{a_i-j}} = \sum_{j=1}^{a_i} \frac{f_{ij}}{g_i^j} ,$$

donde $f_{ij} = t_{a_i-j}$ (observemos que $f_{ia_i} = t_0 \neq 0$, ya que $g_i \nmid f_i$).

Finalmente, reemplazando en (11.6) concluimos que

$$z = \sum_{i=1}^{m} \frac{f_i}{g_i^{a_i}} = \sum_{i=1}^{m}\sum_{j=1}^{a_i} \frac{f_{ij}}{g_i^j}$$

es una descomposición de z como suma de fracciones simples de la forma deseada.

Para probar la unicidad, supongamos que z admite otra descomposición como la del enunciado, digamos

$$\frac{f}{g} = \sum_{i=1}^{n} \sum_{j=1}^{b_i} \frac{h_{ij}}{q_i^j} ,$$

con las mismas hipótesis sobre los q_i, b_i y h_{ij} . Operando, y designando por Q_i el producto $\prod_{j\neq i} q_j^{b_j}$, resulta que

$$\frac{f}{g} = \sum_{i=1}^{n} \frac{\sum_{j=1}^{b_i} h_{ij}\, q_i^{b_i-j}}{q_i^{b_i}} = \sum_{i=1}^{n} \frac{h_i}{q_i^{b_i}} = \frac{\sum_{i=1}^{n} h_i Q_i}{\prod_{i=1}^{n} q_i^{b_i}} ,$$

siendo esta última fracción también irreducible. Para probar esta afirmación, mostraremos que ninguno de los factores irreducibles q_i del denominador divide al numerador. En efecto, tomando congruencias modulo q_i obtenemos

$$\sum_{k=1}^{n} h_k Q_k \equiv h_i Q_i = \left(\sum_{j=1}^{b_i} h_{ij} q_i^{b_i-j} \right) Q_i \equiv h_{ib_i} Q_i \not\equiv 0 ,$$

pues $q_i \nmid h_{ib_i}$ y $q_i \nmid Q_i$. Sigue luego por la unicidad de la representación irreducible de una fracción (el denominador es mónico) que

$$f = \sum_{i=1}^{n} h_i Q_i \tag{11.7}$$

y

$$g = \prod_{i=1}^{n} q_i^{b_i} , \tag{11.8}$$

deduciéndose de (11.8) que $n = m$, $q_i = g_i$ y $b_i = a_i$ para todo i, por el teorema de factorización única en $K[X]$. Usando (11.7) y (11.8), resulta en particular que

$$\sum_{j=1}^{m} h_j Q_j = f = \sum_{j=1}^{m} f_j \left(\prod_{k\neq j} g_k^{a_k} \right) = \sum_{j=1}^{m} f_j Q_j ,$$

de donde sigue, tomando congruencias, que $h_i Q_i \equiv f_i Q_i \,(g_i^{a_i})$ para todo i. Cancelando el factor Q_i (coprimo con $g_i^{a_i}$) obtenemos $h_i \equiv f_i \,(g_i^{a_i})$, lo que nos permite concluir que $h_i = f_i$, ya que el grado de ambos polinomios es menor que el grado de $g_i^{a_i}$.

Finalmente, la unicidad del desarrollo en base g_i nos indica que $h_{ij} = f_{ij}$ para todo par de índices (i, j), como queríamos demostrar. $\Diamond$

Corolario 11.2.5 Toda fracción racional z se expresa unívocamente en la forma

$$z = h + \sum_{i=1}^{n} z_i \,, \tag{11.9}$$

donde h es un polinomio y z_i es una fracción simple para todo i $(n \in \mathbb{N}_0)$.

DEMOSTRACION. Considerando la expresión irreducible f/g de z y designando por h y r el cociente y el resto de la división de f por g en K[X], resulta

$$z = \frac{f}{g} = \frac{hg + r}{g} = h + \frac{r}{g} \,.$$

Obtenemos entonces una descomposición de tipo (11.9) para z aplicando el teorema 11.2.4 a la fracción racional propia r/g (observemos que $n = 0$ si $r = 0$). Dejamos como ejercicio para el lector la tarea de probar la unicidad de la misma. $\Diamond$

EL CASO REAL Como comentamos al iniciar esta sección, la descomposición en fracciones simples de una fracción racional con coeficientes reales brinda un método de integración de la función racional asociada a la misma. Teniendo en cuenta la caracterización de los elementos irreducibles de $\mathbb{R}[X]$ establecida en la proposición 11.1.9, el teorema 11.2.4 adopta en el caso real la siguiente forma:

Proposición 11.2.6 Toda fracción racional propia z con coeficientes reales se descompone unívocamente en $\mathbb{R}(X)$ en la forma

$$z = \sum_{i=1}^{r} \sum_{j=1}^{a_i} \frac{u_{ij}}{(X - \theta_i)^j} + \sum_{i=1}^{s} \sum_{j=1}^{b_i} \frac{v_{ij}X + w_{ij}}{(X^2 + \alpha_i X + \beta_i)^j} \,, \tag{11.10}$$

donde los coeficientes designan números reales y $\alpha_i^2 - 4\beta_i < 0$ para todo i (como en 11.2.4 suponemos $u_{ia_i} \neq 0$ y $v_{ib_i}X + w_{ib_i} \neq 0$ cualquiera sea el índice i). $\Diamond$

Desarrollemos un par de ejemplos. Tengamos en cuenta para ello que, según la demostración del teorema 11.2.4, los denominadores de las fracciones simples que aparecen en la descomposición de una fracción racional f/g son potencias de los factores irreducibles de g. Por lo tanto, para hallar efectivamente dicha descomposición es preciso conocer la factorización de g.

Ejemplo 11.2.7 Descompongamos en $\mathbb{R}(X)$ la fracción

$$z = \frac{f}{g} = \frac{4X^3 + 8X^2 + 4X + 2}{X^4 + 4X^3 + 5X^2 + 2X} \,.$$

Aplicando el criterio de Gauss resulta $g = X(X+2)(X+1)^2$. Luego, z se descompondrá por proposición 11.2.6 en la forma

$$z = \frac{a}{X} + \frac{b}{X+2} + \frac{c}{X+1} + \frac{d}{(X+1)^2} \ ,$$

donde a, b, c y d son números reales. Operando, obtenemos entonces

$$\frac{f}{g} = \frac{a(X+2)(X+1)^2 + bX(X+1)^2 + cX(X+2)(X+1) + dX(X+2)}{g} \ ,$$

o equivalentemente,

$$a(X+2)(X+1)^2 + bX(X+1)^2 + cX(X+2)(X+1) + dX(X+2) = f \, .$$

Para determinar los coeficientes desconocidos comenzamos especializando ambos miembros de la última igualdad en 0, -2 y -1, lo que brinda las relaciones

$$\begin{aligned} 2a &= f(0) = 2 \\ -2b &= f(-2) = -6 \\ -d &= f(-1) = 2 \end{aligned}$$

de donde sigue que $a = 1$, $b = 3$ y $d = -2$. Evaluando luego en 1 resulta que

$$18 = f(1) = 12a + 4b + 6c + 3d = 18 + 6c \, ,$$

y por lo tanto $c = 0$.

En conclusión,

$$z = \frac{1}{X} + \frac{3}{X+2} - \frac{2}{(X+1)^2}$$

es la descomposición única de z como suma de fracciones simples. $\Diamond$

Ejemplo 11.2.8 Sea ahora

$$z = \frac{f}{g} = \frac{2X^5 + 3X^4 - 2X^3 + 5X^2 - 16X + 16}{X^4 + 2X^3 - 2X^2 + 2X - 3} \ .$$

Puesto que z no es una fracción propia aplicamos previamente el algoritmo de la división en $\mathbb{R}[X]$, resultando que $r = 4X^3 - X^2 - 8X + 13$ y $h = 2X - 1$ y son el resto y el cociente de dividir f por g, respectivamente. Luego

$$z = 2X - 1 + \frac{4X^3 - X^2 - 8X + 13}{X^4 + 2X^3 - 2X^2 + 2X - 3} \ ,$$

y debemos descomponer en fracciones simples esta última fracción. Puesto que se ve fácilmente que $g = (X-1)(X+3)(X^2+1)$, tal descomposición será de la forma

$$\frac{r}{g} = \frac{a}{X-1} + \frac{b}{X+3} + \frac{cX+d}{X^2+1} \ ,$$

que podemos determinar resolviendo la ecuación polinómica

$$a(X+3)(X^2+1)+b(X-1)(X^2+1)+(cX+d)(X-1)(X+3)=r\,.$$

Procediendo como en 11.2.7 especializamos primero la igualdad anterior en 1 y en -3, de donde resulta $8a=r(1)=8$ y $-40b=r(-3)=-80$, esto es, $a=1$ y $b=2$.

Evaluando luego en $\imath$, se obtiene la condición

$$(d+c\imath)(-4+2\imath)=r(\imath)=14-12\imath\,,$$

equivalente al sistema lineal de ecuaciones

$$\begin{aligned} c+2d&=-7\\ 2c-d&=6\,, \end{aligned}$$

cuya única solución es $(1,-4)$, como el lector puede verificar inmediatamente.

Por lo tanto, la descomposición buscada es

$$z=2X-1+\frac{1}{X-1}+\frac{2}{X+3}+\frac{X-4}{X^2+1}\,. \quad \Diamond$$

11.2.4. Ejercicios

1. Dados los polinomios $f=X^3+2X^2-X-1$, $g=X^8-X^4+2$, $h=X^3+1$ y $t=2X^4+\imath X^3-X^2-5$, efectuar en $\mathbb{C}(X)$ las siguientes operaciones:

   a) $f/g-ht/g$
   b) $f/h-g/(th)$
   c) $(X/t-g/f)^3$
   d) $\left(g+t^{-2}\right)h^{-1}$.

2. Hallar la representación irreducible de $z=f/g$ en los siguientes casos:

   a) $f=2X^3-6X-4\;;\; g=4X^3-4X^2+4X+12$
   b) $f=(X^3-6X^2+11X-6)(X^2-2)\;;\; g=(X^3+X^2-6X)(2X^2-4)$
   c) $f=X^5-5X^4+4X^3-20X^2+3X-15\;;\; g=3X^3-15X^2+9X-45$.

3. Sean $z_1,z_2,\ldots,z_n$ números complejos distintos y sea f/g la representación irreducible de la fracción

$$z=\sum_{i=1}^{n}\frac{1}{X-z_i}\,.$$

   Determinar los grados de f y g.

4. Si K es un cuerpo numérico y $z = f/g$ es una fracción racional no nula con coeficientes en K, definimos el *grado* de z en la forma

$$Gr(z) = gr(f) - gr(g)\,.$$

Demostrar las siguientes propiedades del grado en $K(X)$ (z y w denotan fracciones racionales no nulas):

a) $Gr(z)$ no depende de la particular representación elegida para z.

b) Si $f \in K[X]$ entonces $Gr(f) = gr(f)$.

c) $z + w = 0$ ó $Gr(z + w) \leq \max\{Gr(z), Gr(w)\}$.

d) $Gr(zw) = Gr(z) + Gr(w)$.

5. Descomponer la fracción $z = f/g$ como suma de fracciones simples en cada uno de los siguientes casos:

a) $f = 4X^5 + X^4 + 3X^3 + 5X^2 - 2X - 3$; $g = X^6 + X^5 - X^4 - X^3$

b) $f = X^{10}$; $g = X^6 + X^5 - 5X^4 - 13X^3 - 19X^2 - 13X - 6$

c) $f = X + 1$; $g = 8X^4 - 4X^3 - 6X^2 + 5X - 1$

d) $f = 3X^4 - X^3 + 4X^2 - 3X + 5$; $g = X^5 - X^4 + 2X^3 - 2X^2 + X - 1$.

6. Demostrar que la descomposición dada por el corolario 11.2.5 es única.

Capítulo 12

Algebra lineal

12.1. Espacios vectoriales

12.1.1. Introducción

Dedicaremos los próximos dos capítulos a establecer los lineamientos básicos del *Algebra lineal*, comenzando con la presentación de su estructura fundamental, la de espacio vectorial.

La materia que nos ocupa se ha convertido desde hace mucho tiempo en una poderosa herramienta de la Matemática, y su vasta utilización en diversas áreas de la misma, como el análisis, la geometría, la teoría de números, etc., y también en otras disciplinas como la Física, la Ingeniería y la Economía, por citar sólo algunas, tornan imprescindible su estudio y manejo fluído.

Con diversos grados de profundidad y extensión, el álgebra lineal se encuentra presente en la *currícula* de cualquier carrera científica, y por eso hemos decidido incluir algo de ella en estas páginas dedicadas a exhibir un panorama general del álgebra. Por razones de brevedad sólo mostraremos aquí algunos aspectos centrales de la teoría, como son por ejemplo las nociones de independencia lineal de vectores y de dimensión de un espacio vectorial, las relaciones algebraicas entre distintos espacios, el cálculo matricial y la resolución de sistemas de ecuaciones lineales. El lector interesado en profundizar estas y otras cuestiones deberá consultar cualquiera de los muchos buenos textos dedicados al tema, algunos de los cuales citamos en la bibliografía ofrecida al final del libro.

Vectores

En sus orígenes históricos los vectores eran flechas, es decir, segmentos orientados en el plano o en el espacio que en Física se usaron, y se usan todavía, para representar fuerzas, velocidades, aceleraciones, etc.. A este tipo de magnitudes se las llama vectoriales, en oposición a las magnitudes escalares que se expresan simplemente mediante un número.

Gráficamente, los vectores quedan caracterizados por su longitud, dirección y sentido. Dos vectores que coincidan en estas tres características se dicen equivalentes, lo cual significa que es posible trasladar uno cualquiera de ellos, sin aplicarle rotaciones ni simetrías, hasta superponerlo con el otro. Físicamente, dos vectores equivalentes representan la misma fuerza (la misma velocidad, la misma aceleración, etc).

En ese contexto gráfico, se aplican a los vectores dos operaciones fundamentales: la suma y el producto por escalares. La suma se obtiene mediante la regla del paralelogramo:

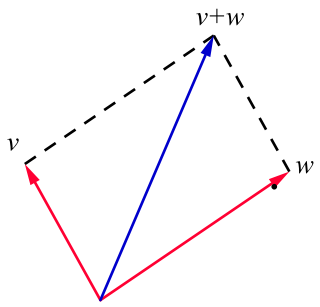


Figura 12.1: Regla del paralelogramo

En general, para sumar dos vectores v y w se toman vectores v' y w' con origen común, respectivamente equivalentes a ellos, y se aplica a éstos la regla del paralelogramo.

El producto de un escalar (número) λ por un vector v prolonga o acorta éste a lo largo de la recta que lo contiene, según sea $|\lambda| > 1$ o $|\lambda| < 1$, cambiando su sentido si λ es negativo.

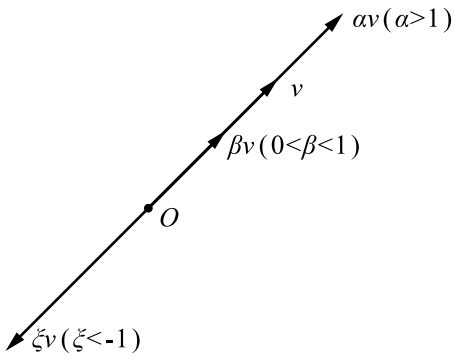


Figura 12.2: Producto por escalares

El medio ambiente en que se sitúan las flechas que mencionamos es $\mathbb{R}^2$ ó $\mathbb{R}^3$, según pensemos los vectores en el plano o en el espacio. En ambos

casos los escalares son números reales, y en ambos casos los vectores se representan gráficamente en el contexto de un sistema de coordenadas cartesianas. Ahora bien, al describir un vector en dicho contexto, debiéramos en principio señalar tanto las coordenadas de su extremo inicial como de su extremo final. Sin embargo, resulta más conveniente desde el punto de vista algebraico fijar al origen de coordenadas como extremo inicial común de todos los vectores. Dado que todo vector v tiene un único equivalente v' con extremo inicial en el origen de coordenadas, no hay ninguna pérdida esencial al adoptar esta norma. Una vez fijada dicha convención, para identificar cada vector sin ambigüedad basta con indicar cuál es su extremo final. Así, el par $(a, b) \in \mathbb{R}^2$ indica también el vector que tiene extremo final en (a, b) y extremo inicial en $(0, 0)$. De este modo, el par (a, b) representa dos objetos gráficos diferentes: un punto o una flecha, no habiendo riesgo de confusión, ya que el contexto nos dirá en cada caso cuál es la interpretación adecuada.

Podemos brindar además una versión algebraica de las dos operaciones que describimos gráficamente arriba. En efecto, consideraciones geométricas elementales nos muestran que si $v = (a, b)$ y $v' = (a', b')$ son vectores de $\mathbb{R}^2$ y λ es un escalar, entonces las operaciones vienen dadas por las fórmulas

$$v + v' = (a + a', b + b') \qquad \text{y} \qquad \lambda v = (\lambda a, \lambda b)\,.$$

Similarmente, en $\mathbb{R}^3$ valen las fórmulas

$$(a, b, c) + (a', b', c') = (a + a', b + b', c + c') \quad \text{y} \quad \lambda(a, b, c) = (\lambda a, \lambda b, \lambda c).$$

Analizando las propiedades que poseen estas operaciones definidas en $\mathbb{R}^2$ y $\mathbb{R}^3$, sencillas de interpretar geométricamente, se arriba a la noción abstracta de espacio vectorial. En dicha estructura interviene un conjunto V, cuyos elementos llamamos vectores, y un cuerpo K cuyos elementos llamamos escalares. En V está definida una operación de suma $(+)$, respecto de la cual V es un grupo abeliano, y existe además una acción de K sobre V, que generaliza el producto de un escalar por un vector que mencionamos al principio. En general, dado un cuerpo K y un conjunto X, cualquier aplicación

$$\vartheta : K \times X \to X$$

se dirá una *acción* de K sobre X. Esto es, una acción asigna a cada $a \in K$ y cada $x \in X$ un elemento $\vartheta((a, x))$ de X, que para agilizar la notación se designa más simplemente por ax. Hecha esta aclaración, precisemos la definición de espacio vectorial:

> Sea V un grupo abeliano y sea K un cuerpo actuando sobre V. Diremos que V es un *espacio vectorial* sobre K respecto a dicha acción si y solo si se satisfacen los siguientes axiomas (las letras griegas indican elementos de K y v y w elementos a V):

(EV_1) $\lambda(v + w) = \lambda v + \lambda w$

(EV_2) $(\alpha+\beta)v = \alpha v + \beta v$
(EV_3) $(\alpha\beta)v = \alpha(\beta v)$
(EV_4) $1\,v = v\,.$

NOTA Observemos que el símbolo + se emplea tanto para designar la suma en el grupo abeliano V como la suma en el cuerpo K, mientras que 0 denotará tanto el elemento neutro de la suma en V como el elemento neutro de la suma en K. No existe de todos modos riesgo de confusión, ya que por el contexto quedará claro el rol que cumple cada uno de los símbolos. Por ejemplo, el lector deberá ser capaz de reconocer dichos roles en la expresión $0\,v + (\lambda+\beta)w + \gamma\,0$.

Digamos también que el carácter axiomático de la definición permite reconocer estructura de espacio vectorial en conjuntos cuyos elementos son de muy diversos tipos, como secuencias finitas, matrices, polinomios, funciones, etc. Si bien siempre brindaremos ejemplos ilustrativos, la idea es que el lector se acostumbre a trabajar formalmente en los espacios vectoriales, olvidando un poco la naturaleza de sus elementos y concentrándose en el manejo de los axiomas y propiedades que iremos demostrando.

En lo que sigue K denotará un cuerpo arbitrario y V un espacio vectorial sobre K (o K-espacio vectorial, como también se lo llama). $\Diamond$

Antes de pasar a ejemplos probaremos algunas propiedades básicas de un espacio vectorial, ilustrando de paso el uso de los axiomas:

Lema 12.1.1 Las siguientes propiedades son válidas en todo espacio vectorial ($\lambda \in K$ y $v \in V$):

1) $\lambda v = 0 \Leftrightarrow \lambda = 0$ ó $v = 0$

2) $-(\lambda v) = (-\lambda)v = \lambda(-v)$

3) $(-1)v = -v\,.$

DEMOSTRACION En el ítem 1), supongamos primero que $\lambda = 0$ y sea $w = 0v$. Tenemos entonces:

$$w = 0v = (0+0)v = 0v + 0v = w + w,$$

de donde

$$0 = w + (-w) = w + w + (-w) = w + 0 = w.$$

En forma completamente análoga se prueba que $\lambda 0 = 0$. Recíprocamente, sea $\lambda v = 0$ y supongamos que $\lambda \neq 0$. Entonces, usando las reglas obtenemos:

$$v = 1v = (\lambda^{-1}\lambda)v = \lambda^{-1}(\lambda v) = \lambda^{-1}0 = 0\,.$$

Las igualdades 2) siguen de 1), ya que

$$(-\lambda)v + \lambda v = (-\lambda+\lambda)v = 0v = 0 = \lambda 0 = \lambda(-v+v) = \lambda(-v) + \lambda v\,.$$

Por último, 3) es un caso particular de 2). $\Diamond$

Ejemplos 12.1.2 En los siguientes ejemplos recorreremos algunas de las principales estructuras de espacio vectorial sobre un cuerpo K. El lector se encargará de verificar en cada caso que el conjunto V es un espacio vectorial respecto a las operaciones definidas (m y n designan números naturales):

1) Como ejemplo muy sencillo comencemos por el espacio vectorial nulo, también llamado trivial, cuyo único elemento es el 0.

2) Generalizando los casos de $\mathbb{R}^2$ y $\mathbb{R}^3$ tenemos el espacio $V = K^n$ de n-uplas de elementos de K, con la estructura aditiva dada por

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

y la acción de K sobre V definida por

$$\lambda(x_1, x_2, \ldots, x_n) = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n)\,.$$

Por razones que resultarán claras más adelante, es éste un ejemplo fundamental de espacio vectorial, que inspira la definición misma de la estructura.

3) Cualquier arreglo de mn elementos de K dispuestos en m hileras horizontales de n elementos cada una se dice una *matriz* de m filas por n columnas con coeficientes en K. Si X es un tal objeto, emplearemos la notación $X = (x_{ij})$ para indicar que x_{ij} es el elemento de K ubicado en la fila i y en la columna j de X (ocasionalmente, designaremos directamente por X_{ij} dicho elemento). Si $V = K^{m\times n}$ designa el conjunto de tales matrices, resulta que V es un K-espacio vectorial con las operaciones de suma y producto por un escalar dadas por

$$\begin{aligned}(x_{ij}) + (y_{ij}) &= (x_{ij} + y_{ij})\,,\\ \lambda(x_{ij}) &= (\lambda x_{ij})\,.\end{aligned}$$

Observemos la fuerte analogía existente entre este ejemplo y el anterior. Para cada índice doble (i, j) el elemento situado en la posición (i, j) de la suma se obtiene sumando en K los correspondientes elementos situados en la posición (i, j), mientras que el producto de un escalar λ por una matriz A es la matriz obtenida multiplicando por λ cada elemento de A.

4) El conjunto $V = K^I$ de funciones de un conjunto no vacío I en K, con las operaciones de suma de funciones y producto de un escalar por una función dadas por

$$\begin{aligned}(f + g)(x) &= f(x) + g(x)\,,\\ (\lambda f)(x) &= \lambda f(x)\,,\end{aligned}$$

donde x denota un elemento genérico de I.

Vale la pena hacer notar que este ejemplo engloba los dos primeros, ya que tanto K^n como $K^{m\times n}$ pueden identificarse con espacios de funciones, a saber, $K^{\mathbb{I}_n}$ y $K^{\mathbb{I}_m\times\mathbb{I}_n}$. En efecto, cada $n-$upla $(x_1, x_2, \ldots, x_n)$ está asociada a una función $g : \mathbb{I}_n \to K$, definida por $g(i) = x_i$, mientras que cada matriz (z_{ij}) corresponde a una función $h : \mathbb{I}_m \times \mathbb{I}_n \to K$, dada por $h((i, j)) = z_{ij}$. Es trivial demostrar que estas correspondencias son biunívocas, y a que a través de dichas identificaciones las operaciones de suma y producto por un escalar definidas en K^n y en $K^{m\times n}$, respectivamente, corresponden a las operaciones definidas arriba en los espacios de funciones.

5) $V = K[X]$. La operación de grupo es la suma usual de polinomios y el producto de un escalar λ por un polinomio f es el polinomio que resulta de multiplicar cada coeficiente de f por λ.

   Asimismo, si fijamos $r \in \mathbb{N}_0$ y consideramos el conjunto formado por el polinomio nulo y todos los polinomios de grado menor o igual que r, vemos fácilmente que éste también es un espacio vectorial respecto a las mismas operaciones. Lo notaremos $K_r[X]$.

6) Si q es un número primo, un caso peculiar del ítem 2) es el espacio de n-uplas $\mathbb{Z}_q^n$, sobre el cuerpo $\mathbb{Z}_q$ de clases de restos módulo q, que tiene una cantidad finita de vectores. En particular, podemos desarrollar en $\mathbb{Z}_q^2$ una geometría plana en la que sólo existe una cantidad finita de puntos y rectas.

   Más particularmente aún, podemos señalar una aplicación interesante del espacio $\mathbb{Z}_2^n$. Todo mensaje, no importa en qué idioma esté escrito, puede representarse a través de un código adecuado como una sucesión finita de ceros y unos. Cada uno de estos dígitos representa un bit de información, y si el mensaje tiene en total n bits, podemos tratarlo como un elemento de $\mathbb{Z}_2^n$.

   Supongamos que enviamos a un receptor distante un mensaje de 5 bits (para no complicar la escritura del ejemplo hemos tomado una cantidad reducida de bits). Digamos que el mensaje es $(1, 1, 1, 0, 0)$, pero que por errores en la transmisión el receptor recibe el mensaje $(1, 1, 0, 0, 0)$. Dado que la resta (o la suma, que módulo 2 es lo mismo) de ambos vectores es $(0, 0, 1, 0, 0)$, concluimos que los errores de transmisión han causado la alteración de un bit de información en el mensaje. Aunque este ejemplo es en realidad muy limitado, es suficiente para sugerirnos que el espacio $\mathbb{Z}_2^n$ puede ser una herramienta teórica muy útil para el análisis de errores en la transmisión de mensajes.

7) Todo cuerpo K es un espacio vectorial sobre sí mismo, donde la acción está dada por el producto en el cuerpo. Siguiendo esta idea, resulta que $\mathbb{R}$ es también un $\mathbb{Q}$-espacio vectorial, en el cual los vectores son los

números reales mientras que los escalares son los números racionales. Análogamente, $\mathbb{C}$ es un $\mathbb{R}$-espacio vectorial y también un $\mathbb{Q}$-espacio vectorial. Claramente, estos ejemplos son casos particulares de la siguiente situación general: si F es un subcuerpo de K entonces K es un espacio vectorial sobre F. $\Diamond$

Combinaciones lineales

Sea $\mathcal{F} = \{v_1, v_2, \ldots, v_n\}$ una familia de vectores de un espacio vectorial V. Diremos que un vector $v \in V$ es *combinación lineal* de la familia $\mathcal{F}$ si y solo si existen escalares $\lambda_1, \lambda_2, \ldots, \lambda_n$ tales que

$$v = \sum_{i=1}^{n} \lambda_i v_i \,.$$

Es evidente que la definición anterior es independiente del ordenamiento dado a los elementos de la familia $\mathcal{F}$, por lo que a menudo diremos simplemente que v es combinación lineal de los vectores $v_1, v_2, \ldots, v_n$.

Ejemplos 12.1.3 Los siguientes ejemplos y observaciones ilustrarán la definición:

1) Veamos si $(2, 4, 0)$ es combinación lineal en $\mathbb{R}^3$ de $v_1 = (1, 1, 2)$ y $v_2 = (1/2, -1/2, 3)$.

   Por definición, se trata de determinar si existen escalares α y β tales que

   $$(2, 4, 0) = \alpha(1, 1, 2) + \beta(1/2, -1/2, 3).$$

   Operando, tenemos que

   $$(2, 4, 0) = (\alpha, \alpha, 2\alpha) + (\beta/2, -\beta/2, 3\beta) = (\alpha + \beta/2, \alpha - \beta/2, 2\alpha + 3\beta),$$

   y puesto que dos ternas son iguales si y sólo si sus respectivas componentes son iguales, resulta que los escalares deben satisfacer el sistema de 3 ecuaciones lineales con dos incógnitas

   $$\begin{aligned} \alpha + \beta/2 &= 2 \\ \alpha - \beta/2 &= 4 \\ 2\alpha + 3\beta &= 0 \,, \end{aligned}$$

   es decir, $(2, 4, 0)$ es combinación lineal de v_1 y v_2 si y sólo si el sistema es resoluble. Puesto que en este caso lo es ($\alpha = 3$ y $\beta = -2$ es una solución), concluimos que la respuesta al problema es afirmativa. Precisamente:

   $$(2, 4, 0) = 3(1, 1, 2) + (-2)(1/2, -1/2, 3) \,.$$

Como detalle adicional, el lector podrá verificar que $(3, -2)$ es la única solución del sistema, vale decir, $(2, 4, 0)$ se escribe de manera única como combinación lineal de v_1 y v_2. También, dejamos como ejercicio probar que $(1, 5, -2)$ *no es* combinación lineal de v_1 y v_2.

2) En el $\mathbb{Q}$-espacio vectorial $\mathbb{Q}[X]$ el polinomio (vector) $5 - X + 6X^2$ es combinación lineal de los polinomios $1 + 2X^2$ y $2 - X$. Encargamos al lector la tarea de determinar los correspondientes escalares.

3) La escritura de un vector como combinación lineal de una familia dada de vectores no necesariamente es única. Por ejemplo en $\mathbb{R}^2$,

$$(-2, 2) = 3(1, -1) + 5(-1, 1) = (-6)(1, -1) + (-4)(-1, 1).$$

4) Si v es un vector no nulo de un espacio vectorial V, el conjunto

$$L = \{\lambda v : \lambda \in K\}$$

de todas las combinaciones lineales de v corresponde gráficamente en los casos $V = \mathbb{R}^2$ ó $V = \mathbb{R}^3$ a la recta que pasa por el origen y por el punto v. Por analogía, se la llama en general la *recta generada* por v.

5) En $\mathbb{R}^2$, designando por e_1 al vector $(1, 0)$ y por e_2 al vector $(0, 1)$, es claro que la recta generada por e_1 es el eje x y la recta generada por e_2 es el eje y. A partir de este hecho se ve fácilmente que todo vector de $\mathbb{R}^2$ se expresa de manera única como combinación lineal de e_1 y e_2, y que los escalares de la combinación lineal correspondiente son las propias componentes del vector, es decir, $(a, b) = a(1, 0) + b(0, 1)$.

El ejemplo que acabamos de describir se generaliza naturalmente a $\mathbb{R}^n$ cualquiera sea el número natural n, y más generalmente aún al espacio K^n de n-uplas con coeficientes en un cuerpo K arbitrario. En efecto, llamando e_i al elemento de $\mathbb{K}^n$ cuya componente i-ésima es 1 y sus demás componentes son nulas, se verifica fácilmente que todo elemento de K^n es combinación lineal de la familia $e_1, e_2, \ldots, e_n$ y que los escalares de la combinación lineal son justamente las componentes del vector. O sea:

$$(a_1, a_2, \ldots, a_n) = \sum_{i=1}^{n} a_i e_i \,.$$

6) Es claro que en todo espacio vectorial V el vector nulo es combinación lineal de cualquier familia $v_1, \ldots, v_n$ de vectores, ya que basta tomar $\lambda_i = 0$ para todo i. Sin embargo, en algunos casos es posible obtener al vector nulo como combinación lineal de $v_1, \ldots, v_n$ sin que necesariamente todos los escalares sean nulos. Por ejemplo,

$$(0, 0, 0) = 2(1, 1, 2) + (-3)(2, 0, -4) + 2(2, -1, -8)$$

en $\mathbb{R}^3$. $\Diamond$

Motivados por el último ejemplo, diremos en general que la combinación lineal

$$\sum_i \lambda_i v_i = 0$$

es *no trivial* si y solo si al menos algún λ_j es distinto de cero. Como veremos pronto, el hecho de que exista o no exista una combinación lineal nula no trivial de una sucesión dada de vectores es una cuestión muy relevante.

12.1.2. Subespacios

Un *subespacio* de un espacio vectorial V es un subconjunto $S \subseteq V$ con la propiedad de que la restricción a S de las operaciones definidas en V determinan en S una estructura de espacio vectorial. La siguiente definición formaliza y resume esta situación:

> Un subconjunto $S \subseteq V$ es un *subespacio* de V si y sólo si se satisfacen las siguientes tres condiciones (las letras v y w designan vectores):

(SEV_1) $0 \in S$

(SEV_2) $v, w \in S \Rightarrow v + w \in S$

(SEV_3) $v \in S \Rightarrow \lambda v \in S$ para todo $\lambda \in K$.

Ejemplos 12.1.4 Veamos algunos ejemplos importantes de la noción de subespacio. Encargamos al lector la tarea de efectuar y completar las demostraciones (m y n designan números naturales):

1) Es inmediato verificar que V y el subconjunto unitario $\{\,0\,\}$ son subespacios de cualquier espacio vectorial V, llamados *subespacios triviales*.

2) *a*) Si $(a_1, \ldots, a_n) \in K^n$, el conjunto S de soluciones $(x_1, \ldots, x_n)$ de la ecuación

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$$

es un subespacio de K^n.

En efecto, si $u = (u_1, \ldots, u_n)$ y $v = (v_1, \ldots, v_n)$ son soluciones, tenemos que

$$\sum_{i=1}^{n} a_i(u_i + v_i) = \sum_{i=1}^{n} a_iu_i + \sum_{i=1}^{n} a_iv_i = 0 + 0 = 0$$

y

$$\sum_{i=1}^{n} a_i(\lambda u_i) = \lambda \sum_{i=1}^{n} a_iu_i = \lambda 0 = 0\,,$$

lo que prueba que $u + v$ y λu son soluciones. Puesto que obviamente $(0, \ldots, 0)$ es solución del sistema, concluimos que S es un subespacio de K^n.

b) Más generalmente, dada una matriz $A = (a_{ij})$ de m filas y n columnas, el conjunto S de soluciones del sistema de m ecuaciones con n incógnitas

$$\sum_{j=1}^{n} a_{ij}\, x_j = 0 \quad (i = 1, 2, \ldots, m)$$

es un subespacio de K^n, hecho que el lector puede verificar sin inconvenientes. Un tal sistema se dice un *sistema lineal homogéneo*, y A se llama la matriz asociada al mismo. Aclaremos que el nombre de homogéneo responde al hecho de que todas la ecuaciones están igualadas a cero, condición necesaria para que el vector nulo sea solución del sistema.

3) Los siguientes subconjuntos S_i de matrices de m filas y n columnas

$$\begin{aligned} S_1 &= \{\, A \,:\, a_{ji} = a_{ij} \;\; \forall\, i, j \leq \min(m, n) \,\} \\ S_2 &= \{\, A \,:\, a_{ji} = -a_{ij} \;\; \forall\, i, j \leq \min(m, n) \,\} \\ S_3 &= \{\, A \,:\, a_{ij} = 0 \;\; \forall\, i > j \,\} \\ S_1 &= \{\, A \,:\, a_{ij} = 0 \;\; \forall\, i < j \,\} \end{aligned}$$

son subespacios de $K^{m\times n}$. Dichas matrices se llaman simétricas, antisimétricas y triangulares superiores e inferiores, respectivamente. Similarmente, los conjuntos de matrices cuadradas

$$\begin{aligned} T_1 &= \{\, A \in K^{n\times n} \,:\, a_{ij} = 0 \;\; \forall\, i \neq j \,\} \\ T_2 &= \{\, A \in K^{n\times n} \,:\, a_{ij} = 0 \;\; \forall\, i \neq j \;\; \wedge \;\; a_{ii} = a_{jj} \;\; \forall \;\; i, j \,\} \end{aligned}$$

son subespacios de $K^{n\times n}$, (matrices diagonales y escalares, respectivamente). Recomendamos al lector fabricarse unos pocos ejemplos numéricos, que le ayudarán a familiarizarse con el aspecto que presentan estos tipos de matrices.

