

Alejandro Bravo Mojica
Hugo Rincón Mejía
César Rincón Orta

ÁLGEBRA SUPERIOR

FACULTAD DE CIENCIAS, UNAM

2013



Álgebra superior
1^a edición, 2006

© D.R. 2013. Universidad Nacional Autónoma de México
Facultad de Ciencias.
Ciudad Universitaria. Delegación Coyoacán
C. P. 04510, México, Distrito Federal
editoriales@ciencias.unam.mx

ISBN: 978-968-32-3750-7

Diseño de portada: Laura Uribe

Prohibida la reproducción parcial o total de la obra por cualquier medio,
sin la autorización por escrito del titular de los derechos patrimoniales

Impreso y hecho en México

Índice general

Prefacio	ix
1 Lógica proposicional	1
1.1 Conceptos primitivos. Verdad, falsedad	1
1.2 Conectivos lógicos y Tablas de verdad	5
1.3 Tautologías y absurdos	8
1.4 Sistemas completos de conectivos	11
1.5 Reglas de inferencia, deducciones	15
1.5.1 Regla del reemplazo	18
1.5.2 Regla de la tautología	20
1.5.3 Negaciones	26
1.5.4 Inferencias no válidas	27
1.6 Reducción al absurdo	30
1.7 Apéndice. Sistemas formales	38
1.7.1 El sistema formal L	39
1.8 El Teorema de la deducción y las hipótesis adicionales	41
1.9 Valuación	53
1.10 Cuantificadores	56
2 Conjuntos y funciones	61
2.1 Axiomas	61
2.1.1 Pertenencia y contención	63
2.1.2 Especificación y existencia	65
2.1.3 No hay un conjunto de todos los conjuntos	66
2.1.4 Intersecciones y complementos	67
2.1.5 Uniones	69
2.1.6 Familias	74
2.1.7 La diferencia simétrica	77

2.1.8	El conjunto potencia	78
2.2	Parejas ordenadas, producto cartesiano y relaciones	79
2.2.1	Axioma de regularidad	84
2.2.2	Órdenes parciales	86
2.2.3	Retículas	90
2.3	Orden en un producto de conjuntos ordenados	92
2.4	Funciones	96
2.4.1	Funciones inyectivas	98
2.4.2	Funciones suprayectivas	101
2.4.3	Funciones biyectivas	105
2.5	Cardinalidad	107
2.5.1	Axioma del infinito	107
2.5.2	Conjuntos infinitos	109
2.6	Imágenes directas e imágenes inversas	112
2.7	Relaciones de equivalencia y particiones	117
2.8	La relación de equivalencia generada por una relación	125
2.9	Operaciones	133
2.9.1	La restricción de una operación	135
2.9.2	Operaciones asociativas	137
2.9.3	Tablas de multiplicar	138
3	El conjunto \mathbb{N} de los números naturales	143
3.1	Introducción	143
3.2	Los axiomas de Peano	144
3.3	Construcción	145
3.4	Definiciones recursivas	149
3.5	Demostraciones inductivas	157
3.6	Conjuntos transitivos	158
3.7	Conjuntos infinitos y conjuntos finitos	161
3.8	El conjunto de los naturales es un conjunto infinito	161
3.9	El orden en los naturales	167
3.10	Recursión	171
3.11	Las propiedades algebraicas de los naturales	174
3.11.1	La suma	174
3.11.2	El producto en \mathbb{N}	183
3.11.3	Potencias	188
3.12	Apéndice. Sobre las definiciones recursivas	190

4 Los números enteros	197
4.1 Construcción y definiciones	198
4.2 El orden en \mathbb{Z}	201
4.2.1 Los enteros positivos	201
4.3 Inmersión de los naturales en los enteros	202
4.4 El producto en \mathbb{Z}	203
4.5 El algoritmo de la división	208
4.6 Divisibilidad y congruencias	210
4.6.1 Subconjuntos de \mathbb{Z} cerrados bajo la resta.	210
4.6.2 El máximo común divisor	216
4.6.3 El mínimo común múltiplo	218
4.7 El Teorema fundamental de la Aritmética	226
4.7.1 El conjunto de primos es infinito	230
4.8 El algoritmo de Euclides	231
4.9 El anillo de los enteros módulo n	234
4.10 Congruencias	240
4.11 Sistemas de congruencias	243
4.11.1 El Teorema chino del residuo	247
4.12 Ecuaciones diofantinas	255
4.13 Sistemas de numeración con bases distintas de 10	262
4.13.1 Algunos criterios de divisibilidad	267
4.14 Los números racionales	270
4.14.1 La suma en \mathbb{Q}	273
4.14.2 El producto en \mathbb{Q}	274
4.14.3 El orden en \mathbb{Q}	275
4.14.4 Inmersión de \mathbb{Z} en \mathbb{Q}	278
5 ¿De cuántas maneras?	281
5.1 ¿Cuántos subconjuntos tiene un conjunto con n elementos?	287
5.1.1 El principio de la pichoneras	291
5.2 Subconjuntos con k elementos de un conjunto con n elementos	292
5.3 Permutaciones	298
5.3.1 Ordenaciones	300
5.4 ¿Cuántas funciones suprayectivas hay de A a B ?	308
5.4.1 Relación de recurrencia para P_m^n	311
5.5 Ejercicios	313

6 El campo de los números reales	331
6.1 Consideraciones generales	331
6.2 Construcción de \mathbb{R} a partir de las cortaduras en \mathbb{Q}	333
6.3 Cortaduras de Dedekind	339
6.4 El producto en \mathbb{R}	348
6.5 Supremos e ínfimos	356
6.5.1 El principio del supremo	357
6.5.2 La recta está completa	360
6.6 Representación decimal de un número real	363
7 El campo \mathbb{C} de los números complejos	369
7.1 La inmersión de \mathbb{R} en \mathbb{C}	375
7.1.1 Modelo	375
7.2 La conjugación	377
7.3 La norma	379
7.4 La ecuación general de segundo grado	381
7.4.1 Sistemas de ecuaciones	385
7.5 Representación geométrica de los números complejos	388
7.5.1 Pasar de coordenadas rectangulares a forma polar	389
7.6 Raíces n -ésimas de un número complejo	395
7.7 El argumento de un número complejo	398
7.8 Algunas transformaciones del plano	399
7.8.1 Contracciones y expansiones	399
7.8.2 Rotaciones	400
7.8.3 Reflexión sobre el eje X	400
7.8.4 Reflexión respecto al origen	401
7.9 La función exponencial compleja	401
7.9.1 Representación geométrica de algunas rectas bajo la transformación E	404
7.9.2 La función logaritmo	406
7.10 Las funciones trigonométricas	409
8 Espacios vectoriales	411
8.1 Conceptos preliminares	411
8.2 Espacios vectoriales	420
8.3 Subespacios	425
8.3.1 Dependencia lineal	431
8.4 Bases	436

8.4.1	Intersección de subespacios y suma de subespacios	444
8.5	Producto punto	445
8.6	Matrices	452
8.6.1	El rango de una matriz	453
8.7	Funciones lineales	456
8.8	La matriz de una función lineal entre $F^n \xrightarrow{T} F^m$	461
8.9	Sistemas de ecuaciones lineales	469
8.9.1	Algunas definiciones	470
8.9.2	Un método para resolver sistemas de ecuaciones lineales	475
8.9.3	Algoritmo para la solución de sistemas de ecuaciones lineales	480
8.10	Matrices reducidas y escalonadas	488
8.11	Determinantes	497
8.11.1	Notaciones para permutaciones	497
8.11.2	La paridad de una permutación	502
8.11.3	Determinantes	504
8.11.4	El desarrollo del determinante respecto a un renglón	506
8.11.5	El determinante de un producto de matrices I	513
8.11.6	Determinantes y rango	518
8.11.7	El determinante de un producto de matrices II	522
8.11.8	Matrices invertibles y determinantes	528
8.11.9	La regla de Cramer	531
8.11.10	Determinantes y funciones multilineales	532
8.11.11	Resumen de las propiedades del determinante	535
9	Polinomios con coeficientes en \mathbb{R}	539
9.1	Construcción y definiciones	540
9.2	Evaluación	547
9.3	Los ideales de $\mathbb{R}[x]$	550
9.3.1	Traslación de la gráfica de un polinomio	560
9.3.2	El método de Horner	565
9.4	Un procedimiento gráfico para resolver algunas desigualdades	569
9.4.1	Procedimiento gráfico para resolver la desigualdad $f(x) < 0$	571
9.4.2	Una aplicación	573
9.5	Reflexión sobre el eje Y	574
9.6	Continuidad	574
9.7	Valores intermedios	578

9.8	Derivadas	582
9.9	Derivadas y multiplicidad	591
9.10	El teorema de Sturm	594
9.11	Regla de los signos de Descartes	605
9.12	Raíces racionales	610
9.13	Coeficientes y raíces	612
9.14	Polinomios de tercer grado	615
9.14.1	El discriminante y número de raíces reales	616
9.15	Polinomios de grado cuatro	623
9.16	Otra construcción de \mathbb{C}	625
A	Una teoría axiomática para \mathbb{R}	629
A.1	Los axiomas	629
A.1.1	Axiomas, Grupo I.	629
B	Las funciones trascendentes	633
B.1	“Un cúmulo de conocimientos previos”	633
B.2	Hipótesis. (Mosaico 1)	639
B.3	La función exponencial (2a. versión)	640
B.4	Funciones trigonométricas	645

Prefacio

La Matemática es una ciencia viva. Cada año incorpora a su acervo miles de teoremas. Cada día se producen nuevos resultados. Aparecen nuevas teorías y se actualizan las que son clásicas. Se mejoran todas. La tecnología aporta nuevos puntos de vista; otra manera de enfocar los temas sustantivos de ésta, que es la más pura expresión de la inteligencia humana. Sin embargo, dentro de esta revolución de nuevas ideas, se distinguen aquellas que por su trascendencia se conservan incólumes. Apenas tocadas por el maquillaje de las nuevas formas de expresión. La Geometría de Euclides, enriquecida con las precisiones de Hilbert, permanece subyacente en una gran parte del conocimiento científico. Y qué decir del Álgebra, el lenguaje universal con el que se expresa la Matemática. Las ciencias de la computación han cambiado sustancialmente el proceso de enseñanza-aprendizaje, pero los conceptos básicos y la lógica con la que deben manejarse siguen siendo vigentes y su importante relevancia se reconoce en el énfasis que se pone en los contenidos curriculares de los primeros cursos de las diferentes licenciaturas que no abandonan la enseñanza de la Geometría ni del Álgebra.

La idea central que nos motivó para escribir este libro fue la de realizar un intento para reunir algunas partes esenciales de ese conocimiento sobre el que se construye y desarrolla el edificio de la Matemática. La experiencia de muchos cursos de álgebra básica que los estudiantes toman en los primeros semestres de sus carreras y que los autores hemos impartido durante varios años en las facultades de Ciencias y de Química de la UNAM, nos llevaron a seleccionar el contenido, y conscientes de que el problema del rigor es uno de los parámetros más importantes en el proceso de la enseñanza-aprendizaje de la Matemática, decidimos mantener éste en un grado de dificultad adecuado para buscar el equilibrio -el justo balance- entre el formalismo deseado y el nivel de conocimientos y habilidades con que -sabemos- ingresan nuestros alumnos a las licenciaturas. Nos queda claro que el aprendizaje de la

Matemática exige la formación de estructuras mentales de la más alta calidad, que obviamente, no pueden generarse de la nada. Para lograr un aprendizaje significativo, es indispensable ante todo, una buena formación previa y se requiere además un esfuerzo mantenido -constante- por parte del estudiioso que debe “hacer suyo el conocimiento”. Que necesita ir modificando sus marcos conceptuales y desarrollando el caudal de habilidades y de herramientas teóricas que le permitan continuar con buen éxito su desarrollo profesional. Enfatizamos aquí la importancia de este esfuerzo, convencidos de que cada resultado, cada definición, cada concepto que el estudiante ignora produce, cuando aparece en un discurso, un “cono de sombra” que oscurece, oculta o distorsiona una parte significativa del desarrollo posterior de teoría, que puede en muchos casos, volverse inentendible para él.

Agradecimientos

Agradecemos a los árbitros por la cuidadosa lectura del texto, y por sus valiosos comentarios y sugerencias.

Agradecemos a Rolando Gómez Macedo por haber señalado multitud de errores tipográficos.

Agradecemos al Dr. Carlos Velarde por las ilustraciones de la transformación geométrica producida por la función exponencial compleja.

A todas las personas que contribuyeron de alguna manera a la realización de este libro, les agradecemos su ayuda e interés.

Capítulo 1

Lógica proposicional

Puede decirse que la Lógica matemática es una teoría analítica del arte de razonar, y uno de sus principales objetivos es sistematizar (codificar) los principios que rigen los razonamientos válidos. Surge de la forma en que usamos el lenguaje para argumentar y persuadir, y se basa en la identificación de las partes esenciales de este lenguaje que se requieren para tal propósito. Es, en este sentido, una *Teoría axiomática intuitiva*, que tiene como uno de sus propósitos fundamentales el de clasificar los razonamientos dentro de dos clases: los válidos y los no válidos.

De una manera informal, diremos que un razonamiento es válido cuando nos permite obtener conclusiones verdaderas si uno ha comenzado con proposiciones verdaderas (las hipótesis). En cambio, un razonamiento que a partir de proposiciones verdaderas produzca conclusiones falsas, no es un razonamiento válido.

En este texto daremos una pequeña introducción al tema de la Lógica Matemática. El lector que quiera profundizar, puede consultar: [7], [18].

1.1 Conceptos primitivos. Verdad, falsedad

Comenzaremos con los conceptos primitivos de Falso (F o 0) y de Verdadero (V o 1). Decimos que ambos conceptos son primitivos porque no los explicamos en términos de conceptos más elementales.

Es claro que el proceso de “explicar” no puede ser infinito, porque entonces nunca podríamos hablar de nada, nos la pasaríamos “explicando” cada concepto usado e inventando nuevos. Para dar una imagen de esto,

intenten decir lo que es una casa. Al hacer esto han tenido que usar algunas palabras, que tendrían que explicarse a su vez, etc.

Por ejemplo, los diccionarios al explicarnos el significado de algo apelan a cierto conocimiento (lo correspondiente al concepto “primitivo”) previo que tiene el lector, porque siempre se cae en descripciones que usan alguna palabra que no se ha explicado.

- **Persona:** un ser humano.
- **Humano :** que consiste o esta producido por gente.
- **Gente:** Un grupo de **personas** con lazos tradicionales comunes.

Ejemplo 1 *Otro ejemplo tomado de un conocido diccionario de la Lengua española:*

- **Aumento:** acrecentamiento de una cosa.
- **Acrecentamiento:** **Aumento.**

Como vimos en el ejemplo previo, si no supiéramos el significado de alguna de las dos palabras, aumento o acrecentamiento, el diccionario nos dejaría en la misma situación.

Así pues, no definiremos lo que significan las palabras falso y verdadero y supondremos que todos tenemos un concepto primitivo de ellas.

En particular, debemos estar de acuerdo en que una afirmación **no puede ser falsa y verdadera a la vez**.

Esperamos que todos estemos de acuerdo en que

“Un perro es un mamífero”

es una afirmación verdadera y en que

“México es el país con mayor número de habitantes”

es una afirmación falsa.

Definición 1 *Diremos que una proposición es cualquier afirmación de la que pueda decidirse si es falsa o verdadera.*

Así que aceptaremos en que “Un perro es un mamífero” y en que “México es el país con mayor número de habitantes” son proposiciones.

En cambio, la frase:

$$\text{“Esta frase es falsa”,} \quad (1.1)$$

que es una frase rara puesto que habla de sí misma, **no es una proposición** ya que no se puede clasificar como falsa o verdadera. (Piénsese en lo siguiente: cuando uno declara que es falsa, inmediatamente algo nos dice que resulta verdadera. Pero en el momento en que la va a dar uno por verdadera, se vuelve falsa).

En las siguientes líneas diga si son proposiciones las afirmaciones siguientes, en caso afirmativo, diga si la proposición correspondiente es falsa o verdadera.

Ejercicio 1

1. Hay un número mayor que todos los demás.
2. Todos los mexicanos hablan español.
3. Todos los españoles hablan en castellano.
4. Si un animal pone huevos entonces es un ave o es un reptil.
5. Si un animal tiene ocho patas entonces es una araña.
6. ¡Viva México!
7. (Orwell) Todos los animales son iguales pero hay algunos que son más iguales que otros.
8. En un rincón del patio adonde nadie va, me gusta ir a subirme en la barda a pensar.
9. Todo cambia, pero no cambia mi amor ni el recuerdo de mi gente.
10. (Santa Teresa) Muero porque no muero.

Ejercicio 2 . *Esta afirmación no es una proposición.*

1. Si $2 + 2 = 4$ entonces $3 \cdot 3 = 9$.

2. Si $3 \cdot 3 = 30$ entonces $2 + 2 = 4$.
3. Si $2 + 2 \neq 4$ entonces $3 \cdot 3 = 10$.
4. Si $2 + 2 \neq 4$ entonces $3 \cdot 3 = 9$.

Ejercicio 3

1. La afirmación siguiente es verdadera.
2. La afirmación anterior es falsa.

Ejercicio 4

1. La afirmación siguiente es falsa.
2. La afirmación anterior es verdadera.

Ejercicio 5

1. La afirmación siguiente es falsa.
2. La afirmación anterior es falsa.

Ejercicio 6 . (*Si encuentra algún inciso difícil, vea el ejercicio 30*).

1. Sócrates era inteligente.
2. Si Sócrates era griego, entonces Sócrates era inteligente.
3. Si Sócrates era tonto entonces Sócrates era tonto.
4. Si Sócrates era inteligente entonces Sócrates era tonto.
5. Si Sócrates era tonto entonces Sócrates era inteligente.

Para denotar las proposiciones usaremos letras como

$$p, q, r, s, t.$$

Para que no se nos acaben, podemos usar mayúsculas

$$(P, Q, R, S)$$

o índices:

$$p_1, p_2, q_3, \dots$$

1.2 Conectivos lógicos y tablas de verdad

Para formar nuevas proposiciones a partir de otras usaremos los conectivos lógicos:

“ \neg ”, “ \wedge ”, “ \vee ”, “ \Rightarrow ”, “ \Leftrightarrow ”.

De tal manera que si p, q son proposiciones, entonces también lo son:

$$\neg p, p \wedge q, p \vee q, p \Rightarrow q, p \Leftrightarrow q,$$

notemos también que nadie prohíbe tomar $p = q$.

La calidad de falsa o verdadera que tiene una proposición como las anteriores, depende de la calidad de falso o verdadero que tengan las proposiciones que la componen.

Definimos los conectivos lógicos mediante “tablas de verdad”.

La negación

El conectivo \neg .

Si p es una proposición, denotaremos su negación por $\neg p$. (se lee: no p).

Definición 2 . “ \neg ” queda definido por medio de la tabla:

p	$\neg p$
0	1
1	0

Esta tabla nos dice que $\neg p$ es una proposición que es verdadera cuando p es falsa, o bien es falsa cuando p es verdadera.

Por ejemplo, la negación de “Un perro es un mamífero” es la proposición (falsa) “Un perro no es un mamífero”.¹

¹Cabe mencionar que en nuestro idioma hay veces que una expresión como “no somos nada” está lejos de significar la negación de que somos nada: $\neg(\text{somos nada})$. Sino que significa que somos nada, dicho con énfasis.

De la misma manera, “ni nadie” significa “nadie”: ¡Ni tú ni nadie lo van a impedir! ¡Pasó!, pidieron los tigres. ¡Ni nunca!, contestaron las mantarayas.

La conjunción

Enseguida definiremos la conjunción de dos proposiciones. Si p, q son proposiciones, su conjunción

$$p \wedge q$$

(léase: p y q) es la proposición que es verdadera cuando tanto p como q son verdaderas.

Definición 3. “ \wedge ” se define mediante la tabla:

p	q	$p \wedge q$
0	0	0
0	1	0
1	0	0
1	1	1

“ \wedge ” se llama conjunción.

$p \wedge q$ es falsa si alguna de las dos proposiciones (o ambas) es falsa.

La disyunción

Cuando p, q son proposiciones, podemos formar su disyunción, $p \vee q$, que se lee: p o q .

$p \vee q$ es falsa sólo cuando tanto p como q son falsas. Así, basta con que una de las proposiciones sea verdadera para que su disyunción sea verdadera.

En este sentido, la disyunción que se usa en la Lógica difiere del uso que se le da a la disyunción en el lenguaje cotidiano, en el que es frecuente que “o” se use en un sentido excluyente. En lenguaje cotidiano, en una expresión del tipo “hoy comeremos carne u hoy comeremos verduras” va implícito que sólo sucederá una de las dos posibilidades.

En cambio, en el lenguaje de la lógica $((2 + 2 = 4) \vee (3 \cdot 3 = 9))$ es verdadera y ambas proposiciones son verdaderas $((2 + 2 = 4), (3 \cdot 3 = 9))$.

El conectivo “ \vee ” se llama disyunción.

Definición 4 “ \vee ” se define mediante la tabla:

p	q	$p \vee q$
0	0	0
0	1	1
1	0	1
1	1	1

$p \vee q$ es falsa cuando tanto p como q lo son.

El conectivo implicación

Dadas las proposiciones p, q , definimos la proposición $p \Rightarrow q$ (p implica q) como la proposición que es cierta cuando p es falsa o q es verdadera. Así que si uno tiene que tanto $p \Rightarrow q$ como p son verdaderas, sería porque q es verdadera.

Definición 5 . El conectivo “ \Rightarrow ”, se define por medio de la tabla:

p	q	$p \Rightarrow q$
0	0	1
0	1	1
1	0	0
1	1	1

$p \Rightarrow q$ se lee: “ p implica q ”,
si p entonces q ;
 “ p sólo si q ”,
 “ q si p ”,
 “ p es condición suficiente para q ”
 “ q es condición necesaria para p ”.

$p \Rightarrow q$ es falsa cuando p es verdadera y q es falsa.

Recuérdese que si p es falsa, entonces la proposición $p \Rightarrow q$ es verdadera (“falso implica lo que sea” o: “de una proposición falsa se puede concluir cualquier proposición”).

Recuérdese también que si q es verdadera, entonces la proposición $p \Rightarrow q$ es verdadera.

Así que

1. Si $\sqrt{2}$ es racional entonces $5 > 7$.
2. Si $3 = 4$ entonces $2 + 2 = 4$,

son ambas proposiciones verdaderas.

El conectivo \Leftrightarrow

Definición 6 . $p \Leftrightarrow q$ tiene la misma tabla de verdad que $(p \Rightarrow q) \wedge (q \Rightarrow p)$:

p	q	$p \Rightarrow q$	$q \Rightarrow p$	$(p \Rightarrow q) \wedge (q \Rightarrow p)$
0	0	1	1	1
0	1	1	0	0
1	0	0	1	0
1	1	1	1	1

(1.2)

Así que omitiendo dos columnas podemos escribir:

p	q	$p \Leftrightarrow q$
0	0	1
0	1	0
1	0	0
1	1	1

(1.3)

Así que $p \Leftrightarrow q$ es verdadera cuando los valores de verdad de p y de q coinciden.

$p \Leftrightarrow q$ se lee:

“ p si y sólo si q ”

“ p es equivalente a q ”

“ p es condición necesaria y suficiente para q ”.

1.3 Tautologías y absurdos

Hagamos ahora algunas tablas de verdad:

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

p	\wedge	$(q$	\vee	$r)$	\Leftrightarrow	$(p$	\wedge	$q)$	\vee	$(p$	\wedge	$r)$
0	0	0	0	0	1		0		0		0	
0	0	0	1	1	1		0		0		0	
0	0	1	1	0	1		0		0		0	
0	0	1	1	1	1		0		0		0	
1	0	0	0	0	1		0		0		0	
1	1	0	1	1	1		0		1		1	
1	1	1	1	0	1		1		1		0	
1	1	1	1	1	1		1		1		1	

$$2. \neg(p \wedge q) \Leftrightarrow (\neg p) \vee (\neg q)$$

\neg	$(p \wedge q)$	\Leftrightarrow	$(\neg p \vee \neg q)$	
1	0	0	0	1
1	0	0	1	1
1	1	0	0	1
0	1	1	1	1
0	1	1	1	0
1	1	0	0	1
0	0	0	0	0

Notemos en las dos tablas anteriores, que las dos proposiciones siempre tienen el valor de verdad 1. Las proposiciones con esta propiedad, que son verdaderas independientemente de los valores de verdad de sus proposiciones componentes, se llaman tautologías.

En el otro extremo, las proposiciones que son falsas independientemente de los valores de verdad de sus proposiciones componentes, se llaman contradicciones o absurdos, frecuentemente las denotamos por ∇ .

Ejercicio 7 . Verifique que $p \vee (q \wedge r) \Leftrightarrow (p \vee q) \wedge (p \vee r)$ es una tautología.

Ejercicio 8 . Verifique que $p \vee (\neg p)$ es una tautología.

Ejercicio 9 . Verifique que $[p \wedge (p \Rightarrow q)] \Rightarrow q$ es una tautología.

Ejercicio 10 . Verifique que $p \Rightarrow p$ es una tautología.

Ejercicio 11 . Verifique que $p \wedge \neg p$ es un absurdo.

Ejercicio 12 . Verifique que $(p \Rightarrow \nabla) \wedge p$ es un absurdo.

Ejercicio 13 . Verifique que $(p \Rightarrow q) \Rightarrow \neg p$ no es ni tautología ni absurdo.

Ejemplos 2 . Algunas otras tautologías importantes son las siguientes:

1. $p \wedge (q \vee r) \Leftrightarrow (p \wedge q) \vee (p \wedge r)$ (distributividad de “ \wedge ” sobre “ \vee ”).
2. $p \vee (q \wedge r) \Leftrightarrow (p \vee q) \wedge (p \vee r)$ (distributividad de “ \vee ” sobre “ \wedge ”).
3. $p \wedge (q \wedge r) \Leftrightarrow (p \wedge q) \wedge r$ (asociatividad).
4. $p \vee (q \vee r) \Leftrightarrow (p \vee q) \vee r$ (asociatividad).

5. $\neg(p \vee q) \Leftrightarrow (\neg p \wedge \neg q)$. (ley de De Morgan).

6. $\neg(p \wedge q) \Leftrightarrow (\neg p \vee \neg q)$ (ley de De Morgan).

7. $(p \vee q) \Leftrightarrow (q \vee p)$ (commutatividad).

8. $(p \wedge q) \Leftrightarrow (q \wedge p)$ (commutatividad).

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

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

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

En vista de las asociatividades de “ \wedge ”, y “ \vee ”, se puede preguntar uno si “ \Rightarrow ” también lo es. Explícitamente nos preguntamos si

$$[p \Rightarrow (q \Rightarrow r)] \Leftrightarrow [(p \Rightarrow q) \Rightarrow r]$$

es una tautología.

Para resolver esto, tenemos dos maneras de proceder: una, haciendo la tabla de verdad completa con sus 8 renglones y la otra es tratar de encontrar valores de verdad que hagan los valores respectivos de $p \Rightarrow (q \Rightarrow r)$ y de $(p \Rightarrow q) \Rightarrow r$, distintos.

Por ejemplo, $p \Rightarrow (q \Rightarrow r)$ es falsa si p es verdadera, q es verdadera y r es falsa. Para estos mismos valores de verdad tenemos que $(p \Rightarrow q) \Rightarrow r$ también es falsa.

Veamos ahora si podemos asignar valores de verdad que hagan falsa $(p \Rightarrow q) \Rightarrow r$ pero que hagan verdadera a $p \Rightarrow (q \Rightarrow r)$.

Una posibilidad es con p falsa, con r falsa y con cualquier valor de verdad para q :

$$[(p \Rightarrow q) \Rightarrow r] \Leftrightarrow [p \Rightarrow (q \Rightarrow r)]$$

0	1	0	0	0	0	0	1		1	
0	1	1	0	0	0	0	1		0	

así vemos que “ \Rightarrow ” no es asociativa.

Proposición 1 . $p \Rightarrow (q \Rightarrow p)$ es una tautología.

Demostración. Obsérvese la siguiente tabla

p	\Rightarrow	$(q \Rightarrow p)$		
0	1	0	1	0
1	1	0	1	1
0	1	1	0	0
1	1	1	1	1

■

Ejercicio 14 . Muestre que $(p \wedge q) \Rightarrow p$ (Simplificación) es una tautología.

Ejercicio 15 . Muestre que $p \Rightarrow (p \vee q)$ (Adición) es una tautología.

1.4 Sistemas completos de conectivos

Revisemos la lista de nuestros conectivos: “ \neg ”, “ \wedge ”, “ \vee ”, “ \Rightarrow ”, “ \Leftrightarrow ”.

Algunos de ellos se pueden definir en términos de los demás: por ejemplo, “ \Leftrightarrow ” se puede definir en términos de “ \Rightarrow ” y de “ \wedge ”, de la manera siguiente:

$$p \Leftrightarrow q \equiv (p \Rightarrow q) \wedge (q \Rightarrow p).$$

(Con el símbolo \equiv indicamos que se está haciendo una definición),

Veamos ahora que los demás conectivos también se pueden definir usando sólo “ \neg ” y “ \wedge ” .

$$p \vee q \Leftrightarrow \neg[\neg(p \vee q)] \Leftrightarrow \neg[\neg p \wedge \neg q].$$

Por lo que podríamos definir el conectivo “ \vee ” de la manera siguiente:

$$p \vee q \equiv \neg[\neg p \wedge \neg q].$$

Ahora,

$$(p \Rightarrow q) \Leftrightarrow \neg[\neg(p \Rightarrow q)] \Leftrightarrow \neg[p \wedge \neg q].$$

Así que “ \vee ”, “ \wedge ”, “ \Rightarrow ” quedan definidos en términos de “ \neg ” y de “ \wedge ”.

No hemos definido todos los posibles conectivos, ¿cuántos conectivos binarios hay? Tantos como tablas de la forma

p	*	q
0	?	0
0	?	1
1	?	0
1	?	1

Es claro que como para cada renglón hay 2 posibles definiciones, se tienen 16 posibles conectivos binarios (¡bastantes más que los 4 que hemos usado: “ \Rightarrow ”, “ \wedge ”, “ \vee ”, “ \Leftrightarrow ”!).

Proposición 2 . *No importa como se defina el conectivo $*$, éste se puede describir en términos de “ \neg ” y de “ \wedge ” :*

Demostración. Ya vimos que “ \vee ”, “ \Rightarrow ”, “ \Leftrightarrow ” se puede describir en términos de “ \neg ” y de “ \wedge ”

Veamos la tabla de verdad de $p \wedge q$:

p	\wedge	q
0	0	0
0	0	1
1	0	0
1	1	1

$p \wedge q$ tiene un 1 en el 4º renglón y ceros en los demás renglones. Análogamente, $\neg p \wedge \neg q$ tiene un 1 en el primer renglón y 0 en los demás. De manera análoga tenemos:

$\neg p$	\wedge	$\neg q$	$\neg p$	\wedge	q	p	\wedge	$\neg q$	p	\wedge	q	p	\wedge	$(\neg q \wedge q)$			
1	1	1	1	0	0	0	0	1	0	0	0	0	0	1	0	0	0
1	0	0	1	1	1	0	0	0	0	0	1	1	0	0	0	0	1
0	0	1	0	0	0	1	1	1	1	0	0	1	0	1	0	1	0
0	0	0	0	0	1	1	0	0	1	1	1	1	0	0	0	0	1

Observando la tabla anterior denotemos r_1, r_2, r_3, r_4, r_5 las proposiciones de la lista.

Consideremos un conectivo $*$.

Si $p * q$ es absurda, entonces es equivalente a $p \wedge q \wedge \neg q$. Si no lo es y tiene 1 en los renglones i_1, \dots, i_k entonces $p * q$ es equivalente con

$$r_{i_1} \vee \dots \vee r_{i_k}.$$

■

Ejemplo 3 . *Por ejemplo, si se definiera*

p	*	q
0	0	0
0	1	1
1	0	0
1	1	1

entonces:

$$p * q \Leftrightarrow r_2 \vee r_4 \Leftrightarrow (\neg p \wedge q) \vee (p \wedge q).$$

En efecto,

$(\neg p \wedge q)$			\vee				$(p \wedge q)$		
1	0	0	0	0	0	0	0	0	0
1	1	1	1	0	0	1	0	0	1
0	0	0	0	1	0	0	1	0	0
0	1	1	1	1	1	1	1	1	1

En vista de lo anterior, decimos que “ \neg ” e “ \wedge ” forman un sistema completo de conectivos, ya que, dada cualquier tabla de verdad con valores asignados, se puede construir una proposición que corresponda a esa tabla de verdad, utilizando únicamente los conectivos “ \neg ” e “ \wedge ”.

Veamos ahora que “ \neg ” e “ \Rightarrow ” es un sistema completo de conectivos:

$$p \wedge q \Leftrightarrow p \wedge \neg(\neg q) \Leftrightarrow \neg[p \Rightarrow (\neg q)].$$

$$p \vee q \Leftrightarrow \neg(\neg(p \vee q)) \Leftrightarrow \neg(\neg p \wedge \neg q) \Leftrightarrow \neg(\neg[\neg p \Rightarrow q]) \Leftrightarrow \neg p \Rightarrow q.$$

Ejercicio 16 . Demuestre que “ \neg ” y “ \vee ” es un sistema completo de conectivos.

Surge de manera natural la siguiente pregunta:

¿Habrá un sistema completo que conste de un sólo conectivo?

Si lo hubiera, tendría que ser binario. (Pues si \otimes es un conectivo unario, entonces dadas dos proposiciones p, q , podríamos obtener las proposiciones

$$\otimes p, \otimes \otimes p, \otimes \otimes \otimes p, \dots$$

y por otra parte las proposiciones

$$\otimes q, \otimes \otimes q, \otimes \otimes \otimes q, \dots$$

pero nunca una proposición del tipo $p \vee q$ ó $p \Rightarrow q$.

Ahora, notemos por ejemplo, que “ \Rightarrow ” no forma por sí mismo un conjunto completo de conectivos. Esto se debe a que

$$p \Rightarrow p$$

es una tautología,

$$(T \Rightarrow p) \Leftrightarrow p,$$

y

$$(p \Rightarrow T) \Leftrightarrow T.$$

Notemos que de esta manera, a partir de p y de “ \Rightarrow ”, las únicas proposiciones que podemos obtener son tautologías o proposiciones equivalentes a p . De esta manera, no podemos obtener $\neg p$.

Ejercicio 17 . *Mostrar que “ \vee ” no es un sistema completo de conectivos.*

Ejercicio 18 . *Mostrar que “ \wedge ” no es un sistema completo de conectivos.*

Intentemos encontrar un sistema completo que conste de un sólo conectivo.

Tomemos

p	★	q
0	1	0
0	1	1
1	1	0
1	0	1

Note que $p★q$ es equivalente a la negación de $p \wedge q$.

Así pues,

$$p★p \Leftrightarrow \neg(p \wedge p) \Leftrightarrow \neg p.$$

Vemos que la negación queda definida en términos de \star .

Además, como

$$p★q \Leftrightarrow \neg(p \wedge q),$$

entonces:

$$\begin{aligned}
 (p★q)★(p★q) &\Leftrightarrow \\
 &\Leftrightarrow \neg[(p★q) \wedge (p★q)] \Leftrightarrow \\
 &\Leftrightarrow \neg[\neg(p \wedge q) \wedge \neg(p \wedge q)] \Leftrightarrow \\
 &\Leftrightarrow \neg[\neg(p \wedge q)] \Leftrightarrow \\
 &\Leftrightarrow p \wedge q.
 \end{aligned}$$

Así que “ \wedge ” queda definido en términos de \star . Como “ \neg ” y “ \wedge ” es un sistema completo de conectivos, entonces \star también lo es.

Ejercicio 19 . Encontrar otro conectivo que por sí mismo sea un sistema completo de conectivos. (¿O ya no hay otro? ¿Por qué?).

Ejercicio 20 . Construya una proposición P compuesta por p, q, r tal que tenga la siguiente tabla de verdad:

p	q	r	P
0	0	0	1
0	0	1	0
0	1	0	1
0	1	1	1
1	0	0	1
1	0	1	0
1	1	0	0
1	1	1	1

Ejercicio 21 . Construya una proposición Q compuesta por p, q, r tal que tenga la siguiente tabla de verdad:

p	q	r	Q
0	0	0	1
0	0	1	0
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	0
1	1	0	1
1	1	1	1

1.5 Reglas de inferencia, deducciones

Una regla de inferencia es una sucesión de proposiciones

$$p_1, p_2, \dots, p_n, q$$

tales que

$$(p_1 \wedge p_2 \wedge \dots \wedge p_n) \Rightarrow q$$

es una tautología.

A las proposiciones p_1, p_2, \dots, p_n se les llama premisas y a q se le llama la conclusión. Una regla de inferencia suele esquematizarse de la siguiente manera:

$$\frac{\begin{array}{c} p_1 \\ p_2 \\ \vdots \\ p_n \end{array}}{\therefore q}$$

$$\frac{p}{\therefore p} \quad (1.4)$$

Ejemplo 4

$$\frac{p \quad q}{\therefore p \wedge q} \quad \text{Conjunción.} \quad (1.5)$$

Ejemplo 5

$$\frac{}{\therefore p \wedge p} \quad , \quad \frac{p \wedge p}{\therefore p} \quad \text{Idempotencia.} \quad (1.6)$$

Ejemplo 6

$$\frac{}{\therefore p \vee p} \quad , \quad \frac{p \vee p}{\therefore p} \quad \text{Idempotencia.} \quad (1.7)$$

Ejemplo 7

$$\frac{p \wedge q}{\therefore p} \quad , \quad \frac{p \wedge q}{\therefore q} \quad \text{Simplificación.} \quad (1.8)$$

Ejemplo 8

$$\frac{p}{\therefore p \vee q} \quad , \quad \text{Adición.} \quad (1.9)$$

Ejemplo 9

$$\frac{p}{\therefore \neg \neg p} \quad , \quad \frac{\neg \neg p}{\therefore p} \quad \text{Doble negación.} \quad (1.10)$$

Ejemplo 10

$$\frac{p \wedge q}{\therefore q \wedge p} , \frac{p \vee q}{\therefore q \vee p} \quad \text{Commutatividad.} \quad (1.11)$$

Ejemplo 11

$$\frac{\begin{array}{c} p \\ p \Rightarrow q \\ \hline \therefore q \end{array}}{\therefore q} \quad \text{Modus ponens.} \quad (1.12)$$

Ejemplo 12

$$\frac{\begin{array}{c} p \vee q \\ \neg q \\ \hline \therefore p \end{array}}{\therefore p} \quad \text{Tollendo ponens (negando, afirmo).} \quad (1.13)$$

Ejemplo 13

$$\frac{\begin{array}{c} p \Rightarrow q \\ \neg q \\ \hline \therefore \neg p \end{array}}{\therefore \neg p} \quad \text{Tollendo tollens.} \quad (1.14)$$

Ejemplo 14

$$\frac{p \Rightarrow q}{\therefore \neg q \Rightarrow \neg p} \quad \text{Contrapuesta.} \quad (1.15)$$

Ejemplo 15

$$\frac{\begin{array}{c} p \Rightarrow q \\ q \Rightarrow r \\ \hline \therefore p \Rightarrow r \end{array}}{\therefore p \Rightarrow r} \quad \text{Silogismo hipotético.} \quad (1.16)$$

Ejemplo 16

$$\frac{p \Rightarrow (q \Rightarrow r)}{\therefore (p \Rightarrow q) \Rightarrow (p \Rightarrow r)} \quad \frac{(p \Rightarrow q) \Rightarrow (p \Rightarrow r)}{\therefore p \Rightarrow (q \Rightarrow r)} \quad \text{.“} \Rightarrow \text{” se distribuye sobre “} \Rightarrow \text{”.} \quad (1.17)$$

Ejemplo 17

$$\frac{\begin{array}{c} (p \Rightarrow q) \vee (r \Rightarrow s) \\ \hline \frac{\begin{array}{c} p \\ r \\ \hline \therefore q \vee s \end{array}}{q \vee s} \end{array}}{q \vee s} \quad \text{Dilema constructivo.} \quad (1.18)$$

Ejemplo 18

$$\begin{array}{c}
 (p \Rightarrow q) \vee (r \Rightarrow s) \\
 \dfrac{\begin{array}{c} \neg q \\ \hline \neg s \end{array} \quad \text{Dilema destructivo.}}{\therefore \neg p \vee \neg r}
 \end{array} \quad (1.19)$$

Ejemplo 19

$$\dfrac{p \Rightarrow (q \Rightarrow r)}{\therefore (p \wedge q) \Rightarrow r} \quad \dfrac{(p \wedge q) \Rightarrow r}{\therefore p \Rightarrow (q \Rightarrow r)} \quad . \text{ Exportación.} \quad (1.20)$$

Verifiquemos esto último:

p	\Rightarrow	$(q$	\Rightarrow	$r)$	\Leftrightarrow	$(p$	\wedge	$q)$	\Rightarrow	r
0	1	0	1	0	1		0		1	0
1	1	0	1	0	1		0		1	0
0	1	1	0	0	1		0		1	0
1	0	1	0	0	1		1		0	0
0	1	0	1	1	1		0		1	1
1	1	0	1	1	1		0		1	1
0	1	1	1	1	1		0		1	1
1	1	1	1	1	1		1		1	1

1.5.1 Regla del reemplazo

Proposición 3 . Si en una proposición compuesta Q , tal que una de sus proposiciones componentes es p , entonces al reemplazar p por una proposición equivalente p' se obtiene una proposición Q' equivalente a Q .

Bosquejo de demostración:

Notemos los siguientes casos particulares:

1. Si $Q =: \neg p$, y $p \equiv p'$ entonces $Q' =: \neg p'$.

p	\equiv	p'	$\neg p$	$\neg p'$	$\neg p \Leftrightarrow \neg p'$
0	1	0	1	1	1
1	1	1	0	0	1

2. Si $Q =: p \vee q$, y $p \equiv p'$ entonces $Q' =: p' \vee q$.

q	p	\equiv	p'	$p \vee q$	$p' \vee q$	$Q \Leftrightarrow Q'$
0	0	1	0	0	0	1
0	1	1	1	1	1	1
1	0	1	0	1	1	1
1	1	1	1	1	1	1

3. Si $Q =: p \wedge q$, y $p \equiv p'$ entonces $Q' =: p' \wedge q$.

q	p	\equiv	p'	$p \wedge q$	$p' \wedge q$	$Q \Leftrightarrow Q'$
0	0	1	0	0	0	1
0	1	1	1	0	0	1
1	0	1	0	0	0	1
1	1	1	1	1	1	1

4. Si $Q =: p \Rightarrow q$, y $p \equiv p'$ entonces $Q' =: p' \Rightarrow q$.

q	p	\equiv	p'	$p \Rightarrow q$	$p' \Rightarrow q$	$Q \Leftrightarrow Q'$
0	0	1	0	1	1	1
0	1	1	1	0	0	1
1	0	1	0	1	1	1
1	1	1	1	1	1	1

5. Si $Q =: q \Rightarrow p$, y $p \equiv p'$ entonces $Q' =: q \Rightarrow p'$.

q	p	\equiv	p'	$q \Rightarrow p$	$q \Rightarrow p'$	$Q \Leftrightarrow Q'$
0	0	1	0	1	1	1
0	1	1	1	1	1	1
1	0	1	0	0	0	1
1	1	1	1	1	1	1

Toda vez que hemos notado los 5 casos anteriores, veamos lo que tienen en común: en Q aparece solamente un conectivo.

Las proposiciones compuestas Q con dos conectivos en las que aparece p se pueden construir a partir de los casos anteriores:

Por ejemplo, a partir de $\neg p$, podemos construir

$$\neg(\neg p), (\neg p) \wedge q, (\neg p) \vee q, (\neg p) \Rightarrow q \text{ y } q \Rightarrow (\neg p).$$

Como ya sabemos que en caso de que $p \Leftrightarrow p'$, $\neg p \Leftrightarrow \neg p'$, podemos sustituir $\neg p$ por $\neg p'$, para obtener una proposición equivalente a la original.

Otro caso:

Si partimos de $(p \Rightarrow q)$, podemos obtener:

1. $\neg(p \Rightarrow q)$,
2. $(p \Rightarrow q) \wedge r, r \wedge (p \Rightarrow q)$,
3. $(p \Rightarrow q) \vee r, r \vee (p \Rightarrow q)$,
4. $(p \Rightarrow q) \Rightarrow r, r \Rightarrow (p \Rightarrow q)$.

Además de las que se obtienen si en lugar de r , ponemos p o ponemos q .

Notemos ahora que al sustituir p por una proposición equivalente p' , obtenemos $p' \Rightarrow q$ (equivalente a $p \Rightarrow q$). Entonces podemos sustituir $(p \Rightarrow q)$ por $(p' \Rightarrow q)$ para obtener una proposición equivalente a la original (usando el caso de un solo conectivo).²

De manera análoga podemos analizar todos los casos en donde Q tiene dos conectivos.

Si Q tuviera tres conectivos entonces $Q =: \neg U$, donde U tiene sólo dos conectivos y en donde se puede sustituir p por p' , o bien $Q =: U \wedge V, Q =: U \vee V, Q =: U \Rightarrow V$, donde U y V son proposiciones con menos de tres conectivos. Así que en U y V , se puede reemplazar p por p' .

Si Q tiene 4 conectivos, se puede reducir a los casos anteriores en los que es válido sustituir p por p' .

Continuando de esta manera uno puede ver que en Q , se puede sustituir p por la proposición equivalente p' . ■

1.5.2 Regla de la tautología

Proposición 4 (Regla de la tautología) . *En una regla de inferencia válida, se puede incluir una tautología dentro de las hipótesis, obteniéndose una regla de inferencia válida.*

²Una demostración completamente rigurosa, se puede hacer por inducción matemática. La inducción se verá en el capítulo acerca de los números naturales. En ese momento, el lector podrá adaptar este argumento y podrá hacer una demostración con toda propiedad.

Demostración. Por definición, una regla de inferencia es una sucesión de proposiciones

$$p_1, p_2, \dots, p_n, q$$

tales que

$$(p_1 \wedge p_2 \wedge \dots \wedge p_n) \Rightarrow q$$

es una tautología.

Por la observación anterior basta ver que

$$(p_1 \wedge p_2 \wedge \dots \wedge p_n) \Leftrightarrow T \wedge (p_1 \wedge p_2 \wedge \dots \wedge p_n)$$

es una tautología si T es una tautología. Notemos que

$$(p_1 \wedge p_2 \wedge \dots \wedge p_n)$$

es falsa o bien es verdadera. Si es falsa también

$$T \wedge (p_1 \wedge p_2 \wedge \dots \wedge p_n)$$

lo es; si es verdadera, también

$$T \wedge (p_1 \wedge p_2 \wedge \dots \wedge p_n)$$

lo es. ■

Proposición 5 . *Sea $(p_1 \wedge p_2 \wedge \dots \wedge p_r) \Rightarrow s$, una tautología, entonces:*

1. Si $(p_1 \wedge p_2 \wedge \dots \wedge p_n) \Rightarrow q$ es una tautología con $n > r$, entonces $p_1 \wedge p_2 \wedge \dots \wedge p_r \wedge s \wedge p_{r+1} \wedge \dots \wedge p_n \Rightarrow q$ es una tautología.
2. Si $p_1 \wedge p_2 \wedge \dots \wedge p_r \wedge s \wedge p_{r+1} \wedge \dots \wedge p_n \Rightarrow q$ es una tautología, entonces $(p_1 \wedge p_2 \wedge \dots \wedge p_n) \Rightarrow q$ es una tautología.

En otras palabras: en una regla válida de inferencia se puede agregar *u omitir, dentro de las hipótesis*, una proposición que se obtenga de las anteriores hipótesis usando una regla válida de inferencia.

Demostración. 1. Supongamos que

$$(p_1 \wedge p_2 \wedge \dots \wedge p_n) \Rightarrow q$$

es una tautología, y sea s una proposición. Observemos que

$$([(p_1 \wedge p_2 \wedge \dots \wedge p_n) \wedge s] \Rightarrow q) \Leftrightarrow$$

$$(s \wedge [(p_1 \wedge p_2 \wedge \dots \wedge p_n)] \Rightarrow q) \Leftrightarrow$$

$$(s \Rightarrow [(p_1 \wedge p_2 \wedge \dots \wedge p_n)] \Rightarrow q) \Leftrightarrow$$

$$(s \Rightarrow \top) \Leftrightarrow \top.$$

(Con \top denotamos una tautología).

2. Afirmamos que si

$$p_1 \wedge p_2 \wedge \dots \wedge p_r \wedge s \wedge p_{r+1} \wedge \dots \wedge p_n \Rightarrow q$$

es una tautología y

$$(p_1 \wedge p_2 \wedge \dots \wedge p_r) \Rightarrow s$$

también lo es, entonces

$$(p_1 \wedge p_2 \wedge \dots \wedge p_n) \Rightarrow q$$

es una tautología.

Hagámonos la siguiente pregunta, ¿cómo podría suceder que $(p_1 \wedge p_2 \wedge \dots \wedge p_n) \Rightarrow q$ no fuera una tautología? Es claro que esto sólo podría pasar si q es falsa y $p_1 \wedge p_2 \wedge \dots \wedge p_n$ es verdadera. Ahora, $p_1 \wedge p_2 \wedge \dots \wedge p_n$ es verdadera si y sólo si cada p_i lo es, en particular las r primeras p_i lo serían. Pero si p_1, p_2, \dots, p_r son verdaderas, entonces también lo es s , en vista de que

$$[(p_1 \wedge p_2 \wedge \dots \wedge p_r) \Rightarrow s]$$

es una tautología. Entonces

$$p_1 \wedge p_2 \wedge \dots \wedge p_r \wedge s \wedge p_{r+1} \wedge \dots \wedge p_n$$

sería verdadera y como $[p_1 \wedge p_2 \wedge \dots \wedge p_r \wedge s \wedge p_{r+1} \wedge \dots \wedge p_n \Rightarrow q]$ es una tautología entonces q también sería verdadera.

El argumento anterior muestra que $((p_1 \wedge p_2 \wedge \dots \wedge p_n) \Rightarrow q)$ es una tautología, tal como queríamos (en caso contrario tendríamos que q es una proposición falsa y verdadera, lo que no hay). ■

Ejemplo 20 . Consideremos el siguiente argumento:

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

¿Es el argumento anterior, un argumento válido?

Una manera de ver que este argumento es válido (es decir, es una regla válida de inferencia) sería viendo que la siguiente proposición es una tautología:

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

La tabla completa ocuparía $2^5 = 32$ renglones y más de $16 \times 32 = 512$, ceros y unos. Hacer esta tabla sería algo fastidioso.

En lugar de hacer esto último, veremos que podemos llegar a la conclusión t , partiendo de las hipótesis y usando reglas válidas de inferencia. Como vimos en la observación anterior, este procedimiento está justificado.

Ejemplo 21 . Queremos ver que de las siguientes premisas (hipótesis) se puede inferir t .

1. $p \Rightarrow (q \Rightarrow r)$.
2. $p \Rightarrow (s \Rightarrow t)$.
3. $p \wedge (q \vee s)$.
4. $\neg r$.
5. $(p \wedge q) \Rightarrow r$ (Exportación 1, reemplazo).
6. $\neg(p \wedge q)$ (Tollendo tollens con 4 y 5).
7. $\neg p \vee \neg q$ (De Morgan, 6, reemplazo).
8. p (Simplificación, 3).
9. $\neg \neg p$ (Doble negación, 8).
10. $\neg q$ (Tollendo ponens, 9,7).

11. $(q \vee s)$ (Simplificación, 3).
12. s (Tollendo ponens, 10, 11).
13. $p \wedge s$ (Conjunción 8, 12).
14. $(p \wedge s) \Rightarrow t$ (Exportación, 2, reemplazo).
15. t (Modus ponens 13, 12).

Observando la lista anterior y usando la observación precedente, podemos concluir que

$$\begin{array}{c}
 p \Rightarrow (q \Rightarrow r) \\
 p \Rightarrow (s \Rightarrow t) \\
 p \wedge (q \vee s) \\
 \hline
 \therefore t
 \end{array}$$

es una regla válida de inferencia.

Ejercicio 22 . Vuelva a hacer el ejemplo anterior, pero usando menos pasos.

Con tal de no hacer toda la tabla de verdad, podríamos hacer lo siguiente: respondamos la pregunta ¿cómo podría ser que

$$[(p \Rightarrow (q \Rightarrow r)) \wedge (p \Rightarrow (s \Rightarrow t)) \wedge (p \wedge (q \vee s)) \wedge (\neg r)] \Rightarrow t$$

no fuera una tautología?

Solamente que t fuera falsa y que

$$(p \Rightarrow (q \Rightarrow r)) \wedge (p \Rightarrow (s \Rightarrow t)) \wedge (p \wedge (q \vee s)) \wedge (\neg r)$$

fuerá verdadera. Entonces r sería falsa, y p verdadera. Con estas suposiciones hacer una tabla ya no es difícil, pues solamente constaría de 4 renglones de ceros y unos:

s	q	p	r	t	$p \Rightarrow (q \Rightarrow r)$	$q \Rightarrow r$	$p \Rightarrow (s \Rightarrow t)$	$(s \Rightarrow t)$	$p \wedge (q \vee s)$	$\neg r$
0	0	1	0	0	1	1	1	1	0	1
0	1	1	0	0	0	0	1	1	1	1
1	0	1	0	0	1	1	0	0	1	1
1	1	1	0	0	0	0	0	0	1	1

Como vemos, no se cumple que todas las hipótesis sean verdaderas. Así que no es posible que la conclusión sea falsa con las hipótesis verdaderas.

Ejercicio 23 . Demuestre que “ \Rightarrow ” se distribuye sobre “ \vee ”. Es decir que

$$\frac{p \Rightarrow (q \vee r)}{\therefore (p \Rightarrow q) \vee (p \Rightarrow r)} \quad y \quad \frac{(p \Rightarrow q) \vee (p \Rightarrow r)}{\therefore p \Rightarrow (q \vee r)}$$

son reglas válidas de inferencia

Ejercicio 24 . Demuestre que “ \Rightarrow ” se distribuye sobre “ \wedge ”.

Es decir que

$$\frac{p \Rightarrow (q \wedge r)}{\therefore (p \Rightarrow q) \wedge (p \Rightarrow r)} \quad y \quad \frac{(p \Rightarrow q) \wedge (p \Rightarrow r)}{\therefore p \Rightarrow (q \wedge r)}$$

son reglas válidas de inferencia.

Ejercicio 25 . Demuestre que “ \vee ” se distribuye sobre “ \Rightarrow ”. Es decir que

$$\frac{p \vee (q \Rightarrow r)}{\therefore (p \vee q) \Rightarrow (p \vee r)} \quad y \quad \frac{(p \vee q) \Rightarrow (p \vee r)}{\therefore p \vee (q \Rightarrow r)}$$

son reglas válidas de inferencia.

Ejercicio 26 . Demuestre que “ \wedge ” **no** se distribuye sobre “ \Rightarrow ”.

Ejemplo 22 . Veamos que el siguiente argumento es válido:

$$\frac{\begin{array}{c} p \Rightarrow (q \vee r) \\ \neg q \\ \hline \therefore p \Rightarrow r \end{array}}{}$$

1. $p \Rightarrow (q \vee r)$
2. $\neg q$
3. $\neg (q \vee r) \Rightarrow \neg p$ (Contrapuesta de 1, reemplazo).
4. $((\neg q) \wedge (\neg r)) \Rightarrow (\neg p)$ (De Morgan, 3, reemplazo).
5. $(\neg q) \Rightarrow ((\neg r) \Rightarrow (\neg p))$ (Exportación, 4, reemplazo).

6. $\neg q \Rightarrow (p \Rightarrow r)$ (Contrapuesta, 5, reemplazo).
7. $(p \Rightarrow r)$ (Modus ponens, 2, 6, reemplazo).

Ejercicio 27 . Muestre que el siguiente argumento es válido:

1. $\neg p \Rightarrow s$.
 2. $p \Rightarrow q$.
 3. $\neg q \vee s$.
 4. $s \Rightarrow (p \Rightarrow r)$.
 5. $s \Rightarrow p$.
- $\therefore r$.

1.5.3 Negaciones

Hacemos aquí un resumen acerca de las negaciones acerca de algunas proposiciones.

La negación de $p \Rightarrow q$

Para empezar, la negación de $p \Rightarrow q$, $\neg(p \Rightarrow q)$ es equivalente a $p \wedge \neg q$. Intuitivamente, uno podría pensar que si $p \Rightarrow q$ significa “si pasa p entonces pasa q ” es natural que lo contrario es “pasa p pero no pasa q ”.

Otra manera de recordarlo es la siguiente: el único caso en que $p \Rightarrow q$ es falsa es cuando p es verdadera y q es falsa. Así que $p \Rightarrow q$ significa lo mismo que “no pasa p o pasa q ” o sea $\neg p \vee q$. Luego la negación de $p \Rightarrow q$ debe coincidir con la negación de $\neg p \vee q$, que es $\neg(\neg p \vee q) \equiv \neg\neg p \wedge \neg q \equiv p \wedge \neg q$, por las Leyes de De Morgan y de la Doble negación.

La negación de $p \Rightarrow q$ es $p \wedge \neg q$.

Ejercicio 28 . Compruebe que en efecto, $p \Rightarrow q$ es equivalente a $\neg p \vee q$.

Ejercicio 29 . Compruebe que en efecto, $\neg(p \Rightarrow q)$ es equivalente a $p \wedge \neg q$.

La negación de $p \wedge q$ y de $p \vee q$

Las leyes de De Morgan nos dicen que

$$\begin{aligned}\neg(p \wedge q) &\equiv \neg p \vee \neg q \\ \neg(p \vee q) &\equiv \neg p \wedge \neg q.\end{aligned}$$

La negación de $p \Leftrightarrow q$

Por definición,

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

así que su negación, de acuerdo con las leyes de De Morgan es

$$\neg(p \Leftrightarrow q) \equiv \neg(p \Rightarrow q) \vee \neg(q \Rightarrow p) \equiv (p \wedge \neg q) \vee (q \wedge \neg p).$$

Lo anterior es lo que uno puede esperar si se acuerda de que $p \Leftrightarrow q$ es verdadera, cuando los valores de verdad de p y de q coinciden.

Ejemplo 23 . *La proposición: “Si F es mexicano entonces habla español” es falsa porque, como sabemos, hay mexicanos que no hablan español, por ejemplo, muchos indígenas chiapanecos hablan idiomas (que no dialectos) mayances como el zeltal, zotzil, chol. Además los recién nacidos no hablan idiomas.*

Ejemplo 24 . *La proposición “Si n es un entero positivo entonces n es el cuadrado de otro entero” es falsa, pues, como veremos más adelante, 2 no es el cuadrado de ningún entero.*

Ejercicio 30 . *Repita el ejercicio 6, negando cada uno de sus incisos (por ejemplo, la negación de 5. Si Sócrates era tonto entonces Sócrates era inteligente, es “Sócrates era tonto y Sócrates no era inteligente” que claramente es una proposición falsa).*

1.5.4 Inferencias no válidas

Pudiera ser que se nos presentara un razonamiento del que no se sabe si es válido. En ese caso, antes de tratar de demostrar la validez, conviene hacer una tabla de verdad para decidir si la conclusión puede ser falsa y las premisas verdaderas.

Ejemplo 25 . Nos preguntamos por la validez del siguiente argumento:

$$\frac{p \Rightarrow q \\ q}{\therefore p.}$$

Veamos si

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

es una tautología.

Tratamos de ver si se puede tener la siguiente situación: p falsa, $(p \Rightarrow q)$ verdadera y q verdadera. En esta situación la respuesta es que sí, sin necesidad de hacer ninguna tabla.

La tabla es

$(p \Rightarrow q) \wedge q]$	p	$((p \Rightarrow q) \wedge q) \Rightarrow p$
1	1	1

Ejemplo 26 . Consideremos el siguiente argumento

$$\frac{(p \Rightarrow q) \vee r \\ p \wedge \neg r}{\therefore q.}$$

Veamos si es posible tener la conclusión falsa con las hipótesis verdaderas:

¿Es posible tener q falsa, $p \wedge \neg r$ verdadera, y $(p \Rightarrow q) \vee r$ verdadera?

Para que $p \wedge \neg r$ sea verdadera se necesita que p sea verdadera y r sea falsa.

Veamos ahora lo que sucede con $(p \Rightarrow q) \vee r$, para estos valores de verdad:

$(p \Rightarrow q) \vee r$
1

Así que el razonamiento debe ser válido y podemos implementar una deducción:

1. $(p \Rightarrow q) \vee r$
2. $p \wedge \neg r$
3. $\neg r$ (Simplificación 2).

4. $p \Rightarrow q$ (Tollendo ponens, 1,3).
5. p (Simplificación 2).
6. q (Modus ponens, 4,5).

Ejemplo 27 . Consideremos el siguiente argumento:

1. $p \Rightarrow q$
2. $p \vee q$

$\therefore q$

Si q es falsa, para que $p \vee q$ sea verdadera necesitamos que p sea verdadera, pero entonces ya no puede ser verdadera $p \Rightarrow q$. Así que el razonamiento debe ser válido, por lo que daremos una deducción de q .

1. $p \Rightarrow q$
2. $p \vee q$
3. $\neg p \vee q$ (Equivalente a 1, reemplazo).
4. $(p \vee q) \wedge (\neg p \vee q)$ (conjunción, 2, 3) .
5. $(p \wedge \neg p) \vee (q)$ (Distributividad, 4).
6. $\neg(p \wedge \neg p)$ (Tautología).
7. q (Tollendo ponens, 5, 6).

Ejercicio 31 . Muestre que el siguiente argumento no es válido

$$\begin{array}{c}
 p \Rightarrow q \\
 s \Rightarrow (p \Rightarrow r) \\
 r \vee \neg q \\
 \hline
 \end{array}
 \frac{s \Rightarrow p}{\therefore r}$$

Ejercicio 32 . Decida si el siguiente argumento es válido. Si no lo es, dé una asignación de valores de verdad que hagan falsa la conclusión pero las hipótesis verdaderas. Si lo es, haga una deducción

1. $p \Rightarrow (q \Rightarrow r)$
2. $p \wedge \neg r$
3. $(p \Rightarrow s) \vee (\neg r \Rightarrow s)$
4. $(\neg q \wedge s) \Rightarrow (q \vee t)$

$$\therefore t$$

Ejercicio 33 . Decida si el siguiente argumento es válido. Si no lo es, dé una asignación de valores de verdad que hagan falsa la conclusión pero las

$$\begin{aligned}
 & p \Rightarrow (q \Rightarrow r) \\
 & p \wedge \neg r \\
 & \text{hipótesis verdaderas. Si lo es, haga una deducción} \quad (p \Rightarrow s) \vee (\neg r \Rightarrow s) \\
 & \quad (\neg q \wedge s) \Rightarrow (q \vee t) \\
 & \quad \therefore t
 \end{aligned}$$

1.6 Reducción al absurdo

Hagamos la siguiente observación.

Observación 1 . Las siguientes proposiciones son equivalentes:

1. $p \wedge q \Rightarrow r \vee s.$
2. $p \Rightarrow \neg q \vee r \vee s.$
3. $p \wedge q \wedge \neg r \Rightarrow s.$

Antes de dar una demostración de esto, notemos que tiene alguna semejanza con el Álgebra elemental (...se pasa al otro lado cambiando de signo...).

Demostración.

$$1) \Rightarrow 2)$$

1. $p \wedge q \Rightarrow r \vee s$
2. $p \Rightarrow (q \Rightarrow (r \vee s))$ (Por exportación).
3. $p \Rightarrow (\neg q \vee (r \vee s))$ (Pues $\neg q \vee (r \vee s)$ es equivalente con $q \Rightarrow (r \vee s)$, ver el ejercicio 28).

2) \Rightarrow 3)

1. $p \Rightarrow \neg q \vee r \vee s$.
2. $p \Rightarrow (\neg q \vee r) \vee s$. (Por asociatividad).
3. $(p \Rightarrow (\neg q \vee r)) \vee (p \Rightarrow s)$. (“ \Rightarrow ” se distribuye sobre “ \vee ”).
4. $\neg\neg(p \Rightarrow (\neg q \vee r)) \vee (p \Rightarrow s)$. (Por doble negación y reemplazo).
5. $\neg(p \Rightarrow (\neg q \vee r)) \Rightarrow (p \Rightarrow s)$. (Por el ejercicio 28).
6. $p \wedge \neg(\neg q \vee r) \Rightarrow (p \Rightarrow s)$. (Por el ejercicio 29).
7. $(p \wedge q \wedge \neg r) \Rightarrow (p \Rightarrow s)$. (Por regla de De Morgan y reemplazo).
8. $(p \wedge p \wedge q \wedge \neg r) \Rightarrow s$. (Por exportación).
9. $(p \wedge q \wedge \neg r) \Rightarrow s$. (Por idempotencia y reemplazo).

3) \Rightarrow 1)

1. $p \wedge q \wedge \neg r \Rightarrow s$.
2. $p \wedge q \Rightarrow (\neg r \Rightarrow s)$. (Por exportación).
3. $p \wedge q \Rightarrow (\neg\neg r \vee s)$. (Por el ejercicio 28).
4. $p \wedge q \Rightarrow (r \vee s)$. (Por la regla de doble negación). ■

Recordemos ahora que denotamos por \top una tautología y por ∇ un absurdo.

Notemos además que

$$(p \wedge \top) \Leftrightarrow p$$

y que

$$\left(q \vee \nabla \right) \Leftrightarrow q$$

son tautologías.

Proposición 6 (*Reducción al absurdo*) *Las siguientes proposiciones son equivalentes:*

1. $p \Rightarrow q$
2. $p \wedge \top \Rightarrow q$
3. $p \wedge \neg q \Rightarrow \nabla_{\circ}$.

Demostración.

La siguiente es una lista de proposiciones equivalentes:

1. $p \Rightarrow q$
2. $p \wedge \top \Rightarrow q$ (Como acabamos de observar, $p \wedge \top \Leftrightarrow p$).
3. $p \Rightarrow q \vee \nabla_{\circ}$ (Por la observación 1).
4. $p \wedge \neg q \Rightarrow \nabla_{\circ}$. (Por la observación 1). ■

La proposición anterior suele enunciarse así:

“Para demostrar $p \Rightarrow q$, una alternativa es la siguiente: supóngase p y lo contrario de lo que se quiere probar (es decir, supóngase $\neg q$), y véase que esto lleva a contradicción”.

Corolario 1 .

$$\frac{\begin{array}{c} p_1 \\ p_2 \\ \vdots \\ p_n \end{array}}{\therefore q}$$

es una regla válida de inferencia si y sólo si

$$(p_1 \wedge p_2 \wedge \dots \wedge p_n) \wedge \neg q$$

nos lleva a contradicción, mediante reglas válidas de inferencia.

Demostración. Lo anterior se sigue de que

$$(p_1 \wedge p_2 \wedge \dots \wedge p_n) \Rightarrow q$$

es equivalente a

$$(p_1 \wedge p_2 \wedge \dots \wedge p_n) \Rightarrow (q \vee \nabla_{\circ})$$

lo que a su vez es equivalente a

$$(p_1 \wedge p_2 \wedge \dots \wedge p_n) \wedge \neg q \Rightarrow \nabla_{\circ},$$

Por la observación 1. ■

Ejemplo 28 . Repitamos el ejemplo 21, pero por reducción al absurdo.

Queremos ver que de las primeras 4 premisas (hipótesis) se puede inferir t .

1. $p \Rightarrow (q \Rightarrow r)$.
2. $p \Rightarrow (s \Rightarrow t)$.
3. $p \wedge (q \vee s)$.
4. $\neg r$.
5. $\neg t$ (Hipótesis adicional reducción al absurdo)..
6. $p \wedge s \Rightarrow t$ (Exportación, a 2).
7. $\neg(p \wedge s)$ (Tollendo tollens, 5, 6).
8. $\neg p \vee \neg s$ (De Morgan, 7).
9. p (Simplificación, 3).
10. $\neg s$ (Tollendo ponens, 8,9).
11. $q \Rightarrow r$ (Modus ponens, 1,9).
12. $\neg q$ (Tollendo tollens, 11,4).
13. $\neg q \wedge \neg s$ (Conjunción, 10,12).
14. $\neg(q \vee s)$ (De Morgan 13).
15. $(q \vee s)$ (Simplificación, 3).
16. $\neg(q \vee s) \wedge (q \vee s)$ (Conjunción, 14, 15).
17. ∇_{\circ} (Equivalente a la anterior). ■

Ejemplo 29 . Repitamos el ejemplo 22, pero por reducción al absurdo.

$$\frac{p \Rightarrow (q \vee r) \quad \neg q}{\therefore p \Rightarrow r}$$

Demostración. 1. $p \Rightarrow (q \vee r)$

2. $\neg q$
3. $\neg(p \Rightarrow r)$
4. $p \wedge \neg r$ (Equivalente a 1.6, en vista del ejercicio 29)
5. p (Simplificación, 1.6).
6. $(q \vee r)$ (Modus ponens, 1.6, 1.6).
7. $\neg r$ (Simplificación, 1.6).
8. q (Tollendo ponens, 1.6, 1.6).
9. $q \wedge \neg q$ (Conjunción, 1.6, 1.6).
10. ∇ (Equivalente a la anterior). ■

Convengamos en que

$$q \Leftarrow p, \text{ significa lo mismo que } p \Rightarrow q$$

Con esta convención, $(q \vee r) \Rightarrow p$, significa lo mismo que $p \Leftarrow (q \vee r)$.