En conexión con el subespacio S_1, se llama *transpuesta* de una matriz $X = (x_{ij}) \in K^{m\times n}$ a la matriz $Y = (y_{ij}) \in K^{n\times m}$ cuyos elementos están definidos por la relación $y_{ij} = x_{ji}$ para todo $(i, j) \in \mathbb{I}_n \times \mathbb{I}_m$. Dicho en forma más coloquial, la transpuesta de X, que notaremos X^t, tiene por filas a las columnas de X y por columnas a las filas de X. Por ejemplo:

$$\begin{pmatrix} 2 & 3 \\ 4 & 1 \\ -1 & 0 \end{pmatrix}^t = \begin{pmatrix} 2 & 4 & -1 \\ 3 & 1 & 0 \end{pmatrix}.$$

Notemos que una matriz $A \in K^{m\times m}$ es simétrica si y solo si $A^t = A$ y es antisimétrica si y solo si $A^t = -A$.

4) Fijado $a \in K$, el conjunto de todos los polinomios $f \in K[X]$ tales que $f(a) = 0$ es un subespacio de $K[X]$. Otros ejemplos son $K_n[X]$ y el conjunto de polinomios constantes, que es en realidad $K_0[X]$.

5) Si $\mathcal{F} = \{v_1, v_2, \ldots, v_n\}$ es una familia de vectores de V, es sencillo probar que el conjunto de todas sus combinaciones lineales es un subespacio de V, que llamaremos *subespacio generado* por $\mathcal{F}$ y notaremos por $gen(v_1, v_2, \ldots, v_n)$.

   Notemos que si un subespacio S de V contiene todos los v_i entonces contiene todas sus combinaciones lineales, por definición de subespacio. Por lo tanto, $gen(v_1, v_2, \ldots, v_n)$ es el menor subespacio de V, en el sentido de la inclusión, que contiene a la familia $\mathcal{F}$, lo que justifica su nombre de subespacio generado por la misma.

   En general, si T es un subespacio de V y $\{u_1, \ldots, u_m\}$ es una familia de vectores de T tal que $T = gen(u_1, \ldots, u_m)$, diremos que $\{u_1, \ldots, u_m\}$ es un *sistema de generadores* de T, o también que los vectores $u_1, \ldots, u_m$ generan T.

6) Como caso particular consideremos el conjunto de soluciones S del sistema de ecuaciones

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

   Si (a, b, c, d) es una solución del sistema, multiplicando por -2 la primera ecuación y sumándola a la segunda obtenemos que $b = -c + d$, de donde deducimos (reemplazando por este valor de b en la primera ecuación) que

$$a = b - c - d = -c + d - c - d = -2c\,.$$

   Por lo tanto, toda solución es de la forma

$$(-2c, -c + d, c, d) = c\,(-2, -1, 1, 0) + d\,(0, 1, 0, 1)\,.$$

   Puesto que es inmediato verificar que $(-2c, -c + d, c, d)$ es solución cualquiera sean c y d en $\mathbb{R}$, concluimos que S es el subespacio de $\mathbb{R}^4$ generado por los vectores $(-2, -1, 1, 0)$ y $(0, 1, 0, 1)$. $\Diamond$

12.1.3. Dependencia e independencia lineal

Para aproximarnos a la la idea que queremos desarrollar, recordemos que todo elemento de $\mathbb{R}^2$ es combinación lineal de $e_1 = (1, 0)$ y $e_2 = (0, 1)$, esto es, e_1 y e_2 generan $\mathbb{R}^2$. En un intento de generalizar este hecho, dados dos vectores v y w, digamos en $\mathbb{R}^2$, podemos preguntarnos si $gen(v, w) = \mathbb{R}^2$.

En general la respuesta a nuestro interrogante es negativa, ya que si por ejemplo $w = \lambda v$ es un múltiplo escalar de v, toda combinación lineal de ambos es entonces de la forma

$$\alpha v + \beta w = \alpha v + \beta\lambda v = (\alpha + \beta\lambda)v$$

y por lo tanto $gen(v, w)$ es la recta generada por v (si $v \neq 0$).

Operando en forma enteramente análoga, el hecho que acabamos de advertir admite la siguiente generalización:

Proposición 12.1.5 Sea V un espacio vectorial y sea $\mathcal{F} = \{v_1, \ldots, v_n\}$ una familia de vectores de V tal que v_k es combinación lineal de la familia $\mathcal{F} - \{v_k\}$ para un cierto índice k. Entonces

$$gen(v_1, v_2, \ldots, v_n) = gen(v_1, \ldots, v_{k-1}, v_{k+1}, \ldots, v_n).$$

Dicho en lenguaje coloquial, si un miembro de una familia de n generadores de un cierto subespacio T es combinación lineal de los otros dicho vector puede omitirse y los $n-1$ restantes también generan T. Por lo tanto, en la búsqueda de un conjunto de generadores de un cierto espacio vectorial lo mejor es hallar uno que tenga la menor cantidad posible de elementos.

DEMOSTRACION Dejamos los detalles a cargo del lector.

Vayamos ahora sí a definir los conceptos de dependencia e independencia lineal, íntimamente relacionados con nuestras consideraciones previas.

> Una familia $\{v_1, v_2, \ldots, v_n\}$ de vectores de un K-espacio vectorial V se dice *linealmente dependiente* (ld) si y solo si alguno de sus miembros es combinación lineal de los restantes. En caso contrario, la familia se dice *linealmente independiente* (li).

Por ejemplo, en $\mathbb{R}^3$ la familia de vectores $v_1 = (1, 2, 1)$, $v_2 = (2, 10, -10)$ y $v_3 = (0, -2, 4)$ es linealmente dependiente, pues $v_2 = 2v_1 + (-3)v_3$, mientras que puede demostrarse fácilmente que la familia$\{(1, 0, 0)(0, 1, 0)(0, 0, 1)\}$ es linealmente independiente.

Más generalmente, observemos que en todo espacio vectorial V cualquier familia de vectores que contenga al vector nulo es linealmente dependiente, ya que 0 es combinación lineal de los restantes, cualesquiera sean estos.

El siguiente resultado nos brindará definiciones equivalentes de ambos conceptos. Las mismas nos permitirán analizar con mayor eficacia la dependencia o independencia lineal de una familia de vectores, ya que reducirán la cuestión al cálculo de las soluciones de un sistema lineal homogéneo.

Proposición 12.1.6 Si V es un espacio vectorial y $\{v_1, v_2, \ldots, v_n\}$ es una familia de vectores de V, las siguientes afirmaciones son equivalentes:

(a) $\{v_1, v_2, \ldots, v_n\}$ es linealmente dependiente.

(b) 0 es combinación lineal no trivial de los vectores $v_1, v_2, \ldots, v_n$.

Notemos que se deduce de la equivalencia de (a) y (b) que $\{v_1, v_2, \ldots, v_n\}$ es linealmente independiente si y solo si la única combinación lineal nula de los vectores $v_1, v_2, \ldots, v_n$ es la trivial. Resulta en particular que si v es un vector no nulo de V la familia unitaria $\{v\}$ es linealmente independiente, por propiedad 1) del lema 12.1.1. Claramente, la familia $\{0\}$ no lo es.

DEMOSTRACION Para probar que (a) implica (b), supongamos sin pérdida de generalidad que v_n es combinación lineal de $v_1, v_2, \ldots, v_{n-1}$, digamos

$$v_n = \sum_{i=1}^{n-1} \lambda_i v_i \,,$$

donde los λ_i son escalares. Equivalentemente,

$$0 = \sum_{i=1}^{n-1} \lambda_i v_i + (-1)v_n \,,$$

y puesto que $-1 \neq 0$, concluimos que la familia de vectores $v_1, v_2, \ldots, v_n$ admite una combinación lineal nula no trivial.

Recíprocamente, supongamos que

$$0 = \sum_{i=1}^{n} \alpha_i v_i$$

es una combinación lineal nula no trivial de $\{v_1 \ldots, v_n\}$ y sea k un índice tal que $\alpha_k \neq 0$. Operando, resulta entonces que

$$\alpha_k v_k = \sum_{i=1}^{k-1} (-\alpha_i) v_i + \sum_{i=k+1}^{n} (-\alpha_i) v_i,$$

de donde sigue, multiplicando por α_k^{-1} y usando los axiomas de espacio vectorial, que

$$v_k = \sum_{i=1}^{k-1} (-\alpha_k^{-1}\alpha_i) v_i + \sum_{i=k+1}^{n} (-\alpha_k^{-1}\alpha_i) v_i \,.$$

Por lo tanto $v_k \in gen(v_1, \ldots v_{k-1}, v_{k+1} \ldots, v_n)$ y $\{v_1, v_2, \ldots, v_n\}$ es linealmente dependiente. $\Diamond$

Ejemplos 12.1.7 En los siguientes casos se apreciará la utilidad de las equivalencias anteriores.

1) Verifiquemos que la sucesión $\{(1,1,1),(1,1,0),(1,0,0)\}$ es linealmente independiente en $\mathbb{R}^3$. Supongamos para ello que λ_1, λ_2 y λ_3 son escalares tales que

$$\lambda_1(1,1,1)+\lambda_2(1,1,0)+\lambda_3(1,0,0)=(0,0,0).$$

Operando, obtenemos que

$$(\lambda_1+\lambda_2+\lambda_3,\lambda_1+\lambda_2,\lambda_1)=(0,0,0),$$

de donde sigue, igualando las componentes, que los escalares deben satisfacer el sistema de ecuaciones

$$\begin{cases}\lambda_1+\lambda_2+\lambda_3=0\\ \lambda_1+\lambda_2=0\\ \lambda_1=0\,.\end{cases}$$

Puesto que se deduce inmediatamente que $\lambda_1=\lambda_2=\lambda_3=0$, concluimos que la única combinación lineal nula de los vectores dados es la trivial, y por lo tanto la familia formada por ellos es linealmente independiente.

2) Analicemos ahora si la sucesión $\left\{1+3X,2+X^2,X^2-6X\right\}$ es linealmente independiente en $\mathbb{R}[X]$. Como antes, sean λ_1, λ_2 y λ_3 números reales tales que

$$\lambda_1(1+3X)+\lambda_2(2+X^2)+\lambda_3(X^2-6X)=0,$$

lo que es equivalente a la igualdad polinomial

$$(\lambda_1+2\lambda_2)+(3\lambda_1-6\lambda_3)X+(\lambda_2+\lambda_3)X^2=0.$$

Teniendo en cuenta la definición de igualdad de polinomios, arribamos esta vez al sistema de ecuaciones

$$\begin{cases}\lambda_1+2\lambda_2=0\\ 3\lambda_1-6\lambda_3=0\\ \lambda_2+\lambda_3=0\,,\end{cases}$$

que tiene infinitas soluciones, ya que de las igualdades $\lambda_1=-2\lambda_2$ y $\lambda_3=-\lambda_2$ deducimos que la terna $(-2a,a,-a)$ es solución cualquiera sea el número real a. En consecuencia, la familia de polinomios dada es linealmente dependiente.

A manera de ejemplo, tomando $a=-1$ resulta que vale la relación

$$2(1+3X)+(-1)(2+X^2)+1(X^2-6X)=0\,.$$

3) La dependencia o independencia lineal de una sucesión de vectores no solo depende de los vectores en sí, sino también del cuerpo de escalares considerado. Tomemos por ejemplo $V = \mathbb{C}$, el cuerpo de los números complejos. Si pensamos a $\mathbb{C}$ como $\mathbb{R}-$espacio vectorial (es decir, si sólo admitimos escalares reales) entonces es fácil comprobar que la sucesión $\{1, \imath\}$ es linealmente independiente. En cambio, la misma es linealmente dependiente si consideramos a $\mathbb{C}$ como $\mathbb{C}$-espacio vectorial, ya que $\imath 1 + (-1)\imath = 0$ es una combinación lineal nula no trivial de los vectores 1 e $\imath$ (nótese que cuando escribimos $\imath 1$ pensamos a $\imath$ como escalar, mientras que cuando escribimos $(-1)\imath$ lo pensamos como vector).

Veamos a continuación algunas propiedades básicas de la dependencia e independencia lineal.

Proposición 12.1.8 Sea V un espacio vectorial y sea $\mathcal{F} = \{v_1, v_2, \ldots, v_n\}$ una familia de vectores de V. Son válidas entonces las siguientes propiedades:

1) $\mathcal{F}$ es linealmente independiente (dependiente) si y sólo si cualquier familia obtenida por permutación de sus miembros es linealmente independiente (dependiente).

2) Si $v_r = v_s$ para algún par de índices distintos r y s entonces $\mathcal{F}$ es linealmente dependiente. Dicho de otro modo, una sucesión linealmente independiente de vectores no puede contener elementos repetidos.

3) $\mathcal{F}$ es linealmente independiente si y solo si toda subfamilia de $\mathcal{F}$ lo es. Como ya veremos, no vale un resultado análogo respecto de la dependencia lineal.

4) Supongamos que $\mathcal{F}$ es linealmente independiente y sea $w \in V$ tal que $\{v_1, \ldots, v_n, w\}$ no lo es. Entonces $w \in gen(v_1, \ldots, v_n)$.

5) Supongamos que $\mathcal{F}$ es linealmente independiente y sea $\{w_1, \ldots, w_{n-1}\}$ la familia de vectores definidos por las relaciones $w_i = v_i - \alpha_i v_n$, donde $\alpha_1, \ldots, \alpha_{n-1}$ es una familia arbitraria de escalares. Entonces la familia $\{w_1, \ldots, w_{n-1}\}$ también es linealmente independiente.

6) Toda familia de $n + 1$ o más vectores en $gen(v_1, v_2, \ldots, v_n)$ es linealmente dependiente.

Demostracion Dejamos la prueba de 1) a cargo del lector, ya que la misma es consecuencia inmediata de las definiciones. Para probar 2), y de acuerdo con 1), podemos suponer sin pérdida de generalidad que $v_1 = v_2$. En tal caso,

$$0 = 1v_1 + (-1)v_2 + \sum_{i=3}^{n} 0\, v_i$$

es una combinación lineal nula no trivial de la familia $\mathcal{F}$, resultando que ella es linealmente dependiente.

Respecto de 3), asumamos por el absurdo que $\mathcal{F}$ admite una subfamilia linealmente dependiente, que sin pérdida de generalidad podemos suponer de la forma $\{v_1, \ldots, v_m\}$, con $m < n$. Entonces, si $\lambda_1, \ldots, \lambda_m$ es una sucesión de escalares (no todos nulos) tales que

$$\sum_{i=1}^{m} \lambda_i v_i = 0,$$

resulta que

$$0 = \sum_{i=1}^{m} \lambda_i v_i + \sum_{i=m+1}^{n} 0\, v_i$$

es una combinación lineal nula no trivial de $\mathcal{F}$, lo que es un contradicción.

En relación al comentario que figura en el enunciado de 3), observemos que si v es un vector no nulo de V entonces $\{v\}$ es li y $\{v, v\}$ es ld, lo que muestra que una familia linealmente dependiente puede contener subfamilias linealmente independientes.

Con respecto a 4), resulta por hipótesis que existen escalares $\lambda_1 \ldots, \lambda_n, \lambda$, no todos nulos, tales que

$$\sum_{i=1}^{n} \lambda_i v_i + \lambda w = 0\,.$$

Si $\lambda = 0$, resultaría que $\lambda_i \neq 0$ para algún i y por lo tanto la familia $\mathcal{F}$ sería linealmente dependiente. En consecuencia $\lambda \neq 0$ y podemos despejar w de la igualdad de arriba, obteniendo

$$w = \sum_{i=1}^{n} (-\lambda^{-1}\lambda_i) v_i\,.$$

Luego $w = gen(v_1, \ldots, v_n)$, como queríamos demostrar.

Probemos ahora 5). Considerando para ello una combinación lineal nula de los w_i, con escalares $\beta_1, \ldots, \beta_{n-1}$, y operando, obtenemos:

$$\begin{aligned} 0 = \sum_{i=1}^{n-1} \beta_i w_i = \sum_{i=1}^{n-1} \beta_i \left(v_i - \alpha_i v_n\right) = \\ = \sum_{i=1}^{n-1} \beta_i v_i - \left(\sum_{i=1}^{n-1} \beta_i \alpha_i\right) v_n\,. \end{aligned}$$

Puesto que por hipótesis esta combinación lineal nula de la familia $\mathcal{F}$ debe ser trivial, resulta en particular que $\beta_i = 0$ para todo i, y por lo tanto la familia $w_1, \ldots, w_{n-1}$ es linealmente independiente.

Para demostrar 6), notemos de entrada que basta probar por 4) que toda familia de $n+1$ vectores en $gen(v_1, \ldots, v_n)$ es linealmente dependiente. Procederemos para ello por inducción en n.

Dejando el caso $n = 1$ a cargo del lector, supongamos que la afirmación es válida para n y consideremos cualquier familia de $n+2$ vectores $w_1, \ldots, w_{n+2}$ en el subespacio generado por los $n+1$ vectores $v_1, \ldots, v_n, v_{n+1}$.

Podría suceder que todos los w_i estuvieran en realidad en $gen(v_1, \ldots, v_n)$, en cuyo caso la familia $w_1, \ldots, w_{n+2}$ sería linealmente dependiente por hipótesis inductiva. Suponiendo entonces que tal situación no se registra, podemos asumir sin pérdida de generalidad que $w_{n+2} \notin gen(v_1, \ldots, v_n)$.

De acuerdo con las hipótesis, para cada índice i $(1 \leq i \leq n+1)$ podemos escribir

$$w_i = a_1^i v_1 + a_2^i + \cdots + a_{n+1}^i v_{n+1},$$

donde los a_j^i son escalares y $a_{n+1}^{n+2} \neq 0$.

Consideremos ahora la familia $u_1, \ldots, u_{n+1}$ de vectores definidos por las relaciones

$$u_i = w_i - (a_{n+1}^{n+2})^{-1} a_{n+1}^i w_{n+2}.$$

Por construcción, resulta que

$$\begin{aligned} u_i &= \sum_{j=1}^{n+1} a_j^i v_j - \sum_{j=1}^{n+1} \left((a_{n+1}^{n+2})^{-1} a_{n+1}^i a_j^{n+2} \right) v_j = \\ &= \sum_{j=1}^{n} \left(a_j^i - (a_{n+1}^{n+2})^{-1} a_{n+1}^i a_j^{n+2} \right) v_j + a_{n+1}^i \, v_{n+1} - a_{n+1}^i v_{n+1} = \\ &= \sum_{j=1}^{n} \left(a_j^i - (a_{n+1}^{n+2})^{-1} a_n^i a_j^{n+2} \right) v_j \,, \end{aligned}$$

esto es, los $n+1$ vectores u_i están en el subespacio generado por $v_1, v_2, \ldots, v_n$. Por hipótesis inductiva resulta entonces que la familia $\{u_1, \ldots, u_{n+1}\}$ es ld, de donde sigue, por (5), que la familia $\{w_1 \ldots, w_{n+1}, w_{n+2}\}$ también lo es, como queríamos demostrar.

Remarquemos el significado de esta última propiedad: en un espacio vectorial generado por n vectores toda familia linealmente independiente tiene a lo sumo n elementos. $\Diamond$

12.1.4. Ejercicios

A lo largo de los siguientes ejercicios la letra V denota un espacio vectorial sobre un cuerpo cualquiera K (m y n designan números naturales y J es un conjunto):

1. Probar que la conmutatividad de la suma en V se deduce de los restantes axiomas de espacio vectorial.

2. Sea $\mathcal{V}$ un conjunto coordinable con V a través de una biyección θ. Probar que las operaciones

$$x + y = \theta^{-1}(\theta(x) + \theta(y))$$
$$\lambda x = \theta^{-1}(\lambda\theta(x))$$

definen una estructura de K-espacio vectorial en $\mathcal{V}$ $(x, y \in \mathcal{V} \wedge \lambda \in K)$.

3. Si K es infinito, probar que todo vector no nulo de V admite infinitos múltiplos escalares distintos. Deducir que sobre un cuerpo infinito todo espacio vectorial no trivial también es infinito.

4. Si $S \subseteq V$, probar que S es un subespacio de V si y solo si se satisfacen las siguientes condiciones:

   a) $S \neq \emptyset$
   b) $u + \alpha w \in S$ cualesquiera sean $u, w \in S$ y $\alpha \in K$.

5. En cada uno de los siguientes casos, decidir si S es un subespacio del espacio vectorial W dado:

   a) $W = K^n$; $S = \{ (x_1, \ldots, x_n) : x_1 = x_n = 0 \}$
   b) $W = K^n$; $S = \{ (x_1, \ldots, x_n) : x_1^2 + x_2 - x_3 = 0 \}$
   c) $W = K^n$; $S = \{ (x_1, \ldots, x_n) : x_1 = 1 - x_n \}$
   d) $W = K^{m\times n}$; $S = \{ (a_{ij}) : a_{11} - a_{mn} = a_{1n} + a_{m1} \}$
   e) $W = \mathbb{R}^{m\times n}$; $S = \left\{ (a_{ij}) : \sum_{j=1}^{n} a_{1j} \geq 0 \right\}$
   f) $W = \mathbb{R}^{\mathbb{R}}$; $S = \{ f : f \text{ es continua} \}$
   g) $W = \mathbb{R}^{\mathbb{R}}$; $S = \{ f : f \text{ es creciente} \}$
   h) $W = \mathbb{R}^J$; $S = \{ f : f(a) = 2f(b) \}$ $(a, b \in J)$
   i) $W = \mathbb{R}^J$; $S = \{ f : f(a)f(b) = 0 \}$ $(a, b \in J)$
   j) $W = \mathbb{R}[X]$; $S = \{ f : f^3(1) = 0 \}$
   k) $W = \mathbb{R}[X]$; $S = \{ f : f \text{ es divisible por } X^2 + 1 \}$.

6. Demostrar que los conjuntos S_i y T_i del ejemplo 3) de 12.1.4 son subespacios de $K^{m\times n}$ y $K^{n\times n}$, respectivamente.

7. Si $v_1, \ldots, v_n$ son vectores de V, probar que $gen(v_1, \ldots, v_n)$ es un subespacio de V.

8. Si $u, w \in V$, mostrar que $gen(2u + w, u + w) = gen(u, w)$.

9. Si q es un primo, calcular el número de rectas de $\mathbb{Z}_q^2$.

10. Sea $X \subseteq V$ y sea $gen(X)$ el conjunto de combinaciones lineales de subfamilias finitas de X.

    a) Probar que $gen(X)$ es un subespacio de V que contiene a X y que todo subespacio T de V que contiene a X también contiene a $gen(X)$.

    b) Si X_1 y X_2 son subconjuntos de V, probar que

    $$X_1 \subseteq X_2 \Rightarrow gen(X_1) \subseteq gen(X_2).$$

    c) Demostrar que $gen(X) = X$ si y solo si X es un subespacio de V.

11. Sea U un $\mathbb{Q}$-espacio vectorial y sea S un subgrupo de la estructura aditiva de U. Probar que S es un subespacio de U.

12. Dados subespacios S y T de V, probar las siguientes propiedades:

    a) $S \cap T$ es un subespacio de V.

    b) $S \cup T$ es un subespacio de V si y sólo si $S \subseteq T$ ó $T \subseteq S$.

    c) El conjunto

    $$S + T = \{u + v : u \in S \ \wedge \ v \in T\}$$

    es un subespacio de V, llamado la *suma* de S y T. Demostrar además que es el menor subespacio, en el sentido de la inclusión, que contiene a S y a T.

13. Si S y T son como en 12), diremos que la suma de S y T es *directa* si y sólo si $S \cap T = (0)$, en cuyo la notaremos $S \oplus T$.

    Probar que la suma es directa si y solo si todo elemento de $S + T$ se expresa *en forma única* como suma de un elemento de S y uno de T.

14. *a*) Probar que V admite la descomposición trivial $V = (0) \oplus V$.

    b) Probar que $\mathbb{R}^3 = gen((2, 1, 0)) \oplus \{(x_1, x_2, x_3) : x_1 - x_2 + x_3 = 0\}$.

    c) Respecto del ejemplo 3) de 12.1.4, analizar en qué casos la suma $S_i + S_j$ es directa.

15. Hallar sistemas de generadores de todos los subespacios del ejercicio 5) y de los subespacios de matrices S_i y T_i del ejemplo 3) de 12.1.4.

16. ¿ Es cierto que si $\mathcal{F}$ es una familia linealmente dependiente de vectores de V entonces cada uno de sus elementos es combinación lineal de los restantes ?

17. En cada uno de los siguientes espacios vectoriales, decidir acerca de la dependencia o independencia lineal de la familia de vectores $\mathcal{F}$ dada:

    a) $W = \mathbb{R}^3$; $\mathcal{F} = \{\, (2,-1,0), (1,1,-3), (-1,-5,2) \,\}$

    b) $W = \mathbb{R}^4$; $\mathcal{F} = \{\, (1,1,2,1), (0,-2,2,4), (-1,0,3,2), (0,2,4,1) \,\}$

    c) $W = \mathbb{R}[X]$; $\mathcal{F} = \left\{\, X + X^2, 1 - 2X, 3 - X + X^2, 2 + X^2 \,\right\}$

    d) $W = \mathbb{R}^{\mathbb{R}}$; $\mathcal{F} = \{\, \cos x, \operatorname{sen} x, e^x \,\}$

    e) $W = K[X]$; $\mathcal{F} = \{\, f_1, \ldots, f_m \,\}$, con $gr(f_i) \neq gr(f_j)$ si $i \neq j$.

18. Sea $\{u_1, \ldots, u_m\}$ una familia linealmente independiente de vectores de un subespacio S de V y sea $u \in V - S$. Probar que $\{u_1, \ldots, u_m, u\}$ es linealmente independiente.

19. Sean S y T subespacios de V tales que $S \cap T = (0)$ y sean $\mathcal{F}$ y $\mathcal{G}$ familias linealmente independientes de vectores de S y T, respectivamente. Probar que la familia $\mathcal{F} \cup \mathcal{G}$ es linealmente independiente.

12.2. Bases y dimensión

12.2.1. Espacios finitamente generados

Un $K-$espacio vectorial V se dice *finitamente generado* si y solo si admite un sistema de generadores finito. Vale decir, todo elemento de V puede expresarse como combinación lineal, con escalares en K, de una cierta familia finita de vectores de V.

Ejemplos 12.2.1 Ilustremos la definición con algunos ejemplos sencillos (m y n designan números naturales):

1) Obviamente, el espacio vectorial nulo es finitamente generado.

2) Ya vimos que el espacio de n-uplas K^n es finitamente generado para todo $n \in \mathbb{N}$, siendo $\{e_1, e_2, \ldots, e_n\}$ un sistema de generadores del mismo.

3) El espacio de polinomios $K[X]$ no es finitamente generado sobre K. En efecto, si $f_1, \ldots, f_m$ es cualquier familia finita de polinomios no nulos y r es el máximo de sus grados, deducimos inmediatamente que toda combinación lineal de los f_i tiene grado a lo sumo r o bien es el polinomio nulo. Resulta en particular que $X^{r+1} \notin gen(f_1, \ldots, f_m)$, lo que prueba que los f_i no generan $K[X]$.

4) Todo subespacio de un espacio vectorial es también un espacio vectorial, por lo que tiene sentido analizar si es o no finitamente generado. Por ejemplo, dado $n \in \mathbb{N}$, el conjunto

$$K_n[X] = \{f \in K[X] \,:\, gr(f) \leq n \text{ ó } f = 0\}$$

es un subespacio finitamente generado de $K[X]$ y $\left\{1, X, \ldots, X^{n-1}\right\}$ es un sistema de generadores del mismo (el lector demostrará estos hechos).

5) $K^{m\times n}$ está finitamente generado por la familia de mn matrices

$$\{E^{rs} : (r,s) \in \mathbb{I}_m \times \mathbb{I}_n\},$$

donde la matriz $E^{r\,s}$ está definida por $E^{rs}_{ij} = \delta_{ir}\delta_{js}$.

En efecto, es muy fácil verificar que toda matriz $A = (a_{ij}) \in K^{m\times n}$ se expresa unívocamente en la forma

$$A = \sum_{(i,j)\in \mathbb{I}_m\times\mathbb{I}_n} a_{ij}\, E^{ij}\,.$$

Similarmente, el conjunto de matrices $\left\{E^{ii} : (1 \leq i \leq n)\right\}$ es un sistema finito de generadores del subespacio de matrices diagonales de $K^{n\times n}$.

6) La condición de generación finita no sólo depende de los vectores sino también de los escalares considerados. Por ejemplo, K es finitamente generado como espacio vectorial sobre sí mismo (cualquier elemento no nulo lo genera), pero no necesariamente es finitamente generado como espacio vectorial sobre un subcuerpo cualquiera. Como ejemplo más cercano citemos el caso de $\mathbb{R}$, que no es finitamente generado sobre $\mathbb{Q}$. No brindaremos aquí una demostración de este hecho, puesto que sus dificultades exceden el nivel de estas páginas, sólo digamos que se trata esencialmente de una cuestión de cardinalidad. $\Diamond$

En lo sucesivo, a menos que se indique lo contrario supondremos que los espacios vectoriales considerados son finitamente generados.

Vimos antes que si una familia de generadores de un espacio vectorial V es linealmente dependiente puede suprimirse alguno de sus vectores de manera que los restantes también generen V. Expresado informalmente, esto significa que los sistemas de generadores linealmente dependientes adolecen de cierta redundancia, por lo que resulta de más interés considerar sistemas de generadores que además sean linealmente independientes. En la siguiente definición les daremos un nombre especial a tales sistemas:

> Si un sistema de generadores de un espacio vectorial V es linealmente independiente diremos que es una *base* de V.

Antes de exhibir algunos ejemplos mostraremos una definición equivalente de base, de relevante significado:

Proposición 12.2.2 Una familia $\{v_1, v_2, \ldots, v_n\}$ de vectores es una base de un espacio vectorial V si y solo si todo elemento de V se expresa *de manera única* como combinación lineal de los v_i.

Demostracion. Puesto que las dos condiciones del enunciado afirman que la familia $\{v_1, v_2, \ldots, v_n\}$ genera V, debemos demostrar que la hipótesis de unicidad en la descomposición es equivalente a la independencia lineal de dicha familia. Asumiendo en primer término que $\{v_1, v_2, \ldots, v_n\}$ es li, supongamos que un cierto vector v de V admite las descomposiciones

$$v \,=\, \sum_{i=1}^{n} \alpha_i \, v_i \,=\, \sum \beta_i \, v_i \,,$$

donde los α_i y los β_i son escalares. Entonces

$$0 = \sum_{i=1}^{n} \alpha_i \, v_i - \sum_{i=1}^{n} \beta_i \, v_i = \sum_{i=1}^{n} (\alpha_i - \beta_i) \; v_i \,,$$

de donde sigue por hipótesis que $\alpha_i - \beta_i = 0$ para todo i. Luego $\alpha_i = \beta_i$ para todo i y la descomposión es única.

Recíprocamente, supongamos que todo elemento de V se descompone en forma única como combinación lineal de los v_i y consideremos una combinación lineal nula de los mismos con escalares λ_i. Observamos entonces que

$$0 = \sum_{i=1}^{n} \lambda_i \, v_i \; = \; \sum_{i=1}^{n} 0 \, v_i \, ,$$

de donde resulta, por la unicidad de la representación, que $\lambda_i = 0$ para todo i. En consecuencia la combinación lineal es trivial y la familia $\{v_1, v_2, \ldots, v_n\}$ es linealmente independiente. $\Diamond$

El resultado que acabamos de probar nos permite definir correctamente la noción de coordenada, que pasamos a enunciar:

> Sea V un espacio vectorial y sea $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ una base de V. Si $v \in V$ se expresa en la forma
>
> $$v = \sum_{i=1}^{n} \lambda_i v_i \, ,$$
>
> diremos que la n-upla $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ es el *vector de coordenadas* de v en la base $\mathcal{B}$.

Por ejemplo, todo elemento de $\mathbb{R}^2$ coincide con su vector de coordenadas en la base $\{e_1, e_2\}$. Por otro lado, el lector puede verificar sin dificultad que el vector de coordenadas de $(1, 18)$ en la base $\{(2, 1), (-1, 3)\}$ es $(3, 5)$.

NOTA Resaltemos el significado del resultado anterior. Fijada una base $\mathcal{B}$ del espacio (ya veremos que existe una), cada vector del espacio está biunívocamente asociado con una secuencia de escalares. Es muy fácil de ver que esta identificación de vectores con secuencias a través de las coordenadas es compatible con las operaciones (a la suma de dos vectores le corresponde la suma de las respectivas secuencias, al vector nulo le corresponde la secuencia nula, etc.), resultando así que un espacio vectorial que admite una base de n elementos es en esencia una copia algebraica de K^n.

Notemos asimismo que en la noción de coordenadas se refleja notablemente el hecho de que una base es un conjunto *ordenado* de vectores. Por ejemplo, es sencillo verificar que $\mathcal{B} = \{(1, 1), ((2, 0)\}$ es una base de $\mathbb{R}^2$ y que el vector de coordenadas de $u = (-1, 3)$ con respecto a $\mathcal{B}$ es $(3, -2)$, mientras que el vector de coordenadas de u con respecto a la base $\{(2, 0), (1, 1)\}$ es $(-2, 3)$. $\Diamond$

Ejemplos 12.2.3 Veamos ejemplos de bases (encargamos al lector verificar la validez de las afirmaciones).

1) La sucesión $\{e_1, e_2, \ldots, e_n\}$ es una base de K^n. Precisamente, toda n-upla de elementos de K se expresa en unívocamente en la forma

$$(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} x_i \, e_i \, .$$

Se la llama *base canónica* de K^n, debido al hecho de que el vector de coordenadas de cada elemento coincide con éste.

2) La familia
$$\left\{E^{ij} : (i,j) \in \mathbb{I}_m \times \mathbb{I}_n\right\}$$
es una base de $K^{m\times n}$ (también llamada canónica) y la matriz
$$\sum_{i=1}^{n} E^{ii}$$
es una base del subespacio de matrices escalares de $K^{n\times n}$.

3) La familia $\{1, X, \ldots, X^n\}$ es una base de $K_n[X]$.

4) La sucesión
$$w_1 = (1, 0, \ldots, 0)\,,\ w_2 = (1, 1, 0, \ldots, 0)\,,\ \ldots,\ x_n = (1, 1, \ldots, 1)$$
también es una base de K^n. Para asegurarse de ello, al lector le bastará demostrar que toda n-upla $(x_1, x_2, \ldots, x_n)$ se expresa de manera única como combinación lineal de los w_k, en la forma
$$(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n-1}(x_i - x_{i+1})\, w_i + x_n\, w_n\,.$$
Por ejemplo, $(3, -1, -3, 3)$ es el vector de coordenadas de $(2, -1, 0, 3)$ en la base $\{w_1, w_2, w_3, w_4\}$ de $\mathbb{R}^4$.

1. La familia $\{1, i\}$ es una base de $\mathbb{C}$ como $\mathbb{R}$-espacio vectorial, mientras que $\{1\}$ es una base de $\mathbb{C}$ como $\mathbb{C}$-espacio vectorial. $\Diamond$

A través del siguiente teorema probaremos dos resultados fundamentales de la teoría, relacionados con la existencia de bases en un espacio vectorial y con el cardinal de las mismas.

Teorema 12.2.4 Todo espacio vectorial no nulo V admite una base y dos bases cualesquiera de V tienen el mismo cardinal.

DEMOSTRACION. Puesto que V es finitamente generado, usando el principio de buena ordenación de los números naturales podemos asegurar la existencia de un sistema de generadores de V de cardinal mínimo s, digamos $\{v_1, v_2, \ldots, v_s\}$. Si esta sucesión de vectores no fuera linealmente independiente, asumiendo sin pérdida de generalidad que v_s es combinación lineal de los restantes vectores resultaría que
$$V = gen\,(v_1, \ldots, v_{s-1}, v_s) = gen\,(v_1, \ldots, v_{s-1})\,,$$

lo que no es posible por la minimalidad de s. Luego $\{v_1, \ldots, v_s)$ es linealmente independiente y por lo tanto es una base de V.

Para probar la segunda afirmación, consideremos dos bases $\{u_1, \ldots, u_m\}$ y $\{w_1, \ldots, w_n\}$ de V. Resulta entonces en particular que $\{u_1, \ldots, u_m\}$ es una familia linealmente independiente contenida en un espacio vectorial generado por n vectores, de donde sigue por propiedad 6) de la proposición 12.1.8 que $m \leq n$. Invirtiendo los roles de los u_i y de los w_i obtenemos de la misma manera que $n \leq m$, de donde concluimos que $m = n$.

Nótese que en el curso de la demostración ha quedado establecido un hecho interesante en sí mismo: en cualquier espacio vectorial finitamente generado todo sistema de generadores de cardinal mínimo es una base. $\Diamond$

12.2.2. Dimensión

A partir del teorema anterior estamos en condiciones de definir un concepto central de la teoría de espacios vectoriales, el concepto de dimensión:

> Si V es un espacio vectorial no nulo, llamaremos *dimensión* de V al cardinal n de cualquiera de sus bases. Emplearemos en tal caso la notación $dim_K V = n$, o simplemente $dim\, V = n$, si es claro por el contexto cuál es el cuerpo de escalares considerado. De modo de considerar todas las situaciones, diremos también que el espacio vectorial nulo tiene dimensión cero.