Ejemplo 30 . “ \Leftarrow ” se distribuye sobre “ \vee ”

Es decir,

$$p \Leftarrow (q \vee r)$$

es equivalente a

$$(p \Leftarrow q) \vee (p \Leftarrow r).$$

Demostración. Para decidir esto, primero veremos que

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

y después veremos que

$$(p \Leftarrow (q \vee r)) \Leftarrow ((p \Leftarrow q) \vee (p \Leftarrow r)).$$

$\Rightarrow (p \Leftarrow (q \vee r)) \Rightarrow ((p \Leftarrow q) \vee (p \Leftarrow r))$ es equivalente a

$$[(p \Leftarrow (q \vee r)) \wedge \neg((p \Leftarrow q) \vee (p \Leftarrow r))] \Rightarrow \nabla_{\circ},$$

en vista de la observación 1.

1. $[(p \Leftarrow (q \vee r)) \wedge \neg((p \Leftarrow q) \vee (p \Leftarrow r))]$
2. $(p \Leftarrow (q \vee r))$
3. $\neg((p \Leftarrow q) \vee (p \Leftarrow r))$
4. $\neg(p \Leftarrow q) \wedge \neg(p \Leftarrow r)$
5. $q \wedge \neg p \wedge r \wedge \neg p$
6. $\neg p$
7. $\neg(q \vee r)$
8. $\neg q \wedge \neg r$
9. $\neg q$
10. q
11. $q \wedge \neg q$
12. ∇_{\circ}

Como hemos dicho, esto demuestra que $(p \Leftarrow (q \vee r)) \Rightarrow ((p \Leftarrow q) \vee (p \Leftarrow r))$.
 $\Leftarrow (p \Leftarrow (q \vee r)) \Leftarrow ((p \Leftarrow q) \vee (p \Leftarrow r))$ es lo mismo que

$$((p \Leftarrow q) \vee (p \Leftarrow r)) \Rightarrow (p \Leftarrow (q \vee r)).$$

La proposición anterior equivale a

$$[((p \Leftarrow q) \vee (p \Leftarrow r)) \wedge \neg(p \Leftarrow (q \vee r))] \Rightarrow \nabla_{\circ},$$

en vista de la observación 1.

1. $[((p \Leftarrow q) \vee (p \Leftarrow r)) \wedge \neg(p \Leftarrow (q \vee r))]$
2. $\neg(p \Leftarrow (q \vee r))$
3. $(q \vee r) \wedge \neg p$
4. $((p \Leftarrow q) \vee (p \Leftarrow r))$
5. $(q \Rightarrow p) \vee (r \Rightarrow p)$
6. $q \vee r$
7. $p \vee p$
8. p
9. $\neg p$
10. $p \wedge \neg p$
11. ∇_{\circ}

Entonces hemos demostrado que “ \Leftarrow ” se distribuye sobre “ \vee ”. ■

Ejercicio 34 . Decidir si

$$((q \vee r) \Rightarrow p) \Leftrightarrow ((q \Rightarrow p) \vee (r \Rightarrow p)).$$

Ejercicio 35 . Escribir en cada paso de los argumentos de arriba, la regla válida de inferencia que lo produjo.

Ejercicio 36 . Mostrar que “ \Leftarrow ” se distribuye sobre “ \wedge ”. Es decir, que $p \Leftarrow (q \wedge r)$ es equivalente a $(p \Leftarrow q) \wedge (p \Leftarrow r)$.

Ejercicio 37 . Mostrar que “ \Leftarrow ” **no** se distribuye sobre “ \Leftarrow ”. Es decir, que $p \Leftarrow (q \Leftarrow r)$ **no** es equivalente a $(p \Leftarrow q) \Leftarrow (p \Leftarrow r)$.

Ejercicio 38 . Mostrar que “ \vee ” se distribuye sobre “ \Leftarrow ”

Ejercicio 39 . Mostrar que “ \wedge ” **no** se distribuye sobre “ \Leftarrow ”

En los siguientes ejercicios hacer deducciones directas.

Ejercicio 40

$$\frac{\begin{array}{c} A \Rightarrow (C \Rightarrow \neg A) \\ \neg(B \vee D) \vee A \\ C \end{array}}{\therefore \neg B \vee \neg D}$$

Ejercicio 41

$$\frac{\begin{array}{c} (E \Rightarrow F) \wedge (G \Rightarrow H) \\ (F \vee G) \Rightarrow (\neg E \wedge \neg H) \end{array}}{\therefore F \Rightarrow \neg G}$$

Ejercicio 42

$$\frac{\begin{array}{c} [(P \vee Q) \wedge (P \vee R)] \Leftarrow S \\ S \wedge \neg P \end{array}}{\therefore Q \wedge R}$$

En los siguientes ejercicios úsese reducción al absurdo.

Ejercicio 43

$$\frac{\begin{array}{c} M \Rightarrow (N \wedge L) \\ L \Rightarrow (N \Rightarrow M) \\ \neg N \end{array}}{\therefore \neg M}$$

Ejercicio 44

$$\frac{\begin{array}{c} P \implies (R \implies \nabla) \\ S \wedge \neg P \end{array}}{\therefore Q \wedge R}$$

Al hacer las siguientes deducciones use la hipótesis adicional adecuada.³

Ejercicio 45

$$\frac{\begin{array}{c} P \vee Q \vee R \implies S \\ S \wedge P \implies R \\ \neg Q \vee \neg R \end{array}}{\therefore P \implies \neg Q}$$

Ejercicio 46

$$\frac{\begin{array}{c} (A \implies (B \wedge C)) \vee (B \implies (C \vee D)) \\ C \implies D \end{array}}{\therefore \neg D \implies (\neg A \vee \neg B)}$$

Ejercicio 47

$$\frac{\begin{array}{c} \neg J \implies \neg K \vee \neg M \\ \neg(N \wedge L) \implies \neg J \end{array}}{\therefore (K \wedge L) \implies (M \implies N)}$$

Ejercicio 48

$$\frac{\begin{array}{c} E \vee \neg G \\ F \implies \neg E \\ \hline H \implies G \end{array}}{\therefore F \implies \neg H}$$

³Por la regla de exportación, $P \implies (Q \implies R)$ es equivalente a $(P \wedge Q) \implies R$. Entonces, para demostrar que $\frac{P}{\therefore Q \implies R}$ es válida, bastaría ver que $\frac{Q}{\therefore R}$ lo es. Así que uno podría agregar la "hipótesis adicional" Q (y llegar a R , en lugar de llegar a $Q \implies R$)

1.7 Apéndice. Sistemas formales

El concepto de “la verdad” es un problema difícil. La pregunta ¿qué es la verdad?, es un problema que se ha dejado a los filósofos y hay grandes disertaciones sobre el tema que no conciernen a los matemáticos.

En realidad, a los matemáticos, lo que más nos interesa es razonar correctamente, porque así podemos escoger el punto de partida de nuestras construcciones. A la manera de un juego, el matemático escoge los principios de los que quiere partir (Axiomas) y aplicando las reglas del juego (las Reglas válidas de inferencia) obtiene nuevas afirmaciones (teoremas). Algunos de los teoremas que obtienen los matemáticos se llaman de otra forma, por su importancia relativa dentro de la Teoría (el conjunto de Axiomas y Teoremas). Así, se pueden llamar Corolarios, Lemas, Proposiciones, Observaciones.

Si se procede de esta manera, los conceptos de verdad y falsedad quedan de lado. Lo que importa es razonar correctamente.

Es claro que, como en todo juego, uno debe escoger los axiomas adecuadamente. Si se escogen los axiomas sin cuidado, puede uno crear un juego insulso y sin gracia. Además, como muchas veces quiere uno que la Matemática corresponda a alguna parte del mundo “real”, uno tiene que ser cuidadoso para que haya una correspondencia entre la teoría y la “realidad”.

Muchas veces, los axiomas conducen a callejones sin salida, bien porque lo que uno obtiene no corresponde con lo que se esperaba, o bien porque al escoger los axiomas, se producen absurdos, tales como demostrar simultáneamente una afirmación y su negación. En este caso, en algún sentido, el juego se ha estropeado, pues desde este momento ya se puede demostrar cualquier afirmación. Es el momento de regresar al principio, y modificar los axiomas, hasta que el juego se desarrolle como es debido (mientras no nos encontremos nuevamente con algún resultado indeseable).

Una teoría se considera más elegante mientras menos axiomas tenga. A nadie se le ocurriría poner como axioma algo que se pueda deducir de los demás axiomas, pero algunas veces esto se hace con el fin de explicar rápidamente algún aspecto interesante de la Teoría.

A veces, con propósitos didácticos se parte de algún hecho que no ha sido demostrado y a partir de él se obtienen consecuencias. Uno debe recordar que las consecuencias deducidas dependen del resultado que se “tomó prestado”. Así, una teoría no tiene que ser forzosamente desarrollada “desde el principio hasta el final”. Hay personas que dejan huecos y después los llenan. En estos casos, se estarán tomando algunos hechos como si fueran axiomas. El

riesgo de proceder así, es que si por alguna razón, se parte de un resultado que después resulta incorrecto, las consecuencias obtenidas a partir de él no tienen ningún valor.

El método de introducir algunos objetos iniciales, algunos axiomas, algunas reglas de inferencia para deducir o construir nuevos objetos, se llama *método formal*, y al sistema de dichos objetos, axiomas y reglas de inferencia, se le llama *sistema formal*.

Enseguida describiremos el sistema formal del Cálculo proposicional.

Definición 7 . *Un sistema formal consta de:*

1. Un alfabeto de símbolos (letras).
2. Un conjunto de palabras formadas con un conjunto finito de letras.
3. Un conjunto de palabras bien formadas (pbf), llamadas axiomas.
4. Un conjunto finito de reglas de deducción que nos permiten construir nuevas palabras bien formadas (pbf) a partir de otras.

1.7.1 El sistema formal L

Definición 8 . *El sistema formal L del Cálculo proposicional está definido de la manera siguiente:*

1. Alfabeto: $\neg, \Rightarrow, (,), p_1, p_2, p_3, \dots$
2. Palabras bien formadas:
 - (a) Cada p_i es una palabra bien formada.
 - (b) Si \mathcal{A} y \mathcal{B} son palabras bien formadas entonces $\neg\mathcal{A}$ y $\mathcal{A} \Rightarrow \mathcal{B}$ son palabras bien formadas.
3. Axiomas:

Si \mathcal{A} , \mathcal{B} , \mathcal{C} son pbf, entonces son axiomas:

Axioma 1 . $\mathcal{A} \Rightarrow (\mathcal{B} \Rightarrow \mathcal{A})$

Axioma 2 . $[\mathcal{A} \Rightarrow (\mathcal{B} \Rightarrow \mathcal{C})] \Rightarrow [(\mathcal{A} \Rightarrow \mathcal{B}) \Rightarrow (\mathcal{A} \Rightarrow \mathcal{C})]$

Axioma 3 . $(\neg A \Rightarrow \neg B) \Rightarrow (\neg \neg A \Rightarrow \neg \neg B)$

4. Reglas de deducción: sólo hay una regla de deducción en L :

Modus ponens (MP):

$$\frac{\begin{array}{c} A \\ \hline A \Rightarrow B \end{array}}{\therefore B}$$

Definición 9 . Una demostración en L es una sucesión

$$A_1, A_2, \dots, A_n$$

tal que cada A_i es o bien un axioma de L , o bien se sigue de dos miembros previos mediante el uso de MP.

Dicha lista se llama una demostración de A_n en L y A_n se llama un teorema de L .

Observación 2 . Por definición, los axiomas de L son teoremas de L . (Su demostración consiste en la lista que consta únicamente del axioma).

Definición 10 . Sea Γ un conjunto de pbf de L . Una sucesión

$$A_1, \dots, A_n$$

es una deducción en L a partir de Γ si

1. cada A_i es un axioma de L ó
2. es un miembro de Γ ó
3. se deduce de miembros anteriores de la lista usando MP.

En estos casos escribiremos

$$\Gamma \vdash_L A_n.$$

En caso de que no se necesiten hipótesis adicionales para demostrar A , entonces A es un teorema de L y escribimos

$$\vdash_L A.$$

Ejemplo 31 . $\vdash_L \mathcal{A} \Rightarrow \mathcal{A}$.

1. $[\mathcal{A} \Rightarrow ((\mathcal{A} \Rightarrow \mathcal{A}) \Rightarrow \mathcal{A})] \Rightarrow [(\mathcal{A} \Rightarrow (\mathcal{A} \Rightarrow \mathcal{A})) \Rightarrow (\mathcal{A} \Rightarrow \mathcal{A})]$ (Axioma 2)
2. $\mathcal{A} \Rightarrow ((\mathcal{A} \Rightarrow \mathcal{A}) \Rightarrow \mathcal{A})$ (Axioma 1)
3. $(\mathcal{A} \Rightarrow (\mathcal{A} \Rightarrow \mathcal{A})) \Rightarrow (\mathcal{A} \Rightarrow \mathcal{A})$ (MP a las dos anteriores)
4. $\mathcal{A} \Rightarrow (\mathcal{A} \Rightarrow \mathcal{A})$ (Axioma 1)
5. $(\mathcal{A} \Rightarrow \mathcal{A})$ (MP a las dos anteriores).

1.8 El Teorema de la deducción y las hipótesis adicionales

Teorema 1 . $\Gamma \vdash_L (\mathcal{B} \Rightarrow \mathcal{A})$ si y solamente si $\Gamma \cup \{\mathcal{B}\} \vdash_L \mathcal{A}$.

Demostración. $\Rightarrow)$ Sea

$$\mathcal{A}_1, \dots, \mathcal{A}_k, \mathcal{B} \Rightarrow \mathcal{A}$$

una deducción de $\mathcal{B} \Rightarrow \mathcal{A}$ a partir de Γ . Entonces

$$\mathcal{A}_1, \dots, \mathcal{A}_k, \mathcal{B}, \mathcal{B} \Rightarrow \mathcal{A}$$

es una deducción a partir de $\Gamma \cup \{\mathcal{B}\}$. Podemos alargar esta lista aplicando MP a los dos últimos elementos de la lista anterior, de tal manera que

$$\mathcal{A}_1, \dots, \mathcal{A}_k, \mathcal{B}, \mathcal{B} \Rightarrow \mathcal{A}, \mathcal{A}$$

es una deducción de \mathcal{A} a partir de $\Gamma \cup \{\mathcal{B}\}$.

$\Leftarrow)$ Supongamos que la afirmación es falsa, así que supongamos que existe una deducción

$$\mathcal{A}_1, \dots, \mathcal{A}_k, \mathcal{B}, \mathcal{A}_{k+1}, \dots, \mathcal{A}_n, \mathcal{A} \quad (1.21)$$

de \mathcal{A} a partir de $\Gamma \cup \{\mathcal{B}\}$ (es decir, $\Gamma \cup \{\mathcal{B}\} \vdash_L \mathcal{A}$). Pero que $\Gamma \vdash_L (\mathcal{B} \Rightarrow \mathcal{A})$. Entre todos los ejemplos de la situación anterior podríamos escoger uno con el menor número de palabras (o de “comas”).

1. a) Si no hubiera palabras $\mathcal{A}_1, \dots, \mathcal{A}_k$ antes de \mathcal{B} , entonces \mathcal{B} sería un axioma, así que

$$\Gamma \vdash_L \mathcal{A}$$

por la definición de deducción. Podríamos alargar la deducción después de \mathcal{A} agregando el axioma $\mathcal{A} \Rightarrow (\mathcal{B} \Rightarrow \mathcal{A})$ y después $(\mathcal{B} \Rightarrow \mathcal{A})$ por MP. Entonces

$$\Gamma \vdash_L (\mathcal{B} \Rightarrow \mathcal{A})$$

contra lo supuesto.

Podemos suponer que \mathcal{B} no es un axioma y que $k \geq 2$.

b) Como

$$\mathcal{A}_1, \dots, \mathcal{A}_k, \mathcal{B}, \mathcal{A}_{k+1}, \dots, \mathcal{A}_n$$

es una deducción más corta que 1.21, por la elección de 1.21 podemos suponer que $\Gamma \vdash_L (\mathcal{B} \Rightarrow \mathcal{A}_n)$. Análogamente, podemos suponer que $\Gamma \vdash_L (\mathcal{B} \Rightarrow \mathcal{A}_{n-1})$, $\Gamma \vdash_L (\mathcal{B} \Rightarrow \mathcal{A}_{n-2})$, ..., $\Gamma \vdash_L (\mathcal{B} \Rightarrow \mathcal{A}_1)$.

\mathcal{A} no puede ser una axioma, pues si lo fuera, tendríamos la lista $\mathcal{A}, \mathcal{A} \Rightarrow (\mathcal{B} \Rightarrow \mathcal{A}), \mathcal{B} \Rightarrow \mathcal{A}$, de donde $\vdash_L \mathcal{B} \Rightarrow \mathcal{A}$, en particular, $\Gamma \vdash_L (\mathcal{B} \Rightarrow \mathcal{A})$, contra lo supuesto.

Como \mathcal{A} no es un axioma, debe ser deducible por MP a partir de miembros anteriores de la lista.

Digamos que de \mathcal{C} y de $\mathcal{C} \Rightarrow \mathcal{A}$, pero para ellos tenemos que $\Gamma \vdash_L (\mathcal{B} \Rightarrow \mathcal{C})$ y $\Gamma \vdash_L \mathcal{B} \Rightarrow (\mathcal{C} \Rightarrow \mathcal{A})$. Entonces tenemos:

$$\begin{aligned} \text{deducción en } \Gamma : & \left\{ \begin{array}{c} \vdots \\ \mathcal{B} \Rightarrow \mathcal{C} \end{array} \right. \\ \text{deducción en } \Gamma : & \left\{ \begin{array}{c} \vdots \\ \mathcal{B} \Rightarrow (\mathcal{C} \Rightarrow \mathcal{A}) \end{array} \right. \\ [\mathcal{B} \Rightarrow (\mathcal{C} \Rightarrow \mathcal{A})] \Rightarrow [(\mathcal{B} \Rightarrow \mathcal{C}) \Rightarrow (\mathcal{B} \Rightarrow \mathcal{A})] & \quad (\text{Axioma}) \\ (\mathcal{B} \Rightarrow \mathcal{C}) \Rightarrow (\mathcal{B} \Rightarrow \mathcal{A}) & \quad \text{MP} \\ \mathcal{B} \Rightarrow \mathcal{A} & \quad \text{MP} \end{aligned}$$

Que es una deducción de $\mathcal{B} \Rightarrow \mathcal{A}$ a partir de Γ . En contra de lo supuesto. ■

En vista del teorema anterior, si queremos demostrar que $\mathcal{B} \Rightarrow \mathcal{A}$ es un teorema en el sistema L , es decir, que

$$\vdash_L \mathcal{B} \Rightarrow \mathcal{A},$$

basta ver que de \mathcal{B} se puede deducir \mathcal{A} , o sea, basta ver que

$$\{\mathcal{B}\} \vdash_L \mathcal{A}.$$

Esto significa que podemos agregar \mathcal{B} como hipótesis adicional.

Como a veces se dice, “para demostrar $\mathcal{B} \Rightarrow \mathcal{A}$, podemos suponer \mathcal{B} , y deducir \mathcal{A} ”.

Ejemplo 32 . $\{(\mathcal{A} \Rightarrow \mathcal{B}), (\mathcal{B} \Rightarrow \mathcal{C})\} \vdash_L \mathcal{A} \Rightarrow \mathcal{C}$. (A esto también le llamaremos Silogismo hipotético)

Demostración

1. $\mathcal{A} \Rightarrow \mathcal{B}$ (Hipótesis adicional)
2. $\mathcal{B} \Rightarrow \mathcal{C}$ (Hipótesis adicional)
3. $[\mathcal{A} \Rightarrow (\mathcal{B} \Rightarrow \mathcal{C})] \Rightarrow [(\mathcal{A} \Rightarrow \mathcal{B}) \Rightarrow (\mathcal{A} \Rightarrow \mathcal{C})]$ (Axioma 2)
4. $(\mathcal{B} \Rightarrow \mathcal{C}) \Rightarrow [\mathcal{A} \Rightarrow (\mathcal{B} \Rightarrow \mathcal{C})]$ (Axioma 1)
5. $[\mathcal{A} \Rightarrow (\mathcal{B} \Rightarrow \mathcal{C})]$ (MP, 2, 4)
6. $(\mathcal{A} \Rightarrow \mathcal{B}) \Rightarrow (\mathcal{A} \Rightarrow \mathcal{C})$ (MP, 5, 3)
7. $(\mathcal{A} \Rightarrow \mathcal{C})$ (MP, 1, 6)

Por el teorema de la deducción tenemos:

$$\{(\mathcal{A} \Rightarrow \mathcal{B})\} \vdash_L (\mathcal{B} \Rightarrow \mathcal{C}) \Rightarrow (\mathcal{A} \Rightarrow \mathcal{C})$$

y

$$\vdash_L (\mathcal{A} \Rightarrow \mathcal{B}) \Rightarrow [(\mathcal{B} \Rightarrow \mathcal{C}) \Rightarrow (\mathcal{A} \Rightarrow \mathcal{C})].$$

Ejemplo 33 . $\{(\neg\neg\mathcal{A})\} \vdash_L \mathcal{A}$.

Demostración

1. $(\neg\neg\mathcal{A}) \Rightarrow (\neg\neg\mathcal{B} \Rightarrow \neg\neg\mathcal{A})$ (Axioma 1)
2. $\neg\neg\mathcal{A}$ (Hipótesis adicional)
3. $\neg\neg\mathcal{B} \Rightarrow \neg\neg\mathcal{A}$ (MP a las dos anteriores)
4. $(\neg\neg\mathcal{B} \Rightarrow \neg\neg\mathcal{A}) \Rightarrow (\neg\mathcal{A} \Rightarrow \neg\mathcal{B})$ (Axioma 3)

5. $(\neg A \Rightarrow \neg B) \Rightarrow (B \Rightarrow A)$ (Axioma 3)
6. $(\neg \neg B \Rightarrow \neg \neg A) \Rightarrow (B \Rightarrow A)$ (Por el Ejemplo 32, que es un teorema de L)⁴
7. $(B \Rightarrow A)$ (MP a 3 y 6) Esto pasa para cualquier B , en particular debe pasar si B es un teorema de L , como $A \Rightarrow A$, por ejemplo
8. $(A \Rightarrow A) \Rightarrow A$ (Por la nota anterior).
9. $A \Rightarrow A$ (Teorema, no hace falta incluir su demostración)
10. A (MP a 8 y 9) ■

Entonces tenemos el siguiente teorema de L :

$$\vdash_L \neg \neg A \Rightarrow A.$$

Ejemplo 34 . $\{A\} \vdash_L \neg \neg A$.

Demostración

1. $\neg \neg \neg A \Rightarrow \neg A$ (Por el ejemplo anterior)
2. $(\neg \neg \neg A \Rightarrow \neg A) \Rightarrow (A \Rightarrow \neg \neg A)$ (Axioma 3)
3. $A \Rightarrow \neg \neg A$ (MP a 1, 2)
4. A (Hipótesis adicional)
5. $\neg \neg A$ (MP a 3, 4)

En estos momentos podemos introducir las siguientes abreviaturas:

1. $A \vee B : \neg A \Rightarrow B$.
2. $A \wedge B : \neg (A \Rightarrow \neg B)$.
3. $A \leftrightarrow B : (A \Rightarrow B) \wedge (B \Rightarrow A)$.

⁴Como ya vimos la demostración de este Teorema, no tiene caso que la repitamos aquí

En sentido estricto “ \vee ”, “ \wedge ”, “ \leftrightarrow ” no son parte del sistema formal L , pero su uso nos permitirá hacer más cortas las demostraciones.

Veamos que dentro del sistema formal L podemos establecer el resto de las reglas de inferencia que ya conocíamos.

Hasta este momento disponemos de MP, de Silogismo hipotético (ejemplo 32), y de Doble negación.

Ejemplo 35 . *Tollendo tollens: $\{\mathcal{A} \Rightarrow \mathcal{B}, \neg \mathcal{B}\} \vdash_L \neg \mathcal{A}$.*

1. $\mathcal{A} \Rightarrow \mathcal{B}$ (Hipótesis adicional).
2. $\neg \neg \mathcal{A} \Rightarrow \mathcal{A}$ (Teorema).
3. $\neg \neg \mathcal{A} \Rightarrow \mathcal{B}$ (Silogismo hipotético).
4. $\mathcal{B} \Rightarrow \neg \neg \mathcal{B}$ (Teorema).
5. $\neg \neg \mathcal{A} \Rightarrow \neg \neg \mathcal{B}$ (Silogismo hipotético).
6. $(\neg \neg \mathcal{A} \Rightarrow \neg \neg \mathcal{B}) \Rightarrow (\neg \mathcal{B} \Rightarrow \neg \mathcal{A})$ (Axioma).
7. $\neg \mathcal{B} \Rightarrow \neg \mathcal{A}$ (MP).
8. $\neg \mathcal{B}$ (Hipótesis adicional).
9. $\neg \mathcal{A}$ (MP).

Ejemplo 36 . *Tollendo ponens: $\{\mathcal{A} \vee \mathcal{B}, \neg \mathcal{B}\} \vdash_L \mathcal{A}$*

1. $\mathcal{A} \vee \mathcal{B} := \neg \mathcal{A} \Rightarrow \mathcal{B}$ (Hipótesis adicional, definición).
2. $\neg \mathcal{B}$ (Hipótesis adicional).
3. $\neg \neg \mathcal{A}$ (TT, ejemplo anterior).
4. $\neg \neg \mathcal{A} \Rightarrow \mathcal{A}$ (Teorema).
5. \mathcal{A} (MP).

Ejemplo 37 . *Commutatividad de \vee : $\{\mathcal{A} \vee \mathcal{B}\} \vdash_L \mathcal{B} \vee \mathcal{A}$*

1. $\mathcal{A} \vee \mathcal{B} := \neg \mathcal{A} \Rightarrow \mathcal{B}$ (Hipótesis adicional, definición).

2. $\mathcal{B} \Rightarrow \neg\neg\mathcal{B}$ (Teorema).
3. $\neg\mathcal{A} \Rightarrow \neg\neg\mathcal{B}$ (SH).
4. $(\neg\mathcal{A} \Rightarrow \neg\neg\mathcal{B}) \Rightarrow (\neg\mathcal{B} \Rightarrow \mathcal{A})$ (Axioma).
5. $(\neg\mathcal{B} \Rightarrow \mathcal{A}) := \mathcal{B} \vee \mathcal{A}$ (MP, definición).

Dejamos como ejercicio: $(\mathcal{A} \wedge \mathcal{B}) \Rightarrow \mathcal{A}$ (Simplificación). Vea el Ejercicio 14).

Ejemplo 38 . $\vdash_L \mathcal{A} \wedge \mathcal{B} \Rightarrow \mathcal{B}$.

1. $\mathcal{A} \wedge \mathcal{B} := \neg(\mathcal{A} \Rightarrow \neg\mathcal{B})$ (Hipótesis adicional).
2. $\neg\mathcal{B} \Rightarrow (\mathcal{A} \Rightarrow \neg\mathcal{B})$ (Axioma).
3. $\neg\neg\mathcal{B}$ (TT).
4. $\neg\neg\mathcal{B} \Rightarrow \mathcal{B}$ (Teorema).
5. \mathcal{B} (MP).

Ejemplo 39 . $\{\mathcal{B} \Rightarrow \neg\mathcal{A}\} \vdash_L \mathcal{A} \Rightarrow \neg\mathcal{B}$.

1. $\mathcal{B} \Rightarrow \neg\mathcal{A}$ (Hipótesis adicional).
2. $\neg\neg\mathcal{B} \Rightarrow \mathcal{B}$ (Teorema).
3. $\neg\neg\mathcal{B} \Rightarrow \neg\mathcal{A}$ (SH, 2,1).
4. $(\neg\neg\mathcal{B} \Rightarrow \neg\mathcal{A}) \Rightarrow (\mathcal{A} \Rightarrow \neg\mathcal{B})$ (Axioma).
5. $(\mathcal{A} \Rightarrow \neg\mathcal{B})$ (MP).

Ejemplo 40 . *Commutatividad de \wedge :* $\{\mathcal{A} \wedge \mathcal{B}\} \vdash_L \mathcal{B} \wedge \mathcal{A}$

1. $\mathcal{A} \wedge \mathcal{B} := \neg(\mathcal{A} \Rightarrow \neg\mathcal{B})$ (Hipótesis adicional, definición).
2. $(\mathcal{B} \Rightarrow \neg\mathcal{A}) \Rightarrow (\mathcal{A} \Rightarrow \neg\mathcal{B})$ (Ejemplo anterior).
3. $(\mathcal{A} \Rightarrow \neg\mathcal{B}) \Rightarrow \neg\neg(\mathcal{A} \Rightarrow \neg\mathcal{B})$ (Teorema).

4. $\neg\neg(\mathcal{B} \Rightarrow \neg\mathcal{A}) \Rightarrow (\mathcal{B} \Rightarrow \neg\mathcal{A})$ (Teorema).
5. $\neg\neg(\mathcal{B} \Rightarrow \neg\mathcal{A}) \Rightarrow \neg\neg(\mathcal{A} \Rightarrow \neg\mathcal{B})$ (SH, 3,4 y 2).
6. $[\neg\neg(\mathcal{B} \Rightarrow \neg\mathcal{A}) \Rightarrow \neg\neg(\mathcal{A} \Rightarrow \neg\mathcal{B})] \Rightarrow [\neg(\mathcal{A} \Rightarrow \neg\mathcal{B}) \Rightarrow (\neg(\mathcal{B} \Rightarrow \neg\mathcal{A}))]$ (Axioma).
7. $\neg(\mathcal{A} \Rightarrow \neg\mathcal{B}) \Rightarrow (\neg(\mathcal{B} \Rightarrow \neg\mathcal{A}))$ (MP, 5, 6).
8. $(\neg(\mathcal{B} \Rightarrow \neg\mathcal{A})) := \mathcal{B} \curlywedge \mathcal{A}$ (MP, definición).

Ejemplo 41 . Adición: $\{\mathcal{A}\} \vdash_L \mathcal{A} \vee \mathcal{B}$.

1. $\mathcal{A} \Rightarrow (\neg\mathcal{B} \Rightarrow \mathcal{A})$ (Axioma).
2. \mathcal{A} (Hipótesis adicional).
3. $(\neg\mathcal{B} \Rightarrow \mathcal{A})$ (MP a 1,2).
4. $\mathcal{A} \vee \mathcal{B}$ (De la anterior, por definición de “ \vee ”).

Ejercicio 49 . Muestre que $\{\mathcal{A} \vee \mathcal{B}\} \vdash_L \mathcal{A} \vee \neg\neg\mathcal{B}$

Ejemplo 42 . $(\neg\mathcal{B} \Rightarrow \mathcal{A}) \vdash_L (\neg\neg\neg\mathcal{B} \Rightarrow \neg\neg\mathcal{A})$

1. $\neg\mathcal{B} \Rightarrow \mathcal{A}$ (Hipótesis adicional).
2. $\mathcal{A} \Rightarrow \neg\neg\mathcal{A}$ (DN).
3. $\neg\mathcal{B} \Rightarrow \neg\neg\mathcal{A}$ (SH).
4. $\neg\neg\neg\mathcal{B} \Rightarrow \neg\mathcal{B}$ (DN).
5. $\neg\neg\neg\mathcal{B} \Rightarrow \neg\neg\mathcal{A}$ (SH).

Ejemplo 43 . De Morgan.

1. (a) $\{\neg(\mathcal{A} \curlywedge \mathcal{B})\} \vdash_L \neg\mathcal{A} \vee \neg\mathcal{B}$
 - i. $\neg(\mathcal{A} \curlywedge \mathcal{B})$ (Hipótesis adicional)
 - ii. $\neg(\neg(\mathcal{A} \Rightarrow \neg\mathcal{B}))$ (Por definición de “ \curlywedge ”)
 - iii. $\mathcal{A} \Rightarrow \neg\mathcal{B}$ (Por doble negación)
 - iv. $\neg\neg\mathcal{A} \Rightarrow \mathcal{A}$ (Por doble negación)
 - v. $\neg\neg\mathcal{A} \Rightarrow \neg\mathcal{B}$ (SH)

- vi. $(\neg\neg A \Rightarrow \neg B) \Rightarrow (B \Rightarrow \neg A)$ (Axioma)
- vii. $B \Rightarrow \neg A$ (MP, a las dos anteriores)
- viii. $\neg\neg B \Rightarrow B$ (Por doble negación)
- ix. $\neg\neg B \Rightarrow \neg A$ (SH)
- x. $\neg B \vee \neg A$ (De la anterior, por definición de “ \vee ”).
- xi. $\neg A \vee \neg B$ (Por la conmutatividad de “ \vee ”).

(b) $\{\neg A \vee \neg B\} \vdash_L \neg(A \wedge B)$

- i. $\neg A \vee \neg B$ (Hipótesis adicional)
- ii. $\neg\neg A \Rightarrow \neg B$ (Por definición de “ \vee ”)
- iii. $A \Rightarrow \neg\neg A$ (Por DN)
- iv. $A \Rightarrow \neg B$ (Por SH)
- v. $\neg(\neg(A \Rightarrow \neg B))$ (Por DN)
- vi. $\neg(A \wedge B)$ (Por definición de “ \wedge ”)

c. $\{\neg(A \vee B)\} \vdash_L \neg A \wedge \neg B$

- i. $\neg(A \vee B)$ (Hipótesis adicional)
- ii. $\neg(\neg A \Rightarrow B)$ (Por definición de “ \vee ”)
- iii. $(\neg B \Rightarrow \neg\neg A) \Rightarrow (\neg A \Rightarrow B)$ (Axioma)
- iv. $\neg(\neg B \Rightarrow \neg\neg A)$ (Tollendo tollens a las dos anteriores)
- v. $\neg B \wedge \neg A$ (Por definición de “ \wedge ”)

d. $\{\neg A \wedge \neg B\} \vdash_L \neg(A \vee B)$

- i. $\neg A \wedge \neg B$ (Hipótesis adicional)
- ii. $\neg(\neg A \Rightarrow \neg\neg B)$ (Por definición de “ \wedge ”)
- iii. $(\neg\neg\neg B \Rightarrow \neg\neg A) \Rightarrow (\neg A \Rightarrow \neg\neg B)$ (Axioma)
- iv. $\neg(\neg\neg\neg B \Rightarrow \neg\neg A)$ (TT)
- v. $(\neg B \Rightarrow A) \Rightarrow (\neg\neg\neg B \Rightarrow \neg\neg A)$ (Ejemplo anterior)
- vi. $\neg(\neg B \Rightarrow A)$ (TT)
- vii. $\neg(B \vee A)$ (Definición de “ \vee ”)
- viii. $(A \vee B) \Rightarrow (B \vee A)$ (Commutatividad de \vee)
- ix. $\neg(A \vee B)$

■

Ejemplo 44 . $\{A \Rightarrow B\} \vdash_L (\neg B \Rightarrow \neg A)$

1. $\neg\neg A \Rightarrow A$ (DN)
2. $A \Rightarrow B$ (Hipótesis adicional)
3. $\neg\neg A \Rightarrow B$ (SH)
4. $B \Rightarrow \neg\neg B$ (DN)
5. $\neg\neg A \Rightarrow \neg\neg B$ (SH)
6. $(\neg\neg A \Rightarrow \neg\neg B) \Rightarrow (\neg B \Rightarrow \neg A)$ (Axioma)
7. $\neg B \Rightarrow \neg A$ (MP)

Ejemplo 45 . Asociatividad de “ \vee ”. $\{A \vee (B \vee C)\} \vdash_L (A \vee B) \vee C$

Queremos demostrar $\{A \vee (B \vee C)\} \vdash_L \neg(A \vee B) \Rightarrow C$. Por el teorema de la deducción, da lo mismo demostrar $\{A \vee (B \vee C), \neg(A \vee B)\} \vdash_L C$

1. $A \vee (B \vee C)$ (Hipótesis adicional)
2. $\neg A \Rightarrow (B \vee C)$ (Definición de “ \vee ”)
3. $\neg(A \vee B)$ (Hipótesis adicional)
4. $\neg A \wedge \neg B$ (De Morgan)
5. $\neg A$ (Simplificación)
6. $B \vee C$ (MP)
7. $\neg B$ (Simplificación)
8. C (TP)

Ejercicio 50 . Deducir la asociatividad de “ \wedge ” $\{A \wedge (B \wedge C)\} \vdash \{ (A \wedge B) \wedge C \}$.

Ejercicio 51 . Mostrar que $\vdash_L B \Rightarrow ((A \Rightarrow \neg B) \Rightarrow \neg A)$.

Sugerencia: Use el teorema de la deducción y muestre que

$$\{B, (A \Rightarrow \neg B)\} \vdash_L \neg A.$$

Ejemplo 46 . *Conjunción: $\{\mathcal{A}, \mathcal{B}\} \vdash_L \mathcal{A} \wedge \mathcal{B}$. Es decir, $\{\mathcal{A}, \mathcal{B}\} \vdash_L \neg(\mathcal{A} \Rightarrow \neg \mathcal{B})$.*

1. \mathcal{B} (Hipótesis adicional)
2. $\mathcal{B} \Rightarrow ((\mathcal{A} \Rightarrow \neg \mathcal{B}) \Rightarrow \neg \mathcal{A})$ (Teorema, por el ejercicio anterior)
3. $(\mathcal{A} \Rightarrow \neg \mathcal{B}) \Rightarrow \neg \mathcal{A}$ (MP)
4. $\mathcal{A} \Rightarrow \neg \neg \mathcal{A}$ (DN)
5. \mathcal{A} (Hipótesis adicional)
6. $\neg \neg \mathcal{A}$ (MP)
7. $\neg(\mathcal{A} \Rightarrow \neg \mathcal{B})$

Ejemplo 47 . “ \wedge ” se distribuye sobre “ \vee ”. $\{\mathcal{A} \wedge (\mathcal{B} \vee \mathcal{C})\} \vdash_L (\mathcal{A} \wedge \mathcal{B}) \vee (\mathcal{A} \wedge \mathcal{C})$.

Es decir, $\{\mathcal{A} \wedge (\mathcal{B} \vee \mathcal{C})\} \vdash_L \neg(\mathcal{A} \wedge \mathcal{B}) \Rightarrow (\mathcal{A} \wedge \mathcal{C})$. Lo que equivale a

$$\{\mathcal{A} \wedge (\mathcal{B} \vee \mathcal{C}), \neg(\mathcal{A} \wedge \mathcal{B})\} \vdash_L (\mathcal{A} \wedge \mathcal{C}).$$

O sea,

$$\{\mathcal{A} \wedge (\mathcal{B} \vee \mathcal{C}), \neg(\mathcal{A} \wedge \mathcal{B})\} \vdash_L \neg(\mathcal{A} \Rightarrow \neg \mathcal{C}).$$

1. $\mathcal{A} \wedge (\mathcal{B} \vee \mathcal{C})$
2. $\neg(\mathcal{A} \Rightarrow \neg(\mathcal{B} \vee \mathcal{C}))$
3. $\neg(\mathcal{A} \wedge \mathcal{B})$
4. $\neg(\mathcal{A} \Rightarrow \neg(\neg \mathcal{B} \Rightarrow \mathcal{C}))$
5. $\neg(\neg \mathcal{A} \vee \neg(\neg \mathcal{B} \Rightarrow \mathcal{C}))$
6. $\neg \neg \mathcal{A} \wedge \neg \neg(\neg \mathcal{B} \Rightarrow \mathcal{C})$
7. $\neg \neg \mathcal{A}$
8. \mathcal{A}
9. $\neg(\neg(\mathcal{A} \Rightarrow \neg \mathcal{B}))$

10. $\mathcal{A} \Rightarrow \neg \mathcal{B}$
11. $\neg \mathcal{B}$
12. $\neg \neg (\neg \mathcal{B} \Rightarrow \mathcal{C})$
13. $\neg \mathcal{B} \Rightarrow \mathcal{C}$
14. \mathcal{C}
15. $\mathcal{A} \wedge \mathcal{C}$
16. $\neg (\mathcal{A} \Rightarrow \neg \mathcal{C})$

Ejercicio 52 . “ \vee ” se distribuye sobre “ \wedge ”. $\{\mathcal{A} \vee (\mathcal{B} \wedge \mathcal{C})\} \vdash_L (\mathcal{A} \vee \mathcal{B}) \wedge (\mathcal{A} \vee \mathcal{C})$.

Ejemplo 48 . Dilema constructivo: $\{(\mathcal{A} \Rightarrow \mathcal{B}) \vee (\mathcal{C} \Rightarrow \mathcal{D}), \mathcal{A}, \mathcal{C}, \neg \mathcal{B}\} \vdash_L \mathcal{B} \vee \mathcal{D}$

Reformulemos con el teorema de la deducción. Como $\mathcal{B} \vee \mathcal{D}$ es por definición $\neg \mathcal{B} \Rightarrow \mathcal{D}$, entonces queremos demostrar que

$$\{(\mathcal{A} \Rightarrow \mathcal{B}) \vee (\mathcal{C} \Rightarrow \mathcal{D}), \mathcal{A}, \mathcal{C}, \neg \mathcal{B}\} \vdash_L \mathcal{D}$$

1. $(\mathcal{A} \Rightarrow \mathcal{B}) \vee (\mathcal{C} \Rightarrow \mathcal{D})$
2. \mathcal{A}
3. $\neg \mathcal{B}$
4. $\mathcal{A} \wedge \neg \mathcal{B} := \neg(\mathcal{A} \Rightarrow \neg \mathcal{B})$
5. $(\mathcal{A} \Rightarrow \mathcal{B}) \Rightarrow (\mathcal{A} \Rightarrow \neg \mathcal{B})$ (¡ Compruébelo!)
6. $\neg(\mathcal{A} \Rightarrow \mathcal{B})$ (TT)
7. $\mathcal{C} \Rightarrow \mathcal{D}$ (TP)
8. \mathcal{C}
9. \mathcal{D} .

Ejercicio 53 . Dilema destructivo: $\{(\mathcal{A} \Rightarrow \mathcal{B}) \vee (\mathcal{C} \Rightarrow \mathcal{D}), \neg \mathcal{B}, \neg \mathcal{D}\} \vdash_L \neg \mathcal{A} \vee \neg \mathcal{C}$.

Ejemplo 49 . $\{\neg(\mathcal{A} \Rightarrow \mathcal{B})\} \vdash_L \mathcal{A} \wedge \neg\mathcal{B}$.

1. $\neg(\mathcal{A} \Rightarrow \mathcal{B})$
2. $(\neg\mathcal{B} \Rightarrow \neg\mathcal{A}) \Rightarrow (\mathcal{A} \Rightarrow \mathcal{B})$ (Axioma)
3. $\neg(\neg\mathcal{B} \Rightarrow \neg\mathcal{A})$
4. $\neg(\neg\mathcal{B} \Rightarrow \neg\mathcal{A}) := \neg\mathcal{B} \wedge \mathcal{A}$
5. $\mathcal{A} \wedge \neg\mathcal{B}$ (Commutatividad).

Ejercicio 54 . $\{\mathcal{A} \wedge \neg\mathcal{B}\} \vdash_L \neg(\mathcal{A} \Rightarrow \mathcal{B})$

Ejemplo 50 . $\{\mathcal{A} \Rightarrow (\mathcal{B} \Rightarrow \mathcal{C})\} \vdash_L (\mathcal{A} \wedge \mathcal{B}) \Rightarrow \mathcal{C}$.

Lo anterior equivale a $\{\mathcal{A} \Rightarrow (\mathcal{B} \Rightarrow \mathcal{C}), (\mathcal{A} \wedge \mathcal{B})\} \vdash_L \mathcal{C}$.

1. $\mathcal{A} \Rightarrow (\mathcal{B} \Rightarrow \mathcal{C})$
2. $\mathcal{A} \wedge \mathcal{B}$
3. \mathcal{A} (Simplificación)
4. $\mathcal{B} \Rightarrow \mathcal{C}$ (MP)
5. \mathcal{B} (Simplificación)
6. \mathcal{C} (MP).

Ejercicio 55 . $\{(\mathcal{A} \wedge \mathcal{B}) \Rightarrow \mathcal{C}\} \vdash_L \mathcal{A} \Rightarrow (\mathcal{B} \Rightarrow \mathcal{C})$.

Ejemplo 51 . *Reducción al absurdo.:*

$$\vdash_L [(\mathcal{A} \wedge \neg\mathcal{B}) \Rightarrow (\mathcal{C} \wedge \neg\mathcal{C})] \Rightarrow (\mathcal{A} \Rightarrow \mathcal{B})$$

Esto, que es una parte de lo que debe entenderse por reducción al absurdo, es equivalente a demostrar:

$$\{[(\mathcal{A} \wedge \neg\mathcal{B}) \Rightarrow (\mathcal{C} \wedge \neg\mathcal{C})], \mathcal{A}\} \vdash_L \mathcal{B} :$$

1. $\mathcal{C} \Rightarrow \mathcal{C}$ (Teorema)

2. $\mathcal{C} \Rightarrow \neg\neg\mathcal{C}$ (DN)
3. $\neg\neg(\mathcal{C} \Rightarrow \neg\neg\mathcal{C})$ (DN)
4. $\neg(\mathcal{C} \wedge \neg\mathcal{C}) := \neg\neg(\mathcal{C} \Rightarrow \neg\neg\mathcal{C})$ (Por definición de “ \wedge ”)
5. $(\mathcal{A} \wedge \neg\mathcal{B}) \Rightarrow (\mathcal{C} \wedge \neg\mathcal{C})$ (Hipótesis adicional)
6. $\neg(\mathcal{A} \wedge \neg\mathcal{B})$ (TT)
7. $\mathcal{A} \Rightarrow \mathcal{B}$ (Equivalente al anterior, por definición de “ \wedge ”)
8. \mathcal{A} (Hipótesis adicional)
9. \mathcal{B} (MP).

Ejercicio 56 . Demuestre la otra parte de la reducción al absurdo:

$$\vdash_L (\mathcal{A} \Rightarrow \mathcal{B}) \Rightarrow [(\mathcal{A} \wedge \neg\mathcal{B}) \Rightarrow (\mathcal{C} \wedge \neg\mathcal{C})].$$

Ejercicio 57 . Vea si hay más reglas válidas de inferencia, y muestre que son teoremas en L , si se interpretan adecuadamente.

1.9 Valuación

Vamos a colgar etiquetas a las palabras bien formadas de L , de la manera siguiente:

1. Las etiquetas son dos: 0, 1.
2. A los axiomas les colgamos la etiqueta 1.
3. Si \mathcal{A} tiene etiqueta 1, entonces $\neg\mathcal{A}$ tendrá etiqueta 0. Si \mathcal{A} tiene etiqueta 0, entonces $\neg\mathcal{A}$ tendrá etiqueta 1.
4. $\mathcal{A} \Rightarrow \mathcal{B}$ tiene etiqueta 1, si \mathcal{B} también la tiene o si \mathcal{A} tiene etiqueta 0.
5. $\mathcal{A} \Rightarrow \mathcal{B}$ tiene etiqueta 0, si \mathcal{B} tiene etiqueta 0, y \mathcal{A} tiene etiqueta 1.

Como la frase “ \mathcal{A} tiene etiqueta r ”, es un poco larga abreviemos esto escribiendo $v(\mathcal{A}) = r$.

Con las reglas anteriores se puede calcular cual es la etiqueta de cualquier palabra bien formada.

Por ejemplo $v(\mathcal{A} \wedge \mathcal{B}) = v(\neg(\mathcal{A} \Rightarrow \neg\mathcal{B})) = \begin{cases} 1 & \text{si } v(\mathcal{A} \Rightarrow \neg\mathcal{B}) = 0 \\ 0 & \text{si } v(\mathcal{A} \Rightarrow \neg\mathcal{B}) = 1 \end{cases}$

Ahora $v(\mathcal{A} \Rightarrow \neg\mathcal{B}) = 0$ si $v(\mathcal{A}) = 1$ y $v(\neg\mathcal{B}) = 0$,

es decir si $v(\mathcal{A}) = 1$ y $v(\mathcal{B}) = 1$.

Ahora $v(\mathcal{A} \Rightarrow \neg\mathcal{B}) = 1$ si $v(\mathcal{A}) = 0$ ó $v(\neg\mathcal{B}) = 1$,

es decir si $v(\mathcal{A}) = 1$ ó $v(\mathcal{B}) = 0$.

Si hacemos una tabla tendremos:

\mathcal{A}	\wedge	\mathcal{B}
0	0	0
0	0	1
1	0	0
1	1	1

¿Les suena conocido?

Ejercicio 58 . Haga una tabla semejante a la anterior para \vee , recuerde que por definición

$$\mathcal{A} \vee \mathcal{B} := \neg\mathcal{A} \Rightarrow \mathcal{B}.$$

Observación 3 . Si $\mathcal{A} \Rightarrow \mathcal{B}$, tiene etiqueta 1, y \mathcal{A} también la tiene, entonces \mathcal{B} tiene etiqueta 1

Esto es una consecuencia de la definición de las etiquetas, si \mathcal{B} tuviera etiqueta 0, entonces por definición, $\mathcal{A} \Rightarrow \mathcal{B}$ hubiera tenido etiqueta 0.

Extendiendo el argumento de arriba, se puede ver que cualquier teorema en L tiene etiqueta 1 (recuérdese que modus ponens es la regla de inferencia de L (las demás son abreviaturas)) y recuerde que los axiomas se etiquetaron con 1. Esto suena bien: “Todos los teoremas tienen 1”, suena como “Todos los teoremas son V” o a “Todos los teoremas son V(erdaderos)”.

Ahora, ¿será cierto que si una palabra bien formada tiene etiqueta 1, entonces es un teorema de L ?

Teorema 2 . Si una palabra bien formada tiene etiqueta 1, entonces es un teorema de L .

Demostración.

Si hubiera una \mathcal{A} con etiqueta 1 que no fuera teorema, la podríamos escoger con el menor número posible de conectivos “ \neg ” y “ \Rightarrow ”. Entonces

\mathcal{A} no puede ser un axioma porque los axiomas son teoremas. ¿Por qué tiene etiqueta 1?

1. Porque $\mathcal{A} = \neg\mathcal{B}$, donde \mathcal{B} tiene etiqueta 0. Pero ¿por qué \mathcal{B} tiene etiqueta 0?

- (a) Porque \mathcal{B} es $\neg\mathcal{C}$, donde \mathcal{C} tiene etiqueta 1. Pero entonces \mathcal{C} es una palabra con etiqueta 1 con menos conectivos que \mathcal{A} . Por la elección de \mathcal{A} , \mathcal{B} tiene que ser un teorema. Pero si \mathcal{B} es un teorema, entonces $\neg\neg\mathcal{B}$ también lo es ($\vdash_L \mathcal{B} \Rightarrow \neg\neg\mathcal{B}$). Este caso queda descartado.
- (b) Porque \mathcal{B} es $(\mathcal{C} \Rightarrow \mathcal{D})$, con \mathcal{C} etiquetada con 1 y \mathcal{D} con 0. Entonces \mathcal{A} es $\neg(\mathcal{C} \Rightarrow \mathcal{D})$. Como $\neg\mathcal{D}$ también tiene menos conectivos que \mathcal{A} , lo mismo que arriba, concluimos que \mathcal{C} y $\neg\mathcal{D}$ tienen que ser teoremas. Pero entonces también lo es $\mathcal{C} \wedge \neg\mathcal{D}$ (conjunción) y por lo tanto también $\neg(\mathcal{C} \Rightarrow \mathcal{D})$. Este caso queda descartado.

2. Debe ser entonces que \mathcal{A} es $\mathcal{B} \Rightarrow \mathcal{C}$, \mathcal{B} etiquetada con 0 ó \mathcal{C} con 1.

- (a) Si \mathcal{C} está etiquetada con 1, como tiene menos conectivos que los que figuran en \mathcal{A} entonces \mathcal{C} debe ser un teorema de L . $\mathcal{C} \Rightarrow (\mathcal{B} \Rightarrow \mathcal{C})$ es un axioma y \mathcal{C} es un teorema, aplicando MP vemos que $\mathcal{A} = \mathcal{B} \Rightarrow \mathcal{C}$ también es un teorema de L . Este caso se descarta.
- (b) Por el caso anterior, podemos suponer que las etiquetas de \mathcal{B} y de \mathcal{C} son 0. Ahora pasa uno de los siguientes casos:
 - i. \mathcal{B} es $\neg\mathcal{D}$, con \mathcal{D} un teorema. Como $\mathcal{D} \Rightarrow (\neg\mathcal{C} \Rightarrow \mathcal{D})$ es un axioma, vemos, aplicando MP, que $\neg\mathcal{C} \Rightarrow \mathcal{D}$ es un teorema, pero entonces también lo son $\neg\mathcal{D} \Rightarrow \neg\neg\mathcal{C}$ y $\neg\mathcal{D} \Rightarrow \mathcal{C}$ que es \mathcal{A} . Así que esto no pasa.
 - ii. \mathcal{B} es $\mathcal{G} \Rightarrow \mathcal{H}$, \mathcal{G} con etiqueta 1 y \mathcal{H} con etiqueta 0. Así, \mathcal{A} es $(\mathcal{G} \Rightarrow \mathcal{H}) \Rightarrow \mathcal{C}$, donde \mathcal{G} , $\neg\mathcal{H}$ y $\neg\mathcal{C}$ son teoremas, porque tienen etiqueta 1 y tienen menos conectivos que \mathcal{A} . Pero entonces también lo son $\mathcal{G} \wedge \neg\mathcal{H}$ por lo tanto también lo es $\neg(\mathcal{G} \Rightarrow \mathcal{H})$ de esto y del axioma

$$\neg(\mathcal{G} \Rightarrow \mathcal{H}) \Rightarrow (\neg\mathcal{C} \Rightarrow (\neg(\mathcal{G} \Rightarrow \mathcal{H}))),$$

tenemos que $\neg C \Rightarrow (\neg (G \Rightarrow H))$ es un teorema, por lo tanto $(G \Rightarrow H) \Rightarrow C$ es un teorema.

No hay lugar para una palabra bien formada con etiqueta 1, que no sea un teorema. ■

En L una palabra bien formada es un teorema si y sólo si tiene etiqueta 1. En vista de que las tablas para la etiquetación de \wedge , \vee , coinciden con las definiciones de “ \wedge ” y “ \vee ” que se dieron en las secciones anteriores y como las reglas válidas de inferencia tiene su contraparte en el Cálculo proposicional L , ya podemos confiar en que si al calcular la tabla para una proposición está resulta una tautología, es porque esta proposición corresponde a un teorema.

Hay muchos temas importantísimos de la Lógica, que aquí no hemos ni siquiera esbozado.

Nuestro interés al respecto es proporcionar al lector una herramienta más para el posterior desarrollo de la Teoría. Nuevamente, sugerimos al lector se dirija a las citas bibliográficas para un mayor estudio.

1.10 Cuantificadores

Como no queremos extendernos demasiado en un Tema que no es propiamente nuestro asunto, daremos una versión muy elemental acerca de los cuantificadores.

Consideremos una proposición $p()$ que una cierta clase de objetos pueda o no tener. Por ejemplo, si n es un número natural, (ver el capítulo 3), $p(n)$ podría significar:

“ $p(n)$ es un número par”

Así que $p()$ representa más que una proposición, una sucesión de proposiciones: $p(0), p(1), p(2), \dots$ o bien:

0 es par, 1 es par, 2 es par,...

La p anterior es un ejemplo de lo que se llama una “función proposicional”.

Ahora, para decir que todos los objetos n en una cierta clase \mathcal{M} , tienen la propiedad p , escribiremos,

$\forall n \in \mathcal{M}, p(n)$
 para toda que sea un elemento de n tiene la propiedad p .

“ \forall ” se llama “cuantificador universal”, y en el sentido en que se usó arriba, podemos considerarla como una abreviatura de “para toda”.

Consideraremos que la anterior es una proposición, que es verdadera en el caso de que todas las n de \mathcal{M} tengan la propiedad p , y falsa cuando

existe un miembro n de \mathcal{M} que no tiene la propiedad p (1.22)

Ahora, es natural que si $p(n)$ significa “ n tiene la propiedad p ”, entonces la negación de esto se debe escribir $\neg p(n)$.

Introducimos el símbolo “ \exists ”, llamado el “cuantificador existencial”, que se puede considerar como una abreviatura de “existe”.

Así, podemos reescribir 1.22, de la manera siguiente:

$\exists n \in \mathcal{M}, \neg p(n)$
 existe elemento de tal que n no tiene la propiedad p . (1.23)

1. $\neg [\forall n \in \mathcal{M}, p(n)]$ es: $\exists n \in \mathcal{M}$, tal que $\neg p(n)$.

2. $\neg [\exists n \in \mathcal{M}, \text{ tal que } p(n)]$ es: $\forall n \in \mathcal{M}, \neg p(n)$.

Ejemplo 52

1. La negación de “Todos los hombres son mortales” es “Existe un hombre que no es mortal”.
2. La negación de “Existe un mamífero que vuela” es “Todos los mamíferos no vuelan”. Claro que aquí la naturaleza de nuestro idioma nos haría expresar la segunda afirmación de la manera siguiente: “Ningún mamífero vuela”.
3. La negación de “ningún número racional es la raíz cuadrada de 2” ($\forall q \in \mathbb{Q}, q \neq \sqrt{2}$) es “existe un número racional q tal que $q = \sqrt{2}$ ”.

Ejemplo 53 . Consideremos las siguientes proposiciones:

1. Para cada natural n , existe un natural $k > n$,
 $(\forall n \in \mathbb{N}), (\exists k \in \mathbb{N} (\text{tal que } k > n))$ tiene la siguiente negación:
 $\exists n \in \mathbb{N} \text{ tal que } \neg(\exists k \in \mathbb{N} \text{ tal que } k > n)$, es decir que
 $\exists n \in \mathbb{N} \text{ tal que } (\forall k \in \mathbb{N}, (k \not> n))$.
 Esto último traducido al lenguaje llano, se leería así:
 “Existe un natural n tal que ningún otro natural k es mayor que n ”
2. Se dice que $\lim_{x \rightarrow 0} f(x) = L$ si
 $\forall \varepsilon > 0, (\exists \delta > 0 \text{ tal que } [\forall x, (|x| < \delta \Rightarrow |f(x) - L| < \varepsilon)])$
 Aquí no nos preocupa qué quiere decir esto. Pero aunque no comprendamos el significado, sí podemos decir cuál es su negación:
 $\exists \varepsilon > 0, \neg(\exists \delta > 0 \text{ tal que } [\forall x, (|x| < \delta \Rightarrow |f(x) - L| < \varepsilon)])$, es decir:
 $\exists \varepsilon > 0, (\forall \delta > 0, \neg[\forall x, (|x| < \delta \Rightarrow |f(x) - L| < \varepsilon)])$, es decir:
 $\exists \varepsilon > 0, (\forall \delta > 0, (\exists x, \text{ tal que } (|x| < \delta) \wedge |f(x) - L| \not< \varepsilon))$.

3. Se dice que d es el máximo común divisor de n y m si:

$$(d \mid n \wedge d \mid m) \wedge (\forall g \in \mathbb{N} ((g \mid n) \wedge (g \mid m)) \Rightarrow [g \mid d]))$$

(Se usa el símbolo “ \mid ” para “divide a”).

¿Cómo se expresaría con símbolos que d no es el máximo común divisor de n y m ?

Respuesta: d no es el máximo común divisor de n y m si

$$\neg(d \mid n \wedge d \mid m) \vee \neg(\forall g \in \mathbb{N} ((g \mid n) \wedge (g \mid m)) \Rightarrow [g \mid d]))$$

es decir $\neg(d \mid n \vee d \mid m) \vee (\exists g \in \mathbb{N} \text{ tal que } \neg((g \mid n) \wedge (g \mid m)) \Rightarrow [g \mid d]))$,
 es decir, si

$$d \nmid n \vee d \nmid m \vee (\exists g \in \mathbb{N} ((g \mid n) \wedge (g \mid m)) \wedge \neg[g \mid d]))$$

o sea

$$d \nmid n \vee d \nmid m \vee (\exists g \in \mathbb{N} ((g \mid n) \wedge (g \mid m)) \wedge [g \nmid d]))$$

Que en español, se lee: d no divide a n ó d no divide a m (d no es un divisor común) ó bien hay una g que siendo divisor común de n y de m , no divide a d (esto es, d no sería un múltiplo de cualquier común divisor, y en consecuencia no sería máximo común divisor).

Ejercicio 59 . Se dice que d es el mínimo común múltiplo de n y m si:

$$([n \mid d] \wedge [m \mid d]) \wedge (\forall k \in \mathbb{N}, [([n \mid k] \wedge [m \mid k]) \Rightarrow (d \mid k)])$$

Exprese en símbolos: “ d no es el mínimo común múltiplo de n y m ”

Ejercicio 60 . Decida los valores de verdad de las proposiciones en los ejemplos 52 y 53.

En los ejercicios siguientes, escribir la negación de los enunciados, sin anteponer un “no”.

Ejercicio 61 . $(\forall x \in A, p(x)) \wedge (\exists z \in A \text{ tal que } q(z))$.

Ejercicio 62 . $\forall x \in A, p(x) \wedge q(x)$.

Ejercicio 63 . $\exists x \in A, p(x) \vee q(x)$.

Ejercicio 64 . $\forall x \in A, p(x) \Rightarrow q(x)$.

Ejercicio 65 . $\forall x \in A, p(x) \Leftrightarrow q(x)$.

Ejercicio 66 . $\exists x \in A \text{ tal que } \forall y \in B, (x, y) \in S$.

Ejercicio 67 . $\forall x \in \mathbb{R}, \forall y \in \mathbb{R}, x + y = y + x$.

Ejercicio 68 . $\exists x \in \mathbb{R}, \forall y \in \mathbb{R}, xy = 0$.

Ejercicio 69 . $\forall x \in \mathbb{R}, \exists y \in \mathbb{R} \text{ tal que } x + y = 0$.

Ejercicio 70 . $\exists y \in \mathbb{R} \text{ tal que } \forall x \in \mathbb{R}, x + y = 0$. ¿Es cierta?

Capítulo 2

Conjuntos y funciones

En este capítulo más que la teoría de los conjuntos presentaremos, una breve introducción al Álgebra de los conjuntos y la manera de usarlos como un lenguaje adecuado para el mejor entendimiento y manejo de algunos de los conceptos de la Matemática. De la misma manera que con el término circunferencia nos referimos a todos los puntos de un plano que equidistan de un punto fijo (el centro de dicha circunferencia), o con el término gráfica de una función f a todos los puntos en el plano de la forma $(x, f(x))$, y el término evento sirve para considerar varios casos posibles en un fenómeno aleatorio, usaremos el término conjunto, en general, para referirnos a varios objetos a la vez (los elementos del conjunto), más aún, con el objeto de tener un lenguaje lo más general posible, usaremos también la palabra conjunto para referirnos a un solo objeto o a ninguno. Esto tendrá sus ventajas como veremos más adelante.¹

2.1 Axiomas

De acuerdo con lo anterior, diremos que un conjunto está formado por objetos a los que llamaremos elementos, o de una manera más precisa los llamaremos sus elementos. Aceptaremos que son los elementos de un conjunto los

¹Lo anterior no tiene por que extrañarnos, después de todo en nuestros estudios anteriores aceptamos cosas como que $4^1 = 4$, o que $4^0 = 1$, o que $4^{\frac{1}{2}} = 2$, ¿y qué sentido tiene el hablar del producto de un cuatro, o del producto de cero cuatros, o del producto de medio cuatros? Simplemente aceptamos como válidas las igualdades anteriores para no tener excepciones en la leyes de los exponentes.

que lo determinan. Así, si dos conjuntos tienen los mismos elementos, para nosotros serán iguales aunque alguien pueda apreciar en uno de ellos alguna característica distintiva.

Axioma 4 (de extensión) . *Dos conjuntos son iguales si y sólo si tienen los mismos elementos.*

Para expresar el hecho de que un objeto x es un elemento de un conjunto A usaremos la notación siguiente:

$$x \in A$$

Y entonces, si A y B son dos conjuntos, otra manera de expresar que $A = B$, es:

$$x \in A \text{ si y sólo si } x \in B.$$

De igual manera que para decir que dos objetos no son iguales usamos el símbolo de igualdad cruzado (\neq), para decir que un objeto x no pertenece a un conjunto A usaremos la notación siguiente:

$$x \notin A.$$

Una forma muy usual de denotar a los conjuntos es simplemente mediante la lista de sus elementos entre llaves, así si:

D es el conjunto formado por 0, $\{0\}$ y $\{2, 3\}$,
 E es el conjunto cuyos elementos son 1, 2 y 3.

F es el conjunto al que pertenecen $\{0, 2\}$, 5 y $\{4\}$.

Entonces

$$D = \{0, \{0\}, \{2, 3\}\},$$

$$E = \{1, 2, 3\},$$

$$F = \{\{0, 2\}, 5, \{4\}\},$$

$$0 \in D,$$

$$0 \notin E,$$

$$1 \in E,$$

$$2 \notin F,$$

$$\{0, 2\} \in F,$$

$$\{0, 2\} \notin D.$$

$$E = \{1, 3, 2\} = \{2, 1, 3\} = \{2, 3, 1\} = \{3, 2, 1\} = \{3, 1, 2\}.$$

Como se puede ver en los ejemplos anteriores, los elementos de un conjunto pueden tener características comunes o ser de lo más disímiles.

2.1.1 Pertenencia y contención

Para denotar un conjunto en la forma descrita anteriormente lo importante, realmente, es que se alisten todos los elementos del conjunto en cuestión, y no el orden en que ésto se haga. Naturalmente los conjuntos infinitos no se pueden denotar así, a menos que de una pequeña lista de sus elementos se sobreentienda cuáles son los demás, como, por ejemplo, si \mathbb{N} es el conjunto de los números naturales:

$$\mathbb{N} = \{0, 1, 2, 3, \dots\},$$

y si \mathbb{Z} es el conjunto de los números enteros:

$$\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}.$$

Las dos relaciones básicas en la Teoría de Conjuntos son, la relación de pertenencia (\in), que es de objeto (elemento) a conjunto, y la relación de contención (\subseteq), que es de conjunto a conjunto y describimos a continuación.

Definición 11 . *Un conjunto A está contenido en un conjunto B, si todos los elementos de A son elementos de B. Lo anterior se denota de la manera siguiente:*

$$\mathbf{A} \subseteq \mathbf{B}.$$

En este caso también podemos decir que A es un subconjunto de B, o que B contiene a A, y denotarlo, $\mathbf{B} \supseteq \mathbf{A}$.

Como ejemplos de contenciones citamos los siguientes:

Ejemplo 54

$$\mathbb{N} = \{0, 1, 2, 3, \dots\} \subseteq \mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}.$$

Ejemplo 55

$$4\mathbb{N} = \{0, 4, 8, 12, \dots\} \subseteq 2\mathbb{N} = \{0, 2, 4, 6, \dots\}.$$

Ejemplo 56

$$6\mathbb{N} = \{0, 6, 12, 18, \dots\} \subseteq 3\mathbb{Z} = \{\dots, -9, -6, -3, 0, 3, 6, 9, \dots\}.$$

Ejemplo 57

$$10\mathbb{Z} = \{..., -30, -20, -10, 0, 10, 20, 30, ...\} \subseteq 5\mathbb{Z}$$

Observación 4 . Es inmediato de la definición anterior que:

1. Para cualquier conjunto A , $A \subseteq A$.
2. Para dos conjuntos A y B cualesquiera, $A = B$ si y sólo si, $A \subseteq B$ y $B \subseteq A$.