Ejemplos 12.2.5 En los siguientes ítems m y n designan números naturales. El lector se encargará de justificar debidamente las siguientes afirmaciones:

1) $dim\, K^n = n$.

2) $dim\, K_n[X] = n$.

3) $dim\, K^{m\times n} = mn$.

4) Sea $(a_1, a_2, \ldots, a_n)$ un elemento no nulo de K^n y sea $S \subseteq K^n$ el subespacio de soluciones de la ecuación

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0\,.$$

   Entonces $dim\, S = n - 1$.

5) Sean U el subespacio de matrices simétricas de $K^{4\times 5}$ y V el subespacio de matrices diagonales de $K^{m\times m}$. Entonces $dim\, U = 14$ y $dim\, V = m$.

6) $dim_{\mathbb{C}}\, \mathbb{C} = 1$ y $dim_{\mathbb{R}}\, \mathbb{C} = 2$. $\Diamond$

Debido a su frecuente uso es útil introducir el siguiente lenguaje:

> Diremos que una familia de vectores $\mathcal{F} = \{v_1, \ldots, v_r\}$ de V *puede extenderse a una base* si existen en V vectores $v_{r+1}, \ldots, v_n$ tales que $\{v_1, \ldots, v_r, v_{r+1}, \ldots, v_n\}$ es una base de V. Dualmente, diremos que de la familia $\mathcal{F}$ *puede extraerse una base* si alguna subfamilia de $\mathcal{F}$ es una base de V.

Interpretaremos que una base satisface ambas definiciones (estaríamos, según el caso, agregando o quitando una familia vacía de vectores). De acuerdo con resultados anteriores, para que $\mathcal{F}$ pueda extenderse a una base es necesario que sea linealmente independiente, mientras que para que pueda extraerse una base de ella es necesario que sus elementos generen V, ya que, obviamente, todo conjunto que contiene a un sistema de generadores también es un sistema de generadores. Veremos a continuación que ambas condiciones también son suficientes.

Proposición 12.2.6 Toda familia linealmente independiente $\mathcal{F}_1$ de vectores de V puede extenderse a una base de V, y de toda familia de generadores $\mathcal{F}_2$ de V puede extraerse una base de V.

Demostracion Sea $\mathcal{F}_1 = \{v_1, \ldots, v_r\}$ y supongamos que V admite una base de m elementos. Como sabemos, debe ser $r \leq m$. Probaremos nuestro primer enunciado por inducción en el entero no negativo $l = m - r$.

Si $l = 0$, esto es, $r = m$, $\mathcal{F}_1$ es una base de V y el resultado sigue. En efecto, en caso contrario existiría un vector $u \in V$ que no es combinación lineal de la familia $\mathcal{F}_1$, y entonces la sucesión $\{v_1, \ldots, v_m, u\}$ sería li, lo que es una contradicción. Obsérvese que hemos utilizado sucesivamente las propiedades 4) y 6) de la proposición 12.1.8.

Suponiendo ahora que $l > 0$ y que $\mathcal{F}_1$ no es una base de V, tomemos un vector $v_{r+1} \in V$ que no sea combinación lineal de la familia $\mathcal{F}_1$, en cuyo caso la familia $\mathcal{E}_1 = \{v_1, \ldots, v_r, v_{r+1}\}$ es linealmente independiente. Puesto que

$$m - (r + 1) = m - r - 1 = l - 1,$$

sigue por un argumento inductivo que $\mathcal{E}_1$ puede ser extendida a una base de V, y por lo tanto también $\mathcal{F}_1$, por ser una subfamilia de $\mathcal{E}_1$. Queda así completada la prueba por inducción.

Para probar la validez de la segunda afirmación, basta elegir una subfamilia de $\mathcal{F}_2$ de cardinal mínimo respecto a la propiedad de generar V. Razonando en forma muy similar a la del teorema anterior, se demuestra que la misma es una base de V. $\Diamond$

NOTA. Quisiéramos que el formalismo de las demostraciones no impida al lector apreciar la sencillez de la idea que encierran. Conservando la notación empleada podemos resumir la primera como sigue: si la familia $\mathcal{F}_1$ es una base, la cuestión ya está resuelta. Si no, se agrega a ella un vector que no sea combinación lineal de sus miembros, obteniéndose de esa manera una

nueva familia linealmente independiente, y se va iterando este proceso hasta obtener una base del espacio. Lo que garantiza que la construcción termina en un número finito de pasos es el hecho de que el cardinal de cualquier familia linealmente independiente está acotado por el cardinal de cualquier sistema de generadores. Precisamente, el proceso finalizará en $m - r$ pasos.

En el segundo caso, si $\mathcal{F}_2$ no es linealmente independiente alguno de sus elementos es combinación lineal de los restantes y podemos descartarlo, obteniendo un nuevo sistema de generadores. Siguiendo en este plan de ir quitando sucesivamente un vector, y suponiendo que $\mathcal{F}_2$ es de cardinal s, es claro que en a lo sumo $s - 1$ pasos obtendremos un sistema de generadores que también es linealmente independiente, como es nuestro propósito. $\Diamond$

Ejemplo 12.2.7 Veamos cómo extender el vector $(1, 2, 0)$ a una base de $\mathbb{R}^3$. Puesto que claramente $(1, 0, 0)$ no es un múltiplo escalar de $(1, 2, 0)$, deducimos que la familia $\{(1, 2, 0), (1, 0, 0)\}$ es linealmente independiente. Observemos por otra parte que cualquier vector que sea combinación lineal de estos dos últimos tendrá su tercera componente nula, resultando en particular que $(0, 0, 1) \notin gen\,((1, 2, 0), (1, 0, 0))$, y por lo tanto la familia $\mathcal{S} = \{(1, 2, 0), (1, 0, 0), (0, 0, 1)\}$ es linealmente independiente.

Sabiendo además que ninguna familia linealmente independiente de $\mathbb{R}^3$ puede contener más de 3 vectores, pues $\mathbb{R}^3$ tiene una base de 3 elementos (la canónica), concluimos $\mathcal{S}$ es una base como la que buscábamos. $\Diamond$

Como corolario inmediato de 12.2.6 obtenemos las siguientes equivalencias, que facilitan la determinación de bases en un espacio vectorial cuya dimensión conocemos previamente.

Teorema 12.2.8 Si $dim\,V = n$ y $\mathcal{F} = \{v_1, v_2, \ldots, v_n\}$ es una familia de vectores de V, las siguientes afirmaciones son equivalentes:

(a) $\mathcal{F}$ es linealmente independiente.

(b) $\mathcal{F}$ es un sistema de generadores de V.

(c) $\mathcal{F}$ es una base de V.

DEMOSTRACION Dejando los detalles de la demostración a cargo del lector, insistimos acerca de la utilidad práctica del hecho: si en un espacio vectorial el cardinal de una cierta familia de vectores coincide con la dimensión del espacio, para probar que dicha familia es una base será suficiente verificar sólo una de las dos condiciones que definen el carácter de base. $\Diamond$

Para cerrar el capítulo, veamos qué puede decirse acerca de los subespacios de un espacio vectorial finitamente generado

Proposición 12.2.9 Si $n \in \mathbb{N}$ y V es un espacio vectorial de dimensión n, todo subespacio S de V es finitamente generado y $dim\,S \leq n$. Además, $dim\,S = n$ si y solo si $S = V$.

DEMOSTRACION El resultado es trivial si S es el subespacio nulo. Si no, puesto que el cardinal de cualquier familia linealmente independiente de vectores de V está acotado por n, podemos elegir en S una familia linealmente independiente de cardinal máximo, digamos $\mathcal{F} = \{u_1, \ldots, u_r\}$. Si $u \in S$, la familia $\{u_1, \ldots, u_r, u\}$ es linealmente dependiente, por la maximalidad de r, de donde deducimos, por propiedad 4) de 12.1.8, que $u \in gen\,(u_1, \ldots, u_r)$. Luego $\mathcal{F}$ es una base de S y $dim\, S = r \leq n$.

Respecto a la segunda afirmación del enunciado, supongamos que S tiene dimensión n y consideremos una base $\mathcal{B} = \{w_1, \ldots, w_n\}$ de S. Puesto que en particular la familia $\mathcal{B}$ es linealmente independiente, sigue de 12.2.8 que $\mathcal{B}$ es un sistema de generadores de V, y por lo tanto

$$S = gen\,(w_1, \ldots, w_n) = V\,.$$

La recíproca es trivial. $\Diamond$

12.2.3. Ejercicios

En lo que sigue V denota un espacio vectorial finitamente generado sobre K y m y n designan números naturales.

1. Si I es un conjunto no vacío, demostrar que K^I es finitamente generado si y solo si I es finito. Probar en tal caso que $dim\, K^I = \#(I)$.

2. Sea $g \in K[X]$ y sea S el conjunto de polinomios con coeficientes en K divisibles por g. Probar que S es un subespacio de $K[X]$ y analizar si es finitamente generado.

3. Determinar la dimensión y exhibir una base de cada uno de los siguientes espacios vectoriales reales:

   a) Los subespacios de matrices S_i y T_i del ejemplo 12.1.4

   b) $W = \{(a_{ij}) \in \mathbb{R}^{n\times n} : \sum_i a_{ii} = 0\}$

   c) $W = \{f \in \mathbb{R}_n[X] : f(1) = 0\}$

   d) El conjunto de soluciones del sistema lineal

   $$\begin{cases} x_1 - x_2 + x_3 - 2x_4 = 0 \\ 4x_1 + 5x_2 - 11x_3 + 4x_4 = 0 \\ 2x_1 + x_2 - 3x_3 = 0 \end{cases}$$

   e) $W = gen\,((1,2,0,1),(0,1,1,-1)) + gen\,((1,1,-1,3),(1,3,1,0))$.

4. *a*) Probar que $\mathcal{B}_1 = \{(1,2,1),(2,1,1),(-1,1,1)\}$ es una base de $\mathbb{R}^3$ y que $\mathcal{B}_2 = \{(1,0,0,0),(1,-1,0,0),(1,1,-1,0),(1,-1,1,-1)\}$ es una base de $\mathbb{R}^4$.

b) Hallar las coordenadas de $(2,3,4)$ en la base $\mathcal{B}_1$ y de $(4,3,2,1)$ en la base $\mathcal{B}_2$.

c) Determinar las componentes de un cierto $u \in \mathbb{R}^4$ sabiendo que el vector de coordenadas de u en la base $\mathcal{B}_2$ es $(4,3,2,1)$.

d) Probar que

$$\mathcal{B} = \left\{E^{11} + E^{12}, E^{12}, E^{21} + E^{22}, E^{22} + E^{11}\right\}$$

es una base de $\mathbb{R}^{2\times 2}$ y determinar las coordenadas de $\begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}$ en la base $\mathcal{B}$.

5. Hallar una base de $\mathbb{R}^4$ que contenga a una base del subespacio de soluciones del sistema lineal

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

6. Sea $\mathcal{F}$ una familia linealmente independiente de vectores de V que no es subfamilia propia de ninguna familia linealmente independiente de elementos de V. Probar que $\mathcal{F}$ es una base de V.

7. Sean S y T subespacios de V y sea $\{u_1, \ldots, u_r\}$ una base de $S \cap T$.

a) Si $\{v_{r+1}, \ldots, v_p\}$ y $\{w_{r+1}, \ldots, w_q\}$ son familias de vectores de S y T, respectivamente, tales que $\{u_1, \ldots, u_r, v_{r+1}, \ldots, v_p\}$ es una base de S y $\{u_1, \ldots, u_r, w_{r+1}, \ldots, w_q\}$ es una base de T, probar que

$$\{u_1, \ldots, u_r, v_{r+1}, \ldots, v_p, w_{r+1}, \ldots, w_q\}$$

es una base de $S + T$.

b) Deducir que $dim(S+T) = dim\, S + dim\, T - dim(S \cap T)$.

c) Concluir que $dim(S \oplus T) = dim\, S + dim\, T$ si la suma es directa.

8. Sean S y T subespacios de V tales que la suma de S y T es directa. Probar que

$$V = S \oplus T \Leftrightarrow dim\, S + dim\, T = dim\, V\,.$$

9. Si S es un subespacio de V, probar que existe un subespacio T de V tal que $V = S \oplus T$ (un tal T se dice un *complemento directo* de S).

10. Si W es un $\mathbb{C}$-espacio vectorial finitamente generado, probar que W también es finitamente generado como $\mathbb{R}$-espacio vectorial y vale que

$$dim_{\mathbb{R}} W = 2\, dim_{\mathbb{C}} W\,.$$

11. Verificar la validez de las siguientes descomposiciones en suma directa (los S_i y T_1 son como en 12.1.4 y W es el subespacio de la parte c) del ejercicio 3):

    a) $K[X] = gen(1) \oplus \{f \in K[X] : f(1) = 0\}$.

    b) $K^{m\times n} = S_1 \oplus S_2$.

    c) $K^{n\times n} = T_1 \oplus W$.

12. Si U es un espacio vectorial, una familia $\mathcal{F}$ (no necesariamente finita) de vectores de U se dice una base de U si y solo si toda subfamilia finita de F es linealmente independiente y todo elemento de U puede expresarse como combinación lineal de alguna subfamilia finita de $\mathcal{F}$.

    a) Probar que $\left\{X^k : k \in \mathbb{N}_0\right\}$ es una base de $K[X]$.

    b) Sea $K^{(\mathbb{N})} = \left\{ f \in K^{\mathbb{N}} : \exists\, s \in \mathbb{N}_0 \text{ tal que } f(k) = 0\ \forall\, k > s \right\}$

        i) Verificar que $K^{(\mathbb{N})}$ es un subespacio de $K^{\mathbb{N}}$.

        ii) Sea para cada $r \in \mathbb{N}$ la función $f_r \in K^{(\mathbb{N})}$ definida por la fórmula $f_r(k) = \delta_{kr}$. Probar que $\{ f_r : r \in \mathbb{N} \}$ es una base de $K^{(\mathbb{N})}$ sobre K.

Capítulo 13

Transformaciones lineales y matrices

13.1. Transformaciones lineales

13.1.1. Definiciones

A lo largo del capítulo, la letra K denotará un cuerpo arbitrario y V y W designarán espacios vectoriales sobre K.

Hemos visto en el capítulo previo que, fijada una base $\mathcal{B} = \{v_1, \ldots, v_n\}$ de V, queda determinada una biyección θ de V en K^n, que asigna a cada elemento de V su vector de coordenadas en la base $\mathcal{B}$. Esto es, θ está definida por la fórmula

$$\theta\left(\sum_{i=1}^{n} \lambda_i v_i\right) = (\lambda_1, \lambda_2, \ldots, \lambda_n).$$

Obviamente, su inversa viene dada por

$$\theta^{-1}(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} x_i v_i.$$

En esa oportunidad comentamos (algo vagamente) que tal correspondencia identificaba algebraicamente ambos espacios. Para darle un carácter más preciso a nuestra afirmación, observemos que si u y w son elementos de V, que se escriben en la base $\mathcal{B}$ en las formas $u = \alpha_1 v_1 + \cdots + \alpha_n v_n$ y $w = \beta_1 v_1 + \cdots + \beta_n v_n$, y λ es un escalar, resulta que

$$\theta(u + w) = \theta\left(\sum_i (\alpha_i + \beta_i) v_i\right) = (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n) =$$
$$= (\alpha_1, \ldots, \alpha_n) + (\beta_1, \ldots, \beta_n) = \theta(u) + \theta(w)$$

y

$$\theta(\lambda u) = \theta\left(\sum_i \lambda\alpha_i v_i\right) = (\lambda\alpha_1, \ldots, \lambda\alpha_n) =$$
$$= \lambda(\alpha_1, \ldots, \alpha_n) = \lambda\,\theta(u)\,,$$

vale decir, θ *preserva* las operaciones de espacio vectorial.

Nos dedicaremos en esta sección a estudiar este tipo de funciones, que llamaremos homomorfismos o transformaciones lineales. Su definición precisa es la siguiente:

> Denominaremos *homomorfismo* o *transformación lineal* de V en W a toda función $f : V \to W$ que satisfaga las dos siguientes condiciones ($u, w \in V$ y $\lambda \in K$):

$TL_1)$ $f(u+w) = f(u) + f(w)$

$TL_2)$ $f(\lambda u) = \lambda f(u)\,.$

Designaremos por $\mathrm{Hom}_K(V, W)$ el conjunto de tales aplicaciones, a las que en ocasiones llamaremos familiarmente *morfismos* de V en W.

Ejemplos 13.1.1 Encargamos al lector la tarea de verificar que las siguientes aplicaciones son transformaciones lineales:

1) $f : V \to W$ definida por $f(v) = 0$ para todo $v \in V$. Se llama la transformación lineal nula y la notaremos $0_{V,W}$.

2) La función identidad $I_V : V \to V$.

3) La aplicación $f : K^2 \to K^2$ definida por $f(x, y) = (ax + by, cx + dy)$, donde a, b, c y d son elementos cualesquiera de K.

4) Más generalmente, si $A = (a_{ij}) \in K^{m\times n}$, la aplicación $t_A : K^n \to K^m$ definida por

$$t_A((x_1, x_2, \ldots, x_n)) = \left(\sum_{k=1}^{n} a_{1k}x_k\,,\ \sum_{k=1}^{n} a_{2k}x_k\,,\ \ldots,\ \sum_{k=1}^{n} a_{mk}x_k\right).$$

   Luego veremos que todo homomorfismo de K^n en K^m es de este tipo.

5) La función de especialización $\mu : K[X] \to K[X]$ dada por $\mu(g) = g(P)$, donde $P \in K[X]$. $\quad\Diamond$

Veamos algunas propiedades básicas de las transformaciones lineales.

Proposición 13.1.2 Sea f un homomorfismo de V en W. Entonces:

1) $f(0) = 0$.

2) $f(-v) = -f(v)$ para todo $v \in V$. Deducimos de esta propiedad que

$$f(v_1 - v_2) = f(v_1) - f(v_2)$$

para todo par de vectores v_1 y v_2 de V.

3) Más generalmente,

$$f\left(\sum_{i=1}^{r} \lambda_i v_i\right) = \sum_{i=1}^{r} \lambda_i f(v_i)$$

para toda familia $v_1, \ldots, v_r$ de elementos de V y toda familia $\lambda_1, \ldots, \lambda_r$ de escalares.

4) La composición de morfismos es un morfismo, esto es, si U es un K-espacio vectorial y $g \in \mathrm{Hom}_K(W, U)$ entonces $g \circ f \in \mathrm{Hom}_K(V, U)$.

5) El conjunto

$$\mathrm{Nu}(f) = \{v \in V : f(v) = 0\}$$

es un subespacio de V, llamado *núcleo* de f, mientras que $\mathrm{Im}(f)$ es un subespacio de W. Como ya veremos, estos conjuntos brindan información valiosa acerca de la función f.

6) $\mathrm{Hom}_K(V, W)$ admite estructura de $K-$espacio vectorial, definiendo la suma y el producto por escalares por las fórmulas

$$(g + h)(v) = g(v) + h(v) \;\text{ y }\; (\lambda g)(v) = \lambda g(v)\,,$$

donde $g, h \in \mathrm{Hom}_(V, W)$, $\lambda \in K$ y $v \in V$.

Demostración Para demostrar 1) basta aplicar f a la igualdad $0 = 0 + 0$, ya que entonces

$$f(0) = f(0 + 0) = f(0) + f(0)$$

y por lo tanto

$$0 = f(0) - f(0) = f(0).$$

El ítem 2) es consecuencia de 1), ya que

$$0 = f(0) = f(v + (-v)) = f(v) + f(-v).$$

En cuanto a la segunda afirmación, tenemos que

$$f(v_1 - v_2) = f(v_1 + (-v_2)) = f(v_1) + f(-v_2) = f(v_1) - f(v_2).$$

La propiedad 3) se prueba sin dificultad por inducción en r, mientras que 4) sigue inmediatamente de la definición. Para probar 5), notemos de entrada

que $0 \in \mathrm{Nu}(f)$, por (1). Suponiendo luego que $v, v' \in \mathrm{Nu}(f)$ y $\alpha \in K$, tenemos que $f(v+v') = f(v)+f(v') = 0+0 = 0$ y $f(\alpha v) = \alpha f(v) = \alpha 0 = 0$. Por lo tanto $v + v'$ y αv pertenecen al núcleo de f y éste es un subespacio de V. Encargamos al lector la tarea (similar) de probar que $\mathrm{Im}(f)$ es un subespacio de W.

En cuanto a 6), una simple (y algo tediosa) aplicación de las definiciones permite demostrar sin problemas que las operaciones definidas en $\mathrm{Hom}_K(V, W)$ satisfacen los axiomas de espacio vectorial, tarea que también encomendamos al lector. $\Diamond$

Ejemplo 13.1.3 Si $A = (a_{ij}) \in K^{m\times n}$ y t_A es la transformación lineal del ejemplo 4) de 13.1.1, es inmediato verificar que el núcleo de t_A es el subespacio de soluciones del sistema lineal homogéneo asociado a la matriz A, esto es, el conjunto de n-uplas $(x_1, x_2, \ldots, x_n)$ en K^n que satisfacen las m ecuaciones

$$\begin{array}{ccccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & 0 \\
a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & 0 \\
\ldots & & \ldots & & \ldots & & \ldots & & \ldots \\
\ldots & & \ldots & & \ldots & & \ldots & & \ldots \\
a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & 0
\end{array}$$

Ya veremos cómo la teoría que iremos desarrollando alrededor de las transformaciones lineales y la conexión entre homomorfismos y sistemas lineales que acabamos de mostrar nos permitirá describir perfectamente las soluciones de dichos sistemas $\Diamond$

Núcleo e Imagen.

Se emplean términos especiales para designar ciertas características de una transformación lineal $f : V \to W$. Así, si f es inyectiva (resp. suryectiva) diremos que f es un *monomorfismo* (resp. un *epimorfismo*). Si f es biyectiva diremos que es un *isomorfismo*. Por ejemplo, la función θ definida al principio del capítulo es un isomorfismo de $\mathbb{V}$ en K^n. En general, si existe un isomorfismo de V en W diremos que los espacios V y W son *isomorfos*, y escribiremos $V \cong W$.

Si $W = V$ se dice que f es un *endomorfismo*, empleándose la notación $\mathrm{End}_K(V)$ para referirse al conjunto de endomorfismos de V. Finalmente, un endomorfismo biyectivo de V se dirá un *automorfismo* de V. Por ejemplo, es evidente que la función identidad I_V es un automorfismo de V.

La siguiente proposición comenzará a revelar el papel preponderante que juegan el núcleo y la imagen de una transformación lineal.

Proposición 13.1.4 Sea $f \in \mathrm{Hom}_K(V, W)$. Entonces f es un monomorfismo si y solo si $\mathrm{Nu}(f) = (0)$ y f es un epimorfismo si y solo si $\mathrm{Im}(f) = W$.

DEMOSTRACIÓN Sólo probaremos la primera equivalencia, ya que la segunda es obvia por definión de epimorfismo.

Puesto que como vimos una transformación lineal aplica el vector nulo de V en el vector nulo de W, es claro que $(0) \subseteq \mathrm{Nu}(f)$ cualquiera sea f. Si f es un monomorfismo, consideremos un elemento cualquiera $v \in \mathrm{Nu}(f)$. Resulta entonces $f(v) = 0 = f(0)$, y siendo f inyectiva concluimos que $v = 0$. Luego $\mathrm{Nu}(f) = (0)$.

Recíprocamente, asumamos que el núcleo de f es trivial y sean $u, v \in V$ tales que $f(u) = f(v)$. Entonces:

$$f(u - v) = f(u) - f(v) = 0,$$

esto es, $u - v \in \mathrm{Nu}(f)$, de donde sigue por hipótesis que $u - v = 0$, ó equivalentemente, $u = v$. Por lo tanto f es inyectiva. $\Diamond$

13.1.2. Teorema de la dimensión

En lo que sigue supondremos, a menos que se indique lo contrario, que todos los espacios vectoriales mencionados son finitamente generados. Estudiaremos en primer término el comportamiento de una transformación lineal con respecto a las nociones de generación y dependencia e independencia lineal.

Proposición 13.1.5 Sea $f \in \mathrm{Hom}_K(V, W)$, sea $\mathcal{F} = \{v_1, \ldots, v_n\}$ una familia de vectores de V y sea $\mathcal{G} = \{f(v_1), \ldots, f(v_n)\}$. Valen entonces las siguientes propiedades:

1) Si $V = gen(v_1, \ldots, v_n)$ entonces $\mathrm{Im}(f) = gen(f(v_1), \ldots, f(v_n))$. Resulta por lo tanto que un epimorfismo de V en W aplica un sistema de generadores de V en un sistema de generadores de W, de donde se deduce en particular que si existe un tal epimorfismo entonces $dim\, V \geq dim\, W$.

2) Si $\mathcal{F}$ es linealmente dependiente entonces $\mathcal{G}$ es linealmente dependiente. Equivalentemente, si $\mathcal{G}$ es linealmente independiente entonces $\mathcal{F}$ es linealmente independiente.

3) Si $\mathcal{F}$ es linealmente independiente y f es un monomorfismo entonces $\mathcal{G}$ es linealmente independiente. En consecuencia, si existe un monoformismo de V en W entonces $dim\, V \leq dim\, W$.

4) Si $\mathcal{F}$ es una base de V y f es un isomorfismo entonces $\mathcal{G}$ es una base de W. Luego, una condición necesaria para que dos espacios vectoriales sean isomorfos es que tengan la misma dimensión.

Demostracion El primer enunciado es consecuencia directa de la linealidad de f, ya que si $w = f(v)$ es un elemento genérico de $\mathrm{Im}(f)$ y $\lambda_1, \ldots, \lambda_n$ son

escalares tales que $v = \sum_{i=1}^{n} \lambda_i v_i$, concluimos que

$$w = f(v) = f\left(\sum_{i=1}^{n} \lambda_i v_i\right) = \sum_{i=1}^{n} \lambda_i f(v_i),$$

esto es, la familia $\mathcal{G}$ genera $\mathrm{Im}(f)$.

También la propiedad 2) se deduce fácilmente de la linealidad de f, ya que en general

$$0 = \sum_{i=1}^{n} \lambda_i v_i \Rightarrow 0 = f(0) = f\left(\sum_{i=1}^{n} \lambda_i v_i\right) = \sum_{i=1}^{n} \lambda_i f(v_i),$$

lo que nos muestra que una combinación lineal nula no trivial de la familia $\mathcal{F}$ induce una combinación lineal nula no trivial de la familia $\mathcal{G}$.

Supongamos ahora que la familia $\mathcal{F}$ es linealmente independiente y que además f es un monomorfismo, y consideremos cualquier combinación lineal nula de la familia $\mathcal{G}$, digamos $\sum_{i=1}^{n} \alpha_i f(v_i) = 0$. Luego,

$$f\left(\sum_{i=1}^{n} \alpha_i v_i\right) = \sum_{i=1}^{n} \alpha_i f(v_i) = 0,$$

lo que implica que $\sum_{i=1}^{n} \alpha_i v_i = 0$, por la caracterización de monomorfismo brindada en la proposición anterior. Por lo tanto $\alpha_i = 0$ para todo i y la familia $\mathcal{G}$ es linealmente independiente.

Finalmente, si f es un isomorfismo y $\mathcal{F}$ es una base de V, resulta por 3) que la familia $\mathcal{G}$ es linealmente independiente y por 1) que sus elementos generan W. Luego $\mathcal{G}$ es una base de W y queda demostrada la propiedad 4). $\Diamond$

El siguiente teorema brinda una relevante fórmula, que relaciona las dimensiones del núcleo y la imagen de un homomorfismo.

Teorema 13.1.6 (de la dimensión) Si $f \in \mathrm{Hom}_K(V, W)$, vale la fórmula

$$dim\, V = dim\, \mathrm{Nu}(f) + dim\, \mathrm{Im}(f)\,.$$

DEMOSTRACION El resultado es obvio si $f = 0_{V,W}$ y también sigue fácilmente si $\mathrm{Nu}(f) = (0)$, ya que entonces f es un monomorfismo e induce por correstricción un isomorfismo de V en $\mathrm{Im}(f)$, resultando por 4) de la proposición 13.1.5 que

$$dim\, V = dim\, \mathrm{Im}(f) = dim\, \mathrm{Nu}(f) + dim\, \mathrm{Im}(f)\,.$$

Supongamos entonces que $dim\, \mathrm{Nu}(f) = r < n = dim\, V$, donde r y n son números naturales. Para probar la fórmula, tomemos una base $\{v_1, \ldots, v_r\}$

de $\mathrm{Nu}(f)$ y extendámosla a una base $\{v_1, \ldots, v_r, v_{r+1}, \ldots, v_n\}$ de V. El teorema quedará demostrado si probamos que $\mathcal{B} = \{f(v_{r+1}), \ldots, f(v_n)\}$ es una base de $\mathrm{Im}(f)$, ya que resultará entonces que

$$dim\,\mathrm{Nu}(f) + dim\,\mathrm{Im}(f) = r + \#(\mathcal{B}) = r + (n-r) = n = dim\,V\,.$$

Es muy sencillo probar que los elementos de $\mathcal{B}$ generan $\mathrm{Im}(f)$, ya que sigue por 1) de la proposición 13.1.5 que

$$\begin{aligned}\mathrm{Im}(f) &= gen\,(f(v_1), \ldots, f(v_n)) = gen\,(0, \ldots, 0, f(v_{r+1}), \ldots, f(v_n)) = \\ &= gen\,(f(v_r), f(v_{r+1}), \ldots, f(v_n))\,.\end{aligned}$$

Respecto a la independencia lineal, consideremos una combinación lineal nula de los elementos de $\mathcal{B}$ con escalares $\alpha_{r+1}, \ldots, \alpha_n$. Tenemos entonces que

$$f\left(\sum_{r+1}^{n} \alpha_i v_i\right) = \sum_{r+1}^{n} \alpha_i f(v_i) = 0,$$

esto es, $\sum_{i=r+1}^{n} \alpha_i v_i \in \mathrm{Nu}(f)$, y por lo tanto existen escalares $\beta_1, \ldots, \beta_r$ tales que

$$\sum_{r+1}^{n} \alpha_i v_i = \sum_{i=1}^{r} \beta_i v_i,$$

ó equivalentemente

$$\sum_{i=1}^{r} (-\beta_i) v_i + \sum_{r+1}^{n} \alpha_i v_i = 0\,.$$

Puesto que $\{v_1, \ldots, v_n\}$ es una base de V esta última combinación lineal debe ser trivial, de donde sigue en particular que $\alpha_i = 0$ para todo i. Luego la familia $\mathcal{B}$ es linealmente independiente y por lo tanto es una base de $\mathrm{Im}(f)$. $\quad \Diamond$

Ejemplo 13.1.7 Caractericemos el núcleo y la imagen del endomorfismo f de $\mathbb{R}^3$ definido por

$$f\,((x_1, x_2, x_3)) = (x_1 - x_2 + x_3, 2x_1 + x_2 - 2x_3, 3x_1 - x_3)\,.$$

Como hemos comentado en el ejemplo 13.1.3, el núcleo de f es el subespacio de soluciones del sistema lineal

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

Si bien más adelante brindaremos un método general de resolución de sistemas lineales, podemos resolver este caso particular mediante cálculos muy sencillos. Observemos en primer término que la segunda ecuación se

obtiene restando la primera de la tercera, por lo que resulta redundante y podemos olvidarnos de ella. Vemos por otra parte que debe verificarse la relación $x_3 = 3x_1$, de la cual deducimos, reemplazando en la primera, que también debe verificarse la igualdad $x_2 = 4x_1$. Luego, podemos expresar todas las incógnitas en función de x_1, resultando que cualquier solución es de la forma $(a, 4a, 3a) = a(1, 4, 3)$, donde a es un número real. Puesto que, recíprocamente, es inmediato verificar que $(a, 4a, 3a)$ es solución del sistema cualquiera sea a, concluimos que el núcleo de f es el subespacio de dimensión uno generado por el vector $(1, 4, 3)$.

Puesto que entonces $dim\,\mathrm{Im}(f) = 2$ (teorema de la dimensión), para hallar una base de la imagen bastará exhibir dos vectores linealmente independientes en ella. Tomando por ejemplo $(1, 2, 3) = f(e_1)$ y $(-1, 1, 0) = f(e_2)$, sigue que los elementos de la imagen de f son de la forma

$$(\alpha - \beta, 2\alpha + \beta, 3\alpha),$$

donde α y β varían libremente sobre $\mathbb{R}$. $\Diamond$

El teorema de la dimensión admite el siguiente corolario, que facilita la tarea de comprobar si una transformación lineal entre espacios de igual dimensión es un isomorfismo:

Corolario 13.1.8 Si $f \in \mathrm{Hom}_K(V, W)$ y $dim\,V = dim\,W$, las siguientes afirmaciones son equivalentes:

(a) f es un monomorfismo.

(b) f es un epimorfismo.

(c) f es un isomorfismo.

Demostración Es consecuencia inmediata del teorema de la dimensión, por lo que dejamos los detalles a cargo del lector. Dada su importancia, hagamos notar que la situación se aplica al caso de endomorfismos de un espacio vectorial de dimensión finita. $\Diamond$

Ejemplos 13.1.9 Como ilustración del corolario, probemos que cualquier endomorfismo φ de V que admita un inverso a derecha es un automorfismo (un inverso a derecha de φ es una aplicación $\sigma \in \mathrm{End}\,V$ tal que $\varphi \circ \sigma = I_V$.

Dado $v \in V$, tenemos por hipótesis que

$$v = I_V(v) = (\varphi \circ \sigma)(v) = \varphi(\sigma(v)),$$

y por lo tanto φ es un epimorfismo. Luego también es un automorfismo, por corolario 13.1.8.

Vale la pena destacar que el resultado es falso si V no es de dimensión finita. Por ejemplo, sea $V = \mathbb{R}[X]$ y sea ∂ el operador de derivación, que claramente es un endomorfismo de V. Si definimos $\sigma : V \to V$ por la fórmula

$$\sigma(f) = \sigma\left(\sum_{i=0}^{m} a_i X^i\right) = \sum_{i=0}^{m} a_i (i+1)^{-1} X^{i+1},$$

es inmediato verificar que σ es un endormorfismo de V, resultando que

$$\theta\left(\sigma(f)\right) = \partial\left(\sum_{i=0}^{m} a_i (i+1)^{-1} X^{i+1}\right) = \sum_{i=0}^{m} a_i (i+1)^{-1} \partial\left(X^{i+1}\right) =$$
$$= \sum_{i=0}^{m} a_i (i+1)^{-1} (i+1) X^i = f\,,$$

vale decir, ∂ verifica las condiciones de nuestro ejemplo. Sin embargo no es un automorfismo de $\mathbb{R}[X]$, ya que su núcleo es el subespacio de polinomios constantes y por consiguiente no es un monomorfismo. $\Diamond$

NOTA La cuestión de la existencia de un isomorfismo entre dos espacios vectoriales tiene especial relevancia, ya que desde el punto de vista algebraico dos espacios vectoriales isomorfos son esencialmente iguales, en el sentido de que tienen exactamente las mismas propiedades y sólo se diferencian en los símbolos que se usan para representar sus elementos y operaciones.

Ahora bien, mostraremos en el desarrollo de la teoría que son válidas las recíprocas de las conclusiones obtenidas en los incisos 1), 3) y 4) de la proposición 13.1.5, acerca de la existencia de monomorfismos, epimorfismos e isomorfismos, de donde resultará en particular que dos espacios vectoriales son isomorfos si y solo si tienen la misma dimensión. Inferiremos como consecuencia de ello que todos los K-espacios vectoriales de dimensión n son esencialmente iguales a K^n, o dicho de otra forma, el espacio de n-uplas con coeficientes en K es el modelo algebraico de todos los K-espacios vectoriales de dimensión n. $\Diamond$

13.1.3. El carácter libre de una base

Estableceremos a continuación otro resultado fundamental: para definir una transformación lineal basta elegir *libremente* las imágenes de los elementos de cualquier base del dominio.

Teorema 13.1.10 Sea $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base de V y sean $w_1, \ldots, w_n$ vectores cualesquiera de W. Existe entonces una *única* transformación lineal $f : V \to W$ tal que $f(v_i) = w_i$ para $i = 1, \ldots, n$.

Se deduce de este hecho que dos homomorfismos entre espacios vectoriales son iguales si y solo si coinciden sobre los elementos de cualquier base del dominio.

DEMOSTRACION Comenzando por la unicidad, supongamos que f y g son morfismos de V en W que satisfacen los requerimientos del enunciado. Si $v \in V$ y $v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$, resulta entonces que

$$g(v) = g\left(\sum_{i=1}^{n} \lambda_i v_i\right) = \sum_{i=1}^{n} \lambda_i g(v_i) = \sum_{i=1}^{n} \lambda_i f(v_i) = f\left(\sum_{i=1}^{n} \lambda_i v_i\right) = f(v),$$

esto es, f y g coinciden en todos los elementos de V y por lo tanto $f = g$. Obsérvese que hemos probado que si dos transformaciones lineales coinciden sobre una base del espacio entonces coinciden en todo el espacio, como consecuencia directa de la linealidad de dichas funciones. Apreciemos también que solo hemos usado que $\mathcal{B}$ es una familia de generadores de V.

Respecto a la existencia de una tal f, tomemos v como antes y definamos

$$f(v) = \sum_{i=1}^{n} \lambda_i w_i \,.$$

Notemos que la buena definición de f está asegurada por el hecho de que todo vector de V se exprese en *forma única* como combinación lineal de los elementos de $\mathcal{B}$.

Para probar que f es lineal, tomemos cualquier par de vectores u y u' de V, de coordenadas $\alpha_1, \ldots, \alpha_n$ y $\alpha'_1, \ldots, \alpha'_n$ en la base $\mathcal{B}$, respectivamente, y sea β un escalar. Entonces

$$f(u+u') = f\left(\sum_{i=1}^{n} (\alpha_i + \alpha'_i) v_i\right) = \sum_{i=1}^{n} (\alpha_i + \alpha'_i) w_i =$$
$$= \sum_{i=1}^{n} \alpha_i w_i + \sum_{i=1}^{n} \alpha'_i w_i = f(u) + f(u')$$

y

$$f(\beta u) = f\left(\sum_{i=1}^{n} \beta \alpha_i v_i\right) = \sum_{i=1}^{n} \beta \alpha_i w_i =$$
$$= \beta \sum_{i=1}^{n} \alpha_i w_i = \beta f(u),$$

lo que prueba que f es un homomorfismo.

Finalmente, es muy fácil probar que f satisface las condiciones del enunciado, ya que

$$f(v_k) = f(\sum_{i=1}^{n} \delta_{ik} v_i) = \sum_{i=1}^{n} \delta_{ik} w_i = w_k$$

para todo $1 \leq k \leq n$. $\quad \Diamond$

Ejemplos 13.1.11 El carácter *libre* que poseen los elementos de cualquier base de un espacio vectorial, reflejado en el teorema anterior, es de gran importancia y utilidad. Antes de exhibir ciertas consecuencias del mismo, lo aplicaremos en algunos ejemplos:

1) Hallemos la expresión general de un endomorfismo f de $\mathbb{R}^3$ tal que $v_1 = (1,-1,1) \in \mathrm{Nu}(f) \cap \mathrm{Im}(f)$ y $v_2 = (1,-1,0) \in \mathrm{Nu}(f)$.

   Suponiendo que un tal f existe, observemos que el núcleo de f debe contener a la familia linealmente independiente $\{v_1, v_2\}$, por lo que $dim\,\mathrm{Nu}(f) \geq 2$. Puesto que f no es nula, ya que $v_1 \in \mathrm{Im}(f)$, sigue por teorema de la dimensión que $\{v_1, v_2\}$ debe ser una base de $\mathrm{Nu}(f)$ y $\{(1,-1,1)\}$ una base de $\mathrm{Im}(f)$. Si extendemos la familia $\{v_1, v_2\}$ a una base $\mathcal{B} = \{v_1, v_2, v_3\}$ de $\mathbb{R}^3$, tomando por ejemplo $v_3 = (1,0,0)$, vemos que la transformación lineal buscada queda definida sobre $\mathcal{B}$ por las relaciones $f(v_1) = f(v_2) = (0,0,0)$ y $f(v_3) = v_1$, que claramente satisface las condiciones requeridas.

   Para hallar la expresión general de f, y siguiendo la demostración del teorema anterior, debemos determinar las coordenadas en la base $\mathcal{B}$ de cualquier vector genérico (x_1, x_2, x_3). Un sencillo cálculo nos muestra que

$$(x_1, x_2, x_3) = x_3 v_1 + (-x_2 - x_3)v_2 + (x_1 + x_2)v_3,$$

   y por lo tanto

$$\begin{aligned} f\left((x_1, x_2, x_3)\right) &= x_3 f(v_1) + (-x_2 - x_3) f(v_2) + (x_1 + x_2) f(v_3) = \\ &= (x_1 + x_2)\,(1,-1,1) = (x_1 + x_2, -x_1 - x_2, x_1 + x_2)\,. \end{aligned}$$

2) Si $dim\, V = m$ y $dim\, W = n$ entonces $dim\,\mathrm{Hom}_K(V,W) = m\,n$.

   Sean $\mathcal{B} = \{v_1, \ldots, v_m\}$ y $\{w_1, \ldots, w_n\}$ bases de V y W, respectivamente, y consideremos, para cada par $(i,j) \in I_m \times I_n$ el morfismo θ_{ij} de V en W definido sobre la base $\mathcal{B}$ en la forma

$$\theta_{ij}(v_k) = \delta_{ik} w_j\,,$$

   es decir, θ_{ij} se anula sobre v_k si $k \neq i$ y $\theta_{ij}(v_i) = w_j$. Probaremos que la familia $\{\,\theta_{ij} : (i,j) \in I_m \times I_n\}$ es una base de $\mathrm{Hom}_K(V,W)$, lo que demostrará nuestra afirmación.

   Respecto a la generación, sea $g \in \mathrm{Hom}_K(V,W)$ y supongamos que

$$g(v_k) = \sum_{j=1}^{n} a_{kj} w_j \quad \text{para } k = 1, 2 \ldots, n\,,$$

   donde los a_{kj} son escalares. Si h es el homomorfismo de V en W definido por

$$h = \sum_{(i,j)} a_{ij}\,\theta_{ij}\,,$$

donde (i,j) recorre $I_m \times I_n$, tenemos que

$$h(v_k) = \left(\sum_{(i,j)} a_{ij}\,\theta_{ij}\right)(v_k) = \sum_{(i,j)} a_{ij}\,\theta_{ij}(v_k) = \sum_{(i,j)} a_{ij}\,\delta_{ik} w_j = \\ = \sum_{j=1}^{n}\left(\sum_{i=1}^{m} a_{ij}\,\delta_{ik} w_j\right) = \sum_{j=1}^{n} a_{kj}\,w_j = g(v_k)$$

para todo índice k. Puesto que g y h coinciden sobre una base sigue por el teorema 13.1.10 que $g = h$, y por lo tanto las transformaciones θ_{ij} generan $\mathrm{Hom}_K(V,W)$.

De forma similar se prueba la independencia lineal, ya que suponiendo que

$$\sum_{(i,j)} b_{ij}\,\theta_{ij} = 0_{V,W},$$

resulta para todo índice k que

$$0 = 0_{V,W}(v_k) = \sum_{(i,j)} b_{ij}\,\theta_{ij}(v_k) = \sum_{(i,j)} b_{ij}\,\delta_{ik} w_j = \\ = \sum_{j=1}^{n}\left(\sum_{i=1}^{m} b_{ij}\,\delta_{ik} w_j\right) = \sum_{j=1}^{n} b_{kj}\,w_j\,.$$

Puesto que $\{w_1,\ldots,w_n\}$ es una base de W, sigue que $b_{kj} = 0$ para todo j. Luego la familia $\{\theta_{ij}\}$ es linealmente independiente y resulta ser una base de $\mathrm{Hom}_K(V,W)$. $\quad\Diamond$

El siguiente corolario del teorema 13.1.10 caracteriza en términos de sus dimensiones la existencia de monomorfismos, epimorfismos o isomorfismos entre dos espacios vectoriales.

Corolario 13.1.12 Sea $dim\,V = m$ y $dim\,W = n$. Entonces:

1) Existe un monomorfismo de V en W si y solo si $m \leq n$.

2) Existe un epimorfismo de V en W si y solo si $m \geq n$.

3) V y W son isomorfos si y solo si $m = n$.

DEMOSTRACION Ya hemos probado en la proposición 13.1.5 que las condiciones sobre las dimensiones son necesarias, por lo que resta probar que también son suficientes. Tomemos para ello cualquier par de bases $\{v_1,\ldots,v_m\}$ y $\{w_1,\ldots,w_n\}$ de V y W, respectivamente.

Si $m \leq n$, consideremos el homomorfismo $f : V \to W$ definido por $f(v_i) = w_i$ para todo $i \leq m$. En tal caso

$$\mathrm{Im}(f) = gen\,(f(v_1),\ldots,f(v_m)) = gen\,(w_1,\ldots,w_m)\,,$$

y por lo tanto $\{w_1, \ldots, w_m\}$ es una base de $\mathrm{Im}(f)$. Luego

$$dim\,\mathrm{Nu}(f) = dim\,V - dim\,\mathrm{Im}(f) = m - m = 0$$

y f es un monomorfismo.

La prueba en el caso de 2) es aún más sencilla. En efecto, suponiendo $m \geq n$ resulta que la transformación lineal $g : V \to W$ definida por

$$g(v_i) = \begin{cases} w_i \text{ si } 1 \leq i \leq n \\ 0 \text{ si } n < i \leq m \end{cases}$$

es un epimorfismo, ya que la imagen de g contiene una base de W.

Finalmente, si $m = n$ existe por la propiedad 1) un monomorfismo h de V en W, que de acuerdo con el corolario 13.1.8 es también un epimorfismo. Luego h es un isomorfismo y el enunciado queda probado. $\Diamond$

13.1.4. Ejercicios

1. Demostrar que $\mathrm{End}_K(V)$ es un anillo, definiendo el producto como la composición de funciones.

2. Sean V y W espacios vectoriales sobre $\mathbb{Q}$ y sea $f : V \to W$ una función tal que $f(v + v') = f(v) + f(v')$ cualesquiera sean v y v' en V. Probar que $f \in \mathrm{Hom}_{\mathbb{Q}}(V, W)$.

3. Sean S y T subespacios de V y W, respectivamente, y sea f un $K-$homomorfismo de V en W. Probar que $f(S)$ es un subespacio de W y $f^{-1}(T)$ es un subespacio de V. Observar los casos $S = V$ y $T = (0)$.

4. En cada uno de los siguientes casos se define una aplicación f entre dos espacios vectoriales V y W. Determinar cuáles de ellas son transformaciones lineales.

   a) $V = \mathbb{R}^2$, $W = \mathbb{R}^3$; $f((x_1, x_2)) = (x_1 - x_2, x_1, x_1 + 2x_2)$
   b) $V = \mathbb{R}^3$, $W = \mathbb{R}^2$; $f((x_1, x_2, x_3)) = (x_1 - x_3, 1 + x_2)$
   c) $V = \mathbb{R}^{2\times 2}$, $W = \mathbb{R}^4$; $f((a_{ij})) = (a_{11} + a_{12}, 0, a_{21}, a_{11} + a_{22})$
   d) $V = \mathbb{R}^{2\times 2}$, $W = \mathbb{R}^2$; $f((a_{ij})) = (a_{12} - a_{11}, -a_{12} + a_{22}a_{11})$
   e) $V = \mathbb{R}^{2\times 2}$, $W = \mathbb{R}$; $f((a_{ij})) = a_{11}a_{22} - a_{12}a_{21}$
   f) $V = K^{m\times n}$, $W = K^{n\times m}$; $f(A)) = A^t$
   g) $V = K[X]$, $W = K[X]$; $f(h) = \partial h$
   h) $V = K[X]$, $W = K^2$; $f(h) = (h(0), h'(1))$
   i) $V = K[X]$, $W = K$; $f(h) = h(a)$ (a un elemento fijo de K)

j) $V = W = \mathbb{C}\,;\, f(z)) = z - 2\overline{z}$ (considerar primero a $\mathbb{C}$ como $\mathbb{R}-$espacio vectorial y luego como $\mathbb{C}-$espacio vectorial)

k) $V = C([0,1])$, $W = \mathbb{R}$; $f(g) = \int_0^1 g(x)\,dx$.

5. Caracterizar el núcleo y la imagen de los homomorfismos del ejercicio 4, determinando en cada caso sus respectivas dimensiones.

6. Respecto del ejercicio 4:

   a) Si f es monomorfismo hallar $g \in \mathrm{Hom}_K(W,V)$ tal que $g \circ f = I_V$.

   b) Si f es epimorfismo hallar $g \in \mathrm{Hom}_K(W,V)$ tal que $f \circ g = I_W$.

   c) Determinar en qué casos f es un isomorfismo.

7. Demostrar que la función inversa de una transformación lineal biyectiva también es una transformación lineal.

8. Demostrar que la relación de isomorfismo $\cong$ de espacios vectoriales es una relación de equivalencia.

9. Si $\alpha \in \mathbb{R}$, la aplicación $\rho_\alpha : \mathbb{R}^2 \to \mathbb{R}^2$ definida por

$$\rho_\alpha(x,y) = (x\cos\alpha - y \text{ sen } \alpha, x \text{ sen } \alpha + y\cos\alpha)$$

   se llama *rotación* de argumento α. Interpretar geométricamente y demostrar las siguientes propiedades:

   a) ρ_α es un automorfismo de $\mathbb{R}^2$

   b) $\rho_\alpha \circ \rho_\beta = \rho_{\alpha+\beta}$

   c) $\rho_0 = I_{\mathbb{R}^2}$

   d) $\rho_\alpha^{-1} = \rho_{-\alpha}$

   e) Dado $\alpha \in \mathbb{R}$ existe un único $\theta \in [2\pi)$ tal que $\rho_\alpha = \rho_\theta$

   f) El conjunto de rotaciones es un subgrupo abeliano de $\mathrm{Aut}(\mathbb{R}^2)$.

10. Sean U, V y W espacios vectoriales sobre K y sean $f \in \mathrm{Hom}_K(U,V)$ y $g \in \mathrm{Hom}_K(V,W)$. Probar:

    a) $g \circ f$ es un monomorfismo si f y g lo son.

    b) $g \circ f$ es un epimorfismo si f y g lo son.

    c) $g \circ f$ es un isomorfismo si f y g lo son.

    d) f es un monomorfismo si $g \circ f$ lo es.

    e) g es un epimorfismo si $g \circ f$ lo es.

11. Sean V y W espacios vectoriales sobre K y sea $f \in \mathrm{Hom}_K(V,W)$. Si g y h son automorfismos de V y W, respectivamente, probar:

    a) $\mathrm{Nu}(h \circ f) = \mathrm{Nu}(f \circ g) = \mathrm{Nu}(f)$

    b) $\mathrm{Im}(h \circ f) = \mathrm{Im}(f \circ g) = \mathrm{Im}(f)$.

12. Sea $f : V \to W$ un homomorfismo de $K-$espacios vectoriales finitamente generados que aplica familias linealmente independientes en familias linealmente independientes. Probar que f es un monomorfismo.

13. Si $f : V \to W$ es un homomorfismo de $K-$espacios vectoriales, probar la equivalencia de las siguientes afirmaciones:

    a) f es un isomorfismo.

    b) $f(\mathcal{B})$ es una base de W para toda base $\mathcal{B}$ de V.

    c) Existe una base $\mathcal{G}$ de V tal que $f(\mathcal{G})$ es una base de W.

14. En cada uno de los siguientes ítems decidir si existe un homomorfismo f que satisfaga las condiciones requeridas:

    a) $f : \mathbb{R}^3 \to \mathbb{R}^2$ tal que

        i) $(1,1)$ y $(1,2)$ pertenecen a la imagen de f

        ii) $f((1,1,1)) = f((1,1,0))$ y $f((0,1,1)) = f((1,0,0))$.

    b) $f : \mathbb{R}^3 \to \mathbb{R}^3$ tal que $(1,2,1) \in \mathrm{Nu}(f) \cap \mathrm{Im}(f)$ y $(1,2,0) \in \mathrm{Im}(f)$.

    c) Un epimorfismo $f : \mathbb{R}^4 \to \mathbb{R}^2$ tal que

    $$\{(1,1,1,1),(1,1,1,0),(1,1,0,0)\} \subseteq \mathrm{Nu}(f).$$

15. En cada uno de los siguientes ítems decidir la misma cuestión del ejercicio anterior. En caso afirmativo, determinar una fórmula general para f.

    a) $f : \mathbb{R}^2 \to \mathbb{R}^3$ tal que $f((2,1)) = (1,-1,0)$.

    b) $f : \mathbb{R}^3 \to \mathbb{R}^2$ tal que $f((2,1,-1)) = (0,1)$, $f((3,0,2)) = (2,-1)$ y $f((1,2,-4)) = (-2,3)$.

    c) $f : \mathbb{R}^4 \to \mathbb{R}^4$ tal que $\mathrm{Nu}(f) = \mathrm{Im}(f)$.

    d) $f : \mathbb{R}^3 \to \mathbb{R}^2$ tal que $f(gen((0,1,-1),(1,2,3))) = gen((1,1))$ y $(2,0) \in \mathrm{Im}(f)$.

    e) $f \in \mathrm{End}(\mathbb{R}^3)$ tal que $f \circ f = 0_{V,V}$ y $f \neq 0_{V,V}$.

16. Sea en $\mathbb{R}^4$ el subespacio $S = gen\,((1,2,-1,0),(3,1,0,1))$.

    a) Hallar $g \in \mathrm{Hom}(\mathbb{R}^4, \mathbb{R}^3)$ tal que $\mathrm{Nu}(g) = S$.

    b) Exhibir un sistema lineal homogéneo con cuatro incógnitas cuyo subespacio de soluciones sea S.

17. Sea f un endomorfismo de V tal que $f \circ f = f$ (un tal endomorfismo se dice un *proyector* de V). Probar que $V = \mathrm{Nu}(f) \oplus \mathrm{Im}(f)$.

18. Sea $f : \mathbb{R}_n[X] \to \mathbb{R}_n[X]$ definida por $f(h) = h'$. Determinar un subespacio S de $\mathbb{R}_n[X]$ tal que

$$\mathbb{R}_n[X] = (\mathrm{Nu}(f) \oplus \mathrm{Im}(f)) \oplus S\,.$$

13.2. Matrices

13.2.1. Producto de matrices

Además de la suma y el producto por escalares que definen la estructura vectorial en los espacios de matrices, existe otra importante operación matricial, estrechamente conectada con el que será el principal tema de exposición de esta sección, a saber, la identificación existente entre los espacios de homomorfismos y los espacios de matrices. La misma no es una operación binaria definida en un cierto conjunto, como muchas de las que hemos estudiado a lo largo de este libro, sino que asocia a cada par de matrices −cuyos tamaños satisfacen un requerimiento adecuado− otra matriz llamada producto de las mismas. Como todo esto es bastante vago y obviamente no queremos parecer misteriosos, pasemos a definir con precisión dicho producto:

> Si m, n y s son números naturales, dadas $A = (a_{ij}) \in K^{m\times n}$ y $B = (b_{ij}) \in K^{n\times s}$ definimos el *producto* de A y B como la matriz $AB \in K^{m\times s}$ definida por
>
> $$(AB)_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}\,.$$

Observaciones Advirtamos en la fórmula de arriba que cuando k varía entre 1 y n los a_{ik} recorren la $i-$ésima fila de A y los b_{kj} recorren la $j-$ésima columna de B (nótese que es fundamental para ello que el número de columnas de A coincida con el número de filas de B). Usando un lenguaje informal, y a efectos de visualizar la definición, podemos decir entonces que el elemento situado en la posición (i, j) del producto AB se obtiene "multiplicando"la $i-$ésima fila de A por la la $j-$ésima columna de B.

Notemos además que en la definición de producto los roles de los factores están bien diferenciados, y que en general BA no estará definido aunque AB lo esté. Precisamente, ambos productos están definidos si y solo si $A \in K^{m\times n}$ y $B \in K^{n\times m}$ para ciertos naturales m y n, en cuyo caso $AB \in K^{m\times m}$ y $BA \in K^{n\times n}$. Si $m = n$, resulta como caso particular importante que el producto de matrices es una operación binaria en $K^{n\times n}$, conjunto que por razones de comodidad notaremos más simplemente $M_n(K)$ y llamaremos espacio de matrices cuadradas de orden n con coeficientes en K $\quad\Diamond$

Ejemplos 13.2.1 Los siguientes ejemplos ilustrarán la definición (los ejemplos numéricos corresponden a matrices con coeficientes reales).

1)

$$\begin{pmatrix} 2 & 3 & -1 \\ 4 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 4 & -1 & -1 \\ 3 & 1 & 0 & 6 \\ 5 & -2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 8 & 13 & -5 & 15 \\ 11 & 17 & -4 & 2 \end{pmatrix}.$$

2) Si $n \in \mathbb{N}$, los elementos E^{ij} de la base canónica de $M_n(K)$ multiplican entre sí según la siguiente fórmula ($r, s, t, u \in \mathbb{I}_n$):

$$E^{rs}E^{tu} = \begin{cases} E^{ru} & \text{si } t = s \\ 0 & \text{en otro caso.} \end{cases}$$

En efecto, aplicando la definición de producto resulta que

$$\left(E^{rs}E^{tu}\right)_{ij} = \sum_{k=1}^{n} E^{rs}_{ik}E^{tu}_{kj} = \sum_{k=1}^{n} \delta_{ir}\delta_{ks}\delta_{kt}\delta_{ju} = \delta_{st}\delta_{ir}\delta_{ju},$$

como queríamos probar (el lector verificará la validez de las igualdades).

Sigue en particular que el producto en $M_n(K)$ no es conmutativo si $n > 1$, ya que por ejemplo $E^{11}E^{12} = E^{12}$ y $E^{12}E^{11} = 0$.

3) Si $A \in K^{m\times n}$ y $\lambda \in K$ entonces

$$\lambda A = \begin{pmatrix} \lambda & 0 & \dots & 0 \\ 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda \end{pmatrix} A\,,$$

donde el factor de la izquierda es la matriz escalar cuadrada de orden m determinada por λ, que notaremos $E_\lambda(m)$, o simplemente E_λ si m está bien determinado. El hecho es fácil de verificar, ya que al multiplicar cualquier fila i de $E_\lambda(m)$ por cualquier columna j de A todos los productos parciales se anulan excepto el $i-$ésimo, cuyo resultado es λa_{ij}. Análogamente, es fácil probar que $\lambda A = AE_\lambda(n)$.

4) En la sección anterior vimos que toda matriz $A = (a_{ij}) \in K^{m\times n}$ induce una transformación lineal $t_A : K^n \to K^m$, definida por

$$t_A\,(x)) = \left(\sum_{k=1}^{n} a_{1k}x_k\,,\ \sum_{k=1}^{n} a_{2k}x_k\,,\ \dots,\ \sum_{k=1}^{n} a_{mk}x_k\right)$$

si $x = (x_1, x_2, \dots, x_n)$.

Ahora bien, identificando en general K^r con $K^{1\times r}$ y con $K^{r\times 1}$ a través de la asignaciones (isomorfismos)

$$(z_1, z_2, \dots, z_r) \mapsto (z_1\ z_2\ \dots\ z_r) \quad \text{y} \quad (z_1, z_2, \dots, z_r) \mapsto \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_r \end{pmatrix},$$

obtenemos una forma matricial para t_A, ya que es inmediato verificar que $t_A((x))$ corresponde bajo estas identificaciones al producto matricial Ax^t, esto es,

$$t_A((x_1, x_2, \ldots, x_n)) = (y_1, y_2, \ldots, y_m) \Leftrightarrow Ax^t = y^t .$$

Más adelante apreciaremos la ventaja de esta interpretación de $t_A(x)$, ya que el uso de ciertas propiedades algebraicas del producto de matrices nos ayudará a estudiar la naturaleza del homomorfismo t_A. $\Diamond$

Probaremos en la siguiente proposición una serie de propiedades algebraicas del producto de matrices.

Proposición 13.2.2 Propiedades del producto matricial:

1) Asociatividad: $A(BC) = (AB)C$, donde $A \in K^{m\times n}$, $B \in K^{n\times r}$ y $C \in K^{r\times s}$.

2) Distributividad: $(A + A')B = AB + A'B$ y $A(B + B') = AB + AB'$, si $A, A' \in K^{m\times n}$ y $B, B' \in K^{n\times r}$.

3) Existencia de elemento neutro: si $r \in \mathbb{N}$, sea

$$I_r = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix} = E_1(r).$$

Entonces

$$A\,I_n = A = I_m\,A$$

para toda matriz $A \in K^{m\times n}$.

Tomando $m = n$ sigue en particular que I_n es el elemento neutro del producto en $M_n(K)$. La llamaremos matriz *identidad* de orden n. A menudo, si n está determinado sin ambigüedad, la notaremos simplemente I.

4) $M_n(K)$ es un anillo (no conmutativo si $n > 1$) respecto a la suma y el producto de matrices.

5) Si $A \in K^{m\times n}$ y si $B \in K^{n\times r}$ entonces $(AB)^t = B^tA^t$.

DEMOSTRACION Solo probaremos la primera y la última propiedad, dejando las restantes a cargo del lector. (obsérvese que la propiedad 3) sigue del ejemplo 3) de 13.2.1 y que la propiedad 4) es consecuencia de las tres primeras).

Respecto a 1), dado $(i,j) \in \mathbb{I}_m \times \mathbb{I}_s$ tenemos

$$\begin{aligned}(A(BC))_{ij} &= \sum_{k=1}^{n} A_{ik}(BC)_{kj} = \sum_{k=1}^{n} A_{ik} \sum_{l=1}^{r} B_{kl}C_{lj} = \sum_{k=1}^{n}\sum_{l=1}^{r} A_{ik}(B_{kl}C_{lj}) = \\ &= \sum_{k=1}^{n}\sum_{l=1}^{r}(A_{ik}B_{kl})C_{lj} = \sum_{l=1}^{r}\left(\sum_{k=1}^{n} A_{ik}B_{kl}\right)C_{lj} = \sum_{l=1}^{r}(AB)_{il}C_{lj} = \\ &= ((AB)C)_{ij}\ ,\end{aligned}$$

como queríamos probar.

Por último, para probar 5) tomemos $(i,j) \in \mathbb{I}_r \times \mathbb{I}_m$. Entonces

$$\begin{aligned}\left((AB)^t\right)_{ij} = (AB)_{ji} &= \sum_{k=1}^{n} A_{jk}B_{ki} = \sum_{k=1}^{n} B_{ki}A_{jk} = \\ &= \sum_{k=1}^{n} B^t_{ik}A^t_{kj} = \left(B^tA^t\right)_{ij}\ ,\end{aligned}$$

y por lo tanto $(AB)^t = B^tA^t$.

Notemos que a lo largo de las demostraciones hemos usado la asociatividad y la conmutatividad del producto en el cuerpo K. $\Diamond$

Matrices inversibles

Puesto que el producto de matrices en $M_n(K)$ admite elemento neutro, tiene perfecto sentido considerar la inversibilidad de una matriz cuadrada de orden n. Precisamente, introducimos la siguiente definición:

> Una matriz $A \in M_n(K)$ se dice *inversible* si y solo si existe una matriz $B \in M_n(K)$ tal que
>
> $$AB = BA = I_n.$$

Observación Una tal matriz B, si existe, es única. En efecto, suponiendo que $C \in M_n(K)$ también satisface las condiciones de la definición, resulta que

$$C = C\,I_n = C(AB) = (CA)B = I_n\,B = B.$$

Debido a este hecho B se dice la matriz *inversa* de A, y se nota A^{-1}. Designaremos por $GL(n,K)$ el conjunto de matrices cuadradas inversibles de orden n con coeficientes en K. $\Diamond$

Ejemplos 13.2.3 Más adelante dispondremos de un criterio para decidir si una matriz cuadrada es inversible. Por ahora ilustraremos la cuestión a través de algunas observaciones y ejemplos elementales (en los ejemplos genéricos supondremos que las matrices pertenecen a $M_n(K)$ para un cierto $n \in \mathbb{N}$).

1) Es claro que la matriz nula no es inversible. Más generalmente, no es inversible ninguna matriz que tenga una fila o columna nula. En cambio, E_λ es inversible para todo $\lambda \in K^*$, siendo $E_\lambda^{-1} = E_{\lambda^{-1}}$. Resulta en particular que $I_n^{-1} = I_n$.

2) $T = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}$ es inversible en $M_2(\mathbb{Q})$. En efecto, planteando la condición

$$\begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

se arriba al sistema de ecuaciones

$$\begin{cases} x + z = 1 \\ 2x + 3z = 0 \\ y + w = 0 \\ 2y + 3w = 1 \,, \end{cases}$$

cuya única solución es

$$(x, y, z, w) = (3, -1, -2, 1),$$

como el lector puede verificar muy sencillamente. Por lo tanto, la matriz

$$S = \begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix}$$

satisface la igualdad que $TS = I_2$. Puesto que es inmediato verificar que también $ST = I_2$, concluimos que T es inversible y que $T^{-1} = S$.

3) La matriz $M = \begin{pmatrix} 3 & 6 \\ 1 & 2 \end{pmatrix}$ no es inversible en $M_2(\mathbb{Q})$, pues con las notaciones del ejemplo anterior la búsqueda de la inversa de M nos conduce en este caso al sistema de ecuaciones

$$\begin{cases} 3x + 6z = 1 \\ 3y + 6w = 0 \\ x + 2z = 0 \\ 2y + 2w = 1 \,, \end{cases}$$

que no es resoluble, ya que multiplicando la tercera ecuación por 3 y restándole la primera llegamos a la contradicción $0 = -1$.

4) La inversa de una matriz inversible es inversible y el producto de matrices inversibles es inversible. En efecto, si $A \in GL(n,K)$ sigue inmediatamente por definición que $\left(A^{-1}\right)^{-1} = A$. Por otra parte, si $B \in GL(n,K)$ tenemos que

$$(B^{-1}A^{-1})AB = B^{-1}(A^{-1}A)B = B^{-1}I_nB = B^{-1}B = I_n$$

y

$$AB(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = I_n\,,$$

lo que muestra que $AB \in GL(n,K)$ y $(AB)^{-1} = B^{-1}A^{-1}$. $\quad \Diamond$

Analizando el ejemplo 2) de 13.2.3, vemos que para hallar la inversa de T sólo planteamos la condición de que la misma fuera una inversa a derecha, verificándose que la misma era también inversa a izquierda. La siguiente proposición nos mostrará que bastará proceder siempre de esa manera.

Proposición 13.2.4 Si $n \in \mathbb{N}$ y $A \in M_n(K)$, las siguientes afirmaciones son equivalentes:

a) A es inversible.

b) Si $X \in M_n(K)$ y $AX = 0$ entonces $X = 0$.

c) Existe $B \in M_n(K)$ tal que $AB = I$.

DEMOSTRACION La implicación *a*) $\Rightarrow$ *b*) es inmediata, ya que si $AX = 0$ tenemos

$$X = IX = (A^{-1}A)X = A^{-1}(AX) = A^{-1}0 = 0\,.$$

Para probar que *b*) implica *c*), consideremos la aplicación

$$\varphi : M_n(K) \to M_n(K)$$

definida por $\varphi(X) = AX$. La misma es aditiva, por la propiedad distributiva del producto de matrices respecto a la suma, y además, dados $\lambda \in K$ y $X \in M_n(K)$ tenemos:

$$\varphi(\lambda X) = \varphi(E_\lambda X) = A(E_\lambda X) = (AE_\lambda)X = (E_\lambda A)X = E_\lambda(AX) = \\ = \lambda(AX) = \lambda\varphi(X),$$

esto es, φ es una transformación lineal. Puesto que por hipótesis φ es un monomorfismo y la dimensión de $M_n(K)$ es finita, sigue por el corolario 13.1.8 que también es un epimorfismo. En particular $I \in \mathrm{Im}(\varphi)$, y por lo tanto existe $B \in M_n(k)$ tal que $I = \varphi(B) = AB$, como queríamos demostrar.

Supongamos por último que vale *c*) y consideremos ahora el endomorfismo ξ de $M_n(K)$ definido por $\xi(X) = XA$. Si $Y \in \mathrm{Nu}(\xi)$, multiplicando a derecha por B la igualdad $YA = 0$ resulta que

$$0 = 0B = (YA)B = Y(AB) = YI = Y\,.$$

Luego ξ es un monomorfismo y por lo tanto un epimorfismo. Razonando como antes, resulta que existe $C \in M_n(K)$ tal que $CA = I$, de donde sigue que

$$C = CI = C(AB) = (CA)B = IB = B\,.$$

En consecuencia A es inversible y $A^{-1} = B$. Finalizada la prueba, señalemos que por razones evidentes de simetría la inversibilidad de A también es equivalente a cualquiera de las condiciones

$*$) Si $X \in M_n(K)$ y $XA = 0$ entonces $X = 0$.

$**$) Existe $B \in M_n(K)$ tal que $BA = I$. $\quad \Diamond$

Rango

Introduciremos ahora una noción que tendrá especial relevancia en la resolución de sistemas de ecuaciones lineales, y estrechamente conectada con la inversibilidad de matrices cuadradas.

Comenzamos fijando una notación: si $A \in K^{m\times n}$, designaremos por F_i la $i-$ésima fila de A y por C_j su $j-$ésima columna. Obsérvese que podemos pensar a cada F_i como un elemento de K^n y a cada C_j como un elemento de K^m. En ocasiones, si es necesario especificar de qué matriz se trata, escribiremos $F_i(A)$ y $C_j(A)$, respectivamente.

Empleando las notaciones descriptas arriba, asignaremos a cada matriz A dos enteros no negativos, llamados el *rango fila* y el *rango columna* de A, que notaremos $rg_f(A)$ y $rg_c(A)$, respectivamente. Sus definiciones son las siguientes:

$$rg_f(A) = dim\,(gen(F_1, F_2, \ldots, F_m))$$
$$rg_c(A) = dim\,(gen(C_1, C_2, \ldots, C_n))\,.$$

En forma equivalente, el rango fila (columna) de una matriz es el máximo número de filas (columnas) linealmente independientes de la misma.

Por ejemplo, en el caso de las matrices con coeficientes racionales

$$A = \begin{pmatrix} 2 & 1 & 1 & 0 \\ 1 & 2 & -4 & 1 \\ 0 & -3 & 9 & -2 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 3 & 2 & 1 \\ -1 & 0 & 4 \\ 2 & 2 & 6 \end{pmatrix},$$

el lector podrá verificar que $rg_f(A) = rg_c(A) = 2$ y $rg_f(B) = rg_c(B) = 3$.

Veamos algunos hechos elementales relacionados con los rangos.

Proposición 13.2.5 Si $A = (a_{ij}) \in K^{m\times n}$ son válidas las siguientes afirmaciones:

1) $rg_f(A) = 0 \Leftrightarrow rg_c(A) = 0 \Leftrightarrow A = 0$

2) $rg_f(E_\lambda(n)) = rg_c(E_\lambda(n)) = n$ si $\lambda \in K^*$. Resulta en particular que $rg_f(I_n) = rg_c(I_n) = n$

3) $rg_f(A) \leq \min(m, n)$ y $rg_c(A) \leq \min(m, n)$

4) $rg_c(A) = dim\,(\mathrm{Im}(t_A))$.

DEMOSTRACION La demostración de los dos primeros enunciados es inmediata y queda como ejercicio para el lector. Para probar 3), observemos que el rango fila de A es la dimensión de un subespacio de K^n, y por lo tanto no puede exceder a n. Por otra parte, dicho subespacio está generado por m vectores, por lo que su dimensión es menor o igual que m. Claramente, un argumento similar vale para el rango columna.

En cuanto a 4), recordando que

$$t_A\left((x_1, x_2, \ldots, x_n)\right) = \left(\sum_{k=1}^{n} a_{1k}x_k\,,\, \sum_{k=1}^{n} a_{2k}x_k\,, \ldots, \sum_{k=1}^{n} a_{mk}x_k\right),$$

resulta que $t_A(e_j) = (a_{1j}, a_{2j}, \ldots, a_{mj})$, esto es, la imagen por t_A del $j-$ésimo vector de la base canónica de K^n es la $j-$ésima columna de A. Puesto que una transformación lineal aplica cualquier base del dominio en un sistema de generadores de la imagen, el resultado sigue. $\Diamond$

13.2.2. Sistemas de ecuaciones lineales

En el capítulo anterior hemos presentado los sistemas de ecuaciones lineales homogéneos, cuyos conjuntos de soluciones son ejemplos importantes de subespacios, aunque no hemos brindado ningún método general para resolverlos. Nos proponemos hacerlo ahora, no solamente por la obvia importancia práctica del asunto, sino también porque el método de resolución nos proporcionará información teórica acerca de los rangos fila y columna de una matriz cualquiera y de la inversibilidad de una matriz cuadrada. Trabajaremos con sistemas más generales que los ya mencionados, en los que no necesariamente todas las ecuaciones están igualadas a cero.