Otra manera de describir conjuntos es mediante alguna o algunas propiedades que todos sus elementos tengan, por ejemplo

$$\{-18, -15, -12, -9, -6, -3, 0, 3, 6, 9, 12, 15, 18\}$$

es el conjunto de (todos) los múltiplos de tres entre -20 y 20 y una forma de denotarlo es:

$$\begin{aligned} & \{-18, -15, -12, 9, 6, 3, 0, 6, 9, 12, 15, 18\} = \\ & = \{x \mid x \text{ es entero múltiplo de } 3, \text{ entre } -20 \text{ y } 20\} \end{aligned}$$

que se lee: El conjunto de las x tales que x es entero múltiplo de 3 entre -20 y 20 . y también:

$$\begin{aligned} & \{-18, -15, -12, -9, -6, -3, 0, 3, 6, 9, 12, 15, 18\} = \\ & = \{x \mid x \in \mathbb{Z}, x \text{ es múltiplo de tres y } -19 \leq x \leq 19\}, \end{aligned}$$

que se lee: el conjunto de las x tales que x está en \mathbb{Z} , es múltiplo de 3 y x es mayor o igual que -19 pero menor igual que 19 .

Naturalmente cuando describimos un conjunto de la manera anterior nos referimos al conjunto de todos los objetos que tienen las propiedades mencionadas.

Las propiedades “ser un número natural menor que 17” y “ser múltiplo de 3 o de 4” determinan el conjunto:

$$\{0, 3, 6, 9, 12, 15, 0, 4, 8, 12, 16\}.$$

La anterior es una forma de escribir el conjunto mencionado, listando los múltiplos de 3 y luego los de 4, con la restricción de que sean naturales menores que 17. Claro que una sería mejor denotarlo

$$\{0, 3, 4, 6, 8, 9, 12, 15, 16\},$$

pero lo importante, realmente, es que aparezcan todos sus elementos en la lista y sólo ellos, ya que un objeto es elemento de un conjunto o no lo es, y si lo es, lo es exactamente una vez y no más, aunque lo digamos muchas veces.

2.1.2 Especificación y existencia

Con el fin de precisar cuándo podemos describir conjuntos a partir de propiedades aceptaremos el siguiente axioma.

Axioma 5 (de especificación) . *Si X es un conjunto y p es una propiedad, los elementos de X que tienen la propiedad p forman un conjunto.*

El conjunto mencionado se denota de la siguiente manera:

$$\{x \in X \mid p(x)\}$$

y se lee: el conjunto de las x en X tales que x tiene la propiedad p . Es decir, $p(x)$ es una forma abreviada de escribir que x tiene la propiedad p , y también se lee simplemente: “ p de x ”. Entonces el conjunto anterior también se lee: “el conjunto de las x en X tales que p de x ”.

Axioma 6 (de existencia) . *Existe algún conjunto.*

Si aceptamos el axioma anterior, y más nos vale, ya que de no ser así, no tiene ningún caso desarrollar esta teoría, consideremos cualquier conjunto A y la propiedad “ser distinto de sí mismo”. Entonces $\{x \in A \mid x \neq x\}$ es un conjunto sin elementos (ningún objeto es distinto de si mismo), y ya que son los elementos de cada conjunto los que lo determinan, este conjunto no depende del conjunto A . Llamaremos a este conjunto sin elementos, “el conjunto vacío” y lo denotaremos con la letra \emptyset .

Una propiedad importante del conjunto vacío \emptyset , es que es subconjunto de cualquier otro conjunto, así si A es un conjunto cualquiera, $\emptyset \subseteq A$. Mas aún, la propiedad anterior caracteriza al conjunto vacío, es decir, si un conjunto X es subconjunto de cualquier conjunto, X debe de ser el conjunto vacío.

Ejercicio 71 . *Muestre que \emptyset es un subconjunto de cualquier otro conjunto B . (Sugerencia: use las propiedades de la implicación para mostrar que $x \in \emptyset \Rightarrow x \in B$, es una proposición verdadera. Como segunda posibilidad, respóndase la siguiente pregunta ¿Qué pasaría si $\emptyset \subseteq B$ fuera falsa?).*

Ejercicio 72 . *Use el ejercicio anterior para demostrar que sólo hay un conjunto vacío. (Sugerencia: use el ejercicio anterior y el axioma de extensión).*

2.1.3 No hay un conjunto de todos los conjuntos

Teorema 3 . *Para cualquier conjunto A , hay un objeto B que no es elemento de A .*

Demostración. Si A es un conjunto cualquiera, a partir de él y de la propiedad “no ser un elemento de sí mismo” podemos formar el conjunto

$$B = \{x \in A \mid x \notin x\},$$

y es fácil de comprobar que $B \notin A$. Si suponemos que $B \in A$, entonces $B \in B$ ó $B \notin B$.

Si $B \in B$, entonces siendo B un elemento de A que pertenece a sí mismo, tendríamos que $B \notin B$, por lo tanto, $B \notin B$.

Ahora, como $B \notin B$, y $B \in A$, entonces $B \in B$, llegamos a la contradicción: $B \notin B$ y $B \in B$.

Hemos demostrado que $B \notin A$. ■

Así, no importa qué conjunto se tome, siempre hay un segundo conjunto que no pertenece al primero.

De acuerdo a lo anterior, no hay un conjunto que tenga a todos los objetos como sus elementos, es decir:

No existe un conjunto que contenga a todos los conjuntos.

La afirmación anterior puede parecerle extraña al lector, ya que posiblemente en sus estudios anteriores haya trabajado con algo a lo que le llamó el Conjunto Universal, sin embargo tal término se utiliza, de acuerdo a lo que dijimos al principio de este trabajo, para referirnos a varios objetos a la vez, en cada caso, a todos los objetos necesarios para estudiar alguna situación específica.

Si ahondar demasiado en esto, diremos que los conjuntos son las colecciones que no son demasiado grandes, por ejemplo no es válido hablar del conjunto de todos los objetos, ya que ellos forman una clase pero no un conjunto.²

Notemos que las propiedades por sí solas no determinan conjuntos. Las propiedades determinan conjuntos cuando son las propiedades de los elementos de un conjunto dado. Aplicando el axioma de especificación a un conjunto y alguna propiedad, se obtiene, como ya hemos visto, un subconjunto del conjunto original, el de los elementos de éste que tienen la propiedad en cuestión.

2.1.4 Intersecciones y complementos

A partir de dos conjuntos, el axioma de especificación permite construir un nuevo conjunto de la siguiente manera: si A y B son conjuntos considere la propiedad “ser un elemento de B ”. Los elementos de A que tienen esa propiedad forman el conjunto:

$$\{x \in A \mid x \in B\},$$

éste está formado por los elementos comunes a los dos conjuntos, se le conoce como la intersección de A y B y se denota $A \cap B$. Si A , B y C son conjuntos, podemos hacer las afirmaciones siguientes:

Observación 5 .

1. $A \cap A = A$.
2. $A \cap B = B \cap A$.
3. $A \cap (B \cap C) = (A \cap B) \cap C$.
4. $A \supseteq A \cap B$, $A \cap B \subseteq B$.

²Para hacer más preciso esto, se han introducido los axiomas de la Teoría de Conjuntos. En estos axiomas se incluyen reglas para producir conjuntos a partir de otros. Por ejemplo. el axioma de la potencia (ver Axioma 11, en la página 78) nos permite afirmar que si A es un conjunto, entonces hay un conjunto $\wp(A)$ cuyos elementos son los subconjuntos de A .

Ahora, aunque en la sucesión de conjuntos $A, \wp(A), \wp\wp(A), \dots$ tengamos cada vez conjuntos más numerosos, podemos estar seguros de que todos ellos son conjuntos porque se han construido de acuerdo a los axiomas.

5. $A = A \cap B$ si y sólo si, $A \subseteq B$.
6. $A \cap \emptyset = \emptyset$.

Naturalmente que otro conjunto que podemos obtener a partir de dos conjuntos A y B es:

$$\{x \in A \mid x \notin B\},$$

que se llama la diferencia de A y B , o A menos B y se denota: $A \setminus B$. Si A , B , y C son conjuntos, las siguientes relaciones son ciertas:

Observación 6 .

1. $A \setminus B = \emptyset$ si y sólo si, $A \subseteq B$. En particular
2. $A \setminus A = \emptyset$.
3. $A \setminus B = A$ si y sólo si $A \cap B = \emptyset$.
4. $A \setminus B = B \setminus A$ si y sólo si, $A = B$.
5. $A \setminus (B \setminus C) = (A \setminus B) \setminus C$ si y sólo si, $A \cap C = \emptyset$.
6. $A \supseteq A \setminus B$.
7. $(A \setminus B) \cap B = \emptyset$.
8. $A \setminus \emptyset = A$.
9. $(A \cap B) \cap (A \setminus B) = \emptyset$.

Convencerse de que las proposiciones en las observaciones 5 y 6 son ciertas no es muy difícil, sin embargo llegar a la conclusión en el inciso 5. de la observación anterior requiere de un pequeño análisis:

Si $A \setminus (B \setminus C) = (A \setminus B) \setminus C$, entonces $A \cap C$ no puede tener elementos: Supongamos que $a \in A \cap C$. Como $B \setminus C$ consta de elementos que no están en C y como $a \in C$, entonces $a \notin B \setminus C$. Como $a \in A$, tendríamos que $a \in A \setminus (B \setminus C)$ y por lo tanto $a \in (A \setminus B) \setminus C$.

Como también $a \in C$, concluimos que $a \notin (A \setminus B) \setminus C$. Contradicciendo la afirmación anterior. Por lo tanto

$$A \cap C = \emptyset.$$

Recíprocamente, si $A \cap C = \emptyset$, dado que tanto $A \setminus (B \setminus C)$ como $(A \setminus B) \setminus C$ constan de elementos de A , para probar que son iguales es suficiente verificar que cada elemento a de A , está en uno de dichos conjuntos si y sólo si está en el otro.

Entonces para $a \in A$:

Si $a \in B$, como $a \notin C$, entonces $a \in B \setminus C$ y así $a \notin A \setminus (B \setminus C)$. Como $a \in B$, entonces $a \notin (A \setminus B)$ y entonces $a \notin (A \setminus B) \setminus C$.

Si $a \notin B$, entonces $a \notin B \setminus C$, y por lo tanto, dado que $a \in A$, entonces $a \in A \setminus (B \setminus C)$. Como $a \notin B$ y $a \in A$, entonces $a \in A \setminus B$ y como $a \notin C$, obtenemos, $a \in (A \setminus B) \setminus C$.

Como ya mencionamos anteriormente, en muchas situaciones consideramos que hay un conjunto del cual son elementos todos los objetos que aparecen en ellas, a dicho conjunto se le llama el conjunto universal (para dicha situación) y se le denota con la letra \mathbb{U} . Entonces ya que cualquier conjunto A (en cuestión) es un subconjunto de \mathbb{U} a la diferencia $\mathbb{U} \setminus A$ se le llama el complemento de A y se le denota A^c , que con frecuencia se lee simplemente A complemento.

Con estas convenciones si A y B son dos conjuntos tenemos la siguiente:

Observación 7

$$1. A \setminus B = A \cap B^c.$$

$$2. A^c = \{x \mid x \notin A\}.$$

Algo que es conveniente aceptar, es que los conjuntos pueden ser, a su vez, elementos de algún otro conjunto, si esto es así, consideremos dos conjuntos A y B cualesquiera, tales que $A \in B$, y la propiedad “ser conjunto”, entonces $\{x \in B \mid x \text{ es conjunto}\}$, es un conjunto no vacío y todos sus elementos son conjuntos, es decir, es un conjunto de conjuntos, para evitar este tipo de frases se acostumbra llamar a los conjuntos de conjuntos, familias de conjuntos.

Ejercicio 73 . *Muestre que $X \subseteq A, X \subseteq B \implies X \subseteq A \cap B$.*

2.1.5 Uniones

Siguiendo la idea de que los conjuntos son las colecciones “no muy grandes”³, si \mathcal{F} es una familia de conjuntos, la colección de los objetos que pertenecen

³Esto más que una idea, es una manera de hablar. Una colección es ”grande” cuando no podemos asegurar que sea un conjunto mediante la aplicación de los axiomas de la

a algún elemento de la familia, esperamos que no sea “muy grande”. De manera más precisa, aceptemos el siguiente :

Axioma 7 (de la unión) . *Si \mathcal{F} es una familia de conjuntos, entonces*

$$\{x \mid x \in A \text{ para algún conjunto } A \in \mathcal{F}\}$$

es un conjunto.

El conjunto anterior se llama la unión de \mathcal{F} y se denota: $\cup \mathcal{F}$.

Observación 8

$$\cup \emptyset = \emptyset.$$

De alguna manera, la unión de una familia de conjuntos es el resultado de reunir en un solo conjunto a los elementos de todos los conjuntos de la familia en cuestión, de hecho, en francés se usa el término *reunión* en vez del término *unión* que usamos en México.

Para tener derecho a hablar de la unión de dos conjuntos es conveniente tener el siguiente:

Axioma 8 (de la pareja) . *Si A y B son conjuntos, entonces $\{A, B\}$ es un conjunto.*

Es decir, existe un conjunto cuyos elementos son precisamente A y B .

Como consecuencia de los dos axiomas anteriores, si A y B son dos conjuntos cualesquiera,

$$\cup \{A, B\} = \{x \mid x \in A \vee x \in B\},$$

es el conjunto formado por los elementos de A junto con los elementos de B . Dicho conjunto se denota también:

$$A \cup B$$

y se le llama la unión de A y B .

teoría de conjuntos.

En cambio, cuando la aplicación de los axiomas nos permite asegurar que una colección es un conjunto decimos que la colección “no es grande”.

Observación 9 . *Para cualesquiera tres conjuntos A , B y C , se pueden demostrar las siguientes afirmaciones:*

1. $\cup \{A\} = A$.
2. $A \cup B = B \cup A$.
3. $A \cup \emptyset = A$.
4. $A \cup (B \cup C) = (A \cup B) \cup C$.
5. $A \cup B \supseteq A, B \subseteq (A \cup B)$.
6. $A \cup B = A$ si y sólo si $A \supseteq B$.
7. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
8. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.
9. $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$.
10. $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$.
11. $A = (A \cap B) \cup (A \setminus B)$.

Una manera de demostrar que dos conjuntos son iguales es, como ya sabemos, (ver el Axioma de extensión) comprobando que tienen los mismos elementos, y para hacer esto es conveniente considerar uno a uno todos diversos casos posibles en los que pueden ocurrir los elementos. Por ejemplo, para demostrar que

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

un elemento x tiene 8 posibilidades:

$$\begin{aligned} & x \in A, x \in B, x \in C \text{ ó} \\ & x \in A, x \in B, x \notin C \text{ ó} \\ & x \in A, x \notin B, x \in C \text{ ó} \\ & x \in A, x \notin B, x \notin C \text{ ó} \\ & x \notin A, x \in B, x \in C \text{ ó} \\ & x \notin A, x \in B, x \notin C \text{ ó} \\ & x \notin A, x \notin B, x \in C \text{ ó} \\ & x \notin A, x \notin B, x \notin C . \end{aligned}$$

Una forma más cómoda de expresar lo anterior es mediante una tabla como la siguiente:

A	B	C
1	1	1
1	1	0
1	0	1
1	0	0
0	1	1
0	1	0
0	0	1
0	0	0

en donde cada uno de los renglones de la tabla expresa cada una de las posibilidades mencionadas, en el mismo orden, así por ejemplo en el renglón tercero el uno, el cero y el uno significan está, no está y está, en A , en B y en C respectivamente. Con la misma notación podemos completar la tabla para convencernos de la veracidad de la igualdad en cuestión:

A	B	C	$B \cap C$	$A \cup (B \cap C)$	$A \cup B$	$A \cup C$	$(A \cup B) \cap$ $\cap (A \cup C)$	$A \cup (B \cap C) =$ $= (A \cup B) \cap (A \cup C)$
0	0	0	0	0	0	0	0	1
0	0	1	0	0	0	1	0	1
0	1	0	0	0	1	0	0	1
0	1	1	1	1	1	1	1	1
1	0	0	0	1	1	1	1	1
1	0	1	0	1	1	1	1	1
1	1	0	0	1	1	1	1	1
1	1	1	1	1	1	1	1	1

Como se puede ver en la tabla, los elementos considerados en cada renglón pertenecen a $A \cup (B \cap C)$ si y sólo si pertenecen a $(A \cup B) \cap (A \cup C)$. Es decir, las dos columnas de los conjuntos que estamos probando que son iguales, coinciden.

Mediante tablas como las anteriores podemos probar también igualdades condicionadas como: $A \setminus (B \setminus C) = (A \setminus B) \setminus C$ si y sólo si $A \cap C = \emptyset$:

\emptyset	A	B	C
0	0	0	0
0	0	0	1
0	0	1	0
0	0	1	1
0	1	0	0
0	1	0	1
0	1	1	0
0	1	1	1

$A \setminus B$	$B \setminus C$	$A \setminus (B \setminus C)$	$(A \setminus B) \setminus C$	$A \cap C$	$A \setminus (B \setminus C) = (A \setminus B) \setminus C$	$A \cap C = \emptyset$
0	0	0	0	0	1	1
0	0	0	0	0	1	1
0	1	0	0	0	1	1
0	0	0	0	0	1	1
1	0	1	1	0	1	1
1	0	1	0	1	0	0
0	1	0	0	0	1	1
0	0	1	0	1	0	0

Como se ve en la tabla, $[A \setminus (B \setminus C) = (A \setminus B) \setminus C] \Leftrightarrow (A \cap C = \emptyset)$ es una tautología.

Observación 10 *Cuando se toma A como el conjunto universal en los incisos 8) y 9) de la observación 9, las igualdades obtenidas se conocen como las leyes de De Morgan, y quedan expresadas:*

1. $(B \cup C)^c = B^c \cap C^c$.
2. $(B \cap C)^c = B^c \cup C^c$.

En lenguaje coloquial, las leyes de De Morgan se expresan de la siguiente manera:

- El complemento de la unión es la intersección de los complementos.
- El complemento de la intersección es la unión de los complementos.

Ejercicio 74 . En un cierto país al 63% de la población le gusta el queso, y al 76% le gustan las manzanas. ¿Qué se puede decir del porcentaje de la población al que le gustan el queso y las manzanas.

Ejercicio 75 . En un grupo de estudiantes hay 71 alumnos a los que les gusta el Álgebra pero no la Geometría, hay 60 a los que les gusta la Geometría pero no les gusta el Cálculo. A 12 estudiantes les gusta el Álgebra, la Geometría y el Cálculo; hay 30 estudiantes a los que les gusta el Cálculo y no les gustan ni el Álgebra ni la Geometría. Hay 151 estudiantes a los que les gusta el Álgebra o el Cálculo. Hay 135 a los que les gusta la Geometría o el Cálculo. A 32 estudiantes les gusta el Álgebra y la Geometría. Hay nueve estudiantes a los que no les gusta ninguna de las tres materias. ¿A cuántos estudiantes les gusta el Álgebra? ¿Cuántos estudiantes hay que les gusta el Álgebra y la Geometría pero no el Cálculo? ¿Cuántos hay a los que les gusta el Álgebra y el Cálculo pero no la Geometría? ¿Cuántos estudiantes hay?

Denotemos con U el conjunto universal. Demuestre las afirmaciones siguientes:

Ejercicio 76 . $A \cup B = U \implies A^c \subseteq B$.

Ejercicio 77 . $A \cap C = \emptyset \implies C \subseteq A^c$.

Ejercicio 78 . $(A \cup X = U) \wedge (A \cap X = \emptyset) \iff X = A^c$.

Ejercicio 79 . $A^c \subsetneq C \implies A \cap C \neq \emptyset$. (recuerde que \subsetneq denota inclusión propia, es decir que $X \subsetneq Y$ significa que $X \subseteq Y$ pero $X \neq Y$).

Ejercicio 80 . $B \subsetneq A^c \implies A \cup B \neq U$.

2.1.6 Familias

Algunas de las propiedades de las operaciones finitas entre conjuntos son válidas también en el caso infinito, sin embargo a fin de simplificar los enunciados es conveniente tener el concepto de familia indicada de conjuntos.

Axioma 9 (de reemplazo) . Si I es un conjunto y para cada $i \in I$, X_i es un conjunto, existe un conjunto, denotado $\{X_i\}_{i \in I}$ cuyos elementos son precisamente los conjuntos X_i .

Se suele decir que $\{X_i\}_{i \in I}$ es una familia de conjuntos indicada por el conjunto I .

En las familias indicadas de conjuntos, por ejemplo $\{X_i\}_{i \in I}$, para dos elementos i, j en I distintos, bien puede suceder que $X_i = X_j$. De hecho esta es una propiedad muy útil de las familias indicadas de conjuntos ya que permite que un mismo conjunto juegue varios papeles simultáneamente. Lo anterior quedará más claro cuando veamos el concepto de función y el de producto cartesiano de conjuntos.

En la situación anterior podemos entender que cada elemento i del conjunto I determina o indica un conjunto (justamente a X_i), así el conjunto I se llama el conjunto de índices. Naturalmente toda familia \mathcal{F} de conjuntos se puede ver como una familia indicada:

para $A \in \mathcal{F}$, tomemos $X_A = A$, aún cuando no sean exactamente lo mismo, para muchos efectos $\{X_A\}_{A \in \mathcal{F}}$ juega el mismo papel que \mathcal{F} .

Cuando en el contexto no haya lugar a confusión, al referirnos a una familia indicada de conjuntos suprimiremos el término “indicada”.

Observación 11 . *Si $\{X_i\}_{i \in I}$ es una familia indicada de conjuntos, con I no vacío y $j \in I$, entonces*

$$\{x \in X_j \mid x \in X_i, \text{ para cada } i \in I\}$$

es el conjunto de los elementos que pertenecen a cada uno de los conjuntos X_i . De acuerdo con la terminología anterior, lo llamaremos la intersección de $\{X_i\}_{i \in I}$ y lo denotaremos:

$$\cap_{i \in I} \{X_i\}.$$

Axioma 10 (de la unión) . *Si $\{X_i\}_{i \in I}$ es una familia de conjuntos, entonces*

$$\{x \mid x \in X_i \text{ para alguna } i \in I\}$$

es un conjunto.

El conjunto anterior está formado por todos los elementos de los conjuntos de la familia, se llamará la unión de la familia $\{X_i\}_{i \in I}$ y se denotará:

$$\cup_{i \in I} \{X_i\}.$$

Algunas de las propiedades de las operaciones entre conjuntos enunciadas anteriormente son válidas cuando se consideran familias de conjuntos, por ejemplo:

Las leyes distributivas de la unión y la intersección para familias de conjuntos son:

Observación 12 . *Si $\{X_i\}_{i \in I}$ es una familia de conjuntos y Y es un conjunto entonces,*

1. $(\bigcup_{i \in I} \{X_i\}) \cap Y = \bigcup_{i \in I} \{X_i \cap Y\}.$
2. $(\bigcap_{i \in I} \{X_i\}) \cup Y = \bigcap_{i \in I} \{X_i \cup Y\}.$

Demostración. Para demostrar 1, podemos proceder comprobando que cada elemento del primer conjunto está en el segundo y recíprocamente. que cada elemento del segundo está en el primero:

\Rightarrow Si $x \in (\bigcup_{i \in I} \{X_i\}) \cap Y$, entonces $x \in \bigcup_{i \in I} \{X_i\}$ y $x \in Y$. Por lo tanto, para alguna $i \in I$, $x \in X_i$. Además, $x \in Y$, por lo que para dicha $i \in I$,

$$x \in X_i \cap Y,$$

de donde concluimos que $x \in \bigcup_{i \in I} \{X_i \cap Y\}$.

\Leftarrow Si $x \in \bigcup_{i \in I} \{X_i \cap Y\}$, entonces para alguna $i \in I$, $x \in X_i \cap Y$, y por lo tanto, $x \in X_i$ para dicha i . Además $x \in Y$, por lo que $x \in \bigcup X_i$ y $x \in Y$, es decir que

$$x \in (\bigcup_{i \in I} \{X_i\}) \cap Y.$$

La demostración de 2, se deja como ejercicio. ■

Ejercicio 81 . *Demuestre que $(\bigcap_{i \in I} \{X_i\}) \cup Y = \bigcap_{i \in I} \{X_i \cup Y\}$.*

Las leyes de De Morgan también se pueden enunciar y demostrar para familias de conjuntos de la manera siguiente:

Si $\{X_j\}_{j \in J}$ es una familia de conjuntos y Y es un conjunto entonces,

- 1) $Y \setminus (\bigcup_{j \in J} \{X_j\}) = \bigcap_{j \in J} \{Y \setminus X_j\}.$
- 2) $Y \setminus (\bigcap_{j \in J} \{X_j\}) = \bigcup_{j \in J} \{Y \setminus X_j\}.$

Cuando Y es el conjunto universal, las dos igualdades anteriores quedan expresadas:

$$1) (\cup_{j \in J} \{X_j\})^c = \cap_{j \in J} \{X_j^c\}.$$

$$2) (\cap_{j \in J} \{X_j\})^c = \cup_{j \in J} \{X_j^c\}.$$

Que como dijimos anteriormente se pueden enunciar diciendo que:

1. El complemento de la unión es la intersección de los complementos.
2. El complemento de la intersección es la unión de los complementos.

2.1.7 La diferencia simétrica

Otra operación entre conjuntos es la diferencia simétrica .

Definición 12 . *Si A y B son dos conjuntos, definimos*

$$A \Delta B \stackrel{\text{def}}{=} (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B),$$

la diferencia simétrica de A y B .

Observación 13 . *Para cualesquier tres conjuntos A , B y C se tienen las igualdades siguientes:*

1. $A \Delta (B \Delta C) = (A \Delta B) \Delta C.$
2. $A \Delta B = B \Delta A.$
3. $A \Delta \emptyset = A, A \Delta A = \emptyset.$
4. $A \cap (B \Delta C) = (A \cap B) \Delta (A \cap C).$

2.1.8 El conjunto potencia

Otra manera de obtener nuevos conjuntos es mediante el:

Axioma 11 (de las partes) . *Si X es un conjunto, los subconjuntos de X son los elementos de un conjunto.*

Al conjunto anterior, la familia de los subconjuntos de X (según la manera que ya mencionamos de referirnos a los conjuntos cuyos elementos son a su vez conjuntos), se le llama la potencia de X o las partes de X y se le denota $\wp(X)$, o 2^X entonces:

$$\wp(X) = \{A \mid A \subseteq X\}.$$

En los ejercicios siguientes, demuestre la validez de los enunciados, acerca de los conjuntos A y B .

Ejercicio 82 . $A \subseteq B \implies \wp(A) \subseteq \wp(B)$.

Ejercicio 83 . $\wp(A) \subseteq \wp(B) \implies A \subseteq B$.

Ejercicio 84 . $A = B \iff \wp(A) = \wp(B)$.

Ejercicio 85

$$1. \quad \wp(A) \cap \wp(B) = \wp(A \cap B).$$

$$2. \quad \wp(A) \cup \wp(B) \subseteq \wp(A \cup B).$$

Ejercicio 86 . *Construya un ejemplo de conjuntos A y B para los que*

$$\wp(A) \cup \wp(B) \not\subseteq \wp(A \cup B).$$

Se dice que un conjunto T es transitivo si $X \in T \implies X \subseteq T$.

Ejercicio 87 . *Muestre que \emptyset es transitivo y que $\{\emptyset\}$ también lo es.*

Ejercicio 88 . *Use axiomas de la Teoría de Conjuntos para construir, a partir del conjunto vacío, 20 conjuntos que no sean transitivos.*

Ejercicio 89 . *Sea N un conjunto. Diga si existe un conjunto cuyos elementos sean los conjuntos de la lista siguiente:*

$$. N, \sigma(N), \sigma(\sigma(N)), \sigma(\sigma(\sigma(N))), \dots$$

¿por qué? (Ver el ejercicio 2.1, en la página 86).

Ejercicio 90 . *Sea $U = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ resuelva la ecuación*

$$\{0, 3, 6, 7\} \Delta X = \{0, 7, 1, 2\},$$

donde Δ denota la diferencia simétrica, y desde luego X denota un subconjunto de U .

2.2 Parejas ordenadas, producto cartesiano y relaciones

Un concepto importante para el desarrollo posterior de la Teoría de los Conjuntos, así como para el estudio de las funciones, el de las estructuras algebraicas, los sistemas coordenados y varias cosas más, es el concepto de pareja ordenada.

Con este término queremos referirnos a dos objetos (elementos), no necesariamente de un mismo conjunto, pero que juegan papeles distintos, distinguiéndolos uno del otro, por ejemplo, especificando de alguna manera cuál es el primero y cuál el segundo.

Así, si A y B son conjuntos, $a \in A$ y $b \in B$, el conjunto

$$\{a, b\} = \{x \in A \cup B \mid x = a \text{ ó } x = b\},$$

no sirve para nuestros fines, ya que como precisamos anteriormente, los elementos de un conjunto lo determinan totalmente. Es decir,

$$\{a, b\} = \{b, a\}.$$

Entonces procederemos de la manera siguiente.

Definición 13 . *Si A y B son conjuntos, $a \in A$ y $b \in B$, la pareja ordenada de a y b es: $(a, b) \doteq \{\{a\}, \{a, b\}\}$.*

Observación 14 . Note que de acuerdo al axioma de especificación,

$$\{a\} = \{x \in A \mid x = a\}$$

y

$$\{a, b\} = \{x \in A \cup B \mid x = a \text{ o } x = b\}$$

son conjuntos. Entonces, por el axioma de la pareja, $\{a\}$ y $\{a, b\}$ forman un conjunto, precisamente el que llamamos (a, b) y ya que

$$\{a\} \in \wp(A \cup B) \text{ y } \{a, b\} \in \wp(A \cup B),$$

entonces

$$(a, b) \in \wp(\wp(A \cup B)).$$

De acuerdo a lo anterior, si A y B son dos conjuntos, entonces para dos elementos $a \in A$ y $b \in B$, se tiene que

$$(a, b) \in \wp(\wp(A \cup B)).$$

Veamos que la definición anterior corresponde a lo que queremos, explícitamente verifiquemos el siguiente teorema.

Teorema 4 . Sean $a, c \in A$, $b, d \in B$, A, B conjuntos. Son equivalentes:

1. $(a, b) = (c, d)$.
2. $(a = c) \wedge (b = d)$.

Demostración. Veamos que $1) \Rightarrow 2)$:

Supongamos que

$$\{\{a\}, \{a, b\}\} \equiv (a, b) = (c, d) \equiv \{\{c\}, \{c, d\}\}$$

Entonces:

$$\{a, b\} = \cup \{\{a\}, \{a, b\}\} = \cup \{\{c\}, \{c, d\}\} = \{c, d\} .$$

Por lo tanto, $\{a, b\} = \{c, d\}$.

También

$$\{a\} = \cap \{\{a\}, \{a, b\}\} = \cap \{\{c\}, \{c, d\}\} = \{c\} .$$

Por lo que $\{a\} = \{c\}$, entonces $a = c$.

Como $\{a, b\} = \{c, d\}$ y $a = c$ entonces

$$\{a, b\} \setminus \{a\} = \{c, d\} \setminus \{c\}.$$

Ahora,

si $\{a, b\} \setminus \{a\} = \{b\}$ entonces $\{c, d\} \setminus \{c\} = \{d\}$, de donde concluimos que $b = d$.

si $\{a, b\} \setminus \{a\} = \emptyset$ entonces $a = b$ y $\{c, d\} \setminus \{c\} = \emptyset$, por lo que $c = d = a = b$, así que también en este caso $b = d$.

En cualquier caso, $a = c$ y $b = d$.

2) \Rightarrow 1) Es claro. ■

Las parejas (a, b) para las que a es un elemento de A y b es un elemento de B forman un conjunto llamado el producto cartesiano de A y B , a saber:

$$A \times B \doteq \{(a, b) \in \wp(A \cup B) \mid a \in A \text{ y } b \in B\}.$$

$A \times B$ y también se llama, simplemente, el producto de A y B , o A por B , o A cruz B .

Observación 15 . *Algunas de las propiedades del producto cartesiano de conjuntos y las operaciones mencionadas anteriormente son las siguientes: Si A , B , C y D son conjuntos:*

1. $A \times (B \cup C) = (A \times B) \cup (A \times C)$.
2. $(A \cup B) \times C = (A \times C) \cup (B \times C)$.
3. $A \times (B \cap C) = (A \times B) \cap (A \times C)$.
4. $(A \cap B) \times C = (A \times C) \cap (B \times C)$.
5. $(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$.
6. $A \times B = \emptyset$, si y sólo si $A = \emptyset$ o $B = \emptyset$.
7. $(A \setminus B) \times C = (A \times C) \setminus (B \times C)$.

Ejercicio 91 . *Demostrar los incisos de la observación anterior.*

El concepto de producto cartesiano de conjuntos y el de pareja ordenada permiten tratar en forma precisa el de *relación*. Por ejemplo si en el conjunto $X = \{2, 3, \dots, 9\}$ consideramos la relación “ser divisor de”, dicha relación y el conjunto

$$\left\{ \begin{array}{l} (2, 2), (2, 4), (2, 6), (2, 8), (3, 3), (3, 6), (3, 9), \\ (4, 4), (4, 8), (5, 5), (6, 6), (7, 7), (8, 8), (9, 9) \end{array} \right\}$$

proporcionan la misma información.

Generalizando lo anterior podemos aceptar como definición la siguiente:

Definición 14

1. Si A y B son conjuntos, una *relación* R de A en B es un subconjunto del producto cartesiano $A \times B$
2. A una relación de un conjunto A en sí mismo simplemente le llamaremos *una relación en A* .
3. Al conjunto $\{x \in A \mid \exists y \in B \text{ tal que } (x, y) \in R\}$ se llama el *dominio de la relación R* y al conjunto B el *contradominio* de la misma.

Algunos ejemplos de relaciones son los siguientes:

Ejemplo 58 y Ejemplo 59

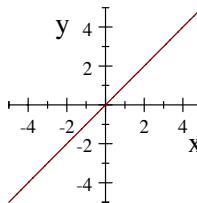
1. Si A es un conjunto, la diagonal de $A \times A$, $D(A)$ es el conjunto

$$\{(a, a) \mid a \in A\}.$$

Nótese que esta relación no es más que la relación de igualdad en A . Es decir:

$$\{(a, b) \in A \times A \mid a = b\} = \{(a, a) \mid a \in A\}.$$

El nombre diagonal de $A \times A$ se debe a que si A es un segmento en la Recta Real, $A \times A$ es un cuadrado en el Plano Cartesiano una de cuyas diagonales, como segmento geométrico, es precisamente $D(A)$.



$$\{(x, y) \in \mathbb{R} \times \mathbb{R} \mid x = y\}$$

2. Si A es un conjunto, la relación de pertenencia de los elementos de A a los subconjuntos de A es: $\{(x, X) \in A \times \wp(A) \mid x \in X\}$.
3. Si \mathcal{F} es una familia de conjuntos, la relación de contención entre los elementos de \mathcal{F} es:

$$\{(A, B) \in \mathcal{F} \times \mathcal{F} \mid \mathcal{A} \subseteq \mathcal{B}\}.$$

4. Si R es una relación de A en B , la relación inversa de R denotada R^{-1} es la relación

$$\{(b, a) \in B \times A \mid (a, b) \in R\}.$$

5. Si R es una relación de A en B , $A_1 \subseteq A$ y $B_1 \subseteq B$: entonces

- (a) La restricción de R a $A_1 \times B_1$ es

$$\{(a, b) \in R \mid a \in A_1, b \in B_1\}$$

y se denota $R|_{A_1}^{B_1}$.

$$\begin{aligned} R &\subseteq A \times B \\ R|_{A_1}^{B_1} &\subseteq A_1 \times B_1. \end{aligned}$$

- (b) A la restricción de R a $A_1 \times B$ se le llama simplemente la restricción de R a A_1 y se le denota con: $R|_{A_1}$.

Entonces:

$$R|_{A_1} = \{(a, b) \in R \mid a \in A_1\}.$$

- (c) A la restricción de R a $A \times B_1$ se le conoce como la correstricción de R a B_1 y para referirnos a ella usaremos: $R|^{B_1}$.

Por lo tanto:

$$R|^{B_1} = \{(a, b) \in R \mid b \in B_1\}.$$

Observación 16 . *En cualquier familia de conjuntos \mathcal{F} , la relación de contención \subseteq tiene las siguientes propiedades:*

1. La contención es reflexiva, es decir, para cada $A \in \mathcal{F}$, $A \subseteq A$.
2. La contención es antisimétrica, es decir, para cualesquiera dos conjuntos A y B en \mathcal{F} , si $A \subseteq B$ y $B \subseteq A$, entonces $A = B$.
3. La contención es transitiva, es decir, para cualesquiera tres conjuntos A , B y C en \mathcal{F} , si $A \subseteq B$ y $B \subseteq C$, entonces $A \subseteq C$.

2.2.1 Axioma de regularidad

El siguiente axioma tiene entre otros, el propósito de evitar que un conjunto pueda ser elemento de sí mismo.

Axioma 12 (de regularidad) . *No existe una lista infinita de conjuntos tales que*

$$\cdots \in a_3 \in a_2 \in a_1.$$

Observación 17 . *Como consecuencia, si A es un conjunto, entonces $A \notin A$. Pues si $A \in A$ podríamos escribir una lista*

$$\cdots A \in A \in A \in A,$$

que no termina, pues siempre podríamos agregar otro “ $A \in$ ”.

Otra consecuencia, es que

$$A \in B \Rightarrow B \notin A,$$

pues si $A \in B$ y $B \in A$, entonces podríamos escribir una lista

$$\cdots A \in B \in A \in B \in A$$

que no termina, pues siempre podemos agregar un nuevo “ $A \in B \in$ ”.

Ejercicio 92 . Use el axioma de regularidad para demostrar que no existe un “conjunto de todos los conjuntos”.

Ejercicio 93 . Demuestre que si A, B, X son conjuntos tales que

$$1. A \cup X = B \cup X \text{ y}$$

$$2. A \cap X = B \cap X,$$

$$\text{entonces } A = B.$$

Ejercicio 94 . Use el axioma de regularidad para demostrar que un conjunto no puede ser elemento de sí mismo, de la manera siguiente: considere el conjunto $\{A\}$ y ahora use el hecho de que el conjunto anterior debe contener un elemento ajeno con $\{A\}$. Entonces $\{A\} \cap A = \emptyset$. Concluya.

Ejercicio 95. Use el axioma de regularidad para demostrar que $(A, B) \neq A$.

Ejercicio 96 . Use el axioma de regularidad para demostrar que $(A, B) \neq B$.

Definición 15 . Un conjunto X es transitivo si

$$A \in B \in X \implies A \in X.$$

Ejercicio 97 . Muestre que un conjunto X es transitivo \iff

$$B \in X \implies B \subseteq X.$$

Ejercicio 98 . Muestre que los conjuntos

$$\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}$$

son transitivos.

Ejercicio 99 . Muestre que si A es transitivo y $B \subseteq A$, entonces $A \cup \{B\}$ es transitivo.

Ejercicio 100 . *Demuestre que si A es un conjunto transitivo entonces el sucesor de A ,*

$$\sigma(A) = A \cup \{A\} \quad (2.1)$$

es un conjunto transitivo.

Ejercicio 101 . *Muestre que $\{0, 1, 2, \{1, 2\}\}$ y $\{0, 1, 2, \{1, 2\}, \{0, 2\}\}$ son conjuntos transitivos.*

Ejercicio 102 . *Encuentre el menor conjunto transitivo que contiene a*

$$\{\{2, 4\}, \{1\}, \{\{1\}, \{2, 3\}\}\}.$$

Ejercicio 103 . *Demuestre que si A y B son conjuntos transitivos entonces*

$$A \cap B \text{ y } A \cup B$$

también son transitivos.

Ejercicio 104 . *Encuentre un conjunto transitivo que no sea un número natural y con el menor número posible de elementos.*

2.2.2 Órdenes parciales

En realidad la contención entre los conjuntos de un familia no es más que un ejemplo de lo que se conoce como orden parcial.

Definición 16 . *Diremos que una relación \leq , en un conjunto A , es un orden parcial (en A), si :*

1. La relación es reflexiva, es decir, para cada elemento $x \in A$, $x \leq x$.
2. La relación es antisimétrica, es decir, para cualesquiera dos elementos x y y en A , si $x \leq y$ y $y \leq x$, entonces $x = y$.
3. La relación es transitiva, es decir, para cualesquiera tres elementos x , y y z en A , si $x \leq y$ y $y \leq z$, entonces $x \leq z$.

Puesto que hemos usado el signo de menor o igual para denotar un orden parcial, leeremos la expresión $x \leq y$ de la manera siguiente: “equis menor o igual a ye”.

Para referirnos a un conjunto y a un orden parcial en él, diremos simplemente: “un conjunto ordenado parcialmente”.

Ejemplo 60 . *Un ejemplo de orden parcial en el conjunto de los números naturales es la relación “divide a”.*

Observación 18 . *La relación inversa de un orden parcial \leq en un conjunto A es también un orden parcial, como demostraremos a continuación. También se le llama el orden dual de \leq y se denota \leq^* .*

Demostración. \leq^* es un orden ya que:

- i) Si $a \in A$, $a \leq^* a$, puesto que $a \leq a$.
- ii) Si $a \leq^* b$ y $b \leq^* a$, entonces por la definición de \leq^* , $b \leq a$ y $a \leq b$, y de la antisimetría de \leq concluimos que $a = b$.
- iii) Si $a \leq^* b$ y $b \leq^* c$, entonces $b \leq a$ y $c \leq b$, por la transitividad de \leq sabemos que $c \leq a$ o sea que $a \leq^* c$. ■

Ejemplo 61 . *El orden dual de la relación “divide a” o “ser divisor de” en el conjunto de los Números Naturales es la relación “ser múltiplo de”.*

Observación 19 . *En cierta manera todo orden parcial es la contención entre conjuntos, en el siguiente sentido: si \leq es un orden parcial en un conjunto A , entonces para cada $a \in A$, consideremos el conjunto de los elementos de A menores o iguales a a , y denotémoslo con $\langle a \rangle$, es decir: $\langle a \rangle = \{x \in A \mid x \leq a\}$.*

Entonces para cualesquiera dos elementos a y b en A , $a \leq b$, si y sólo si, $\langle a \rangle \subseteq \langle b \rangle$.

Así, los conjuntos A y $\{\langle a \rangle \in \wp(A) \mid a \in A\}$ tienen las mismas propiedades, vistos como conjuntos parcialmente ordenados.

Observación 20 . *Si vemos la familia $\wp(A)$ de los subconjuntos de un conjunto A como un conjunto ordenado parcialmente por la contención entre conjuntos, entonces para cualesquiera dos conjuntos B y C , el conjunto $B \cup C$ es el conjunto “menor” entre los que son mayores o iguales que B y C .*

Es decir:

1. Si $B \in \wp(A)$ y $C \in \wp(A)$, entonces $B \cup C \in \wp(A)$
2. $B \leq B \cup C$ y $C \leq B \cup C$.
3. Si $D \in \wp(A)$, $B \leq D$ y $C \leq D$, entonces, $B \cup C \leq D$.

De igual manera, $B \cap C$, es el conjunto “mayor” en $\wp(A)$ de los que son menores o iguales que B y que C .

A un conjunto ordenado parcialmente y con las propiedades anteriores le llamaremos retícula.

A fin de enunciar lo anterior de una manera más precisa haremos las definiciones siguientes:

Definición 17 . *Sea A un conjunto parcialmente ordenado y sea $B \subseteq A$.*

1. Diremos que $m \in A$ es una cota superior para B si

$$b \leq m, \quad \forall b \in B.$$

2. Diremos que $\beta \in B$ es el menor elemento de B si

$$\beta \leq b, \quad \forall b \in B.$$

3. Diremos que $s \in A$ es el supremo de B si s es el menor elemento en el conjunto de las cotas superiores de B . Escribiremos

$$s = \sup(B).$$

Observación 21 . *Cuando b y c son elementos de un conjunto ordenado parcialmente A , el supremo de $\{b, c\}$ es un elemento denotado $b \vee c$ de A , tal que:*

1. $b \leq b \vee c$ y $c \leq b \vee c$, y además
2. si $b \leq d$ y $c \leq d$, entonces, $b \vee c \leq d$.

En la misma situación anterior, el ínfimo de $\{b, c\}$ es un elemento de A denotado $b \wedge c$ con las propiedades siguientes:

1. $b \wedge c \leq b$ y $b \wedge c \leq c$, y
2. si $d \leq b$ y $d \leq c$, entonces, $d \leq b \wedge c$.

Proposición 7 . *Si \mathcal{R} es una relación de orden en A y $B \subseteq A$, entonces $\mathcal{R}|_B = \mathcal{R} \cap (B \times B)$ es una relación de orden en B .*

$$\begin{array}{ccc} \mathcal{R} & \hookrightarrow & A \times A \\ \uparrow & & \uparrow \\ \mathcal{R} \cap (B \times B) & \hookrightarrow & B \times B \end{array}$$

Así pues, cuando hablamos de un subconjunto ordenado B de un conjunto ordenado parcialmente A , sobreentenderemos que nos referimos a él como conjunto ordenado parcialmente con dicha restricción.

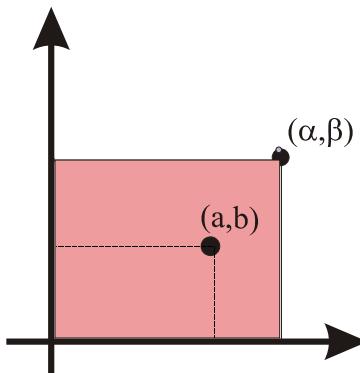
Ejemplo 62 Si

$$A = B = \{x \mid x \text{ es un número real y } 0 \leq x \leq 1\}$$

y $(a, b), (\alpha, \beta) \in A \times B$, entonces

$$(a, b) \leq (\alpha, \beta)$$

si y sólo si (a, b) está en el rectángulo que tiene como vértices opuestos al origen del plano cartesiano $(0, 0)$ y al punto (α, β) .



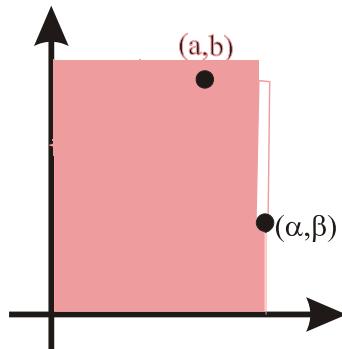
Definición 18 El otro orden al que nos referimos lo denotaremos \leq_l , y está definido por:

$$(a, b) \leq_l (\alpha, \beta) \text{ si } a < \alpha \text{ ó } (a = \alpha \text{ y } b \leq \beta)$$

Para $A \times B$, (a, b) y (α, β) como en el ejemplo anterior,

$$(a, b) \leq (\alpha, \beta)$$

si y sólo si el punto (a, b) es el punto (α, β) o está a la izquierda o exactamente abajo de él.



2.2.3 Retículas

Definición 19 . Diremos que un conjunto ordenado parcialmente A es una retícula superior si cada par de elementos de A tiene supremo en A .

De forma similar un conjunto ordenado parcialmente A es una retícula inferior si cualesquiera dos elementos de A tienen ínfimo en A .

Por último, un conjunto ordenado parcialmente que es tanto retícula superior como inferior se le llama simplemente una retícula.

Ejemplo 63 . Si A es un conjunto con más de un elemento, la familia

$$\{B \subseteq A \mid B \neq \emptyset\}$$

con la contención es una retícula superior que no es retícula inferior. En la misma situación $\{B \subseteq A \mid B \neq A\}$ es una retícula inferior que no es retícula superior.

Definición 20 . $\beta \in A$ es una cota inferior de (B, \leq) , si β es menor o igual que todos los elementos de B .

Definición 21 . *Un elemento m de un conjunto parcialmente ordenado B es el mayor elemento de $C \subseteq B$, si $m \in C$ y $c \leq m, \forall c \in C$.*

Definición 22 . *Un elemento m de un conjunto parcialmente ordenado B es máximo si $m \leq b \implies m = b, \forall b \in B$.*

Recordemos que a la menor cota superior de B , (si la hay) se le llama el supremo de B , y se denota $\sup(B)$ o bien $\bigvee B$.

Observación 22 . *Cuando un conjunto B contiene a su supremo, éste es único: ya que si α_0 y β_0 son ambos supremos de B , entonces $\alpha_0 \leq \beta_0$ (β_0 es cota superior para B y α_0 es la menor de las cotas superiores) pero también $\beta_0 \leq \alpha_0$, así que por la antisimetría de un orden parcial tenemos que $\alpha_0 = \beta_0$.*

Recordemos que a la mayor cota inferior de B , (si la hay) se le llama el ínfimo de B y se denota $\bigwedge B$ o $\inf(B)$.

Observación 23 . $\beta_0 \in B$ es el ínfimo de B , si:

1. Para cada $a \in B$, $\beta_0 \leq a$.
2. Si $\beta \in A$ y para cada $a \in B$, $\beta \leq a$. Entonces, $\beta \leq \beta_0$.

Como en el caso de los supremos, cuando un conjunto tiene ínfimo, éste es único.

Definición 23

1. Un conjunto A ordenado parcialmente es una retícula completa superiormente si cada subconjunto de A tiene supremo.
2. Es una retícula completa inferiormente si cada subconjunto de A tiene ínfimo.
3. Es una retícula completa si lo es superior e inferiormente.

Ejemplo 64 . *Si A es un conjunto, la familia de subconjuntos de A , $\wp(A)$ es una retícula completa con la contención. ya que si $\mathcal{S} \subseteq \wp(A)$, entonces $\bigvee \mathcal{S} = \bigcup \mathcal{S}$ y $\bigwedge \mathcal{S} = \bigcap \mathcal{S}$.*

Ejemplo 65 . Si A es un conjunto infinito, la familia de los subconjuntos finitos de A , ordenada parcialmente con la contención entre conjuntos,

$$\wp_f A = \{B \subseteq A \mid B \text{ es finito}\},$$

es una retícula pero no es completa, ya que $\wp_f A$ no tiene supremo.

Ejemplo 66 . La familia de subconjuntos infinitos de A $\{B \subseteq A \mid B \text{ es infinito}\}$ es una retícula superior completa pero no es una retícula inferior.

2.3 Orden en un producto de conjuntos ordenados

Cuando A y B son dos conjuntos ordenados parcialmente, en el producto cartesiano de los conjuntos, $A \times B$, se pueden definir de manera natural dos órdenes parciales. Con el fin de simplificar la notación, mientras no haya lugar a confusión, usaremos el mismo símbolo, \leq , para todos los conjuntos ordenados parcialmente con los que estemos tratando.

Para A y B dos conjuntos ordenados parcialmente se tiene un orden parcial, \leq en $A \times B$, definido de la manera siguiente:

Definición 24 . Sean $(a, b), (\alpha, \beta) \in A \times B$. Definimos $(a, b) \leq (\alpha, \beta)$ si $a \leq \alpha$ y $b \leq \beta$

Nótese que al decir $a \leq \alpha$ nos estamos refiriendo al orden en A , con $b \leq \beta$ al orden en B y al escribir $(a, b) \leq (\alpha, \beta)$, \leq denota el orden que estamos definiendo en $A \times B$.

El otro orden al que nos referimos lo denotaremos \leq_l , y está definido por:

$$(a, b) \leq_l (\alpha, \beta) \text{ si } a \leq \alpha \text{ ó } (a = \alpha \text{ y } b \leq \beta).$$

El orden anterior se llama lexicográfico, o alfabético, en analogía con la situación de que en un diccionario va antes “casa” que “masa” y a que va primero “masa” que “mata”.

1. Decimos que dos elementos a, b en un conjunto ordenado parcialmente A , son comparables, si:

$$a \leq b \text{ ó } b \leq a.$$

2. Dos elementos que no son comparables se dice que son incomparables o incompatibles.

Definición 25 . *Un subconjunto C de A es una cadena, si cualesquiera dos elementos de C son comparables.*

Ejemplo 67 . *En el conjunto de los Números Naturales ordenados con la relación “divide a”, el conjunto de las potencias de 2, $\{1, 2, 4, 8, 16, \dots\}$, es una cadena. En general, si p es un número primo, el conjunto de potencias de p ,*

$$\{1, p^2, p^3, \dots\},$$

es una cadena con la relación “divide a”.

Definición 26 . *Un subconjunto B de A se llama una anticadena, si cualesquiera dos elementos de B son incomparables.*

Ejemplo 68 . *En el conjunto ordenado de los naturales con la relación “divide a” del ejemplo anterior, el conjunto de los números primos, $\{2, 3, 5, 7, 11, \dots\}$ es una anticadena..*

Ejemplo 69 . *Si A es un conjunto finito con n elementos, en la familia, $\wp(A)$, de los subconjuntos de A ordenados parcialmente con la inclusión, si m es un entero y $2 \leq m \leq n - 1$, la familia de los subconjuntos de A con m elementos,*

$$\{B \subseteq A \mid B \text{ tiene } m \text{ elementos}\},$$

es una anticadena.

Un conjunto ordenado parcialmente en el que cualesquiera dos de sus elementos son comparables, es decir una cadena, se llama también un conjunto ordenado linealmente, u ordenado totalmente.

Al orden parcial correspondiente se le llama, en este caso, un orden lineal o total.

Observación 24 . *En general un conjunto A con una relación \leq es un conjunto ordenado totalmente por \leq si dicha relación es:*

1. Total, es decir, si $x, y \in A$, entonces $x \leq y$ ó $y \leq x$.
2. Antisimétrica, o sea, si $x, y \in A$, $x \leq y$ y $y \leq x$, entonces $x = y$.

3. Transitiva, es decir, si $x, y, z \in A$, $x \leq y$ y $y \leq z$, entonces $x \leq z$.

Ejercicio 105 . Si A y B son dos conjuntos ordenados totalmente, $A \times B$ está ordenado totalmente por \leq_l .

Observación 25 . Note que el hecho de que la relación sea total implica que es reflexiva, y que una relación R en un conjunto A es total si y sólo si $R \cup R^{-1} = A \times A$.

Recordemos las siguientes definiciones.

Una relación R en un conjunto A es reflexiva si y sólo si $D(A) \subseteq R$. (Ver 1, página 82).

Una relación R en un conjunto A es antisimétrica si y sólo si $R \cap R^{-1} \subseteq D(A)$.(ver 4, página 83).

Una caracterización de la transitividad, análoga a las anteriores la podemos enunciar en términos de la composición de relaciones, concepto que además es muy útil en varios temas de la Matemática, en particular en el estudio de las funciones, que como veremos son casos particulares de relaciones.

Definición 27 . Si A , B y C son conjuntos, R es una relación de A en B y S es una relación de B en C , la composición de R y S es la relación

$$S \circ R = \{(a, c) \mid \text{para alguna } b \in B, (a, b) \in R \text{ y } (b, c) \in S\}.$$

Para referirnos a $S \circ R$ diremos “erre compuesta con ese” o “erre seguida de ese”.

Observación 26 . Una relación R es transitiva si y sólo si $R \circ R \subseteq R$.

Anteriormente mencionamos que la restricción de un orden parcial, es a su vez un orden parcial, y al analizar esto podemos hacernos la pregunta siguiente:

¿Las propiedades de una relación, las hereda una restricción de la relación?

Proposición 8 . La restricción de una relación reflexiva es reflexiva.

Demostración. Si R es una relación reflexiva en un conjunto A y B es un subconjunto de A , entonces, para cada $b \in B$, se tiene que $(b, b) \in R$ pues R es reflexiva y $b \in A$.

Por lo tanto $(b, b) \in R \cap (B \times B) = R|_B$. ■

Proposición 9 . *La restricción de una relación antisimétrica es antisimétrica.*

Demostración. En efecto, sea A un conjunto, R una relación antisimétrica en A y $B \subseteq A$.

Si $(a, b) \in (R|_B)$ y $(b, a) \in (R|_B)$, entonces $(a, b) \in R$ y $(b, a) \in R$, por lo que $a = b$.

Otra manera de verlo es la siguiente:

$$(R|_B) = R \cap (B \times B), \text{ y } (R|_B)^{-1} = R^{-1} \cap (B \times B), \text{ así que}$$

$$\begin{aligned} (R|_B) \cap (R|_B)^{-1} &\subseteq (R \cap R^{-1}) \cap (B \times B) \subseteq \\ &\subseteq D(A) \cap (B \times B) = D(B). \end{aligned}$$

■

De manera análoga a la anterior se puede demostrar que:

Proposición 10 . *La restricción de una relación transitiva es transitiva y la restricción de una relación total es total.*

Como consecuencia de lo anterior, la restricción de un orden total es un orden total, y por lo tanto cada subconjunto de un conjunto ordenado totalmente lo consideraremos de manera natural como un conjunto ordenado totalmente.

Definición 28 . *Recordemos que si A y B son conjuntos y R es una relación de A en B , entonces el conjunto de los elementos de A que están relacionados con algún elemento de B se llama el dominio de la relación R . Así, el dominio de R es:*

$$Dom(R) = \{a \in A \mid \text{para alguna } b \in B, (a, b) \in R\}.$$

Contra lo que podría esperarse, se le llama el contradominio de R al conjunto B .

En cambio al conjunto de los elementos de B que están relacionados con algún elemento de A se le llama la imagen de la relación R . Así, la imagen de R , es:

$$Im(R) = \{b \in B \mid \text{para alguna } a \in A, (a, b) \in R\}.$$

Al lector le puede parecer ocioso hablar de una relación de A en B para referirnos a un conjunto de parejas ordenadas, sólo porque cada una éstas tiene un elemento en A y otro en B , si lo mismo podemos decir del dominio y la imagen de la relación, (cada una de las parejas en la relación tiene un elemento en el dominio y otro en la imagen) o de cualesquiera dos conjuntos, si uno contiene al dominio y el otro a la imagen. Sin embargo, cuando dos conjuntos tienen estructuras matemáticas similares, algunas relaciones especiales de uno en el otro, permiten comparar dichas estructuras y encontrar propiedades de una de ellas conociendo algunas de la otra.

2.4 Funciones

Possiblemente las relaciones más importantes en la Matemática sean las funciones, ellas permiten precisar muchos conceptos, relacionar, como dijimos antes, unas estructuras con otras y describir tanto hechos de la Matemática misma, como una gran cantidad de fenómenos de la naturaleza. Los casos más sencillos y conocidos de funciones son las fórmulas, para las cuales, cuando se determina el valor de los datos se obtiene un único resultado, por ejemplo en la fórmula del área de un triángulo, base por altura sobre dos, los valores de la base y de la altura determinan el valor del área. Diremos que una relación es unívoca cuando “el primer elemento de una pareja determina al segundo”, es decir cuando ningún objeto es el primer elemento de dos parejas distintas en la relación en cuestión, de manera más precisa:

Definición 29 . *Una relación R es unívoca si*

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

Definición 30 . *Si A y B son dos conjuntos, y f es una relación unívoca de A en B , tal que el dominio de f es A entonces la terna ordenada (A, f, B) es una función.*

Es decir, $(A, f \subseteq A \times B, B)$ es una función de A en B , si:

para cada $a \in A$ hay un único elemento $b \in B$ tal que $(a, b) \in f$.

Si al elemento b del renglón anterior lo denotamos $f(a)$ podemos entender a las funciones de la manera siguiente:

Una función f de un conjunto A en un conjunto B es una regla que determina para cada elemento a de A , un elemento único $f(a)$ de B .

En la literatura muchas veces se dice lo anterior así: una función es una regla que asocia a cada elemento del dominio uno y sólo un elemento del contradominio. Al elemento $f(a)$ se le llama el valor de la función f en a o la imagen de a bajo f , y en el lenguaje verbal se refiere uno a él como “ f de a ”. Con las convenciones anteriores, si (A, f, B) es una función:

$$f = \{(x, f(x)) \mid x \in A\}.$$

Por costumbre, a veces en lugar de escribir (A, f, B) , escribimos simplemente f para describir la función.

Notación 1 . Se suele escribir $a \xrightarrow{f} b$ para indicar que $b = f(a)$.

Observación 27 . En vista de lo anterior, si A y B son dos conjuntos y f y g son dos funciones de A en B , $f = g$, si y sólo si

$$\{(x, f(x)) \mid x \in A\} = \{(x, g(x)) \mid x \in A\},$$

es decir, si y sólo si para cada $x \in A$, $f(x) = g(x)$.

Así pues dos funciones son iguales si y sólo si tienen el mismo dominio, el mismo contradominio, y las reglas de ambas son iguales en todos los elementos del dominio común.

Otra manera de decir lo anterior es la siguiente: dos funciones son iguales si tienen el mismo dominio, el mismo contradominio, y los mismos valores en todos los elementos de su dominio.

Para referirnos a una función f de un conjunto A en un conjunto B , escribiremos simplemente:

$$f : A \rightarrow B$$

o bien,

$$A \xrightarrow{f} B$$

Ejemplo 70 . Si A es un conjunto, la igualdad en A es una función de A en A a la que suele llamársele la identidad en A y denotársele 1_A , con el lenguaje que acabamos de convenir:

$$1_A : A \rightarrow A$$

es la función tal que para cada $a \in A$, $1_A(a) = a$.

Ejemplo 71 . Cuando un conjunto B es subconjunto de un conjunto A , se tiene de manera natural una función de B en A , llamada la inclusión de B en A y denotada i cuya regla de correspondencia es:

$$\text{Para cada } b \in B, i(b) = b$$

Nótese que la identidad es un caso particular de la inclusión.

2.4.1 Funciones inyectivas

Las inclusiones forman parte de una familia más grande de funciones, la de las funciones inyectivas. Una función es inyectiva si los valores de la función son distintos para elementos distintos de su dominio. Otra forma de decir lo anterior es:

Definición 31 . Una función $f : A \rightarrow B$ es inyectiva si para cualesquiera dos elementos a, b en A ,

$$(f(a) = f(b)) \Rightarrow (a = b).$$

El hecho de que una función $f : A \rightarrow B$ sea inyectiva lo expresaremos muchas veces usando la flecha \rightarrowtail , es decir escribiremos $f : A \rightarrowtail B$.

Desde el punto de vista conjuntista cada función inyectiva “copia el dominio en el contradominio”, ya que si $f : A \rightarrowtail B$ es una función inyectiva, cada elemento $a \in A$ tiene una copia $f(a)$ en B y así la imagen de f , $f(A) = \{f(x) \mid x \in A\}$ es una copia de A donde “a cada uno de sus elementos, si éste se llamaba a ahora se le ha cambiado el nombre por el de por el de $f(a)$ ”.

Con las convenciones anteriores, recordemos que si A es un conjunto ordenado parcialmente hemos definido para cada $a \in A$: $\langle a \rangle = \{x \in A \mid x \leq a\}$, entonces mediante la función $s : A \rightarrowtail \wp(A)$ tal que para cada $a \in A$, $s(a) = \langle a \rangle$, obtenemos en $\wp(A)$ una copia, $\{\langle a \rangle \mid a \in A\}$, de A .

Más aún, esta copia no lo es sólo desde el punto de vista conjuntista, ya que como

$$a \leq b \text{ si y sólo si } \langle a \rangle \subseteq \langle b \rangle,$$

entonces $\{\langle a \rangle \mid a \in A\}$ es una copia de A como conjunto ordenado.

Observación 28 . Si $f : A \rightarrowtail B$ y $g : B \rightarrow C$ son funciones, f y g como relaciones que son, se pueden componer, y su composición $g \circ f$ es, como ya sabemos, una relación de A en C , más aún, $(A, g \circ f, C)$ es una función ya que:

1. Si $a \in A$, como

$$A = \text{Dom}(f) = \{x \in A \mid \exists y \in B, (x, y) \in f\},$$

hay una $b \in B$, para la cual $(a, b) \in f$. Como

$$b \in B = \text{Dom}(g) = \{y \in B \mid \exists z \in C, (y, z) \in g\},$$

hay una $c \in C$, tal que $(b, c) \in g$.

De aquí, podemos concluir que $(a, c) \in g \circ f$ y por lo tanto $a \in \text{Dom}(g \circ f)$.

2. Con lo anterior hemos demostrado que $\text{Dom}(g \circ f) = A$ y puesto que $g \circ f$ es una relación de A en B , sólo falta probar que $g \circ f$ es unívoca. si $(a, c_1) \in g \circ f$ y

$$(a, c_2) \in g \circ f = \{(x, z) \in A \times C \mid \exists y \in B, (x, y) \in f \text{ y } (y, z) \in g\},$$

entonces hay una $b_1 \in B$ para la cual $(a, b_1) \in f$ y $(b_1, c_1) \in g$, y también hay una $b_2 \in B$, tal que $(a, b_2) \in f$ y $(b_2, c_2) \in g$. Ahora el hecho de que f sea unívoca obliga a que $b_1 = b_2$ y de aquí necesariamente $c_1 = c_2$ ya que g es unívoca. De donde concluimos que $g \circ f$ es unívoca. ■

Si ya aceptamos que una función f de un conjunto A en un conjunto B es una regla que para cada elemento $a \in A$ determina un único elemento $f(a) \in B$, con este lenguaje podemos rehacer la demostración anterior de la manera siguiente:

Si $f : A \rightarrow B$ y $g : B \rightarrow C$ son funciones, $g \circ f$ es una función de A en C . En efecto si $a \in A$, f determina un único elemento $f(a) \in B$, el cual a su vez determina un único elemento $g(f(a)) \in C$.

Observación 29 . *Nótese que de acuerdo a las convenciones que hemos hecho acerca de la notación y la terminología para las funciones, para cada $a \in A$ el único elemento de C que la composición $(g \circ f)$ determina es precisamente $g(f(a))$; es decir $(g \circ f)(a) = g(f(a))$.*

Otra manera de decir lo anterior es:

Si $f : A \rightarrow B$ y $g : B \rightarrow C$ son funciones, entonces $g \circ f : A \rightarrow C$ es la función cuya regla de correspondencia es,

$$\text{para cada } a \in A, (g \circ f)(a) = g(f(a)).$$

Las funciones inyectivas tienen respecto a la composición (de funciones) una propiedad similar a la que tienen los números distintos de cero respecto a la multiplicación, en el siguiente sentido:

Proposición 11 . *Si $f : A \rightarrow B$, $g : A \rightarrow B$ y $h : B \rightarrow \mathbf{C}$ son funciones y $h \circ f = h \circ g$, entonces $f = g$.*

Demostración. Como dos funciones son iguales si y sólo si tienen el mismo dominio, el mismo contradominio y la misma regla de correspondencia, para comprobar que f y g son iguales basta demostrar que tienen la misma regla de correspondencia.

Pero para cada $a \in A$, tenemos que

$$h(f(a)) = (h \circ f)(a) = (h \circ g)(a) = h(g(a)),$$

y del hecho de que h es una función inyectiva, se concluye que

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

así que $f = g$. ■

Lo anterior a veces se enuncia diciendo que las funciones inyectivas se pueden cancelar por la izquierda. En cambio no es cierto que las funciones inyectivas se puedan cancelar por la derecha como se ve en ejemplo siguiente.

Ejemplo 72 . *Si A y B son conjuntos $A \neq \emptyset$, $A \neq B$, y $A \subseteq B$ (lo que expresaremos también: $\emptyset \neq A \subsetneq B$), tomemos*

$$i : A \hookrightarrow B,$$

la inclusión de A en B y escojamos $a \in A$, para definir

$$\begin{array}{rccc} f : & B & \rightarrow & B \\ & A \ni b & \longmapsto & b \\ & A \not\ni b & \longmapsto & a, \end{array}$$

entonces,

$$1_B \circ i = i = f \circ i \text{ aunque } f \neq 1_B.$$

Hemos visto que las funciones inyectivas se cancelan por el lado izquierdo. Ahora veremos que el recíproco también es cierto:

Teorema 5 . *Si $f : A \rightarrow B$ es cancelable por la izquierda, entonces f es inyectiva.*

Demostración. Por contrapuesta, si f no es inyectiva, es porque existen $a \neq c \in A$ tales que $f(a) = f(c)$. Tomemos un conjunto con un único elemento, por ejemplo $\{\emptyset\}$.

Definamos $h : \{\emptyset\} \rightarrow A$, por $h(\emptyset) = a$ y definamos $k : \{\emptyset\} \rightarrow A$, por: $k(\emptyset) = c$.

Las funciones h y k son distintas pues $a \neq c$, pero $f \circ h : \{\emptyset\} \rightarrow B$, es tal que $(f \circ h)(\emptyset) = f(h(\emptyset)) = f(a)$

mientras que $f \circ k : \{\emptyset\} \rightarrow B$, es tal que $(f \circ h)(\emptyset) = f(k(\emptyset)) = f(c)$.

Entonces $\{\emptyset\} \xrightarrow{f \circ h} B = \{\emptyset\} \xrightarrow{f \circ k} B$, pero $h \neq k$, es decir, f no es cancelable por la izquierda. ■

Hemos demostrado que una función $f : A \rightarrow B$ es inyectiva si y sólo si es cancelable por la izquierda. ¿A qué corresponderá el hecho de que $f : A \rightarrow B$ sea cancelable por la derecha?