Precisamente, dada $A = (a_{ij}) \in K^{m\times n}$ y dados elementos $b_1, b_2, \ldots, b_m$ de K, un esquema $\mathcal{S}$ del tipo

$$\begin{array}{ccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\
a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & b_2 \\
\ldots & & \ldots & & \ldots & & \ldots & & \ldots \\
\ldots & & \ldots & & \ldots & & \ldots & & \ldots \\
a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & b_m
\end{array}$$

se dice un sistema lineal de m ecuaciones y n incógnitas con coeficientes en el cuerpo K.

Naturalmente, una $n-$upla $(c_1, c_2, \ldots, c_n) \in K^n$ se dice una *solución* de $\mathcal{S}$ si y solo si sus componentes satisfacen todas las ecuaciones, esto es,

$$\sum_{j=1}^{n} a_{ij}\, c_j = b_i \text{ para } i = 1, 2 \ldots, m\,.$$

El sistema se dirá *compatible* o *incompatible* según admita o no soluciones. En el primer caso, suele decirse que es compatible determinado si admite una única solución y compatible indeterminado si admite más de una. Si bien no usaremos demasiado esta terminología, por no creerla del todo adecuada, la mencionamos por ser de uso corriente.

Si $b_i = 0$ para todo i diremos que $\mathcal{S}$ es *homogéneo*, mientras que se dice *no homogéneo* si $b_k \neq 0$ para algún k. Existe una notoria distinción entre ambos tipos, ya que un sistema homogéneo siempre es compatible (el vector nulo es solución), mientras que un sistema no homogéneo puede ser compatible o incompatible (véase por caso los sistemas de los ejemplos 13.2.3), aunque desde el punto de vista estructural existe entre ellos una diferencia más importante: el conjunto de soluciones de un sistema lineal homogéneo es un subespacio de K^n, mientras que el de un sistema no homogéneo en ningún caso lo es (el vector nulo no es solución si algún b_i es distinto de cero).

La matriz A se dice la matriz del sistema, mientras que la matriz $(A \mid b)$ de m filas y $n+1$ columnas

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{pmatrix}$$

se llama la matriz *ampliada* del sistema. El hecho de que todo sistema de ecuaciones esté asociado a una matriz nos permite expresarlo en ocasiones en una forma más compacta y sugerente, ya que las m igualdades que lo definen equivalen a la igualdad matricial

$$A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Por ejemplo, esta forma de escribir nos permite apreciar que el sistema es compatible si y solo si el vector $b = (b_1, b_2, \ldots, b_m)$ pertenece a la imagen de la transformación lineal t_A asociada a la matriz A.

OPERACIONES ELEMENTALES Hemos visto hasta aquí algunos casos sencillos de resolución de sistemas, y en todos ellos hemos encontrado sus soluciones (o hemos declarado su incompatibilidad) manipulando las ecuaciones en

una forma adecuada. Naturalmente, es fundamental que esas manipulaciones preserven el conjunto de soluciones del sistema, sin introducir alguna nueva y sin perder alguna existente. En general, dos sistemas de ecuaciones que tienen las mismas soluciones se dicen equivalentes. Por lo tanto, toda operación que realicemos para resolver un sistema debe transformar éste en otro equivalente.

Designaremos por E_i la $i-$ésima ecuación de $\mathcal{S}$. Si $r, s \in \mathbb{I}_m$ y $\lambda \in K$, notaremos λE_r la ecuación obtenida multiplicando los coeficientes de E_r por λ, y designaremos por $E_r + E_s$ la ecuación que se obtiene sumando miembro a miembro los coeficientes de E_r y E_s. Esto es:

$$\lambda E_r \,:\, \sum_{j=1}^{n} \lambda a_{rj} x_j = \lambda b_r \quad \text{y} \quad E_r + E_s \,:\, \sum_{j=1}^{n} (a_{rj} + a_{sj}) x_j = b_r + b_s \,.$$

Mostraremos en el siguiente lema un grupo de operaciones que aplicadas a las ecuaciones de un sistema preservan sus soluciones, y que como veremos más adelante, permiten simplificarlo notablemente. Las llamaremos *operaciones elementales* y serán la base del método general de resolución que estableceremos luego.

Lema 13.2.6 El conjunto de soluciones de $\mathcal{S}$ permanece invariante si se aplica cualquiera de las tres siguientes operaciones ($r, s \in \mathbb{I}_m$ y $\lambda \in K$):

i) Intercambiar E_r y E_s entre sí.

ii) Reemplazar E_r por λE_r ($\lambda \neq 0$).

iii) Reemplazar E_r por $E_r + \lambda E_s$ ($r \neq s$).

Se sobreentiende que en todos los casos las demás ecuaciones permanecen sin cambios.

DEMOSTRACION El resultado es obvio para la primera operación. Respecto a las otras dos, notemos que en ambos casos el nuevo sistema $\mathcal{S}'$ sólo difiere del original en la $r-$ésima ecuación, por lo que bastará probar que toda solución $(c_1, c_2, \ldots, c_n)$ de $\mathcal{S}$ es solución de la $r-$ésima ecuación de $\mathcal{S}'$. La prueba es bien sencilla, ya que

$$\sum_{j=1}^{n} (\lambda a_{rj}) c_j = \lambda \sum_{j=1}^{n} a_{rj} c_j = \lambda b_r$$

en el caso ii), y

$$\sum_{j=1}^{n} (a_{rj} + \lambda a_{sj}) c_j = \sum_{j=1}^{n} a_{rj} c_j + \lambda \sum_{j=1}^{n} a_{sj} c_j = b_r + \lambda b_s$$

en el caso iii).

Inversamente, la demostración anterior también prueba que toda solución del nuevo sistema es solución del original. En efecto, E_r se obtiene en ii) multiplicando la $r-$ésima ecuación de $\mathcal{S}'$ por λ^{-1}, y en iii), sumándole la $s-$ésima ecuación de $\mathcal{S}'$ multiplicada por $-\lambda$.

Obsérvese que en la demostración ha quedado tácitamente establecido que cada una de las transformaciones elementales admite una transformación inversa (que también es elemental), razón por la cual al aplicar cualquiera de ellas a un sistema de ecuaciones se obtiene otro que resulta equivalente al dado. $\Diamond$

Un sistema lineal está determinado por los coeficientes de sus ecuaciones, o sea, por su matriz ampliada. Debido a ello, para resolver un sistema es suficiente y mucho más cómodo trabajar directamente con su matriz. En tal caso, debemos tener en cuenta que las operaciones de ecuaciones indicadas arriba corresponden a operaciones de filas en la matriz ampliada. Precisamente, las siguientes operaciones entre las filas F_i de $(A \mid b)$ (o de cualquier otra matriz), serán llamadas *operaciones elementales de filas*:

i) Intercambiar F_r y F_s entre sí.

ii) Reemplazar F_r por λF_r $(\lambda \neq 0)$.

iii) Reemplazar F_r por $F_r + \lambda F_s$ $(r \neq s)$.

En cualquier caso, la matriz obtenida se dirá *equivalente* a la matriz dada.

NOTA Observemos que en la resolución de un sistema homogéneo la última columna permanecerá nula a lo largo de cualquier aplicación sucesiva de las operaciones anteriores. Por lo tanto, lo más práctico es omitirla y trabajar sólo con la matriz A del sistema. Por supuesto, al final debemos recordar que todas las ecuaciones están igualadas a cero. $\Diamond$

En ejemplos anteriores hemos calculado los rangos fila y columna de varias matrices, y en todo ellos resultaron ser iguales. No se trata de una casualidad, pues ya veremos más adelante que ambos siempre coinciden. En la siguiente proposición mostraremos un resultado conducente a probar tal hecho.

Proposición 13.2.7 Los rango fila y columna de una matriz son invariantes por aplicación de cualquiera de las operaciones elementales de fila.

DEMOSTRACION Sea $B \in K^{m\times n}$ y sea C obtenida de B por aplicación de alguna de las operaciones elementales de fila. Para estudiar sus respectivos rangos columna, consideremos el subespacio S de soluciones del sistema

lineal homogéneo

$$B\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

que de acuerdo con el lema anterior coincide con el subespacio de soluciones del sistema

$$C\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Puesto que como vimos en la proposición 13.2.5 el rango columna de B es la dimensión de la imagen de la transformación lineal

$$t_B : K^n \longrightarrow K^m$$

asociada a B, cuyo núcleo es S, aplicando el teorema de la dimensión resulta que

$$\begin{aligned} rg_c(B) = dim\,(\mathrm{Im}(t_B)) = n - dim\,(\mathrm{Nu}(t_B)) = n - dim\,(S) = \\ = n - dim\,(\mathrm{Nu}(t_C)) = dim\,(\mathrm{Im}(t_C)) = rg_c(C)\,, \end{aligned}$$

como queríamos demostrar.

En cuanto a los rangos fila, es claro que $rg_f(C) = rg_f(B)$ si C se obtiene de B por transposición de filas. Para tratar los otros dos casos, y para disponer de una notación más fluida, designemos por $v_1, \ldots, v_m$ los vectores de K^n correspondientes a las filas de B, siendo entonces el rango fila de B la dimensión de $U = gen(v_1, \ldots, v_m)$.

Si C se obtiene de B por una operación de tipo ii), digamos multiplicando la $k-$śima fila por un escalar α no nulo, resultará que $rg_f(C)$ es la dimensión del subespacio $U_2 = gen(v_1, \ldots, \alpha v_k, \ldots, v_m)$. La igualdad de los rangos sigue entonces de la igualdad de los subespacios U y U_2, ya que todo generador de U_2 pertenece a U y viceversa (notemos que $v_k = (\alpha^{-1})\alpha v_k$).

En el caso de una operación de tipo iii), suponiendo que C se obtiene reemplazando $F_k(B)$ por $F_k(B) + \gamma F_s(B)$, donde $\gamma \in K$ y $s \neq k$, la igualdad de las correspondientes dimensiones también sigue de la igualdad de los subespacios U y $U_3 = gen(v_1, \ldots, v_k + \gamma v_s, \ldots, v_m)$. En efecto, todo generador de U_3 pertenece a U (lo que es obvio) y todo generador de U pertenece a U_3, pues $v_k = (v_k + \gamma v_s) + (-\gamma)v_s$. Luego $rg_f(C) = rg_f(B)$ y la prueba finaliza. $\Diamond$

Ejemplo 13.2.8 Vamos a resolver, aplicando las operaciones elementales, el sistema lineal homogéneo con coeficientes reales

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

Para simplificar la matriz del sistema, iremos transformándola mediante la aplicación de operaciones elementales de fila. Cada paso será indicado con una flecha, aclarando que en algunos pasos se aplicarán sucesivamente dos o más operaciones de la misma índole. El proceso es el siguiente:

$$\begin{pmatrix} 2 & 1 & 3 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 1 \\ 1 & -1 & 3 & -1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -1 & 3 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 1 \\ 2 & 1 & 3 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -1 & 3 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 3 & -3 & 3 \end{pmatrix} \longrightarrow$$

$$\longrightarrow \begin{pmatrix} 1 & 0 & 2 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

En el primer paso hemos permutado entre sí la primera y la cuarta fila. En el segundo hemos sumado a la cuarta fila la primera multiplicada por -2 y en el tercero hemos sumado sucesivamente a las filas 1, 3 y 4 la segunda fila multiplicada por 1, -1 y -3, respectivamente. Finalmente, en el cuarto paso sumamos a las filas 1 y 4 la tercera fila multiplicada por 1 y -3, respectivamente. Volviendo al lenguaje de ecuaciones, concluimos que el sistema original es equivalente al sistema

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

(obsérvese que hemos suprimido la última ecuación, ya que al corresponder a una fila nula de la matriz ampliada es satisfecha por todo elemento del espacio y por lo tanto es redundante). Dada la forma sencilla que tiene éste, podemos expresar las variables x_1, x_2 y x_4 en función de x_3, resultando que todo elemento del espacio S de soluciones del sistema es de la forma $(-2x_3, x_3, x_3, 0)$. Puesto que es inmediato verificar que todo vector de este tipo es solución, concluimos que

$$x \in S \Leftrightarrow \text{existe } \lambda \in \mathbb{R} \text{ tal que } x = (-2\lambda, \lambda, \lambda, 0) = \lambda(-2, 1, 1, 0),$$

esto es, S es el subespacio de dimensión uno de R^4 generado por el vector $(-2, 1, 1, 0)$ $\quad \Diamond$

Matrices escalonadas reducidas.

Es indudable que pudimos resolver con suma facilidad el sistema anterior debido al peculiar aspecto de su matriz. La buena noticia es que mediante las operaciones elementales de fila podemos transformar cualquier matriz en una que tenga esas características. Precisaremos a través de la siguiente definición a qué tipo de características nos estamos refiriendo.

> Diremos que una matriz $T = (t_{ij})$ es *escalonada reducida* (ER) si y solo si se verifican las siguientes condiciones:
>
> ER$_1$) Si $F_k(T) = 0$ entonces $F_i(T) = 0$ para todo $i > k$
>
> ER$_2$) El primer elemento no nulo de cualquier fila no nula de T es 1
>
> ER$_3$) Si el primer elemento no nulo de $F_k(T)$ está en la columna s de T entonces $t_{is} = 0$ para todo $i \neq k$
>
> ER$_4$) Si $k < l$ y $F_k(T)$ y $F_l(T)$ son no nulas, el primer coeficiente no nulo de $F_l(T)$ está a la derecha del primer coeficiente no nulo de $F_k(T)$ (no necesariamente en columnas consecutivas).

Como resumen más coloquial, para dar una idea visual de la definición, digamos que las filas nulas de una matriz ER (si existen) se encuentran en la parte inferior de la misma, y que si una columna contiene el primer elemento no nulo de una fila entonces los restantes elementos de la columna son nulos. Suponiendo por último que las filas no nulas de T son las r primeras, resulta por la segunda y la cuarta condición que el primer elemento no nulo de cada una de ellas es 1 y que estos r unos se disponen de izquierda a derecha en "escalera"descendente.

Ejemplos 13.2.9 Ilustremos la definición con algunos ejemplos.

1) Es inmediato verificar que la matriz nula y cualquier matriz identidad son escalonadas reducidas.

2) La matriz $\begin{pmatrix} \mathbf{1} & 0 & 2 & 0 \\ 0 & \mathbf{1} & -1 & 0 \\ 0 & 0 & 0 & \mathbf{1} \\ 0 & 0 & 0 & 0 \end{pmatrix}$ del ejemplo anterior es escalonada reducida.

   Obsérvese que $r = 3$ en este caso y que hemos resaltado en negrita los "unos" mencionados en el comentario que sigue a la definición. La

tercera columna no contiene el primer elemento no nulo de ninguna fila y por lo tanto no existe ninguna restricción sobre sus elementos.

3) Las matrices matrices no son escalonadas reducidas:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} \; ; \; \begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 0 & 0 \end{pmatrix} \; ; \; \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \; ; \; \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{pmatrix} \; .$$

El lector podrá verificar que en cada una de ellas no se cumple exactamente una de las cuatro condiciones de la definición. $\Diamond$

Retornando a las ecuaciones, pensemos en un sistema homogéneo (ya veremos cómo lidiar con el caso no homogéneo) cuya matriz es escalonada reducida. Suponiendo como antes que las filas no nulas son las r primeras, y descartadas las filas nulas, cada F_i tendrá su primer uno en una cierta columna j_i, siendo $j_1 < j_2 < \ldots < j_r$. En términos de las incógnitas, resulta por la condición ER_3) de la definición que cada una de dichas variables x_{j_i} *aparece exactamente una vez en el sistema*, a saber, con coeficiente 1 en la $i-$ésima ecuación. Podemos entonces resolver el sistema en forma análoga a la del ejemplo 13.2.8, expresando las r variables $x_{j_1}, x_{j_2}, \ldots, x_{j_r}$ en función de las $n - r$ restantes, que variarán libremente.

Antes de probar que toda matriz puede llevarse a la forma escalonada reducida mediante operaciones de fila (proceso de *triangulación*), lo que de acuerdo con lo anterior nos brindará un método general para resolver sistemas lineales, probemos una importante propiedad de las matrices escalonadas reducidas, relacionada con sus rangos fila y columna.

Proposición 13.2.10 Sea $T = (t_{ij}) \in K^{m\times n}$ una matriz escalonada reducida con r filas no nulas. Entonces

$$rg_f(T) = rg_c(T) = r\,.$$

DEMOSTRACION El resultado es trivial si $r = 0$, ya que en ese caso T es nula. Suponiendo $r > 0$, designemos por $v_1, v_2, \ldots, v_r$ los vectores de K^n correspondientes a las primeras r filas de T, esto es, $(v_k)_l = t_{kl}$, donde $(v_k)_l$ denota la $l-$ésima componente de v_k. Si $\alpha_1, \alpha_2, \ldots, \alpha_r$ es una familia de escalares y j_i designa como antes la columna en la que aparece el primer elemento no nulo de $F_i(T)$ $(1 \le i \le r)$, resulta que la componente j_i-ésima de $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_r v_r$ es

$$\left(\sum_{k=1}^{r} \alpha_k v_k\right)_{j_i} = \sum_{k=1}^{r} \alpha_k (v_k)_{j_i} = \sum_{k=1}^{r} \alpha_k t_{kj_i} = \sum_{k=1}^{r} \alpha_k \delta_{ki} = \alpha_i\,,$$

por definición de los j_i.

Resulta en particular que si la combinación lineal es nula entonces $\alpha_i = 0$ para todo i y por lo tanto la familia $\{v_1, v_2, \ldots, v_r\}$ es linealmente independiente. Luego $rg_f(T) = r$, ya que las restantes filas de T son nulas.

En cuanto al rango columna, si $w_1, \ldots, w_n$ son los vectores de K^m definidos por las columnas de T, resulta en este caso que $(w_k)_l = t_{lk}$. Sigue de la condición ER_3) de matriz escalonada reducida que $w_{j_1}, w_{j_2}, \ldots, w_{j_r}$ son los vectores de la base canónica (e_k) de K^m correspondientes a los índices $j_1, j_2, \ldots, j_r$, y en consecuencia $rg_c(T) \geq r$.

Por otra parte, dado cualquier índice j entre 1 y n, y teniendo en cuenta que las últimas $m - r$ filas de T son nulas, resulta que

$$w_j = \sum_{i=1}^{m} t_{ij} e_i = \sum_{i=1}^{r} t_{ij} e_i \,,$$

lo que muestra que $gen(w_1, w_2, \ldots, w_m)$ está contenido en un espacio vectorial de dimensión r y por lo tanto $rg_c(T) \leq r$. Luego $rg_c(T) = r$, como queríamos demostrar $\quad \Diamond$

Probemos ahora sí que toda matriz puede triangularse.

Proposición 13.2.11 Toda matriz puede transformarse mediante operaciones elementales de fila en una matriz escalonada reducida.

DEMOSTRACION No probaremos formalmente el enunciado, sino que delinearemos la sucesión de pasos a seguir para triangular una matriz $A \in K^{m \times n}$. Puesto que la matriz nula es escalonada reducida, supondremos $A \neq 0$.

Como primer paso determinamos

$$j_1 = \min \{j : a_{kj} \neq 0 \text{ para algún } k\} \,,$$

es decir, C_{j_1} es la primera columna no nula de A. Luego, intercambiando F_1 y F_k (si $k > 1$) y dividiendo la primera fila por a_{kj_1}, obtenemos una matriz B tal que $b_{1j_1} = 1$ (por supuesto, si existiera un indice i tal que $a_{ij_1} = 1$ simplemente permutamos F_1 y F_i). Por último, para terminar la primera etapa, anulamos el resto de los elementos de la columna j_1 de B reemplazando $F_i(B)$ por $F_i(B) - a_{ij_1} F_1(B)$ para todo $i > 1$ tal que $a_{ij_1} \neq 0$. En resumen, al final de este tramo obtenemos una matriz C, equivalente a A, tal que $c_{ij} = 0$ para todo $j < j_1$, $c_{1j_1} = 1$ y $c_{ij_1} = 0$ para todo $i > 1$.

Si $F_i(C) = 0$ para todo $i > 1$ resulta que C es escalonada reducida y el proceso termina. Si no, tomamos

$$j_2 = \min \{j > j_1 : a_{kj} \neq 0 \text{ para algún } k \geq 2\} \,,$$

y procedemos en forma completamente análoga a la de la primera etapa. Vale decir, obtenemos una matriz D equivalente a C (luego a A) tal que $d_{ij} = c_{ij}$ si $j \leq j_1$, $d_{ij=0}$ si $i \geq 2$ y $j_1 < j < j_2$, $d_{2j_2} = 1$ y $d_{ij_2} = 0$ para todo $i \neq 2$, lo que completa la segunda etapa.

Luego el proceso continúa siempre de la misma manera: si $F_i(D) = 0$ para todo $i > 2$ entonces D es escalonada reducida y llegamos al final. Si no, elegimos el menor $j_3 > j_2$ tal que $d_{kj_3} \neq 0$ para algún $k \geq 3 \ldots$, y así sucesivamente, hasta que en alguna etapa la matriz alcanzada sea escalonada reducida. Obsérvese que ello debe ocurrir en a lo sumo n etapas, ya que $\{j_1, j_2, \ldots\}$ es una sucesión estrictamente creciente de números naturales acotada por n. $\Diamond$

Ejemplo 13.2.12 Vamos a aplicar el método descripto al caso de la matriz

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

En cada etapa señalaremos las operaciones realizadas y exhibiremos a continuación la matriz obtenida. El lector se encargará de verificar que los cálculos son correctos.

Primera etapa: Intercambiamos la primera y la cuarta fila, de manera de ubicar un 1 en la posición $(1, 1)$, y eliminamos los elementos no nulos del resto de la primera columna pivoteando con dicho elemento

$$M_1 = \begin{pmatrix} 1 & 0 & -3 & -1 & 5 \\ 0 & 3 & 15 & 6 & -21 \\ 0 & -2 & -10 & -4 & 17 \\ 0 & 1 & 5 & 2 & -7 \end{pmatrix}.$$

Segunda etapa: Similarmente a la primera etapa, intercambiamos la segunda y la cuarta fila y eliminamos los elementos no nulos del resto de la segunda columna

$$M_2 = \begin{pmatrix} 1 & 0 & -3 & -1 & 5 \\ 0 & 1 & 5 & 2 & -7 \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Tercera etapa: Análogamente, dividimos la tercera fila por 3 y eliminamos los restantes elementos no nulos de la quinta columna

$$M_3 = \begin{pmatrix} 1 & 0 & -3 & -1 & 0 \\ 0 & 1 & 5 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Finaliza entonces el proceso de triangulación, resultando que M_3 es la forma escalonada reducida de M. Obsérvese que en este caso tenemos $r = 3$, $j_1 = 1$, $j_2 = 2$ y $j_3 = 5$. $\Diamond$

La proposición anterior, que podría calificarse de índole práctica, nos permite obtener el siguiente y relevante resultado teórico:

Corolario 13.2.13 $rg_f(A) = rg_c(A)$ para toda matriz A. Cualquiera de estos números será llamado simplemente el *rango* de A, y lo notaremos $rg(A)$.

DEMOSTRACION Basta aplicar las proposiciones 13.2.7, 13.2.10 y 13.2.11, ya que si T es una matriz escalonada reducida equivalente a A tenemos:

$$rg_f(A) = rg_f(T) = rg_c(T) = rg_c(A)\,. \quad \Diamond$$

Método de resolución.

Ya estamos en condiciones de establecer un método general de resolución de sistemas de ecuaciones lineales, que naturalmente incluirá una forma adecuada de describir sus soluciones (si las hay). Consideremos para ello el sistema genérico $\mathcal{S}$ con coeficientes en K dado por las ecuaciones

$$\begin{array}{ccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\
a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & b_2 \\
\ldots & & \ldots & & \ldots & & \ldots & & \ldots \\
\ldots & & \ldots & & \ldots & & \ldots & & \ldots \\
a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & b_m
\end{array}$$

y analicemos separadamente el caso homogéneo y el caso no homogéneo:

CASO HOMOGENEO Esta situación ya ha sido tratada anteriormente, así que simplemente la recordaremos y puntualizaremos algunos detalles. Se procede operando sobre las filas de la matriz A del sistema hasta obtener una matriz escalonada reducida $T = (t_{ij})$ equivalente a A, que tendrá en particular el mismo rango r que A. Podemos en tal caso expresar ciertas variables $x_{j_1}, x_{j_2}, \ldots, x_{j_r}$ en función de las $n - r$ restantes, donde j_i es el número de columna en la que aparece el primer elemento no nulo de la fila i $(1 \leq i \leq r)$.

Precisamente, y suponiendo para simplificar la notación que las variables j_i son las r primeras, las soluciones se obtienen asignando libremente valores arbitrarios $\lambda_{r+1}, \lambda_{r+2}, \ldots, \lambda_n$ a las $n - r$ últimas variables y despejando x_i de la $i-$ésima ecuación para cada $i \leq r$.

Concretamente,

$$x_i = \sum_{j=r+1}^{n} (-t_{ij})\lambda_j\,,$$

resultando que una $n-$upla $x = (x_1, x_2, \ldots, x_n)$ es solución de $\mathcal{S}$ si y solo si es de la forma

$$x = (\sum_{j=r+1}^{n} (-t_{1j})\lambda_j \,, \ldots, \sum_{j=r+1}^{n} (-t_{rj})\lambda_j \,, \lambda_{r+1} \,, \lambda_{r+2} \,, \ldots, \lambda_n) =$$
$$= \sum_{k=r+1}^{n} \lambda_k(-t_{1k} \,, -t_{2k} \,, \ldots \,, -t_{rk} \,, 0 \,, \ldots \,, 0 \,, \overset{k}{1} \,, 0 \,, \ldots \,, 0) \,.$$

Designando por w_k cada una de las $n-$uplas que intervienen en esta última combinación lineal, queda probado entonces que los w_k generan el subespacio S de soluciones de $\mathcal{S}$. Como además la familia $\{w_{r+1}, w_{r+2}, \ldots, w_n\}$ es linealmente independiente, como es muy fácil probar, concluimos que la misma es una base de S.

En particular, hemos obtenido una fórmula para determinar la dimensión del subespacio S de soluciones de un sistema lineal homogéneo de m ecuaciones y n incógnitas con matriz A, a saber:

Fórmula 13.2.14

$$dim(S) = n - rg(A) \,.$$

NOTA Si bien hemos destacado la fórmula de arriba, efectuamos el desarrollo anterior con el objetivo principal de mostrar al lector cómo hallar una base del subespacio de soluciones, ya que en realidad podríamos haber deducido su dimensión *a priori*. En efecto, recordando que S es el núcleo de la transformación lineal t_A asociada a A y usando el teorema de la dimensión resulta que

$$dim(S) = dim\,(\mathrm{Nu}(t_A)) = n - dim\,(\mathrm{Im}(t_A)) = n - rg(A) \,. \quad \Diamond$$

Ejemplo 13.2.15 Resolvamos el sistema lineal con coeficientes reales

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

cuya matriz (de rango 3) es la matriz M del ejemplo 13.2.12. Puesto que M_3 es la forma escalonada reducida de M, el sistema dado es equivalente al sistema

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

en el que podemos expresar x_1, x_2 y x_5 en función de x_3 y x_4. Precisamente, asignando valores a y b cualesquiera a estas dos variables resulta que toda solución es de la forma

$$(3a + b, -5a - 2b, a, b, 0) = a(3, -5, 1, 0, 0) + b(1, -2, 0, 1, 0)$$

esto es, el subespacio de soluciones es de dimensión 2 y está generado por los vectores $(3, -5, 1, 0, 0)$ y $(1, -2, 0, 1, 0)$ (notará el lector que hemos reproducido en este ejemplo particular todos los pasos del caso general). $\Diamond$

La fórmula obtenida para la dimensión del subespacio de soluciones de un sistema lineal homogéneo tiene un par de consecuencias interesantes, que detallamos en el siguiente enunciado.

Corolario 13.2.16 Conservando las notaciones precedentes, son válidos los siguientes hechos:

1) $\mathcal{S}$ admite solución única (la trivial) si y solo si $rg(A) = n$.

2) Si $m < n$ entonces $\mathcal{S}$ admite soluciones no triviales.

DEMOSTRACION Ambas afirmaciones se deducen de la fórmula 13.2.14. La cuestión es inmediata en el caso de la primera, y en cuanto a la segunda tenemos que

$$dim(S) = n - rg(A) \geq n - m > 0\,,$$

y por lo tanto S es un subespacio no trivial de K^n. $\Diamond$

CASO NO HOMOGENEO Para resolver el caso no homogéneo se procede en forma muy similar, trabajando con la matriz ampliada $A \mid b$. Precisamente, extendemos a la matriz ampliada cada una de las operaciones de filas necesarias para obtener la forma escalonada reducida T de A. Así, si en el primer paso intercambiamos las fila i y k de A entonces también intercambiamos los coeficientes b_i y b_k, si multiplicamos la fila $i-$ésima de A por un escalar α entonces reemplazamos b_i por αb_i, y si reemplazamos $F_k(A)$ por $F_k(A) + \lambda F_i(A)$, también reemplazamos b_k por $b_k + \lambda b_i$. Procediendo sucesivamente de esta manera arribamos finalmente al sistema equivalente

$$T \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix},$$

donde los c_i son elementos de K. Suponiendo como siempre que las últimas $n - r$ filas de T son nulas, se presentan dos situaciones posibles, a saber:

i) $c_k \neq 0$ para algún $k > r$. En tal caso el sistema es incompatible, ya que claramente la $k-$ésima ecuación no admite ninguna solución (todos los coeficientes del miembro de la izquierda son nulos y el de la derecha no).

ii) $c_k = 0$ para todo $k > r$. Entonces el sistema es compatible, ya que eliminando las últimas $n - r$ ecuaciones, por ser redundantes, podemos proceder como en el caso homogéneo, expresando r variables en función de $n - r$ parámetros libres. $\Diamond$

Antes de exhibir un ejemplo de resolución de un sistema lineal no homogéneo, mostraremos una forma más estructural de describir el conjunto S de soluciones de $\mathcal{S}$, obviamente si éste es compatible. Si bien S no es un subespacio en el caso no homogéneo, veremos que está fuertemente conectado con un cierto subespacio de K^n.

En efecto, supongamos conocer una solución $c = (c_1, c_2, \ldots, c_n)$ de $\mathcal{S}$ y consideremos una solución cualquiera $x = (x_1, x_2, \ldots, x_n)$. Resulta entonces que

$$\sum_{j=1}^{n} a_{ij}(x_j - c_j) = \sum_{j=1}^{n} a_{ij}x_j - \sum_{j=1}^{n} a_{ij}c_j = b_i - b_i = 0\,,$$

para todo $1 \leq i \leq m$, lo que significa que $x - c$ pertenece al subespacio S_0 de soluciones del sistema homogéneo

$$A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

llamado *sistema homogéneo asociado* a $\mathcal{S}$. Escribiendo $x = (x - c) + c$, sigue que todo elemento de S es suma de un elemento de S_0 y la solución fija c.

Recíprocamente, si $y \in S_0$ tenemos que

$$\sum_{j=1}^{n} a_{ij}(y_j + c_j) = \sum_{j=1}^{n} a_{ij}y_j + \sum_{j=1}^{n} a_{ij}c_j = 0 + b_i = b_i$$

para todo $1 \leq i \leq m$ y por lo tanto $y + c \in S$.

En resumen, hemos demostrado la igualdad

$$S = S_0 + c\,, \tag{13.1}$$

donde $S_0 + c = \{\, y + c : y \in S_0 \,\}$.

Vale decir, las soluciones de un sistema lineal se obtienen sumando una solución particular del mismo a cada una de las soluciones del sistema homogéneo asociado. Si K es infinito, deducimos de este hecho que todo sistema lineal compatible con coeficientes en K tiene una única solución o tiene infinitas soluciones, según que S_0 sea el subespacio nulo o sea un subespacio no trivial de K^n.

Aunque vale en cualquier caso, agreguemos que la fórmula 13.1 tiene real interés si el sistema es no homogéneo, ya que $S_0 = S$ en el caso homogéneo y la igualdad se verifica trivialmente tomando $c = 0$.

Ejemplo 13.2.17 Apliquemos el método general para resolver el sistema

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

cuya matriz M es la del ejemplo 13.2.12. De acuerdo con los cálculos que hemos efectuado para triangular M, el lector puede verificar que es equivalente resolver el sistema

$$\begin{cases} x_1 - 3x_3 - x_4 = -6 \\ x_2 + 5x_3 + 2x_4 = 27 \\ x_5 = 3 \\ 0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 = 0\,. \end{cases}$$

Estamos entonces en el caso ii) y el sistema es compatible. Para resolverlo efectivamente, eliminamos la última ecuación y expresamos las variables x_1, x_2 y x_5 en función de x_3 y x_4, resultando que la forma general de las soluciones es

$$\begin{aligned} x \;&=\; (3\alpha + \beta - 6, -5\alpha - 2\beta + 27, \alpha, \beta, 3) \;= \\ &= \alpha(3, -5, 1, 0, 0) + \beta(1, -2, 0, 1, 0) + (-6, 27, 0, 0, 3)\,, \end{aligned}$$

donde los parámetros α y β varían libremente sobre $\mathbb{R}$.

Observemos que $c = (-6, 27, 0, 0, 3)$ en este ejemplo (solución que corresponde a tomar $\alpha = \beta = 0$) y que S_0 es el subespacio de dimensión 2 de $\mathbb{R}^5$ generado por los vectores $(3, -5, 1, 0, 0)$ y $(1, -2, 0, 1, 0)$.

Más generalmente, proponemos al lector que demuestre como ejercicio que el sistema lineal

$$M \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix}$$

es compatible si y solo si $3b_1 - b_2 - b_4 = 0$. Equivalentemente,

$$\mathrm{Im}(t_M) = \left\{(y_1, y_2, y_3, y_4) \in \mathbb{R}^4 : 3y_1 - y_2 - y_4 = 0\right\}\,. \quad \Diamond$$

Para completar el tema estableceremos un par de resultados de importancia teórica, que nos mostrarán que es posible analizar la compatibilidad de un sistema aún sin resolverlo completamente (emplearemos las notaciones del caso general).

Proposición 13.2.18 Valen las siguientes afirmaciones:

1) $\mathcal{S}$ es compatible si y solo si $rg(A \mid b) = rg(A)$.

2) Si $m = n$, $\mathcal{S}$ tiene solución única si y solo si $rg(A) = n$.

DEMOSTRACION Supongamos que $\mathcal{S}$ es compatible y sea $c = (c_1, c_2, \ldots, c_n)$ una solución. Designando por b el vector determinado por la columna de los coeficientes independientes y usando la misma notación de la proposición 13.2.10 resulta que

$$b_i = \sum_{j=1}^{n} a_{ij}c_j = \sum_{j=1}^{n} c_j(w_j)_i = (\sum_{j=1}^{n} c_j w_j)_i$$

para todo $1 \leq i \leq m$, lo que significa que b es combinación lineal (con escalares c_j) de los vectores w_j correspondientes a las columnas de A. Luego,

$$\begin{aligned} rg(A) = rg_c(A) = dim\,(gen(w_1, w_2, \ldots, w_n)) = \\ = dim\,(gen(w_1, w_2, \ldots, w_n, b)) = rg_c(A \mid b) = rg(A \mid b), \end{aligned}$$

lo que prueba que la condición de la afirmación 1) es necesaria. Para demostrar que es suficiente, observemos que sigue de la igualdad de los rangos que b es combinación lineal de los w_j, digamos $b = \lambda_1 w_1 + \lambda_2 + w_2 + \cdots + \lambda_n w_n$. Igualando componentes resulta que

$$b_i = \sum_{k=1}^{n} \lambda_k(w_k)i = \sum_{k=1}^{n} \lambda_k a_{ik} = \sum_{k=1}^{n} a_{ik}\lambda_k$$

para todo $1 \leq i \leq m$. Luego $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ es una solución y el sistema es compatible.

Para probar 2), supongamos en primer término que $rg(A) = n$. Puesto que toda columna de A es también una columna de $A \mid b$ y puesto que ésta tiene n filas, valen las relaciones

$$n = rg(A) \leq rg(A \mid b) = rg_f(A \mid b) \leq n\,,$$

lo que asegura que $rg(A \mid b) = rg(A)$ y por lo tanto el sistema es compatible. Deducimos además de las fórmulas 13.1 y 13.2.14 que $\mathcal{S}$ admite una única solución, ya que $dim(S_0) = n - rg(A) = 0$ y en consecuencia S_0 es nulo.

La recíproca es inmediata, ya que

$$\#(S) = 1 \Leftrightarrow dim(S_0) = 0 \Leftrightarrow rg(A) = n\,. \quad \Diamond$$

13.2.3. Inversibilidad y rango

Fijado un número natural n, dedicaremos el resto de la sección a determinar condiciones necesarias y suficientes para que una matriz cuadrada de orden n sea inversible, y mostraremos cómo hallar (si existe) su inversa.