2.4.2 Funciones suprayectivas

Definición 32 . *Una función $f : A \rightarrow B$ es suprayectiva si para cada $b \in B$, existe $a \in A$ tal que $b = f(a)$.*

Otra manera de decir lo mismo es pidiendo que la imagen de f coincida con su contradominio ($f(A) = B$).

Teorema 6 . *Una función suprayectiva es cancelable por la derecha.*

Demostración. Supongamos que $f : A \rightarrow B$ es suprayectiva, y que

$$g \circ f = h \circ f,$$

con $g : B \rightarrow C$ y $h : B \rightarrow C$. Queremos demostrar que $B \xrightarrow{g} C = B \xrightarrow{h} C$.

Si $B = \emptyset$, entonces $g \subseteq \emptyset \times C = \emptyset$, por lo que $g = \emptyset$. Análogamente, $h = \emptyset$.

Supongamos ahora que $B \neq \emptyset$.

Tomemos $b \in B$, como f es suprayectiva, existe $a \in A$ tal que $f(a) = b$. Entonces

$$g(b) = g(f(a)) = (g \circ f)(a) \underbrace{=}_{g \circ f = h \circ f} (h \circ f)(a) = h(f(a)) = h(b).$$

Como $g(b) = h(b)$, $\forall b \in B$, entonces $g = h$. ■

Veamos ahora que una función cancelable por la derecha es suprayectiva.

Teorema 7 . *Si $f : A \rightarrow B$ es cancelable por la derecha, entonces f es suprayectiva.*

Demostración. Por contrapuesta, si $f : A \rightarrow B$ no es suprayectiva, definamos las siguientes funciones de B a $\{0, 1\} : B \xrightarrow{\hat{0}} \{0, 1\}$ tal que $b \mapsto 0$, $\forall b \in B$, y $B \xrightarrow{g} \{0, 1\}$ por $b \mapsto 0$, si $b \in f(A)$, $b \mapsto 1$, si $b \notin f(A)$.

Es claro que

$$g \circ f = \hat{0} \circ f$$

pero $g \neq \hat{0}$, ya que para $b \in B \setminus f(A)$, $g(b) = 1 \neq 0 = \hat{0}(b)$. ■

Definición 33 . *Sean $A \xrightarrow{f} B$ y $B \xrightarrow{g} A$ dos funciones supongamos que $A \xrightarrow{g \circ f} A = A \xrightarrow{Id_A} A$, diremos que g es inverso por la izquierda de f y que f es inverso por la derecha de g .*

Teorema 8 . *Son equivalentes para $\emptyset \neq A \xrightarrow{f} B$:*

1. f es inyectiva,
2. f tiene inverso por la izquierda.

Demostración. 1 \Rightarrow 2) Consideremos $B = f(A) \cup (B \setminus f(A))$.

Escojamos un elemento $x \in A$ ($A \neq \emptyset$), definimos

$$B = f(A) \cup (B \setminus f(A)) \xrightarrow{g} A$$

$$f(a) \mapsto a, b \mapsto x \text{ si } b \notin f(A).$$

Entonces $g \circ f(a) = g(f(a)) = a = Id_A(a)$, $\forall a \in A$, por lo que $B \xrightarrow{g} A$ es inverso izquierdo para f .

2) \Rightarrow 1) Supongamos que g es inverso por la izquierda de f , entonces dados $a \neq c \in A$, tenemos que

$$g(f(a)) = Id_A(a) = a \neq c = Id_A(c) = g(f(c)),$$

entonces $f(a) \neq f(c)$. ■

Análogamente una función que tiene inverso por la derecha es suprayectiva, porque si tiene inverso por la derecha es cancelable por la derecha.

Ejercicio 106 . Demuestre que si $f : A \rightarrow B$ es una función con inverso por la derecha, entonces es cancelable por la derecha.

Axioma 13 . Una función suprayectiva $f : A \rightarrow B$ tiene inverso por la derecha.

El axioma anterior se conoce como “Axioma de Elección”.

Observación 30 . Si $\{A_x\}_{x \in I}$ es una familia no vacía ($I \neq \emptyset$) de conjuntos no vacíos ($A_x \neq \emptyset, \forall x \in I$) ajenos dos a dos ($I_x \cap I_y = \emptyset, x \neq y \in I$) entonces, considerando la función suprayectiva

$$\cup \{A_x\}_{x \in I} \xrightarrow{f} I,$$

$$b \mapsto y \text{ si } b \in A_y,$$

el axioma anterior nos asegura la existencia de un inverso derecho para f . Notemos que si $g : I \rightarrow \cup \{A_x\}_{x \in I}$ es un inverso derecho para f , entonces $f(g(z)) = z$, así que $g(z)$ es un elemento de A_z . Puede decirse que g es una función que “escoge” un elemento de cada A_x , por eso, g se llama una función de elección”.

Teorema 9 . La composición de funciones es asociativa.

Demostración. Sean $A \xrightarrow{f} B$, $B \xrightarrow{g} C$ y $C \xrightarrow{h} D$ entonces

$$A \xrightarrow{(h \circ g) \circ f} D = A \xrightarrow{h \circ (g \circ f)} D,$$

pues para cada $a \in A$,

$$((h \circ g) \circ f)(a) = (h \circ g)(f(a)) = h(g(f(a))),$$

pero también

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

■

Teorema 10 . Si $A \xrightarrow{f} B$ y $B \xrightarrow{g} C$ son funciones inyectivas entonces

$$A \xrightarrow{g \circ f} C$$

también es inyectiva.

Demostración. Sean $B \xrightarrow{h} A$ y $C \xrightarrow{k} B$ inversos izquierdos de f y de g respectivamente. Demostraremos que $C \xrightarrow{h \circ k} A$ es inverso por la izquierda de $A \xrightarrow{g \circ f} C$:

$$\begin{aligned} (h \circ k) \circ (g \circ f) &= h \circ (k \circ (g \circ f)) = \\ &= h \circ ((k \circ g) \circ f) = h \circ (Id_B \circ f) = \\ &= h \circ f = Id_A. \end{aligned}$$

■

Ejercicio 107 . Demuestre que la composición de funciones inyectivas es inyectivas

1. Usando directamente la definición de inyectividad.
2. Mostrando que la composición de las funciones inyectivas es cancelable por la izquierda.

Teorema 11 . Si $A \xrightarrow{f} B$ y $B \xrightarrow{g} C$ son funciones suprayectivas entonces

$$A \xrightarrow{g \circ f} C$$

también es suprayectivas.

Demostración. Sea $c \in C$, como g es suprayectiva, existe $b \in B$ tal que $c = g(b)$. Como f es suprayectiva, existe $a \in A$ tal que $b = f(a)$.

Por lo tanto $(g \circ f)(a) = g(f(a)) = g(b) = c$.

Por lo tanto $C = (g \circ f)(A)$. ■

Ejercicio 108 . Demuestre que la composición de funciones suprayectivas es suprayectiva.

1. Mostrando un inverso derecho para la composición
2. Mostrando que la composición de las funciones suprayectivas es cancelable por la derecha.

2.4.3 Funciones biyectivas

Definición 34 . Una función $A \xrightarrow{f} B$ es biyectiva cuando es inyectiva y suprayectiva.

Definición 35 . Una función $A \xrightarrow{f} B$ es invertible cuando existe $B \xrightarrow{g} A$ que es inverso izquierdo y derecho de f , en este caso decimos que g es inverso de f .

Definición 36 . Una función $A \xrightarrow{f} B$ es cancelable cuando es cancelable por la izquierda y por la derecha.

Teorema 12 . Son equivalentes para $A \xrightarrow{f} B$:

1. f es biyectiva.
2. f es invertible.
3. f es cancelable.
4. f tiene inverso por la izquierda e inverso por la derecha.

Demostración. Como ya sabemos que son equivalentes para una función f ser inyectiva, cancelable por la izquierda y tener inverso por la izquierda y también que son propiedades equivalentes para una función ser suprayectiva, ser cancelable por la derecha y tener inverso por la derecha, tenemos que 1), 2) y 3) son equivalentes.

Es claro que 2) \Rightarrow 4).

4) \Rightarrow 2) Supongamos que $B \xrightarrow[k]{h} A$ son tales que h es inverso izquierdo de $A \xrightarrow{f} B$ y que k es inverso derecho de f , entonces

$$\begin{aligned} h \circ f &= Id_A \\ f \circ k &= Id_B, \end{aligned}$$

Entonces

$$k = Id_A \circ k = (h \circ f) \circ k = h \circ (f \circ k) = h \circ Id_B = h.$$

Lo que muestra que $k = h$ es un inverso de f . ■

Ejercicio 109 . Demuestre que si $A \xrightarrow{f} B$ y $B \xrightarrow{g} C$ son funciones biyectivas, entonces $A \xrightarrow{g \circ f} C$ es biyectiva.

Teorema 13 . Son equivalentes para $A \xrightarrow{f} B$:

1. f es invertible.
2. f tiene un único inverso izquierdo.
3. f tiene un único inverso derecho.

Demostración. 1) \Rightarrow 2). Si $h \circ f = Id_A = k \circ f$ entonces $h = k$ pues las funciones invertibles son cancelables.

1) \Rightarrow 3). Es análogo a 1) \Rightarrow 2).

Ahora basta demostrar la equivalencia entre 2) y 3) :

2) \Rightarrow 3). Supongamos que

$$f \circ g = Id_B = f \circ h,$$

digamos que $k : B \rightarrow A$ es el inverso izquierdo de f . Entonces

$$\begin{aligned} g &= Id_A \circ g = (k \circ f) \circ g = \\ &= k \circ (f \circ g) = k \circ Id_B = k \circ (f \circ h) = \\ &= (k \circ f) \circ h = Id_A \circ h = h, \end{aligned}$$

por lo que el inverso derecho es único.

3) \Rightarrow 2). Es análogo a 2) \Rightarrow 3).

3) \Rightarrow 1). Suponiendo 3), tenemos que f tiene inverso derecho. Por 3) \Rightarrow 2), tenemos que f también tiene inverso izquierdo. ■

Corolario 2 . La inversa de una función $A \xrightarrow{f} B$, cuando existe, es única y se denota $B \xrightarrow{f^{-1}} A$.

Además, como $f \circ f^{-1} = Id_B$, y $f^{-1} \circ f = Id_A$, es claro que $(f^{-1})^{-1} = f$, así que f^{-1} también es invertible.

Teorema 14 . Sean $A \xrightarrow{f} B$, $B \xrightarrow{g} C$ dos funciones. Entonces

1. $A \xrightarrow{g \circ f} C$ suprayectiva $\Rightarrow g$ es suprayectiva.

2. $A \xrightarrow{g \circ f} C$ inyectiva $\Rightarrow f$ es inyectiva.

Demuestración. 1. Si $A \xrightarrow{g \circ f} C$ es suprayectiva, entonces tiene un inverso derecho $h : C \rightarrow A$. Como

$$Id_C = (g \circ f) \circ h = g \circ (f \circ h),$$

vemos que $f \circ h$ es un inverso derecho de g . Por lo tanto g es suprayectiva.

2. Si $A \xrightarrow{g \circ f} C$ es inyectiva, entonces $g \circ f$ es cancelable por la izquierda. Ahora,

$$\begin{aligned} (f \circ h = f \circ k) &\Rightarrow (g \circ f \circ h = g \circ f \circ k), \\ ((g \circ f) \circ h = (g \circ f) \circ k) &\Rightarrow (h = k). \end{aligned}$$

Así que también f es cancelable por la izquierda, por lo que f es inyectiva.

■

2.5 Cardinalidad

2.5.1 Axioma del infinito

Observemos que si A es un conjunto, el axioma de las parejas nos asegura que hay un conjunto $\{A, A\} = \{A\}$.

Ahora, el axioma de regularidad nos permite asegurar que $A \neq \{A\}$ (pues $A \in \{A\}$ pero $A \notin A$).

Nuevamente por el axioma de las parejas tenemos que hay un conjunto $\{A, \{A\}\}$ y el axioma de la unión nos dice que

$$\cup \{A, \{A\}\} = A \cup \{A\},$$

es un conjunto. Notemos que $A \cup \{A\}$ tiene exactamente un elemento más que A (a saber: A).

El conjunto $A \cup \{A\}$ será importante para nosotros y lo denotaremos $\sigma(A)$.

Los axiomas de la Teoría de conjuntos nos permiten afirmar la existencia de los siguientes conjuntos:

$$\emptyset, \sigma(\emptyset) = \emptyset \cup \{\emptyset\} = \{\emptyset\}, \sigma(\sigma(\emptyset)) = \{\emptyset, \{\emptyset\}\}, \sigma(\sigma(\sigma(\emptyset))), \dots$$

en la lista anterior, cada conjunto es elemento del conjunto siguiente, pero por el axioma de regularidad, ningún conjunto de la lista pertenece a

un conjunto precedente. Esto nos permite afirmar que cada conjunto de la lista es distinto de los que le preceden.

Nos preguntamos si la lista interminable anterior serán los elementos de un conjunto. Como hasta el momento no hay nada que nos lo asegure introducimos la siguiente definición.

Definición 37 . *Un conjunto A es inductivo si:*

1. $\emptyset \in A$.
2. $c \in A \Rightarrow \sigma(c) \in A$.

Axioma 14 (del infinito) . *Existe un conjunto inductivo.*

La razón de que se llame así al axioma, es porque un conjunto inductivo tiene que contener todos los elementos de la lista interminable de elementos distintos:

$$\emptyset, \sigma(\emptyset), \sigma(\sigma(\emptyset)), \sigma(\sigma(\sigma(\emptyset))), \dots$$

Definición 38 . *Decimos que A y B tienen la misma cardinalidad si existe $A \xrightarrow{f} B$ biyectiva. Expresaremos este hecho escribiendo $|A| = |B|$. También diremos que A y B tienen el mismo número de elementos.*

Escribiremos $|A| \leq |B|$ si existe $A \xrightarrow{f} B$ inyectiva.

Observemos que esto es equivalente a que haya $B \xrightarrow{g} A$, función suprayectiva.

Es un teorema muy importante de la Teoría de Conjuntos (Teorema de Cantor-Bernstein-Schröeder) que si $|A| \leq |B|$ y $|B| \leq |A|$ entonces $|A| = |B|$.

Aunque la notación parece sugerir el teorema anterior, el resultado no es nada obvio. Se afirma que si hay dos funciones f, g de A a B , una inyectiva y la otra suprayectiva, entonces debe haber una función biyectiva de A a B .

Notación 2 . *Escribiremos $|A| < |B|$ para decir que $|A| \leq |B|$ pero $|A| \neq |B|$. Esto sucede cuando hay una función inyectiva de A a B pero no hay ninguna función suprayectiva de A a B . (Usando el teorema de Cantor-Bernstein-Schröeder).*

Ejemplo $|\emptyset| < |\{\emptyset\}|$.

Teorema 15 . Para cualquier conjunto A , se tiene que $|A| < |\wp(A)|$.

Demostración. La función

$$A \rightarrow \wp(A)$$

$$a \mapsto \{a\}$$

es inyectiva, pues si $a \neq b$ entonces $\{a\} \neq \{b\}$.

Entonces $|A| \leq |\wp(A)|$, pero no puede haber una función suprayectiva $g : A \rightarrow \wp(A)$.

Si $g : A \rightarrow \wp(A)$ fuera una función suprayectiva, podríamos considerar el conjunto

$$B = \{a \in A \mid a \notin g(a)\} \in \wp(A).$$

Bajo la hipótesis de que g es suprayectiva, existiría un elemento $c \in A$ tal que $B = g(c)$.

Se tiene que $c \in B$ ó $c \notin B$. ¿Cuál es el caso?

Si $c \in B$, es porque $c \notin g(c) = B$.

Por lo tanto $c \notin B$.

Pero $c \notin B \Rightarrow \neg(c \notin g(c))$. Es decir $c \in B$. \circlearrowleft

La contradicción anterior muestra que no puede haber una suprayectión de A a $\wp(A)$.

Por lo que $|A| < |\wp(A)|$. ■

2.5.2 Conjuntos infinitos

Definición 39

1. Se dice que el conjunto A es infinito si existe un subconjunto propio B de A tal que $|A| = |B|$.
2. Un conjunto que no es infinito se llama finito.

Ejemplo 73 . \emptyset es finito, ya que \emptyset no tiene subconjuntos propios.

Ejemplo 74 . Un conjunto inductivo es infinito:

Supongamos que N es inductivo y consideremos la función

$$N \xrightarrow{\sigma} N.$$

$$a \mapsto \sigma(a) = a \cup \{a\}.$$

Observemos primero que la función σ es inyectiva: pues si

$$a \cup \{a\} = b \cup \{b\},$$

entonces $a \in b \cup \{b\}$ y $b \in a \cup \{a\}$.

$a \in b$, $b \in a$ es contradictorio con el axioma de regularidad (... $b \in a \in b \in a \in b \in a$ continúa indefinidamente).

$a \in b$, $b = a$ contradice el axioma de regularidad.

Por lo tanto $a \notin b$, por lo que $a = b$.

Entonces

$$A \xrightarrow{\sigma|_{\sigma(A)}} \sigma(A)$$

es una biyección entre A y $\sigma(A)$. Además, $\sigma(A)$ es un subconjunto propio de A : $\emptyset \in A \setminus \sigma(A)$ ($\emptyset \in A$ porque A es inductivo y $\emptyset \neq \sigma(a)$ porque $\sigma(a)$ tiene como elemento a a).

Observación 31. Si A es infinito y $|A| = |B|$, entonces B es infinito.

Demostración. Supongamos que $C \subsetneq A$ y que $C \xrightarrow{f} A$ y $A \xrightarrow{g} B$ son biyecciones. Consideremos

$$\begin{array}{ccc} C & \xrightarrow{f} & A \\ \downarrow g|_C^{g(C)} & & \downarrow g \\ g(C) & & B \end{array}$$

$C \subsetneq A \Rightarrow g(C) \subsetneq B$, pues si $x \in A \setminus C$ entonces $g(x) \in B \setminus g(C)$.

Ahora, son biyecciones $B \xrightarrow{g^{-1}} A$, $A \xrightarrow{f^{-1}} C$ y $C \xrightarrow{g|_C^{g(C)}} g(C)$. Entonces

$$B \xrightarrow{g|_C^{g(C)} \circ f^{-1} \circ g^{-1}} g(C)$$

es una biyección entre B y uno de sus subconjuntos propios. Por lo tanto B es infinito. ■

Teorema 16. Si A es un conjunto finito entonces $s(A)$ también es un conjunto finito.

Demostración. Supongamos que A es finito.

Si $s(A) \xrightarrow{h} C$ fuera una biyección entre $s(A) = A \cup \{A\}$ y un subconjunto propio de $s(A)$, podemos considerar

$$A \xrightarrow{h|_A} C \setminus \{h(A)\},$$

que es inyectiva y suprayectiva.

i) Si $C \setminus \{h(A)\}$ es un subconjunto de A entonces $C \setminus \{h(A)\} = A$, pues A es finito.

Como

$$A \subseteq C = A \cup \{h(A)\},$$

y como

$$C \subsetneq A \cup \{A\}$$

se sigue que $C = A$ y que $h(A) \in A$. Pero entonces tenemos una biyección entre $s(A)$ y A así que $s(A)$ es finito e infinito ∇ .

ii) Entonces $C \setminus \{h(A)\}$ no es un subconjunto de A . Pero como $C \subseteq A \cup \{A\}$, concluimos que $A \in C$ (en caso contrario $C \subseteq A$) y además $h(A) \neq A$.

Así que

$$A \in C \setminus \{h(A)\},$$

por lo que $A = h(x) \neq h(A)$, con $x \neq A$.

La función

$$C \setminus \{h(A)\} \rightarrow C \setminus \{h(x)\} \text{ tal que}$$

$$c \mapsto c \text{ si } c \neq h(x)$$

$$h(x) \mapsto h(A)$$

es una biyección.

Entonces

$$A \xrightarrow{h|_A} C \setminus \{h(A)\} \rightarrow C \setminus \{h(x)\}$$

es una biyección entre A y $C \setminus \{h(x)\} = C \setminus \{A\} \subseteq A$.

Resta notar que $C \setminus \{A\} \subsetneq A$.

Como $C \subsetneq s(A)$ entonces $\exists y \in s(A) = A \cup \{A\}$ tal que $y \notin C$, por lo que $y \neq A$. Entonces $y \in A$.

Como $y \in A \setminus (C \setminus \{A\})$ entonces

$$(C \setminus \{A\}) \subsetneq A$$

y tenemos una biyección entre A y un subconjunto propio ∇ .

La contradicción viene de suponer que $s(A)$ es infinito. ■

2.6 Imágenes directas e imágenes inversas

Si $A \xrightarrow{f} B$ es una función, podemos definir una función

$$\begin{array}{ccc} \wp(A) & \xrightarrow{f^*} & \wp(B) \\ X & \mapsto & f(X) = \{y \in B \mid y = f(x), x \in X\}. \end{array}$$

Observación 32. Si $A \xrightarrow{f} B$ es una función biyectiva, entonces $\wp(A) \xrightarrow{f^*} \wp(B)$ también es biyectiva.

Demostración. Es inmediato que $\wp(A) \xrightarrow{f^*} \wp(B)$ tiene inversa:

$$\wp(B) \xrightarrow{(f^{-1})^*} \wp(A).$$

Para ver esto, tomemos $X \in \wp(A)$, entonces

$$\begin{aligned} (f^{-1})^*(f^*(X)) &= (f^{-1})^*(\{y \in B \mid y = f(x), x \in X\}) = \\ &= \{x \in A \mid x = f^{-1}(y), y = f(x), x \in X\} = \\ &= \{x \in A \mid x = f^{-1}(f(x)), x \in X\} = X = Id_{\wp(A)}(X). \end{aligned}$$

Por lo que

$$(f^{-1})^* \circ f^* = Id_{\wp(A)},$$

simétricamente

$$f^* \circ (f^{-1})^* = Id_{\wp(B)}.$$

■

Así que

$$|A| = |B| \Rightarrow |\wp(A)| = |\wp(B)|.$$

Teorema 17 . *Sea $A \xrightarrow{f} B$ una función, $X, Y \in \wp(A)$. Entonces $\wp(A) \xrightarrow{f^*} \wp(B)$ satisface:*

1. $X \subseteq Y \Rightarrow f^*(X) \subseteq f^*(Y)$.
2. $f^*(X \cup Y) = f^*(X) \cup f^*(Y)$.
3. $f^*(X \cap Y) \subseteq f^*(X) \cap f^*(Y)$ y la inclusión puede ser propia.

Demostración. 1. Es claro de la definición.

2. Como $X \subseteq X \cup Y$ entonces $f^*(X) \subseteq f^*(X \cup Y)$. Análogamente $f^*(Y) \subseteq f^*(X \cup Y)$ por lo tanto

$$f^*(X) \cup f^*(Y) \subseteq f^*(X \cup Y).$$

Ahora, $z \in f^*(X \cup Y) = \{f(x) \mid x \in X \cup Y\} \Rightarrow z = f(x)$ con $x \in X$ ó $x \in Y$. Por lo tanto $z = f(x) \in f^*(X) \cup f^*(Y)$.

3. $(X \cap Y \subseteq X) \Rightarrow f^*(X \cap Y) \subseteq f^*(X)$, análogamente, $f^*(X \cap Y) \subseteq f^*(Y)$. Por lo tanto

$$f^*(X \cap Y) \subseteq f^*(X) \cap f^*(Y).$$

Mostramos que la inclusión puede ser propia con un ejemplo:

Para

$$\begin{array}{ccc} \{a, b\} & \xrightarrow{f} & \{1\} \\ a & \mapsto & 1 \\ b & \mapsto & 1 \end{array}$$

tenemos que

$$\begin{array}{ccc} \wp(\{a, b\}) & \xrightarrow{f^*} & \wp(\{1\}) \\ \emptyset & \mapsto & \emptyset \\ \{a\} & \mapsto & \{1\} \\ \{b\} & \mapsto & \{1\} \\ \{a, b\} & \mapsto & \{1\} \end{array}$$

Vemos que

$$f^*(\{a\} \cap \{b\}) = f^*(\emptyset) = \emptyset \subsetneq f^*(\{a\}) \cap f^*(\{b\}) = \{1\}.$$

■ Del comportamiento de f^* respecto a uniones e intersecciones arbitrarias podemos decir lo siguiente:

Teorema 18 . *Sea $A \xrightarrow{f} B$ una función y sea $\{I_z\}_{z \in J}$ una familia de subconjuntos de A , entonces*

$$1. \quad f^*(\cup\{I_z\}_{z \in J}) = \cup\{f^*(I_z)\}_{z \in J}.$$

$$2. \quad f^*(\cap\{I_z\}_{z \in J}) \subseteq \cap\{f^*(I_z)\}_{z \in J}.$$

Demostración. 1.

$$\begin{aligned} y \in f^*(\cup\{I_z\}_{z \in J}) &\Leftrightarrow y = f(x), \text{ para alguna } x \in \cup\{I_z\}_{z \in J} \Leftrightarrow \\ &\Leftrightarrow y = f(x), \text{ para alguna } x \in I_z, \text{ para alguna } z \in J \Leftrightarrow \\ &\Leftrightarrow y \in f^*(I_z), \text{ para alguna } z \in J \Leftrightarrow \\ &\Leftrightarrow y \in \cup\{f^*(I_z)\}_{z \in J}. \end{aligned}$$

2.

$$\begin{aligned} y \in f^*(\cap\{I_z\}_{z \in J}) &\Rightarrow y = f(x), \text{ para alguna } x \in \cap\{I_z\}_{z \in J} \Rightarrow \\ &\Rightarrow y = f(x), \text{ con } x \in I_z, \forall z \in J \Rightarrow \\ &\Rightarrow y \in f^*(I_z), \forall z \in J \Rightarrow \\ &\Rightarrow y \in \cap\{f^*(I_z)\}_{z \in J}. \end{aligned}$$

Por lo tanto $f^*(\cap\{I_z\}_{z \in J}) \subseteq \cap\{f^*(I_z)\}_{z \in J}$. ■

Teorema 19 . *Son equivalentes para $A \xrightarrow{f} B$:*

1. $A \xrightarrow{f} B$ es suprayectiva.

2. $\wp(A) \xrightarrow{f^*} \wp(B)$ es suprayectiva.

Demostración. 1). \Rightarrow 2)

Si $A \xrightarrow{f} B$ entonces tiene inverso derecho $g : B \rightarrow A$, veamos que $g^* : \wp(B) \rightarrow \wp(A)$ es inverso derecho para $f^* : \text{Sea } Y \in \wp(B)$ entonces

$$\begin{aligned} f^*(g^*(Y)) &= f^*(\{g(y) \mid y \in Y\}) = \{f(g(y)) \mid y \in Y\} = \\ &= \{(f \circ g)(y) \mid y \in Y\} = \{Id_B(y) \mid y \in Y\} = Y = Id_{\wp(B)}(Y). \end{aligned}$$

2). \Rightarrow 1). Supongamos que f^* es suprayectiva. Sea $y \in B$ como $\{y\} = f^*(X)$ para algún subconjunto X de A , es claro que $f(x) = y$ para cada $x \in X$. Por lo tanto $y \in f(X)$, para cada $y \in B$. ■

El teorema anterior hace natural que nos planteemos lo correspondiente para una función inyectiva.

Teorema 20 . Son equivalentes para $A \xrightarrow{f} B$:

1. $A \xrightarrow{f} B$ es inyectiva.
2. $\wp(A) \xrightarrow{f^*} \wp(B)$ es inyectiva.

Demostración. 1) \Rightarrow 2)

Si $A \xrightarrow{f} B$ es inyectiva entonces tiene inverso izquierdo $B \xrightarrow{g} A$.

Sea $X \subseteq A$, entonces

$$\begin{aligned} g^* f^* (X) &= g(\{f(x) \mid x \in X\}) = \{(gf)(x) \mid x \in X\} = \\ &= \{Id_A(x) \mid x \in X\} = X = Id_{\wp(A)}(X). \end{aligned}$$

2) \Rightarrow 1)

Supongamos ahora que f^* es inyectiva. Supongamos que $a \neq b \in X$, entonces $\{a\} \neq \{b\}$.

Por lo tanto $\{f(a)\} = f^*\{a\} \neq f^*\{b\} = \{f(b)\}$ de donde tenemos que $f(a) \neq f(b)$. ■

Definición 40 . Si $A \xrightarrow{f} B$ y $Y \subseteq B$, definimos

$$f^{-1}(Y) = \{x \in A \mid f(x) \in Y\}.$$

Notemos que $f^{-1}(Y)$ existe aunque f no sea biyectiva, es decir aunque no exista la función inversa de f .

Si $A \xrightarrow{f} B$ es una función podemos definir una función

$$\begin{array}{ccc} \wp(B) & \xrightarrow{f_*} & \wp(A) \\ Y & \mapsto & f^{-1}(Y) \end{array}.$$

f_* se comporta mejor que f^* respecto a preservar uniones e intersecciones.

Teorema 21 . Si $X, Y \in B$, $\wp(B) \xrightarrow{f_*} \wp(A)$ entonces

1. $X \subseteq Y \Rightarrow f_*(X) \subseteq f_*(Y)$.
2. $f_*(X \cap Y) = f_*(X) \cap f_*(Y)$.
3. $f_*(X \cup Y) = f_*(X) \cup f_*(Y)$.

Demostración. 1. Notemos que $a \in f_*(X) \Leftrightarrow f(a) \in X$.

Así que si $X \subseteq Y$ y $a \in f_*(X)$, entonces $f(a) \in Y$, por lo que $a \in f_*(Y)$.

2. $X \cap Y \subseteq X \Rightarrow f_*(X \cap Y) \subseteq f_*(X)$. Análogamente, $f_*(X \cap Y) \subseteq f_*(Y)$, por lo tanto

$$f_*(X \cap Y) \subseteq f_*(X) \cap f_*(Y).$$

Recíprocamente, $a \in f_*(X) \cap f_*(Y)$ es equivalente a $f(a) \in X \wedge f(a) \in Y$ que equivale a $f(a) \in X \cap Y$. Por último $f(a) \in X \cap Y \Leftrightarrow a \in f_*(X \cap Y)$.

3. $X \subseteq X \cup Y \Rightarrow f_*(X) \subseteq f_*(X \cup Y)$.

Análogamente, $f_*(Y) \subseteq f_*(X \cup Y)$, por lo tanto

$$f_*(X) \cup f_*(Y) \subseteq f_*(X \cup Y).$$

Ahora,

$$a \in f_*(X \cup Y) \Rightarrow f(a) \in X \cup Y.$$

$$f(a) \in X \cup Y \Rightarrow f(a) \in X \vee f(a) \in Y.$$

Así:

$$f(a) \in X \Rightarrow a \in f_*(X),$$

$$f(a) \in Y \Rightarrow a \in f_*(Y).$$

Como resultado, $a \in f_*(X \cup Y) \Rightarrow a \in f_*(X) \cup f_*(Y)$. ■

f_* también se comporta bien respecto a uniones e intersecciones arbitraria,

Teorema 22 . *Sea $A \xrightarrow{f} B$ una función y sea $\{I_x\}_{x \in J}$ una familia de subconjuntos de B . Entonces*

$$1. f_*(\cap \{I_x\}_{x \in J}) = \cap \{f_*(I_x)\}_{x \in J}.$$

$$2. f_*(\cup \{I_x\}_{x \in J}) = \cup \{f_*(I_x)\}_{x \in J}.$$

Demostración. 1. $z \in f_*(\cap \{I_x\}_{x \in J}) \Leftrightarrow f(z) \in \cap \{I_x\}_{x \in J} \Leftrightarrow$

$$\Leftrightarrow f(z) \in I_x, \forall x \in J \Leftrightarrow z \in f_*(I_x), \forall x \in J \Leftrightarrow$$

$$\Leftrightarrow z \in \cap \{f_*(I_x)\}_{x \in J}.$$

$$2. z \in f_*(\cup \{I_x\}_{x \in J}) \Leftrightarrow f(z) \in \cup \{I_x\}_{x \in J} \Leftrightarrow$$

$\Leftrightarrow f(z) \in I_x$, para alguna $x \in J \Leftrightarrow z \in f_*(I_x)$, para alguna $x \in J \Leftrightarrow z \in \cup \{f_*(I_x)\}_{x \in J}$. ■

Teorema 23 . *Son equivalentes:*

1. $A \xrightarrow{f} B$ es inyectiva.
2. $\wp(B) \xrightarrow{f_*} \wp(A)$ es suprayectiva.

Demostración. 1) \Rightarrow 2) Supongamos que f es inyectiva y supongamos que $X \in \wp(A)$, veremos que

$$X = f_*(f^*(X)) :$$

Es claro que $X \subseteq f^{-1}(f(X)) = f_*(f^*(X))$.

Por otra parte si $y \in f^{-1}(f(X))$ entonces $f(y) \in f(X)$. Esto significa que $f(y) = f(x)$, para alguna $x \in X$. Como f es inyectiva, entonces $y = x \in X$. Por lo tanto

$$f^{-1}(f(X)) \subseteq X.$$

2) \Rightarrow 1) Si $A \xrightarrow{f} B$ no fuera inyectiva entonces habrían dos elementos distintos a, c de A tales que $f(a) = f(c)$, pero en ese caso $\{a\} \notin f_*(\wp(B))$:

Si $\{a\} = f_*(Y)$, con $Y \subseteq B$, entonces $f(a) \in Y$. Así también $f(a) = f(c) \in Y$, por lo que $c \in f_*(Y) = \{a\}$, contra la hipótesis de que $a \neq c$. ■

2.7 Relaciones de equivalencia y particiones

En esta sección veremos un concepto que es de fundamental importancia dentro de las Matemáticas, el concepto de relación de equivalencia.

Definición 41 . Una relación R en un conjunto A es de equivalencia si

1. Es reflexiva.
2. Es simétrica, es decir $(a, b) \in R \Rightarrow (b, a) \in R$.
3. Es transitiva, es decir que $R \circ R \subseteq R$, o lo que es lo mismo: si

$$(a, b) \in R \wedge (b, c) \in R \Rightarrow (a, c) \in R.$$

Notemos que la condición de simetría se puede expresar de la siguiente forma:

$$R \subseteq R^{-1}.$$

Notemos que si R es simétrica entonces R^{-1} también lo es:

$$(a, b) \in R^{-1} \Rightarrow (b, a) \in R,$$

$$(b, a) \in R \Rightarrow (a, b) \in R,$$

$$(a, b) \in R \Rightarrow (b, a) \in R^{-1}.$$

Por lo que $R^{-1} \subseteq (R^{-1})^{-1} = R$.

Así que es lo mismo pedir que $R \subseteq R^{-1}$ a pedir que $R = R^{-1}$.

Ejemplos 75

1. La relación diagonal $D(A)$ es una relación de equivalencia en A ,

- (a) Es reflexiva pues $D(A) \subseteq D(A)$.
- (b) Es simétrica pues $D(A) = D(A)^{-1}$.
- (c) Es transitiva porque $D(A) \circ D(A) \subseteq D(A)$.

2. La relación $A \times A$ es una relación de equivalencia en A pues

- (a) $D(A) \subseteq A \times A : (a, a) \in A \times A, \forall a \in A$.
- (b) $(A \times A)^{-1} = A \times A$.
- (c) $(A \times A) \circ (A \times A) \subseteq A \times A$.

3. Si $A \xrightarrow{f} B$ es una función definamos $R \subseteq A \times A$ por:

$$R = \{(a, c) \mid f(a) = f(c)\}.$$

Esta relación es

- (a) Reflexiva pues $(a, a) \in R \Leftrightarrow f(a) = f(a)$.

- (b) Simétrica pues

$$(a, c) \in R \Leftrightarrow f(a) = f(c) \Leftrightarrow f(c) = f(a) \Leftrightarrow (c, a) \in R.$$

- (c) Transitiva, pues si $(f(a) = f(b) \wedge f(b) = f(c))$ entonces $f(a) = f(c)$.

4. En $\wp(\{0, 1, 2\})$ definimos la relación R por $(A, B) \in R$ si $|A| = |B|$, entonces R es una relación de equivalencia porque es un caso particular del ejemplo anterior tomando

$$\wp(\{0, 1, 2\}) \xrightarrow{\perp\perp} \{0, 1, 2, 3\}.$$

Por ejemplo, $|\{0, 2\}| = |\{1, 2\}|$ por lo que $(\{0, 2\}, \{1, 2\}) \in R$.

Notación 3 . *Frecuentemente se emplea la notación aRb para decir que $(a, b) \in R$. Además se emplea $a \not R b$ para decir que $(a, b) \notin R$.*

1. Si R es una relación de equivalencia en el conjunto A , y $x \in A$, definimos la clase de equivalencia de x , $[x]_R$ por

$$[x]_R = \{y \in A \mid xRy\}.$$

2. El conjunto de clases de equivalencia se denota

$$A/R = \{[x]_R \mid x \in A\} \subseteq \wp(A).$$

Ejemplos 76

1. Tomemos la relación $D(A)$ en $A = \{0, 1, 2, 3\}$ entonces las clases de equivalencia son

$$[0]_{D(A)} = \{0\}, [1]_{D(A)} = \{1\}, [2]_{D(A)} = \{2\}, [3]_{D(A)} = \{3\},$$

por lo que

$$A/D(A) = \{\{0\}, \{1\}, \{2\}, \{3\}\}.$$

2. El conjunto de clases de equivalencia de la relación $A \times A$ es A , así que $A/(A \times A) = \{A\}$ (una sola clase de equivalencia).

Tomando, $A = \{0, 1, 2, 3\}$, tenemos que

$$A/(A \times A) = \{\{0, 1, 2, 3\}\},$$

$$[0] = [1] = [2] = [3] = \{0, 1, 2, 3\}.$$

3. Si consideramos $A \xrightarrow{f} B$ y la relación de equivalencia definida en A por

$$xRy \Leftrightarrow f(x) = f(y)$$

entonces la clase de equivalencia de $a \in A$ es

$$[a]_R = \{b \in A \mid f(b) = f(a)\}.$$

4. Consideremos la relación R : “tener la misma cardinalidad que” en $\wp(A)$, para $A = \{0, 1, 2\}$.

Entonces

$$[\emptyset] = \{\emptyset\}$$

(el único subconjunto de A sin elementos es \emptyset).

$$[\{1\}] = \{\{0\}, \{1\}, \{2\}\},$$

$$[\{1, 2\}] = \{\{0, 2\}, \{1, 2\}, \{0, 1\}\}$$

y

$$[\{0, 1, 2\}] = \{\{0, 1, 2\}\}.$$

Observación 33 . *Sea R una relación de equivalencia en A notemos que el conjunto A/R de clases de equivalencia tiene las siguientes propiedades:*

1. $\cup A/R = \cup \{[a]_R \mid a \in A\} = A$.
2. $([a]_R \neq [b]_R \in A/R) \Rightarrow ([a]_R \cap [b]_R = \emptyset)$.
3. $\forall z \in A/R, z \neq \emptyset$.

Demostración. 1. Notemos primero que cada clase de equivalencia es un subconjunto de A : $[a]_R \subseteq A$, por lo que

$$\cup \{[a]_R \mid a \in A\} \subseteq A.$$

Recíprocamente $a \in A \Rightarrow a \in [a]_R$ (cada elemento de A pertenece a su propia clase de equivalencia). Por lo tanto

$$A \subseteq \cup \{[a]_R \mid a \in A\}.$$

2. Por contrapuesta:

Si $c \in [a]_R \cap [b]_R$ entonces cRa y cRb así que aRc y cRb . Por transitividad tenemos que aRb . Así que $b \in [a]_R$.

Tomemos un elemento en $[b]_R$, digamos d . Entonces bRd . Así que bRd y aRb por lo que aRd , es decir $d \in [a]_R$. Con esto tenemos que $[b]_R \subseteq [a]_R$.

Intercambiando a con b en el argumento anterior, tenemos que $[a]_R \subseteq [b]_R$.

3. Si $z \in A/R$ entonces $z = [a]_R$ para alguna $a \in A$, por lo tanto $a \in z$, y así $z \neq \emptyset$. ■

Definición 42 . Una familia \mathcal{P} de subconjuntos de A ($\mathcal{P} \subseteq \wp(\mathcal{A})$) es una partición de A si

1. $\cup \mathcal{P} = \mathcal{A}$ (Cada elemento de A pertenece a un elemento de la familia \mathcal{P}).
2. $Z \neq Y \in \mathcal{P} \Rightarrow Z \cap Y = \emptyset$. (Partes distintas son ajenas).
3. $Y \in \mathcal{P} \Rightarrow Y \neq \emptyset$. (Cada parte es no vacía).

Ejemplo 77 . Tomemos las particiones de $\{0, 1, 2\}$:

En una parte: $\{\{0, 1, 2\}\}$.

En dos partes: $\{\{0\}, \{1, 2\}\}, \{\{1\}, \{0, 2\}\}, \{\{2\}, \{1, 0\}\}$.

En tres partes: $\{\{0\}, \{1\}, \{2\}\}$.

Ejemplo 78 . La partición vacía \emptyset es una partición de \emptyset :

Para empezar, $\emptyset \subseteq \wp(\emptyset)$

$$\cup \emptyset = \emptyset,$$

$$x \neq y \in \emptyset \Rightarrow x \cap y = \emptyset,$$

$$z \in \emptyset \Rightarrow z \neq \emptyset.$$

Ejercicio 110 . Demuestre que la partición vacía es la única partición del conjunto vacío. Así que el número de particiones de \emptyset es 1.

Los conceptos de relación de equivalencia y de partición están estrechamente ligados. Ya vimos que el conjunto de clases de equivalencia de elementos de A respecto de una relación de equivalencia R en A forma una partición de A .

Recíprocamente, dada una partición \mathcal{P} de A podemos definir la relación $\approx_{\mathcal{P}}$ en A por:

$$a \approx_{\mathcal{P}} b \text{ si } a, b \in Y, \text{ para } Y \in \mathcal{P}.$$

Es decir, $a \approx_P b$ si a y b están en la misma parte de la partición.

Otra manera de describir la relación \approx_P es

$$\approx_P = \bigcup \{Y \times Y\}_{Y \in \mathcal{P}} \subseteq A \times A.$$

Ejemplo 79 . *Dada la partición*

$$P = \{\{0, 1\}, \{2\}, \{3, 4, 5\}\}$$

de

$$A = \{0, 1, 2, 3, 4, 5\}$$

tenemos que

$$\begin{aligned} \approx_P &= \bigcup \{\{0, 1\} \times \{0, 1\}, \{2\} \times \{2\}, \{3, 4, 5\} \times \{3, 4, 5\}\} = \\ &= \bigcup \left\{ \begin{array}{c} \{(0, 0), (0, 1), (1, 0), (1, 1)\}, \\ \{(2, 2)\}, \\ \{(3, 3), (3, 4), (3, 5), (4, 3), (4, 4), (4, 5), \\ (5, 3), (5, 4), (5, 5)\} \end{array} \right\} = \\ &= \left\{ \begin{array}{c} (0, 0), (0, 1), (1, 0), (1, 1), (2, 2), \\ (3, 3), (3, 4), (3, 5), (4, 3), (4, 4), (4, 5), (5, 3), (5, 4), (5, 5) \end{array} \right\}. \end{aligned}$$

Denotemos por $Eq(A)$ el conjunto de relaciones de equivalencia definidas en A . Note que en efecto es un conjunto, pues dada una relación de equivalencia R en A , se tiene que $R \subseteq (A \times A)$, así que $R \in \wp(A \times A)$ por lo que

$$Eq(A) = \{R \in \wp(A \times A) \mid R \text{ es una relación de equivalencia en } A\}.$$

Denotemos ahora por $Part(A)$ el conjunto de particiones de A , de nuevo, notemos que $Part(A)$ es en efecto un conjunto, pues si \mathcal{P} es una partición en A entonces sus elementos son subconjuntos de A , es decir que $\mathcal{P} \subseteq \wp(A)$ por lo que $\mathcal{P} \in \wp(\wp(A))$ así que

$$Part(A) = \{\mathcal{P} \in \wp(\wp(A)) \mid \mathcal{P} \text{ es una partición de } A\}.$$

Lema 1 . *Si \mathcal{P} es una partición de A entonces $\approx_{\mathcal{P}}$ es una relación de equivalencia en A .*

Además las clases de equivalencia de $\approx_{\mathcal{P}}$ son los elementos de la partición \mathcal{P} .

Demostración. Comencemos ahora con una partición de A , \mathcal{P} . Demostremos que $\approx_{\mathcal{P}}$ es una relación de equivalencia en A :

Como \mathcal{P} es una partición de A , tenemos que dada $a \in A$, $\exists Y \in \mathcal{P}$ tal que $a \in Y$ (cada elemento de A pertenece a una de las partes de \mathcal{P}). Por lo tanto $(a, a) \in Y \times Y \subseteq \cup\{X \times X \mid X \in \mathcal{P}\} = \approx_{\mathcal{P}}$. Por lo tanto $D(A) \subseteq \approx_{\mathcal{P}}$ y así $\approx_{\mathcal{P}}$ es reflexiva.

$$(a, b) \in \approx_{\mathcal{P}} = \cup\{X \times X \mid X \in \mathcal{P}\} \Rightarrow (a, b) \in Z \times Z, \text{ para alguna } Z \in \mathcal{P}.$$

Entonces $(b, a) \in Z \times Z$, por lo que $(b, a) \in \approx_{\mathcal{P}} \Leftrightarrow (a, b) \in (\approx_{\mathcal{P}})^{-1}$. Por lo tanto $\approx_{\mathcal{P}} \subseteq (\approx_{\mathcal{P}})^{-1}$, por lo que $\approx_{\mathcal{P}}$ es simétrica.

Si $(a, b), (b, c) \in \approx_{\mathcal{P}}$ entonces existen $X, Y \in \mathcal{P}$ tales que $a, b \in X$ y $b, c \in Y$. Así, $b \in X \cap Y$, lo que implica que $X = Y$ (en una partición, partes distintas son ajenas). Entonces $a, c \in X = Y$, por lo que $(a, c) \in X \times X$. Entonces $(a, c) \in \approx_{\mathcal{P}}$. Vemos pues que

$$(a, b), (b, c) \in \approx_{\mathcal{P}} \Rightarrow (a, c) \in \approx_{\mathcal{P}},$$

es decir que $\approx_{\mathcal{P}}$ es transitiva.

Por último, si \mathcal{P} es una partición de A , y $a \in A$ entonces existe una **única** $Y \in \mathcal{P}$ tal que $a \in Y$.

Como $a \approx_{\mathcal{P}} b$ si $a, b \in Y$, para $Y \in \mathcal{P}$, es claro que

$$[a]_{\approx_{\mathcal{P}}} = Y,$$

donde Y es el elemento de \mathcal{P} que contiene a a . ■

Teorema 24 . *La función*

$$\begin{array}{ccc} \text{Eq}(A) & \xrightarrow{p} & \text{Part}(A) \\ R & \mapsto & A/R \end{array}$$

es una biyección con inversa

$$\begin{array}{ccc} \text{Part}(A) & \xrightarrow{q} & \text{Eq}(A) \\ \mathcal{U} & \mapsto & \approx_{\mathcal{P}} \end{array}.$$

Demostración. Que p es una función se sigue de la observación 33. Veamos que q es la función inversa de p :

Sea R una relación de equivalencia en A , entonces

$$q(p(R)) = q(A/R) = \approx_{A/R}.$$

Queremos demostrar que $R = \approx_{A/R}$.

Hemos visto que

$$\approx_{A/R} = \cup \{[a]_R \times [a]_R\}_{a \in A},$$

así pues:

$(a, b) \in R \Rightarrow b \in [a]_R$. Por lo tanto $(a, b) \in [a]_R \times [a]_R$, de donde se tiene que $(a, b) \in \approx_{A/R}$.

$$\therefore R \subseteq \approx_{A/R}.$$

Recíprocamente,

$(x, y) \in \cup \{[a]_R \times [a]_R\}_{a \in A} \Rightarrow (x, y) \in [a]_R \times [a]_R$, para alguna $a \in A$. Así que $x \in [a]_R, y \in [a]_R$, es decir que

$$xRa \wedge aRy,$$

por lo que xRy , por lo que $(x, y) \in R$.

$$\therefore \approx_{A/R} \subseteq R.$$

Con esto hemos demostrado que $q(p(R)) = R = Id_{Eq(A)}(R)$.

Ahora, del Lema anterior tenemos que si \mathcal{U} es una partición de A , entonces $q(\mathcal{U})$ es una relación de equivalencia en A , queremos demostrar que $p(q(\mathcal{U})) = \mathcal{U}$.

$$p(q(\mathcal{U})) = A/q(\mathcal{U}) = \{[a]_{q(\mathcal{U})} \mid a \in A\}.$$

$(Y \in \mathcal{U}) \Rightarrow Y = [b]_{q(\mathcal{U})}$, si $b \in Y$. Por lo tanto $Y = [b]_{q(\mathcal{U})} \in A/q(\mathcal{U})$. De donde tenemos que

$$\mathcal{U} \subseteq A/q(\mathcal{U}).$$

Recíprocamente $[a]_{q(\mathcal{U})} = Y$, si $a \in Y$, con $Y \in \mathcal{U}$. Por lo tanto

$$A/q(\mathcal{U}) \subseteq \mathcal{U}.$$

Por lo tanto

$$p(q(\mathcal{U})) = \mathcal{U} = Id_{Part(A)}(\mathcal{U}).$$

Con esto hemos demostrado que q es inversa de p . ■

Observación 34 . Si \approx es una relación de equivalencia en A entonces

$$\begin{array}{ccc} A & \xrightarrow{p} & A/\approx \\ a & \longmapsto & [a]_{\approx} \end{array}$$

es una función suprayectiva.

2.8 La relación de equivalencia generada por una relación

Observación 35 . Si $R \subseteq A \times A$ entonces $R \cup D(A)$ es la menor relación reflexiva en A que contiene a R .

Ejercicio 111 . Si $R, S \subseteq A \times B$ son relaciones de A a B entonces

1. $(R \cup S)^{-1} = (R^{-1} \cup S^{-1}) \subseteq B \times A$.
2. $(R \cap S)^{-1} = (R^{-1} \cap S^{-1}) \subseteq B \times A$.

Observación 36 . Si $R \subseteq A \times A$ entonces $R \cup R^{-1}$ es la menor relación simétrica que incluye a R .

Demostración. Es claro que $R \subseteq (R \cup R^{-1})$, y que $R \cup R^{-1}$ es simétrica. Ahora, si $R \subseteq S$ y S es simétrica, entonces

$$(b, a) \in R^{-1} \Rightarrow (a, b) \in R,$$

$$(a, b) \in R \Rightarrow (a, b) \in S,$$

$$(a, b) \in S \Rightarrow (b, a) \in S.$$

$$\therefore R^{-1} \subseteq S$$

Ahora, $(R \subseteq S \wedge R^{-1} \subseteq S) \Rightarrow (R \cup R^{-1}) \subseteq S$. ■

Si $R \subseteq A \times A$ consideremos $R \circ R$ y hagamos

$$R_1 =: R \cup (R \circ R),$$

luego hagamos

$$R_2 =: R_1 \cup (R_1 \circ R_1),$$

y así sucesivamente, es decir que

$$R_{s(k)} =: R_k \cup (R_k \circ R_k),$$

donde $s(k)$ denota el sucesor de k .

Es claro que

$$R \subseteq R_1 \subseteq R_2 \subseteq \dots \subseteq R_k \subseteq R_{s(k)} \subseteq \dots$$

Consideremos $R_\infty = \cup \{R, R_1, R_2, \dots, R_k, R_{s(k)}, \dots\}$ mostraremos que R_∞ es la menor (en el sentido de la contención) relación transitiva en A que contiene a R .

Proposición 12 . *Sean R y R_∞ como en el párrafo anterior, entonces R_∞ es la menor relación transitiva en A que contiene a R .*

Demostración. Es claro que $R \subseteq R_\infty$.

Además R_∞ es transitiva, pues podemos notar que $R_\infty \circ R_\infty \subseteq R_\infty$:

Si $(a, c) \in (R_\infty \circ R_\infty)$ entonces existe $b \in A$ tal que $(a, b) \in R_\infty, (b, c) \in R_\infty$ pero entonces $(a, b) \in R_n$ y $(b, c) \in R_m$, donde R_n y R_m aparecen en la lista

$$R \subseteq R_1 \subseteq R_2 \subseteq \dots \subseteq R_k \subseteq R_{s(k)} \subseteq \dots$$

supongamos que R_n aparece primero, entonces $R_n \subseteq R_m$, por lo que

$$(a, b), (b, c) \in R_m.$$

Así que $(a, c) \in R_m \circ R_m \subseteq R_{s(m)} \subseteq R_\infty$.

$$\therefore (R_\infty \circ R_\infty) \subseteq R_\infty.$$

Es decir que R_∞ es transitiva.

Por último es claro que si $R \subseteq T$ y T es una relación transitiva en A entonces

$$R \circ R \subseteq T \circ T \subseteq T,$$

por lo que

$$R_1 = R \cup (R \circ R) \subseteq T,$$

de la misma manera

$$R_2 = R_1 \cup (R_1 \circ R_1) \subseteq T,$$

y

$$R \subseteq R_1 \subseteq R_2 \subseteq \dots \subseteq R_k \subseteq R_{s(k)} \subseteq \dots \subseteq T$$

por lo que $R_\infty \subseteq T$. ■

Observación 37 . Si $R \subseteq A \times B, S \subseteq B \times C$, entonces $(S \circ R)^{-1} = R^{-1} \circ S^{-1}$:

Demuestra. $(z, x) \in (S \circ R)^{-1} \Leftrightarrow (x, z) \in S \circ R \Leftrightarrow \exists y \in B \text{ tal que } (x, y) \in R \wedge (y, z) \in S \Leftrightarrow \exists y \in B \text{ tal que } (y, x) \in R^{-1} \wedge (z, y) \in S^{-1} \Leftrightarrow (z, x) \in R^{-1} \circ S^{-1}$. ■

Observación 38 . Si R es una relación simétrica, entonces R_∞ también lo es.

Demuestra. R simétrica $\Rightarrow R_1$ es simétrica:

$R_1 = R \cup (R \circ R)$, por lo que

$$R_1^{-1} = R^{-1} \cup (R \circ R)^{-1} = R^{-1} \cup (R^{-1} \circ R^{-1}) = R \cup (R \circ R) = R_1,$$

por lo que R_1 es simétrica.

Con el mismo argumento, R_2, R_3, \dots son simétricas.

Ahora, $(a, b) \in R_\infty \Rightarrow (a, b) \in R_n$, para alguna R_n en la lista

$$R \subseteq R_1 \subseteq R_2 \subseteq \dots \subseteq R_k \subseteq R_{s(k)} \subseteq \dots,$$

como R_n es simétrica, entonces $(b, a) \in R_n \subseteq R_\infty$.

$$\therefore (a, b) \in R_\infty \Rightarrow (a, b) \in R_\infty^{-1}.$$

■

Teorema 25 . Si $R \subseteq A \times A$ es una relación en A entonces

$$(D(A) \cup R \cup R^{-1})_\infty$$

es la menor relación de equivalencia definida en A que contiene a R .

Demuestra. Es claro que $R \subseteq (D(A) \cup R \cup R^{-1})_\infty$.

Además $D(A) \subseteq (D(A) \cup R \cup R^{-1})_\infty$, por lo que $(D(A) \cup R \cup R^{-1})_\infty$ es reflexiva.

Como $(D(A) \cup R \cup R^{-1})^{-1} = (D(A)^{-1} \cup R^{-1} \cup (R^{-1})^{-1}) = D(A) \cup R \cup R^{-1}$, tenemos que $D(A) \cup R \cup R^{-1}$ es simétrica, así que por el teorema y la observación previas, tenemos que $(D(A) \cup R \cup R^{-1})_\infty$ es simétrica y transitiva. Es decir,

$$(D(A) \cup R \cup R^{-1})_\infty$$

es de equivalencia.

Para demostrar que $(D(A) \cup R \cup R^{-1})_\infty$ es la menor relación de equivalencia que contiene a R basta demostrar que si $R \subseteq S$ y S es de equivalencia, entonces

$$D(A) \cup R \cup R^{-1} \subseteq S.$$

$D(A) \subseteq S$ pues S es una relación reflexiva en A .

$R \subseteq S$ por hipótesis.

$R \subseteq S \Rightarrow R^{-1} \subseteq S$, porque S es simétrica,

$$\therefore D(A) \cup R \cup R^{-1} \subseteq S.$$

■

Ejemplo 80 . Sea

$$R = \{(0,0), (0,1), (1,2), (3,4)\} \subseteq \{0,1,2,3,4,5\} \times \{0,1,2,3,4,5\}.$$

Entonces $R^{-1} = \{(0,0), (1,0), (2,1), (4,3)\}$ y

$$D(A) = \{(0,0), (1,1), (2,2), (3,3), (4,4), (5,5)\}.$$

Por lo tanto $X = D(A) \cup R \cup R^{-1} =$

$$\begin{aligned} & \{(0,0), (0,1), (1,2), (3,4)\} \cup \\ & \cup \{(0,0), (1,0), (2,1), (4,3)\} \cup \\ & \cup \{(0,0), (1,1), (2,2), (3,3), (4,4), (5,5)\} \end{aligned}$$

$$= \left\{ (0,0), (0,1), (1,2), (3,4), (1,0), (2,1), \dots, (4,3), (1,1), (2,2), (3,3), (4,4), (5,5) \right\}.$$

$$X_1 = X \circ X = \left\{ \begin{array}{l} (0,0), (0,1), (0,2) \\ (1,1), (1,2), (1,0), \\ (2,2), (2,1), (2,0) \\ (3,3), (3,4), \\ (4,4), (4,3) \\ (5,5) \end{array} \right\},$$

$$X_1 \circ X_1 = \left\{ \begin{array}{l} (0,0), (0,1), (0,2) \\ (1,1), (1,2), (1,0), \\ (2,2), (2,1), (2,0) \\ (3,3), (3,4), \\ (4,4), (4,3) \\ (5,5) \end{array} \right\}, \text{ entonces } X_1 = X_1 \circ X_1, \text{ y } X_1 \text{ ya es una}$$

relación de equivalencia.

Sería más fácil plantear el ejemplo anterior en términos de particiones: ¿Cuál es la partición de $\{0, 1, 2, 3, 4, 5\}$ con partes más pequeñas tal que 0, 1 y 2 están en la misma parte, y 3 está en la misma parte que 4? Planteado este problema es claro que la solución es $\{\{0, 1, 2\}, \{3, 4\}, \{5\}\}$ y de aquí que la menor relación de equivalencia que contiene a R es:

$$(\{0, 1, 2\} \times \{0, 1, 2\}) \cup (\{3, 4\} \times \{3, 4\}) \cup (\{5\} \times \{5\}) =$$

$$= \left\{ \begin{array}{l} (0, 0), (0, 1), (0, 2), \\ (1, 0), (1, 1), (1, 2), \\ (2, 0), (2, 1), (2, 2), \\ (3, 3), (3, 4), (4, 3), (4, 4), \\ (5, 5) \end{array} \right\}.$$

Ejercicio 112 . Demuestre que si una relación R en A es reflexiva, entonces $R \subseteq R \circ R$.

Ejercicio 113 . Demuestre que si R es una relación de orden y una relación de equivalencia en A entonces $R = D(A)$.

Ejercicio 114 . Sea $\{R_i\}_{i \in J}$ una familia de relaciones reflexivas en A , demuestre que $\cap \{R_i\}_{i \in J}$ es reflexiva.

Ejercicio 115 . Sea $\{R_i\}_{i \in J}$ una familia de relaciones simétricas en A , demuestre que $\cap \{R_i\}_{i \in J}$ es simétrica.

Sugerencia: demuestre que $(\cap \{R_i\}_{i \in J})^{-1} = (\cap \{R_i^{-1}\}_{i \in J})$.

Ejercicio 116 . Sea $\{R_i\}_{i \in J}$ una familia de relaciones transitivas en A , demuestre que $\cap \{R_i\}_{i \in J}$ es transitiva.

Sugerencia: demuestre que $(\cap \{R_i\}_{i \in J}) \circ (\cap \{R_i\}_{i \in J}) = (\cap \{R_i\}_{i \in J})$.

Ejercicio 117 . Sea $\{R_i\}_{i \in J}$ una familia de relaciones de equivalencia en A demuestre que $\cap \{R_i\}_{i \in J}$ es de equivalencia.

Ejercicio 118 . Use los ejercicios anteriores para dar otra demostración que dada una relación R en A existe una menor relación de equivalencia en que contiene a R .

Notación 4 . Denotemos por $\text{Rel}(A, B)$ el conjunto de relaciones de A a B . Como una relación de A a B no es otra cosa que un subconjunto de $A \times B$ (es decir, un elemento de $\wp(A \times B)$), después de todo tenemos que

$$\text{Rel}(A, B) = \wp(A \times B).$$

Así que tenemos definido un orden en $\text{Rel}(A, B)$, a saber, la contención. Es decir que

$$R \leq S \Leftrightarrow R \subseteq S.$$

Observación 39 . Si (A, \leq) es un conjunto parcialmente ordenado y $A \xrightarrow{f} B$ es una función biyectiva, entonces podemos definir un orden parcial \preceq en B por:

$$b_1 \preceq b_2 \text{ si } f^{-1}(b_1) \leq f^{-1}(b_2).$$

Consideremos $\text{Eq}(A)$ el conjunto de relaciones de equivalencia en el conjunto A y $\text{Part}(A)$ el conjunto de particiones de A como ya sabemos que

$$\begin{array}{ccc} \text{Eq}(A) & \xrightarrow{p} & \text{Part}(A) \\ R & \mapsto & A/R \end{array}$$

es una biyección, podemos dotar de orden a $\text{Part}(A)$:

$$P \leq Q \Leftrightarrow p^{-1}(P) \subseteq p^{-1}(Q).$$

Cuando $P \leq Q$ decimos que P es más fina que Q y la razón es la siguiente:

$$P \leq Q \Leftrightarrow A/(\approx_P) \subseteq A/(\approx_Q).$$

Por lo tanto si a y b están en la misma parte de P entonces también están en la misma parte de Q . Esto quiere decir que si $a \in Y \in P$ y $a \in Z \in Q$, entonces $Y \subseteq Z$. Así que las partes de Q están formadas por la unión de algunas de las partes de P .

Así que la mayor partición de A es $\{A\}$ mientras que la menor (“la más fina”) es $\{\{a\} \mid a \in A\}$.

Ejemplo 81 . Consideremos las dos particiones de $\{0, 1, 2, 3, 4, 5\}$:

$$\{\{0, 1\}, \{2, 3\}, \{4\}, \{5\}\}$$

y

$$\{\{0, 1, 2, 3\}, \{4, 5\}\},$$

entonces

$$\{\{0, 1\}, \{2, 3\}, \{4\}, \{5\}\} \preceq \{\{0, 1, 2, 3\}, \{4, 5\}\}.$$

Ejercicio 119 . Sean P y Q dos particiones de A demuestre que

$$P \preceq Q \Leftrightarrow (\forall Y \in P, \exists Z \in Q \text{ tal que } Y \subseteq Z).$$

Ejercicio 120 . Suponga que $P \preceq Q$ y sea $Z \in Q$, demuestre que $\{Y \in P \mid Y \subseteq Z\}$ es una partición de Z .

Ejercicio 121 . Sean P, Q particiones de A definamos $P \wedge Q$ como la mayor de las particiones de A tales que son más finas que P y que Q . Demuestre que

$$P \wedge Q = \{Y \cap Z \mid Y \in P, Z \in Q, Y \cap Z \neq \emptyset\}.$$

Ejercicio 122 . Consideremos las dos particiones de $\{0, 1, 2, 3, 4, 5\}$:

$$\{\{0, 1\}, \{2, 3\}, \{4\}, \{5\}\}$$

y

$$\{\{0, 1, 2\}, \{3, 4, 5\}\}.$$

Describa $\{\{0, 1\}, \{2, 3\}, \{4\}, \{5\}\} \wedge \{\{0, 1, 2\}, \{3, 4, 5\}\}$.

Ejercicio 123 . Demuestre que si R es una relación de equivalencia en A entonces existe un conjunto B y una función suprayectiva $A \xrightarrow{f} B$ tal que $aRc \Leftrightarrow f(a) = f(c)$.

Ejercicio 124 . Muestre que si $A \xrightarrow{f} B$ es una función suprayectiva entonces hay una relación de equivalencia R en A y una función biyectiva $\hat{f} : A/R \rightarrow B$ tal que $\hat{f} \circ p = f$, donde $\begin{array}{ccc} A & \xrightarrow{p} & A/R \\ a & \mapsto & [a]_R \end{array}$.

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ & \searrow^p & \uparrow^{\hat{f}} \\ & & A/R \end{array}$$

Ejercicio 125 . Suponga que $A \xrightarrow{f} B$ es una biyección, entonces

$$f^* : \wp(A) \rightarrow \wp(B)$$

es una biyección con inverso

$$f_* : \wp(B) \rightarrow \wp(A).$$

Demuestre que la función

$$\begin{array}{ccc} \text{Part}(A) & \rightarrow & \text{Part}(B) \\ P & \longmapsto & \{f^*(X) \mid X \in P\} \end{array}$$

es una biyección.

Sugerencia: $f^* : \wp(A) \rightarrow \wp(B)$ biyección $\Rightarrow (f^*)^* : \wp(\wp(A)) \rightarrow \wp(\wp(B))$ es una biyección con inverso $(f^*)_*$. Note también que $\text{Part}(A) \subseteq \wp(\wp(A))$.

Ejercicio 126 . En el ejercicio anterior demuestre que

$$\left((f^*)^*|_{\text{Part}(A)} \right)^{-1} = ((f^*)_*)|_{\text{Part}(B)} = \left(((f^{-1})^*)^* \right)|_{\text{Part}(B)}.$$

Notación 5 Sean A y B conjuntos, denotemos por

$$B^A = \left\{ A \xrightarrow{f} B \mid f \text{ es función} \right\}.$$

Observación 40 Notemos que B^A es un conjunto:

Si $A \xrightarrow{f} B$ es una función, entonces $f \in \text{Rel}(A, B) = \wp(A \times B)$, luego $B^A = \{(A, f, B) \in \{A\} \times \wp(A \times B) \times \{B\} \mid (A, f, B) \text{ es una función}\}$.

Ejemplo 82 . Si $A = \{1, 2, 3\}$, $B = \{0, 1\}$ y

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ 1 & \mapsto & 0 \\ 2 & \mapsto & 0 \\ 3 & \mapsto & 1 \end{array}$$

entonces

$$f = \{(1, 0), (2, 0), (3, 1)\} \subseteq A \times B,$$

por lo que $f \in \wp(A \times B)$.

Ejemplo 83 . Definamos una relación $\hat{=}$ en B^A por $A \xrightarrow{f} B \hat{=} A \xrightarrow{g} B$ si $f(A) = g(A)$. Claramente esta relación es de equivalencia.

Se puede ver también de esta manera: si $\begin{array}{ccc} B^A & \xrightarrow{\Psi} & \wp(B) \\ f & \mapsto & f(A) \end{array}$ entonces $f \hat{=} g \Leftrightarrow \Psi(f) = \Psi(g)$.

Ejemplo 84 . Tomemos $A = \{0, 1, 2\}$ y $B = \{5, 6\}$ entonces

1.

$$\begin{array}{ccc} \{0, 1, 2\} & \xrightarrow{f} & \{5, 6\} \\ 0 & \mapsto & 5 \\ 1 & \mapsto & 5 \\ 2 & \mapsto & 6 \end{array} \quad \begin{array}{ccc} \{0, 1, 2\} & \xrightarrow{g} & \{5, 6\} \\ 0 & \mapsto & 6 \\ 1 & \mapsto & 6 \\ 2 & \mapsto & 5 \end{array}$$

pues $f(\{0, 1, 2\}) = \{5, 6\} = g(\{0, 1, 2\})$,

2. además $[f]_{\cong} = \left\{ A \xrightarrow{k} B \mid k \text{ es suprayectiva} \right\}$.

3. Note que hay tantas clases de equivalencia en B^A como elementos en $\wp(B) \setminus \{\emptyset\} = \{\{5\}, \{6\}, \{5, 6\}\}$.

Ejercicio 127 . Sea

$$R = \{(0, 1), (2, 3), (3, 4), (5, 5)\} \subseteq \{0, 1, 2, 3, 4, 5, 6\} \times \{0, 1, 2, 3, 4, 5, 6\}$$

1. Encuentre la menor relación reflexiva en $A =: \{0, 1, 2, 3, 4, 5, 6\}$ que incluye a R .
2. Encuentre la menor relación simétrica en $A =: \{0, 1, 2, 3, 4, 5, 6\}$ que incluye a R .
3. Encuentre la menor relación transitiva en $A =: \{0, 1, 2, 3, 4, 5, 6\}$ que incluye a R .
4. Encuentre la menor relación de equivalencia en $A =: \{0, 1, 2, 3, 4, 5, 6\}$ que incluye a R .