Considerando la identificación natural existente entre K^n y $K^{n\times 1}$ a través del isomorfismo

$$\vartheta\,((z_1, z_2, \ldots, z_n)) = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix},$$

emplearemos indistintamente la letra z para referirnos tanto a la $n-$upla de componentes z_i como a la matriz columna $\vartheta(z)$. Por ejemplo, dada $A \in M_n(K)$, el sistema lineal de matriz A y coeficientes independientes $b_1, b_2, \ldots, b_n$ será notado concisamente $AX = b$. Asimismo, dadas $n-$uplas $Z_1, Z_2, \ldots, Z_n$ indicaremos por $(Z_1\ Z_2\ \ldots\ Z_n)$ la matriz cuadrada de orden n cuya $j-$ésima columna es Z_j.

En el siguiente teorema caracterizaremos los elementos inversibles de $M_n(K)$.

Teorema 13.2.19 Si $A \in M_n(K)$, los siguientes enunciados son equivalentes:

a) A es inversible

b) El sistema $AX = b$ admite solución única cualquiera sea $b \in K^{n\times 1}$

c) $rg(A) = n$.

DEMOSTRACION Ya hemos probado en la proposición 13.2.18 la equivalencia de las afirmaciones *b*) y *c*), por lo que bastará probar que *a*) es equivalente a cualquiera de ellas, por ejemplo a *b*).

Suponiendo que A es inversible, dada cualquier $n-$upla z tenemos que

$$Az = b \Leftrightarrow A^{-1}(Az) = A^{-1}b \Leftrightarrow (A^{-1}A)z = A^{-1}b \Leftrightarrow z = A^{-1}b\,,$$

esto es, el sistema $AX = b$ admite la única solución $X = A^{-1}b$.

Recíprocamente, supongamos que *b*) es verdadera y sean $e_1, e_2, \ldots, e_n$ los elementos de la base canónica de K^n. Como por hipótesis existen $n-$uplas $X_1, X_2, \ldots, X_n$ tales que $AX_j = e_j$ para todo $1 \leq j \leq n$, designando por C la matriz $(X_1\ X_2\ \ldots\ X_n)$ resulta que

$$AC = (AX_1\ AX_2\ \ldots\ AX_n) = (e_1\ e_2\ \ldots\ e_n) = I_n\,.$$

Luego A es inversible, por proposición 13.2.4. $\Diamond$

Corolario 13.2.20 Una matriz cuadrada es inversible si y solo si su forma escalonada reducida es la matriz identidad.

DEMOSTRACION Sea $A \in M_n(K)$. Puesto que el rango es invariante por operaciones de fila, si A es equivalente a I_n resulta que $rg(A) = rg(I_n) = n$ y en consecuencia A es inversible, por el teorema anterior.

Sea ahora A inversible y sea T la forma escalonada reducida de A. Observando que T no tiene filas nulas, por ser $rg(T) = rg(A) = n$, designemos como de costumbre por j_k el número de columna en el que se encuentra el primer 1 de la fila k de A $(1 \leq k \leq n)$. Puesto que deben satisfacerse las desigualdades

$$1 \leq j_1 < j_2 < \ldots < j_n \leq n\,,$$

deducimos que $j_k = k$ para todo k, por ser los j_i números naturales. Teniendo en cuenta la definición de matriz escalonada reducida concluimos que $T = I_n$, como queríamos demostrar. $\Diamond$

Cálculo de la inversa

Mostraremos a continuación un método de cálculo de la inversa de una matriz cuadrada, conocido como *método de Gauss*, que básicamente es un proceso de triangulación. En realidad, para aplicarlo no es necesario saber *a priori* si la matriz es inversible. Si no lo es, ello quedará en evidencia en algún paso del proceso, mientras que si es inversible al final del algoritmo obtendremos la inversa.

MATRICES ELEMENTALES Comenzamos introduciendo un cierto conjunto de matrices cuadradas estrechamente vinculadas con las operaciones elementales de fila (m denotará un número natural fijo, r y s elementos distintos de $\mathbb{I}_m$ y λ un escalar no nulo):

- I^{rs} es la matriz que resulta de permutar las filas r y s de $I = I_m$.
- I^r_λ se obtiene reemplazando 1 por λ en la posición (r, r) de I.
- I^{rs}_λ se obtiene reemplazando 0 por λ en la posición (r, s) de I.

Las matrices anteriores se dicen *elementales* .

En el siguiente lema apreciaremos su conexión con el proceso de triangulación de una matriz.

Lema 13.2.21 Propiedades de las matrices elementales:

1) Si $A \in K^{m\times n}$, toda operación elemental de fila efectuada sobre A se obtiene multiplicando A por una matriz elemental adecuada, según el siguiente detalle (F_i designará la $i-$ésima fila de A):

   i) Si T se obtiene de A permutando F_r y F_s entonces $T = I^{rs}A$

   ii) Si T se obtiene de A multiplicando F_r por λ entonces $T = I^r_\lambda A$

   iii) Si T se obtiene de A reemplazando F_r por $F_r + \lambda F_s$ entonces $T = I^{rs}_\lambda A$.

2) Toda matriz elemental es inversible y su inversa también es una matriz elemental.

DEMOSTRACION En cada uno de los ítems de 1) calcularemos el coeficiente de la posición (i, j) del producto y comprobaremos que coincide con T_{ij} para todo par (i, j). Comenzamos por i), efectuando las operaciones en cada uno de los siguientes casos:

$$i \neq r, s \quad ; \quad (I^{rs}A)_{ij} = \sum_{k=1}^{m} I^{rs}_{ik} a_{kj} = \sum_{k=1}^{m} \delta_{ik} a_{kj} = a_{ij} = T_{ij}$$

$$i = r \quad ; \quad (I^{rs}A)_{rj} = \sum_{k=1}^{m} I^{rs}_{rk} a_{kj} = \sum_{k=1}^{m} \delta_{sk} a_{kj} = a_{sj} = T_{rj}$$

$$i = s \quad ; \quad (I^{rs}A)_{sj} = \sum_{k=1}^{m} I^{rs}_{sk} a_{kj} = \sum_{k=1}^{m} \delta_{rk} a_{kj} = a_{rj} = T_{sj} \, .$$

En la situación ii) sólo debemos considerar dos casos, a saber:

$$i \neq r \quad ; \quad (I_\lambda^r A)_{ij} = \sum_{k=1}^{m} (I_\lambda^r)_{ik} a_{kj} = \sum_{k=1}^{m} \delta_{ik} a_{kj} = a_{ij} = T_{ij}$$
$$i = r \quad ; \quad (I_\lambda^r A)_{rj} = \sum_{k=1}^{m} (I_\lambda^r)_{rk} a_{kj} = \sum_{k=1}^{m} \lambda \delta_{rk} a_{kj} = \lambda a_{rj} = T_{rj} \, .$$

Similarmente, en iii) también debemos verificar dos casos:

$$i \neq r \quad ; \quad (I_\lambda^{rs} A)_{ij} = \sum_{k=1}^{m} (I_\lambda^{rs})_{ik} a_{kj} = \sum_{k=1}^{m} \delta_{ik} a_{kj} = a_{ij} = T_{ij}$$
$$i = r \quad ; \quad (I_\lambda^{rs} A)_{rj} = \sum_{k=1}^{m} (I_\lambda^{rs})_{rk} a_{kj} = (I_\lambda^{rs})_{rr} a_{rj} + (I_\lambda^{rs})_{rs} a_{sj} =$$
$$= a_{rj} + \lambda a_{sj} = T_{rj} \, .$$

La propiedad 2) es una consecuencia inmediata del punto 1), ya que conociendo el efecto de multiplicar a izquierda una matriz cualquiera por una matriz elemental resulta que valen las igualdades

$$I^{rs} I^{rs} = I_\lambda^r I_{\lambda^{-1}}^r = I_\lambda^{rs} I_{-\lambda}^{rs} = I_m \, .$$

Luego toda matriz elemental es inversible y su inversa también es elemental, siendo $(I^{rs})^{-1} = I^{rs}$, $(I_\lambda^r)^{-1} = I_{\lambda^{-1}}^r$ y $(I_\lambda^{rs})^{-1} = I_{-\lambda}^{rs}$. $\quad \Diamond$

Método de Gauss

Supongamos que queremos determinar si una matriz cuadrada A de orden m es inversible, y en caso de serlo, calcular su inversa. Comenzamos para ello a triangular A mediante operaciones elementales de fila, obteniendo en el proceso una secuencia $A_1, A_2, \ldots, A_q$ de matrices equivalentes a A, siendo A_q escalonada reducida.

Si alguna de las matrices A_i tiene una fila nula entonces A *no es inversible*, pues $rg(A) = rg(A_i) < m$, y la cuestión ya queda resuelta. En caso contrario A es inversible y A_q es la matriz identidad I_m, como vimos en el corolario 13.2.20. Entonces, teniendo en cuenta que toda operación elemental de fila sobre una matriz se obtiene multiplicándola a izquierda por una matriz elemental, resulta que existen matrices elementales $T_1, T_2, \ldots, T_q$ tales que

$$I_m \, = \, T_q \ldots T_2 \, T_1 \, A$$

y por lo tanto $A^{-1} = T_q \ldots T_2 \, T_1$.

Por otro lado, de la igualdad $T_q \ldots T_2 \, T_1 = T_q \ldots T_2 \, T_1 \, I_m$ deducimos que A^{-1} se obtiene partiendo de la matriz identidad y efectuando en orden *las mismas operaciones de fila que transforman A en la matriz identidad.*

En la práctica, y en eso consiste el método de Gauss, se trabaja simultáneamente con dos series de matrices, una comienza en A y la otra en la matriz identidad. En la primera se realizan las operaciones de filas necesarias para obtener la forma escalonada reducida de A, y cada una de ellas se aplica al mismo tiempo a la segunda serie. Si algún término de la primera serie tiene una fila nula entonces A no es inversible y allí finaliza el proceso. Si no, en la primera serie se arriba a la identidad al cabo de q etapas y la matriz obtenida en la $q-$ésima etapa de la segunda serie es la inversa de A.

Ejemplo 13.2.22 Ilustremos el método con la matriz

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 3 & -2 \\ -4 & 6 & 3 \end{pmatrix}.$$

En este caso, a través del proceso de triangulación simultánea obtenemos las siguientes series de matrices (el lector tomará debida nota de las operaciones realizadas en cada paso):

$$A \longrightarrow \begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 0 \\ 0 & 14 & -1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 0 \\ 0 & 14 & -1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \longrightarrow$$

$$\longrightarrow \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I\,;$$

$$I \longrightarrow \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 4 & 0 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & 1 & 0 \\ 2 & -1 & 0 \\ 4 & 0 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} -3 & 2 & 0 \\ 2 & -1 & 0 \\ -24 & 14 & 1 \end{pmatrix} \longrightarrow$$

$$\longrightarrow \begin{pmatrix} -3 & 2 & 0 \\ 2 & -1 & 0 \\ 24 & -14 & -1 \end{pmatrix} \longrightarrow \begin{pmatrix} 21 & -12 & -1 \\ 2 & -1 & 0 \\ 24 & -14 & -1 \end{pmatrix} = A^{-1}. \quad \Diamond$$

13.2.4. Matriz de una transformacion lineal

A lo largo de este capítulo hemos señalado algunas conexiones existentes entre homomorfismos y matrices. Por ejemplo, vimos que toda matriz A de

m filas y n columnas induce una transformación lineal de K^n en K^m, que notamos t_A, y que el espacio de homomorfismos $\mathrm{Hom}_K(V,W)$ es isomorfo al espacio de matrices $K^{r\times s}$ si $dim\, V = r$ y $dim\, W = s$.

Recordando que t_A puede interpretarse como un producto matricial, precisamente

$$t_A\left((x_1,x_2,\ldots,x_n)\right) = (y_1,y_2,\ldots,y_m) \Leftrightarrow A\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix},$$

veamos que podemos generalizar su definición de manera de abarcar situaciones más amplias.

Concretamente, supongamos que V y W son espacios de dimensiones n y m y que $\mathcal{B}$ y $\mathcal{B}'$ son bases de V y W, respectivamente. La matriz A induce entonces una aplicación $f_A : V \longrightarrow W$ definida por la relación

$$f(v) = w \Leftrightarrow Ax^t = y^t, \tag{13.2}$$

donde x es el vector de coordenadas de x en la base $\mathcal{B}$ e y es el vector de coordenadas de w en la base $\mathcal{B}'$.

Es trivial probar, usando las propiedades estructurales del producto de matrices, que f_A es un morfismo de espacios vectoriales y que $f_A = t_A$ si $V = K^n$, $W = K^m$ y $\mathcal{B}$ y $\mathcal{B}'$ son las respectivas bases canónicas.

Por ejemplo, sean $V = W = \mathbb{R}^2$ y

$$A = \begin{pmatrix} 2 & 1 \\ 3 & 0 \end{pmatrix}.$$

Tomando $\mathcal{B}$ como la base canónica de $\mathbb{R}^2$ y $\mathcal{B}' = \{(1,-1),(-1,2)\}$, resulta que $f_A\left((2,1)\right) = (-1,7)$, pues

$$A\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$$

y $5(1,-1) + 6(-1,2) = (-1,7)$.

En el siguiente e importante teorema demostraremos que, fijadas las bases $\mathcal{B}$ y $\mathcal{B}'$, toda transformación lineal de V en W es del tipo anterior.

Teorema 13.2.23 Si f es una transformación lineal de V en W existe una única matriz $A \in K^{m\times n}$ tal que $f = f_A$.

DEMOSTRACION Si $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ y $\mathcal{B}' = \{w_1, w_2, \ldots, w_m\}$, consideremos la matriz $A \in K^{m\times n}$ cuyos elementos a_{ij} están determinados por las relaciones

$$f(v_j) = \sum_{i=1}^{m} a_{ij} w_i \qquad (j = 1, 2, \ldots n).$$

Si $v \in V$ y $x = (x_1, x_2, \ldots, x_n)$ es el vector de coordenadas de v en la base $\mathcal{B}$, tenemos entonces que

$$f(v) = f\left(\sum_{j=1}^{n} x_j v_j\right) = \sum_{j=1}^{n} x_j f(v_j) = \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij} w_i$$
$$= \sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij} x_j w_i = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij} x_j\right) w_i = \sum_{i=1}^{m} (Ax^t)_{i1} w_i,$$

lo que prueba que Ax^t es, con la identificación usual, el vector de coordenadas de $f(v)$ en la base $\mathcal{B}'$. Luego $f(v) = f_A(v)$ para todo v y por lo tanto $f = f_A$.

Una vez demostrada la existencia de una matriz como la que postula el enunciado, probemos su unicidad. Supongamos para ello que $C = (c_{ij})$ es cualquier otra matriz de m filas y n columnas tal que $f = f_C$. Puesto que $f(v_k) = f_C(v_k)$ para todo $k \leq n$ y el vector de coordenadas de v_k en la base $\mathcal{B}$ es el $k-$ésimo vector e_k de la base canónica de K^n sigue que el vector de coordenadas de $f(v_k)$ en la base $\mathcal{B}'$ corresponde a

$$\left(Ce_k^t\right)^t = e_k C^t = (c_{1k}\, c_{2k}\, \ldots\, c_{mk})\ ,$$

esto es,

$$f(v_k) = \sum_{i=1}^{m} c_{ik} w_i\,.$$

Luego $c_{ik} = a_{ik}$ para todo i y para todo k y por lo tanto $C = A$. $\quad \Diamond$

El hecho de que podamos calcular los valores de la función f por medio de la relación (13.2) indica que la matriz A contiene toda la información necesaria para determinar f. Ello por supuesto no debe extrañarnos, ya que sabemos que un homomorfismo entre espacios vectoriales de dimensión finita queda completamente definido conociendo las imágenes de un número finito de vectores del dominio. Dada su importancia y utilidad (ya veremos que A brinda información adicional sobre f), introducimos la siguiente definición:

> La matriz A se denomina la matriz de f *con respecto a las bases* $\mathcal{B}$ *y* $\mathcal{B}'$. La notaremos $M_{\mathcal{B}\mathcal{B}'}(f)$.

NOTA Observemos que la matriz $M_{\mathcal{B}\mathcal{B}'}(f)$ se escribe por columnas, en el sentido de que los vectores de coordenadas de $f(v_1), f(v_2), \ldots, f(v_n)$ respecto de la base $\mathcal{B}'$ son las columnas de la misma. Se trata de una convención, y bien podríamos hacerlo por filas con el mismo provecho. Las diferencias entre ambas formas de trabajar son puramente formales, y las dos opciones tienen sus ventajas y desventajas notacionales. Naturalmente, habiendo elegido hacerlo por columnas nos mantendremos siempre en esa tesitura.

Notemos también que $M_{\mathcal{B}\mathcal{B}'}(f) \in K^{m\times n}$, por lo que es cuadrada si y solo si $dim\, W = dim\, V$. Para simplificar la notación en el caso particular de que f sea un endomorfismo de V y $\mathcal{B}'$ sea igual a $\mathcal{B}$, escribiremos $M_{\mathcal{B}}(f)$ en vez $M_{\mathcal{B}\mathcal{B}}(f)$, y la llamaremos la matriz de f con respecto a la base $\mathcal{B}$. $\quad \Diamond$

Ejemplos 13.2.24 Ilustremos el nuevo concepto resaltando algunos casos especiales y exhibiendo algunos ejemplos de cálculo concreto. En los dos primeros ítems $\mathcal{B}$ y $\mathcal{B}'$ denotan bases cualesquiera de V y W, respectivamente.

1) Si $f \in \mathrm{Hom}_K(V,W)$ entonces $M_{\mathcal{B}\mathcal{B}'}(f) = 0$ si y solo si $f = 0_{V,W}$.

2) Si g es un endomorfismo de V entonces $M_{\mathcal{B}}(g) = I$ si y solo si $g = I_V$.

3) Si $A \in K^{m\times n}$ entonces $M_{\mathbb{E}\mathbb{E}'}(t_A) = A$, donde $\mathbb{E}$ y $\mathbb{E}'$ designan las bases canónicas de K^n y K^m, respectivamente. Encargamos al lector la demostración de este hecho y de las afirmaciones de los ejemplos precedentes.

4) Consideremos el homomorfismo $h : \mathbb{R}^3 \to \mathbb{R}^2$ definido por

$$h\left((x_1,x_2,x_3)\right) = (2x_1 - x_3, x_1 + 4x_2 + x_3)$$

y calculemos su matriz con respecto a la base canónica $\mathbb{E}$ de $\mathbb{R}^3$ y la base $\mathcal{B} = \{(1,1),(2,3)\}$ de $\mathbb{R}^2$. Operando, tenemos que

$$\begin{aligned} h(e_1) &= (2,1) = 4(1,1) - (2,3) \\ h(e_2) &= (0,4) = -8(1,1) + 4(2,3) \quad \text{y} \\ h(e_3) &= (-1,1) = -5(1,1) + 2(2,3)\,, \end{aligned}$$

como el lector puede verificar inmediatamente. Luego,

$$M_{\mathbb{E}\mathcal{B}}(h) = \begin{pmatrix} 4 & -8 & -5 \\ -1 & 4 & 2 \end{pmatrix}.$$

En cambio, invitamos al lector a comprobar que

$$M_{\mathcal{B}_1\mathcal{B}_2}(h) = \begin{pmatrix} 5 & 17 & 19 \\ -1 & -5 & -6 \end{pmatrix}$$

si $\mathcal{B}_1 = \{(1,0,0),(1,1,0),(1,1,1)\}$ y $\mathcal{B}_2 = \{(1,0),(3,-1)\}$.

5) Para ilustrar el hecho de que la matriz de una transformación lineal determina completamente a ésta, hallemos la forma general de un endomorfismo h de $\mathbb{R}^2$ cuya matriz en la base $\mathcal{B}$ del ejemplo anterior es

$$M_{\mathcal{B}}(h) = \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}.$$

Dado un elemento genérico (x_1,x_2) de $\mathbb{R}^2$, la resolución de un sencillo sistema lineal de dos ecuaciones con dos incógnitas nos muestra que vale la igualdad

$$(x_1,x_2) = (3x_1 - 2x_2)(1,1) + (x_2 - x_1)(2,3),$$

lo que significa que $(3x_1 - 2x_2, x_2 - x_1)$ es el vector de coordenadas de (x_1, x_2) respecto de la base $\mathcal{B}$. Por lo tanto

$$\begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 3x_1 - 2x_2 \\ x_2 - x_1 \end{pmatrix} = \begin{pmatrix} 5x_1 - 3x_2 \\ 7x_1 - 4x_2 \end{pmatrix}$$

es la forma transpuesta del vector de coordenadas de $h((x_1, x_2))$ respecto de la base $\mathcal{B}$, esto es,

$$h((x_1, x_2)) = (5x_1 - 3x_2)(1, 1) + (7x_1 - 4x_2)(2, 3).$$

Luego

$$h((x_1, x_2)) = (19x_1 - 11x_2, 26x_1 - 15x_2)$$

es la forma general de h que queríamos hallar. El lector puede verificar que el resultado es correcto aplicando la fórmula obtenida a los vectores de la base $\mathcal{B}$. $\Diamond$

Conservando la notación anterior, demostraremos en la siguiente proposición algunas hechos relacionados con la matriz de una transformación lineal.

Proposición 13.2.25 Son válidas las siguientes propiedades:

1) Si $f, g \in \mathrm{Hom}_K(V, W)$ entonces $M_{\mathcal{B}\mathcal{B}'}(f + g) = M_{\mathcal{B}\mathcal{B}'}(f) + M_{\mathcal{B}\mathcal{B}'}(g)$.

2) Si $f \in \mathrm{Hom}_K(V, W)$ y $\lambda \in K$ entonces $M_{\mathcal{B}\mathcal{B}'}(\lambda f) = \lambda M_{\mathcal{B}\mathcal{B}'}(f)$.

3) La aplicación
$$\chi : \mathrm{Hom}_K(V, W) \to K^{m \times n}$$
definida por $\chi(h) = M_{\mathcal{B}\mathcal{B}'}(h)$ es un isomorfismo.

4) Sea U un espacio vectorial y sea $\mathcal{B}'' = \{u_1, u_2, \ldots, u_s\}$ una base de U. Si $f \in \mathrm{Hom}_K(V, W)$ y $g \in \mathrm{Hom}_K(W, U)$ entonces
$$M_{\mathcal{B}\mathcal{B}''}(g \circ f) = M_{\mathcal{B}'\mathcal{B}''}(g)\, M_{\mathcal{B}\mathcal{B}'}(f).$$

DEMOSTRACION Para probar 1) designemos por $A = (a_{ij})$ y $C = (c_{ij})$ las matrices de f y g con respecto a las bases dadas. Entonces, dado cualquier índice $j \leq n$ tenemos que

$$(f + g)(v_j) = f(v_j) + g(v_j) = \sum_{i=1}^{m} a_{ij} w_i + \sum_{i=1}^{m} c_{ij} w_i =$$
$$= \sum_{i=1}^{m} (a_{ij} + c_{ij}) w_i = \sum_{i=1}^{m} (A + C)_{ij} w_i,$$

lo que prueba que $M_{\mathcal{B}\mathcal{B}'}(f+g) = A + C$.

La prueba de 2) es enteramente análoga, ya que

$$(\lambda f)(v_j) = \lambda f(v_j) = \lambda \sum_{i=1}^{m} a_{ij} w_i = \sum_{i=1}^{m} \lambda a_{ij} w_i = \sum_{i=1}^{m} (\lambda A)_{ij} w_i \,.$$

Con respecto a 3), observemos que las propiedades 1) y 2) aseguran que χ es un homomorfismo de espacios vectoriales, que además es un monomorfismo, pues ya vimos que $M_{\mathcal{B}\mathcal{B}'}(h) = 0$ si y solo si $h = 0_{V,W}$. Puesto que $\mathrm{Hom}_K(V,W)$ y $K^{m\times n}$ tienen la misma dimensión (ver ejemplo 13.1.11), concluimos que χ es un isomorfismo.

Probemos finalmente que la composición de homomorfismos se corresponde con el producto de matrices, como postula 4). Para ello, designemos por A y D las matrices $M_{\mathcal{B}\mathcal{B}'}(f)$ y $M_{\mathcal{B}'\mathcal{B}''}(g)$, respectivamente. Resulta entonces que

$$\begin{aligned}(g \circ f)(v_j) = g(f(v_j)) = g\left(\sum_{k=1}^{m} a_{kj} w_k\right) = \sum_{k=1}^{m} a_{kj} g(w_k) = \\ = \sum_{k=1}^{m} a_{kj} \sum_{i=1}^{s} d_{ik} u_i = \sum_{k=1}^{m} \sum_{i=1}^{s} d_{ik} a_{kj} u_i = \\ = \sum_{i=1}^{s} \left(\sum_{k=1}^{m} d_{ik} a_{kj}\right) u_i = \sum_{i=1}^{s} (DA)_{ij} u_i \,,\end{aligned}$$

y por lo tanto $M_{\mathcal{B}\mathcal{B}''}(g \circ f) = DA$, como queríamos probar. $\Diamond$

NOTA Como caso particular de 3), notemos que si $W = V$ y $\mathcal{B}' = \mathcal{B}$ la aplicación $\chi(f) = M_{\mathcal{B}}(f)$ establece un isomorfismo de $\mathrm{End}_K(V)$ en $M_n(K)$. Puesto que $\chi(h \circ g) = \chi(h)\chi(g)$ cualesquiera sean los endomorfismos g y h de V, por la propiedad 4), resulta además que χ es en tal caso un isomorfismo de anillos. Como consecuencia de la similitud estructural que ello supone, muchas cuestiones referidas a endomorfismos de un espacio vectorial de dimensión finita pueden tratarse y resolverse a nivel matricial, y recíprocamente, ciertos problemas de matrices se resuelven más adecuadamente empleando recursos de la teoría de endomorfismos de un espacio vectorial de dimensión finita. $\Diamond$

Matriz de cambio de base

Dado un espacio vectorial V, interesa sin duda conocer qué relación existe entre las coordenadas $x_1, x_2, \ldots, x_n$ de cualquier vector u respecto a cierta base $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ de V y sus coordenadas $y_1, y_2, \ldots, y_n$ respecto a otra base $\mathcal{B}' = \{w_1, w_2, \ldots, w_n\}$ de V.

Consideremos para resolver esta cuestión la matriz C de la función identidad de V respecto de las bases $\mathcal{B}$ y $\mathcal{B}'$, cuyos coeficientes vienen dados por

las relaciones

$$v_j = I_V(v_j) = \sum_{i=1}^{n} c_{ij} w_i \quad (j = 1, 2, \ldots, n).$$

Esto es, los elementos de la $j-$ésima columna de C son las coordenadas de v_j en la base $\mathcal{B}'$. Puesto que

$$I_V\left(\sum_{i=1}^{n} x_i v_i\right) = I_V(u) = u = \sum_{i=1}^{n} y_i w_i,$$

sigue de la teoría general que las coordenadas x_i de u en la base $\mathcal{B}$ y las coordenadas y_i de u en la base $\mathcal{B}'$ se relacionan por la fórmula

$$C \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

En consonancia con este hecho, C recibe el nombre de *matriz de cambio de base* (de $\mathcal{B}$ a $\mathcal{B}'$ en este caso). La notaremos $C(\mathcal{B}, \mathcal{B}')$.

Ejemplo 13.2.26 Determinemos en $\mathbb{R}^3$ las coordenadas de $u = (2, -5, 13)$ respecto de la base $\mathbb{E}_1 = \{(1, -1, 1), (2, -2, 1), (-4, 3, 0)\}$, cuyos elementos notaremos w_1, w_2 y w_3.

Puesto que conocemos el vector de coordenadas de u respecto de la base canónica $\mathbb{E}$ (el mismo u), bastará multiplicar éste por la matriz de cambio de base $C(\mathbb{E}, \mathbb{E}_1)$, cuyas columnas se obtienen expresando cada vector de la base canónica como combinación lineal de los w_i. Resolviendo en cada caso un sistema lineal de tres ecuaciones con tres incógnitas, arribamos a las igualdades

$$\begin{aligned} e_1 &= 3w_1 - 3w_2 - w_3 \\ e_2 &= 4w_1 - 4w_2 - w_3 \quad \text{y} \\ e_3 &= 2w_1 - w_2\,, \end{aligned}$$

como puede verificar el lector. Por lo tanto, el vector de coordenadas de u en la base $\mathbb{E}_1$ es

$$\begin{pmatrix} 3 & 4 & 2 \\ -3 & -4 & -1 \\ -1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ -5 \\ 13 \end{pmatrix} = \begin{pmatrix} 12 \\ 1 \\ 3 \end{pmatrix},$$

esto es, $u = 12w_1 + w_2 + 3w_3$. $\quad \Diamond$

Estableceremos en la siguiente proposición algunas fórmulas útiles referidas a las matrices de cambio de base y a las relaciones existentes entre matrices de una misma transfomación lineal (supondremos que $\mathcal{B}$ y $\mathcal{B}_1$ son bases de V y que $\mathcal{B}'$ y $\mathcal{B}'_1$ son bases de W).

Proposición 13.2.27 Valen las siguientes propiedades:

1) La matriz de cambio de base $C(\mathcal{B}, \mathcal{B}_1)$ es inversible. Precisamente,

$$C(\mathcal{B}, \mathcal{B}_1)^{-1} = C(\mathcal{B}_1, \mathcal{B})\,.$$

2) Si $f \in \mathrm{Hom}_K(V, W)$ entonces

$$M_{\mathcal{B}_1\mathcal{B}'_1}(f) = C(\mathcal{B}', \mathcal{B}'_1)\, M_{\mathcal{B}\mathcal{B}'}(f)\, C(\mathcal{B}_1, \mathcal{B})\,.$$

3) Si $f \in \mathrm{End}_K(V)$ entonces

$$M_{\mathcal{B}_1}(f) = C(\mathcal{B}_1, \mathcal{B})^{-1}\, M_{\mathcal{B}}(f)\, C(\mathcal{B}_1, \mathcal{B})\,.$$

DEMOSTRACION Aplicando la propiedad 4) de la proposición 13.2.25 al caso $V = W = U$, $\mathcal{B}' = \mathcal{B}_1$ y $\mathcal{B}'' = \mathcal{B}$ resulta que

$$I = M_{\mathcal{B}\mathcal{B}}(I_V) = M_{\mathcal{B}\mathcal{B}}(I_V \circ I_V) = M_{\mathcal{B}_1\mathcal{B}}(I_V)\, M_{\mathcal{B}\mathcal{B}_1}(I_V) = C(\mathcal{B}_1, \mathcal{B})\, C(\mathcal{B}, \mathcal{B}_1)$$

lo que claramente prueba 1).

Aplicando reiteradamente la misma propiedad también podemos demostrar 2), ya que escribiendo $f = I_W \circ f \circ I_V = I_W \circ (f \circ I_V)$ tenemos

$$\begin{aligned} M_{\mathcal{B}_1\mathcal{B}'_1}(f) &= M_{\mathcal{B}'\mathcal{B}'_1}(I_W)\, M_{\mathcal{B}_1\mathcal{B}'}(f \circ I_V) = C(\mathcal{B}', \mathcal{B}'_1)\, M_{\mathcal{B}\mathcal{B}'}(f)\, M_{\mathcal{B}_1\mathcal{B}}(I_V) = \\ &= C(\mathcal{B}', \mathcal{B}'_1)\, M_{\mathcal{B}\mathcal{B}'}(f)\, C(\mathcal{B}_1, \mathcal{B})\,. \end{aligned}$$

Notemos finalmente que la última fórmula es un caso particular de la segunda, tomando $W = V$, $\mathcal{B}' = \mathcal{B}$ y $\mathcal{B}'_1 = \mathcal{B}_1$. $\quad \Diamond$

Ejemplo 13.2.28 Si $\mathbb{E}_1$ es la base de $\mathbb{R}^3$ del ejemplo 13.2.26, calculemos la matriz respecto de $\mathbb{E}_1$ del endomorfismo g definido por

$$g\left((x_1, x_2, x_3)\right) = (x_1 + x_2 + x_3, 2x_1 - x_2 - 3x_3, x_2 + 5x_3).$$

Aplicando las propiedades 1) y 3) de la proposición anterior resulta que

$$M_{\mathbb{E}_1}(g) = C(\mathbb{E}_1, \mathbb{E})^{-1}\, M_{\mathbb{E}}(g)\, C(\mathbb{E}_1, \mathbb{E}) = C(\mathbb{E}, \mathbb{E}_1)\, M_{\mathbb{E}}(g)\, C(\mathbb{E}_1, \mathbb{E}).$$

Respecto a los tres últimos factores de arriba, es claro que

$$M_{\mathbb{E}}(g) = \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & -3 \\ 0 & 1 & 5 \end{pmatrix} \quad \text{y} \quad C(\mathbb{E}_1, \mathbb{E}) = \begin{pmatrix} 1 & 2 & -4 \\ -1 & -2 & 3 \\ 1 & 1 & 0 \end{pmatrix},$$

mientras que ya vimos en el ejemplo mencionado que

$$C(\mathbb{E}, \mathbb{E}_1) = \begin{pmatrix} 3 & 4 & 2 \\ -3 & -4 & -1 \\ -1 & -1 & 0 \end{pmatrix}.$$

Por lo tanto, la matriz que queríamos calcular es

$$\begin{pmatrix} 3 & 4 & 2 \\ -3 & -4 & -1 \\ -1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & -3 \\ 0 & 1 & 5 \end{pmatrix} \begin{pmatrix} 1 & 2 & -4 \\ -1 & -2 & 3 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 11 & 21 & -41 \\ -7 & -18 & 44 \\ -1 & -4 & 12 \end{pmatrix}. \quad \Diamond$$

Isomorfismos y matrices inversibles

La matriz de un homomorfismo no sólo es útil para determinar sus valores, sino que también encierra importante información acerca de algunas de sus características. El lector tendrá una prueba de ello en los enunciados que siguen a continuación.

Proposición 13.2.29 Si f es un homomorfismo de V en W y A es su matriz respecto de ciertas bases $\mathcal{B}$ y $\mathcal{B}'$ de V y W, son válidos los siguientes hechos:

1) $dim\,\mathrm{Im}(f) = rg(A)$.

2) Si $dim\,V = dim\,W$ entonces f es un isomorfismo si y solo si A es inversible.

Demostracion Recordemos en primer término que si $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$, la aplicación $\theta : V \to K^n$ definida por

$$\theta\left(\sum_{i=1}^{n} \lambda_i v_i\right) = (\lambda_1, \lambda_2, \ldots, \lambda_n),$$

que asigna a cada elemento de V su vector de coordenadas respecto de la base $\mathcal{B}$, es un isomorfismo. Para probar la primera afirmación del enunciado, caracterizaremos convenientemente el núcleo de f.

En general, dado $v \in V$, y designando por x su vector de coordenadas respecto de la base $\mathcal{B}$, sabemos que el vector de coordenadas y de $f(v)$ respecto de la base $\mathcal{B}'$ satisface la condición $Ax^t = y^t$, resultando en particular que $v \in \mathrm{Nu}(f)$ si y solo si $\theta(v)$ pertenece al subespacio S de soluciones del sistema lineal homogéneo $AX = 0$.

Aplicando el teorema de la dimensión, y teniendo en cuenta que la restricción de θ establece un isomorfismo entre el núcleo de f y S, obtenemos:

$$dim\,\mathrm{Im}(f) = n - dim\,\mathrm{Nu}(f) = n - dim\,S = n - (n - rg(A)) = rg(A),$$

como queríamos demostrar.

La validez de 2) es consecuencia de 1), a través de la siguiente cadena de equivalencias:

$$f \text{ es un isomorfismo } \Leftrightarrow f \text{ es un epimorfismo } \Leftrightarrow dim\,\mathrm{Im}(f) = n \Leftrightarrow$$
$$\Leftrightarrow rg(A) = n \Leftrightarrow A \text{ es inversible.} \quad \Diamond$$

NOTA A manera de resumen brindaremos a continuación una lista de enunciados equivalentes entre sí. Si bien todas las equivalencias han sido ya demostradas en estas páginas, dada su importancia queremos resaltarlas especialmente. Supondremos que h es un homomorfismo entre dos espacios vectoriales de la misma dimensión n y que C es la matriz de h respecto de ciertas bases del dominio y el codominio. En tal caso, las siguientes afirmaciones son equivalentes:

1. h es un isomorfismo.

2. $rg(C) = n$.