2.9 Operaciones

Entre las funciones más importantes se encuentran las operaciones.

Definición 43 . Una operación en un conjunto A es una función

$$f : A \times A \rightarrow A.$$

Por costumbre, en lugar de $f(a, c)$ se escribe afc , y también por costumbre se usan símbolos como $\circ, *, \times, \div, +, -$ para denotar operaciones. Así uno escribe $f \circ g$ en lugar de $\circ(f, g)$ y $2 + 3$ en lugar de escribir $+(2, 3)$.

1. La conjunción

$$\begin{array}{ccc} \{0, 1\} \times \{0, 1\} & \xrightarrow{\wedge} & \{0, 1\} \\ (0, 0) & \longmapsto & 0 \\ (0, 1) & \longmapsto & 0 \\ (1, 0) & \longmapsto & 0 \\ (1, 1) & \longmapsto & 1 \end{array} .$$

es una operación en $\{0, 1\}$.

2. La intersección \cap es una operación en $\wp(A)$:

$$\begin{array}{ccc} \wp(A) \times \wp(A) & \xrightarrow{\cap} & \wp(A) \\ (X, Y) & \longmapsto & X \cap Y \end{array} .$$

3. La unión es una operación en $\wp(A)$:

$$\begin{array}{ccc} \wp(A) \times \wp(A) & \xrightarrow{\cup} & \wp(A) \\ (X, Y) & \longmapsto & X \cup Y \end{array} .$$

4. Tomemos A^A el conjunto de las funciones de A en A , entonces la composición es una operación en A^A :

$$\begin{array}{ccc} A^A \times A^A & \xrightarrow{\circ} & A^A \\ (f, g) & \longmapsto & f \circ g \end{array} ,$$

pues si $A \xrightarrow{f} A$ y $A \xrightarrow{g} A$ son funciones de A en A entonces $A \xrightarrow{f \circ g} A$ también es una función de A en A .

5. Si A es un conjunto entonces

$$\begin{array}{ccc} A \times A & \rightarrow & A \\ (a, c) & \mapsto & a \end{array}$$

es una operación.

6. Si A es un conjunto entonces

$$\begin{array}{ccc} A \times A & \rightarrow & A \\ (a, c) & \mapsto & c \end{array}$$

es una operación.

7. Si (A, \leq, \wedge, \vee) es una retícula, entonces

$$\begin{array}{ccc} A \times A & \xrightarrow{\vee} & A \\ (a, c) & \mapsto & a \vee c \end{array}$$

y

$$\begin{array}{ccc} A \times A & \xrightarrow{\wedge} & A \\ (a, c) & \mapsto & a \wedge c \end{array}$$

son operaciones en A .

2.9.1 La restricción de una operación

Notemos que si $* : A \times A \rightarrow A$ es una operación en A , y $B \subseteq A$ entonces $B \times B \subseteq A \times A$, así podemos considerar la composición de la función inclusión

$$B \times B \hookrightarrow A \times A$$

con la operación

$$A \times A \xrightarrow{*} A,$$

obteniendo

$$\begin{array}{ccccccc} B \times B & \hookrightarrow & A \times A & \xrightarrow{*} & A \\ (b_1, b_2) & \longmapsto & (b_1, b_2) & \longmapsto & b_1 * b_2 \end{array},$$

notemos que esta composición no es una operación en B porque el codominio no es B sino A .

Si se tuviera que el producto de dos elementos de B fuera otra vez un elemento de B , entonces podríamos correstringir (abreviatura de “podríamos tomar la corrección”) la función anterior a B y entonces

$$\begin{array}{ccccccc} B \times B & \hookrightarrow & A \times A & \xrightarrow{*} & B \\ (b_1, b_2) & \longmapsto & (b_1, b_2) & \longmapsto & b_1 * b_2 \end{array}$$

sí sería una operación en B .

Definición 44 . Si $* : A \times A \rightarrow A$ es una operación en A , y $B \subseteq A$ es tal que

$$(b_1, b_2) \in B \times B \Rightarrow b_1 * b_2 \in B,$$

diremos que B es cerrado bajo $*$.

Ejemplo 85 . Consideremos la siguiente operación en $\{0, 1, 2, 3, 4, 5\}$, descrita por la siguiente tabla⁴

*	0	1	2	3	4	5
0	0	1	2	3	4	5
1	1	2	3	4	5	0
2	2	3	4	5	0	1
3	3	4	5	0	1	2
4	4	5	0	1	2	3
5	5	0	1	2	3	4

Notemos que el subconjunto $\{0, 2, 4\}$ es cerrado bajo la operación:

$$\begin{aligned} 0 * 0 &= 0 & 0 * 2 &= 2 & 0 * 4 &= 4 \\ 2 * 0 &= 2 & 2 * 2 &= 4 & 2 * 4 &= 0, \\ 4 * 0 &= 4 & 4 * 2 &= 0 & 4 * 4 &= 2 \end{aligned}$$

así que podemos tomar

$$\{0, 2, 4\} \times \{0, 2, 4\} \xrightarrow{\ast|_{\{0, 2, 4\} \times \{0, 2, 4\}}} \{0, 2, 4\}.$$

Ejercicio 128 . Sea A un conjunto, hemos notado que la composición es una operación en A^A , demuestre que

1. El conjunto $\{f \in A^A \mid f \text{ es inyectiva}\}$ es un subconjunto de A^A cerrado bajo \circ .

⁴En una tabla de multiplicar, la convención usual es:

$$\begin{array}{ccccccc} * & \cdots & & j & & \cdots & \\ \vdots & \ddots & & \vdots & & \ddots & \\ i & \cdots & & i * j & & \cdots & \\ \vdots & \ddots & & \vdots & & \ddots & \end{array}$$

2. El conjunto $\{f \in A^A \mid f \text{ es suprayectiva}\}$ es un subconjunto de A^A cerrado bajo \circ .
3. El conjunto $\{f \in A^A \mid f \text{ es biyectiva}\}$ es un subconjunto de A^A cerrado bajo \circ .

Ejercicio 129 . Sea A un conjunto y $B \subseteq A$ demuestre que $\{C \subseteq A \mid C \cap B = \emptyset\}$ es un subconjunto de $\wp(A)$ que es cerrado bajo \cup .

2.9.2 Operaciones asociativas

Definición 45 . Decimos que la operación $* : X \times X \rightarrow X$ es *asociativa* si

$$x * (y * z) = (x * y) * z, \quad \forall x, y, z \in X.$$

En este caso la pareja ordenada $(X, *)$ se llama *semigrupo*.

Ejemplos 86 . Son semigrupos:

1. $(\wp(X), \cap)$,
 2. $(\wp(X), \cup)$,
 3. $(\{f : X \rightarrow X \mid f \text{ es una función}\}, \circ)$
1. Denotemos con \mathbb{Z} el conjunto de los enteros $(\mathbb{Z}, -)$ no es un semigrupo:

$$1 = 1 - 0 = 1 - (1 - 1) \neq (1 - 1) - 1 = -1.$$

2. Consideremos la operación de diferencia de conjuntos en

$$\wp(\{0, 1\}) = \{\emptyset, \{0\}, \{1\}, \{0, 1\}\},$$

observando que

$$\{0, 1\} = \{0, 1\} \setminus \emptyset = \{0, 1\} \setminus (\{0, 1\} \setminus \{0, 1\}) \neq (\{0, 1\} \setminus \{0, 1\}) \setminus \{0, 1\} = \emptyset.$$

vemos que $(\wp(\{0, 1\}), \setminus)$ no es un semigrupo.

Definición 46 . La operación $* : A \times A \rightarrow A$ es *comutativa* si $a_1 * a_2 = a_2 * a_1, \forall a_1, a_2 \in A$.

Ejercicio 130 . Demuestre que si B es un subconjunto de A cerrado bajo la operación $*$ entonces

1. $* : A \times A \rightarrow A$ asociativa $\Rightarrow *|_{B \times B}^B : B \times B \rightarrow B$ es asociativa.
2. $* : A \times A \rightarrow A$ comutativa $\Rightarrow *|_{B \times B}^B : B \times B \rightarrow B$ es comutativa.

2.9.3 Tablas de multiplicar

Definición 47 . Sea $*$ una operación en un conjunto finito $\{a_1, a_2, \dots, a_n\}$, la tabla de multiplicar de $*$, es el arreglo cuadrado

*	a_1	a_2	\cdots	a_i	\cdots	a_j	\cdots	a_n
a_1	$a_1 * a_1$			$a_1 * a_i$		$a_1 * a_j$		$a_1 * a_n$
a_2								
\vdots								
a_i	$a_i * a_1$					$a_i * a_j$		
\vdots								
a_j	$a_j * a_1$			$a_j * a_i$				
\vdots								
a_n	$a_n * a_1$					$a_n * a_j$		$a_n * a_n$

Ejemplo 87 . En el conjunto $\{0, 1\}$ se puede definir 16 operaciones:

Calculemos el número elementos en

$$\left\{ \{0, 1\} \times \{0, 1\} \xrightarrow{f} \{0, 1\} \mid f \text{ es una función} \right\}.$$

Notemos lo siguiente: cada uno de los cuatro elementos de $\{0, 1\} \times \{0, 1\}$ tiene que ir a dar a 0 ó a 1 bajo una función de las de arriba. Entonces debe ser claro que hay $2 * 2 * 2 * 2 = 16$ elementos en el conjunto de funciones cuya cardinalidad estamos calculando.

De estas 16 operaciones hay 8 asociativas. De las que mencionamos algunas:

1.

*	0	1
0	0	0
1	0	0

es asociativa.

2. Por la misma razón,

*	0	1
0	1	1
1	1	1

es asociativa.

3. La disyunción lógica

\vee	0	1
0	0	1
1	1	0

es asociativa.

4. La conjunción lógica

\wedge	0	1
0	0	0
1	0	1

es asociativa.

5. Definamos $*$ por: $x * y = y \quad \forall x, y \in \{0, 1\}$. Es claro que las dos maneras de poner paréntesis en

$$x * y * z,$$

nos produce el mismo resultado: z . La tabla correspondiente es

*	0	1
0	0	1
1	0	1

6. Análogamente,

*	0	1
0	0	0
1	1	1

Hasta este momento hemos escrito 6 de las 8 operaciones asociativas que se pueden definir en $\{0, 1\}$.

Ejercicio 131 . Encuentre las otras dos operaciones asociativas que se pueden definir en $\{0, 1\}$.

Para mostrar una operación que no es asociativa en el conjunto $\{0, 1\}$, tomemos la tabla de la implicación, “ \Rightarrow ”:

\Rightarrow	0	1
0	1	1
1	0	1

No es asociativa, pues

$$[0 \Rightarrow (0 \Rightarrow 0)] = [0 \Rightarrow 1] = 1,$$

mientras que

$$[(0 \Rightarrow 0) \Rightarrow 0] = [1 \Rightarrow 0] = 0.$$

Definición 48 *Sea $*$ una operación asociativa en S .*

1. $e \in S$ es un neutro izquierdo para $*$, si $e * x = x, \forall x \in S$.
2. $e \in S$ es un neutro derecho para $*$, si $x * e = x, \forall x \in S$.
3. $e \in S$ es un neutro para $*$, si e es un neutro izquierdo y derecho para $*$.

Observación 41 . *Si e es un neutro izquierdo para $*$ y f es un neutro derecho para la misma operación, entonces $e = f$.*

Demostración. $e = e * f = f$. La primera igualdad se da porque f es neutro derecho y la segunda porque e es neutro izquierdo. ■

Observación 42 . *Si e, f son dos neutros izquierdos distintos para una operación $*$, entonces $*$ no tiene neutro.*

Ejemplo 88 . *Un semigrupo con dos neutros izquierdos:*

*	0	1
0	0	1
1	0	1

Nótese que no hay neutro derecho.

Ejercicio 132 . *Sea $* : A \times A \rightarrow A$ una operación en A , sea $c \in A$ entonces $\{c\} \times A \subseteq A \times A$ y $A \times \{c\} \subseteq A \times A$ por lo que tenemos las funciones*

$$\{c\} \times A \hookrightarrow A \times A \xrightarrow{*} A$$

y

$$A \times \{c\} \hookrightarrow A \times A \xrightarrow{*} A.$$

Demuestre que

1. $\{c\} \times A \hookrightarrow A \times A \xrightarrow{*} A$ es inyectiva $\Leftrightarrow (c * x = c * y \Rightarrow x = y)$, (c es cancelable por la izquierda).
2. $A \times \{c\} \hookrightarrow A \times A \xrightarrow{*} A$ es inyectiva $\Leftrightarrow (x * c = y * c \Rightarrow x = y)$, (c es cancelable por la derecha).

Ejercicio 133 Mantengamos la notación del ejercicio anterior, demuestre que

1. $\{c\} \times A \hookrightarrow A \times A \xrightarrow{*} A$ es suprayectiva \Leftrightarrow

$$(\forall z \in A, \exists x \in A \text{ tal que } c * x = z),$$

(esto es algo así como que cada z se puede dividir por c por la izquierda, por lo que llamaremos a c divisor izquierdo).

2. $A \times \{c\} \hookrightarrow A \times A \xrightarrow{*} A$ es suprayectiva \Leftrightarrow

$$(\forall z \in A, \exists x \in A \text{ tal que } x * c = z),$$

(esto es algo así como que cada z se puede dividir por c por la derecha).

Ejemplo 89 . Consideremos la siguiente operación en $\{0, 1, 2, 3\}$,

*	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

encuentre:

1. Los elementos cancelables por la izquierda.
2. Los elementos cancelables por la derecha.
3. Los elementos divisores izquierdos.
4. Los elementos divisores derechos.

Ejercicio 134 . Suponga que en la tabla de la operación $*$ sucede que en la columna que tiene z hasta arriba todos los elementos son diferentes

$$\begin{array}{cccc}
 * & \dots & z & \dots \\
 \vdots & \ddots & \vdots & \dots \\
 \dots & & x & \dots \\
 & \ddots & \vdots & , \quad x \neq y. \\
 \dots & & y & \\
 & & \vdots &
 \end{array}$$

Muestre que esto es equivalente a que z sea cancelable por la derecha.

Ejercicio 135 . Describa la condición similar a la anterior que caracterice el hecho de que z sea cancelable por la izquierda.

Ejercicio 136 . Suponga que en la tabla de la operación $* : A \times A \rightarrow A$ se tiene que en la columna que tiene c hasta arriba aparecen todos los elementos de A

$$\begin{array}{ccccc}
 * & \dots & & c & \dots \\
 \vdots & \ddots & & \vdots & \dots \\
 \dots & & & z & \dots \\
 & \ddots & & \vdots & , \\
 & & \overbrace{\quad \quad \quad \quad}^{\text{en esta columna}} & & \\
 & & \text{aparecen todos} & & \\
 & & \text{los elementos de } A & &
 \end{array}$$

muestre que esto es equivalente a que c sea un divisor derecho.

Capítulo 3

El conjunto \mathbb{N} de los números naturales

3.1 Introducción

En este capítulo se presenta una construcción del conjunto \mathbb{N} de los números naturales basada en las propiedades que intuitivamente suponemos para ellos y apoyada en los conocimientos previos de lógica y teoría de conjuntos que se han desarrollado en los capítulos anteriores.

\mathbb{N} es el conjunto ordenado $\{0, 1, 2, \dots\}$ y como sabemos, en él podemos efectuar sumas y multiplicaciones, que son operaciones binarias cuyas propiedades son conocidas y queremos formalizar.

Podemos reconocer como características fundamentales en \mathbb{N} : primero, la posibilidad de pasar de una manera precisa de cada número “al que le sigue” y que este paso es tal que a números distintos corresponden sucesores distintos, es decir, notamos la existencia de una función inyectiva

$$\sigma : \mathbb{N} \rightarrow \mathbb{N}$$

que asigna a cada número natural n , un único elemento $\sigma(n)$ - también en \mathbb{N} - llamado “el sucesor de n ”, y segundo, el hecho de que aplicando iteradamente el proceso de “tomar sucesores” a partir de un elemento distinguido “cero”, se puede alcanzar -teóricamente- cualquier número natural m dado de antemano.

Es decir: \mathbb{N} es, precisamente, la colección de números que se obtienen a partir de un objeto inicial cero, y pasando de cada número n a otro, determinado en forma unívoca $\sigma(n)$ llamado el sucesor de n .

Estas observaciones -que tienen que estar incluídas en toda teoría axiomática para \mathbb{N} - constituyen la base sobre la que G. Peano presentó su teoría, que toma como conceptos no definidos, al conjunto total \mathbb{N} , y a su elemento inicial cero. La única relación primitiva que consideró es la función sucesor. Los axiomas, que se conocen con su nombre, son los siguientes:

3.2 Los axiomas de Peano

Axioma 15 . *0 es un número natural.*

Axioma 16 . *Si n es un número natural, existe un único número natural $\sigma(n)$ que es el sucesor de n .*

Axioma 17 . *Para todo número natural n , $\sigma(n) \neq 0$.*

Axioma 18 . *Para todos los números naturales n y m , si $\sigma(n) = \sigma(m)$ entonces $n = m$.*

Axioma 19 . *Si S es un subconjunto de \mathbb{N} tal que $0 \in S$ y $\sigma(n) \in S$ para cada $n \in S$, entonces $S = \mathbb{N}$.*

Que pueden resumirse sustituyendo los primeros cuatro por:

$$\exists \sigma : \mathbb{N} \rightarrow \mathbb{N} \setminus \{0\} \text{ biyectiva, y}$$

conservando el 5º axioma.

Cuando se presupone la existencia de \mathbb{N} , y no se desea hacer consideraciones sobre lo que sus elementos puedan ser, se puede partir de aquí y proceder a definir las relaciones usuales entre ellos y a demostrar sus propiedades. Posteriormente, con base en este conjunto, por medio de ampliaciones adecuadas, pueden construirse los enteros, los racionales, los reales y los complejos. Sin embargo, a la luz de las paradojas de la teoría de los conjuntos que afectan a la mayoría de los sistemas axiomáticos informales se ha intentado comenzar de otra manera, estableciendo un sistema lógico lo suficientemente rico para construir con él una teoría axiomática para los conjuntos y proceder, a partir de ésta, a la formalización de las estructuras numéricas fundamentales. Un sistema axiomático informal es el que presupone alguna teoría de conjuntos y algún sistema lógico, sin establecerlos (formalmente).

Siguiendo este segundo camino, y apoyándonos en el material que hemos considerado en los capítulos anteriores, procederemos a construir un modelo para \mathbb{N} , en el que -como ya se dijo-, los postulados de Peano resultan consecuencia de las definiciones, o son teoremas que se demuestran casi trivialmente.

Surge aquí el problema de determinar si ambas presentaciones -la axiomática de Peano y la que se deriva de la teoría de conjuntos- son equivalentes en el sentido de producir estructuras comparables. Es un resultado conocido el que asegura que “excepto por isomorfismos, existe un único sistema numérico que satisface los axiomas de Peano” que por esta razón, resultan categóricos y que por lo tanto muestran que la diferencia entre dos modelos cualesquiera para \mathbb{N} , es simplemente cuestión de representación.

3.3 Construcción

Recordemos:

Definición 49 . *Si x es un conjunto, el sucesor de x es el conjunto*

$$x \cup \{x\} =: \sigma(x).$$

Observación 43 . *Por definición $A \in \sigma(A)$. Además, $A \subseteq \sigma(A)$.*

Si consideramos el conjunto vacío, \emptyset , entonces su sucesor es

$$\emptyset \cup \{\emptyset\} = \{\emptyset\},$$

un conjunto con un solo elemento.

Si consideramos el conjunto $\{2, 4, \{1, 3\}\}$, entonces su sucesor es

$$\{2, 4, \{1, 3\}, \{2, 4, \{1, 3\}\}\},$$

un conjunto con cuatro elementos.

Definición 50 . *Un conjunto A , cuyos elementos son conjuntos, es inductivo si y sólo si:*

1. $\emptyset \in A$ y
2. $x \in A \Rightarrow \sigma(x) \in A$, para cada conjunto x .

El axioma de infinito que dice que: existe al menos un conjunto inductivo.

Axioma 20 (de infinito) . *Existe un conjunto inductivo.*

Puede verse que si A es un conjunto inductivo, entonces cuenta entre sus elementos con los de la cadena

$$\{\emptyset, \sigma(\emptyset), \sigma(\sigma(\emptyset)), \dots\}$$

y que esta cadena forma, por si sola, un conjunto que es inductivo y que por ser subconjunto de todo conjunto inductivo resulta mínimo en el sentido de la contención. Finalmente, por su semejanza con la sucesión $0, 1, 2, \dots$ observamos que podría ser un buen candidato para modelar \mathbb{N} . Una manera de tener un modelo adecuado, consiste en construir un conjunto que siendo inductivo sea mínimo en el sentido de la contención y nos gustaría construirlo tomando la intersección de todos los conjuntos inductivos.

Es evidente que la intersección I de una familia A_i de conjuntos inductivos, es un conjunto inductivo.

En efecto, 0 está en cada intersecando, luego está en I . Y si x es un elemento de la intersección es porque x está en cada miembro de la familia y por lo tanto, $\sigma(x)$ también.

Además I está contenido en cada A_i y en ese sentido es mínimo, pero como la colección de todos los conjuntos inductivos podría no ser un conjunto, no podemos basar la construcción en “tomar la intersección de todos ellos” como era nuestro deseo.

El proceso que seguiremos consistirá en tomar un conjunto inductivo A arbitrario, y trabajar con él.

Sea pues A un conjunto inductivo dado -que existe en vista del axioma correspondiente-. Sea β la colección de los subconjuntos de A que son inductivos, y \mathbb{N}_A la intersección de todos ellos. Observamos que el axioma de las partes nos asegura que es un conjunto -de conjuntos- y que por lo tanto \mathbb{N}_A está bien definido, es inductivo y está contenido en todo subconjunto inductivo de A . Para hacer ver que es el que se necesita, demostraremos el siguiente.

Teorema 26 . \mathbb{N}_A está contenido en todo conjunto inductivo.

Demostración. Sea B un conjunto inductivo, y considérese

$$C = A \cap B.$$

Entonces C está contenido en B y también está contenido en A y es inductivo, luego

$$\mathbb{N}_A \subseteq C \subseteq B.$$

■

Corolario 3 . \mathbb{N}_A es el menor de los conjuntos inductivos, y por lo tanto, es el único con esta propiedad es independiente del conjunto A con el que principiamos, por lo que, en lo sucesivo suprimiremos el subíndice A .

Conclusión: Existe un único conjunto inductivo que está contenido en todo conjunto inductivo, al que llamaremos \mathbb{N} .

Teorema 27 . En \mathbb{N} valen los axiomas de Peano.

Demostración. En efecto, definiendo al conjunto vacío - cuando se considera como número - como “cero” e identificando a la función

$$\begin{aligned} \sigma : \mathbb{N} &\rightarrow \mathbb{N} \\ x &\longmapsto x \cup \{x\} \end{aligned}$$

con la función “sucesor” de los conjuntos, los axiomas 1 y 2 -cero es un número natural y el sucesor de cada número natural es un número natural - son consecuencias directas de la definición de conjunto inductivo.

El axioma 17 -cero no es sucesor de número natural alguno- se sigue de que cero es vacío, mientras que de la definición, todo sucesor tiene al menos un elemento.

De esta misma definición es inmediato que si el sucesor de n es igual al sucesor de m ,

$$n \cup \{n\} = m \cup \{m\},$$

entonces n es igual a m . (Note que si $n \neq m$, entonces $n \in m$ y $m \in n$, cosa que iría en contra del Axioma de regularidad).

Finalmente, el axioma 19 es consecuencia directa del hecho de que \mathbb{N} está contenido en todo inductivo.

En efecto, las hipótesis de este axioma, $S \subseteq \mathbb{N}$, $0 \in S$ y $n \in S \Rightarrow \sigma(n) \in S$ garantizan que S es inductivo y por lo tanto $\mathbb{N} \subseteq S$. ■

El axioma 19 es el fundamento del llamado “Primer Principio de Inducción” o “quinto postulado de Peano”, que dice que si P es una propiedad que cada número natural puede o no tener pero la tiene el cero, y además es

hereditaria -cada vez que la tenga un número, la tiene necesariamente el que le sigue- entonces todo número natural tiene la propiedad P .

Recordemos que una propiedad P para los elementos de un conjunto A es una función

$$P : A \rightarrow \{0, 1\}$$

tal que $P(x) = 1$ (o simplemente $P(x)$) se interpreta diciendo que “ x tiene la propiedad P ” mientras que $P(x) = 0$ (o también $\neg P(x)$) dice que “ x no la tiene”.

El primer principio de inducción asegura que:

Si P es una propiedad para \mathbb{N} tal que

1. $P(0)$.

2. $\forall k \in \mathbb{N}, P(k) \Rightarrow P(\sigma(k))$,

entonces $\forall n \in \mathbb{N}, P(n)$.

En efecto, si

$$S = \{n \in \mathbb{N} \mid P(n)\},$$

por (1), $0 \in S$ y (2) dice que

$$\forall k \in \mathbb{N}, k \in S \Rightarrow \sigma(k) \in S,$$

y por lo tanto $S = \mathbb{N}$.

Recuerde que cada propiedad P - o función característica- para un conjunto A , define a un subconjunto B , a saber el de los elementos de A que tienen la propiedad P y que recíprocamente, a cada subconjunto B de A corresponde -al menos- la propiedad de pertenecer a B .

Explícitamente, si $B \subseteq A$,

$$P_B : A \rightarrow \{0, 1\}$$

es la función (característica) definida por:

$$P_B(n) = \begin{cases} 0 & \text{si } n \notin B \\ 1 & \text{si } n \in B. \end{cases}$$

Observe que si para un conjunto $S \subseteq \mathbb{N}$, su función característica tiene las propiedades

1. $P(0)$

2. $\forall k, P(k) \Rightarrow P(\sigma(k))$,

entonces, en vista del 5º postulado, S resulta igual a \mathbb{N} , y que si en lugar del 5º se postulara esta última afirmación, el mencionado 5º postulado sería un teorema inmediato y en vista de la obvia equivalencia entre el 5º postulado de Peano y el primer principio de inducción, nos parece justificada la frecuente costumbre de intercambiar sus nombres, -costumbre que en lo sucesivo, usaremos libremente.

El 5º postulado de Peano (o primer principio de inducción) se utiliza principalmente para definir expresiones como x^n ó $n!$ que incluyen variables numéricas que toman valores en \mathbb{N} , así como para demostrar que ciertas propiedades se cumplen para todos los números naturales. Cuando se usa para definir, las definiciones que resultan se llaman “recursivas” y las demostraciones son “demostraciones por inducción”. Dedicaremos los siguientes apartados para cada una de estas aplicaciones:

3.4 Definiciones recursivas

Supongamos que se desea explicar el significado de una sucesión de expresiones $\{\eta(n)\}$ -una para cada número natural- y ante la imposibilidad de definir explícitamente cada una de ellas, deseáramos recurrir al 5º. postulado. Lo que tendríamos que hacer, siguiendo el método canónico (el que usó Peano) sería:

1º) Precisar el significado $\eta(0)$ (base), y

2º) suponiendo definida la expresión $\eta(k)$, diseñar alguna manera de hacer explícito el significado de $\eta(\sigma(k))$ (paso inductivo). Podríamos suponer este “diseño” como una función

$$f : X \rightarrow X$$

del conjunto de significados en él mismo.

Sin embargo, algunas veces no basta saber de que punto se parte ($\eta(k)$) para describir al siguiente punto ($\eta(\sigma(k))$) sino que es preciso también tomar en cuenta el orden del paso que se está dando. (Así por ejemplo para la función “factorial” una vez que se ha definido. $\eta(0) = 0!$ como 1, se procede a construir $\eta(1) = 1!$ Pasando de 1 a 1. Es decir que si se usara una sola función $f : X \rightarrow X$ para transformar un significado en otro, tendríamos que aceptar que $f(1) = 1$ pero entonces $2! = \eta(2)$ sería $f(\eta(1)) = f(1) = 1$ lo que evidentemente no es lo que deseamos).

En estos casos puede usarse una función

$$\varphi : (\mathbb{N} \times X) \rightarrow X$$

(que en este caso sería $\varphi(0, 1) = 1$, $\varphi(1, 1) = 2$ y en general

$$\varphi(\sigma(k), \sigma(n)) = \sigma(k)\varphi(k, n)).$$

O bien, una familia de funciones $\{f_i : X \rightarrow X\}_{i \in \mathbb{N}}$ de modo que $\eta(\sigma(k)) = f_k(k)$.

Esta segunda forma, es la que corresponde al Teorema de recursión generalizada -recursión fuerte- que afirma que :

Teorema 28 (Recursión generalizada) . (o recursión fuerte). *Para todo conjunto X con un elemento distinguido x_0 y toda familia $\{f_i : X \rightarrow X\}_{i \in \mathbb{N}}$ de funciones, existe una única función*

$$\eta : \mathbb{N} \rightarrow X$$

tal que

$$\eta(0) = x_0$$

$$\text{y para toda } k \in \mathbb{N}, \eta(\sigma(k)) = f_k(\eta(k)).$$

Nótese que lo que se requiere es:

1. Un elemento “distinguido” $x_0 \in X$ que se tomará como el significado (0).
2. Una familia $\{f_i : X \rightarrow X\}_{i \in \mathbb{N}}$ de funciones tal que para toda $k \in \mathbb{N}$, $\eta(\sigma(k)) = f_k(\eta(k))$.

Explícitamente, si $\eta(0) = x_0$, entonces

$$\eta(1) = f_0(x_0) = x_1,$$

$$\eta(2) = f_1(x_1) = x_2,$$

y así sucesivamente.

Daremos la demostración de este teorema al final de capítulo.

A reserva de discutir posteriormente (ver sección 3.10 y apéndice 3.12) la validez de este procedimiento de definición, daremos dos ejemplos que lo ilustran.

Consideremos en detalle el problema de explicar lo que debe entenderse por g^n en donde g es un elemento cualquiera de un conjunto con una operación asociativa con elemento neutro, y n es un número natural. G podría ser el conjunto \mathbb{R}^* de los números reales diferentes de cero, la operación, en este caso podría ser la multiplicación usual y por supuesto, el neutro, el uno.

Lo que se tiene en mente es formalizar por medio de una definición recursiva, los significados de las expresiones de la forma g^n , tales como g^2 , que desde luego, queremos interpretar como el producto (en G) de 2 factores iguales a g ; g^3 como $g^2 \cdot g$, etcétera.

Decidamos primero cómo debe definirse g^0 . De acuerdo a nuestro deseo, g^n debe corresponder al producto de n factores iguales a g y, por lo tanto, debemos escoger g^0 como el “producto vacío” (resultado de multiplicar entre sí, cero factores).

Es fácil convencerse de que en toda estructura algebraica con una operación asociativa con neutro, el producto vacío debe definirse como el neutro de la operación, con objeto de que la asociatividad funcione a ultranza:

$$\begin{aligned} ((ab)c)d &= (ab)(cd) = (a(bc))d = \dots = (abc)(d) = \\ &= (abcd)() = abcd = (abcd)(1). \end{aligned}$$

Así por ejemplo, para los números naturales, la suma vacía es cero y el producto vacío es 1. Entre los subconjuntos de un conjunto dado X , la unión vacía es \emptyset y la intersección vacía es el total ($A \cap X = A, \forall A \subseteq X$).

(Pedimos al lector que se convenza de la conveniencia de aceptar los argumentos anteriores), y en vista de ellos definimos $g^0 = e$ (el neutro de G). Se completa la definición como sigue:

$$g^{\sigma(k)} = g^k \cdot g.$$

En este caso, $\forall k \in \mathbb{N}$, f_k es la función $f : G \rightarrow G$ que nos permite pasar de un significado (g^k) al siguiente ($g^{\sigma(k)}$) que viene a ser la multiplicación por g .

Es decir: $f_k(x) = f(x) = xg$.

Como un segundo ejemplo, considérese el problema de construir la tabla de sumar del 2. Es decir, se desea definir una función $s_2 : \mathbb{N} \rightarrow \mathbb{N}$ (en donde, en lugar de $s_2(n)$ se usa la conocida expresión $2 + n$). Se define:

$$2 + 0 = 2$$

$$2 + \sigma(k) = \sigma(2 + k).$$

Entonces

$$2 + 1 = 2 + \sigma(0) = \sigma(2 + 0) = \sigma(2) = 3$$

$$2 + 2 = 2 + \sigma(1) = \sigma(2 + 1) = \sigma(3) = 4,$$

etcétera.

Ahora nuestra función de “significados”, también única, es la función sucesor. En efecto, para precisar el significado de la expresión $2 + \sigma(k)$, se aplica esta función -sucesor- al significado $2 + k$.

Con objeto de continuar nuestro estudio de los números naturales, hacemos notar que los ejemplos anteriores son casos particulares del uso del Teorema de recursión débil que asegura que para cada conjunto X con elemento distinguido x_0 , y para cada función $f : X \rightarrow X$, existe una única sucesión

$$\eta : \mathbb{N} \rightarrow \mathbb{X}$$

tal que :

$$\eta(0) = x_0 \text{ y } \forall k \in \mathbb{N}, \eta(\sigma(k)) = f(\eta(k)).$$

En Álgebra es frecuente utilizar diagramas para representar colecciones de conjuntos y funciones, así un esquema como el que sigue:

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ \downarrow h & & \downarrow k \\ C & \xrightarrow{g} & D \end{array}$$

es otra manera de decir que

$$f : A \rightarrow B, h : A \rightarrow C, g : C \rightarrow D \text{ y } k : B \rightarrow D$$

son funciones para las que están definidas las composiciones $g \circ h$ y $k \circ f$. Se dice que “el diagrama conmuta” si

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

-la “llegada de A a D , es independiente del camino”-. Convención que se considera válida para diagramas más complicados.

Se usa el diagrama $\{\cdot\} \xrightarrow{a} X$ para designar a la única función de un conjunto “ $\{\cdot\}$ ” con un solo elemento “ \cdot ”, en X , tal que la imagen de \cdot es a .

Con este acuerdo, el Teorema de recursión débil equivale a afirmar que commuta el siguiente diagrama:

$$\begin{array}{ccc}
 \mathbb{N} & \xrightarrow{\sigma} & \mathbb{N} \\
 \bar{0} \nearrow & \downarrow \eta & \downarrow \eta \\
 \{0\} & & \\
 \bar{a} \searrow & & \\
 X & \xrightarrow{f} & X
 \end{array}$$

Y el de recursión fuerte:

$$\begin{array}{ccc}
 \mathbb{N} & \xrightarrow{\sigma} & \mathbb{N} \\
 \bar{0} \nearrow & \downarrow (Id_{\mathbb{N}}, \eta) & \downarrow (Id_{\mathbb{N}}, \eta) \\
 \{0\} & & \\
 \xrightarrow{(0, x_0)} & \downarrow & \\
 \mathbb{N} \times X & \xrightarrow{F} & \mathbb{N} \times X
 \end{array}$$

$$(k, x) \xrightarrow{\quad} (\sigma(k), f_k(x))$$

Como lo indica el diagrama,

$$F(k, x) = (\sigma(k), f_k(x))$$

y

$$(Id_{\mathbb{N}}, \eta)(k) = (k, \eta(k)).$$

Que el diagrama commute significa que

1. $(0, \eta(0)) = (Id_{\mathbb{N}}, \eta)(0) = (0, x_0)$, de donde $\eta(0) = x_0$ y
2. $F \circ (Id_{\mathbb{N}}, \eta)(k) = F(k, \eta(k)) = (\sigma(k), f_k(\eta(k))) = (Id_{\mathbb{N}}, \eta) \circ \sigma(k) = (\sigma(k), \eta(\sigma(k)))$, es decir

$$\eta(\sigma(k)) = f_k(\eta(k)).$$

Ilustramos el uso del primer diagrama con los ejemplos siguientes:

Ejemplo 90 . Sean. $X = \mathbb{N}$, $x_0 = 7$, y $f : \mathbb{N} \rightarrow \mathbb{N}$ la función sucesor. Entonces:

$$\eta(0) = 7$$

$$\eta(\sigma(k)) = \sigma(\eta(k))$$

como puede leerse en el dibujo:

$$\begin{array}{ccc} \mathbb{N} & \xrightarrow{\sigma} & \mathbb{N} \\ \{0\} \xrightarrow[{}_{x_0}]{}^7 & \downarrow \eta & \downarrow \eta \\ \mathbb{N} & \xrightarrow{f=\sigma} & \mathbb{N}. \end{array}$$

Nótese que en este caso,

$$\eta(0) = 7$$

$$\eta(1) = \sigma(\eta(0)) = \sigma(7) = 8$$

$$\eta(2) = \sigma(8) = 9$$

y que, por lo tanto η no es otra cosa que la tabla de sumar del 7.

$$\begin{array}{ccccccc} 0 & \mapsto & 1 & \mapsto & 2 & & \\ \mathbb{N} & \xrightarrow{\sigma} & \mathbb{N} & \xrightarrow{\sigma} & \mathbb{N} & & \\ \downarrow \eta & & \downarrow \eta \\ \mathbb{N} & \xrightarrow{f=\sigma} & \mathbb{N} & \xrightarrow{f=\sigma} & \mathbb{N} & & \\ 7 & \mapsto & 8 & \mapsto & 9 & & \end{array}$$

En efecto si en vez de $\eta()$ se escribe $7 + ()$, se obtiene:

$$7 + 0 = 7$$

$$7 + 1 = 8$$

$$7 + 2 = 9$$

En general,

$$7 + \sigma(k) = \sigma(7 + k)$$

que junto con $7 + 0 = 7$ constituye la definición recursiva de la tabla de sumar del siete.

Es obvio que para definir la tabla de sumar de cualquier número natural n , puede usarse el procedimiento anterior, definiendo x_0 como n .

Hemos demostrado el siguiente:

Teorema 29 . Para cada número natural n , existe una única función

$$s_n : \mathbb{N} \rightarrow \mathbb{N}$$

tal que

$$s_n(0) = n,$$

y

$$\forall k \in \mathbb{N}, s_n(\sigma(k)) = \sigma(s_n(k)).$$

Es decir $n + 0 = n$ y $\forall k \in \mathbb{N}, n + \sigma(k) = \sigma(n + k)$.

Ejemplo 91 . Supongamos ahora que $X = \mathbb{N}$, y $f : \mathbb{N} \rightarrow \mathbb{N}$ es la función sumar 3. $f(n)$ significa $n + 3$, que suponemos ya conocida.

Si escogemos para x_0 el valor 0, resulta:

$$\eta(0) = 0$$

y

$$\forall k \in \mathbb{N}, \eta(\sigma(k)) = \eta(k) + 3.$$

o sea:

$$\eta(0) = 0$$

$$\eta(1) = \eta(0) + 3 = 0 + 3 = 3$$

$$\eta(2) = \eta(1) + 3 = 3 + 3 = 6$$

$$\eta(3) = \eta(2) + 3 = 6 + 3 = 9$$

que, como se ve, es la tabla de multiplicar del 3.

En efecto, si en vez que $\eta(\)$ se usa $(\) \cdot 3$, resulta:

$$0 \cdot 3 = 0$$

$$1 \cdot 3 = 0$$

$$2 \cdot 3 = 6$$

y en general,

$$\sigma(k) \cdot 3 = k \cdot 3 + 3.$$

Como en el ejemplo 91, podemos generalizar el resultado para cada número natural n , sustituyendo la $f : \mathbb{N} \rightarrow \mathbb{N}$ por la función de sumar de n , y en este caso, $\eta(0) = 0$ y $\forall k \in \mathbb{N}, \eta(\sigma(k)) = \eta(k) + n$.

Cambiando η por “ p_n ”, queda demostrado que:

Teorema 30 . Para cada número natural n , existe una única función

$$p_n : \mathbb{N} \rightarrow \mathbb{N}$$

-multiplicar por n - tal que

$$p_n(0) = 0$$

y

$$\forall k \in \mathbb{N}, p_n(\sigma(k)) = p_n(k) + n.$$

O bien, $n \cdot 0 = 0$, y $\forall n \in \mathbb{N}, n \cdot (\sigma(k)) = n \cdot k + n$.

Ejemplo 92 . Sea $X = \mathbb{N}$, $f = p_5$.

Tomemos una x_0 provisional -a reserva de cambiarla posteriormente- y hagamos funcionar nuestra “máquina recursiva”.

Así,

$$\eta(0) = x_0 \text{ y } \eta(\sigma(k)) = \eta(k) \cdot 5.$$

Entonces:

$$\eta(0) = x_0$$

$$\eta(1) = x_0 \cdot 5$$

$$\eta(2) = (x_0 \cdot 5) \cdot 5 = x_0 \cdot 25$$

$$\eta(3) = (x_0 \cdot 25) \cdot 5 = x_0 \cdot 125.$$

Notemos aquí la conveniencia de definir x_0 como 1 y en ese caso resulta ser la “exponencial base 5” o sea

$$5^0 = 1, 5^{\sigma(k)} = 5^k \cdot 5.$$

Como antes, generalizamos para todo natural n distinto de cero, usando

$$f = p_n,$$

e insistimos en que $\forall a \neq 0, a^0 = 1$ por definición.

3.5 Demostraciones inductivas

El 5º postulado de Peano suele utilizarse de manera análoga a la forma en que se usó en las definiciones recursivas, para demostrar proposiciones de la forma:

“todos los números naturales tienen cierta propiedad P ”.

Supongamos que P es una propiedad que cada número natural puede o no tener. (Ser primo, ser par, ...) y que deseamos comprobar que todos los números naturales la tienen. Simbolicemos con $p(n)$ la expresión

“ n tiene la propiedad p ”.

Si ahora demostramos que

1. $p(0)$ (cero tiene la propiedad p), y que:
2. $\forall k \in \mathbb{N}, p(k) \Rightarrow p(\sigma(k)) \in \mathbb{N}$ (La propiedad p es “hereditaria”), habremos demostrado que

$$\forall n \in \mathbb{N}, p(n),$$

que es lo que deseábamos hacer.

En efecto, si

$$S = \{n \in \mathbb{N} \mid p(n)\},$$

por 1), $0 \in S$ y por 2),

$$\forall k \in \mathbb{N}, k \in S \Rightarrow \sigma(k) \in S,$$

y por lo tanto (5º postulado) $S = \mathbb{N}$.

Observemos una vez más que en 2) lo que se afirma es la veracidad de una implicación, a saber:

$$p(k) \Rightarrow p(\sigma(k)),$$

lo que no dice que $p(k)$ es cierta ni que $p(\sigma(k))$ lo sea sino que asegura que la propiedad “ p ” se hereda al pasar de un número k cualquiera a su sucesor. No se comete el error de “circularidad” cuando se supone $p(k)$ (hipótesis de inducción) para demostrar $p(\sigma(k))$, sino que se está haciendo uso legítimo del meta-teorema de la deducción, (ver sección 1.8).

Queremos repetir aquí, que al igual que en el caso de las definiciones recursivas, el método propuesto (usar el 5º postulado), en aquel lugar para

“definir”, y en éste para “demostrar”, de ninguna manera es único. Siendo \mathbb{N} un conjunto, como lo es, pueden demostrarse algunas propiedades universales que se refieren a sus elementos, con la técnica usual de la teoría de los conjuntos, a saber:

Se toma un elemento arbitrario del conjunto de que se trate, del cual sólo es válido suponer que tiene la propiedad que lo caracteriza como elemento de éste. Si a partir de esta única suposición podemos concluir que el elemento considerado tiene la propiedad que nos interesa, aceptaremos haber demostrado que todos la tienen.

En la aritmética de \mathbb{N} , son frecuentes las ocasiones en que se desea demostrar alguna proposición (abierta) en la que aparecen dos (o más) números naturales m y n , y que debe cumplirse para cualesquiera dos de ellos. ($\forall m, n \in \mathbb{N}, p(n, m)$). En estos casos se puede utilizar un procedimiento “híbrido” que consiste en fijar uno de ellos y hacer inducción sobre el otro. En este caso, lo que se está haciendo puede resumirse de la manera siguiente:

Proposición 13 . *Para demostrar que*

$$\forall m, n \in \mathbb{N}, p(n, m),$$

se toma $m \in \mathbb{N}$ un número natural arbitrario, (del que sólo suponemos que es un número natural). Entonces se demuestra por medio de la inducción, que $\forall n \in \mathbb{N}, p(n, m)$.

Obviamente, este procedimiento, no es inductivo “puro” y sin embargo suele dársele también este nombre y es común verlo aparecer ejemplificando el uso de la inducción para hacer demostraciones. Por supuesto que en la mayor parte de estos casos, también se puede proceder haciendo una doble (triple o múltiple) inducción, pero queremos enfatizar aquí, que el tratar de mantenerse dentro de una única línea de razonamiento (geometría sin álgebra; topología sin análisis, y en este caso “inducción pura”), lejos de parecernos una virtud, la consideramos una limitación y como tal, un defecto.

3.6 Conjuntos transitivos

Recordemos la definición del conjunto transitivo de la página 78.

Definición 51 . *Un conjunto X es transitivo si*

$$Z \in Y \in X \implies Z \in X.$$

O equivalentemente

$$Y \in X \implies Y \subseteq X.$$

Ejemplo 93 . Consideremos el conjunto

$$3 = \{0, 1, 2\}$$

Note que cada uno de los elementos de 3 es también un subconjunto de 3.

No ejemplo 94 . En cambio, $\{2, 3\}$ no es un conjunto transitivo porque $2 \in \{2, 3\}$, pero $2 \notin \{2, 3\}$ ($1 \in 2$ pero $1 \notin \{2, 3\}$).

Ejercicio 137 . Demostrar que el menor subconjunto transitivo que contiene a $\{2, 3\}$ es $\{0, 1, 2, 3\}$.

Ejercicio 138 . ¿Cuál es el menor conjunto transitivo que contiene a $\{2, \{1, 3\}\}$?

Teorema 31 . Todo número natural es un conjunto transitivo.

Demostración. Denotemos $A = \{n \in \mathbb{N} \mid n \text{ es transitivo}\}$. Veremos que este conjunto es inductivo.

i) Base de la inducción.

Como \emptyset no tiene elementos, la proposición $y \in \emptyset$ siempre es falsa. Así que la proposición

$$y \in \emptyset \implies y \subseteq \emptyset,$$

es cierta. Por lo tanto, \emptyset es transitivo.

ii) Paso iInductivo.

Supongamos que $k \in \mathbb{N}$ es transitivo, queremos demostrar que $\sigma(k)$ es transitivo también.

Si $a \in b \in \sigma(k)$, como $\sigma(k) = k \cup \{k\}$, entonces $(b \in k) \vee (b = k)$.

Si $b \in k$, entonces $a \in b \in k$, y como k es transitivo por hipótesis, tendríamos que $a \in k$. Pero como $k \subseteq \sigma(k)$, entonces $a \in \sigma(k)$.

Si $b = k$, entonces $a \in k \subseteq \sigma(k)$.

En cualquier caso, $a \in \sigma(k)$. Es decir que

$$a \in b \in \sigma(k) \implies a \in \sigma(k),$$

así que $\sigma(k)$ es un conjunto transitivo. ■

Teorema 32 . *El conjunto de los naturales es un conjunto transitivo. (Por favor note la diferencia entre las afirmaciones: \mathbb{N} es transitivo, cada elemento de \mathbb{N} es transitivo).*

Demostración. Necesitamos comprobar que

$$a \in b \in \mathbb{N} \implies a \in \mathbb{N}, \forall b \in \mathbb{N} \quad (3.1)$$

es decir, que los elementos de los naturales son también naturales.

Consideremos 3.1 como una proposición que se dice del natural b . Abreviemos esta proposición $P(b)$.

Notemos que lo que queremos demostrar es que

$$B =: \{n \in \mathbb{N} \mid P(n)\}$$

es todo el conjunto \mathbb{N} . (Es decir, con esto demostraremos que el conjunto de los naturales que tienen la propiedad P son todos).

i) Base de la inducción.

Escribamos la proposición $P(\emptyset)$:

$$a \in \emptyset \in \mathbb{N} \implies a \in \mathbb{N}.$$

Esta proposición es verdadera puesto que

$$a \in \emptyset \in \mathbb{N} \quad \text{es una abreviatura de} \quad (a \in \emptyset) \wedge (\emptyset \in \mathbb{N}),$$

que es falsa.

ii) Paso inductivo.

Queremos demostrar que $P(k) \implies P(\sigma(k))$, $\forall k \in \mathbb{N}$.

Supongamos que $k \in \mathbb{N}$ a es tal que $P(k)$. Es decir que los elementos de k son números naturales. Queremos demostrar que $P(\sigma(k))$ es verdadera.

$$P(\sigma(k)) : a \in \sigma(k) \in \mathbb{N} \implies a \in \mathbb{N}.$$

Supongamos que $a \in \sigma(k)$, queremos demostrar que a es un natural.

$$a \in \sigma(k) = k \cup \{k\} \implies (a \in k) \vee (a = k).$$

Si $a \in k$, entonces $a \in \mathbb{N}$, por hipótesis.

Si $a = k$, entonces $a \in \mathbb{N}$, pues $k \in \mathbb{N}$. ■

3.7 Conjuntos infinitos y conjuntos finitos

Definición 52 . *Se dice que un conjunto X es infinito, si existe una biyección entre X y uno de sus subconjuntos propios. (X es infinito si existe $X \xrightarrow{f} Y$ biyectiva, $Y \subsetneq X$).*

Recordemos que un conjunto X es finito, si no es infinito.

Ejemplo 95 . *El conjunto vacío es un conjunto finito.*

Demostración. Simplemente notemos que por definición, para que un conjunto X sea infinito, necesita tener subconjuntos propios. Como \emptyset no tiene subconjuntos propios, entonces no puede ser infinito. ■

Teorema 33 . *Cada número natural es finito.*

Demostración. Denotemos

$$F = \{n \in \mathbb{N} \mid n \text{ es finito}\},$$

queremos demostrar que $F = \mathbb{N}$.

i) Base de la inducción.

Como hemos demostrado en el ejemplo de arriba, $\emptyset \in F$.

ii) Paso inductivo.

Supongamos que k es finito, queremos demostrar que $\sigma(k)$ también es finito. Esto es exactamente lo que se demuestra en el teorema 16, de la página 110. ■

3.8 El conjunto de los naturales es un conjunto infinito

\mathbb{N} es infinito en vista del siguiente teorema:

Teorema 34 . *La función*

$$\sigma : \mathbb{N} \longrightarrow \mathbb{N} \setminus \{0\}$$

es una biyección.

Demostración. Notemos primero que la función σ es inyectiva:

Supongamos que $\sigma(n) = \sigma(m) = m \cup \{m\}$.

Como $m \in \sigma(m)$, entonces $m \in \sigma(n) = n \cup \{n\}$. Por lo tanto $m \in n \vee m = n$.

Por simetría, $n \in m \vee n = m$.

En la observación 17 notamos que ningún conjunto puede ser elemento de sí mismo. Por otra parte, no puede suceder que $[(m \in n) \wedge (n \in m)]$, pues esto implicaría que $m \in m$, pues los naturales como ya vimos, son conjuntos transitivos).

La única posibilidad es $n = m$.

Veamos ahora que la función $\sigma : \mathbb{N} \longrightarrow \mathbb{N} \setminus \{0\}$ es suprayectiva.

Para empezar, notemos que $0 \notin \text{Im}(\sigma)$:

Esto se debe a que $0 = \emptyset$, pero $k \in \sigma(k)$, $\forall k \in \mathbb{N}$.

Demostraremos que $\{0\} \cup \text{Im}(\sigma) = \mathbb{N}$, demostrando que $\{0\} \cup \text{Im}(\sigma)$ es un subconjunto inductivo del conjunto de los números naturales.

Base de la inducción: es claro que $0 \in \{0\} \cup \text{Im}(\sigma)$.

Paso inductivo: si $k \in \{0\} \cup \text{Im}(\sigma)$, entonces es obvio que $\sigma(k) \in \text{Im}(\sigma)$.

Por lo tanto, $[(\{0\} \cup \text{Im}(\sigma)) = \mathbb{N}] \wedge [\text{Im}(\sigma) \subseteq \mathbb{N} \setminus \{0\}]$, de aquí tenemos que todo natural $\neq 0$ pertenece a $\text{Im}(\sigma)$, es decir que $\mathbb{N} \setminus \{0\} \subseteq \text{Im}(\sigma)$. Como también tenemos la inclusión recíproca, podemos concluir que

$$\text{Im}(\sigma) = \mathbb{N} \setminus \{0\}.$$

■

Definición 53 . Se dice que el conjunto A tiene a lo más tantos elementos como el conjunto B , si existe una función inyectiva de A a B .

En esta situación, escribiremos $|A| \leq |B|$.

Con otras palabras: el cardinal del conjunto A , $|A|$, es menor o igual que el cardinal del conjunto B , $|B|$.

Desde luego, la definición anterior está inspirada en el hecho de que si $A \subseteq B$, es obvio que $|A| \leq |B|$, y en el hecho de que una función inyectiva es una especie de generalización de la función inclusión.

Recordemos que dos conjuntos A, B tienen el mismo número elementos, si existe una biyección entre ambos conjuntos.

Ejemplo 96 . Un conjunto X es infinito cuando tiene el mismo número de elementos que uno de sus subconjuntos propios. Por ejemplo, \mathbb{N} y $\mathbb{N} \setminus \{0\}$ tienen el mismo número de elementos.

Como hemos notado antes, una función inyectiva $f : A \rightarrow B$ tiene un inverso derecho $g : B \rightarrow A$, que resulta ser suprayectiva. Recíprocamente, dada una función suprayectiva $g : B \rightarrow A$, (la doble punta de la flecha anterior expresa que la función es suprayectiva) existe una función inyectiva $f : A \rightarrow B$, que es inverso por la izquierda de g . (Como ya hemos notado, esta última afirmación es equivalente al axioma de elección).

De esta manera, tenemos que son equivalentes las afirmaciones:

1. $|A| \leq |B|$.
2. $\exists f : A \rightarrow B$.
3. $\exists g : B \rightarrow A$.

Proposición 14 . *Las siguientes afirmaciones son equivalentes para una función $f : A \rightarrow A$, A un conjunto **finito**.*

1. f es biyectiva.
2. f es inyectiva.
3. f es suprayectiva.

Demostración. Obviamente 1) \implies 2) \wedge 3).

Debería ser claro que basta demostrar la equivalencia entre 2) y 3).

Supongamos 2). Si f no fuera suprayectiva, entonces $f(A)$ sería un subconjunto propio de A , y así

$$\begin{aligned} f|^{f(A)} : A &\rightarrow f(A) \\ a &\mapsto f(a) \end{aligned}$$

sería una biyección entre A y el subconjunto propio de A . Esto contradiría el hecho de que A es finito.

Esta contradicción muestra que si f es inyectiva, entonces también tiene que ser suprayectiva.

Supongamos 3), es decir, supongamos que $A \xrightarrow{f} A$ es suprayectiva. Tomemos $A \xrightarrow{g} A$ un inverso derecho para f , por el inciso anterior, g tiene que ser suprayectiva, es decir que tiene que ser una biyección. Pero entonces de $f \circ g = Id_A$, se sigue que

$$f = f \circ g \circ g^{-1} = Id_A \circ g^{-1} = g^{-1},$$

donde vemos que f es biyectiva. ■

Lema 2 Si X es un conjunto infinito, entonces $\forall x \in X$, el conjunto $X \setminus \{x\}$ es infinito también.

Demostración. Tomemos una biyección $X \xrightarrow{f} A$, $A \subsetneq X$, y tomemos $x \in X$.

Entonces la restricción $f|_{X \setminus \{x\}}^{A \setminus \{f(x)\}}$ en el diagrama

$$\begin{array}{ccc} X & \xrightarrow{f} & A \\ \uparrow & & \uparrow \\ X \setminus \{x\} & \xrightarrow{f|_{X \setminus \{x\}}^{A \setminus \{f(x)\}}} & A \setminus \{f(x)\} \end{array},$$

también es una biyección.

Demostraremos ahora que $A \setminus \{f(x)\}$ es un subconjunto propio de $X \setminus \{x\}$.

Por contradicción, supongamos que $A \setminus \{f(x)\} = X \setminus \{x\}$.

Si $c \in X \setminus A$, entonces $c \in (X \setminus \{x\}) \setminus (A \setminus \{f(x)\})$. En caso contrario, tendríamos que

$$[c \notin (X \setminus \{x\})] \vee [c \in A \setminus \{f(x)\}].$$

Es decir tendríamos que $[c \notin X] \vee [c \in \{x\}] \vee [(c \in A \wedge c \neq f(x))]$. Como $c \in X \setminus A$, tendríamos que

$$c \in \{x\},$$

es decir que $c = x$. Así resulta que el único elemento de $X \setminus A$ es x .

Como estamos suponiendo que $A \setminus \{f(x)\} = X \setminus \{x\}$, y como x es el único elemento en $X \setminus A$, entonces $X = A \cup \{x\}$.

De aquí se sigue que $A \setminus \{f(x)\} = X \setminus \{x\} = A$, por lo que $f(x) \notin A$, absurdo.

Como $X \setminus \{x\} \xrightarrow{f|_{X \setminus \{x\}}^{A \setminus \{f(x)\}}} A \setminus \{f(x)\}$ es una biyección entre $X \setminus \{x\}$ y un subconjunto propio, concluimos que $X \setminus \{x\}$ es infinito. ■

Corolario 4 . Si F es un conjunto finito y z es un objeto (un conjunto) que no pertenece a F , entonces $F \cup \{z\}$ también es finito.

Lema 3 . Si $X \subseteq Y$ y X es infinito, entonces Y también es infinito.

Demostración. Tomemos $Y = (Y \setminus X) \cup X$. Supongamos que

$$f : X \longrightarrow A,$$

es una biyección entre X y un subconjunto propio de X . Definamos

$$g : Y \longrightarrow Y$$

mediante:

$$\begin{array}{rcl} g : & Y & \longrightarrow & Y \\ & y & \mapsto & y & \text{si } y \in (Y \setminus X) \\ & x & \mapsto & f(x) & \text{si } x \in X \end{array}$$

g es una biyección entre Y y $(Y \setminus X) \cup A$, que es un subconjunto propio de $(Y \setminus X) \cup X = Y$. ■

Teorema 35 . *El conjunto de los naturales es el conjunto infinito más pequeño.*

Demostración. Ya hemos visto que \mathbb{N} es un conjunto infinito.

Supongamos que X es un conjunto infinito. Entonces $X \neq \emptyset$, pues $\emptyset = 0$, y ya vimos que todos los naturales son conjuntos finitos.

Podemos escoger entonces un elemento $a_0 \in X$, y considerar el conjunto $X \setminus \{a_0\}$.

Como vimos en el Lema 2, $X \setminus \{a_0\}$ sigue siendo infinito, en particular no es vacío así que contiene un elemento que denotaremos a_1 .

Podemos repetir el argumento con el conjunto $X \setminus \{a_0, a_1\}$, para obtener un elemento $a_2 \in X \setminus \{a_0, a_1\}$.

De esta manera tenemos una sucesión¹

$$a_0, a_1, a_2, \dots$$

de elementos distintos de X .

Podemos pensar esta sucesión como una función inyectiva

$$\begin{array}{rcl} \mathbb{N} & \xrightarrow{h} & X \\ k & \mapsto & a_k \end{array},$$

y podemos pensar que el conjunto $\text{Im}(h) = h(\mathbb{N}) \subseteq X$, es una copia del conjunto de los números naturales. ■

Así podemos decir que cualquier conjunto infinito contiene una copia del conjunto de los números naturales. Es en este sentido en el que decimos que el conjunto de los números naturales es el menor conjunto infinito.

¹Una sucesión es una función cuyo dominio es \mathbb{N} .

Teorema 36 . *Si F es un conjunto finito, entonces hay un natural n tal que $|F| = |n|$.*

Demostración. Si $F = \emptyset$, entonces $|F| = |0|$.

Si $F \neq \emptyset$, podemos tomar $a_0 \in F$. Si $F \setminus \{a_0\} \neq \emptyset$, podemos tomar $a_1 \in F \setminus \{a_0\}$. Repetimos este argumento, y notamos que debe terminar, pues en otro caso encontraríamos una sucesión de elementos distintos de F :

$$a_0, a_1, \dots$$

lo que nos daría una inyección $\mathbb{N} \rightarrow F$, contradiciendo que F es finito (F contendría una copia de los naturales, que es un conjunto infinito).

Si el proceso termina en a_k , es porque $F \setminus \{a_0, a_1, \dots, a_k\} = \emptyset$, de donde tendríamos que $F = \{a_0, a_1, \dots, a_k\}$, que claramente tiene tantos elementos como $\{0, 1, \dots, k\} = \sigma(k)$. ■

En los siguientes ejercicios, indicaremos una construcción alternativa del conjunto de los números naturales.

Defina el sucesor de un conjunto X de la siguiente manera:

$$s(X) = \{X\}.$$

Diga que un conjunto X es inductivo si satisface las siguientes dos condiciones:

1. $\emptyset \in X$.
2. $A \in X \implies s(A) \in X$.

Introduzca el siguiente axioma de infinito: EXISTE UN CONJUNTO INDUCTIVO.

Ejercicio 139 *Demuestre que la intersección de una familia de conjuntos inductivos es también un conjunto inductivo.*

Ejercicio 140 *Considere un conjunto inductivo (que existe por el axioma anterior) y considere el conjunto de los elementos de este conjunto que pertenecen a cualquier otro conjunto inductivo. Demuestre que este conjunto es un conjunto inductivo y que es un subconjunto de cualquier otro conjunto inductivo. Denote por \mathbb{N} a este conjunto, defina $0 := \emptyset$ y demuestre que:*

Ejercicio 141 . *La función*

$$\begin{array}{rccc} s : & \mathbb{N} & \longrightarrow & \mathbb{N} \\ & k & \longmapsto & \{k\} \end{array}$$

es inyectiva, y su imagen es $\mathbb{N} \setminus \{\emptyset\}$.

3.9 El orden en los naturales

1. $n < m$ si $n \in m$.
2. $n \leq m$ si $(n \in m) \vee (n = m)$.

Observación 44 . $0 \leq n, \quad \forall n \in \mathbb{N}$.

Demostración. Consideremos la siguiente proposición:

$$P(n) : 0 \leq n.$$

Demostrar que todos los naturales tienen la propiedad P , es lo mismo que demostrar que

$$\{n \in \mathbb{N} \mid P(n)\}$$

es todo el conjunto \mathbb{N} .

Demostraremos que el conjunto

$$\{n \in \mathbb{N} \mid P(n)\}$$

es inductivo.

Base de la Inducción.

$P(0)$ es cierta, puesto que $0 = 0$.

Paso Inductivo.

Supongamos $P(k)$, es decir que $0 \leq k$. Por lo tanto, $0 \in k \vee 0 = k$.

Si $0 \in k$, entonces como $k \subseteq \sigma(k)$, tendríamos que $0 \in \sigma(k)$.

Si $0 = k$, entonces $0 = k \in \sigma(k)$.

En cualquier caso, $0 \in \sigma(k)$, es decir, $0 < \sigma(k)$.

Por lo tanto $P(\sigma(k))$ es cierta. ■

Lema 4 . $n < m \implies \sigma(n) < \sigma(m)$.

Demostración. Consideremos la proposición

$$n < m \implies \sigma(n) < \sigma(m),$$

como una afirmación que se hace acerca de m .

Es decir consideremos

$$P(m) : n < m \implies \sigma(n) < \sigma(m).$$

Demostremos que el conjunto $\{m \in \mathbb{N} \mid P(m)\}$ es inductivo.

Base de la inducción.

Queremos demostrar que

$$n \in 0 \implies \sigma(n) \in \sigma(0),$$

$n \in 0 = \emptyset$, es una proposición falsa, por lo que la implicación es verdadera.

Paso inductivo.

Supongamos que es cierta $P(k) : P(k) : n < k \implies \sigma(n) < \sigma(k)$.

Queremos demostrar que también

$P(\sigma(k)) : P(\sigma(k)) : n < \sigma(k) \implies \sigma(n) < \sigma(\sigma(k))$ es cierta.

De no ser así, sucederían:

1) $n < k \implies \sigma(n) < \sigma(k)$.

2) $n < \sigma(k)$.

3) $\neg[\sigma(n) < \sigma(\sigma(k))]$.

De 3) tendríamos que $\sigma(n) \notin \sigma(\sigma(k)) = \sigma(k) \cup \{\sigma(k)\}$. En particular, $\sigma(n) \neq \sigma(k)$. (Por lo tanto $n \neq k$).

De 2), tendríamos que $n \in k \cup \{k\}$. Así que $n \in k \vee n = k$. Como la segunda posibilidad ha sido descartada, tenemos que $n \in k$. Pero entonces de 1) se sigue que $\sigma(n) < \sigma(k)$.

Así que $\sigma(n) \in \sigma(k) \in \sigma(\sigma(k))$, y dado que los naturales son conjuntos transitivos, tendríamos que $\sigma(n) \in \sigma(\sigma(k))$, en contradicción con 3).

La contradicción demuestra que $P(k) \wedge \neg P(\sigma(k))$ es falsa, por lo que $\neg P(k) \vee P(\sigma(k))$ es verdadera. Esto es equivalente a $P(k) \implies P(\sigma(k))$.

■

Lema 5 . $\forall n \in \mathbb{N}, \quad m < n \implies \begin{cases} \sigma(m) < n \\ \vee \\ \sigma(m) = n \end{cases}.$

Demostración. Escribamos

$$P(n) : m < n \implies \begin{cases} \sigma(m) < n \\ \vee \\ \sigma(m) = n \end{cases}$$

Demostraremos que $\{n \in \mathbb{N} \mid P(n)\} = \mathbb{N}$. Por este efecto, hasta demostrar que $\{n \in \mathbb{N} \mid P(n)\}$ es un conjunto inductivo.

Base de la inducción.

$$P(0) : m < 0 \implies \begin{cases} \sigma(m) < 0 \\ \vee \\ \sigma(m) = 0 \end{cases}$$

es una proposición verdadera, porque la

proposición $m < 0$ es falsa.

Paso inductivo.

Supongamos que $P(k)$ es cierta. Si $P(\sigma(k))$ fuera falsa entonces existiría $m \in \mathbb{N}$, tal que:

- 1) $m \in \sigma(k)$,
- 2) $\sigma(m) \notin \sigma(k)$,
- 3) $\sigma(m) \neq \sigma(k)$.

$$m \in \sigma(k) = k \cup \{k\} \implies m \in k \vee m = k.$$

Si $m \in k$, entonces por la hipótesis inducción, tendríamos que $\sigma(m) < k \vee \sigma(m) = k$. Lo que implicaría que $\sigma(m) < \sigma(k)$, en contra del inciso 2).

Nos queda la posibilidad de que $m = k$, pero implica que $\sigma(m) = \sigma(k)$, en contra de 3).

Concluimos que si $P(k)$ es cierta, entonces $P(\sigma(k))$ también es cierta.

■

Corolario 5 . $\forall n \in \mathbb{N}, (m < n) \wedge \neg(\sigma(m) < n) \implies \sigma(m) = n$.

Hemos observado que una de las consecuencias del axioma de la teoría de conjuntos que prohíbe la existencia de sucesiones descendentes de pertenencias infinitas (no existen sucesiones de la forma: $\dots a_n \in a_{n-1} \in \dots \in a_1 \in a_0$), es que un conjunto no puede ser elemento de sí mismo. Tampoco puede pasar que para dos conjuntos A, B se tenga que cada uno es elemento del otro, pues en ese caso uno podría escribir la sucesión infinita

$$\dots B \in A \in B \in \dots \in B \in A,$$

que no puede ocurrir dentro de la Teoría de Conjuntos.

Teorema 37 . $\forall n \in \mathbb{N}, \forall m \in \mathbb{N}$ vale una y sólo una de las siguientes afirmaciones:

1. $n = m$.
2. $n \in m$.
3. $m \in n$.

Demostración. Como hemos observado en el párrafo anterior, las proposiciones 1), 2) y 3) no pueden ocurrir simultáneamente.

Consideremos la proposición

$$Q(n) : \forall m \in \mathbb{N}, \quad (n = m) \vee (n \in m) \vee (m \in n).$$

Demostraremos que $\{n \in \mathbb{N} \mid Q(n)\}$ es inductivo.

Base de la inducción. Como ya hemos observado, $0 \leq m, \forall m \in \mathbb{N}$. Por lo tanto $Q(0)$ es verdadera.

Paso inductivo.

Supongamos que $Q(k)$ es cierta pero que $Q(\sigma(k))$ es falsa.

Entonces $\exists m \in \mathbb{N}$ tal que:

$$(\sigma(k) \neq m) \wedge (\sigma(k) \notin m) \wedge (m \notin \sigma(k)).$$

Sin embargo, por hipótesis

$$(k = m) \vee (k \in m) \vee (m \in k).$$

Si $k = m$ entonces $m = k \in \sigma(k)$. Esta opción queda descartada.

Si $m \in k$ entonces $m \in k \in \sigma(k)$, por lo que $m \in \sigma(k)$. Esta opción también se descarta.

Sólo nos queda la posibilidad de que $k \in m$.

Pero ahora, de $k \in m$ y de $(\sigma(k) \notin m)$ se sigue, en vista del corolario anterior, que $\sigma(k) = m$, ∇ . ■

Teorema 38 (Principio del Buen Orden) . Todo subconjunto no vacío de \mathbb{N} tiene un primer elemento (es decir un elemento menor que todos los demás).

Demostración. Sea $\emptyset \neq A \subsetneq \mathbb{N}$. Si A no tuviera un primer elemento, podríamos escoger un elemento $a_0 \in A$. Como este elemento no es el primero, existiría otro elemento que va antes, llamémoslo a_1 .

Entonces

$$a_1 \in a_0,$$

como tampoco a_1 es el primer elemento de A , debe haber un elemento que va antes de a_1 , llamémoslo a_2 .

Entonces

$$a_2 \in a_1 \in a_0,$$

podemos proseguir indefinidamente con este argumento, contradiciendo el axioma de regularidad. ■

Ejemplo 97 . 0 es el primer elemento de \mathbb{N} .

3.10 Recusión

Demostraremos el Teorema de recursión débil:

Teorema 39 . Si $a \in X$ y $X \xrightarrow{f} X$ es una función, entonces existe una única función $\eta : \mathbb{N} \longrightarrow \mathbb{X}$ tal que:

1. $\eta(0) = a$.
2. $\eta(\sigma(n)) = f(\eta(n))$.

$$\begin{array}{ccccccccccccc} 0 & \xrightarrow{\sigma} & 1 & \xrightarrow{\sigma} & 2 & \xrightarrow{\sigma} & 3 & \xrightarrow{\sigma} & \dots & n & \xrightarrow{\sigma} & & \sigma(n) \\ \downarrow^{\eta} & & \downarrow^{\eta} & & \downarrow^{\eta} & & \downarrow^{\eta} & & & \downarrow^{\eta} & & & \downarrow^{\eta} \\ a & \xrightarrow{f} & f(a) & \xrightarrow{f} & ff(a) & \xrightarrow{f} & fff(a) & \xrightarrow{f} & \dots & \eta(n) & \xrightarrow{f} & f(\eta(n)) = \eta(\sigma(n)) \end{array}$$

Demostración. Consideremos

$$\mathfrak{R} = \{R \subseteq \mathbb{N} \times X \mid (0, a) \in R, \quad (n, x) \in R \implies (\sigma(n), f(x)) \in R\}.$$

Demostraremos que $\cap \mathfrak{R} \in \mathfrak{R}$:

Para empezar, es claro que $\cap \mathfrak{R} \subseteq \mathbb{N} \times \mathbb{X}$.

Además, como $(0, a) \in R \quad \forall R \in \mathfrak{R}$, entonces $(0, a) \in \cap \mathfrak{R}$.

Si $(n, x) \in \cap \mathfrak{R}$, entonces $(n, x) \in R \quad \forall R \in \mathfrak{R}$. Pero entonces

$$(\sigma(n), f(x)) \in R \quad \forall R \in \mathfrak{R}.$$

Es decir que $(\sigma(n), f(x)) \in \cap \mathfrak{R}$.

Una vez que hemos visto que $\cap \mathfrak{R} \in \mathfrak{R}$, demostraremos que $\mathbb{N} \xrightarrow{\cap \mathfrak{R}} \mathbb{X}$ es una función.

Denotemos $\eta := \cap \mathfrak{R}$.

Necesitamos demostrar dos cosas: la primera, es que $dom(\eta) = \mathbb{N}$; la segunda, es que $[(n, x), (n, y) \in \eta] \implies [x = y]$.

Veamos que $dom(\eta) = \mathbb{N}$, para esto basta ver que $dom(\eta)$ es un conjunto inductivo.

Como $(0, a) \in \eta$, entonces $0 \in dom(\eta)$.

Supongamos que $n \in dom(\eta)$. Esto quiere decir que $\exists x \in X$ tal que $(n, x) \in \eta$. Pero entonces también $(\sigma(n), f(x)) \in \eta$, por lo que $\sigma(n) \in dom(\eta)$. Así, $dom(\eta) = \mathbb{N}$.

Para demostrar que $[(n, x), (n, y) \in \eta] \implies [x = y]$, supongamos lo contrario, es decir supongamos que $\exists n \in \mathbb{N}$ tal que

$$(n, x), (n, y) \in \eta, \text{ con } x \neq y \in X.$$

Por el Principio del Buen Orden, podríamos tomar la menor n con la propiedad anterior, hagámoslo.

Vamos a ver que $n \neq 0$. Pues si $(0, b) \in \eta$, con $b \neq a$, entonces $\eta \setminus \{(0, b)\} \in \mathfrak{R}$:

$(0, a) \in \eta \setminus \{(0, b)\}$, pues $(0, a) \in \eta$ y $(0, a) \neq (0, b)$.

Además,

$$(n, x) \in \eta \setminus \{(0, b)\} \implies (\sigma(n), f(x)) \in \eta \setminus \{(0, b)\};$$

pues $(\sigma(n), f(x)) \in \eta$ y $(\sigma(n), f(x)) \neq (0, b)$, ya que $\sigma(n) \neq 0$ (0 no es un sucesor).

Si recordamos la definición de η , tendremos que

$$\eta = \cap \mathfrak{R} \subseteq \eta \setminus \{(0, b)\} \subseteq \eta,$$

de donde tendríamos que $\eta \setminus \{(0, b)\} = \eta$, ∇ .

La contradicción anterior demuestra que $n \neq 0$.

Hemos visto que $n \neq 0$.

Entonces $(n, x) \neq (n, y) \in \eta$. Como $n \neq 0$, entonces $n = \sigma(m), m \in \mathbb{N}$. Como $m \in \text{dom}(\eta), \exists u \in X$ tal que $(m, u) \in \eta$. Entonces

$$(\sigma(m), f(u)) \in \eta.$$

$(\sigma(m), f(u))$ tiene que ser distinto de (n, x) o de (n, y) . Supongamos que $(\sigma(m), f(u)) \neq (n, x)$.

Entonces

$$f(u) \neq x. \quad (3.2)$$

Consideremos $\eta \setminus \{(n, x)\}$, veremos que $\eta \setminus \{(n, x)\} \in \mathfrak{R}$.

Para empezar, $(0, a) \in \eta \setminus \{(n, x)\}$, pues $(0, a) \in \eta$ y $0 \neq n$.

Ahora veamos que

$$(k, w) \in \eta \setminus \{(n, x)\} \implies (\sigma(k), f(w)) \in \eta \setminus \{(n, x)\} :$$

en caso contrario,

$$(k, w) \in \eta \setminus \{(n, x)\} \wedge (\sigma(k), f(w)) = (n, x) :$$

pero entonces

$$(\sigma(k), f(w)) = (n, x) = (\sigma(m), x), \quad (3.3)$$

así que $\sigma(k) = \sigma(m)$, por lo que $k = m$ (recuérdese que σ es una función inyectiva).

Además, $(k, w) = (m, w)$. Pero $(m, u) \in \eta$, dada la manera en que escogimos a n , y dado que $m < n$, tenemos que

$$[(m, w), (m, u) \in \eta] \implies (w = u).$$

Entonces, de la ecuación 3.3, tenemos que

$$x = f(w) = f(u).$$

Esto contradice que $x \neq f(u)$ (3.2).

Esta contradicción termina la demostración del existencia de la función η .

Para demostrar la unicidad, supongamos que $\mu : \mathbb{N} \longrightarrow X$ es otra función que satisface:

- 1) $\mu(0) = a$, y
- 2) $\mu(\sigma(n)) = f(\mu(n))$,

usando inducción es inmediato que

$$\eta(n) = \mu(n) \quad \forall n \in \mathbb{N}.$$

■ Notemos que en lugar de

$$\begin{array}{ccccccccc} 0 & \xrightarrow{\sigma} & 1 & \xrightarrow{\sigma} & 2 & \xrightarrow{\sigma} & \dots & n & \xrightarrow{\sigma} \\ \downarrow \eta & & \downarrow \eta & & \downarrow \eta & & & \downarrow \eta & \\ a & \xrightarrow{f} & f(a) & \xrightarrow{f} & ff(a) & \xrightarrow{f} & \dots & \eta(n) & \xrightarrow{f} f(\eta(n)) = \eta(\sigma(n)) \end{array}$$

podemos escribir simplemente,

$$\begin{array}{ccc} n & \xrightarrow{\sigma} & \sigma(n) \\ \downarrow \eta & & \downarrow \eta \\ \eta(n) & \xrightarrow{f} & f(\eta(n)) = \eta(\sigma(n)) \end{array},$$

pero aquí se ha omitido la información de que $\eta(0) = a$. Para completar esta información, escribimos el diagrama

$$\begin{array}{ccccc} & & \mathbb{N} & \xrightarrow{\sigma} & \mathbb{N} \\ & \nearrow \bar{0} & \downarrow \eta & & \downarrow \eta \\ \{0\} & & X & \xrightarrow{f} & X \end{array}$$

Usaremos el teorema de Recursión para demostrar las propiedades algebraicas del conjunto de los números naturales.

3.11 Las propiedades algebraicas de los naturales

3.11.1 La suma

Definición 54 . Usaremos el teorema de Recursión para definir la función “sumar m ”, $s_m : \mathbb{N} \longrightarrow \mathbb{N}$: (piénsese que $s_m(k) = m + k$).

1. $s_m(0) = m$,
2. $s_m(\sigma(n)) = \sigma(s_m(n))$.

Alternativamente, s_m es la única función $\mathbb{N} \rightarrow \mathbb{N}$ tal que hace commutativo diagrama siguiente:

$$\begin{array}{ccc}
 \mathbb{N} & \xrightarrow{\sigma} & \mathbb{N} \\
 \bar{0} \nearrow & \downarrow s_m & \downarrow s_m \\
 \{0\} & & \\
 \bar{m} \searrow & \downarrow & \downarrow \\
 \mathbb{N} & \xrightarrow{\sigma} & \mathbb{N}
 \end{array}$$

Notemos que para definir s_m , se ha aplicado el teorema de Recursión a la función $\sigma : \mathbb{N} \rightarrow \mathbb{N}$, y al elemento especial $m \in \mathbb{N}$.

Observación 45 . s_0 es la función $Id_{\mathbb{N}}$.

Demostración. Por definición, s_0 es la única función que hace commutativo el diagrama

$$\begin{array}{ccc}
 \mathbb{N} & \xrightarrow{\sigma} & \mathbb{N} \\
 \bar{0} \nearrow & \downarrow s_0 & \downarrow s_0 \\
 \{0\} & & \\
 \bar{0} \searrow & \downarrow & \downarrow \\
 \mathbb{N} & \xrightarrow{\sigma} & \mathbb{N}
 \end{array}$$

Comparemos este diagrama con el siguiente:

$$\begin{array}{ccc}
 \mathbb{N} & \xrightarrow{\sigma} & \mathbb{N} \\
 \bar{0} \nearrow & \downarrow Id_{\mathbb{N}} & \downarrow Id_{\mathbb{N}} \\
 \{0\} & & \\
 \bar{0} \searrow & \downarrow & \downarrow \\
 \mathbb{N} & \xrightarrow{\sigma} & \mathbb{N}
 \end{array}$$

Por la unicidad en el teorema de Recursión, tenemos que $s_0 = Id_{\mathbb{N}}$. Es decir, que $s_0(k) = k \quad \forall k \in \mathbb{N}$. ■

Teorema 40 . $s_1 = \sigma : \mathbb{N} \longrightarrow \mathbb{N}$.

Demostración. Nuevamente, sólo tendremos que comparar dos diagramas: por definición, $s_1 : \mathbb{N} \longrightarrow \mathbb{N}$ es la única función que hace commutativo el diagrama siguiente:

$$\begin{array}{ccc}
 \mathbb{N} & \xrightarrow{\sigma} & \mathbb{N} \\
 \bar{0} \nearrow & \downarrow s_1 & \downarrow s_1 \\
 \{0\} & & \\
 \bar{1} \searrow & \downarrow & \downarrow \\
 \mathbb{N} & \xrightarrow{\sigma} & \mathbb{N}
 \end{array}$$

pero también

$$\begin{array}{ccc}
 \mathbb{N} & \xrightarrow{\sigma} & \mathbb{N} \\
 \bar{0} \nearrow & \downarrow \sigma & \downarrow \sigma \\
 \{0\} & & \\
 \bar{1} \searrow & \downarrow & \downarrow \\
 \mathbb{N} & \xrightarrow{\sigma} & \mathbb{N}
 \end{array}$$

es commutativo, puesto que $\sigma(0) = 1$ y $\sigma \circ \sigma = \sigma \circ \sigma$.

Por la unicidad en el teorema de Recursión, tenemos que $s_1 = \sigma$. Es decir que $s_1(k) = \sigma(k)$ ($1 + k = \sigma(k)$). ■

Teorema 41 . *La suma de naturales es asociativa, es decir que*

$$n + (m + k) = (n + m) + k \quad \forall n, m, k \in \mathbb{N}.$$

Demostración. Lo que queremos demostrar es que

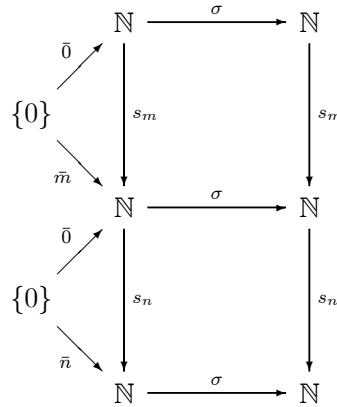
$$s_n(s_m(k)) = s_{n+m}(k) \quad \forall n, m, k \in \mathbb{N}.$$

Ahora, por definición de composición de funciones, es claro que lo que queremos demostrar es que

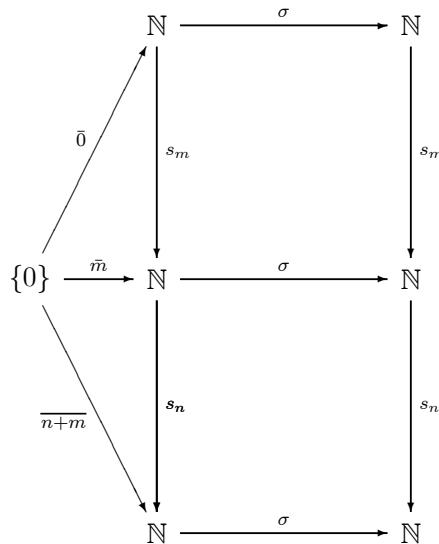
$$(s_n \circ s_m)(k) = s_{n+m}(k) \quad \forall n, m, k \in \mathbb{N}.$$

O bien, que las funciones $s_n \circ s_m$ y s_{n+m} son la misma función.

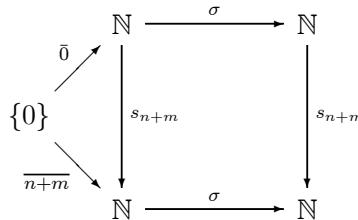
Comparemos el diagrama que define a la función s_{n+m} , con el siguiente diagrama:



o bien:



El diagrama que define a la función s_{n+m} es:



es decir que s_{n+m} es la única función de \mathbb{N} en sí mismo tal que:

1. $s_{n+m}(0) = n + m$ y
- 2) $s_{n+m} \circ \sigma = \sigma \circ s_{n+m}$.

Pero si observamos, $s_n \circ s_m$ también tiene las dos propiedades:

Por una parte, $(s_n \circ s_m)(0) = s_n(s_m(0)) = s_n(m) = n + m$.

Y además $(s_n \circ s_m) \circ \sigma = s_n \circ (s_m \circ \sigma) = s_n \circ (\sigma \circ s_m) = (s_n \circ \sigma) \circ s_m = (\sigma \circ s_n) \circ s_m = \sigma \circ (s_n \circ s_m)$.

Por la unicidad en teorema de Recursión, tenemos que $s_n \circ s_m = s_{n+m}$, lo que es equivalente a la asociatividad de la suma, como notamos al principio de este argumento. ■

Teorema 42 . *La suma de naturales es conmutativa, es decir que*

$$n + m = m + n, \quad \forall n, m \in \mathbb{N}.$$

Demostración. Queremos demostrar que $s_n(m) = s_m(n)$, $\forall n, m \in \mathbb{N}$.

Demostraremos esto por inducción sobre n .

Consideremos la proposición

$$P(n) : n + m = m + n, \quad \forall m \in \mathbb{N},$$

demostraremos que esta proposición es cierta para cualquier natural n .

Explícitamente, demostraremos que el conjunto

$$X =: \{n \in \mathbb{N} \mid P(n)\}$$

es inductivo.

Base de la inducción.

Recordemos que $s_0 = Id_{\mathbb{N}}$ y que por definición, $s_m(0) = m$. Entonces $0 + m = s_0(m) = m = s_m(0)$. Como esto pasa para toda $m \in \mathbb{N}$, tenemos que $0 \in X$.

Paso inductivo.

Supongamos que $n + m = m + n$, $\forall m \in \mathbb{N}$, quisiéramos demostrar que también

$$\sigma(n) + m = m + \sigma(n), \quad \forall m \in \mathbb{N}.$$

$$\begin{aligned} \text{Pero } \sigma(n) + m &= (1 + n) + m = 1 + (n + m) = \sigma(n + m) = \sigma(m + n) = \\ &= 1 + (m + n) = (1 + m) + n = \sigma(m) + n. \end{aligned}$$

En las ecuaciones de arriba, se ha usado que $\sigma = s_1$, la asociatividad de la suma, y la hipótesis de inducción. Para terminar el argumento basta demostrar que

$$\sigma(m) + n = m + \sigma(n),$$

que puede expresarse en la forma siguiente:

$$s_{\sigma(m)}(n) = s_m(\sigma(n)) = (s_m \circ \sigma)(n).$$

De aquí, que basta demostrar que

$$s_{\sigma(m)} = s_m \circ \sigma.$$

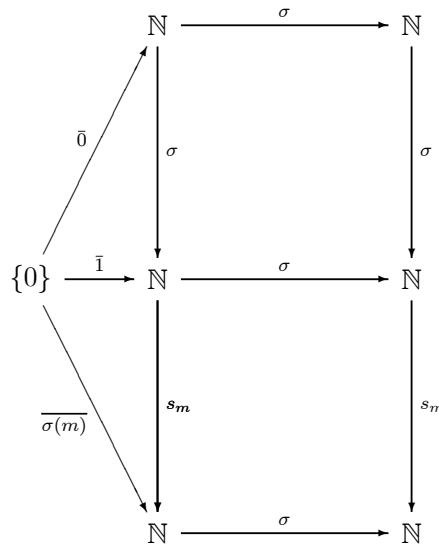
Nuestro problema se reduce a comparar dos diagramas.

$$\begin{array}{ccc} \mathbb{N} & \xrightarrow{\sigma} & \mathbb{N} \\ \bar{0} \nearrow & & \downarrow s_{\sigma(m)} \\ \{0\} & & \downarrow s_{\sigma(m)} \\ \bar{\sigma(m)} \searrow & & \mathbb{N} \xrightarrow{\sigma} \mathbb{N} \end{array}$$

$s_{\sigma(m)}$ es la única función que satisface:

- 1) $s_{\sigma(m)}(0) = \sigma(m)$.
- 2) $s_{\sigma(m)} \circ \sigma = \sigma \circ s_{\sigma(m)}$.

Escribiendo el siguiente diagrama



notamos que $(s_m \circ \sigma)(0) = (\sigma \circ s_m)(0) = \sigma(s_m(0)) = \sigma(m)$.

Además, $(s_m \circ \sigma) \circ \sigma = (\sigma \circ s_m) \circ \sigma = \sigma \circ (s_m \circ \sigma)$.

Concluimos que $(s_m \circ \sigma) = s_{\sigma(m)}$, que era lo que nos hacía falta. ■

Observación 46 . Son equivalentes las siguientes dos afirmaciones:

1. s_m es una función inyectiva.
 2. m es cancelable (respecto a la suma). (Es decir: $m + k = m + l \implies k = l$).

Demuestra. s_m es inyectiva $\Leftrightarrow (s_m(k) = s_m(l) \Rightarrow (k = l)) \Leftrightarrow$
 $\Leftrightarrow (m + k = m + l \Rightarrow k = l) \Leftrightarrow$
 $\Leftrightarrow m$ es cancelable respecto a la suma. ■

Teorema 43 . $\forall n \in \mathbb{N}$, s_n es inyectiva.

Demostración. Demostraremos esto por inducción.

Base de la inducción.

Hemos visto que $s_0 = Id_{\mathbb{N}}$, que es una función inyectiva.

Paso inductivo.

Si s_k es una función inyectiva entonces también lo es $s_{\sigma(k)}$, pues como hemos visto, $s_{\sigma(k)} = s_k \circ \sigma$, una composición de funciones inyectivas. ■

Teorema 44 . $n + m = 0 \implies (n = 0) \wedge (m = 0)$.

Demostración. Por contrapuesta, si $n \neq 0 \vee m \neq 0$, entonces $n = \sigma(n')$ $\vee m = \sigma(m')$. Por lo tanto,

$$n + m = \begin{cases} \sigma(n') + m = (1 + n') + m = \sigma(n' + m) \\ \vee \\ m + \sigma(m') = m + (1 + m') = \sigma(n + m') \end{cases}$$

es distinto de 0, ya que es el sucesor de algún natural. ■

Teorema 45 . Para $n, m \in \mathbb{N}$, se cumple que:

1. $n < m \Leftrightarrow (\exists k \in \mathbb{N} \setminus \{0\}, \quad n + k = m)$.
2. $n \leq m \Leftrightarrow (\exists k \in \mathbb{N}, \quad n + k = m)$.

Demostración. 1) \implies) Por inducción sobre n .

Consideremos la proposición

$$P(n) : n < m \implies (\exists k \in \mathbb{N} \setminus \{0\}, \quad n + k = m).$$

Demostraremos que $A = \{n \in \mathbb{N} \mid P(n)\} = \mathbb{N}$.

Es claro que si $0 < m$ entonces $0 + m = m$; por lo tanto $0 \in A$.

Supongamos ahora que $j < m \implies \exists a \neq 0$, tal que $j + a = m$.

Si tuviéramos que $\sigma(j) < m$, entonces

$$j < \sigma(j) < m,$$

por lo que $j + a = m$, con $a \neq 0$. Así, $a = \sigma(a')$, para alguna $a' \in \mathbb{N}$. Entonces

$$\begin{aligned} m &= j + \sigma(a') = j + (1 + a') = (j + 1) + a' = \\ &= \sigma(j) + a'. \end{aligned}$$

Notemos además que $a' \neq 0$, porque en caso contrario, $m = \sigma(j)$, pero $\sigma(j) < m$.

Entonces, $\sigma(j) < m \implies \exists a' \in \mathbb{N} \setminus \{0\}$ tal que $\sigma(j) + a' = m$. Esto coloca a $\sigma(j)$ como un elemento de A .

Por lo tanto, A es un subconjunto inductivo de \mathbb{N} , por lo que $A = \mathbb{N}$.

\Leftarrow) Sea $Q(n)$ la proposición recíproca de $P(n)$, es decir,

$$Q(n) : (\exists k \in \mathbb{N} \setminus \{0\}, n + k = m) \implies (n < m).$$

Demostraremos que el conjunto

$$B =: \{n \in \mathbb{N} \mid Q(n)\}$$

es todo \mathbb{N} . Denotemos $\mathbb{N}^+ = \mathbb{N} \setminus \{0\}$.

Base de la inducción. $0 + k = m$, con $k \in \mathbb{N}^+ \implies (m = k) \wedge (0 < m)$ (0 es menor que todo natural distinto de 0, ver la observación 44).

Paso inductivo. Supongamos que $r \in B$. Si $\sigma(r) \notin B$, entonces

$$\sigma(r) + k = m, \text{ pero } \neg(\sigma(r) < m), k \neq 0.$$

es decir,

$$(1 + r) + k = m, \text{ pero } (\sigma(r) \notin m), k \neq 0.$$

Como $(1 + r) + k = r + (1 + k) = m$, por hipótesis de inducción tenemos que

$$r < m.$$

pero $[(r < m) \wedge (\sigma(r) \notin m)] \implies \sigma(r) = m$ (lema 5, en la página 169).

Pero entonces,

$$(1 + r) + k = \sigma(r) + k = m = (1 + r) + 0,$$

por lo que $k = 0 \nabla$. (Recuerde que los naturales son cancelables respecto a la suma, teorema 43).

Esa contradicción muestra que si $r \in B$ entonces $\sigma(r) \in B$.

Por lo tanto, B es inductivo, por lo tanto $B = \mathbb{N}$.

La parte 2) se sigue inmediatamente de 1). ■

Ejercicio 142 . Demuestre la parte 2) en el enunciado del teorema anterior.

Lema 6 . $\{n \in \mathbb{N} \mid k < n < \sigma(k)\} = \emptyset$.

Demostración. Por el teorema anterior, $\exists a, b \in \mathbb{N}^+$, tales que $k + a = n$ y $n + b = \sigma(k)$.

Entonces

$$1 + k = \sigma(k) = n + b = k + a + b,$$

cancelando k tenemos que

$$1 = a + b.$$

Pero $a = 1 + a'$, $b = 1 + b'$, $a', b' \in \mathbb{N}$.

Por lo tanto,

$$1 = 1 + a' + 1 + b',$$

de donde tenemos que

$$0 = a' + (1 + b')$$

pero esto implica que $1 + b' = b = 0 \nabla$. (Se acaba de usar el teorema 44). ■

Note que el teorema que acabamos de demostrar es una buena razón para decir que $\sigma(n)$ es el “sucesor” de n .

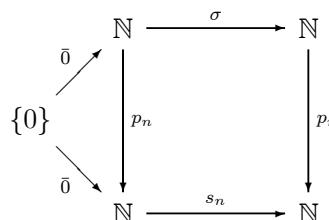
3.11.2 El producto en \mathbb{N}

Usaremos el teorema de Recursión para definir el producto en el conjunto de los naturales.

Definición 55 . Definimos la función “multiplicar por n ”, $p_n : \mathbb{N} \longrightarrow \mathbb{N}$, por medio de las dos propiedades siguientes:

1. $p_n(0) = 0$.
2. $p_n \circ \sigma = s_n \circ p_n$.

Alternativamente, p_n es la única función que hace commutativo el diagrama:



$$1. \ p_0 = \bar{0} : \mathbb{N} \longrightarrow \mathbb{N} .$$

$$2. \ p_1 = Id : \mathbb{N} \longrightarrow \mathbb{N} .$$

Demostración. Demostraremos 1) y dejaremos 2) como ejercicio. Únicamente tenemos que comparar los siguientes dos diagramas:

$$\begin{array}{ccc}
 \mathbb{N} & \xrightarrow{\sigma} & \mathbb{N} \\
 \bar{0} \nearrow \swarrow \{0\} & \downarrow p_0 & \downarrow p_0 \\
 \mathbb{N} & \xrightarrow{s_0} & \mathbb{N}
 \end{array}$$

y

$$\begin{array}{ccc}
 \mathbb{N} & \xrightarrow{\sigma} & \mathbb{N} \\
 \bar{0} \nearrow \swarrow \{0\} & \downarrow \bar{0} & \downarrow \bar{0} \\
 \mathbb{N} & \xrightarrow{s_0} & \mathbb{N}
 \end{array}$$

Notemos que ambos diagramas comutan, por lo tanto $p_0 = \bar{0}$. ■

Ejercicio 143 . Demuestra el inciso 2) del Lema anterior.

Teorema 46 . El producto de los naturales se distribuye sobre la suma.

Demostración. Lo que queremos demostrar es que

$$n \cdot (m + k) = nm + nk, \quad n, m, k \in \mathbb{N}, \quad \forall n, m, k \in \mathbb{N} .$$

Consideremos la ecuación anterior como una propiedad del natural k .

Por inducción sobre k .

Base de la inducción.

Si $k = 0$, entonces ambos lados de la ecuación valen nm .

Paso inductivo.

Supóngala cierta para k y tratemos de probarla para $k+1$.

$$\begin{aligned}
 n \cdot (m + (k + 1)) &= \\
 &= n \cdot ((m + k) + 1) && \text{por definición de } p_n \\
 &= nm + nk + n && \text{por hipótesis de inducción} \\
 &= nm + n(k + 1) . \quad \blacksquare
 \end{aligned}$$

Teorema 47 . *El producto de los naturales es asociativo.*

Demostración. Queremos demostrar que

$$(n \cdot m) \cdot k = n \cdot (m \cdot k), \quad \forall n, m, k \in \mathbb{N}$$

es decir queremos demostrar que

$$p_{n \cdot m}(k) = p_n(p_m(k)) \quad \forall n, m, k \in \mathbb{N}.$$

esta es equivalente a demostrar que las funciones

$$p_{n \cdot m} \text{ y } p_n \circ p_m$$

coinciden.

Comparemos el diagrama que define a $p_{n \cdot m}$:

$$\begin{array}{ccc} \mathbb{N} & \xrightarrow{\sigma} & \mathbb{N} \\ \bar{0} \nearrow & \downarrow p_{nm} & \downarrow p_{nm} \\ \{0\} & & \\ \bar{0} \searrow & & \\ & \mathbb{N} & \xrightarrow{s_{nm}} \mathbb{N} \end{array}$$

con el diagrama

$$\begin{array}{ccccc} \mathbb{N} & \xrightarrow{\sigma} & \mathbb{N} & & \\ \bar{0} \nearrow & \downarrow p_m & \downarrow p_m & & \\ \{0\} & \xrightarrow{\bar{0}} & \mathbb{N} & \xrightarrow{s_m} & \mathbb{N} \\ & & \bar{0} \searrow & & \\ & & \mathbb{N} & \xrightarrow{s_{nm}} & \mathbb{N} \\ & & \downarrow p_n & & \downarrow p_n \\ & & \mathbb{N} & & \end{array}$$

Para ver que el diagrama anterior es conmutativo, basta demostrar que

$$\begin{array}{ccc} \mathbb{N} & \xrightarrow{s_m} & \mathbb{N} \\ \downarrow p_n & & \downarrow p_n \\ \mathbb{N} & \xrightarrow{s_{nm}} & \mathbb{N} \end{array}$$

commuta.

Lo anterior es equivalente a decir que $\forall k \in \mathbb{N}$,

$$(p_n \circ s_m)(k) = (s_{nm} \circ p_n)(k),$$

es decir que

$$n \cdot (m + k) = nm + nk.$$

Pero esto es la propiedad distributiva. ■

Teorema 48 . *El producto en \mathbb{N} es conmutativo.*

Demostración. Queremos demostrar que

$$nm = mn, \quad \forall n, m \in \mathbb{N}.$$

Consideremos la proposición anterior, como una proposición acerca de n , $P(n)$. Vamos a demostrarla por inducción.

Base de la inducción.

Si $n = 0$ entonces

$0m = \bar{0}(m) = 0$. Además $m0 = 0$, por definición de p_m .

Paso inductivo.

Supongamos que la afirmación es cierta para n ($nm = mn$). Entonces, queremos demostrar que

$$(1 + n)m = m(n + 1).$$

Como $m(n + 1) = mn + m = nm + m$, basta entonces demostrar que para toda m :

$$Q(m) : (1 + n)m = nm + m,$$

Base de la inducción. $(1 + n) \cdot 0 = 0 = 0 \cdot n + 0$.

Paso inductivo.

Supongamos $Q(k)$, entonces

$$\begin{aligned} (1+n)(k+1) &= (1+n)k + 1 + n = \\ &= nk + k + 1 + n = (nk + n) + (k + 1) = \\ &= n(k+1) + (k+1). \blacksquare \end{aligned}$$

Lema 7 . $(n < m) \wedge (a \in \mathbb{N}^+) \Rightarrow (an < am)$.

Demostración. Como $n < m$, $\exists r \in \mathbb{N}^+$ tal que $n + r = m$.

Entonces $am = a(n+r) = an + ar$. Como $ar \neq 0$, pues es un producto de naturales distintos de 0, entonces $an < am$. ■

Lema 8 . $(an = am) \wedge (a \in \mathbb{N}^+) \Rightarrow (n = m)$.

Demostración. Se sigue inmediatamente del Lema anterior y de la propiedad de tricotomía (por ejemplo piense que sucedería si $n < m$). ■

Teorema 49 (Segundo principio de inducción matemática) . *Sea*

$$A \subseteq \mathbb{N},$$

tal que

$$(k < n \Rightarrow k \in A) \Rightarrow (n \in A), \quad \forall n \in \mathbb{N}, \quad (3.4)$$

entonces $A = \mathbb{N}$.

Demostración. Veamos que significa la proposición de arriba para $n = 0$:

$$(k < 0 \Rightarrow k \in A) \Rightarrow (0 \in A),$$

$(k < 0 \Rightarrow k \in A)$ es una proposición verdadera dado que $k < 0$ ($k \in \emptyset$) es una proposición falsa. Así que decir que

$$(k < 0 \Rightarrow k \in A) \Rightarrow (0 \in A),$$

es verdadera es lo mismo que decir que $0 \in A$.

Note que $[(k < 0 \Rightarrow k \in A) \Rightarrow (0 \in A)]$ es equivalente a $0 \in A$.

Si hubiera un natural $n \notin A$, también habría un elemento menor con la propiedad de no pertenecer a A .

Supongamos que n es el menor natural que no pertenece a A .

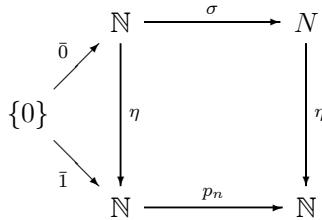
De esta manera, todos los naturales menores que n pertenecen a A . Pero por hipótesis, esto implica que $n \in A$, ∇ . ■

Nota 1 . *El segundo principio de inducción, es conocido también como principio de inducción transfinita o como principio inducción fuerte. Es muy importante señalar que es necesario demostrar la base de inducción al hacer uso de este principio.*

3.11.3 Potencias

En los siguientes ejercicios, indicaremos como definir potencias.

Ejercicio 144 . Use el Teorema de recursión para definir una función mediante el diagrama siguiente:



y note que

$$\eta(0) = 1, \eta(1) = n, \eta(2) = n \cdot n, \eta(3) = n \cdot n \cdot n, \dots$$

por lo que tenemos que es natural usar la notación

$$n^k = \eta(k).$$

Ejercicio 145 . Demuestre que $0^0 = 1$ y que $0^k = 0$ si $k \neq 0$.

Ejercicio 146 . Demuestre que $1^k = 1$, para toda $k \in \mathbb{N}$.

Ejercicio 147 . Muestre con un ejemplo que no vale en general que

$$n^m = m^n.$$

Ejercicio 148 . Muestre que no vale en general que

$$n^{(m^k)} = (n^m)^k.$$

Ejercicio 149 . Demuestre que

$$n^{(m+k)} = n^m \cdot n^k.$$

Ejercicio 150 . Demuestre que

$$(n^m)^k = n^{(m \cdot k)}.$$

Ejercicio 151 . *Se dice que un número natural n es par si $n = 2k$. En este caso decimos que k es la mitad de n y escribimos*

$$k = \frac{n}{2}.$$

Demuestre que $n(n+1)$ es un número par, para cada $n \in \mathbb{N}$.

Ejercicio 152 . *Demuestre que*

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

Ejercicio 153 . *Demuestre que*

$$n(n+1)(n+2)$$

es un múltiplo de 3, para cada $n \in \mathbb{N}$.

Ejercicio 154 . *Demuestre que*

$$(n)(n+1)(2n+1)$$

es un múltiplo de 6, para cada $n \in \mathbb{N}$.

Ejercicio 155 . *Demuestre que*

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

Ejercicio 156 . *Demuestre que si un conjunto X tiene n elementos entonces su conjunto potencia $\wp(X)$ tiene 2^n elementos.*

Ejercicio 157 . *Demuestre que el número de funciones de un conjunto A con n elementos a un conjunto B con m elementos es m^n . De tal manera que si denotamos*

$$B^A = \{f : A \longrightarrow B \mid f \text{ es una función}\}$$

podemos escribir

$$|B^A| = |B|^{|A|}.$$

Ejercicio 158 . *Demuestre que si el número de elementos del conjunto A es igual al número de elementos del conjunto B , entonces A tiene tantos subconjuntos como subconjuntos tiene B . Es decir:*

$$|A| = |B| \implies |\wp(A)| = |\wp(B)|.$$

Ejercicio 159 . *Demuestre que el conjunto de los naturales pares es un conjunto infinito (Por ejemplo, demuestre que hay tantos naturales como naturales pares).*

Ejercicio 160 . *Demuestre que si X es un conjunto finito entonces existe un número natural n que tiene tantos elementos como X . En este caso decimos que X tiene n elementos (por ejemplo 2 tiene 2 elementos).*

Ejercicio 161 . *Los nombres de los primeros siete naturales son: cero, uno, dos, tres, cuatro, cinco, seis. Demuestre, usando las definiciones que $2 + 2 = 4$ y que $2 \cdot 3 = 6$.*

Ejercicio 162 . *Demostrar que para toda $n \in \mathbb{N}$, $n > 0$:*

1. $3^{2n+2} + 2^{6n+1}$ es múltiplo de 11.
2. $3^{4n+2} + 2 * 4^{2n+1}$ es múltiplo de 11.
3. $2^{2n-1}3^{n+2} + 1$ es múltiplo de 11.
4. $3^{2n+2} - 8n - 9$ es múltiplo de 64.
5. $5^{5n+1} + 4^{5n+2} + 3^{5n}$ es múltiplo de 11.

3.12 Apéndice. Sobre las definiciones recursivas

El uso del procedimiento recursivo para definir sucesiones de significados, se basó aparentemente sólo en el quinto postulado de Peano. En ese momento se afirmó que si se tiene un elemento distinguido x_0 que puede considerarse como el significado de la expresión $\eta(n)$ para $n = 0$ (base) y se cuenta con una manera ($f : X \longrightarrow X$) de pasar del significado de cualquier número k al

significado del que sigue k' , (paso inductivo) entonces quedará bien definido el significado de la expresión para cada número natural. Es decir, existe

$$\begin{aligned}\eta : \mathbb{N} &\longrightarrow X \\ \eta(0) &= x_0 \\ \eta(k') &= f(\eta(k)).\end{aligned}$$

En efecto, se dijo, si

$$S = \{n \in \mathbb{N} \mid \eta(n) \text{ está definido}\},$$

por el paso 1 (base) $0 \in S$ y como para cada $k \in \mathbb{N}$ se puede pasar (mediante f) de $\eta(k)$ a $\eta(k')$, resulta que si $k \in S$ entonces $k' \in S$. Por lo tanto $S = \mathbb{N}$.

En la aparente obviedad del argumento anterior se esconde el hecho de que para justificar ese resultado se necesitaron todos los axiomas de Peano y no sólo el quinto (el uso de los primeros dos axiomas es muy claro) y para demostrar esta afirmación se construirán ejemplos que prueban que si en un conjunto A con un elemento inicial 0 y una función sucesor $\sigma : A \longrightarrow A$ no se cumple alguno de los otros tres axiomas de Peano, entonces, para algún conjunto X con elemento distinguido x_0 y función $f : X \longrightarrow X$, puede no existir la función $\eta : A \longrightarrow X$ con las condiciones básicas requeridas:

$$\begin{aligned}\eta(0) &= x_0 \\ \eta(\sigma(k)) &= f(\eta(k)).\end{aligned}$$

Esto justificará la necesidad de la demostración que se dio del teorema de Recursión.

Ejemplo 98 . Sean $A = \{0, 1\}$,

$$\begin{aligned}\sigma : A &\longrightarrow A \\ \sigma(0) &= 1 \\ \sigma(1) &= 1.\end{aligned}$$

En A se cumplen todos los axiomas de Peano, excepto el que asegura que σ es inyectiva.

Ahora para $X = A$, y $a_0 = 0$ y $f = \sigma$, puede verse que entonces no existe ninguna $\eta : A \rightarrow X$ tal que $\eta(0) = 0$ y $\eta(\sigma(k)) = f(\eta(k))$, ya que si tal η existiera, se tendría:

$$\eta(0) = 0 \therefore \eta(1) = \eta(\sigma(0)) = f(\eta(0)) = f(0) = \sigma(0) = 1$$

y

$$\eta(1) = \eta(\sigma(1)) = f(\eta(1)) = f(1) = 0$$

$\therefore 0 = 1$, contradicción.

Ejemplo 99 . Sean $A = \{0, 1\}$

$$\sigma : A \rightarrow A$$

$$\sigma(0) = 1$$

$$\sigma(1) = 0,$$

obviamente en A valen los axiomas de Peano efecto el que asegura que 0 no es sucesor.

Ejemplo 100 . Si se definen $X = A$, $a_0 = 0$ y

$$f : X \rightarrow X$$

$$0 \mapsto 1$$

$$1 \mapsto 1,$$

no existe $\eta : A \rightarrow X$ tal que $\eta(0) = a_0$ y tal que

$$\eta(\sigma(k)) = f(\eta(k)), \forall k \in A,$$

ya que si η fuera tal,

$$\eta(0) = 0$$

$$\eta(1) = \eta(\sigma(0)) = f(\eta(0)) = f(0) = 1$$

$$\eta(0) = \eta(\sigma(1)) = f(\eta(1)) = f(1) = 1,$$

contradicción.

Ejemplo 101 . Sean $A = \mathbb{N} \cup \{*\}$ y $\sigma(n) = \begin{cases} 1 + n & \text{si } n \in \mathbb{N} \\ * & \text{si } n = * \end{cases}$.

En A valen los axiomas de Peano excepto el quinto. Si ahora $X = \mathbb{N}$, $a_0 = 0$ y $f : X \rightarrow X$ es la función sucesor, no existe $\eta : A \rightarrow \mathbb{N}$ tal que $\eta(0) = 0$ y tal que

$$\forall k, \eta(\sigma(k)) = f(\eta(k)) \quad (3.5)$$

ya que para $n = *$, $\eta(*) \in \mathbb{N}$. Pero

$$\eta(*) = \eta(\sigma(*)) = f(\eta(*)) = 1 + (\eta(*)) \in \mathbb{N},$$

lo que es absurdo, pues se estaría asegurando que existe un número natural $(\eta(*)$) que es igual a su sucesor.

En los ejemplos que se dieron en el texto puede verse que el procedimiento inductivo que usamos permite construir los significados de tantos números n como se quiera -comenzando con cero y prosiguiendo de uno en uno- lo que induce a pensar que el método descrito, garantiza la existencia de la sucesión completa de significados $\eta(n)$, ya que si S es el conjunto de los números naturales para los cuales la expresión ha quedado definida, la base (definición explícita de $\eta(0)$) afirma que $0 \in S$. Y por el paso inductivo que permite pasar del significado (k) a $(s(k))$ se puede asegurar que si k está en S entonces $s(k)$ también; y por lo tanto en S se cumplen las hipótesis del 5º postulado de Peano, por lo que S debe ser igual a \mathbb{N} . Sin embargo, al razonar así, se comete un vicio de circularidad al utilizar una notación funcional (η) para un conjunto que todavía no existe, que está en proceso de construcción y que incluso podría no existir. El 5º postulado asegura que si hubiera una $\eta : \mathbb{N} \rightarrow \mathbb{X}$ como se prescribe en los pasos 1 y 2, sería única, pero para garantizar su existencia se requiere además, del concurso de los otros postulados de Peano.

Los ejemplos anteriores demuestran que los axiomas de Peano son indispensables para poder asegurar que existe la sucesión completa de significados para las expresiones que deseamos definir y con objeto de legitimar el procedimiento que hemos usado para construirla, haremos ver que la presencia de tales axiomas es también suficiente. Este argumento está contenido en los axiomas de recursión, y dado que el primero -que como se verá, resulta equivalente a los axiomas de Peano- es un caso particular del segundo, (recursión fuerte), sólo demostraremos este último,

Recordemos:

Teorema 50 (de Recursión Fuerte). $\forall X$ conjunto con un elemento distinguido x_0 y $\forall \{f_i\}_{i \in \mathbb{N}}$ familia de funciones de X en X , $\exists! \eta : \mathbb{N} \rightarrow X$ tal que²

1. $\eta(0) = x_0$, y
2. $\forall k \in \mathbb{N}, \eta(1+k) = f_k(\eta(k))$.

La demostración de este teorema es efectiva (se construye el objeto que el teorema asegura que existe), en este caso $\eta : \mathbb{N} \rightarrow X$.

Demostración. Sea

$$R = \{M \subseteq \mathbb{N} \times X \mid (0, x_0) \in M \text{ y } [(k, x) \in M \Rightarrow (1+k, f_k(x)) \in M]\} \quad (3.6)$$

$R \neq \emptyset$ ya que evidentemente

$$(\mathbb{N} \times X) \in R. \quad (3.7)$$

Sea $\eta = \cap R$.

Debemos demostrar que:

- 1º) η tiene dominio \mathbb{N} .
- 2º) η satisface 1) y 2).
- 3º) η es función.
- 4º) η es única con estas propiedades.

1º. Sea $S \subseteq \mathbb{N}$ el dominio de η .

$(0, x_0) \in M, \forall M \in R$ implica que $(0, x_0) \in \eta$ por lo tanto $0 \in S$.

Supongamos que $k \in S$, es decir, $\exists x \in X$ tal que $(k, x) \in \eta$ por lo que $(k, x) \in M, \forall M \in R$, lo que implica que $(1+k, f_k(x)) \in M, \forall M \in R$, por lo tanto $(1+k, f_k(x)) \in \eta$, por lo que $1+k \in S$, de donde $S = \mathbb{N}$.

2º. De la argumentación anterior se sigue obviamente que η satisface 1) y 2).

3º. Sea ahora $T \subseteq \mathbb{N}$ el conjunto de elementos del dominio de η que tienen una sola imagen.

$(0, x_0) \in \eta$ y supóngase que $(0, y) \in \eta, x_0 \neq y$. Entonces $\mu = \eta \setminus \{(0, y)\}$ satisface 1) y 2). En efecto, $(0, x_0) \in \mu$ y si $(k, x) \in \mu, ((k, x) \in \mu), (1+k, f_k(x)) \in \mu$ y como $1+k \neq 0$, entonces $(1+k, f_k(x))$ no es el que se quitó, por lo que $(1+k, f_k(x)) \in \mu$ de donde μ es un elemento de R y como tal, contiene a η , pero $\mu \subsetneq \eta$. El absurdo surge de suponer que existe tal y .

² $\exists! \eta$ significa “existe una única η ”

Es decir, $0 \in T$.

Supóngase que $k \in T$, o sea, (k, x) es la única pareja de η que tiene primera coordenada k . Entonces $(1+k, f_k(x)) \in \eta$, y supóngase que $\exists z, z \neq f_k(x)$ tal que $(1+k, z) \in \eta$. Sea ahora $\lambda = \eta \setminus \{(1+k, z)\}$. $(0, x_0) \in \lambda$ (porque $0 \neq 1+k$). Sea $(h, y) \in \lambda$. Debemos demostrar que $(1+h, f_l(y)) \in \lambda$.

Ahora, si $h = k$ no hay problema ya que entonces $1+h = 1+k$ y por lo tanto $(1+h, f_l(y))$ no es la que se quitó y por tanto está en λ .

Ahora si $h \neq k$, entonces $y = x$, la pareja que debemos mostrar que está en λ es $(1+k, f_k(x))$ que no hemos quitado, es decir, T satisface: $0 \in T$ y $k \in T \Rightarrow 1+k \in T$, por lo tanto $T = \mathbb{N}$.

4º. Sea $\theta : \mathbb{N} \rightarrow X$ una función que satisface 1) y 2), y sea

$$U = \{n \in \mathbb{N} \mid \theta(n) = \eta(n)\}. \quad (3.8)$$

Por demostrar que $U = \mathbb{N}$.

Por 1) $0 \in U$. Supóngase que $k \in U$, es decir, $\theta(k) = \eta(k)$,

Por 2) $\eta(1+k) = f_k(\eta(k)) = f_k(\eta(k)) = \theta(1+k)$. ■

Capítulo 4

Los números enteros

Las ecuaciones de la forma $a = b + x$ no siempre tienen solución en el conjunto \mathbb{N} de los números naturales -no siempre pueden resolverse-.

Con objeto de evitar esta limitación, se extiende el conjunto \mathbb{N} a otro, en el que todas las ecuaciones de la forma anterior puedan resolverse, que contenga a \mathbb{N} y que sea mínimo -en el sentido de la contención- con estas propiedades.

Se obtiene así el conjunto \mathbb{Z} de los números enteros, con un subconjunto isomorfo a los naturales, y que tiene una estructura que resultó fundamental para la Teoría de los números, básica para la Aritmética, y que sirvió de modelo para los anillos que se conocen como dominios enteros (el adjetivo “enteros” se debe precisamente al modelo).¹

¹Un anillo es una quinteta ordenada $(R, +, \cdot, 0, 1)$ tal que:

1. R es un conjunto,
2. $+, \cdot$ son operaciones asociativas en R , con neutros respectivos $0, 1$.
3. $+$ es conmutativa.
4. $\forall r, s, t \in R$ se tiene que
 - (a) $r \cdot (s + t) = (r \cdot s) + (r \cdot t)$.
 - (b) $(s + t) \cdot r = s \cdot r + t \cdot r$.

Un dominio entero es un anillo donde el producto es conmutativo, donde $1 \neq 0$, y donde se pueden cancelar factores distintos de 0.

Un campo es un anillo con producto conmutativo en donde $1 \neq 0$ y en donde cada elemento distinto de 0, posee inverso multiplicativo.

Los campos son dominios enteros.

Los primeros párrafos de este capítulo toman en cuenta las observaciones anteriores y explican porqué se escogió la relación de equivalencia con la que se construyen los enteros -las clases de equivalencia inducidas por la relación-.

4.1 Construcción y definiciones

Supongamos que $n > m$, $n, m \in \mathbb{N}$. Entonces

$$\exists k \in \mathbb{N}^+. \exists. n = m + k.$$

(. \exists . se lee: “tal que”). Así que k es la solución de la ecuación

$$n = m + x.$$

Dados dos naturales n, m , la ecuación $n = m + x$, tiene solución en \mathbb{N} $\Leftrightarrow n \geq m$.

Cuando $n < m$, la ecuación $n = m + x$ no tiene solución en los naturales. Esto se debe a que

$$\forall k \in \mathbb{N}, m + k \geq m > n,$$

por lo que $m + k \neq n, \forall k \in \mathbb{N}$.

Por otra parte, 3 es solución de

$$7 = 4 + x,$$

pero también es solución de

$$16 = 13 + x.$$

¿Cuándo pasa que $a = b + x$ tiene las mismas soluciones que $c = d + x$?

Observemos lo siguiente:

Si $a = b + k$ y $c = d + k$, para la misma $k \in \mathbb{N}$, entonces

$$\begin{array}{rcl} a & = & b + k \\ d + k & = & c \end{array},$$

así que sumando cada lado, tenemos que $a + d + k = b + k + c$. Cancelando k , obtenemos

$$a + d = b + c.$$

Recíprocamente, si

$$a + d = b + c$$

y

k es solución de $a = b + x$,

entonces

$$a = b + k.$$

Sumando d , tenemos $a + d = b + k + d$, así que $b + c = b + k + d$. Cancelando b obtenemos $c = d + k$. Por lo que

k también es solución de $c = d + x$.

Entonces, la solución de $a = b + x$, es la solución de $c = d + x$ ($a > b$, $c > d$).

Podemos observar lo siguiente:

$a = b + x$ y $c = d + x$ tienen la misma solución $\iff a + d = b + c$.

Notemos que la ecuación $a = b + x$ está determinada por la pareja (a, b) . Inclusive podemos pensar que (a, b) es una abreviatura para la ecuación. Siguiendo con esta manera de pensar ¿por qué no decir que (a, b) está relacionada con (c, d) ($(a, b) \sim (c, d)$), en lugar de decir que las ecuaciones $a = b + x$ y $c = d + x$ tienen la misma solución?

Ya sabemos que esto pasa si y sólo si $b + d = a + c$.

O bien, comenzando de nuevo, ¿por qué no definir la relación \sim en $\mathbb{N} \times \mathbb{N}$ por

$$(a, b) \sim (c, d) \text{ si y sólo si } a + d = b + c.$$

Una vez demostrado que \sim es una relación de equivalencia, denotando

$$[(a, b)]_\sim = \{(c, d) \mid (a, b) \sim (c, d)\},$$

podemos pensar que éste es el conjunto de todas las ecuaciones que tienen la misma solución, y de hecho podríamos identificar $[(a, b)]_\sim$ con la solución de $a = b + x$.

En vista de lo anterior, hacemos la siguiente definición.

Definición 56

1. Un número entero es una clase $[(a, b)]_\sim$ definida por:

$$[(a, b)]_\sim = \{(c, d) \mid (a, b) \sim (c, d)\} = \{(c, d) \in \mathbb{N} \times \mathbb{N} \mid a + d = b + c\}.$$

2. El conjunto de los enteros se denota por \mathbb{Z} , y por definición,

$$\mathbb{Z} = \mathbb{N} \times \mathbb{N} / \sim = \{ [(a, b)]_{\sim} \mid (a, b) \in \mathbb{N} \times \mathbb{N} \}.$$

Vamos ahora a definir la suma de los enteros para que parezca más natural llamarlos “números”.

Supongamos que $m > n$ y que $a > b$. Escribamos

$$a = b + k \text{ y } m = n + l.$$

Sumando, tenemos que $a + m = b + n + (k + l)$. Así que $k + l$ es la solución de

$$a + m = b + n + x$$

(esta ecuación corresponde a la pareja $(a + m, b + n)$). Esto sugiere que se puede definir la suma de enteros de la manera siguiente:

Definición 57 . Se define $\hat{+} : \mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{Z}$, mediante,

$$[(a, b)]_{\sim} \hat{+} [(m, n)]_{\sim} = [(a + m, b + n)]_{\sim}.$$

(Hay que ver que esta definición es buena).²

² $[(a, b)]_{\sim} \hat{+} [(m, n)]_{\sim} = [(a + m, b + n)]_{\sim}$ es una buena definición:

Supongamos que $[(a, b)]_{\sim} = [(\alpha, \beta)]_{\sim}$ y que $[(m, n)]_{\sim} = [(\mu, \nu)]_{\sim}$.

Queremos demostrar que

$$[(a, b)]_{\sim} \hat{+} [(m, n)]_{\sim} = [(\alpha, \beta)]_{\sim} \hat{+} [(\mu, \nu)]_{\sim}.$$

Tenemos que $a + \beta = \alpha + b$ y que $m + \nu = n + \mu$.

Entonces,

$$(a + \beta) + (m + \nu) = (\alpha + b) + (n + \mu),$$

es decir,

$$(a + m) + (\beta + \nu) = (b + n) + (\alpha + \mu),$$

es decir,

$$[(a + m, b + n)]_{\sim} = [(\alpha + \mu, \beta + \nu)]_{\sim}.$$

Es una consecuencia de la asociatividad de la suma de los naturales, que esta nueva suma también es asociativa.

También debe ser claro que $\hat{+}$ tiene neutro:

$$[(0, 0)]_{\sim} = [(k, k)]_{\sim}.$$

Por otra parte, si sumamos $[(a, b)]_{\sim}$ con $[(b, a)]_{\sim}$, obtenemos

$$[(a + b, b + a)]_{\sim} = [(0, 0)]_{\sim}.$$

Por lo que decimos que

$[(b, a)]_{\sim}$ es el inverso aditivo de $[(a, b)]_{\sim}$,

y escribimos

$$[(b, a)] = -[(a, b)]_{\sim}.$$

4.2 El orden en \mathbb{Z}

4.2.1 Los enteros positivos

En \mathbb{Z} definimos el orden definiendo la clase \mathcal{P} de los positivos:

$$[(a, b)]_{\sim} \in \mathcal{P} \Leftrightarrow a > b.$$

Donde $>$ es el orden en \mathbb{N} .

Podemos notar que esta es una buena definición. Es decir que si $[(a, b)]_{\sim} = [(\alpha, \beta)]_{\sim}$ entonces $a > b \implies \alpha > \beta$. Esto se debe a que por una parte $a + \beta = b + \alpha$ y por otra $a = b + d$ con $d \in \mathbb{N} \setminus \{0\}$. Entonces $b + d + \beta = b + \alpha$, cancelando b tenemos que $d + \beta = \alpha$, es decir que $\alpha > \beta$.

Observación 47 . *Notemos que es una consecuencia de la tricotomía en el orden de los naturales que para cada $z \in \mathbb{Z}$ sucede una y sólo una de las siguientes proposiciones:*

1. $[(a, b)]_{\sim} \in \mathcal{P}$.
2. $[(a, b)]_{\sim} = [(0, 0)]_{\sim} (\Leftrightarrow a = b)$.
3. $-[(a, b)]_{\sim} = [(b, a)]_{\sim} \in \mathcal{P}$.

Así que hacemos la siguiente definición.

$$[(a, b)]_{\sim} > [(m, n)]_{\sim} \text{ si } [(a, b)]_{\sim} - [(m, n)]_{\sim} =: [(a, b)]_{\sim} \hat{+} (-(m, n)) \in \mathcal{P}.$$

Proposición 15. $[(a, b)]_{\sim} > [(m, n)] \Leftrightarrow [(a, b)]_{\sim} \hat{+} [(n, m)]_{\sim} = [(a + n, b + m)]_{\sim} \in \mathcal{P} \Leftrightarrow a + n > b + m, \text{ en } \mathbb{N}.$

De paso notemos que

$$[(a, b)]_{\sim} \in \mathcal{P} \Leftrightarrow [(a, b)]_{\sim} > [(0, 0)]_{\sim}.$$

En lo que sigue, para simplificar la notación escribiremos $\overline{(a, b)}$ en lugar de $[(a, b)]_{\sim}$.

Proposición 16. $\forall z \in \mathbb{Z}, \text{ pasa una y sólo una de las siguientes afirmaciones:}$

1. $z = \overline{(n, 0)}$, con $n \in \mathbb{N}^+$.
2. $z = \overline{(0, 0)}$.
3. $z = \overline{(0, n)}$, con $n \in \mathbb{N}^+$.

Lo anterior es una consecuencia de la tricomía del orden en \mathbb{N} .

Ejercicio 163. *Demuestre la proposición 16.*

4.3 Inmersión de los naturales en los enteros

Observación 48. *Notemos que la función*

$$\begin{array}{ccc} \mathbb{N} & \xrightarrow{\gamma} & \mathbb{N} \times \mathbb{N} / \sim \\ n & \mapsto & [(n, 0)]_{\sim} \end{array}$$

es una función inyectiva que respeta la suma, el neutro aditivo y el orden.

Demostración. Inyectividad) Si $n \neq m$, entonces $[(n, 0)]_{\sim} \neq [(m, 0)]_{\sim}$, pues $n + 0 \neq 0 + m$.

Se respeta la suma) $\gamma(n + m) = [(n + m, 0)]_{\sim} = [(n, 0)]_{\sim} \hat{+} [(m, 0)]_{\sim}$.

Se respeta el orden) Si $n < m$, entonces $[(n, 0)]_{\sim} < [(m, 0)]_{\sim}$, pues $n + 0 < 0 + m$. ■

Por comodidad de notación, denotaremos $\overline{(n, m)} = [(n, m)]_{\sim}$.

Observación 49 . *La biyección*

$$\begin{array}{ccc} \mathbb{N} & \xrightarrow{\quad} & \left\{ \overline{(n, 0)} \mid n \in \mathbb{N} \right\} \\ n & \mapsto & \overline{(n, 0)} \end{array}$$

permite identificar los naturales con los enteros no negativos.

Además,

$$\overline{(n, 0)} \hat{+} \overline{(0, n)} = \overline{(n, n)} = \overline{(0, 0)}.$$

Así que $\overline{(0, n)}$ es el inverso aditivo de $\overline{(n, 0)}$ y escribiremos

$$\overline{(0, n)} = -\overline{(n, 0)}.$$

Si n, m son naturales, y $n < m$ (entonces $\exists k \in \mathbb{N}^+$ tal que $n + k = m$) y consideramos la ecuación

$$\overline{(n, 0)} = \overline{(m, 0)} \hat{+} \overline{(x, y)} = \overline{(m + x, y)}$$

que es equivalente con

$$n + y = m + x$$

(una de cuyas soluciones es $y = k$ y $x = 0$), tenemos que

$$\begin{array}{ccccccc} \overline{(n, 0)} & = & \overline{(m, 0)} & + & \overline{(0, k)} \\ \downarrow & & \downarrow & & \downarrow \\ n & & m & + & -k \end{array}$$

y así la ecuación

$$n = m + x$$

tiene la solución

$$\overline{(0, k)} = -\overline{(k, 0)} \text{ en } \mathbb{Z}.$$

4.4 El producto en \mathbb{Z}

Supongamos que tenemos $a, b, c, d, k, l \in \mathbb{N}$ tales que

$$a = b + k \tag{4.1}$$

(es decir que k es solución de $a = b + x$)

$$c = d + l \quad (4.2)$$

(es decir que l es solución de $c = d + x$), entonces, multiplicando la primera ecuación por l tenemos que

$$al = bl + kl,$$

sumando bd ,

$$al + bd = bl + bd + kl \quad (4.3)$$

notemos ahora que multiplicando 4.2 por b , tenemos que $bc = bd + bl$, sustituyendo en 4.3,

$$al + bd = bc + kl \quad (4.4)$$

sumemos ahora ad :

$$ad + al + bd = ad + bc + kl \quad (4.5)$$

Notemos ahora que

$$a(\underbrace{d + l}_c) + bd = ad + al + bd = ad + bc + kl \quad (4.6)$$

de donde

$$ac + bd = ad + bc + kl \quad (4.7)$$

así que kl es solución de $ac + bd = ad + bc + x$. Esto sugiere que debemos definir

$$\overline{(a, b)} \cdot \overline{(c, d)} \doteqdot \overline{(ac + bd, ad + bc)}.$$

Definición 58

$$\begin{array}{ccc} \mathbb{Z} \times \mathbb{Z} & \xrightarrow{\cdot} & \mathbb{Z} \\ \overline{(a, b)}, \overline{(c, d)} & \mapsto & \overline{(ac + bd, ad + bc)} \end{array}.$$

Observación 50 . *El producto está bien definido.*

Demostración. Supongamos que

$$\overline{(a, b)} = \overline{(\alpha, \beta)} \text{ y que } \overline{(c, d)} = \overline{(\lambda, \delta)}.$$

Entonces

$$a + \beta = b + \alpha \text{ y } c + \delta = d + \lambda.$$

Entonces

$$\overline{(a, b) \cdot (c, d)} = \overline{(ac + bd, ad + bc)}$$

y

$$\overline{(\alpha, \beta)(\lambda, \delta)} = \overline{(\alpha\lambda + \beta\delta, \alpha\delta + \beta\lambda)}.$$

Queremos demostrar que

$$\overline{(ac + bd, ad + bc)} = \overline{(\alpha\lambda + \beta\delta, \alpha\delta + \beta\lambda)}.$$

Es decir queremos demostrar que

$$ac + bd + \alpha\delta + \beta\lambda = ad + bc + \alpha\lambda + \beta\delta.$$

Pero

$$\begin{aligned} ac + bd + \alpha\delta + \beta\lambda &= ad + bc + \alpha\lambda + \beta\delta \Leftrightarrow \\ &\Leftrightarrow \beta c + ac + bd + \alpha\delta + \beta\lambda = \beta c + ad + bc + \alpha\lambda + \beta\delta \\ &\Leftrightarrow (\beta + a) c + bd + \alpha\delta + \beta\lambda = \beta c + ad + bc + \alpha\lambda + \beta\delta \\ &\Leftrightarrow (b + \alpha) c + bd + \alpha\delta + \beta\lambda = \beta c + \beta\delta + ad + bc + \alpha\lambda \\ &\Leftrightarrow (b + \alpha) c + bd + \alpha\delta + \beta\lambda = \beta(c + \delta) + ad + bc + \alpha\lambda \Leftrightarrow \\ &\Leftrightarrow (b + \alpha) c + bd + \alpha\delta + \beta\lambda = \beta(d + \lambda) + ad + bc + \alpha\lambda \Leftrightarrow \\ &\Leftrightarrow \alpha c + bd + \alpha\delta = \beta d + ad + \alpha\lambda \Leftrightarrow \\ &\Leftrightarrow \alpha(c + \delta) + bd = (\beta + a)d + \alpha\lambda \Leftrightarrow \\ &\Leftrightarrow \alpha(d + \lambda) + bd = (\beta + a)d + \alpha\lambda \Leftrightarrow \\ &\Leftrightarrow \alpha d + bd = (\alpha + b)d. \end{aligned}$$

■

Observación 51 . En particular, $\overline{(a, 0)} \cdot \overline{(0, b)} = \overline{(0, ab)}$, $\overline{(0, b)} \cdot \overline{(0, d)} = \overline{(bd, 0)}$, $\overline{(a, 0)} \cdot \overline{(c, 0)} = \overline{(ac, 0)}$. Además, \cdot es conmutativo.

Ahora podemos demostrar las siguientes propiedades del producto de \mathbb{Z} .

Proposición 17 . El producto

$$\begin{array}{ccc} \mathbb{Z} \times \mathbb{Z} & \xrightarrow{\cdot} & \mathbb{Z} \\ \left(\overline{(a, b)}, \overline{(c, d)} \right) & \mapsto & \overline{(ac + bd, ad + bc)} \end{array}$$

tiene las siguientes propiedades:

1. Está bien definido.
2. Es asociativo.
3. Es commutativo.
4. Tiene neutro.
5. Los únicos elementos que tienen inverso multiplicativo son 1 y -1 .
6. El producto se distribuye sobre la suma.
7. Se pueden cancelar factores distintos de cero.
8. Producto de positivos es positivo.
9. La multiplicación por un positivo respeta el orden.
10. La multiplicación por un negativo (es decir el inverso aditivo de un positivo) invierte el orden.

En adelante suprimiremos el punto para las multiplicaciones y simplificaremos el símbolo de suma en \mathbb{Z} , denotándola con $+$.

Demostración. 1) Que el producto en \mathbb{Z} está bien definido lo acabamos de ver.

2)

$$\overline{(a, b)} \left(\overline{(c, d)} \, \overline{(n, m)} \right) = \overline{(a, b)(cn + dm, cm + dn)} =$$

$$= \overline{(a(cn + dm) + b(cm + dn), a(cm + dn) + b(cn + dm))} =$$

$$= \overline{(acn + adm + bcm + bdn, acm + adn + bcn + bdm)}$$

Por otra parte,

$$\left(\overline{(a, b)(c, d)} \right) \overline{(n, m)} = \overline{(ac + bd, ad + bc)(n, m)} =$$

$$= \overline{((ac + bd)n + (ad + bc)m, (ac + bd)m + (ad + bc)n)} =$$

$$= \overline{(acn + adm + bcm + bdn, acm + adn + bcn + bdm)}.$$

3) $\overline{(a, b)(c, d)} = \overline{(ac + bd, ad + bc)}$ y $\overline{(a, b)(c, d)} = \overline{(ca + db, cb + da)}.$

Así que la commutatividad del producto de los enteros es una consecuencia de la commutatividad de las operaciones en \mathbb{N} .

4) $\overline{(a, b)(1, 0)} = \overline{(a + b0, a0 + b1)} = \overline{(a, b)}.$

5) $\overline{(1, 0)} = \overline{(a, b)(c, d)} = \overline{(ac + bd, ad + bc)} \Rightarrow$

$$\begin{aligned} ac + bd &= 1 \\ ad + bc &= 0 \end{aligned}$$

tomando en cuenta que $a, b, c, d \in \mathbb{N}$, entonces de $ad + bc = 0$, se tiene que $ad = 0$ y $bc = 0$. Si $a = 0$, entonces $bd = 1$. Por lo que $b = 1 = d$. Así que $(a, b) = (0, 1) = -1_{\mathbb{Z}}$ (aquí, $1_{\mathbb{Z}} = (1, 0)$). Si $a \neq 0$, entonces $d = 0$, así que $ac = 1$. Por consiguiente, $a = 1 = c$. De donde se tiene que $b = 0 + b = d + b = ad + bc = 0$. Por lo tanto $\overline{(a, b)} = \overline{(1, 0)}.$

6)

$$\overline{(a, b)} \left(\overline{(c, d)} + \overline{(m, n)} \right) = \overline{(a, b)(c + m, d + n)} =$$

$$= \overline{(a(c + m) + b(d + n), a(d + n) + b(c + m))} =$$

$$= \overline{(ac + am + bd + bn, ad + an + bc + bm)}.$$

Por otra parte,

$$\overline{(a, b)(c, d)} + \overline{(a, b)(m, n)} = \overline{(ac + bd, ad + bc)} + \overline{(am + bn, an + bm)} =$$

$$= \overline{(ac + am + bd + bn, ad + an + bc + bm)}.$$

7) Supongamos que $\overline{(a, b)(c, d)} = \overline{(a, b)(m, n)}$ y que $\overline{(a, b)} \neq \overline{(0, 0)}$ (que es lo mismo que decir que $a \neq b$).