3. El sistema lineal $CX = b$ admite solución única para todo $b \in K^{n\times 1}$.

4. El sistema lineal $CX = b$ admite solución única para algún $b \in K^{n\times 1}$.

5. C es inversible. $\Diamond$

13.2.5. Ejercicios

1. *a*) Probar que AA^t y A^tA son matrices simétricas cualquiera sea la matriz A.

   b) Probar que los subespacios de matrices diagonales, escalares y triangulares de $M_n(K)$ son cerrados respecto del producto de matrices. ¿ Y el subespacio de matrices simétricas ?

2. El conjunto

$$C\left(M_n(K)\right) = \{A \in M_n(K) : AX = XA \text{ para todo } X \in M_n(K)\}$$

   se denomina el *centro* de $M_n(K)$. Probar las siguientes propiedades:

   a) $C\left(M_n(K)\right)$ es un subespacio de $M_n(K)$ cerrado respecto del producto de matrices.

   b) $C\left(M_n(K)\right)$ contiene al subespacio de matrices escalares.

   c) $A \in C\left(M_n(K)\right)$ si y solo si $AE^{ij} = E^{ij}A$ para todo $(i,j) \in \mathbb{I}_n^2$.

   d) $A \in C\left(M_n(K)\right)$ si y solo si A es escalar.

3. Si $A \in K^{m\times n}$ y $B \in K^{n\times s}$, probar que $t_{AB} = t_A \circ t_B$.

4. Si $A, B \in M_n(K)$, probar las siguientes propiedades:

   a) A es inversible si y solo si A^t lo es.
   b) AB es inversible si y solo si A y B lo son.
   c) $t_{I_n} = I_{K^n}$.
   d) A es inversible si y solo si t_A es un automorfismo de K^n.

5. Caracterizar en $M_n(K)$ los conjuntos de matrices diagonales, escalares y triangulares inversibles.

6. Probar que $rg(A) = rg(A^t)$ para toda matriz A.

7. Demostrar las siguientes propiedades:

   a) Si $A \in K^{m\times n}$ y $B \in K^{n\times s}$ entonces $rg(AB) \leq rg(A)$.
   b) Si $A \in K^{m\times n}$ y $B \in GL(n, K)$ entonces $rg(AB) = rg(A)$.
   c) Si $A \in GL(m, K)$ y $B \in K^{m\times n}$ entonces $rg(AB) = rg(B)$.

8. Determinar la forma escalonada reducida de las matrices

$$\begin{pmatrix} 0 & 2 & 4 & 2 \\ 0 & -1 & -2 & -1 \\ 0 & 3 & 6 & 6 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 0 & 1 & -1 & 3 & -2 \\ 2 & 2 & -1 & 0 & 5 \\ -4 & -3 & 1 & 3 & -12 \end{pmatrix}.$$

9. Resolver los siguientes sistemas de ecuaciones lineales:

   a) $\begin{cases} 5x_1 - 3x_3 + 7x_4 = 0 \\ 9x_1 - x_2 - 4x_3 + 13x_4 = 0 \\ x_1 - x_2 + x_3 + 2x_4 = 0 \end{cases}$

   b) $\begin{cases} 5x_1 - 3x_3 + 7x_4 = -2 \\ 9x_1 - x_2 - 4x_3 + 13x_4 = 3 \\ x_1 - x_2 + x_3 + 2x_4 = 2 \end{cases}$

   c) $\begin{cases} 2x_1 - 2x_2 + 2x_3 - x_4 + 2x_5 = 1 \\ x_1 - 3x_2 + x_3 + x_4 + x_5 = 0 \\ 4x_1 - 8x_2 + 4x_3 + x_4 + 4x_5 = 0 \end{cases}$

   d) $\begin{cases} x_1 - x_2 - 2x_4 = 4 \\ x_1 + x_2 + x_3 + x_4 = 2 \\ 2x_1 + x_3 - x_4 = 6 \end{cases}$

$e)$ $\begin{cases} 2x_1 + 2x_2 + 3x_3 = 0 \\ x_1 - 3x_2 + x_3 = 0 \\ x_1 + x_2 + x_3 = 0 \\ x_1 + 4x_3 = 0 \end{cases}$

$f)$ $\begin{cases} 2x_1 + 3x_2 = 0 \\ x_1 + x_2 = -1 \\ x_1 + 3x_2 = 3\,. \end{cases}$

10. Si $\alpha \in \mathbb{R}$, clasificar el sistema

$$\begin{cases} 2x_1 + x_3 = 0 \\ x_1 - \alpha x_2 + x_3 = -2 \\ 2x_1 + 8\alpha x_2 + 3\alpha x_3 = 6 \end{cases}$$

según los valores de α.

11. Sea V un K-espacio vectorial de dimensión m y sea $\mathcal{F}$ una familia de n vectores de V $(m < n)$. Probar usando el corolario 13.2.16 que $\mathcal{F}$ es linealmente dependiente.

12. Determinar cuáles de las siguientes matrices son inversibles y calcular (cuando existan) sus inversas:

$a)$ $\begin{pmatrix} -2 & -4 & 0 \\ 3 & 1 & -7 \\ 8 & 6 & -2 \end{pmatrix}$ $\quad b)$ $\begin{pmatrix} -1 & -2 & 9 \\ 1 & -1 & 2 \\ 3 & 0 & -5 \end{pmatrix}$

$c)$ $\begin{pmatrix} 0 & 2 & 4 & 1 \\ 0 & 1 & 2 & 0 \\ -3 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$ $\quad d)$ $\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \end{pmatrix}$.

13. En los dos siguientes casos determinar los $\alpha \in \mathbb{R}$ para los cuales la matriz dada resulta inversible:

$a)$ $\begin{pmatrix} 2 & \alpha+8 & 1 \\ 0 & -1 & \alpha \\ 1 & 2 & 3 \end{pmatrix}$ $\quad$ $b)$ $\begin{pmatrix} 1 & \alpha-1 & 1 \\ -3 & 2 & 1 \\ \alpha & -1 & 0 \end{pmatrix}$

14. $a)$ Demostrar que una matriz cuadrada es inversible si y solo si es producto de matrices elementales.

    $b)$ Expresar las matrices inversibles del ejercicio 12 como producto de matrices elementales.

15. En cada uno de los siguientes casos hallar la matriz del homomorfismo f respecto de las bases $\mathcal{B}$ y $\mathcal{B}'$ dadas:

    $a)$ $f : \mathbb{R}^2 \to \mathbb{R}^3$; $f((x_1, x_2)) = (x_1 - x_2, x_1 + 2x_2, x_2)$

    i) $\mathcal{B}$ y $\mathcal{B}'$ las bases canónicas
    ii) $\mathcal{B} = \{(1,-1),(2,2)\}$, $\mathcal{B}' = \{(1,1,0),(2,-3,1),(0,5,1)\}$
    iii) $\mathcal{B} = \{(2,2),(1,-1)\}$, $\mathcal{B}' = \{(2,-3,1),(0,5,1),(1,1,0)\}$.

    $b)$ $f : \mathbb{R}^{2\times 3} \to \mathbb{R}^{3\times 2}$; $f(A) = A^t$

    $\mathcal{B}$ y $\mathcal{B}'$ las bases canónicas (ordenadas lexicográficamente)

    $c)$ $f : \mathbb{C}^2 \longrightarrow C^2$; $f((z_1, z_2)) = (z_1 + 2\imath z_2, \imath z_1 - z_2)$

    i) $\mathcal{B}$ y $\mathcal{B}'$ la base canónica de $\mathbb{C}^2$
    ii) $\mathcal{B} = \mathcal{B}' = \{(1,0),(\imath,0),(0,1),(0,\imath)\}$ (considerando a $\mathbb{C}^2$ como $\mathbb{R}-$espacio vectorial).

    $d)$ $f : \mathbb{R}_3[X] \to \mathbb{R}_3[X]$; $f(h) = h'$

    i) $\mathcal{B}$ y $\mathcal{B}'$ las bases canónicas
    ii) $\mathcal{B} = \left\{1, X^3, X^2, X\right\}$, $\mathcal{B}' = \left\{2, X+1, X^2-1, X^3+1, X^4\right\}$.

16. Sea $f \in \mathrm{Hom}(\mathbb{R}^3, \mathbb{R}^4)$ y sean $\mathcal{B} = \{v_1, v_2, v_3\}$ y $\mathcal{B}' = \{w_1, w_2, w_3, w_3\}$ bases de $\mathbb{R}^3$ y $\mathbb{R}^4$ tales que

$$M_{\mathcal{B}\mathcal{B}'}(f) = \begin{pmatrix} 2 & 1 & 2 \\ -3 & 2 & -5 \\ -1 & 1 & -1 \\ 1 & -2 & 1 \end{pmatrix}.$$

a) Hallar las coordenadas respecto de la base $\mathcal{B}'$ de $f(v_1 - v_2 + 2v_3)$.

b) Hallar bases de $\mathrm{Nu}(f)$ y $\mathrm{Im}(f)$.

c) Caracterizar el conjunto $\{v \in V : f(v) = w_1 - w_2 + 2w_3 - w_4\}$.

d) Decidir si existen bases $\mathcal{B}_1$ y $\mathcal{B}'_1$ bases de $\mathbb{R}^3$ y $\mathbb{R}^4$ tales que

$$M_{\mathcal{B}_1\mathcal{B}'_1}(f) = \begin{pmatrix} 2 & 1 & 0 \\ 3 & 6 & 3 \\ -1 & 4 & 3 \\ -5 & 11 & 3 \end{pmatrix}.$$

17. Si $A \in M_n(K)$ y $B \in GL(n, K)$, probar que $rg(BAB^{-1}) = rg(A)$.

18. Sea $f : V \to W$ un isomorfismo y sean $\mathcal{B}$ y $\mathcal{B}'$ bases de V y W, respectivamente. Probar que

$$M_{\mathcal{B}'\mathcal{B}}(f^{-1}) = (M_{\mathcal{B}\mathcal{B}'}(f))^{-1} .$$

19. Sea $f : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $f((x_1, x_2)) = (15x_1 - 24x_2, 8x_1 - 13x_2)$. Hallar una base $\mathcal{B}$ de $\mathbb{R}^2$ tal que $M_{\mathcal{B}}(f)$ sea diagonal.

20. Sea f un endomorfismo de un espacio vectorial V de dimensión n. Probar que f es un proyector si y solo si existe una base $\mathcal{B}$ de V tal que

$$(M_{\mathcal{B}}(f))_{kj} = \begin{cases} 1 & \text{si } k \leq m \wedge j = k \\ 0 & \text{en otro caso,} \end{cases}$$

donde m es un entero no negativo menor o igual que n.

Capítulo 14

Nociones de Algebra abstracta

14.1. Estructuras algebraicas

14.1.1. Introducción

A lo largo de capítulos anteriores hemos definido y empleado conceptos tales como grupos, anillos, cuerpos, etc., para referirnos en todos los casos a conjuntos dotados de una o más operaciones que satisfacen ciertas propiedades específicas. Tales conjuntos, que con su correspondiente sistema axiomático responden al nombre genérico de *estructuras algebraicas*, son el principal objeto de estudio de la llamada originariamente Algebra moderna o abstracta, denominaciones ya caídas en desuso.

Los comienzos de la disciplina datan aproximadamente de la última mitad del siglo XVIII, con la aparición en diversos campos de estructuras de grupo, que dieron lugar más tarde al desarrollo de una teoría general y abstracta, actualmente muy vasta. Podemos mencionar por ejemplo la aparición de los grupos en la Aritmética, donde Euler (1760) introdujo la noción de clase de congruencia, en la teoría de ecuaciones algebraicas, con los trabajos de Galois, Lagrange y Abel sobre los grupos de permutaciones de las raíces de un polinomio, y en la Geometría, donde Mőebius (1830) comenzó a referirse a grupos de transformaciones actuando sobre objetos geométricos. A la fecha, los grupos y muchas otras estructuras algebraicas son una herramienta indispensable del Algebra y de la Matemática en general, y su estudio sistemático forma parte del entrenamiento básico de los estudiantes de esta ciencia y otras afines.

En este capítulo ofreceremos un panorama de las principales estructuras algebraicas abstractas, destacando algunas de sus características esenciales y exhibiendo un buen número de ejemplos. El lector debiera interpretar estas páginas como una especie de diccionario introductorio al lenguaje de la teoría, con miras a un ulterior estudio más minucioso de la misma.

14.1.2. Estructuras algebraicas básicas

Brindaremos a partir de aquí una lista de las principales estructuras algebraicas, algunas de las cuales ya han sido tratadas en este libro. Salvo mención expresa en otro sentido, las operaciones consideradas de aquí en más serán operaciones binarias.

Semigrupos

> Sea S un conjunto y sea $(x, y) \mapsto x * y$ una operación asociativa en S admitiendo un elemento neutro e. Diremos entonces que S es un *semigrupo* respecto a dicha operación, o también que $(S, *)$ es un semigrupo. Si además la operación es conmutativa, $(S, *)$ se dirá un semigrupo conmutativo.

Ejemplos 14.1.1 Ejemplos conocidos de semigrupos son el conjunto $\mathbb{N}_0$ de enteros no negativos, respecto a la suma usual, y el conjunto $\mathbb{Z}$ de números enteros respecto al producto usual, siendo 0 y 1 los correspondientes elementos neutros. Menos familiar, la operación $a * b = \max\{a, b\}$ también define en $\mathbb{N}_0$ una estructura de semigrupo, ya que claramente es asociativa y admite 0 como elemento neutro. Vemos así un primer caso de dos estructuras distintas de semigrupo definidas sobre el mismo conjunto.

Para exhibir un ejemplo de semigrupo no conmutativo, tomemos un número natural n y consideremos el conjunto $M_n(\mathbb{R})$ de matrices cuadradas de orden n con coeficientes reales. Resulta entonces que $M_n(\mathbb{R})$ es un semigrupo respecto al producto usual de matrices, ya que éste es asociativo y admite la matriz identidad I_n como elemento neutro, resultando conmutativo sólo en el caso trivial $n = 1$.

Dado un conjunto no vacío X, otro ejemplo importante es el del conjunto X^X de funciones de X en X, que tiene estructura de semigrupo respecto a la composición de funciones, ya que hemos visto en el capítulo 1 que dicha operación es asociativa y admite la función identidad I_X como elemento neutro.

Por último, sea $\mathcal{P}$ el conjunto de secuencias finitas de elementos de un conjunto X cualquiera. Vale decir, los elementos de $\mathcal{P}$ son de la forma

$$\alpha = x_1 x_2 \dots x_m \,,$$

donde los x_i son elementos de X y $m \geq 0$ (el caso $m = 0$ corresponde a la secuencia vacía ϕ). Si definimos una operación entre secuencias por simple yuxtaposición de las mismas, en la forma

$$x_1 x_2 \dots x_r * y_1 y_2 \dots y_s = x_1 x_2 \dots x_r y_1 y_2 \dots y_s \,,$$

sigue fácilmente que $\mathcal{P}$ es un semigrupo respecto a esta operación, con ϕ como elemento neutro. Es claro además que $(\mathcal{P}, *)$ no es conmutativo si $\#(X) > 1$, ya que tomando dos elementos distintos a, b de X resulta

$$a * b = ab \neq ba = b * a. \quad \Diamond$$

Grupos

Si G es un semigrupo y todo elemento de G es inversible diremos que G es un *grupo*.

Detallando las condiciones de la definición, resulta que $(G, *)$ es un grupo si y sólo si se satisfacen los siguientes axiomas:

(G_1) $x * (y * z) = (x * y) * z$ cualesquiera sean $x, y, z \in G$

(G_2) Existe $e \in G$ tal que $x * e = e * x = x$ para todo $x \in G$

(G_3) Para todo $x \in G$ existe $y \in G$ tal que $x * y = y * x = e$.

Si la operacion es conmutativa se dice que G es un grupo conmutativo o *abeliano*, en homenaje al noruego Niels Henrik Abel (1802-1827), que demostró la imposibilidad de resolver por radicales la ecuación algebraica general de quinto grado, mediante el estudio del grupo de permutaciones de sus raíces.

Un grupo G se dice *finito* o *infinito* según sea finito o infinito el conjunto G. En general, el cardinal de G se llama el *orden* de G, y lo notaremos $|\, G \,|$. Digamos de paso que todo conjunto unitario $X = \{\, x \,\}$ admite una única estructura de grupo, definida por la operación $x * x = x$. Se lo llama *grupo trivial*.

Antes de pasar a ejemplos más estimulantes, acordemos algunas cuestiones formales. Generalmente, cuando nos refiramos a un grupo abstracto G indicaremos la operación en la forma multiplicativa $x \,.\, y$, o más simplemente, en la forma xy. Por abuso de lenguaje, a veces llamaremos "producto" a la operación en un grupo abstracto. Ocasionalmente emplearemos notación aditiva, con el símbolo $+$ de la suma para designar la operación en G, pero sólo en el caso de que G sea abeliano. Si la notación es multiplicativa, designaremos por 1 el elemento neutro y por x^{-1} el inverso de x, cuya existencia asegura el axioma (G_3). En cambio, si la notación es aditiva usaremos el símbolo 0 para el elemento neutro y designaremos el inverso de x por $-x$.

Ejemplos 14.1.2 La siguiente es una lista de grupos, algunos de los cuales han sido ya descriptos en páginas precedentes. En varios de los casos dejamos a cargo del lector la tarea de probar en detalle que la operación dada define una estructura de grupo en el correspondiente conjunto.

1) Los conjuntos numéricos $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$ tienen estructura de grupo respecto a la suma usual, mientras que $\mathbb{Q}^*$, $\mathbb{R}^*$ y $\mathbb{C}^*$ son grupos respecto al producto usual. El lector podrá verificar sin dificultad que el conjunto $\mathbb{R}_{>0}$ de números reales positivos y la circunferencia unitaria $\mathbb{S}^1$ (conjunto de números complejos de módulo 1) también tienen estructura de grupo respecto al producto.

2) Como ya sabemos, todo espacio vectorial es un grupo abeliano con respecto a la suma de vectores. Se agregan entonces a nuestros ejemplos de grupo los espacios de $n-$uplas, los espacios de matrices y los espacios de homomorfismos entre dos espacios vectoriales.

3) Si $n \in \mathbb{N}$, el conjunto $\mathbb{Z}_n$ de clases de congruencia módulo n es un grupo respecto a la suma de clases, como demostramos en la proposición 6.1.4, mientras que el conjunto G_n de raíces $n-$ésimas de 1 es un grupo respecto al producto en $\mathbb{C}$, hecho que se deduce de la proposición 8.3.5. Más adelante veremos que estas estructuras de grupo son esencialmente iguales, a pesar de que los conjuntos que las albergan son distintos.

   El próximo ejemplo servirá como fuente de ejemplos de grupos.

4) Sea S un semigrupo (notado multiplicativamente) y sea $\mathcal{U}(S)$ el conjunto de elementos inversibles de S. Entonces $\mathcal{U}(S)$ es un grupo.

   Para demostrar nuestra afirmación, debemos probar primero que la restricción del producto a $\mathcal{U}(S)$ determina una operación en $\mathcal{U}(S)$, o dicho en otras palabras, que el producto de elementos inversibles es inversible. Para ello, consideremos elementos inversibles u y v y sean u^{-1} y v^{-1} sus inversos. Designando por e el elemento neutro resulta entonces

$$\begin{aligned}(uv)(v^{-1}u^{-1}) = ((uv)v^{-1})u^{-1} = (u(vv^{-1}))u^{-1} = \\ = (ue)u^{-1} = uu^{-1} = e,\end{aligned}$$

   lo que prueba que $v^{-1}u^{-1}$ es un inverso a derecha de de uv. Procediendo en forma análoga obtenemos también que $(v^{-1}u^{-1})(uv) = e$, de donde concluimos que uv es inversible y $v^{-1}u^{-1}$ es su inverso. Por lo tanto, la restricción del producto es una operación binaria en $\mathcal{U}(S)$, que obviamente es asociativa y admite elemento neutro, ya que $e \in \mathcal{U}(S)$.

   Resta luego probar que todo elemento de a de $\mathcal{U}(S)$ tiene inverso en $\mathcal{U}(S)$. Ahora bien, a admite por definición un inverso a^{-1} en S, resultando por obvias razones de simetría que a es a su vez el inverso de a^{-1}. Luego $a^{-1} \in \mathcal{U}(S)$ y nuestra demostración concluye.

   Veamos algunos casos particulares de esta situación:

   a) Si $n \in \mathbb{N}$, el grupo lineal general $GL(n, \mathbb{R})$ es el grupo de elementos inversibles del semigrupo $M_n(\mathbb{R})$. Como vimos en el capítulo 13, sus elementos admiten la caracterización:

$$A \in GL(n, \mathbb{R}) \Leftrightarrow rg(A) = n\,.$$

   b) Si $S = X^X$, donde X es un conjunto no vacío, $\mathcal{U}(S)$ es el conjunto de biyecciones o permutaciones de X, esto es, las funciones inversibles respecto a la composición. Se lo denomina *grupo simétrico* del conjunto X, y lo notaremos $\mathbb{S}_X$.

Si $n \in \mathbb{N}$, el grupo de permutaciones de $\mathbb{I}_n$ se acostumbra a notar más sencillamente $\mathbb{S}_n$, y sus elementos suelen representarse en la forma

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & i & \dots & n \\ x_1 & x_2 & \dots & x_i & \dots & x_n \end{pmatrix},$$

donde la notación indica que $\sigma(i) = x_i$, respondiendo a la idea de que una permutación de n elementos está dada por un determinado ordenamiento de los mismos. Por ejemplo, si $n = 4$ el esquema

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

representa a la permutación φ definida por $\varphi(1) = 4$, $\varphi(2) = 1$, $\varphi(3) = 3$ y $\varphi(4) = 2$. A manera de ilustración, el lector podrá comprobar la validez de las siguientes igualdades en $\mathbb{S}_5$:

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

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 4 & 2 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 5 & 3 & 1 \end{pmatrix}.$$

c) Además de ser un grupo respecto a la suma de clases, vimos que $\mathbb{Z}_n$ $(n \in \mathbb{N})$ admite una estructura de semigrupo respecto al producto de las mismas. Por lo tanto, el conjunto de elementos inversibles para dicho producto es un grupo, que notaremos Z_n^*. De acuerdo con la proposición 6.2.4 podemos describirlo en la siguiente forma (recordemos que (x) indica la clase de x módulo n):

$$(a) \in Z_n^* \Leftrightarrow (a : n) = 1\,.$$

Por ejemplo, $Z_{18}^* = \{1, 5, 7, 11, 13, 17\}$ y $Z_p^* = Z_p - \{0\}$ si p es primo.

Los grupos que hemos considerado en los ejemplos anteriores son conmutativos, con excepción de los grupos de matrices y de permutaciones del ejemplo 4), que sólo resultan abelianos en los casos $n = 1$ y $\#(X) \leq 2$, respectivamente.

En cuanto al orden, $\mathbb{Z}_n$ y G_n son grupos finitos de orden n, y también son finitos $\mathbb{S}_n$ y $\mathbb{Z}_n^*$, de órdenes $n!$ y $\varphi(n)$, respectivamente. Todos los demás grupos exhibidos son infinitos. $\Diamond$

Subgrupos

Sea G un grupo y sea S un subconjunto de G. Decimos que S es un *subgrupo* de G si y solo si se satisfacen las siguientes condiciones (las letras designan elementos de G):

(SG_1) $1 \in S$

(SG_2) $x, y \in S \Rightarrow xy \in S$

(SG_3) $x \in S \Rightarrow x^{-1} \in S\,.$

Observemos que la condición (SG_2) afirma que la restrición de la operación de grupo a S define una operación en S, obviamente asociativa, mientras que las condiciones (SG_1) y (SG_3) indican que la misma define una estructura de grupo en S. Por lo tanto, podemos interpretar la noción de subgrupo de la siguiente forma: es un grupo S incluido en un grupo G y la operación en S está definida por restricción de la operación en G. Por ejemplo, $\mathbb{Z}$ es un subgrupo de $\mathbb{Q}$ (respecto a la suma), mientras que $\mathbb{Q}^*$ no es subgrupo de $\mathbb{R}$, a pesar de existir una relación de inclusión entre los conjuntos subyacentes, ya que el primero es grupo respecto al producto y el segundo lo es respecto a la suma.

Ejemplos 14.1.3 Comencemos con algunos ejemplos de carácter general.

1) Dado un grupo G, es inmediato probar que $\{1\}$ y G son subgrupos de G, llamados los subgrupos *triviales* de G. Más interesante es el caso del subconjunto

$$S = \{a \in G \,:\, ax = xa \text{ para todo } x \in G\}\,,$$

esto es, el conjunto de elementos que conmutan con todos los elementos del grupo. Dejamos al lector la tarea de demostrar que S es un subgrupo de G, que llamaremos el *centro* de G y notaremos $\mathcal{C}(G)$. Observemos que $\mathcal{C}(G)$ es un subgrupo abeliano de G y que G es abeliano si y solo si $\mathcal{C}(G) = G$.

2) Si G es un grupo y $g \in G$, sea

$$T = \{\, g^m : m \in \mathbb{Z}\,\}\,,$$

donde las potencias no negativas se definen inductivamente en la forma

$$g^k = \begin{cases} 1 & \text{si } k = 0 \\ g g^{k-1} & \text{si } k > 0\,, \end{cases}$$

mientras que $g^{-k} = \left(g^{-1}\right)^k$ si $k > 0\,.$

El lector verificará que si m es positivo entonces $g^m = g\, g\, \ldots\, g$, donde el producto contiene m factores. Puesto que dados $m, n \in \mathbb{Z}$ es fácil

demostrar la validez de las fórmulas $g^{m+n} = g^m g^n$ y $(g^m)^{-1} = g^{-m}$, sigue que T es un subgrupo de G, que llamaremos el subgrupo *cíclico* generado por g. Lo notaremos $\langle\, g \,\rangle$, y tiene la propiedad de ser el menor subgrupo de G (en el sentido de la inclusión) que contiene a g. Subrayemos que en el caso de notación aditiva g^m se indica mg.

3) Ampliando un ejemplo anterior, señalemos que cada miembro de la cadena de grupos

$$\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$$

es subgrupo de todos los subsiguientes, y lo mismo ocurre en la cadena de grupos

$$G_2 \subset \mathbb{Q}^* \subset \mathbb{R}^* \subset \mathbb{C}^* .$$

Notemos asimismo que $\mathbb{R}_{>0}$ es un subgrupo de $\mathbb{R}^*$ y que G_n es subgrupo de $\mathbb{S}^1$ cualquiera sea $n \in \mathbb{N}$.

4) Como sabemos, son subgrupos de la estructura aditiva de $\mathbb{R}^{m\times n}$ los subconjuntos de matrices simétricas, antisimétricas, triangulares (superiores e inferiores), diagonales y escalares. Respecto a la estructura multiplicativa en $GL(n, \mathbb{R})$, mencionemos los subgrupos de matrices diagonales y escalares inversibles.

Un ejemplo de subgrupo en el grupo de permutaciones de un conjunto no vacío X es el subconjunto

$$H = \{\, \sigma \in \mathbb{S}_X \,:\, \sigma(y) = y \text{ para todo } y \in A \,\},$$

donde $A \subseteq X$. Si A y X son finitos, digamos de cardinales m y n, respectivamente, sigue inmediatamente que $|\, H \,| = (n - m)!$

5) Examinando las correspondientes tablas de operaciones, se comprueba sin dificultad que el conjunto

$$\{(0), (8), (16), (24), (4), (12), (20)\}$$

es un subgrupo de $\mathbb{Z}_{28}$ y el conjunto

$$\{(1), (5), (21), (25), (27), (31), (47), (51)\}$$

es un subgrupo de $\mathbb{Z}_{52}^*$.

Hablando más generalmente, señalemos que $\mathbb{Z}_n\, (n \in \mathbb{N})$ admite un subgrupo de orden d para cada divisor d de n, a saber el subgrupo cíclico generado por (n/d). Se obtiene una versión multiplicativa de este último hecho aplicando la propiedad 4) de la proposición 8.3.5, ya que a partir de ella deducimos que G_d es un subgrupo de G_n.

Encargamos al lector la tarea de verificar todas las afirmaciones. $\Diamond$

Anillos

Un *anillo* es un conjunto R dotado de dos operaciones, llamadas suma y producto y notadas con los símbolos usuales, que satisfacen los siguientes requerimientos (las letras indican elementos de R):

(A_1) $(R, +)$ es un grupo abeliano

(A_2) $(R, .)$ es un semigrupo

(A_3) $a(b + c) = ab + ac$ y $(b + c)a = ba + ca$.

Como de costumbre, designamos por 0 y 1 los respectivos elementos neutros, que suponemos distintos. Las condiciones (A_3) se llaman propiedades distributivas del producto respecto a la suma. Si el producto en R es conmutativo diremos que R es un anillo conmutativo. Asimismo, supongamos que para cualquier par de elementos $x, y \in R$ se verifica la siguiente propiedad:

$$xy = 0 \iff x = 0 \vee y = 0 .$$

Diremos entonces que R es un anillo *íntegro*. Si además es conmutativo, se dirá un *dominio de integridad* (el lector probará sin mayor dificultad que la parte "si" del bicondicional anterior es válida en cualquier anillo).

Ejemplos 14.1.4 Veamos varios ejemplos ilustrativos:

1) Los conjuntos numéricos $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$ son dominios de integridad respecto a la suma y producto usuales. Otro ejemplo, de gran interés en Aritmética, es el anillo de *enteros de Gauss*, a saber:

$$\mathbb{Z}[\imath] = \{ a + b\imath \,:\, a, b \in \mathbb{Z} \} ,$$

que también es un dominio de integridad respecto a la suma y el producto de números complejos.

2) Existen muchas formas de construir otros anillos a partir de un anillo R. Como ejemplos importantes citaremos aquí el anillo de polinomios $R[X]$ con coeficientes en R, que ya hemos definido en el capítulo 9, y el anillo $M_n(R)$ de matrices cuadradas de orden n con coeficientes en $R\,(n \in \mathbb{N})$, con las operaciones de suma y producto usuales de matrices.

   En el caso de los polinomios, es fácil probar que $R[X]$ es conmutativo (íntegro) si y solo si R es conmutativo (íntegro), mientras que $M_n(R)$ no es conmutativo ni íntegro si $n > 1$ (el lector recordará cómo multiplican entre sí los elementos de la base canónica).

3) Si $n \in \mathbb{N}$ tenemos el conjunto $\mathbb{Z}_n$ de clases de congruencia módulo n, que es un anillo finito y conmutativo con las operaciones de suma y producto de clases. Resulta un sencillo ejercicio de Aritmética probar que $\mathbb{Z}_n$ es un dominio de integridad si y solo si n es un número primo.

4) Si X es un conjunto no vacío, el conjunto de partes $\mathbb{P}(X)$ es un anillo conmutativo respecto a las operaciones de diferencia simétrica (suma) e intersección (producto), como se desprende de las propiedades que hemos establecido en la primera sección del capítulo 1. Resulta en este caso que $\emptyset$ y X son los correspondientes elementos neutros y que $\mathbb{P}(X)$ no es un dominio de integridad si $\#(X) > 1$, ya que si U es cualquier subconjunto unitario de X tenemos:

$$U\,U^c = U \cap U^c = \emptyset = 0,$$

a pesar de que los "factores" U y U^c son no nulos (no vacíos).

Con notación algo más abstracta, observemos que dado cualquier elemento x del anillo valen las igualdades $x^2 = x$ y $2x = 0$. En general, un anillo cuyos elementos satisfacen estas relaciones (en realidad la segunda de deduce de la primera) se llama un *anillo de Boole*. $\Diamond$

Subanillos

Un subconjunto S de un anillo R se dice un *subanillo* de R si y solo si se satisfacen las siguientes condiciones:

(SA_1) S es un subgrupo de la estructura aditiva de R

(SA_2) S es cerrado para el producto

(SA_3) $1 \in S$.

En otras palabras, S es cerrado para la suma y el producto, contiene ambos elementos neutros y los inversos aditivos de todos sus elementos. Equivalentemente, S es un anillo incluido en R, con las operaciones definidas por restricción.

Por ejemplo, y como en el caso de grupos, cada anillo de la cadena

$$\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$$

es subanillo de los que le siguen, y vale la misma relación entre los correspondientes anillos de polinomios, en la cadena $\mathbb{Z}[X] \subset \mathbb{Q}[X] \subset \ldots$ Respecto a los anillos de polinomios, es inmediato verificar que el conjunto de polinomios constantes es un subanillo de $R[X]$, cualquiera sea el anillo R.

Mencionemos en el anillo de matrices $M_n(R)$ los subanillos de matrices escalares, diagonales y triangulares, y en $\mathbb{P}(X)$ el subanillo finito $\{\,\emptyset, X\,\}$.

A manera de ejemplos genéricos de subanillos de un anillo R anotemos que R mismo lo es, así como también el *centro* de R, a saber, el conjunto

$$\mathcal{C}(R) = \{a \in A : ax = xa \text{ para todo } x \in R\},$$

que claramente es un anillo conmutativo. Por último, sea $b \in R$ y consideremos el conjunto S de expresiones polinomiales en b con coeficientes enteros,

es decir,

$$S = \left\{ \sum_{i=0}^{n} m_i b^i \right\},$$

donde n es un entero no negativo (variable) y los m_i son números enteros. Es fácil ver que S es un subanillo de R (encargamos al lector los detalles de la demostración), notado $\mathbb{Z}[\,b\,]$ y llamado el subanillo de R generado por b, debido al hecho de ser el menor subanillo de R (en el sentido de la inclusión) que contiene a b. Por ejemplo, el subanillo de $\mathbb{C}$ generado por $\imath$ es el anillo de enteros de Gauss. $\Diamond$

NOTA Existe otro tipo importante de subestructura en un anillo, llamada ideal.

> Precisamente, dado un anillo R y un subconjunto $\mathcal{I}$ de R, diremos que $\mathcal{I}$ es un *ideal a izquierda* (*a derecha*) de R si y solo si se satisfacen las siguientes condiciones:

(ID_1) $\mathcal{I}$ es un subgrupo de la estructura aditiva de R

(ID_2) $ax \in \mathcal{I}$ ($xa \in \mathcal{I}$) para todo $a \in R$ y para todo $x \in \mathcal{I}$.

Si $\mathcal{I}$ es un ideal a izquierda y a derecha decimos que $\mathcal{I}$ es un ideal *bilátero*. Por ejemplo, $\{\,0\}$ y R son ideales biláteros en cualquier anillo R. Obviamente todas estas nociones coinciden si R es conmutativo, en cuyo caso hablamos simplemente de ideales de R. Por ejemplo, el conjunto

$$n\,\mathbb{Z} = \{\, kn \,:\, k \in \mathbb{Z} \,\}$$

de múltiplos enteros de n es un ideal de $\mathbb{Z}$ cualquiera sea $n \in \mathbb{Z}$. $\Diamond$

Cuerpos

> Un anillo conmutativo K se dice un *cuerpo* si y solo si todo elemento no nulo de K es inversible respecto al producto. Equivalentemente, el conjunto $K^* = K - \{\,0\}$ es un grupo respecto al producto.

Casos conocidos por el lector de esta estructura son naturalmente los cuerpos numéricos $\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$. Como ejemplo de otro tipo, recordemos del capítulo 6 que si p es un primo toda clase de congruencia no nula módulo p es inversible respecto al producto, lo que brinda una colección de ejemplos de cuerpos finitos, a saber los cuerpos $\mathbb{Z}_p$ de clases residuales.

Subcuerpos

> Dado un cuerpo K y un subconjunto F de K, decimos que F es un *subcuerpo* de K si y solo si se satisfacen las siguientes condiciones:

(SC_1) F es un subanillo de K
(SC_2) $x^{-1} \in F$ para todo $x \in F - \{0\}$.

Vemos que aparece nuevamente la misma idea: un cuerpo F es un subcuerpo de un cuerpo K si y solo si sus operaciones están definidas por restricción de las correspondientes operaciones de K. Por ejemplo, $\mathbb{Q}$ es un subcuerpo de $\mathbb{R}$ y de $\mathbb{C}$, mientras que $\mathbb{R}$ es un subcuerpo de $\mathbb{C}$.

Más novedosamente, si m es un número natural que no es un cuadrado perfecto probaremos que el conjunto F de números reales de la forma

$$x = a + b\sqrt{m} \quad (a, b \in \mathbb{Q}) \tag{14.1}$$

es un subcuerpo de $\mathbb{R}$. Notemos por ejemplo que $\sqrt{m} = 0 + 1\sqrt{m} \in F$ y que $\mathbb{Q} \subset F$, ya que si $r \in \mathbb{Q}$ podemos escribir $r = r + 0\sqrt{m}$. Previo a comprobar que F satisface las condiciones de la definición, probaremos que cada elemento de F se expresa de una única manera en la forma (14.1). Para ello, supongamos que

$$x = a + b\sqrt{m} = c + d\sqrt{m} \quad (a, b, c, d \in \mathbb{Q})$$

son dos formas posibles de escribir un elemento de F, de donde sigue que

$$(b - d)\sqrt{m} = c - a. \tag{14.2}$$

Puesto que en general el producto de un racional $s \neq 0$ por un irracional t es irracional (ya que de otro modo tendríamos $t = sts^{-1} \in \mathbb{Q}$), deducimos de (14.2) que $b - d = c - a = 0$, y por lo tanto hay un único par (a, b) de números racionales asociado al elemento x.