Entonces $\overline{(ac + bd, ad + bc)} = \overline{(am + bn, an + bm)}$, por lo que $ac + bd + an + bm = ad + bc + am + bn$. Como $a \neq b$, supongamos que $a > b$ (el caso $a < b$

es simétrico) (En este caso $a = b + z$, $z \in \mathbb{N}$ y denotamos $z = a - b \in \mathbb{N}$). Entonces

$$\begin{aligned} ac + bd + an + bm &= ad + bc + am + bn \Leftrightarrow \\ &\Leftrightarrow (a - b)c + (a - b)n = (a - b)d + (a - b)m \Leftrightarrow \\ &\Leftrightarrow (a - b)(c + n) = (a - b)(d + m) \Leftrightarrow \\ &\Leftrightarrow c + n = d + m \\ &\Leftrightarrow \overline{(c, d)} = \overline{(m, n)}. \end{aligned}$$

8) Si $\overline{(a, b)} \in \mathcal{P}$, $\overline{(c, d)} \in \mathcal{P}$, entonces $a > b, c > d$. Así que $\overline{(a, b)(c, d)} = \overline{(ac + bd, ad + bc)} \in \mathcal{P}$.

Esto último es porque $(ac + bd) > (ad + bc)$, pues tenemos que

$$a > b, c > d \geq 0 \text{ y } c + d > 0,$$

de donde $a(c + d) > b(c + d)$.

9) y 10) Son inmediatas. ■

4.5 El algoritmo de la división

Definición 59 . Si $b \in \mathbb{Z}$,

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

Teorema 51 . Si $a, b \in \mathbb{Z}$ y $b \neq 0$, entonces $\exists! (q, r) \in \mathbb{Z} \times \mathbb{Z}$, $0 \leq r < |b|$, tales que

$$a = bq + r.$$

Lo anterior se lee de la manera siguiente: para cada a entero y b entero distinto de cero, existen enteros únicos q y r (r mayor o igual que cero y menor que el valor absoluto de b) tales que $a = bq + r$.

Demostración. Como la proposición anterior es un poco larga, es mejor escribir el diagrama siguiente:

$$n \overline{q} \\ r \quad 0 \leq r < n.$$

Consideremos el subconjunto de los enteros

$$\{a - bq \mid q \in \mathbb{Z}\}.$$

Por simplificar, denotemos el conjunto anterior, de la siguiente manera:

$$a - b\mathbb{Z}.$$

Primero veremos que en este conjunto hay enteros positivos.

Si

$$a - b\mathbb{Z} \subseteq \{z \in \mathbb{Z} \mid z \leq 0\},$$

entonces el conjunto de los inversos de estos enteros estaría incluído en $P \cup \{0\}$, el conjunto de los enteros no negativos. Es decir,

$$-a + b\mathbb{Z} \subseteq P \cup \{0\}.$$

Por el principio del buen orden, en $-a + b\mathbb{Z}$ habría un elemento menor, $k = -a + bz$, digamos. Pero es claro que $-a + bz - |b| < k$ ∇ .

Entonces en $a - b\mathbb{Z}$ existen elementos ≥ 0 .

Escojamos el menor elemento no negativo de $a - b\mathbb{Z}$ y llamémosle r . Entonces

$$0 \leq r = a - bq.$$

Si $r = 0$ entonces hemos terminado. Veamos pues, que en el caso de que $r > 0$, sucede también que $r < |b|$. En caso contrario, tendríamos que $a - bq = r \geq |b|$, pero en este caso, restando $|b|$ tenemos que

$$0 \leq a - bq - |b| = r - |b| < r \nabla.$$

Esta contradicción a la elección de r , muestra que si $r > 0$, entonces $r < |b|$. Por lo tanto

$$0 \leq r < |b| \text{ y } a = bq + r,$$

como queríamos.

Unicidad.

Supongamos que $a = bq + r$ y que también $a = bq' + r'$, con $0 \leq r, r' < |b|$. Notemos que basta demostrar que $r = r'$, pues en este caso,

$$bq + r = bq' + r' \Rightarrow bq = bq' \Rightarrow q = q'.$$

Si los enteros r, r' son distintos, uno de ellos tendría que ser el menor. Supongamos sin perder generalidad que

$$r < r'.$$

Entonces

$$0 < r' - r < r' < |b|.$$

Pero también

$$0 < r' - r = bq' - a - bq + a = b(q - q') = |b| |(q - q')| \geq |b| \nabla.$$

La última desigualdad se sigue de que $q - q'$ no es 0, así que tiene que ser ≥ 1 . ■

4.6 Divisibilidad y congruencias

En \mathbb{Z} se define la resta de la manera siguiente: $z - w = z + (-w)$.

Ejercicio 164 . *Muestre que la resta en \mathbb{Z} no es conmutativa, ni asociativa.*

4.6.1 Subconjuntos de \mathbb{Z} cerrados bajo la resta.

Nos interesan los subconjuntos no vacíos de \mathbb{Z} cerrados bajo la resta.

Observación 52 . *Si un subconjunto de \mathbb{Z} cerrado bajo la resta tiene algún elemento distinto de 0, entonces también tiene un elemento positivo.*

Demostración. Sea

$$\emptyset \neq S \subseteq \mathbb{Z}$$

cerrado bajo la resta.

Sea $z \in S, z \in \mathbb{Z} \setminus \{0\}$, entonces $0 = z - z \in S$. Por lo tanto $-z = 0 - z \in S$. Como z y $-z$ están ambas en S , entonces S contiene un elemento positivo. ■

Observación 53 . *Si un subconjunto S de \mathbb{Z} es cerrado bajo la resta, entonces S es cerrado bajo la suma.*

Demostración. Si S es vacío, no hay nada que demostrar. (¿Por qué?).

Sea $\emptyset \neq S \subseteq \mathbb{Z}$ cerrado bajo la resta. Si z_1, z_2 son elementos de S , entonces (como se vió arriba) $-z_2 \in S$. Así que

$$z_1 + z_2 = z_1 - (-z_2) \in S.$$

(Note que $-z_2 + z_2 = 0 \Rightarrow z_2 = -(-z_2)$). ■

Ejercicio 165 . Muestre que la operación vacía ($\emptyset \times \emptyset \longrightarrow \emptyset$) es asociativa.

En lo que siguearemos uso de la siguientes notaciones:

Notación 6

1. $\mathbb{Z} x = x \mathbb{Z} =: \{zx \mid z \in \mathbb{Z}\}$, si x es un entero.
2. $w + \mathbb{Z} x =: \{w + zx \mid z \in \mathbb{Z}\}$, $w, x \in \mathbb{Z}$.
3. $\mathbb{Z} a + \mathbb{Z} b =: \{z_1a + z_2b \mid z_1, z_2 \in \mathbb{Z}\}$, $a, b \in \mathbb{Z}$. $z_1a + z_2b$ se llama una combinación entera de a y b .

Observación 54 . Si un subconjunto S de \mathbb{Z} es cerrado bajo la resta, entonces cuando $x \in S$, todo múltiplo de x también está en S . Es decir que

$$x \in S \Rightarrow \mathbb{Z} x \subseteq S.$$

Demostración. Sea $\emptyset \neq S \subseteq \mathbb{Z}$ cerrado bajo la resta.

Primero veremos, por inducción, que los múltiplos naturales de x pertenecen a S .

Si x es un elemento de S , entonces (como se vió arriba) $0 \in S$.

Si $nx \in S$, entonces $nx + x \in S$.

Lo anterior muestra por inducción, que un múltiplo natural de x pertenece a S . Por otra parte, el inverso aditivo de un elemento en S también pertenece a S . Por lo tanto, todos los múltiplos enteros de x pertenecen a S . ■

Observación 55 . Si un subconjunto $S \neq \emptyset$ de \mathbb{Z} es cerrado bajo la resta, entonces $\exists n \in \mathbb{N}$ tal que $S = n \mathbb{Z}$.

Demostración. Si $S = \{0\}$, entonces $S = 0 \mathbb{Z}$.

Si $S \neq \{0\}$, entonces, como vimos en la observación de arriba, en S hay un elemento positivo. Así que $S \cap \mathcal{P} \neq \emptyset$, donde \mathcal{P} denota el conjunto de enteros positivos. Por el principio del Buen Orden, $S \cap \mathcal{P}$ tiene un elemento menor, al que llamamos n . Así que $n \in S$ y como vimos en la observación anterior, $n \mathbb{Z} \subseteq S$.

Tomemos ahora un elemento cualquiera z de S , apliquémosle el algoritmo de la división como en el diagrama

$$n \overline{)z} \quad , \quad r \quad 0 \leq r < n$$

es decir que $z = nq + r$, o bien, $r = z - nq$. Podemos observar que z y nq son dos elementos de S , así que r también pertenece a S .

Si r fuera positivo entonces sería menor que el menor positivo de S que es $n \overset{\circ}{\nabla}$. Por lo tanto $r = 0$, que es lo mismo que decir que $z = nq$. Por lo tanto $S \subseteq n \mathbb{Z}$. ■

Definición 60. Un subconjunto I de \mathbb{Z} no vacío y cerrado bajo la resta se llama un ideal de \mathbb{Z} .

Ejercicio 166. Muestre que son equivalentes para $I \subseteq \mathbb{Z}$:

1. I es un ideal.

- (a) $0 \in I$.
- (b) I es cerrado bajo la suma.
- (c) $a \in I \implies a \mathbb{Z} \subseteq I$.

Teorema 52. Si $\{S_\alpha\}_{\alpha \in X}$ es una familia de subconjuntos no vacíos de \mathbb{Z} cerrados bajo la resta, entonces $\cap \{S_\alpha\}_{\alpha \in X}$ también es un subconjunto no vacío cerrado bajo la resta.

Demostración. De la última observación se sigue que

$$0 \in S_\alpha, \forall \alpha \in X.$$

Por lo tanto

$$0 \in \cap \{S_\alpha\}_{\alpha \in X}.$$

Ahora, si

$$n, m \in \cap \{S_\alpha\}_{\alpha \in X}, \text{ entonces } n - m \in S_\alpha, \forall \alpha \in X.$$

Así que

$$n - m \in \cap \{S_\alpha\}_{\alpha \in X}.$$

■

1. De lo anterior se sigue que para cada subconjunto S de \mathbb{Z} existe un subconjunto no vacío de \mathbb{Z} cerrado bajo la resta, mínimo con la propiedad de contener a S . A saber:

$$\cap \{Y \mid S \subseteq Y \subseteq \mathbb{Z}, \emptyset \neq Y, Y \text{ cerrado bajo la resta}\}.$$

2. Notar que \mathbb{Z} es un subconjunto de \mathbb{Z} cerrado bajo la resta que contiene a cualquier $S \subseteq \mathbb{Z}$.

Notación 7 . Por abreviar, escribamos

$$\langle S \rangle = \cap \{Y \mid S \subseteq Y, \emptyset \neq Y \text{ y } Y \text{ cerrado bajo la resta}\}.$$

1. $\langle \{2\} \rangle = 2\mathbb{Z}$.
2. $\langle \{0\} \rangle = 0\mathbb{Z} = \{0\}$.
3. $\langle \{n\} \rangle = \langle \{-n\} \rangle = n\mathbb{Z}$.
4. $\langle \emptyset \rangle = \{0\} = 0\mathbb{Z}$.
5. $\langle \{15, 6\} \rangle = 3\mathbb{Z}$:
Es claro que $15, 6 \in 3\mathbb{Z}$, y es claro que $3\mathbb{Z}$ es cerrado bajo la resta, por lo tanto $\langle \{15, 6\} \rangle \subseteq 3\mathbb{Z}$. Ahora, todo subconjunto \mathcal{S} de \mathbb{Z} cerrado bajo la resta debe contener a $3 = 15 - (2 \cdot 6)$, y así debe contener también a todo múltiplo entero de 3, es decir, $3\mathbb{Z} \subseteq \mathcal{S}$. En particular, $3\mathbb{Z} \subseteq \langle \{15, 6\} \rangle$.
6. $\langle \{n, m\} \rangle = \{n \cdot z_1 + m \cdot z_2 \mid z_1, z_2 \in \mathbb{Z}\}$.

Es claro que

$$\{n \cdot z_1 + m \cdot z_2 \mid z_1, z_2 \in \mathbb{Z}\}$$

es un subconjunto de \mathbb{Z} cerrado bajo la resta que contiene a n y a m . Por lo tanto³

$$\langle\{n, m\}\rangle \subseteq \{n \cdot z_1 + m \cdot z_2 \mid z_1, z_2 \in \mathbb{Z}\}.$$

Por otra parte, un subconjunto cerrado bajo la resta que contenga tanto a n como a m , debe contener a los múltiplos enteros de cada una y en consecuencia, debe contener las sumas de ellos. En particular,

$$\{n \cdot z_1 + m \cdot z_2 \mid z_1, z_2 \in \mathbb{Z}\} \subseteq \langle\{n, m\}\rangle.$$

Por la observación 55,

$$\langle\{n, m\}\rangle = d\mathbb{Z}, \text{ para alguna } d \geq 0.$$

Notemos que $d = 0 \Leftrightarrow n = 0 = m$.

Veremos algunas propiedades de esta d en el teorema 53.

Definición 61 . *Definimos la relación “|” en \mathbb{Z} por:*

$$n \mid m \Leftrightarrow m \in n\mathbb{Z}.$$

(Es decir: n divide a m si m es un múltiplo de n).

Observación 56 . $n \mid m \Leftrightarrow m\mathbb{Z} \subseteq n\mathbb{Z}$.

Demostración. $\Rightarrow)$ $n \mid m \Rightarrow m \in n\mathbb{Z}$, pero como $n\mathbb{Z}$ es un ideal, $m \in n\mathbb{Z} \Rightarrow m\mathbb{Z} \subseteq n\mathbb{Z}$.

$\Leftarrow)$ Si $m\mathbb{Z} \subseteq n\mathbb{Z}$ entonces $m = m \cdot 1 \in n\mathbb{Z}$. Así que $n \mid m$. ■

Proposición 18 . *La relación “|” en \mathbb{Z} es*

1. reflexiva.
2. transitiva.
3. $\forall a, b \in \mathbb{Z},$
 $((a \mid b) \wedge (b \mid a)) \Rightarrow b \in \{a, -a\}.$

³Recuérdese que como consecuencia de la definición de $\langle S \rangle$, $\langle S \rangle$ es el menor subconjunto no vacío de \mathbb{Z} cerrado bajo la resta que contiene a S ..

Demostración. 1) $a = 1 \cdot a$.

2) $((a \mid b) \wedge (b \mid c)) \Rightarrow (b = ax \wedge c = by)$, para algunas $x, y \in \mathbb{Z}$. Entonces

$$c = by = (ax)y = a(xy),$$

por lo que $a \mid c$.

3) $(b = ax \wedge a = by, x, y \in \mathbb{Z}) \Rightarrow a = by = axy$.

Entonces

$$\begin{cases} a = 0 \text{ y entonces } b = 0 \text{ } (0 \mid b \Rightarrow b = 0). \text{ Ó} \\ a \neq 0 \text{ en cuyo caso } 1 = xy. \end{cases}$$

En el segundo caso, $x \in \{1, -1\}$. Ya que

$$x > 1 \Rightarrow y > 0,$$

por lo que $xy > y > 0 \Rightarrow xy \geq y + 1 > 1$. Por lo tanto $x \leq 1$. Análogamente, se tiene que $x \geq -1$. Como $x \neq 0$, entonces $x \in \{1, -1\}$.

Luego

$$b \in \{a, -a\}.$$

■

Ejercicio 167 . Demostrar que la restricción de la relación “ \mid ” a \mathbb{N} , es una relación de orden.

Teorema 53 . Si $n \neq 0$ o $m \neq 0$, y $\langle\{n, m\}\rangle = d\mathbb{Z}$, ($d \geq 0$), entonces d tiene las siguientes propiedades:

1. $d > 0$.
2. $d \mid n \wedge d \mid m$.
3. $(k \mid n) \wedge (k \mid m) \Rightarrow k \mid d$.

Demostración. Como $d \neq 0$ y $d \geq 0$, entonces $d > 0$.

Ahora, es claro que

$$n \in d\mathbb{Z} \Leftrightarrow d \mid n,$$

por lo tanto

$$d \mid n.$$

Análogamente,

$$d \mid m.$$

Ahora,

$$((k \mid n) \wedge (k \mid m)) \Leftrightarrow ((n \in k\mathbb{Z}) \wedge (m \in k\mathbb{Z})),$$

también,

$$\begin{aligned} ((n \in k\mathbb{Z}) \wedge (m \in k\mathbb{Z})) &\Rightarrow \{n, m\} \subseteq k\mathbb{Z} \Rightarrow \\ &\Rightarrow d\mathbb{Z} = \langle \{n, m\} \rangle \subseteq k\mathbb{Z} \Rightarrow \\ &\Rightarrow d \in k\mathbb{Z} \Rightarrow \\ &\Rightarrow k \mid d. \end{aligned}$$

■

4.6.2 El máximo común divisor

Definición 62 . *El entero $d \geq 0$, tal que*

$$d\mathbb{Z} = n\mathbb{Z} + m\mathbb{Z} \tag{4.8}$$

se llama el máximo común divisor de n y m .

La razón de esta nomenclatura es el teorema 53

Notación 8 . *Denotaremos por $(n; m)$ al máximo común divisor de n y m .*

Notemos que $(n; m) > 0$ si $n \neq 0$ ó $m \neq 0$ y que $(0; 0) = 0$.

Observación 57 . $\forall n, m \in \mathbb{Z}, \exists \alpha, \beta \in \mathbb{Z}$ tales que $(n; m) = \alpha n + \beta m$.

Basta ver la ecuación 4.8, para convencerse.

Definición 63 . *Se dice que n, m son primos relativos si $(n; m) = 1$.*

Observación 58 . *n, m son primos relativos si y sólo si $\exists \alpha, \beta \in \mathbb{Z}$ tales que $\alpha n + \beta m = 1$.*

Demostración. \Rightarrow) Se sigue de la observación previa.

\Leftarrow) Si $\alpha n + \beta m = 1$, entonces $\forall z \in \mathbb{Z}, z = \alpha nz + \beta mz \in n\mathbb{Z} + m\mathbb{Z}$.

Por lo tanto

$$1\mathbb{Z} = \mathbb{Z} = n\mathbb{Z} + m\mathbb{Z}$$

■

Teorema 54 . Sean $a, b, c \in \mathbb{Z}$. Si $a \mid bc$ y $(a; b) = 1$ entonces $a \mid c$.

Demostración. Escribamos

$$1 = \alpha a + \beta b,$$

entonces

$$c = c\alpha a + cb\beta,$$

como $a \mid bc$, entonces $bc = ax$, para algún $x \in \mathbb{Z}$. Entonces

$$c = a\alpha a + cb\beta = a\alpha a + ax\beta = a(c\alpha + x\beta).$$

■

Teorema 55 . Sean $a, b, c \in \mathbb{Z}$. Si

1. $a \mid c$, $b \mid c$ y
2. $(a; b) = 1$

entonces $ab \mid c$.

Demostración. Escribamos

$$1 = \alpha a + \beta b$$

entonces

$$c = \alpha ac + \beta bc.$$

Podemos escribir $c = ac'$, $c = bc''$, sustituyendo, obtenemos

$$c = \alpha abc'' + \beta bac' = ab(\alpha c'' + \beta c').$$

■

Ejercicio 168 . Encuentre:

1. $(121; -33)$.
2. $(78696; 19332)$.

Ejercicio 169 . Muestre que dos enteros consecutivos son primos relativos.

Ejercicio 170 . Muestre que si $(a; b) = 1$, entonces $(a^n; b^m) = 1$, $\forall n, m \in \mathbb{N}$.

4.6.3 El mínimo común múltiplo

Como hemos notado antes, la intersección de subconjuntos de \mathbb{Z} cerrados bajo la resta es también cerrada bajo la resta. En particular, si $a, b \in \mathbb{Z}$, tenemos que $a \mathbb{Z} \cap b \mathbb{Z} = m \mathbb{Z}$, para algún entero m que se puede tomar ≥ 0 . Si a y b son cero, entonces $m = 0$. Si alguna de a y b no es cero, entonces $m > 0$. Así que m tiene las siguientes propiedades:

Observación 59 . *Si $a \neq 0$ o $b \neq 0$, y $a \mathbb{Z} \cap b \mathbb{Z} = m \mathbb{Z}$, $m \geq 0$, entonces*

- $m > 0$.
- m es un múltiplo común de a y de b : $a \mid m$ y $b \mid m$ (es decir, $m \in a \mathbb{Z}$ y $m \in b \mathbb{Z}$).
- m es un mínimo múltiplo común de a y de b , es decir que además de que m es un múltiplo común de a y de b :

$$[a \mid n \text{ y } b \mid n] \Rightarrow m \mid n.$$

Demostración. El primer inciso es claro. Para el segundo, nótese nada más que

$$m \in m \mathbb{Z} = a \mathbb{Z} \cap b \mathbb{Z}.$$

Para el tercero:

$$(a \mid n \text{ y } b \mid n)$$

equivale a

$$n \in a \mathbb{Z} \cap b \mathbb{Z} = m \mathbb{Z}.$$

Por lo tanto $m \mid n$. ■

Denotando con $[a; b]$ al mínimo común múltiplo de a y b , podemos escribir el resultado anterior como $a \mathbb{Z} \cap b \mathbb{Z} = [a; b] \mathbb{Z}$.

Lema 9 . *La relación “ \mid ” es una relación de orden en $\mathbb{Z}^+ \cup \{0\}$.*

Demostración. Tenemos que ver que “ \mid ” es reflexiva, antisimétrica y transitiva.

Reflexividad) $a \in a \mathbb{Z}$, $\forall a \in \mathbb{Z}^+ \cup \{0\}$.

Transitividad)

$$\begin{aligned}
 (a \mid b) \wedge (b \mid c) &\Rightarrow \\
 \Rightarrow (c \in b \mathbb{Z}) \wedge (b \in a \mathbb{Z}) &\Rightarrow \\
 \Rightarrow c \in b \mathbb{Z} \subseteq a \mathbb{Z} &\Rightarrow \\
 \Rightarrow c \in a \mathbb{Z} &\Rightarrow \\
 \Rightarrow a \mid c.
 \end{aligned}$$

Antisimetría) Si $a, b \in \mathbb{Z}^+ \cup \{0\}$, $a \mid b$ y $b \mid a$, entonces $a = b$:

$$\begin{aligned}
 a &= bz_1, b = az_2 \Rightarrow \\
 \Rightarrow a &= az_2z_1 \Rightarrow \\
 \Rightarrow 1 &= z_2z_1 \Rightarrow \\
 \Rightarrow z_1 &\in \{1, -1\}.
 \end{aligned}$$

Por lo tanto $a = b$ ó $a = -b$. Pero $a, b \in \mathbb{Z}^+ \cup \{0\} \Rightarrow a = b$. ■

Definición 64 . $a \mathbb{Z} b \mathbb{Z} := \{az_1bz_2 \mid z_1, z_2 \in \mathbb{Z}\}$.

Lema 10 . $a \mathbb{Z} b \mathbb{Z} = ab \mathbb{Z}$.

Demostración. Notemos que $a \mathbb{Z} b \mathbb{Z}$ contiene a ab y es cerrado bajo la resta, por lo que es un ideal y además $ab \mathbb{Z} \subseteq a \mathbb{Z} b \mathbb{Z}$.

Recíprocamente, $az_1bz_2 = ab(z_1z_2) \in ab \mathbb{Z}$. ■

En vista de lo anterior, el producto de dos ideales es un ideal.

Lema 11 . $a \mathbb{Z} (b \mathbb{Z} + c \mathbb{Z}) = a \mathbb{Z} b \mathbb{Z} + a \mathbb{Z} c \mathbb{Z}$.

Demostración. \subseteq)⁴

$$\begin{aligned}
 a \mathbb{Z} (b \mathbb{Z} + c \mathbb{Z}) &= a \mathbb{Z} ((b; c) \mathbb{Z}) = \\
 &= a(b; c) \mathbb{Z} = a(\alpha b + \beta c) \mathbb{Z} = \\
 &= (a\alpha b + a\beta c) \mathbb{Z} \subseteq a\alpha b \mathbb{Z} + a\beta c \mathbb{Z} = \\
 &= a\alpha \mathbb{Z} b \mathbb{Z} + a\beta \mathbb{Z} c \mathbb{Z} \subseteq a \mathbb{Z} b \mathbb{Z} + a \mathbb{Z} c \mathbb{Z}.
 \end{aligned}$$

\supseteq)

$$a \mathbb{Z} b \mathbb{Z} \subseteq a \mathbb{Z} (b \mathbb{Z} + c \mathbb{Z})$$

⁴Recuérdese que $(b; c) \in b \mathbb{Z} + c \mathbb{Z}$.

puesto que

$$b \mathbb{Z} \subseteq (b \mathbb{Z} + c \mathbb{Z}).$$

También

$$a \mathbb{Z} c \mathbb{Z} \subseteq a \mathbb{Z} (b \mathbb{Z} + c \mathbb{Z}),$$

pues

$$c \mathbb{Z} \subseteq (b \mathbb{Z} + c \mathbb{Z}).$$

es cerrado bajo la suma, así que

$$a \mathbb{Z} b \mathbb{Z} + a \mathbb{Z} c \mathbb{Z} \subseteq a \mathbb{Z} (b \mathbb{Z} + c \mathbb{Z}).$$

■

Teorema 56 . *Si $a > 0$, b , c son enteros, entonces $(ab; ac) = a(b; c)$.*

Demostración.

$$\begin{aligned} a(b; c) \mathbb{Z} &= a \mathbb{Z} (b; c) \mathbb{Z} = \\ &= a \mathbb{Z} ((b; c) \mathbb{Z}) = a \mathbb{Z} (b \mathbb{Z} + c \mathbb{Z}) = \\ &= a \mathbb{Z} b \mathbb{Z} + a \mathbb{Z} c \mathbb{Z} = ab \mathbb{Z} + ac \mathbb{Z} = \\ &= (ab; ac) \mathbb{Z}. \end{aligned}$$

Por lo tanto

$$a(b; c) \mid (ab; ac)$$

y recíprocamente.

Como $a(b; c)$ y $(ab; ac)$ son no negativos, entonces

$$a(b; c) = (ab; ac).$$

■

Teorema 57 . $a \mathbb{Z} (b \mathbb{Z} \cap c \mathbb{Z}) = a \mathbb{Z} b \mathbb{Z} \cap a \mathbb{Z} c \mathbb{Z}$.

Demostración. Es claro que 0 pertenece a los dos conjuntos de la ecuación.

⊆) Supongamos que para $z_2, z_3 \in \mathbb{Z}$, se tiene que $bz_2 = cz_3$ y supongamos que $az_1 \in a \mathbb{Z}$, entonces

$$az_1 bz_2 = az_1 cz_3 \in a \mathbb{Z} b \mathbb{Z} \cap a \mathbb{Z} c \mathbb{Z}.$$

⊇) Tomemos un elemento distinto de 0 en $a \mathbb{Z} b \mathbb{Z} \cap a \mathbb{Z} c \mathbb{Z}$. Digamos que tomamos

$$az_1bz_2 = az_3cz_4 \neq 0,$$

entonces $a \neq 0$, así que cancelando, obtenemos

$$z_1bz_2 = z_3cz_4 \in (b \mathbb{Z} \cap c \mathbb{Z}),$$

así que

$$az_1bz_2 = az_3cz_4 \in a(b \mathbb{Z} \cap c \mathbb{Z}) \subseteq a \mathbb{Z} (b \mathbb{Z} \cap c \mathbb{Z})$$

■

Lema 12 . *Si $a > 0$, b , c son enteros, entonces $[ab; ac] = a[b; c]$.*

Demostración. $[ab; ac] \mathbb{Z} = ab \mathbb{Z} \cap ac \mathbb{Z} = a \mathbb{Z} (b \mathbb{Z} \cap c \mathbb{Z}) = a \mathbb{Z} ([b; c] \mathbb{Z}) = (a[b; c]) \mathbb{Z}$. ■

Observación 60

$$x | y \Leftrightarrow -x | y \Leftrightarrow x | -y \Leftrightarrow -x | -y.$$

Demostración. $xz = y \Leftrightarrow (-x)(-z) = y \Leftrightarrow (x)(-z) = -y \Leftrightarrow (-x)(z) = -y$. ■

Como consecuencia de lo anterior, tenemos que el conjunto de divisores (y de múltiplos) de y coincide con el de $-y$. En particular, podemos hacer la siguiente observación.

Observación 61

$$(x; y) = (-x; y) = (x; -y) = (-x; -y),$$

y también

$$[x; y] = [-x; y] = [x; -y] = [-x; -y].$$

Notemos que si $b \neq 0$ y $a | b$ entonces $ac = b$, para alguna única $c \in \mathbb{Z}$. Pues $ac = ad = b \neq 0 \implies a \neq 0$ y así $c = d$.

Por esta razón podemos denotar $c = \frac{b}{a}$ al único entero x con la propiedad de que $ax = b$ (Cuando $a | b$ y $b \neq 0$).

Teorema 58 . $(a; b) [a; b] = ab$.

Demostración. Si $a = 0$ ó $b = 0$, entonces $[a; b] = 0 = ab$.

Supongamos entonces que $a \neq 0$ y $b \neq 0$.

Por la observación 61 podemos suponer que $a > 0$ y $b > 0$.

Como

$$ab \mid (a; b) [a; b] \iff \frac{ab}{[a; b]} \mid (a; b) \iff \frac{ab}{[a; b]} \mid a \text{ y } \frac{ab}{[a; b]} \mid b \iff (ab \mid a [a; b]$$

$$\text{y } ab \mid b [a; b]) \iff (b \mid [a; b] \text{ y } a \mid [a; b]).$$

Por lo tanto tenemos que $ab \mid (a; b) [a; b]$ o lo que es lo mismo: $(a; b) [a; b] \mathbb{Z} \subseteq ab \mathbb{Z}$.

Por otra parte, como $(a; b) \mid a$, digamos que $a = (a; b) a_1$ (y que $b = (a; b) b_1$).

Entonces

$$ab = (a; b) a_1 (a; b) b_1,$$

así que podemos denotar

$$\frac{ab}{(a; b)} = a_1 (a; b) b_1 = ab_1 = a_1 b.$$

Como se ve, $\frac{ab}{(a; b)}$ es un múltiplo común de a y de b , por lo tanto también

es múltiplo de $[a; b]$. Así que $\exists w \in \mathbb{Z}$, $w > 0$ tal que $\frac{ab}{(a; b)} = w [a; b]$.

Multiplicando por $(a; b)$, tenemos que

$$ab = w [a; b] (a; b) \in wab \mathbb{Z}$$

(pues como vimos arriba, $(a; b) [a; b] \in ab \mathbb{Z}$). Así,

$$ab = wabz \Rightarrow wz = 1 \Rightarrow w = 1 \Rightarrow ab = [a; b] (a; b).$$

■

Observación 62

$$(a \mathbb{Z} + b \mathbb{Z}) + c \mathbb{Z} = a \mathbb{Z} + (b \mathbb{Z} + c \mathbb{Z}) = (a \mathbb{Z} + c \mathbb{Z}) + b \mathbb{Z}.$$

Por lo tanto

$$((a; b); c) = (a; (b; c)) = ((a; c); b).$$

Así que denotaremos a este número $(a; b; c)$.

Ejercicio 171 . *Demostrar que $(a; b; c)$ es el mayor divisor común de a, b, c .*

Ejercicio 172 . *Tomando $a = 6, b = 10, c = 30$, compruebe que $((a; b); c) = (a; (b; c)) = ((a; c); b)$.*

Observación 63

$$(a \mathbb{Z} \cap b \mathbb{Z}) \cap c \mathbb{Z} = a \mathbb{Z} \cap (b \mathbb{Z} \cap c \mathbb{Z}) = (a \mathbb{Z} \cap c \mathbb{Z}) \cap b \mathbb{Z}.$$

Por lo tanto

$$[[a; b]; c] = [a; [b; c]] = [[a; b]; c].$$

Así que denotaremos a este número

$$[a; b; c].$$

Teorema 59 (Propiedad modular). *Si $b \mathbb{Z} \subseteq a \mathbb{Z}$, entonces*

$$a \mathbb{Z} \cap (b \mathbb{Z} + c \mathbb{Z}) = b \mathbb{Z} + (a \mathbb{Z} \cap c \mathbb{Z}).$$

Demostración. \supseteq)

$$b \mathbb{Z} \subseteq a \mathbb{Z} \text{ y } b \mathbb{Z} \subseteq b \mathbb{Z} + c \mathbb{Z}.$$

Por lo tanto

$$b \mathbb{Z} \subseteq a \mathbb{Z} \cap (b \mathbb{Z} + c \mathbb{Z}).$$

También

$$a \mathbb{Z} \cap c \mathbb{Z} \subseteq a \mathbb{Z} \text{ y } a \mathbb{Z} \cap c \mathbb{Z} \subseteq (b \mathbb{Z} + c \mathbb{Z}),$$

así que

$$a \mathbb{Z} \cap c \mathbb{Z} \subseteq a \mathbb{Z} \cap (b \mathbb{Z} + c \mathbb{Z}).$$

Entonces

$$b \mathbb{Z} + (a \mathbb{Z} \cap c \mathbb{Z}) \subseteq a \mathbb{Z} \cap (b \mathbb{Z} + c \mathbb{Z}).$$

\subseteq)

Si $az_1 = bz_2 + cz_3$, entonces

$$cz_3 = az_1 - bz_2 \in a \mathbb{Z} + b \mathbb{Z} = a \mathbb{Z}.$$

Por lo tanto $cz_3 \in a \mathbb{Z} \cap c \mathbb{Z}$, y entonces

$$az_1 = bz_2 + cz_3 \in b \mathbb{Z} + (a \mathbb{Z} \cap c \mathbb{Z}).$$

■

Además, como caso particular tenemos que

$$(a\mathbb{Z} + b\mathbb{Z}) \cap (b\mathbb{Z} + c\mathbb{Z}) = b\mathbb{Z} + ((a\mathbb{Z} + b\mathbb{Z}) \cap c\mathbb{Z}).$$

Esto se sigue de que $a\mathbb{Z} + b\mathbb{Z} = (a; b)\mathbb{Z}$ y de que $b\mathbb{Z} \subseteq (a; b)\mathbb{Z}$.

Teorema 60 . $(a; [b; c]) = [(a; b); (a; c)]$ (*Calcular un máximo común divisor se distribuye sobre un mínimo común múltiplo*).

Demostración. $(a; [b; c])\mathbb{Z} = a\mathbb{Z} + [b; c]\mathbb{Z} = a\mathbb{Z} + (b\mathbb{Z} \cap c\mathbb{Z})$.

Si pudiéramos demostrar que

$$a\mathbb{Z} + (b\mathbb{Z} \cap c\mathbb{Z}) = (a\mathbb{Z} + b\mathbb{Z}) \cap (a\mathbb{Z} + c\mathbb{Z}),$$

habríamos terminado porque

$$(a\mathbb{Z} + b\mathbb{Z}) \cap (a\mathbb{Z} + c\mathbb{Z}) = (a; b)\mathbb{Z} \cap (a; c)\mathbb{Z} = [(a; b); (a; c)]\mathbb{Z}.$$

Veamos entonces que

$$a\mathbb{Z} + (b\mathbb{Z} \cap c\mathbb{Z}) = (a\mathbb{Z} + b\mathbb{Z}) \cap (a\mathbb{Z} + c\mathbb{Z}):$$

\subseteq)

Supongamos que $bz_2 = cz_3$, entonces

$$az_1 + bz_2 = az_1 + cz_3 \in (a\mathbb{Z} + b\mathbb{Z}) \cap (a\mathbb{Z} + c\mathbb{Z}).$$

\supseteq)

$$\begin{aligned} (a\mathbb{Z} + b\mathbb{Z}) \cap (a\mathbb{Z} + c\mathbb{Z}) &= \\ &= (a; b)\mathbb{Z} \cap (a\mathbb{Z} + c\mathbb{Z}) = \\ &= a\mathbb{Z} + ((a; b)\mathbb{Z} \cap c\mathbb{Z}) = \\ &= a\mathbb{Z} + [(a; b); c]\mathbb{Z}. \end{aligned}$$

⁵Basta demostrar que

$$[(a; b); c] \in a\mathbb{Z} + (b\mathbb{Z} \cap c\mathbb{Z}) = a\mathbb{Z} + ([b; c]\mathbb{Z}).$$

⁵Se usó la observación previa al enunciado de este Teorema.

Recuérdese que el máximo común divisor de dos número es una combinación entera de ellos.

Podemos escribir que

$$[(a; b); c] = xa \frac{c}{((a; b); c)} + by \frac{c}{((a; b); c)}.$$

Observemos ahora que $\frac{bc}{((a; b); c)}$ es un múltiplo común de b y de c , y entonces también lo es de $[b; c]$. Entonces

$$by \frac{c}{((a; b); c)} = [b; c] w,$$

para algún entero w . Por lo tanto

$$[(a; b); c] = a \left(x \frac{c}{((a; b); c)} \right) + [b; c] w.$$

■

Ahora es fácil demostrar el siguiente teorema.

Teorema 61 . $[a; (b; c)] = ([a; b]; [a; c])$ (*Calcular un mínimo común múltiplo se distribuye sobre un máximo común divisor*).

Demostración. Es claro que $a \mid [a; b]$ y $a \mid [a; c]$.

Por lo tanto

$$a \mid ([a; b]; [a; c]).$$

También

$$(b; c) \mid [a; b] \text{ y } (b; c) \mid [a; c]$$

por lo que

$$(b; c) \mid ([a; b]; [a; c]).$$

Como $([a; b]; [a; c])$ es un múltiplo común de a y de $(b; c)$, entonces

$$[a; (b; c)] \mid ([a; b]; [a; c]).$$

Para demostrar el recíproco, basta ver que

$$\begin{aligned} [a; (b; c)] &\in ([a; b]; [a; c]) \mathbb{Z} = \\ &= [a; b] \mathbb{Z} + [a; c] \mathbb{Z} = \\ &= (a \mathbb{Z} \cap b \mathbb{Z}) + (a \mathbb{Z} \cap c \mathbb{Z}). \end{aligned}$$

Pero

$$[a; (b; c)] (a; b; c) = a (b; c) = a(bx + cy) = abx + acy.$$

Entonces

$$[a; (b; c)] = \frac{abx}{(a; b; c)} + \frac{acy}{(a; b; c)}.$$

Es claro que

$$\frac{abx}{(a; b; c)} = a \left(\frac{bx}{(a; b; c)} \right) = b \left(\frac{ax}{(a; b; c)} \right) \in a \mathbb{Z} \cap b \mathbb{Z},$$

mientras que

$$\frac{acy}{(a; b; c)} \in a \mathbb{Z} \cap c \mathbb{Z}.$$

■

4.7 El Teorema fundamental de la Aritmética

Definición 65 . Decimos que $p \in \mathbb{Z}$ es primo si tiene exactamente cuatro divisores.

Observación 64 . 1 (-1) sólo tiene los divisores 1 y -1 :

Demostración. $1 = qx \Rightarrow 1 = |q| |x|$, con $|q|, |x| \geq 1$.

Si $|q| > 1$, entonces $1 = |q| |x| > |x| \geq 1$. ■

Por lo tanto $|q| = 1$. ■

Observación 65 . En vista de lo anterior, 1 no es primo, ya que sólo tiene dos divisores: $1, -1$.

Observación 66 . Notemos que si p es un primo entonces sus cuatro divisores tienen que ser necesariamente $1, -1, p$ y $-p$:

Demostración. Si p fuera primo, y $1, -1, p$ y $-p$ no fueran cuatro divisores de p , sería porque $p = 1$ ó $p = -1$, ninguno de los cuales es primo.

Notación 9 . Denotemos por \wp el conjunto de enteros primos.

Notemos que si $p \in \wp$, entonces $p \neq 1$ como vimos en el argumento de arriba.

Nótese que 0 tampoco es primo pues 0 tiene una infinidad de divisores

$$(z \mid 0, \forall z \in \mathbb{Z}).$$

En cambio, 2 es primo (ejercicio), 3 es primo, 5 es primo.

En este momento se puede preguntar uno lo siguiente: ¿Cuántos primos hay? ¿Un número infinito o finito?

Para responder esta pregunta, haremos algunas observaciones.

- Si p es un número primo y $z \in \mathbb{Z}$, entonces $(p; z) \in \{1, p\}$.

Esto se sigue de que $(p; z)$ es uno de los cuatro divisores de p y de que $(p; z) > 0$.

- Si p es un número primo y $p \nmid z \in \mathbb{Z}$, entonces $(p; z) = 1$.

Es claro, por el inciso anterior.

Observación 67 . Si $p \in \wp$ y $p \mid ab$, $a, b \in \mathbb{Z}$, entonces $p \mid a$ ó $p \mid b$.

Demostración. Supongamos que $p \in \wp$ y $p \mid ab$. Si $p \nmid a$, entonces por la observación anterior, $(p; a) = 1$. Por lo tanto (teorema 54) $p \mid b$. ■

Observación 68 . Si $p = ab$, $ab \in \mathbb{Z}$, y $a > 0$, entonces $a = 1$ ó $a = p$.

Demostración. Como $a \mid p$, entonces $(p; a) = a$.⁶ Pero un divisor positivo de p es 1 ó p . Por lo tanto $a = 1$ ó $a = p$. ■

Teorema 62 (Fundamental de la Aritmética) . Si $a \in \mathbb{Z}^+$, $a > 1$, entonces $\exists! k \in \mathbb{N}$, $\exists! p_1, p_2, \dots, p_k \in \wp$. Tales que $a = p_1 \cdot p_2 \cdot \dots \cdot p_k$.

Demostración. Existencia.

Por inducción sobre a .

Base.

Si $a = 2$, entonces $k = 1$ y

$$\{p_1\} = \{2\}.$$

⁶Como $a \mid p \Leftrightarrow p \mathbb{Z} \subseteq a \mathbb{Z}$, tenemos que $p \mathbb{Z} + a \mathbb{Z} \subseteq a \mathbb{Z} + a \mathbb{Z} = a \mathbb{Z}$.

Paso inductivo.

Supongamos ahora que $a > 2$, y suponemos que la afirmación es cierta para todos los enteros mayores que 1 menores que a .

Si $a \in \wp$, entonces $k = 1$ y

$$\{p_1\} = \{a\}.$$

Si a no es un primo, entonces, como es mayor que 1, entonces tiene los divisores $1, -1, a, -a$ y algún otro, b digamos, que podemos suponer positivo ($-b$ también es divisor de a) y que es distinto de 1 y de a .

Así pues, $a = bc$, con $c > 0$, dado que a y b son positivos.

Como b no es a , entonces c no es 1. Por lo tanto $c \geq 2$. Entonces $b < a$ ($[b \geq a \text{ y } c \geq 2] \Rightarrow a = bc \geq 2a$ ∇) y también $c < a$.

Podemos aplicar la hipótesis de inducción a b y c para concluir que ambos se factorizan como producto de primos. Así que a también se factoriza como producto de primos.

Unicidad) Supongamos que hay un entero positivo que se factoriza como producto de primos de dos maneras distintas. Entonces el conjunto

$$\{a \in \mathbb{Z} \mid a > 1, a \text{ tiene dos factorizaciones en primos}\} \neq \emptyset,$$

por lo tanto (principio del buen orden) tendría un elemento menor, m digamos. Entonces $m = p_1 \cdot p_2 \cdot \dots \cdot p_k = q_1 \cdot q_2 \cdot \dots \cdot q_l$, con $k \neq l$ o con $p_1, p_2, \dots, p_k \neq q_1, q_2, \dots, q_l$, donde suponemos que las p_i y las q_j se escriben en orden no decreciente.

Como

$$p_1 \mid q_1 \cdot q_2 \cdot \dots \cdot q_l,$$

entonces $p_1 \mid q_1$ ó $(p_1, q_1) = 1$.

Ahora,

$$\begin{cases} p_1 \mid q_1 \Rightarrow p_1 = q_1 \text{ y} \\ (p_1, q_1) = 1 \Rightarrow p_1 \mid q_2 \cdot \dots \cdot q_l. \end{cases}$$

⁷Así que

$$\begin{cases} p_1 = q_1 \text{ o} \\ p_1 \mid q_2 \cdot \dots \cdot q_l. \end{cases}$$

⁷Un divisor de un primo p es 1 o es el primo p . Por lo tanto p_1 es 1 o q_1 . Pero $p_1 \neq 1$, pues p_1 es primo.

Repetiendo el argumento, tenemos que

$$\left\{ \begin{array}{l} p_1 = q_1 \text{ o} \\ p_1 = q_2 \text{ o} \\ p_1 \mid q_3 \cdot \dots \cdot q_l. \end{array} \right.$$

Al final podemos escribir que

$$p_1 = q_1 \text{ o } p_1 = q_2 \text{ o } p_1 = q_3 \text{ o } \dots \text{ o } p_1 = q_l.$$

Así que p_1 es uno de los factores primos en $q_1 \cdot q_2 \cdot \dots \cdot q_l$, digamos que $p_1 = q_j$. Por simetría, $q_1 = p_r$. Ahora,

$$p_r \leq q_1 \leq q_j = p_1 \leq p_r.$$

Así que $p_1 = q_1$ y cancelando en las factorizaciones de m , tenemos que

$$p_2 \cdot \dots \cdot p_k = q_2 \cdot \dots \cdot q_l$$

son dos factorizaciones de $\frac{m}{p_1} < m$. Por la elección de m , tenemos que

$$k = l$$

y que

$$p_2 = q_2, p_3 = q_3, \dots, p_k = q_k.$$

Además

$$p_1 = q_1.$$

Contradicciendo que teníamos dos factorizaciones distintas de m . ■

Ejercicio 173 . Encuentre la factorización en primos de

1. 100.
2. 130.
3. 1960.
4. 109.
5. 713.

Ejercicio 174 . Encuentre los conjuntos de divisores positivos para cada número en el ejercicio anterior.

Ejercicio 175 . Hallar el menor múltiplo positivo de 945 que sea un cuadrado.

Ejercicio 176 . Hallar el número de divisores de 2160 y calcular su suma.

4.7.1 El conjunto de primos es infinito

Teorema 63 . *El conjunto de primos es infinito.*

Demostración. Sean p_1, p_2, \dots, p_k , los primeros k primos y consideremos el número

$$p_1 \cdot p_2 \cdot \dots \cdot p_k + 1,$$

notemos que este número no es divisible por ninguno de los primos

$$p_1, p_2, \dots, p_k,$$

lo que muestra que

$$p_i \nmid p_1 \cdot p_2 \cdot \dots \cdot p_k + 1.$$

⁸ Por el teorema fundamental de la Aritmética,

$$p_1 \cdot p_2 \cdot \dots \cdot p_k + 1$$

tiene un divisor primo diferente de los de la lista

$$p_1, p_2, \dots, p_k.$$

Esto muestra que hay más de k primos (cualquiera que sea k). Por lo tanto, el número de primos es infinito. ■

Ejercicio 177 . *Muestre que si n es un entero mayor que 1 no es un primo, entonces existe un primo positivo p tal que $p^2 \leq n$.*

Ejercicio 178 . *Determine cuáles de los siguientes números son primos:*

1. 503.
2. 943.
3. 1511.

⁸Note que $p_1 \left| \frac{p_2 * \dots * p_k}{p_1 * p_2 * \dots * p_k + 1} \right.$

y $p_1 \left| \frac{q}{p_1 * p_2 * \dots * p_k + 1} \right.$ no pueden ocurrir simultáneamente, por la unicidad en el

Teorema 51, en la página 208 (Algoritmo de la división).

4. $2^{13} - 1$.

5. 899.

Ejercicio 179 . Demuestre que el único conjunto de 3 impares positivos consecutivos primos es $\{3, 5, 7\}$.

1. Demostrar que si $2^n - 1$ es primo, con $n \in \mathbb{N}$ entonces n es impar o n es 2.
2. Demostrar que si $2^n - 1$ es primo, con $n \in \mathbb{N}$ entonces n es un primo.

Ejercicio 180 . Demuestre que si n es un número natural y $2^n + 1$ es un primo, entonces n tiene que ser una potencia de 2.

Ejercicio 181 . Demostrar que si $n \geq 2$, existe p primo tal que $n < p < n!$.

Ejercicio 182 . Demostrar que ningún primo de la forma $4n + 3$ es suma de dos cuadrados. (Sugerencia: escriba los cuadrados de \mathbb{Z}_4)

Ejercicio 183 . $2^{2^n} + 1 =: F_n$ se llama el n -ésimo número de Fermat.

1. Demuestre que $(F_n; F_m) = 1$ si $n \neq m$.
2. Use lo anterior para deducir que hay una infinidad de primos. (Recuerde el Teorema fundamental de la Aritmética).

4.8 El algoritmo de Euclides

Veamos ahora como calcular el máximo común divisor de dos enteros. Por las observaciones anteriores, podemos suponer que ambos enteros son positivos, a y b , digamos.

Lema 13 . Si $a = qb + r$, entonces $(a; b) = (b; r)$.

Demostración.

$$\begin{aligned}
 (a; b) \mathbb{Z} &= a \mathbb{Z} + b \mathbb{Z} = \\
 &= (qb + r) \mathbb{Z} + b \mathbb{Z} \subseteq b \mathbb{Z} + r \mathbb{Z} + b \mathbb{Z} = b \mathbb{Z} + r \mathbb{Z} = \\
 &= b \mathbb{Z} + (a - qb) \mathbb{Z} \subseteq a \mathbb{Z} + b \mathbb{Z} = (a; b) \mathbb{Z}.
 \end{aligned}$$

Por lo tanto

$$b \mathbb{Z} + r \mathbb{Z} = (a; b) \mathbb{Z}.$$

Por lo que

$$(b; r) \mathbb{Z} = b \mathbb{Z} + r \mathbb{Z} = (a; b) \mathbb{Z}.$$

■

Teorema 64 (*Algoritmo de Euclides*). *Sean $a, b > 0$. Considere la situación,*

$$b \overline{q_1} \atop r_1 \quad 0 \leq r_1 < b,$$

entonces $r_1 = a - bq_1$.

Teorema 65 . *Si $r_1 > 0$, hágase la división*

$$r_1 \overline{q_2} \atop r_2 \quad 0 \leq r_2 < r_1$$

si $r_2 > 0$, hágase la división

$$r_2 \overline{q_3} \atop r_3 \quad 0 \leq r_3 < r_2,$$

continúese de esta manera para obtener una sucesión,

$$b > r_1 > r_2 > r_3 > \dots \geq 0.$$

Entonces $(a; b)$ es el último residuo distinto de cero en la sucesión anterior.

Demostración. Por definición,

$$r_{i+1} \overline{q_{i+2}} \atop r_{i+2} \quad 0 \leq r_{i+2} < r_{i+1}.$$

Así que

$$(r_i; r_{i+1}) = (r_{i+1}; r_{i+2}).$$

Por lo tanto

$$(a; b) = (b; r_1) = (r_1; r_2) = \dots = (r_i; r_{i+1}) = (r_{i+1}; r_{i+2}) = \dots = (r_k; 0).$$

Aquí hay que notar que la sucesión

$$b > r_1 > r_2 > r_3 > \dots$$

termina (por el principio del buen orden) y termina en 0.

Ahora sólo resta notar que si r_k es el último residuo $\neq 0$, entonces

$$(r_k; 0) = r_k.$$

■

Definición 66 . $a \stackrel{n}{\equiv} b$ si $a - b \in n\mathbb{Z}$.

(Se lee: a es congruente con b módulo n).

Teorema 66 . $\stackrel{n}{\equiv}$ es una relación de equivalencia en \mathbb{Z} .

Demostración. Reflexividad)

$a \stackrel{n}{\equiv} a$, pues $a - a = 0 \in n\mathbb{Z}$.

Simetría)

$$a \stackrel{n}{\equiv} b \Rightarrow a - b \in n\mathbb{Z} \Rightarrow -(a - b) \in n\mathbb{Z} \Rightarrow b - a \in n\mathbb{Z} \Rightarrow b \stackrel{n}{\equiv} a.$$

Transitividad)

$$\begin{aligned} a \stackrel{n}{\equiv} b \stackrel{n}{\equiv} c &\Rightarrow \\ &\Rightarrow (a - b \in n\mathbb{Z} \text{ y } b - c \in n\mathbb{Z}) \Rightarrow \\ &\Rightarrow ((a - b) + (b - c) \in n\mathbb{Z}) \Rightarrow \\ &\Rightarrow (a - c \in n\mathbb{Z}) \\ &\Rightarrow a \stackrel{n}{\equiv} c. \end{aligned}$$

■

Observación 69 . Si denotamos $[a]_n$, la clase de equivalencia de a , entonces

$$\begin{aligned}[a]_n &= \left\{ b \in \mathbb{Z} \mid b \stackrel{n}{\equiv} a \right\} = \{b \in \mathbb{Z} \mid b - a \in n\mathbb{Z}\} = \\ &= \{b \in \mathbb{Z} \mid b - a = nz \text{ p. a. } z \in \mathbb{Z}\} = \{a + nz \mid z \in \mathbb{Z}\} = \\ &\stackrel{\text{def}}{=} a + n\mathbb{Z}.\end{aligned}$$

Observación 70 . En la situación

$$n \overline{\overline{a}}^q_r, \quad 0 \leq r < n,$$

tenemos que $a = nq + r$, por lo que $a - r \in n\mathbb{Z}$. Así que $a \stackrel{n}{\equiv} r$.

Es decir, un entero es congruente módulo n con el residuo que deja al ser dividido entre n .

4.9 El anillo de los enteros módulo n

Como consecuencia de la observación anterior, tenemos para el conjunto de clases de congruencia:

$$\mathbb{Z} \not{\equiv}^n = \{[0]_n, [1]_n, \dots, [n-1]_n\}.$$

Como la notación $\mathbb{Z} \not{\equiv}^n$ es algo aparatosa, escribiremos simplemente \mathbb{Z}_n . (Note que n se usa aquí como subíndice, no confundir con \mathbb{Z}^n , el conjunto de múltiplos de n).

Enseguida dotamos a \mathbb{Z}_n de suma y de producto.

Definición 67

1. $\mathbb{Z}_n \times \mathbb{Z}_n \xrightarrow{+} \mathbb{Z}_n$ se define por $[a]_n + [b]_n \stackrel{\text{def}}{=} [a + b]_n$
2. $\mathbb{Z}_n \times \mathbb{Z}_n \xrightarrow{\cdot} \mathbb{Z}_n$ se define por $[a]_n \cdot [b]_n \stackrel{\text{def}}{=} [a \cdot b]_n$.

Veamos que las definiciones anteriores son buenas, en el sentido de que no dependen de los representantes escogidos en las clases de congruencia.

Demostración. Buena definición de $+$:

$$\begin{aligned}
 a \stackrel{n}{\equiv} \alpha, b \stackrel{n}{\equiv} \beta &\Rightarrow \\
 \Rightarrow a - \alpha \in n\mathbb{Z}, b - \beta \in n\mathbb{Z} &\Rightarrow \\
 \Rightarrow a - \alpha + b - \beta \in n\mathbb{Z} &\Rightarrow \\
 \Rightarrow a + b - (\alpha + \beta) \in n\mathbb{Z} &\Rightarrow \\
 \Rightarrow a + b \stackrel{n}{\equiv} \alpha + \beta &\Rightarrow \\
 \Rightarrow [a + b]_n = [\alpha + \beta]_n. &
 \end{aligned}$$

Buena definición de \cdot :

$$\begin{aligned}
 a \stackrel{n}{\equiv} \alpha, b \stackrel{n}{\equiv} \beta &\Rightarrow \\
 \Rightarrow a - \alpha \in n\mathbb{Z}, b - \beta \in n\mathbb{Z} &\Rightarrow \\
 \Rightarrow b \cdot (a - \alpha) \in n\mathbb{Z} &
 \end{aligned}$$

y

$$\begin{aligned}
 \alpha \cdot (b - \beta) &\in n\mathbb{Z} \Rightarrow \\
 \Rightarrow b \cdot a - b \cdot \alpha + \alpha \cdot b - \alpha \cdot \beta &\in n\mathbb{Z} \Rightarrow \\
 \Rightarrow a \cdot b - \alpha \cdot \beta \in n\mathbb{Z} &\Rightarrow \\
 \Rightarrow a \cdot b \stackrel{n}{\equiv} \alpha \cdot \beta. &
 \end{aligned}$$

■ Es fácil ver que las operaciones que hemos definido son asociativas, conmutativas y con neutros respectivos: $[0]_n$, $[1]_n$.

Cada elemento $[r]_n$ de \mathbb{Z}_n tiene inverso aditivo: $[-r]_n$

Además el producto se distribuye sobre la suma:

$$\begin{aligned}
 [r]_n \cdot ([s]_n + [t]_n) &= [r]_n \cdot ([s + t]_n) = \\
 &= [r \cdot (s + t)]_n = \\
 &= [rs + rt]_n = \\
 &= [rs]_n + [rt]_n = \\
 &= [r]_n [s]_n + [r]_n [t]_n.
 \end{aligned}$$

Así que

$(\mathbb{Z}_n, +, \cdot, [0]_n, [1]_n)$ es un anillo.

Ver nota en la página 197.

Veamos ahora cuáles son los elementos invertibles (bajo el producto) de \mathbb{Z}_n . Estos elementos también se conocen como unidades.

Teorema 67 . $[a]_n$ es una unidad de \mathbb{Z}_n si y sólo si $(a; n) = 1$.

Demostración. \Rightarrow)

Si $[b]_n$ es inverso multiplicativo de $[a]_n$, entonces

$$[1]_n = [a]_n \cdot [b]_n = [ab]_n ,$$

así que $ab \equiv 1$, es decir, $ab - 1 \in n\mathbb{Z}$.

Entonces

$$ab - 1 = nz, z \in \mathbb{Z} .$$

Luego,

$$1 = ab + nz \in a\mathbb{Z} + n\mathbb{Z} \Rightarrow \mathbb{Z} = a\mathbb{Z} + n\mathbb{Z} = (a; n)\mathbb{Z} .$$

Por lo que $(a; n) = 1$.

\Leftarrow)

$$\begin{aligned} (a; n) = 1 \Rightarrow 1 &= ab + nz \Rightarrow \\ \Rightarrow [1]_n &= [ab + nz]_n = \\ &= [ab]_n + [nz]_n = \\ &= [a]_n [b]_n + [n]_n [z]_n = \\ &= [a]_n [b]_n + [0]_n [z]_n = \\ &= [a]_n [b]_n + [0z]_n = \\ &= [a]_n [b] . \end{aligned}$$

Es decir, $[a]_n$ es una unidad. ■

El conjunto de unidades de \mathbb{Z}_n se suele denotar \mathbb{Z}_n^\times .

Teorema 68 . \mathbb{Z}_n es un campo si y sólo si n es un primo.

Demostración. Como \mathbb{Z}_n es un anillo commutativo, \mathbb{Z}_n es un campo si y sólo si $\mathbb{Z}_n^\times = \mathbb{Z}_n \setminus \{[0]_n\}$ (es decir si todo elemento distinto de 0 tiene inverso multiplicativo).

Esto sucede si y sólo si $[a]_n$ es una unidad, para toda $[a] \neq [0]_n$, es decir, si y sólo si $(a; n) = 1 \forall a \in \{1, \dots, n-1\}$. Lo anterior sucede si y sólo si los únicos divisores positivos de n son 1 y n , es decir si y sólo si n es un primo.

■

Ejemplo 102 . \mathbb{Z}_{17} es un campo porque 17 es primo. El inverso multiplicativo de $[12]_{17}$, se puede encontrar expresando 1 como combinación entera de 12 y 17 (usando el Algoritmo de Euclides).

$$\begin{array}{r} 1 \\ 12 \overline{) 17} \\ 5 \\ \hline 2 \end{array} \quad \begin{array}{r} 2 \\ 5 \overline{) 12} \\ 2 \\ \hline 1 \end{array} \quad \begin{array}{r} 2 \\ 2 \overline{) 5} \\ 1 \\ \hline \end{array}$$

entonces

$$\begin{aligned} 1 &= 5 - 2 \cdot 2 = 5 - 2 \cdot (12 - (5 \cdot 2)) = \\ &= 5 \cdot 5 - 2 \cdot 12 = 5 \cdot (17 - 12) - 2 \cdot 12 = \\ &= 17 \cdot 5 - 7 \cdot 12 \end{aligned}$$

tomando residuos módulo 17 tenemos que el inverso de $[12]_{17}$ es

$[-7]_{17} = [10]_{17}$. Note que en efecto $12 \cdot 10 \stackrel{17}{\equiv} 1$:

$$\begin{array}{r} 7 \\ 17 \overline{) 120} \\ 1 \end{array}$$

Ejercicio 184 . Escriba las tablas para las operaciones en \mathbb{Z}_2 , \mathbb{Z}_3 , \mathbb{Z}_4 , \mathbb{Z}_5 .

Ejercicio 185 . Encuentre los inversos multiplicativos de los elementos distintos de 0 en \mathbb{Z}_7 .

Ejercicio 186 . Encuentre las unidades de \mathbb{Z}_{18} , encuentre un elemento no nulo que no tenga inverso multiplicativo.

Definición 68

$$\mathbb{N}^{(\mathbb{N})} = \{f : \mathbb{N} \rightarrow \mathbb{N} \mid \text{soporte}(f) \text{ es finito}\}.$$

Donde

$$\text{soporte}(f) = \{k \in \mathbb{N} \mid f(k) \neq 0\}.$$

Entonces $\mathbb{N}^{(\mathbb{N})}$ es el conjunto de las sucesiones de naturales casi nulas (las sucesiones que eventualmente se anulan).

Definición 69 . Definiremos \preceq en $\mathbb{N}^{(\mathbb{N})}$ por $f \preceq g$ si $f(k) \leq g(k) \forall k \in \mathbb{N}$.

Ejercicio 187 . Demostrar que \preceq es reflexiva, antisimétrica y transitiva. Es decir, \preceq es una relación de orden en $\mathbb{N}^{(\mathbb{N})}$.

Como el conjunto de primos positivos es infinito, pero es un subconjunto de \mathbb{N}^+ , podemos numerarlos: 2, 3, 5, 7, 11, ... Digamos que $p_0 = 2, p_1 = 3, p_2 = 5, \dots$ etc.

Ahora podemos definir

$$\begin{array}{ccc} \mathbb{N}^{(\mathbb{N})} & \xrightarrow{\varphi} & \mathbb{N}^+ \\ f & \longmapsto & p_0^{f(0)} p_1^{f(1)} p_2^{f(2)} \dots \end{array} .$$

Por ejemplo, consideremos

$$f : 2, 3, 0, 5, 7, 0, \bar{0}, \dots$$

(la barra sobre 0 indica que este se repite a partir de ahí) Entonces

$$\varphi(f) = 2^2 \cdot 3^3 \cdot 5^0 \cdot 7^5 \cdot 11^7 \cdot 13^0 \cdot 17^0 \in \mathbb{N}^+.$$

El Teorema fundamental de la Aritmética nos garantiza que φ es suprayectiva, y la unicidad de la factorización nos dice que φ es inyectiva. Es decir φ es una biyección con inverso φ^{-1} .

Además,

$$f \preceq g \Leftrightarrow \varphi(f) \mid \varphi(g)$$

De aquí se siguen las dos afirmaciones que dejamos como ejercicios al lector:

Ejercicio 188 . Sean $a, b \in \mathbb{N}^+$ demostrar que $(a; b) = \varphi(\varphi^{-1}(a) \wedge \varphi^{-1}(b))$. (Recordar la definición de \wedge en la página 91).

Ejercicio 189 . Sean $a, b \in \mathbb{N}^+$ demostrar que $(a; b) = \varphi(\varphi^{-1}(a) \vee \varphi^{-1}(b))$. (Recordar la definición de \wedge en la página 91).

Ejemplo 103 . Sean

$$\begin{aligned} a &= 2^3 3^2 5^2 7^3 = 617400, \\ b &= 2^2 3^5 5^1 7^2 = 238140, \end{aligned}$$

entonces, según el ejercicio 188,

$$\begin{aligned} (a; b) &= \varphi(\varphi^{-1}(a) \wedge \varphi^{-1}(b)) = \\ &= \varphi(\varphi^{-1}(2^3 3^2 5^2 7^3) \wedge \varphi^{-1}(2^2 3^5 5^1 7^2)) = \\ &= \varphi((3, 2, 2, 3, 0, 0, \dots) \wedge (2, 5, 1, 2, 0, 0, \dots)) = \\ &= \varphi((2, 2, 1, 2, 0, 0, \dots)) = 2^2 3^2 5^1 7^2 = 8820. \end{aligned}$$

Comprobémoslo usando el algoritmo de Euclides:

$$\begin{array}{r} 2 \\ 238140 \left| \begin{array}{r} 617400 \\ 141120 \end{array} \right. \end{array}$$

$$\begin{array}{r} 1 \\ 141120 \left| \begin{array}{r} 238140 \\ 97020 \end{array} \right. \end{array}$$

$$\begin{array}{r} 1 \\ 97020 \left| \begin{array}{r} 141120 \\ 44100 \end{array} \right. \end{array}$$

$$\begin{array}{r} 2 \\ 44100 \left| \begin{array}{r} 97020 \\ 8820 \end{array} \right. \end{array}$$

$$\begin{array}{r} 5 \\ 8820 \left| \begin{array}{r} 44100 \\ 0 \end{array} \right. \end{array}$$

Ejercicio 190 . Encuentre el máximo común divisor y el mínimo común múltiplo de los siguientes conjuntos:

1. $\{20, -15, 22, -10\}.$
2. $\{168, 842, 252\}.$

4.10 Congruencias

Consideremos la congruencia

$$x \stackrel{n}{\equiv} a \quad (4.9)$$

El conjunto de todas las soluciones de esta congruencia, es decir el conjunto de todos los enteros que son congruentes con a módulo n es

$$\{a + nz \mid z \in \mathbb{Z}\} = a + n\mathbb{Z}.$$

Consideremos ahora

$$ax \stackrel{n}{\equiv} b \quad (4.10)$$

Esta congruencia tiene solución si y sólo si

$$\begin{aligned} \exists s \in \mathbb{Z} \text{ tal que } as \stackrel{n}{\equiv} b \text{ si y sólo si } as - b \in n\mathbb{Z} \text{ si y sólo si } b \in -as + n\mathbb{Z} &\Leftrightarrow \\ \Leftrightarrow b \in -as + n\mathbb{Z} \subseteq a\mathbb{Z} + n\mathbb{Z} = (a; n)\mathbb{Z} &\Leftrightarrow \\ \Leftrightarrow (a; n) \mid b. \end{aligned}$$

Ahora, supongamos que en efecto $(a; n) \mid b$ y tratemos de resolver $ax \stackrel{n}{\equiv} b$.

Escribamos de nuevo esta congruencia en la forma equivalente

$$ax - b = nz, z \in \mathbb{Z}.$$

Como $(a; n)$ divide a a, n , y b , entonces podemos escribir

$$(a \not\mid (a; n)) x - b \not\mid (a; n) = (n \not\mid (a; n)) z, \quad z \in \mathbb{Z} \quad (4.11)$$

que se transforma fácilmente en

$$(a \not\mid (a; n)) x \stackrel{n \not\mid (a; n)}{\equiv} b \not\mid (a; n) \quad (4.12)$$

Notemos ahora lo siguiente:

$$(a \not\mid (a; n)) y (n \not\mid (a; n)) \quad (4.13)$$

son primos relativos:

$$(a; n)((a \not\mid (a; n)); (n \not\mid (a; n))) = (a; n) \cdot 1,$$

así que cancelando $(a; n)$, obtenemos

$$((a \not\mid (a; n)); (n \not\mid (a; n))) = 1.$$

En esta situación, $(a \diagup (a; n))$ es una unidad en $\mathbb{Z}_{(n \diagup (a; n))}$ ⁹ y por lo tanto $\exists t \in \mathbb{Z}$ tal que $(a \diagup (a; n)) t \stackrel{n \diagup (a; n)}{\equiv} 1$. Por lo tanto obtenemos la congruencia

$$x \stackrel{n \diagup (a; n)}{\equiv} t(b \diagup (a; n)) \quad (4.14)$$

que tiene las soluciones $t(b \diagup (a; n)) + (n \diagup (a; n))k$, $k \in \mathbb{Z}$.