Si $u = a + b\sqrt{m}$ y $v = c + d\sqrt{m}$ son elementos genéricos de F, sigue de las igualdades

$$\begin{aligned} u + v &= (a + c) + (b + d)\sqrt{m}\,, \\ uv &= (ac + mbd) + (ad + bc)\sqrt{m}\,, \\ -u &= (-a) + (-b)\sqrt{m}\,, \end{aligned}$$

que $u+v$, uv y $-u$ también son elementos de F, pues todos los números que aparecen entre paréntesis son racionales. Puesto que además 0 y 1 pertenecen a F, por ser racionales, concluimos que F es un subanillo de $\mathbb{R}$.

Mostraremos finalmente que F es un cuerpo, probando que el inverso multiplicativo de un elemento no nulo $a+b\sqrt{m}$ de F admite una expresión de tipo (14.1). En efecto, sigue de la igualdad $(a+b\sqrt{m})(a-b\sqrt{m}) = a^2 - mb^2$ que

$$(a + b\sqrt{m})^{-1} = \frac{a - b\sqrt{m}}{a^2 - mb^2} = \left(\frac{a}{a^2 - mb^2}\right) + \left(\frac{-b}{a^2 - mb^2}\right)\sqrt{m} \in F\,,$$

pues los coeficientes de esta última expresión son racionales. Notemos que $a - b\sqrt{m}$, y por ende $a^2 - mb^2$, son distintos de 0, ya que en caso contrario resultaría $a = b = 0$, por la unicidad en la escritura de los elementos de F.

No es difícil demostrar que F es el menor subcuerpo de $\mathbb{R}$ que contiene a $\mathbb{Q}$ y a la raíz cuadrada de m. Debido a ello se lo nota $\mathbb{Q}(\sqrt{m})$. $\quad \Diamond$

Algebras

Sea A un anillo y supongamos además que A es un espacio vectorial sobre un cuerpo K, con la estructura aditiva dada por la suma del anillo.

> Diremos que A es un *álgebra* sobre K, o que A es una $K-$álgebra, si y solo si el producto por escalares y el producto en A están ligados por las relaciones
>
> $$(\lambda\, a)\, b = \lambda\, (a\, b) = a\, (\lambda\, b)\,,$$
>
> cualesquieran sean $\lambda \in K$ y $a, b \in A$.

Una $K-$álgebra A se dice conmutativa o no conmutativa según que el producto en A sea conmutativo o no conmutativo. Notemos que en cualquier caso las condiciones de la definición aseguran que los elementos de la forma $\lambda 1$ pertenecen al centro de A para todo $\lambda \in K$ (1 denota aquí el elemento neutro del producto en A), ya que si $x \in A$ tenemos:

$$(\lambda\, 1)\, x = \lambda\, (1\, x) = \lambda\, (x\, 1) = x\, (\lambda\, 1)\,.$$

Como ejemplos típicos de $K-$álgebras mencionemos el álgebra de polinomios $K[X]$, el álgebra $M_n(K)$ de matrices de orden n con coeficientes en K y el álgebra $\mathrm{End}_K(K^n)$ de endomorfismos de K^n, con el producto definido por la composición de funciones. La primera es conmutativa, mientras que las dos últimas son no conmutativas si $n > 1$.

Como un último ejemplo importante, señalemos que si K es subcuerpo de un cuerpo F éste admite una estructura natural de $K-$álgebra conmutativa, definiendo la acción de un escalar λ sobre un elemento x de F como el producto de dichos elementos en F. $\Diamond$

Previsiblemente, una $K-$*subálgebra* de una $K-$álgebra A es un subespacio de A, cerrado con respecto al producto y conteniendo a 1.

14.1.3. Ejercicios

1. Sea S un semigrupo notado multiplicativamente y sean $m, n \in \mathbb{N}_0$.

   a) Si $x_1, x_2, \ldots, x_n$ son elementos de S, definir inductivamente el producto $x_1 x_2 \ldots x_n$

   b) Definir x^n para cualquier $x \in S$ y probar la validez de la fórmulas $x^{m+n} = x^m x^n$ y $(x^m)^n = x^{mn}$

   c) Si $x, y \in S$, probar que $(xy)^n = x^n y^n$ si S es conmutativo.

2. Sea G un grupo.

a) Si $a \in G$, probar que las transformaciones $x \mapsto ax$, $x \mapsto xa$ y $x \mapsto axa^{-1}$ son biyecciones de G

b) Si G es finito, demostrar que existe $m \in \mathbb{N}$ tal que $g^m = 1$ para todo $g \in G$.

3. Probar que $\mathbb{S}_X$ no es abeliano si $\#(X) > 2$.

4. Sea G un grupo y sea S un subconjunto de G. Probar que S es un subgrupo de G si y solo si

$$\prod_i x_i^{a_i} \in S$$

para toda familia finita (x_i) de elementos de S y toda familia (a_i) de enteros.

5. Un grupo G se dice cíclico si y solo si existe $g \in G$ tal que $\langle\, g \,\rangle = G$.

a) Probar que todo grupo cíclico es conmutativo

b) Probar que $\mathbb{Z}$, $\mathbb{Z}_n$, G_n y $\mathbb{S}_2$ son grupos cíclicos, y que $\mathbb{Q}$ no lo es.

6. Probar que

$$G_\infty = \bigcup_{n \in \mathbb{N}} G_n$$

es un subgrupo de $\mathbb{S}^1$.

7. Un subgrupo H de un grupo G se dice un *subgrupo invariante* si y solo si $ghg^{-1} \in H$ para todo par de elementos $h \in H$ y $g \in G$. Indicaremos la situación en la forma $H \triangleright G$.

a) Si G es un grupo, probar que sus subgrupos triviales son invariantes, y que todo subgrupo contenido en el centro de G es invariante. Deducir que todo subgrupo de un grupo abeliano es invariante

b) Sea G un grupo. Dado un subgrupo H de G, probar que $H \triangleright G$ si y solo si $gHg^{-1} = H$ para todo $g \in G$, donde

$$gHg^{-1} = \left\{ ghg^{-1} : h \in H \right\}$$

c) Sean $Y \subseteq X$ conjuntos no vacíos tales que $\#(X - Y) > 1$. Probar que el subgrupo

$$\{\, \sigma \in \mathbb{S}_X \,:\, \sigma(y) = y \text{ para todo } y \in Y \,\}$$

no es invariante.

8. Si A es un anillo, probar que $a\,0 = 0\,a = 0$ para todo $a \in A$. ¿Qué puede decirse de un anillo en el cual los elementos neutros de la suma y el producto coinciden ?

9. Dado un anillo A, demostrar que si $a^2 = a$ para todo $a \in A$ entonces $2a = 0$ para todo $a \in A$. Probar que todo anillo de Boole es conmutativo.

10. Si A es un anillo, demostrar que el conjunto de elementos de A inversibles con respecto al producto es un grupo. Se llama el *grupo de unidades* de A y se lo nota $\mathcal{U}(A)$.

    Determinar $\mathcal{U}(A)$ en los casos $A = \mathbb{Z}$, $A = \mathbb{Z}[\imath]$ y A un anillo de Boole cualquiera.

11. Sea X un conjunto no vacío y sea A un anillo. Demostrar que el conjunto $B = A^X$ de funciones de X en A admite una estructura de anillo definiendo la suma y el producto de funciones en la forma

    $$(f+g)(x) = f(x) + g(x) \quad \text{y} \quad (fg)(x) = f(x)g(x),$$

    donde $f, g \in B$ y $x \in X$. Probar que B es conmutativo si y solo si A lo es y caracterizar $\mathcal{U}(B)$ en los casos $A = \mathbb{Z}$ y $A = \mathbb{R}$.

12. Demostrar que todo cuerpo es un dominio de integridad.

13. Demostrar que Z_n es un dominio de integridad si y solo si n es primo.

14. *a*) Si X es un conjunto no vacío, probar que $\mathbb{P}(X)$ admite un subanillo de orden 2.
    b) Probar que $\mathbb{Z}$ no tiene subanillos propios y que $\mathbb{Q}$ no tiene subcuerpos propios
    c) Si A es un anillo conmutativo y $n \in \mathbb{N}$, probar que el centro de $M_n(A)$ es el subanillo de matrices escalares.

15. Si A es un anillo conmutativo y $\mathcal{J}$ es un ideal de A, probar la equivalencia de las siguientes afirmaciones:
    a) $\mathcal{J} \cap \mathcal{U}(A) \neq \emptyset$
    b) $1 \in \mathcal{J}$
    c) $\mathcal{J} = A$.

16. Probar que un cuerpo admite exactamente dos ideales.

17. Si K es un cuerpo y X es un conjunto no vacío, probar que el espacio vectorial K^X es una $K-$álgebra, con el producto definido como en el ejercicio 11. Demostrar que $C([0,1])$ es una subálgebra de $\mathbb{R}^{[0,1]}$ y caracterizar sus elementos inversibles.

14.2. Homomorfismos

14.2.1. Homomorfismos de grupos

Luego de haber exhibido diversas estructuras algebraicas, estudiaremos ahora ciertas relaciones funcionales entre estructuras del mismo tipo, comenzando por los grupos. Para motivar un poco el tema, consideremos en el conjunto $\mathcal{M} = \{a, b, c, d\}$ las operaciones definidas por las tablas

$\times$	a	b	c	d
a	a	b	c	d
b	b	a	d	c
c	c	d	a	b
d	d	c	b	a

y

$*$	a	b	c	d
a	c	d	a	b
b	d	c	b	a
c	a	b	c	d
d	b	a	d	c

.

Vemos por simple inspección de las mismas que ambas operaciones definen estructuras de grupo abeliano en $\mathcal{M}$. Estrictamente hablando se trata de estructuras diferentes, definidas por operaciones distintas, pero una atenta mirada a las tablas revela que dichas estructuras son esencialmente iguales, en el sentido de que la segunda se obtiene por simple alteración del nombre de las letras en las entradas de la primera, específicamente reemplazando a, b, c y d por c, d, a y b, respectivamente. Dicho en términos más formales, resulta que la permutación σ de $\mathcal{M}$ definida por $\sigma(a) = c$, $\sigma(b) = d$, $\sigma(c) = a$ y $\sigma(d) = b$ verifica la relación

$$\sigma(x \times y) = \sigma(x) * \sigma(y)$$

cualesquiera sean $x, y \in \mathcal{M}$, lo cual significa que la estructura $(\mathcal{M}, *)$ se obtiene "transportando" la estructura $(\mathcal{M}, \times)$ a través de la biyección σ.

Teniendo en cuenta este sencillo ejemplo, estudiemos en general aquellas relaciones funcionales entre grupos que preservan sus operaciones, a las que llamaremos *homomorfismos de grupos* y que definimos de la siguiente manera:

> Sean G y H grupos y sea $f : G \to H$ una aplicación. Diremos que f es un homomorfismo de grupos si y solo si
>
> $$f(ab) = f(a)f(b)$$
>
> cualesquiera sean $a, b \in G$.

Designaremos por $\mathrm{Hom}(G, H)$ el conjunto de homomorfismos de grupos de G en H. Ocasionalmente, emplearemos el término *morfismos* para referirnos a tales funciones.

Por ejemplo, son homomorfismos de grupos las aplicaciones $g : \mathbb{C}^* \to \mathbb{R}^*$ y $h : \mathbb{R} \to \mathbb{R}_{>0}$ definidas por $g(z) = |z|$ y $h(x) = e^x$, respectivamente.

Señalemos también que $\mathrm{Hom}(G,H) \neq \emptyset$ cualesquiera sean los grupos G y H, ya que la función $f : G \to H$ definida por $f(x) = 1$ para todo $x \in G$ es un morfismo de grupos, llamado *morfismo trivial*.

En la siguiente proposición destacamos algunas propiedades elementales de los homomorfismos de grupos.

Proposición 14.2.1 Si G y H son grupos y $f \in \mathrm{Hom}(G,H)$ son válidas las siguientes propiedades:

1) $f(1) = 1$

2) $f(x^{-1}) = f(x)^{-1}\ \forall\, x \in G$. Más generalmente,

$$f\left(\prod_{i=1}^{s} x_i^{m_i}\right) = \prod_{i=1}^{s} f(x_i)^{m_i},$$

cualesquiera sean los elementos x_i en G y los enteros m_i

3) Si S es un subgrupo de G, el conjunto

$$f(S) = \{\, f(x) : x \in S\}$$

es un subgrupo de H, llamado imagen directa de S por f, mientras que si T es un subgrupo de H el conjunto

$$f^{-1}(T) = \{\, x \in G : f(x) \in T\}$$

es un subgrupo de G, que llamamos imagen inversa de T por f

4) Sea K un grupo y sea $g \in \mathrm{Hom}(H,K)$. Entonces $g \circ f \in \mathrm{Hom}(G,K)$

5) Si f es biyectiva entonces $f^{-1} \in \mathrm{Hom}(H,G)$.

Demostracion Queda a cargo del lector. $\Diamond$

Se emplean términos especiales para designar ciertas características de los morfismos de grupo. Así, si $f \in \mathrm{Hom}(G,H)$ es inyectiva (resp. suryectiva) diremos que f es un *monomorfismo* (resp. un *epimorfismo*). Si f es biyectiva, diremos que f es un *isomorfismo*.

Por ejemplo, la función exponencial es un isomorfismo de grupos de $\mathbb{R}$ en $\mathbb{R}_{>0}$, siendo su inversa la función logaritmo natural, que resulta ser entonces un isomorfismo de grupos de $\mathbb{R}_{>0}$ en $\mathbb{R}$. En general, si existe un isomorfismo de G en H diremos que G y H son *isomorfos* y escribiremos $G \cong H$.

Si $H = G$ se dice que f es un endomorfismo, empleándose la notación $\mathrm{End}(G)$ para referirse al conjunto de endomorfismos de G. Finalmente, un endomorfismo biyectivo de G se dirá un *automorfismo* de G. Por ejemplo, es evidente que la función identidad I_G es un automorfismo de G. Sigue de este hecho y de las propiedades 4) y 5) de la proposición 14.2.1 que el conjunto de automorfismos de G es un grupo respecto a la composición de funciones. Se lo llama el *grupo de automorfismos* de G y se lo nota $\mathrm{Aut}(G)$.

Núcleo e imagen

Si G y H son grupos y f es un homomorfismo de G en H, los conjuntos

$$\mathrm{Nu}(f) = \{\, x \in G : f(x) = 1 \,\} \quad \text{y} \quad \mathrm{Im}(f) = \{\, f(x) : x \in G \,\}$$

se llaman respectivamente el *núcleo* y la *imagen* de f.

Por ejemplo, el núcleo del morfismo de $\mathbb{C}^*$ en $\mathbb{R}^*$ dado por $g(z) = |\, z \,|$ es la circunferencia unitaria S^1, mientras que su imagen es el subgrupo $\mathbb{R}_{>0}$ de números reales positivos.

Volviendo a la definición, sigue de 14.2.1 que $\mathrm{Nu}(f) = f^{-1}(\{\, 1\})$ es un subgrupo de G y que $\mathrm{Im}(f) = f(G)$ es un subgrupo de H. Obsérvese que estamos usando genéricamente notación multiplicativa, pero el lector no debe perder de vista que en cualquier caso el núcleo se define como la imagen inversa del elemento neutro de H. Por ejemplo, en el caso de notación aditiva sería $\mathrm{Nu}(f) = \{\, x \in G : f(x) = 0 \,\}$.

Lema 14.2.2 f es un epimorfismo si y solo si $\mathrm{Im}(f) = G$ y f es un monomorfismo si y solo si $\mathrm{Nu}(f) = \{\, 1\}$.

Demostracion La primera afirmación es evidente a partir de las definiciones.

En cuanto a la segunda, sigue de la propiedad 1) de la proposición 14.2.1 que $1 \in \mathrm{Nu}(f)$ cualquiera sea f. Si f es un monomorfismo, consideremos un elemento cualquiera $x \in \mathrm{Nu}(f)$. Resulta entonces $f(x) = 1 = f(1)$, y siendo f inyectiva concluimos que $x = 1$. Luego $\mathrm{Nu}(f) = \{\, 1\}$.

Recíprocamente, asumamos que el núcleo de f es trivial y sean $a, b \in G$ tales que $f(a) = f(b)$. Entonces

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

esto es, $ab^{-1} \in \mathrm{Nu}(f)$. Sigue luego por hipótesis que $ab^{-1} = 1$, ó equivalentemente, $a = b$. En consecuencia f es inyectiva. $\diamondsuit$

Ejemplos 14.2.3 Ilustremos las definiciones y propiedades anteriores.

1) Si G es un grupo, la aplicación $f(x) = x^{-1}$ es un endomorfismo de G si y solo si G es abeliano. En efecto, si G es conmutativo y $x, y \in G$, tenemos

$$f(xy) = (xy)^{-1} = y^{-1}x^{-1} = x^{-1}y^{-1} = f(x)f(y),$$

y por lo tanto g es un morfismo. Inversamente, supongamos esto último y sean $x, y \in G$. Entonces

$$\begin{aligned} xyx^{-1} &= x(y^{-1})^{-1}x^{-1} = xf(y^{-1})f(x) = xf(y^{-1}x) = \\ &= x\left(y^{-1}x\right)^{-1} = xx^{-1}y = y, \end{aligned}$$

de donde se obtiene (multiplicando a derecha por x) que $xy = yx$. Más generalmente, usando un argumento inductivo el lector puede probar que si G es abeliano la aplicación $x \mapsto x^n$ es un endomorfismo de G cualquiera sea $n \in \mathbb{Z}$.

En el caso general, es fácil probar que la función $\iota_a : G \to G$ definida por $\iota_a(x) = axa^{-1}$ es un automorfismo de G cualquiera sea $a \in G$, siendo $\iota_{a^{-1}}$ su inverso. Tales automorfismos, llamados *interiores*, determinan un subgrupo de $\mathrm{Aut}(G)$, y tienen real interés en el caso no abeliano, ya que todos ellos se reducen a la identidad si G es abeliano.

2) Las siguientes funciones son homomorfismos de grupos:

   a) $f_1 : \mathbb{R} \to \mathbb{C}^*$, $f_1(x) = e^{\imath x}$.

   b) $f_2 : \mathbb{Q} \to \mathbb{C}^*$, $f_2(x) = e^{2\pi\imath x}$.

   c) $f_3 : M_n(K) \to K$, $f_3(A) = t_r(A)$.

   Encomendamos al lector verificar todos los detalles, así como la tarea de caracterizar el núcleo y la imagen de cada uno de los morfismos f_i.

3) Si M y M' son grupos abelianos (que notaremos aquí aditivamente), el conjunto $\mathrm{Hom}(M, M')$ admite una estructura natural de grupo abeliano, definiendo la suma de dos morfismos f y g de M en M' como la función

$$f + g : M \to M'$$

   dada por $(f + g)(x) = f(x) + g(x)$.

   Es mera rutina probar que $f + g$ es un homomorfismo de M en M', y que se satisfacen en $\mathrm{Hom}(M, M')$, respecto de esta operación, los correspondientes axiomas de grupo abeliano.

   Tomemos por ejemplo $M = M' = \mathbb{Z}$. Por un lado, dado $n \in \mathbb{Z}$ es claro que la función definida por $h_n(x) = nx$ es un endomorfismo de $\mathbb{Z}$ (es la version aditiva de la función $x \mapsto x^n$ exhibida en 1), verificándose además la relación $h_{m+n} = h_m + h_n$, para todo par de enteros m y n.

   Por otro lado, si f es cualquier endomorfismo de $\mathbb{Z}$ sea $k = f(1)$. Entonces, dado $x \in \mathbb{Z}$, tenemos:

$$f(x) = f(x\,1) = xf(1) = xk = kx = h_k(x),$$

   y por lo tanto $f = h_k$. En consecuencia,

$$\mathrm{End}(\mathbb{Z}) = \{\, h_n : n \in \mathbb{Z} \,\}\,.$$

   En términos de estructura, hemos probado que la aplicación

$$\Gamma : \mathbb{Z} \to \mathrm{End}(\mathbb{Z})$$

definida por $\Gamma(n) = h_n$ es un epimorfismo de grupos. Puesto que es inyectiva, ya que $h_a = h_b$ implica $a = h_a(1) = h_b(1) = b$, concluimos que es un isomorfismo. Luego, $\mathrm{End}(\mathbb{Z}) \cong \mathbb{Z}$.

Para cerrar este ítem, veamos que hay un único morfismo de $\mathbb{Q}$ en $\mathbb{Z}$. En efecto, sea $f \in \mathrm{Hom}(\mathbb{Q}, \mathbb{Z})$ y sea $x \in \mathbb{Q}$. Entonces, dado un primo p, tenemos:

$$f(x) = f\left(p\,\frac{x}{p}\right) = pf\left(\frac{x}{p}\right),$$

y por lo tanto $p \mid f(x)$. Puesto que esta relación es válida para todo primo p concluimos que $f(x) = 0$, vale decir, f es el morfismo nulo. Hemos probado así que el grupo $\mathrm{Hom}(\mathbb{Q}, \mathbb{Z})$ es trivial.

4) Como último ejemplo, precisemos una afirmación que hicimos páginas atrás, acerca de que los grupos de clases de congruencia módulo n y de raíces $n-$ésimas de la unidad son esencialmente iguales, cualquiera sea $n \in \mathbb{N}$. Concretamente, probaremos que son isomorfos. Para ello, fijada una raíz $n-$ésima primitiva w, consideremos la aplicación

$$F : \mathbb{Z}_n \to G_n$$

dada por la fórmula $F((k)) = w^k$, y veamos primero que está bien definida, es decir, que $F((k))$ no depende del representante elegido para la clase. Supongamos a tal efecto que t es congruente con k módulo n, digamos $t = k + qn$ $(q \in \mathbb{Z})$. Tenemos entonces:

$$F((t)) = w^t = w^{k+qn} = w^k (w^n)^q = w^k = F((k)),$$

lo que asegura la buena definición de F.

Es inmediato también probar que F es un morfismo de grupos, ya que

$$F((a) + (b)) = F((a+b)) = w^{a+b} = w^a w^b = F((a))\,F((b)),$$

cualesquiera sean $a, b \in \mathbb{Z}$.

Puesto que por definición de raíz primitiva todos los elementos de G_n son potencias de w, resulta que F es una función suryectiva, y siendo $\mathbb{Z}_n$ y G_n grupos finitos del mismo orden, concluimos que también es inyectiva. En definitiva F es un isomorfismo y $\mathbb{Z}_n \cong G_n$. $\quad \Diamond$

14.2.2. Homomorfismos de otras estructuras

El concepto de homomorfismo de grupos se extiende a otras estructuras, en el sentido de transformaciones que preservan las diversas operaciones involucradas en ellas. Con las diferencias de rigor la situación es muy similar al caso de grupos, por lo que sólo brindaremos una somera descripción de los diversos tipos de homomorfismos. En realidad, ellos son en particular morfismos de grupos abelianos, por lo que podemos aplicar con respecto

a los mismos las propiedades establecidas en el apartado anterior y emplear parecida notación y vocabulario (monomorfismo, epimorfismo, etc.). Comencemos por la estructura de anillo.

Homomorfismos de anillos.

Sean A y B anillos y sea $f : A \to B$ una aplicación. Diremos que f es un *homomorfismo de anillos* si y solo si se satisfacen las siguientes condiciones:

1) f es un homomorfismo de grupos
2) $f(aa') = f(a)f(a') \; \forall \; a, a' \in A$
3) $f(1) = 1$.

Naturalmente, la condición 1) está referida a las estructuras aditivas de A y B. Como se suele decir, un morfismo de anillos es entonces una función aditiva y multiplicativa que aplica el 1 en el 1. Observemos al respecto que la función idénticamente nula satisface 1) y 2) pero no 3), lo que indica que esta última condición es independiente de las dos primeras. Conservando la notación de la definición, es fácil probar por inducción que

$$f\left(\sum_i a_i a_i'\right) = \sum_i f(a_i)f(a_i')$$

para cualquier par de familias finitas (a_i) y (a_i') de elementos de A, resultando como corolario que $f(a^m) = f(a)^m$ para todo $a \in A$ y para todo entero no negativo m. Asimismo, sigue inmediatamente de la definición que $\mathrm{Nu}(f)$ es un ideal bilátero de A y que $\mathrm{Im}(f)$ es un subanillo de B.

Si A y B son cuerpos, diremos que f es un *homomorfismo de cuerpos*. En tal caso, dado $x \in A^*$ tenemos

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

de donde deducimos que $f(x) \neq 0$ y $f(x)^{-1} = f(x^{-1})$. Sigue en particular que todo homomorfismo de cuerpos es un monomorfismo.

Ejemplos 14.2.4 Veamos algunos ejemplos básicos.

1) Si A es un anillo y $n \in \mathbb{N}$, es inmediato verificar que las aplicaciones

$$g : A \to A[X] \quad \text{y} \quad h : A \to M_n(A)$$

definidas respectivamente por $g(a) = a\,X^0$ y $h(a) = a\,I_n$ son monomorfismos de anillos. Como consecuencia de ello, podemos identificar en los dos casos el dominio con la imagen, resultando entonces que ambos anillos contienen una copia algebraica de A, a saber, el subanillo de polinomios constantes en $A[X]$ y el subanillo de matrices escalares en $M_n(A)$.

2) Si B es un anillo, existe un único homomorfismo de anillos de $\mathbb{Z}$ en B.

En efecto, observemos en primer término que la aplicación $\mu : \mathbb{Z} \to B$ definida por $\mu(k) = k1$ lo es. Por otra parte, dado cualquier morfismo f de $\mathbb{Z}$ en B y dado $k \in \mathbb{Z}$, tenemos:

$$f(k) = f(k1) = kf(1) = k1 = \mu(k),$$

y por lo tanto $f = \mu$. Por ejemplo, la proyección canónica $k \mapsto (k)$, que asigna a cada entero k su clase de congruencia módulo $n\,(n \in \mathbb{N})$ es el único morfismo de anillos de $\mathbb{Z}$ en $\mathbb{Z}_n$.

3) Veamos ahora que la función identidad es el único endomorfismo de cuerpo de $\mathbb{R}$. Para ello, consideremos cualquier endomorfismo f y observemos primero que $f(m) = m$ para todo $m \in \mathbb{Z}$, hecho que se deduce de 2). Probaremos a continuación que f tiene el mismo comportamiento sobre los racionales.

Sea pues $x \in \mathbb{Q}$, digamos $x = a/b$, con a y b enteros $(b \neq 0)$. Entonces

$$bf(x) = f(bx) = f(a) = a$$

y por lo tanto $f(x) = a/b = x$, como queríamos demostrar.

En el próximo paso mostraremos que f es estrictamente creciente, para lo cual bastará probar que conserva signos, ya que en ese caso tendríamos

$$x > y \Leftrightarrow x - y > 0 \Leftrightarrow f(x-y) > 0 \Leftrightarrow f(x) - f(y) > 0 \Leftrightarrow f(x) > f(y),$$

cualesquiera sean $x, y \in \mathbb{R}$. Tomemos entonces un número real no nulo r. Si $r > 0$ tenemos

$$f(r) = f\left(\left(\sqrt{r}\,\right)^2\right) = f(\sqrt{r})^2 > 0,$$

pues f es un monomorfismo, mientras que si $r < 0$ resulta

$$f(r) = f(-(-r)) = -f(-r) < 0$$

por lo anterior, quedando así probado que f conserva signos.

Para completar nuestra afirmación, supongamos que $f(c) \neq c$ para algún $c \in \mathbb{R}$, y asumamos sin pérdida de generalidad que $c < f(c)$ (si fuera $f(c) < c$ resultaría $-c < f(-c)$). Entonces, tomando cualquier número racional q tal que $c < q < f(c)$, obtenemos $f(c) < f(q) = q$, lo que es una contradicción. En consecuencia $f = I_{\mathbb{R}}$.

Nótese que hemos probado también que el único endomorfismo de cuerpo de $\mathbb{Q}$ es la identidad. De todos modos, este no es un hecho general. Por ejemplo, la conjugación es un automorfismo de cuerpo de $\mathbb{C}$ distinto de la identidad.

4) A diferencia del caso de grupos, puede ocurrir que no exista ningún homomorfismo (de anillos) de un anillo A en otro B, como por ejemplo de $\mathbb{Q}$ en $\mathbb{Z}_7$, ya que todo morfismo de cuerpos es un monomorfismo, y no existen funciones inyectivas de un conjunto infinito en otro finito. Veamos que por razones de otra índole, tampoco existen homomorfismos de anillos de $M_2(\mathbb{Z})$ en $\mathbb{Z}$.

   Usemos para ello la base conónica $\left\{E^{ij}\right\}$ de $M_2(\mathbb{R})$. Suponiendo que f es un homomorfismo de anillos de $M_2(\mathbb{Z}$ en $\mathbb{Z}$, tenemos que

$$f(E^{11})f(E^{22}) = f(E^{11}E^{22}) = f(0) = 0,$$

   de donde sigue que $f(E^{11}) = 0$ ó $f(E^{22}) = 0$. Considerando el caso $f(E^{11}) = 0$, resulta

$$f(E^{22}) = f(E^{21}E^{12}) = f(E^{21}E^{11}E^{12}) = f(E^{21})f(E^{11})f(E^{12}) = 0,$$

   lo que lleva a una contradicción, ya que por otro lado tendríamos

$$1 = f(I_2) = f(E^{11} + E^{22}) = f(E^{11}) + f(E^{22}) = 0.$$

   En el caso $f(E^{22}) = 0$ se procede de manera análoga. $\Diamond$

Homomorfismos de álgebras.

Si K es un cuerpo y A y B son $K-$álgebras, un *homomorfismo de $K-$álgebras* de A en B es un morfismo de anillos $f : A \to B$ tal que $f(\lambda\, a) = \lambda f(a)$ para todo $\lambda \in K$ y para todo $a \in A$. En otras palabras, f es tanto un homomorfismo de espacios vectoriales como de anillos.

Ejemplos 14.2.5 Veamos dos ejemplos importantes de esta clase de transformaciones (como siempre K denota un cuerpo).

1) Generalizando una situación que hemos estudiado en el capítulo 10, dados una $K-$álgebra A y cualquier elemento α en A existe un único morfismo de $K-$álgebras

$$\epsilon_\alpha : K[X] \to A$$

   que aplica X en α, definido en la forma

$$\epsilon_\alpha(\sum_i \lambda_i X^i) = \sum_i \lambda_i \alpha^i.$$

   Se demuestra sin dificultad que ϵ_α es un morfismo de $K-$álgebras, llamado *morfismo de especialización* en α, siendo claro que $\epsilon_\alpha(X) = \alpha$. La buena definición de ϵ_α y la unicidad a la que nos referíamos arriba

se deducen del hecho de que todo elemento de $K[X]$ se expresa unívocamente como combinación lineal de potencias de X, por definición de polinomio. Esto significa por otra parte que todo morfismo de $K-$álgebras con dominio en $K[X]$ queda completamente determinado por la imagen de X, y por lo tanto es una especialización. Si $f \in K[X]$, se acostumbra a escribir $\epsilon_\alpha(f) = f(\alpha)$.

Consideremos por ejemplo el endomorfismo de $K[X]$ definido por especializacion en $X+1$. El mismo resulta un automorfismo y ϵ_{X-1} es su inversa, ya que

$$\epsilon_{X-1} \circ \epsilon_{X+1}(X) = \epsilon_{X-1}(X+1) = X - 1 + 1 = X,$$

y por lo tanto $\epsilon_{X-1} \circ \epsilon_{X+1} = I_{K[X]}$. En forma completamente análoga se prueba que $\epsilon_{X+1} \circ \epsilon_{X-1} = I_{K[X]}$. Deducimos de estos hechos que todo polinomio en X puede expresarse de manera única como "polinomio" en $X+1$ o en $X-1$.

Por su importancia en el estudio de los endomorfismos de un espacio vectorial de dimensión finita, mencionemos también los morfismos de especialización en matrices

$$f(T) = \sum_i \lambda_i T^i,$$

donde $T \in M_n(K)$.

Tomando por ejemplo $n = 2$, $T = \begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}$ y $f = 4X + 3$ resulta

$$f(T) = 4T + 3I_2 = \begin{pmatrix} 3 & 8 \\ 4 & 7 \end{pmatrix},$$

mientras que si $g = X^2 - X - 2$ tenemos

$$g(T) = T^2 - T - 2I_2 = 0,$$

como el lector puede comprobar.

2) Si $n \in \mathbb{N}$, la aplicación

$$\chi : \mathrm{End}_K(K^n) \to M_n(K)$$

descripta en la Nota que sigue a la proposición 13.2.25 es un isomorfismo de $K-$álgebras. $\Diamond$

14.2.3. Ejercicios

1. Demostrar que un grupo G es abeliano si y solo si la aplicación $x \mapsto x^2$ es un endomorfismo de G.

2. Demostrar la proposición 14.2.1

3. Sea $f : G \to H$ un morfismo de grupos y sean S y T subgrupos invariantes de G y H, respectivamente. Demostrar que $f(S) \triangleright \mathrm{Im}(f)$ y que $f^{-1}(T) \triangleright G$. Deducir que $\mathrm{Nu}(f)$ es un subgrupo invariante de G.

4. Si G es un grupo, probar que la aplicación $a \mapsto \iota_a$ es un morfismo de grupos de G en $\mathrm{Aut}(G)$ cuyo núcleo es el centro de G.

5. *a*) Si G es un grupo, probar que $\mathrm{End}(G)$ es un semigrupo respecto a la composición de funciones y caracterizar $\mathcal{U}(\mathrm{End}(G))$
   b) Probar que $\mathrm{End}(G)$ admite una estructura natural de anillo si G es abeliano.

6. Probar que $\mathrm{Hom}(Z, M) \cong M$ para todo grupo abeliano M.

7. Si $n > 1$, demostrar que no existen homomorfismos de anillos de $\mathbb{Z}_n$ en $\mathbb{Z}$.

Bibliografía

[1] Tom Apostol, *Introducción a la Teoría analítica de números*, Editorial Reverté, Barcelona, 1984.

[2] David Bressoud, *Factorization and Primality Testing*, Springer-Verlag, New York, 1988.

[3] Ruel Churchill, James Brown, Roger Verhey, *Variables complejas y sus aplicaciones*, McGraw-Hill, México DF,1970.

[4] Jay Davore, *Probabilidades y Estadística en Ciencias e Ingenierías*, Ediciones Paraninfo, México, 2006.

[5] Leonard Dickson, *History of the Theory of Numbers*, Chelsea, New York, 1950.

[6] Enzo Gentile, *Aritmética Elemental en la Formación Matemática*, Red Olímpica, Buenos Aires, 1991.

[7] Ronald Graham, Donald Knuth, Oren Patashnik, *Concrete Mathematics*, Addison-Wesley, Reading, MA, 1994.

[8] Paul Halmos, Teoría intuitiva de conjuntos, Ed. CECSA, México D.F.,1965.

[9] Israel Herstein, *Topics in Algebra*, Wiley and Sons, New York, 1975.

[10] Kenneth Hoffman, Ray Kunze, *Algebra Lineal*, Prentice Hall Hispanoamérica, México, 1973.

[11] Nathan Jacobson, *Basic Algebra I*, Dover Ed., San Francisco, 1985.

[12] David Johnson, *Elements of logic via numbers and sets*, Springer-Verlag, London, 1998.

[13] Neal Koblitz, *A Course in Number Theory and Cryptography*, Springer-Verlag, 1994.

[14] Ivan Niven, *Matemática de las opciones o cómo contar sin contar*, Red Olímpica, Buenos Aires, 1995.

[15] Ivan Niven, Herbert Zuckerman, Hugh Montgomery, *An Introduction to the Theory of Numbers*, Wiley and Sons, New York 1991.

[16] Willard Quine, *Los métodos de la Lógica*, Planeta-Agostini, Barcelona, 1986.

[17] Paulo Ribenboim, *13 Lectures on Fermat's Last Theorem*, Springer-Verlag, New York, 1979.

[18] Walter Rudin, *Principles of Mathematical Analysis*, McGraw-Hill, New York, 1964.

[19] Pierre Samuel, *Théorie algébrique des nombres*, Hermann, Paris, 1967.

[20] Simon Singh, *El último teorema de Fermat*, Grupo Editorial Norma, Bogotá, 1999.

[21] Michael Spivak, *Calculus*, Editorial Reverté, México, 1999.

Índice alfabético