Desde luego, todas estas son soluciones de nuestra congruencia original

$$ax \stackrel{n}{\equiv} b \quad (4.15)$$

Pero las soluciones para $k = 0, 1, \dots, (a; n) - 1$ son incongruentes (no son congruentes) módulo n . La solución para $k = (a; n)$ es congruente módulo n con la que corresponde a $k = 0$. En resumen, escribimos las observaciones anteriores en el siguiente teorema.

Teorema 69 . *La congruencia*

$$ax \stackrel{n}{\equiv} b \quad (4.16)$$

tiene solución si y sólo si $(a; n) \mid b$. En este caso hay $(a; n)$ soluciones incongruentes módulo n , y las soluciones son

$$x = (b \diagup (a; n))t + (n \diagup (a; n))k, \quad k \in \{0, 1, \dots, (a; n) - 1\},$$

donde la clase de t en $\mathbb{Z}_{(n \diagup (a; n))}$ es un inverso de la clase de $(a \diagup (a; n))$ (en $\mathbb{Z}_{(n \diagup (a; n))}$).

Ejemplo 104 . *La congruencia*

$$15x \stackrel{18}{\equiv} 9 \quad (4.17)$$

tiene solución pues

$$(15; 18) = 3 \mid 9.$$

Además debe tener 3 soluciones incongruentes módulo 18. En efecto:

$$15x \stackrel{18}{\equiv} 9 \quad (4.18)$$

⁹Estrictamente, lo que es una unidad en $\mathbb{Z}_{(n \diagup (a; n))}$ es la clase de congruencia de $a \diagup (a; n)$ módulo $n \diagup (a; n)$.

es equivalente (tiene las mismas soluciones) que

$$5x \stackrel{6}{\equiv} 3 \quad (4.19)$$

Como $1 = 6 - 5$, vemos que -1 es un inverso de 5 en \mathbb{Z}_6 , así que multiplicando por -1 la congruencia anterior, obtenemos

$$x \stackrel{6}{\equiv} -3 \quad (4.20)$$

de donde obtenemos las soluciones $x = -3$, $x = 3$, $x = 9$.

La siguiente solución es $x = 15$, pero 15 ya es congruente con -3 módulo 18 . En efecto,

$$15(-3) = -45. - 45 - 9 = -54 = -3 \cdot 18.$$

Y por último,

$$15(3) = 45. 45 - 9 = 36 = 2 \cdot 18.$$

Otro ejemplo:

Ejemplo 105

$$42x \stackrel{3 \cdot 5 \cdot 7}{\equiv} 63$$

tiene solución, pues $(42; 3 \cdot 5 \cdot 7) = 3 \cdot 7 \mid 63$, además tiene 21 soluciones incongruentes módulo $3 \cdot 5 \cdot 7 = 105$.

En efecto,

$$42x \stackrel{3 \cdot 5 \cdot 7}{\equiv} 63$$

tiene las mismas soluciones que

$$2x \stackrel{5}{\equiv} 3$$

Multiplicando por 3 esta congruencia, obtenemos

$$x \stackrel{5}{\equiv} 9 \stackrel{5}{\equiv} 4.$$

Así que las soluciones buscadas son

$$x = 4 + 5k, k \in \{0, \dots, 20\}.$$

Por ejemplo, tomemos $k = 17$, entonces

$$x = 4 + 5 \cdot 17 = 89.$$

Ahora $42 \cdot 89 = 3738$, y $3738 \stackrel{3 \cdot 5 \cdot 7}{\equiv} 63$ Es decir, $(3738 - 63) = 35 \cdot (3 \cdot 5 \cdot 7)$, y así $3738 \stackrel{3 \cdot 5 \cdot 7}{\equiv} 63$.

Si tomamos $k = 13$, entonces

$$x = 4 + 5 \cdot 13 = 69; \quad 42 \cdot 69 = 2898; \quad 2898 - 63 = 2835; \quad 2835 / (3 \cdot 5 \cdot 7) = 27.$$

Por lo que $2898 - 63 = (3 \cdot 5 \cdot 7) \cdot 27$, es decir, $2898 \stackrel{3 \cdot 5 \cdot 7}{\equiv} 63$.

Ejercicio 191 . Demostrar que todo entero es congruente módulo 7 con un número en el siguiente conjunto: $\{191, 7, 54, 31, 36, 20, 765\}$.

Ejercicio 192 . Demostrar que todo primo mayor que 5 es de la forma $30m + n$ con $n \in \{1, 7, 11, 13, 17, 19, 23, 29\}$.

Ejercicio 193 . Muestre que si $a \stackrel{m}{\equiv} b$ entonces $(a; m) = (b; m)$.

Ejercicio 194 . Encuentre el residuo de las siguientes divisiones:

$$1. \quad 1^5 + 2^5 + \dots + 1080^5 \text{ dividido entre 7.}$$

$$2. \quad 1! + 2! + 3! + \dots + (10^{10})! \text{ dividido entre 24.}$$

$$3. \quad \binom{3}{3} + \binom{4}{3} + \binom{5}{3} + \dots + \binom{102}{3} \text{ dividido entre 7.}$$

4.11 Sistemas de congruencias

Podemos tratar de encontrar las soluciones de un sistema de congruencias como

$$\begin{aligned} x &\stackrel{m}{\equiv} a \\ x &\stackrel{n}{\equiv} b \end{aligned} \tag{4.21}$$

es decir, queremos saber cuando existe una solución de la primera congruencia que resuelva también la segunda.

Como las soluciones de la primera congruencia son los enteros en $a + m\mathbb{Z}$, es decir los enteros $x = a + mz$, veamos cuando sucede que una de estas soluciones resuelve la segunda congruencia:

$$a + mz \stackrel{n}{\equiv} b \tag{4.22}$$

Esta congruencia tiene las mismas soluciones que la congruencia

$$mz \stackrel{n}{\equiv} b - a \quad (4.23)$$

Ya sabemos que esta última congruencia tiene solución si y sólo si $(m; n) \mid (b - a)$, y además tiene las siguientes $(m; n)$ soluciones incongruentes módulo n :

$$z = ((b - a) / (m; n))t + (n / (m; n))k,$$

donde t es un inverso multiplicativo de $m / (m; n)$ módulo $n / (m; n)$ y se escoge $k \in \{0, 1, \dots, (m; n) - 1\}$.

Sustituyendo estas soluciones para z en 4.22 obtenemos las soluciones comunes.

Ejemplo 106 . Resolver

$$\begin{aligned} 3x &\stackrel{45}{\equiv} 12 \\ 5x &\stackrel{7}{\equiv} 2 \end{aligned} \quad (4.24)$$

El sistema de congruencias anterior tiene el mismo conjunto de soluciones que

$$\begin{aligned} x &\stackrel{15}{\equiv} 4 \\ 5x &\stackrel{7}{\equiv} 2 \end{aligned} \quad (4.25)$$

5 que es el coeficiente de x en la segunda congruencia, es una unidad en \mathbb{Z}_7 : $1 = 3 \cdot 5 - 2 \cdot 7$. De donde vemos que 3 es inverso de 5 módulo 7. Así, que multiplicando la segunda congruencia en 4.25 por 3, obtenemos el sistema

$$\begin{aligned} x &\stackrel{15}{\equiv} 4 \\ x &\stackrel{7}{\equiv} 6 \end{aligned} \quad (4.26)$$

La solución de la primera congruencia es

$$x = 4 + 15z, \quad z \in \mathbb{Z}, \quad (4.27)$$

ahora veamos cuales de estas soluciones resuelven también la segunda congruencia, es decir, resolvamos

$$4 + 15z \stackrel{7}{\equiv} 6 \quad (4.28)$$

que equivale a $15z \equiv 7 \pmod{6-4}$.

Como $15 \equiv 1$, obtenemos la congruencia equivalente $z \equiv 2$, cuyas soluciones son $z = 2 + 7k$, $k \in \mathbb{Z}$. Sustituyendo en 4.27 obtenemos

$$x = 4 + 15(2 + 7k), \quad k \in \mathbb{Z}$$

Es decir,

$$x = 34 + 105k, \quad k \in \mathbb{Z},$$

o bien,

$$x \equiv^{105} 34$$

En efecto,

$$\begin{aligned} 34 &\equiv^{15} 4 \\ 34 &\equiv^7 6 \end{aligned} \tag{4.29}$$

Otro ejemplo:

Ejemplo 107

$$\begin{aligned} 3x &\equiv^6 9 \\ 10x &\equiv^{35} 45 \end{aligned} \tag{4.30}$$

es equivalente con

$$\begin{aligned} x &\equiv^2 3 \\ 2x &\equiv^7 9 \end{aligned} \tag{4.31}$$

que se obtuvo al dividir la primera congruencia entre $3 = (3; 6)$ y la segunda entre $5 = (10; 35)$.

La primera congruencia equivale a $x \equiv^2 1$, y como 4 es inverso multiplicativo de 2 módulo 7, si multiplicamos por 4 la segunda congruencia y después reducimos módulo 7, obtenemos

$$x \equiv^7 36 \equiv^7 1$$

Entonces 4.30 es equivalente con

$$\begin{aligned} x &\equiv 1 \\ x &\equiv 1 \end{aligned} \tag{4.32}$$

Es claro que 1 es una solución de ambas congruencias. Si s fuera otra solución, entonces

$$\begin{aligned} s &\equiv 1 \\ s &\equiv 1 \end{aligned} \tag{4.33}$$

por lo que $2 \mid s - 1$ y $7 \mid s - 1$. Entonces (puesto que 2 y 7 son primos relativos) $14 \mid s - 1$, es decir,

$$x \equiv 1$$

nos da todas las soluciones del sistema 4.30, así que el conjunto de soluciones de 4.30 es $1 + 14\mathbb{Z}$.

Tomemos, por ejemplo, $z = -7$, entonces $1 + 14 \cdot (-7) = -97$. Así que -97 debe ser solución de

$$\begin{aligned} 3x &\equiv 9 \\ 10x &\equiv 45 \end{aligned} \tag{4.34}$$

$$3(-97) = -291,$$

$$-291 - 9 = -300 = -50 \cdot 6.$$

Entonces -97 resuelve la primera congruencia.

$$\text{Ahora, } 10(-97) = -970.$$

$-970 - 45 = 35 \cdot (-29) = -1015$. Por lo tanto -97 resuelve también la segunda congruencia.

Podemos resumir la discusión anterior en el siguiente teorema.

Teorema 70 .

1. El sistema de congruencias

$$\begin{aligned} x &\equiv a \\ x &\equiv b \end{aligned} \tag{4.35}$$

tiene solución si y sólo si $(m; n) \mid a - b$.

2. El sistema de congruencias

$$\begin{aligned} ax &\stackrel{m}{\equiv} b \\ cx &\stackrel{n}{\equiv} d \end{aligned} \tag{4.36}$$

tiene solución si cada una de las congruencias tiene solución, y si el sistema resultante al dividir la primera congruencia entre $(a; m)$ y la segunda entre $(c; n)$ el sistema resultante tiene solución.

Es decir,

en $\mathbb{Z}_{m/(a;m)}$ y u es inverso de $(c/(c;n))$ en $\mathbb{Z}_{n/(c;n)}$

$\Leftrightarrow (m/ (a; m) ; n/ (c; n)) \mid (t (b/ (a; m)) - u (d/ (c; n)))$ (donde t es inverso de $a/ (a; m)$ en $\mathbb{Z}_{n/ (c; n)}$ y u es inverso de $c/ (c; n)$ en $\mathbb{Z}_{n/ (c; n)}$).

Debemos aclarar aquí que la última condición, aunque correcta, se escribe únicamente por completar la caracterización. No se pretende ni se recomienda que se aprenda de memoria, pues siempre es mucho más fácil proceder a partir del sistema original,

$$\begin{aligned} ax &\stackrel{m}{\equiv} b \\ cx &\stackrel{n}{\equiv} d \end{aligned} \tag{4.37}$$

simplificándolo, a un sistema donde los coeficientes de x sean 1, y una vez hecho esto ya es inmediata la solución del sistema.

4.11.1 El Teorema chino del residuo

Teorema 71 . *El sistema de congruencias*

$$\begin{aligned} x &\stackrel{m_1}{\equiv} a_1 \\ x &\stackrel{m_2}{\equiv} a_2 \\ &\vdots \\ x &\stackrel{m_k}{\equiv} a_k \end{aligned} \tag{4.38}$$

tiene solución si y sólo si cada par de congruencias tiene solución común, es decir, si y sólo si

$$(m_i; m_j) \mid (a_i - a_j), \forall i, j \in \{1, \dots, k\}. \quad (4.39)$$

Además, cuando el sistema tiene solución, entonces todas las soluciones son congruentes módulo $[m_1; m_2; \dots; m_k]$.

Demostración. \Leftarrow) En este sentido de la demostración, veremos que si se cumple la condición 4.39, entonces el sistema tiene solución y que además todas las soluciones son congruentes módulo $[m_1; m_2; \dots; m_k]$.

Por inducción sobre k , el número de congruencias.

Base. Si $k = 1$, no hay nada que demostrar, y para $k = 2$, tenemos que $(m_1; m_2) \mid a_1 - a_2$ y el sistema

$$\begin{aligned} x &\stackrel{m_1}{\equiv} a_1 \\ x &\stackrel{m_2}{\equiv} a_2 \end{aligned}$$

tiene solución por el teorema 70, resta ver que dos soluciones son congruentes mod $[m_1; m_2]$. Si s, t son soluciones, entonces $s - a_1 \in m_1 \mathbb{Z}$, $t - a_1 \in m_1 \mathbb{Z}$, $s - a_2 \in m_2 \mathbb{Z}$, $t - a_2 \in m_2 \mathbb{Z}$, luego

$$\begin{aligned} s - t &= (s - a_1) - (t - a_1) \in m_1 \mathbb{Z} \text{ y} \\ &= (s - a_2) - (t - a_2) \in m_2 \mathbb{Z}. \end{aligned}$$

Por lo tanto $(s - t) \in m_1 \mathbb{Z} \cap m_2 \mathbb{Z} = [m_1; m_2] \mathbb{Z}$, es decir que $s \stackrel{[m_1; m_2]}{\equiv} t$.

Paso inductivo. Supongamos que $k > 2$, y que la afirmación se cumple para sistemas de congruencias con $k - 1$ congruencias. Notemos que la condición 4.39 se cumple para las primeras $k - 1$ congruencias, si se cumple para todas.

Por hipótesis de inducción, supondremos que hay una solución s del sistema de las $k - 1$ primeras congruencias y que las demás soluciones son congruentes módulo $[m_1; m_2; \dots; m_{k-1}]$, así que el sistema

$$\begin{aligned} x &\stackrel{m_1}{\equiv} a_1 \\ x &\stackrel{m_2}{\equiv} a_2 \\ &\vdots \\ x &\stackrel{m_{k-1}}{\equiv} a_{k-1} \end{aligned} \quad (4.40)$$

es equivalente a

$$x \stackrel{[m_1; m_2; \dots; m_{k-1}]}{\equiv} s$$

por lo tanto 4.38 es equivalente a

$$\begin{aligned} x &\stackrel{[m_1; m_2; \dots; m_{k-1}]}{\equiv} s \\ x &\stackrel{m_k}{\equiv} a_k \end{aligned} \quad (4.41)$$

Para ver que este sistema tiene solución, basta demostrar que

$$([m_1; m_2; \dots; m_{k-1}]; m_k) \mid (s - a_k). \quad (4.42)$$

Ahora, recordemos que el máximo común divisor se distribuye sobre el mínimo común múltiplo, así que

$$([m_1; m_2; \dots; m_{k-1}]; m_k) = [(m_1; m_k); (m_2; m_k); \dots; (m_{k-1}; m_k)]$$

por lo tanto, para comprobar 4.42, basta demostrar que $(m_i; m_k) \mid (s - a_k)$, $\forall i < k$.

Pero s es solución de $x \stackrel{m_i}{\equiv} a_i$ (por lo tanto, $s = a_i + km_i$) y por otra parte $(m_i; m_k) \mid (a_i - a_k)$. Así, tenemos que $(m_i; m_k) \mid (s - km_i - a_k) = (s - a_k) - km_i$. Así, $(s - a_k) \stackrel{(m_i; m_k)}{\equiv} km_i \stackrel{(m_i; m_k)}{\equiv} 0$.

Con esto vemos que 4.41 tiene solución.

Además todas las soluciones de 4.41 son congruentes módulo

$$[[m_1; m_2; \dots; m_{k-1}]; m_k] = [m_1; m_2; \dots; m_{k-1}; m_k].$$

\Rightarrow) En la otra dirección, tenemos que para que el sistema completo de congruencias tenga solución, se necesita que cada pareja de congruencias la tenga, y por lo tanto se tiene la condición 4.39. Ahora, si t_1, t_2 fueran dos soluciones del sistema, entonces $t_1 \stackrel{m_i}{\equiv} t_2, \forall i \in \{1, \dots, k\}$.

Entonces $m_i \mid t_1 - t_2, \forall i \in \{1, \dots, k\}$, por lo que $[m_1; m_2; \dots; m_{k-1}; m_k] \mid t_1 - t_2$, es decir, $t_1 \stackrel{[m_1; m_2; \dots; m_{k-1}; m_k]}{\equiv} t_2$. ■

Ejemplo 108 . Consideremos el sistema de congruencias

$$\begin{aligned} x &\stackrel{2 \cdot 3 \cdot 5 \cdot 7}{\equiv} 14 \\ x &\stackrel{2^2 \cdot 5}{\equiv} 24 \\ x &\stackrel{2^5 \cdot 3}{\equiv} 29 \\ x &\stackrel{9 \cdot 5}{\equiv} 44 \end{aligned}.$$

Notemos que este sistema se puede resolver: la primera congruencia (es compatible con la segunda, es decir, se pueden resolver simultáneamente) pues

$$(2 \cdot 3 \cdot 5 \cdot 7; 2^2 \cdot 5) = 10 \text{ divide a } 14 - 24 = -10.$$

La primera congruencia es compatible con la tercera pues $(2 \cdot 3 \cdot 5 \cdot 7; 25 \cdot 3) = 15$ divide a $14 - 29 = -15$. La primera congruencia es compatible con la última pues $(2 \cdot 3 \cdot 5 \cdot 7; 9 \cdot 5) = 15$ que divide a $14 - 44 = -30$.

Compatibilidad de la

segunda con tercera: $(2^2 \cdot 5; 25 \cdot 3) = 5 \mid 24 - 29 = -5$.

segunda con cuarta: $(2^2 \cdot 5; 9 \cdot 5) = 5 \mid 24 - 44 = -20$.

tercera con cuarta: $(25 \cdot 3; 9 \cdot 5) = 15 \mid 29 - 44 = -15$.

Consideremos el sistema de las dos primeras congruencias:

$$\begin{aligned} x &\stackrel{2 \cdot 3 \cdot 5 \cdot 7}{\equiv} 14 \\ x &\stackrel{2^2 \cdot 5}{\equiv} 24 \end{aligned}$$

Encontrémosle una solución particular: todas las soluciones de la primera congruencia son de la forma $14 + (2 \cdot 3 \cdot 5 \cdot 7)z$, $z \in \mathbb{Z}$ (esto es lo mismo que decir que el conjunto de soluciones de la primera congruencia es $14 + (2 \cdot 3 \cdot 5 \cdot 7) \mathbb{Z}$). Alguna de estas soluciones resuelve también la segunda:

$$x = 14 + 210z \stackrel{2^2 \cdot 5}{\equiv} 24$$

o bien,

$$210z \stackrel{2^2 \cdot 5}{\equiv} 24 - 14$$

es decir,

$$210z \stackrel{20}{\equiv} 10$$

Como $(210; 20) = 10$, dividamos entre 10 todos los números de la congruencia anterior (incluyendo el módulo) para obtener

$$21z \stackrel{2}{\equiv} 1 \text{ que equivale a } z \stackrel{2}{\equiv} 1$$

por lo tanto, podemos tomar $z = 1$ y vemos que una solución común a las dos primeras congruencias es $14 + (2 \cdot 3 \cdot 5 \cdot 7) \cdot 1 = 224$.

Las dos primeras congruencias son equivalentes a

$$x \stackrel{420}{\equiv} 224$$

ya que $[2 \cdot 3 \cdot 5 \cdot 7; 2^2 \cdot 5] = 420$.

Repitamos el procedimiento para el sistema de las dos últimas congruencias:

$$\begin{aligned} x &\stackrel{25 \cdot 3}{\equiv} 29 \\ x &\stackrel{9 \cdot 5}{\equiv} 44 \end{aligned}$$

$x = 29 + (25 \cdot 3)z \stackrel{9 \cdot 5}{\equiv} 44 \Leftrightarrow 75z \stackrel{9 \cdot 5}{\equiv} (44 - 29) = 15$. Como $(75; 9 \cdot 5) = 15$, dividiendo la última congruencia entre 15 (incluyendo el módulo) obtenemos $5z \stackrel{3}{\equiv} 1$ si multiplicamos por 2 obtenemos: $z \stackrel{3}{\equiv} 2$. Tomemos $z = 2$, obteniendo $x = 29 + (25 \cdot 3)2 = 179$. Tenemos que el sistema de las dos últimas congruencias equivale a la congruencia

$$x \stackrel{225}{\equiv} 179$$

ya que $[25 \cdot 3; 9 \cdot 5] = 225$.

Por último, resta resolver simultáneamente el sistema

$$\begin{aligned} x &\stackrel{420}{\equiv} 224 \\ x &\stackrel{225}{\equiv} 179 \end{aligned}$$

$x = 224 + 420z \stackrel{225}{\equiv} 179 \Leftrightarrow 420z \stackrel{225}{\equiv} 179 - 224 = -45$, como $(420; 225) = 15$, tenemos que resolver $(420/15)z \stackrel{(225/15)}{\equiv} (-45/15)$, es decir

$$28z \stackrel{15}{\equiv} -3,$$

esta congruencia equivale a $-2z \stackrel{15}{\equiv} -3$, así que $2z \stackrel{15}{\equiv} 3$, multiplicando por 8 y tomando residuos módulo 15, tenemos que

$$z \stackrel{15}{\equiv} 9,$$

tomando $z = 9$, obtenemos la solución $x = (224 + (420 \cdot 9)) = 4004$, entonces el sistema original es equivalente a la congruencia

$$x \stackrel{6300}{\equiv} 4004$$

pues $[210; 20; 75; 45] = 6300$.

Así que 4004 y $4004 - 6300 = -2296$ son soluciones del sistema

$$\begin{aligned} x &\stackrel{2 \cdot 3 \cdot 5 \cdot 7}{\equiv} 14 \\ x &\stackrel{2^2 \cdot 5}{\equiv} 24 \\ x &\stackrel{25 \cdot 3}{\equiv} 29 \\ x &\stackrel{9 \cdot 5}{\equiv} 44 \end{aligned}$$

En efecto:

$$(4004 - 14) / (2 \cdot 3 \cdot 5 \cdot 7) = 3990 / 210 = 19.$$

$$(4004 - 24) / (2^2 \cdot 5) = 3980 / 20 = 199.$$

$$(4004 - 29) / (25 \cdot 3) = 3975 / 25 \cdot 3 = 477.$$

$$(4004 - 44) / (9 \cdot 5) = 88.$$

También:

$$(-2296 - 14) / (2 \cdot 3 \cdot 5 \cdot 7) = -11.$$

$$(-2296 - 24) / (2^2 \cdot 5) = -116.$$

$$(-2296 - 29) / (25 \cdot 3) = -31.$$

$$(-2296 - 44) / (9 \cdot 5) = -52.$$

Notemos que como caso particular del teorema anterior, obtenemos el Teorema chino del residuo:

Teorema 72 (*Teorema chino del residuo*). *El sistema de congruencias*

$$\begin{aligned} x &\stackrel{m_1}{\equiv} a_1 \\ x &\stackrel{m_2}{\equiv} a_2 \\ &\vdots \\ x &\stackrel{m_k}{\equiv} a_k \end{aligned} \tag{4.43}$$

tiene solución si cada par de módulos son primos relativos dos a dos, esto es, si $(m_i; m_j) = 1$, $i \neq j$. En este caso todas las soluciones son congruentes módulo $m_1 \cdot m_2 \cdot \dots \cdot m_k$.

Demarcación. Simplemente observemos que se aplica el teorema anterior: $(m_i; m_j) = 1 \mid a_i - a_j$, $i \neq j$. Todas las soluciones son congruentes módulo $[m_1; m_2; \dots; m_k] = m_1 \cdot m_2 \cdot \dots \cdot m_k$. ■

Ejemplo 109 . Consideremos el sistema de congruencias

$$\begin{aligned} x &\stackrel{83}{\equiv} 32 \\ x &\stackrel{110}{\equiv} 70 \\ x &\stackrel{135}{\equiv} 30 \end{aligned}$$

¹⁰83 es primo, $110 = (2)(5)(11)$, $135 = (3)^3(5)$. Notemos que no estamos dentro de las hipótesis del teorema Chino, pues $(110; 135) = 5$ ($(110; 135) = 5(22; 27) = 5(2 \cdot 11; 3^3) = 5 \cdot 1 = 5$), sin embargo, como las dos últimas congruencias

$$\begin{aligned} x &\stackrel{110}{\equiv} 70 \\ x &\stackrel{135}{\equiv} 30 \end{aligned} \tag{4.44}$$

son equivalentes a: $x = 70 + k \cdot 110 \stackrel{135}{\equiv} 30$ es decir a $k \cdot 110 \stackrel{135}{\equiv} -40$ que tiene las mismas soluciones que $k \cdot 22 \stackrel{27}{\equiv} -8$. Ahora, 16 es el inverso de 22 módulo 27: $16 \cdot 22 = 352$ y $352/27 = 13\frac{1}{27}$. Por lo tanto, multiplicando $k \cdot 22 \stackrel{27}{\equiv} -8$ por 16, obtenemos $k \stackrel{27}{\equiv} -8 \cdot 16 = -128 \stackrel{27}{\equiv} 7$. Por lo tanto una solución particular de 4.44 es $x = 70 + 7 \cdot 110 = 840$. En resumidas cuentas, 4.44, es equivalente a la congruencia $x \stackrel{2970}{\equiv} 840$.¹¹

Ahora sí, el sistema

$$\begin{aligned} x &\stackrel{83}{\equiv} 32 \\ x &\stackrel{2970}{\equiv} 840 \end{aligned}$$

satisface las condiciones de Teorema chino del residuo: 83 es primo y $83 \nmid 2970 = (2)(3)^3(5)(11)$.

Resolvamos: $x = 32 + 83z \stackrel{2970}{\equiv} 840$ equivale a $83z \stackrel{2970}{\equiv} 840 - 32 = 808$. 2147 es el inverso multiplicativo de 83 módulo 2970 (Ahora que, $2147 \stackrel{2970}{\equiv} 2147 - 2970 = -823$).¹² Multipliquemos $83z \stackrel{2970}{\equiv} 808$ por -823 :

$$z \stackrel{2970}{\equiv} 808 \cdot (-823) = -664984$$

$-664984 \stackrel{2970}{\equiv} 296$, pues $(-664984 - 296)/2970 = -224$. Sustituyendo $z = 296$ en $x = 32 + 83z$, obtenemos $x = 32 + 83 \cdot (296) = 24600$.

Así que una solución particular al sistema es 24600 y el sistema original es equivalente con la congruencia

$$x \stackrel{83 \cdot 2970}{\equiv} 24600$$

¹⁰ Ejemplo de Ch'in Chiu-Shao en su libro de 1247 y por quien se califica de "chino" al teorema de arriba.

¹¹ $2970 = [110; 135]$.

¹² Veámoslo:

$83 \cdot (-823) = -68309$; $-68309/2970 = -23 + \frac{1}{2970}$. Es decir que $83 \cdot (-823) = (-23) \cdot 2970 + 1 \stackrel{2970}{\equiv} 1$.

o bien

$$x \stackrel{246510}{\equiv} 24600$$

Comprobación: veamos que 24600 resuelve cada una de las siguientes congruencias:

$$\begin{aligned} x &\stackrel{83}{\equiv} 32 \\ x &\stackrel{110}{\equiv} 70 \\ x &\stackrel{135}{\equiv} 30 \end{aligned}$$

En efecto, $(24600 - 32)/83 = 296$, $(24600 - 70)/110 = 223$ y $(24600 - 30)/135 = 182$. Note que la siguiente solución positiva es $24600 + 246510 = 271110$.

Otro caso aún más particular del teorema 72, es el siguiente:

Teorema 73 . Si p_1, \dots, p_k son primos positivos distintos, entonces el sistema de congruencias

$$\begin{aligned} x &\stackrel{p_1}{\equiv} a_1 \\ x &\stackrel{p_2}{\equiv} a_2 \\ &\vdots \\ x &\stackrel{p_k}{\equiv} a_k \end{aligned} \tag{4.45}$$

tiene solución y todas las soluciones son congruentes módulo $p_1 \cdot \dots \cdot p_k$.

Ejemplo 110 . Resolvamos el sistema

$$\begin{aligned} x &\stackrel{3}{\equiv} 1 \\ x &\stackrel{5}{\equiv} 2 \\ x &\stackrel{7}{\equiv} 3 \end{aligned} \tag{4.46}$$

$x = 1 + 3z \stackrel{5}{\equiv} 2 \Leftrightarrow 3z \stackrel{5}{\equiv} 1 \Leftrightarrow 2 \cdot 3z \stackrel{5}{\equiv} 2 \Leftrightarrow z \stackrel{5}{\equiv} 2$. Así que $x = 1 + 3 \cdot 2 = 7$ es una solución de las dos primeras congruencias. Por lo tanto las dos primeras congruencias son equivalentes a la congruencia $x \stackrel{15}{\equiv} 7$, cuyas soluciones son $7 + 15w, w \in \mathbb{Z}$.

Ahora, $x = 7 + 15w \stackrel{7}{\equiv} 3 \Leftrightarrow 15w \stackrel{7}{\equiv} 3 \Leftrightarrow w \stackrel{7}{\equiv} 3$. Sustituyendo $w = 3$ en $x = 7 + 15 \cdot w$, obtenemos $x = 7 + 15(3) = 52$. Por lo tanto 4.46 equivale a

$$x \stackrel{105}{\equiv} 52$$

En efecto: $(52 - 1)/3 = 17, (52 - 2)/5 = 10, (52 - 3)/7 = 7$.

Ejercicio 195 . Una banda de 17 ladrones roba un gran saco de billetes. Tratan de repartir los billetes equitativamente, pero sobran 3 billetes. Dos de los ladrones empiezan a pelear por el sobrante hasta que uno dispara al otro. El dinero se redistribuye, pero esta vez sobran 10 billetes. De nuevo empieza la pelea y otro ladrón resulta muerto. Cuando el dinero se redistribuye, no sobra nada. ¿Cuál es la menor cantidad posible de billetes que los ladrones robaron?

Ejercicio 196 . Hallar 4 enteros consecutivos que sean múltiplos de 5, 7, 9 y 11 respectivamente.

Ejercicio 197 . La producción diaria de huevos en una granja es inferior a 75. Cierta noche el recolector informa que la cantidad de huevos recogida es tal que contada de tres en tres sobran 2, contadas de cinco en cinco sobran 4 y contadas de 7 en 7 sobran 5. El capataz dice que no es posible ¿Quién tiene razón?

4.12 Ecuaciones diofantinas

Una ecuación de la forma

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = c \quad (4.47)$$

con coeficientes $a_1, a_2, \dots, a_n \in \mathbb{Z}$, $c \in \mathbb{Z}$ se llama diofantina. Las soluciones que nos interesan son las soluciones enteras.

La más sencilla de las ecuaciones diofantinas del tipo anterior es

$$ax = c \quad (4.48)$$

que tiene solución sólo si $a \mid c$, y en este caso, la solución es el entero $\frac{c}{a}$.

Consideremos ahora el siguiente caso de ecuación diofantina.

$$ax + by = c \quad (4.49)$$

Esta ecuación es tan sencilla que podemos sin más decir cuando tiene solución.

Teorema 74 . La ecuación 4.49 tiene solución $\Leftrightarrow (a; b) \mid c$.

Demostración. $\Rightarrow)$ Si $\exists (s, t) \in \mathbb{Z} \times \mathbb{Z}$ tal que $c = as + bt \in a\mathbb{Z} + b\mathbb{Z} = (a; b)\mathbb{Z}$, entonces $c \in (a; b)\mathbb{Z}$. Es decir, $(a; b) \mid c$.

$\Leftarrow)$ Si $(a; b) \mid c$, entonces $ax + by = c$ tiene las mismas soluciones que

$$\frac{a}{(a; b)}x + \frac{b}{(a; b)}y = \frac{c}{(a; b)}. \quad (4.50)$$

Ahora, como $\left(\frac{a}{(a; b)}; \frac{b}{(a; b)}\right) = 1^{13}$, entonces $\exists \alpha, \beta \in \mathbb{Z}$ tal que

$$\alpha \frac{a}{(a; b)} + \beta \frac{b}{(a; b)} = 1.$$

Multiplicando por $\frac{c}{(a; b)}$, obtenemos

$$\frac{a}{(a; b)}\alpha \frac{c}{(a; b)} + \frac{b}{(a; b)}\beta \frac{c}{(a; b)} = \frac{c}{(a; b)}.$$

Así, obtenemos una solución: $\left(\alpha \frac{c}{(a; b)}, \beta \frac{c}{(a; b)}\right)$. ■

Obsérvese que $ax + by = c$ tiene solución si y sólo si la congruencia $ax \stackrel{b}{\equiv} c$ tiene solución, de donde obtenemos nuevamente el criterio para la existencia de soluciones.

Del teorema anterior, podemos hacer la siguiente afirmación.

Lema 14 . *La ecuación diofantina 4.49 tiene las mismas soluciones que 4.50.*

Demostración. Nótese que cuando no hay soluciones, entonces $(a; b) \nmid c$ y la ecuación 4.50 ni siquiera tiene coeficientes enteros.

Ya hemos visto que las soluciones de 4.49 son también soluciones de 4.50.

Recíprocamente, si

$$\frac{a}{(a; b)}s + \frac{b}{(a; b)}t = \frac{c}{(a; b)}$$

¹³ $(a; b) \left(\frac{a}{(a; b)}; \frac{b}{(a; b)}\right) = \left((a; b) \frac{a}{(a; b)}; (a; b) \frac{b}{(a; b)}\right) = (a; b) \cdot 1 \Rightarrow \left(\frac{a}{(a; b)}; \frac{b}{(a; b)}\right) = 1.$

con $s, t \in \mathbb{Z}$, es claro que multiplicando por $(a; b)$ obtenemos

$$as + bt = c,$$

es decir que (s, t) también es solución de 4.49. ■

Por el teorema anterior, basta saber resolver ecuaciones diofantinas reducidas, es decir, con coeficientes primos relativos.

Ejemplo 111 . *Resolver*

$$2520x + 1188y = 108$$

Aplicando el algoritmo de Euclides:

$$\begin{array}{r} 2 \\ 1188 \overline{)2520} \end{array} \quad \begin{array}{r} 8 \\ 144 \overline{)1188} \end{array} \quad \begin{array}{r} 4 \\ 36 \overline{)144} \end{array},$$

vemos que el máximo común divisor de 2520 y 1188 es 36.

Ahora $36 \overline{)108}$, así que la ecuación

$$2520x + 1188y = 108$$

tiene solución y tiene las mismas soluciones que

$$\frac{2520}{36}x + \frac{1188}{36}y = \frac{108}{36},$$

es decir que

$$70x + 33y = 3.$$

Ahora queremos expresar 1 como combinación entera de 70 y 33. Nuevamente, usamos el algoritmo de Euclides:

$$\begin{array}{r} 8 \\ 4 \overline{)33} \end{array} \quad \begin{array}{r} 2 \\ 33 \overline{)70} \end{array}$$

Entonces $1 = 33 - 8 \cdot 4 = 33 - 8 \cdot (70 - 2 \cdot 33) = -8 \cdot 70 + (17 \cdot 33)$, así:

$$-8 \cdot 70 + (17 \cdot 33) = 1.$$

Multipliquemos por 3:

$$-24 \cdot 70 + (51 \cdot 33) = 3.$$

Así, obtenemos la solución $(-24, 51)$.

Comprobación:

$$2520(-24) + 1188(51) = 60588 - 60480 = 108.$$

Lema 15 . Si $a, b \in \mathbb{Z} \setminus \{0\}$ son primos relativos entonces las soluciones de

$$ax + by = 0$$

son

$$\{(-bz, az) \in \mathbb{Z} \times \mathbb{Z} \mid z \in \mathbb{Z}\}.$$

Demuestracción. $ax + by = 0 \Leftrightarrow ax = -by$. Si (s, t) es una solución, entonces

$$as = -bt.$$

Como $a \mid bt$ y $(a; b) = 1$, entonces $a \mid t$. Por lo tanto $t = az$, para alguna $z \in \mathbb{Z}$. Entonces

$$as = -baz,$$

por lo que $s = -bz$. Vemos pues, que una solución es de la forma $(-bz, az)$ con $z \in \mathbb{Z}$.

Recíprocamente, tomemos una pareja $(-bw, aw)$, con $w \in \mathbb{Z}$, entonces

$$a(-bw) = -b(aw).$$

■

Para poder expresar el siguiente teorema, es conveniente definir una suma en $\mathbb{Z} \times \mathbb{Z}$. Esto se hace de manera natural sumando “coordenada a coordenada”: $(a, b) + (c, d) = (a + c, b + d)$.

Ejercicio 198 . Demuestre que la suma anterior es asociativa, conmutativa, con neutro $(0, 0)$ y donde cada elemento (a, b) tiene inverso aditivo: $(-a, -b)$. De tal manera que $(a, b) - (c, d) = (a, b) + (-c, d) = (a - c, b - d)$.

Teorema 75 . La ecuación

$$ax + by = c \quad (4.51)$$

tiene conjunto de soluciones

$$S = (s, t) + \{(\alpha, \beta) \mid (\alpha, \beta) \text{ es solución de } ax + by = 0\},$$

¹⁴ donde (s, t) es una solución particular de 4.51.

Demostración. Denotemos $S_H = \{(\alpha, \beta) \mid (\alpha, \beta) \text{ es solución de } ax + by = 0\}$.

\subseteq) Sea $(u, v) \in S$, entonces $(u, v) = (s, t) + ((u, v) - (s, t))$. Basta notar que $(u, v) - (s, t) \in S_H$:

$$(u, v) - (s, t) = (u - s, v - t). \text{ Ahora}$$

$$a(u - s) + b(v - t) = (au + bv) - (as + bt) = c - c = 0,$$

por lo tanto $(u, v) - (s, t) \in S_H$.

\supseteq) $(s, t) + (\alpha, \beta)$ con $a\alpha + b\beta = 0 \Rightarrow (s, t) + (\alpha, \beta) = (s + \alpha, t + \beta)$ satisface:

$$a(s + \alpha) + b(t + \beta) = (as + bt) + (a\alpha + b\beta) = c + 0 = 0.$$

Por lo tanto $(s, t) + (\alpha, \beta) \in S$. ■

Teorema 76 . La ecuación

$$ax + by = c \quad (4.52)$$

tiene conjunto de soluciones

$$S = (s, t) + \{(\alpha, \beta) \mid (\alpha, \beta) \text{ es solución de } ax + by = 0\},$$

donde (s, t) es una solución particular y (α, β) es solución de

$$\hat{a}x + \hat{b}y = 0 \quad (4.53)$$

$$\text{donde } \hat{a} = \frac{a}{(a; b)} \text{ y } \hat{b} = \frac{b}{(a; b)}.$$

¹⁴ Naturalmente, $(s, t) + \{(\alpha, \beta) \mid (\alpha, \beta) \text{ es solución de } ax + by = 0\}$ denota el conjunto de sumas $\{(s, t) + (\alpha, \beta) \mid (\alpha, \beta) \text{ es solución de } ax + by = 0\}$.

Demostración. Se sigue del teorema anterior y de que

$$ax + by = 0$$

y

$$\hat{a}x + \hat{b}y = 0$$

comparten soluciones. ■

Ejemplo 112 . *Encontrar las soluciones de*

$$2x + 3y = 1$$

Una solución es $(-1, 1)$.

La solución de

$$2x + 3y = 0$$

es

$$(3z, -2z), \quad z \in \mathbb{Z}.$$

Por lo tanto la solución es

$$(-1, 1) + (3z, -2z) = (-1 + 3z, 1 - 2z), \quad z \in \mathbb{Z}.$$

Por ejemplo, $(14, -9)$ es una solución:

$$2 \cdot 14 + 3 \cdot (-9) = 28 - 27 = 1.$$

Ejemplo 113 . *Dos mercancías cuestan respectivamente \$71 y \$83 el kilo. ¿Qué cantidades enteras se pueden comprar con \$1670?*

$$71x + 83y = 1670$$

$$\begin{array}{r} 1 \\ 11 \overline{)12} \\ 1 \end{array} \quad \begin{array}{r} 5 \\ 12 \overline{)71} \\ 11 \end{array} \quad \begin{array}{r} 1 \\ 71 \overline{)83} \\ 12 \end{array},$$

por lo que

$$\begin{aligned} 1 &= 12 - 11 = (83 - 71) - (71 \cdot (-5) \cdot 12) \\ &= 83 - 2 \cdot 71 + 5 \cdot 12 = 83 - 2 \cdot 71 + 5 \cdot (83 - 71) \\ &= 6 \cdot 83 - 7 \cdot 71 \end{aligned}$$

Por lo tanto una solución es $(-7 \cdot 1670, 6 \cdot 1670)$, mientras que la solución de $71x + 83y = 0$ es $(83z, -71z)$, $z \in \mathbb{Z}$.

Las soluciones son entonces

$$(-7 \cdot 1670 + 83z, 6 \cdot 1670 - 71z), \quad z \in \mathbb{Z}.$$

Ahora debemos tomar en cuenta que queremos soluciones no negativas, así que:

$$-7 \cdot 1670 + 83z \geq 0 \Leftrightarrow 83z \geq 7 \cdot 1670 \Leftrightarrow z \geq \frac{1670 \cdot 7}{83} > 140.$$

y

$$\begin{aligned} 6 \cdot 1670 - 71z &\geq 0 \Leftrightarrow 6 \cdot 1670 \geq 71z \Leftrightarrow (6 \cdot 1670) / 71 = \frac{10020}{71} \geq z \\ &\Leftrightarrow 141 \frac{9}{71} \geq z \Leftrightarrow 141 \geq z \end{aligned}$$

La única solución es con $z = 141$.

$$(-7 \cdot 1670 + 83 \cdot 141, 6 \cdot 1670 - 71 \cdot 141) = (13, 9).$$

En efecto:

$$71 \cdot 13 + 83 \cdot 9 = 1670.$$

Ejercicio 199 . Hallar 4 enteros consecutivos que sean múltiplos de 5, 7, 9 y 11 respectivamente.

Ejercicio 200 . Sean a, b, c enteros tales que $a^2 + b^2 = c^2$. Demostrar que

1. ab es par.
2. ab es múltiplo de 3.
3. a o b es múltiplo de 4.
4. Un elemento de $\{a, b, c\}$ es múltiplo de 5.

4.13 Sistemas de numeración con bases distintas de 10

Notemos que 1999 es una abreviatura para el número

$$1 \cdot 10^3 + 9 \cdot 10^2 + 9 \cdot 10^1 + 9 \cdot 10^0$$

y que en general, el número decimal

$$c_k c_{k-1} \cdots c_1 c_0$$

es una abreviatura para

$$c_k \cdot 10^k + c_{k-1} \cdot 10^{k-1} + \cdots + c_1 \cdot 10^1 + c_0 \cdot 10^0,$$

donde c_0, c_1, \dots, c_k dígitos, es decir, pertenecen a $\{0, 1, \dots, 9\}$.

Consideremos el número $N = c_k c_{k-1} \cdots c_1 c_0$. Supongamos por un momento que no tenemos a la vista la representación decimal de N ¿cómo se obtiene? Notemos que c_0 es el residuo al dividir N entre 10:

$$10 \overline{) \begin{array}{r} c_k \cdot 10^{k-1} + c_{k-1} \cdot 10^{k-2} + \cdots + c_1 \cdot 10^0 \\ \hline N \\ c_0 \end{array}}$$

Ahora, por la misma razón, c_1 es el residuo al dividir $(N - c_0)/10$ entre 10:

$$10 \overline{) \begin{array}{r} c_k \cdot 10^{k-2} + c_{k-1} \cdot 10^{k-3} + \cdots + c_2 \cdot 10^0 \\ \hline c_k \cdot 10^{k-1} + c_{k-1} \cdot 10^{k-2} + \cdots + c_1 \cdot 10^0 \\ c_1 \end{array}}$$

etc.

Podemos hacer un esquema para este proceso:

$$\begin{array}{c|ccccc} & & & N & \\ & & & c_0 & \\ c_k \cdot 10^{k-1} + c_{k-1} \cdot 10^{k-2} + \cdots + c_1 \cdot 10^0 & & & c_1 & \\ c_k \cdot 10^{k-2} + c_{k-1} \cdot 10^{k-3} + \cdots + c_2 \cdot 10^0 & & & c_2 & \\ \vdots & & & \vdots & \\ c_k \cdot 10 + c_{k-1} & & & c_{k-1} & \\ c_k & & & c_k & \end{array}.$$

En lugar de tomar como base del sistema de numeración al número 10, podemos usar cualquier otro número natural mayor que 1, por ejemplo el 2, el 3 el 7 el 16.

Supongamos que queremos expresar 1999 en base dos, es decir queremos escribir

1999 en la forma

$$c_k 2^k + c_{k-1} 2^{k-1} + \cdots + c_1 2^1 + c_0 2^0,$$

con $c_k, c_{k-1}, \dots, c, c_0 \in \{0, 1\}$, el conjunto de dígitos binarios (bits).

Entonces

$$\begin{array}{r}
 62 & 124 & 249 & 499 & 999 \\
 2 \overline{)124} & 2 \overline{)249} & 2 \overline{)499} & 2 \overline{)999} & 2 \overline{)1999} \\
 0 & 1 & 1 & 1 & 1 \\
 \hline
 & & & & \\
 0 & 1 & 3 & 7 & 15 \\
 2 \overline{)1} & 2 \overline{)3} & 2 \overline{)7} & 2 \overline{)15} & 2 \overline{)31} \\
 1 & 1 & 1 & 1 & 1 \\
 \hline
 & & & & \\
 & & & & 0
 \end{array}$$

entonces la expresión binaria de 1999 es (leyendo los residuos de izquierda a derecha conforme los fuimos encontrando:

$$11111001111,$$

otra vez, es más cómodo el esquema

$$\begin{array}{r|c}
 1999 & 1 \\
 999 & 1 \\
 499 & 1 \\
 249 & 1 \\
 124 & 0 \\
 62 & 0 \\
 31 & 1 \\
 15 & 1 \\
 7 & 1 \\
 3 & 1 \\
 1 & 1 \\
 0 & 1
 \end{array}$$

Podemos hacer otro ejemplo, escribamos 1999 en base siete:

$$\begin{array}{r|l} 1999 & 4 \\ 285 & 5 \\ 40 & 5 \\ 5 & 5 \\ 0 & \end{array},$$

Así, $1999_{diez} = 5554_{siete}$:

$$5 \cdot 7^3 + 5 \cdot 7^2 + 5 \cdot 7 + 4 = 1715 + 245 + 35 + 4 = 1999.$$

Formalicemos un poco.

Teorema 77 . *Sea $b \in \mathbb{N} \setminus \{0, 1\}$. Para cualquier natural $N > 0$, existen $k \in \mathbb{N}$, $d_0, d_1, \dots, d_k \in \{0, 1, \dots, b-1\}$ únicos, tales que*

$$N = d_k b^k + \dots + d_1 b + d_0.$$

Demostración. Existencia.

Usaremos el segundo principio de inducción (sobre N).

Base de la inducción. Si $N = 1$, entonces $k = 0$ y $d_0 = 1$.

Paso inductivo. Supongamos que $N > 0$ y que la afirmación es válida para naturales menores que N .

Aplicemos el algoritmo de la división:

$$b \overline{)N}^q, \quad r \quad 0 \leq r < b$$

hagamos $d_0 = r$, y como $q = \frac{N-r}{b} < N$ (porque $b > 1$). Entonces aplicando la hipótesis de inducción a q tenemos que

$$\begin{aligned} q &= d_k b^{k-1} + d_{k-1} b^{k-2} + \dots + d_2 b + d_1, \\ \text{con } d_1, d_2, \dots, d_k &\in \{0, 1, \dots, b-1\}. \end{aligned}$$

Ahora, como $N = qb + r$, entonces

$$N = d_k b^k + \dots + d_1 b + d_0.$$

Con $d_0 = r, d_1, d_2, \dots, d_k \in \{0, 1, \dots, b-1\}$.

Unicidad. Si hubiera un natural positivo con más de dos representaciones, entonces por el principio del buen orden habría también un natural con esta propiedad menor que todos los demás.

Si este fuera el caso denotemos N el menor natural con dos representaciones en base b :

$$c_l b^k + \dots + c_1 b + c_0 = M = d_k b^k + \dots + d_1 b + d_0 \quad (4.54)$$

con $d_0, d_1, d_2, \dots, d_k, c_0, c_1, c_2, \dots, c_l \in \{0, 1, \dots, b-1\}$.

Como ya vimos, c_0 y (también d_0) es el residuo al dividir M entre b . Por lo tanto, $c_0 = d_0$. Por otra parte $\frac{M - c_0}{b} < M$, así que $\frac{M - c_0}{b}$ tiene expresión única en base b (esto es así por la manera en que escogimos M). Entonces

$$c_l b^{l-1} + \dots + c_1 = \frac{M - c_0}{b} = d_k b^{k-1} + \dots + d_1.$$

es la misma expresión. Por lo tanto

$$l = k \text{ y } c_1 = d_1, c_2 = d_2, \dots, c_l = d_k.$$

Esto, junto con $c_0 = d_0$ muestra que 4.54 no son dos representaciones distintas sino que son la misma. ■

Uno puede hacer operaciones aritméticas con los algoritmos usuales en otras bases. Claro que si va uno a usar bases distintas de 10, convendría hacer primero las tablas de sumar y de multiplicar correspondientes a las que uno tuvo que aprender en el segundo año de primaria.

Por ejemplo, he aquí las tablas de sumar y multiplicar en base 7:

+siete	0	1	2	3	4	5	6
0	0	1	2	3	4	5	6
1	1	2	3	4	5	6	10
2	2	3	4	5	6	10	11
3	3	4	5	6	10	11	12
4	4	5	6	10	11	12	13
5	5	6	10	11	12	13	14
6	6	10	11	12	13	14	15

\cdot siete	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6
2	0	2	4	6	11	13	15
3	0	3	6	12	15	21	24
4	0	4	11	15	22	26	31
5	0	5	13	21	26	34	42
6	0	6	15	24	31	42	51

Por ejemplo, podemos hacer

$$\begin{array}{r}
 & 2 & 3 & 5 \\
 & \times & 3 & 7 \\
 \hline
 & 1 & 6 & 4 & 5 \\
 & + & 7 & 0 & 5 \\
 \hline
 & 8 & 6 & 9 & 5
 \end{array}$$

en base siete:

$$\begin{array}{r}
 235 \\
 33 \left| \begin{array}{r} 4 \\ 5 \\ 4 \\ 0 \end{array} \right. , \text{ así } 235 = 454_{\text{siete}}. \\
 \hline
 37 \\
 5 \left| \begin{array}{r} 2 \\ 5 \\ 0 \end{array} \right. , \text{ así } 37 = 52_{\text{siete}}
 \end{array}$$

Ahora,

$$\begin{array}{r}
 & 4 & 5 & 4 & \text{siete} \\
 & \times & 5 & 2 & \text{siete} \\
 \hline
 & 1 & 2 & 4 & 1 \\
 & + & 3 & 2 & 6 & 6 \\
 \hline
 & 3 & 4 & 2 & 3 & 1
 \end{array}$$

En efecto, $34231_{\text{siete}} = 3 \cdot 7^4 + 4 \cdot 7^3 + 2 \cdot 7^2 + 3 \cdot 7 + 1 = 8695$.

También podemos hacer divisiones:

$$\begin{array}{r}
 1332 \\
 15 \overline{)19990} \\
 4990 \\
 490 \\
 40 \\
 10
 \end{array}$$

pasemos el numerador y el divisor a base 7:

$$\begin{array}{r}
 19990 \quad 5 \\
 2855 \quad 6 \\
 407 \quad 1 \quad 15 \\
 58 \quad 2 \quad 2 \quad 1 \\
 8 \quad 1 \quad 0 \quad 2 \\
 1 \quad 1 \\
 0
 \end{array}$$

Hagamos ahora la división teniendo a la vista las tablas de sumar y multiplicar en base 7:

$$\begin{array}{r}
 3 \quad 6 \quad 1 \quad 2 \\
 2 \quad 1 \quad | \quad 1 \quad 1 \quad 2 \quad 1 \quad 6 \quad 5 \\
 - \quad 6 \quad 3 \\
 1 \quad 6 \quad 1 \\
 - \quad 1 \quad 5 \quad 6 \\
 2 \quad 6 \\
 - \quad 2 \quad 1 \\
 5 \quad 5 \\
 - \quad 4 \quad 2 \\
 1 \quad 3
 \end{array}$$

así que el residuo de la división es $13_{siete} = 7 + 3 = 10_{diez}$. El cociente es $3612_{siete} = 3 \cdot 7^3 + 6 \cdot 7^2 + 1 \cdot 7 + 2 = 1332_{diez}$.

4.13.1 Algunos criterios de divisibilidad

Notemos que como $10 \stackrel{9}{\equiv} 1$, entonces $10^k \stackrel{9}{\equiv} 1^k = 1$, para cada $k \in \mathbb{N}$. Entonces un número que en sistema decimal se escriba

$$d_k \cdots d_1 d_0$$

es congruente con

$$d_k 1^k + \cdots + d_1 1 + d_0$$

módulo 9.

Así, tenemos la siguiente observación:

Observación 71

$$d_k \cdots d_1 d_0 \stackrel{9}{\equiv} d_k + \cdots + d_1 + d_0.$$

Ejemplo 114

$$7480287475 \stackrel{9}{\equiv} 7 + 4 + 8 + 0 + 2 + 8 + 7 + 4 + 7 + 5 \stackrel{9}{\equiv} 52 \stackrel{9}{\equiv} 7.$$

En efecto:

$$\begin{array}{r} 83114352 \\ 9 \overline{)7480287475} \\ 28 \\ \hline 10 \\ 12 \\ \hline 38 \\ 27 \\ \hline 47 \\ 25 \\ \hline 7 \end{array}$$

La observación anterior es la justificación del método para comprobar las operaciones, que a muchos de nosotros nos enseñaron en la primaria. En realidad este método consistía únicamente en realizar las operaciones módulo 9, aprovechando la sencillez del algoritmo para encontrar el residuo módulo 9 de un número que está escrito en base 10. He aquí algunos ejemplos:

$$\begin{array}{r} 128 \quad 2 \\ \times \quad 35 \quad \times 8 \\ \hline 640, \quad 16 \stackrel{9}{\equiv} 7 \\ 3840 \\ \hline 4480 \quad 4480 \stackrel{9}{\equiv} 16 \stackrel{9}{\equiv} 7 \end{array}$$

Así que podemos estar seguros de que la multiplicación está bien hecha, “módulo 9”.

$$\begin{array}{r} 3840 \quad 6 \\ + 4480 \quad 7 \\ \hline 640, \quad + \quad 1 \\ 353 \quad \quad \quad 2 \\ \hline 9313 \quad 16 \stackrel{9}{\equiv} 7 \\ \boxed{9313 \stackrel{9}{\equiv} 7} \end{array}$$

Ahora notemos que $10 \stackrel{11}{\equiv} -1$, entonces $10^{2n} \stackrel{11}{\equiv} (-1)^{2n} \stackrel{11}{\equiv} 1$ pero $10^{2n+1} \stackrel{11}{\equiv} (-1)^{2n+1} \stackrel{11}{\equiv} -1$.

Observación 72

$$d_k \cdots d_1 d_{0_{diez}} \stackrel{11}{\equiv} (d_0 + d_2 + d_4 + \dots) - (d_1 + d_3 + d_5 + \dots)$$

Ejemplo 115

$$7480287475 \stackrel{11}{\equiv} (5 + 4 + 8 + 0 + 4) - (7 + 7 + 2 + 8 + 7) = -10 \stackrel{11}{\equiv} 1.$$

En efecto:

$$\begin{array}{r} 68002614 \\ 11 \overline{)7480287475} \\ 88 \\ \hline 028 \\ 67 \\ \hline 14 \\ 37 \\ \hline 45 \\ 1 \end{array}$$

Ejercicio 201 ¿Cuál es el dígito x que hace que $748x287475$ deje residuo

1. 3 al dividirse entre 9?
2. 7 al dividirse entre 11?
3. 72 al dividirse entre 99? ¿o no hay?

Podemos encontrar criterios de divisibilidad parecidos si los números están escritos en bases distintas de 10. Veamos algunos ejemplos.

Ejemplo 116 . $8 \stackrel{7}{\equiv} 1$ por lo que si un número escrito en base 8 es $a_n a_{n-1} \cdots a_2 a_1$ entonces

$$\begin{aligned} a_n a_{n-1} \cdots a_2 a_1 a_{0_{ochos}} &= a_n \cdot 8^n + a_{n-1} \cdot 8^{n-1} + \cdots + a_1 \cdot 8^1 + a_0 \cdot 8^0 \stackrel{7}{\equiv} \\ &\stackrel{7}{\equiv} a_n + a_{n-1} + \cdots + a_2 + a_1 + a_0. \end{aligned}$$

Por ejemplo $1234567_{ochos} \stackrel{7}{\equiv} 1 + 2 + 3 + 4 + 5 + 6 + 7 = 28_{diez} = 34_{ochos} \stackrel{7}{\equiv} 7 \stackrel{7}{\equiv} 0$.

$$1 * 8^6 + 2 * 8^5 + 3 * 8^4 + 4 * 8^3 + 5 * 8^2 + 6 * 8 + 7 = 342\,391.$$

Y en efecto, $342\,391 = 7 * 48\,913 \stackrel{7}{\equiv} 0$.

Ejercicio 202 . Encuentre el residuo de 2347_{diez} al dividirse entre 7, pasando a base 8 y usando el ejemplo anterior.

Ejercicio 203 . Demuestre que como $8 \stackrel{9}{\equiv} -1$, entonces

$$a_n \cdots a_1 a_0 \stackrel{9}{\equiv} (a_0 + a_2 + \cdots) - (a_1 + a_3 + \cdots).$$

Ejercicio 204 . Encuentre el residuo de 1234567_{ocho} al dividirse entre 9, usando el ejercicio anterior.

Ejercicio 205 . Dado que $10^4 \stackrel{73}{\equiv} -1$, hallar un criterio de divisibilidad entre 73.

Ejercicio 206 . Hallar un criterio de divisibilidad por 13 en base 1000.

Ejercicio 207 . Enunciar criterios de divisibilidad entre 14, 18, 19 y 21.

4.14 Los números racionales

En esta sección daremos una construcción del campo de los números racionales. Definiremos un número racional como una clase de equivalencia de una función aditiva y suprayectiva $n \mathbb{Z} \xrightarrow{f_{m,n}} m \mathbb{Z}$ (la función que manda nz a mz , $n \neq 0$), entre dos ideales de \mathbb{Z} . La clase de equivalencia $n \mathbb{Z} \xrightarrow{f_{m,n}} m \mathbb{Z}$ se denotará $\frac{m}{n}$.

Con esta definición, la suma de dos números racionales se define como una suma de funciones (eliendo adecuadamente los representantes de las clases) y el producto se define por medio de la composición de funciones (también eliendo de manera adecuada los representantes).

Veremos que esta construcción coincide con la usual, que consiste en tomar clases de equivalencia en $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$, donde la relación está dada por $(a, b) \sim (m, n)$ si $an = bm$.

El punto de partida de la construcción presentada aquí es observar que $\mathbb{Z} \xrightarrow{f_{n,1}} n \mathbb{Z}$, (multiplicar por n) está relacionada con el número n .

$$z \mapsto zn$$

Sin embargo, $\begin{array}{ccc} 2\mathbb{Z} & \xrightarrow{f_{2n,2}} & 2n\mathbb{Z} \\ 2z & \longmapsto & 2zn \end{array}$ también es multiplicar por n y de hecho

$$\begin{array}{ccc} \mathbb{Z} & \xrightarrow{f_{n,1}} & n\mathbb{Z} \\ \uparrow \text{inclusión} & & \uparrow \text{inclusión} \\ 2\mathbb{Z} & \xrightarrow{f_{2n,2}} & 2n\mathbb{Z} \end{array}$$

commuta.

Definición 70 . $n\mathbb{Z} \xrightarrow{f_{m,n}} m\mathbb{Z}$, $n \neq 0$, es la función que manda nz a mz , $z \in \mathbb{Z}$.

Proposición 19 . $f_{m,n}$ está determinada por la propiedad de ser aditiva,

$$f(x+y) = f(x) + f(y),$$

y por la de enviar m a n .

Demostración. Demostraremos por inducción, que $f_{m,n}(nz) = mz$ si $z \geq 1$.

Base: $n1 = n \mapsto m = m1$.

Paso inductivo: $n(k+1) = nk + n \mapsto f_{m,n}(nk) + f_{m,n}(n) = mk + m = m(k+1)$. Con esto tenemos que $f_{m,n}(nz) = mz$, para toda $z \geq 1$.

Ahora, como $f_{m,n}$ es aditiva, entonces $f_{m,n}(0) = f_{m,n}(0+0) = f_{m,n}(0) + f_{m,n}(0)$, de donde tenemos que $f_{m,n}(0) = 0$.

Además,

$$0 = f_{m,n}(0) = f_{m,n}(nz - nz) = f_{m,n}(nz + (-nz)) = f_{m,n}(nz) + f_{m,n}(n(-z)),$$

de donde tenemos que $f_{m,n}(n(-z)) = -f_{m,n}(nz) = -(mz) = m(-z)$, si $z < 0$.

En resumen, $f_{m,n}(nz) = mz$, para toda $z \in \mathbb{Z}$. ■

Consideremos ahora el conjunto $\{n\mathbb{Z} \xrightarrow{f_{m,n}} m\mathbb{Z} \mid (m,n) \in \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})\}$, en el que definiremos la relación \sim de la manera siguiente.

Definición 71 . $f_{m,n} \sim f_{r,s}$ si $f_{m,n}, f_{r,s}$ coinciden en la intersección de sus dominios, es decir si $(f_{m,n})_{|n\mathbb{Z} \cap s\mathbb{Z}} = (f_{r,s})_{|n\mathbb{Z} \cap s\mathbb{Z}}$.

Lema 16 . $f_{m,n} \sim f_{r,s}$ si y sólo si $ms = nr$.

Demostración. Como $n \mathbb{Z} \xrightarrow{f_{m,n}} m \mathbb{Z}$ y $s \mathbb{Z} \xrightarrow{f_{r,s}} r \mathbb{Z}$, entonces la intersección de los dominios es $n \mathbb{Z} \cap s \mathbb{Z} = [n; s] \mathbb{Z}$ (ver página 218).

\Rightarrow) Si $f_{m,n} \sim f_{r,s}$ tenemos que $f_{m,n}(ns) = f_{r,s}(ns)$, ya que $ns \in n \mathbb{Z} \cap s \mathbb{Z}$.

Entonces $ms = f_{m,n}(ns) = f_{r,s}(ns) = nr$, por las definiciones de $f_{m,n}$ y $f_{r,s}$.

\Leftarrow) Supongamos ahora que $ms = nr$. Para ver que $f_{m,n}$ y $f_{r,s}$ coinciden en $n \mathbb{Z} \cap s \mathbb{Z}$, basta ver que coinciden en $[n; s]$.

Como $\frac{n}{(n; s)}s = [n; s]$ (ver el teorema 58, en la página 221) tenemos que

$$f_{m,n}([n; s]) = f_{m,n}\left(\frac{n}{(n; s)}s\right) = f_{m,n}\left(n\frac{s}{(n; s)}\right) = m\frac{s}{(n; s)} \text{ y}$$

$$f_{r,s}([n; s]) = f_{r,s}\left(\frac{n}{(n; s)}s\right) = \frac{n}{(n; s)}r.$$

Basta notar que $m\frac{s}{(n; s)} = \frac{n}{(n; s)}r$, ya que $ms = nr$. ■

Teorema 78 . La relación \sim es de equivalencia en

$$\left\{ n \mathbb{Z} \xrightarrow{f_{m,n}} m \mathbb{Z} \mid (m, n) \in \mathbb{Z} \times (\mathbb{Z} \setminus \{0\}) \right\}.$$

Demostración. $f_{m,n} \sim f_{m,n}$ ya que $mn = nm$.

$f_{m,n} \sim f_{r,s} \Leftrightarrow ms = nr$. Por otra parte, $f_{r,s} \sim f_{m,n} \Leftrightarrow rn = sm$. De donde se ve que $f_{m,n} \sim f_{r,s} \Leftrightarrow f_{r,s} \sim f_{m,n}$.

Si $f_{m,n} \sim f_{r,s} \sim f_{y,z}$ entonces $ms = nr$ y $rz = sy$. Multiplicando la primera ecuación por z tenemos que $msz = nrz$. En nrz podemos sustituir rz por sy , para obtener $msz = nsy$, como $s \neq 0$, podemos cancelar s para obtener $mz = ny$, es decir que $f_{m,n} \sim f_{y,z}$. ■

En vista del teorema anterior $\left\{ n \mathbb{Z} \xrightarrow{f_{m,n}} m \mathbb{Z} \mid (m, n) \in \mathbb{Z} \times (\mathbb{Z} \setminus \{0\}) \right\}$, queda partido en clases de equivalencia.

Definición 72 . $\mathbb{Q} = \frac{\left\{ n \mathbb{Z} \xrightarrow{f_{m,n}} m \mathbb{Z} \mid (m, n) \in \mathbb{Z} \times (\mathbb{Z} \setminus \{0\}) \right\}}{\sim}$.

Si denotamos por $\frac{a}{b}$ a la clase de equivalencia de $f_{a,b} : b\mathbb{Z} \longrightarrow a\mathbb{Z}$. (Notar que $b \neq 0$, por definición), podemos escribir:

$$\mathbb{Q} = \left\{ \frac{a}{b} \mid a \in \mathbb{Z}, b \in \mathbb{Z} \setminus \{0\} \right\}.$$

Y además se tiene que $\frac{a}{b} = \frac{c}{d}$ si y sólo si $ad = bc$.

4.14.1 La suma en \mathbb{Q}

Queremos definir ahora las operaciones en \mathbb{Q} .

Para sumar $\frac{a}{b}, \frac{c}{d}$ notemos que las funciones $f_{a,b}$ y $f_{c,d}$ no comparten dominios. La intersección de los dominios es $b\mathbb{Z} \cap d\mathbb{Z} = [b;d]$, en donde ambas funciones están definidas. Como $[b;d] = \frac{bd}{(b;d)}$, tenemos que $f_{a,b}\left(\frac{bd}{(b;d)}\right) = \frac{ad}{(b;d)}$ mientras que $f_{c,d}\left(\frac{bd}{(b;d)}\right) = \frac{bc}{(b;d)}$, la suma es $\frac{ad}{(b;d)} + \frac{bc}{(b;d)} = a\frac{d}{(b;d)} + c\frac{b}{(b;d)}$.

Esto sugiere definir $\frac{a}{b} + \frac{c}{d}$ como la clase de equivalencia de la función tal que

$$[b;d] \mapsto a\frac{d}{(b;d)} + c\frac{b}{(b;d)},$$

o mejor, multiplicando por $(b;d)$, con la clase de equivalencia de la función

$$bd \mapsto ad + cb.$$

Esta función es $f_{bd,ad+bc}$, cuya clase de equivalencia es $\frac{ad+bc}{bd}$.

Definición 73. La suma en \mathbb{Q} está dada por: $\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}$.

Proposición 20. La definición anterior es buena, ya que si $\frac{a}{b} = \frac{\alpha}{\beta}$ y $\frac{c}{d} = \frac{\lambda}{\mu}$ entonces $\frac{ad+bc}{bd} = \frac{\alpha\mu+\beta\lambda}{\beta\mu}$.

Demostración. Por hipótesis, $a\beta = b\alpha$ y $c\mu = d\lambda$.

Ahora $(ad+bc)\beta\mu = ad\beta\mu+bc\beta\mu = (a\beta)d\mu+b\beta(c\mu)$, usando las hipótesis tenemos que $(ad+bc)\beta\mu = (b\alpha)d\mu+b\beta(d\lambda) = bd(\alpha\mu+\beta\lambda)$. ■

Ejercicio 208 . Demuestre la suma definida en \mathbb{Q} es conmutativa, asociativa, con neutro $\frac{0}{1}$ y que $\frac{-a}{b}$ es el inverso aditivo de $\frac{a}{b}$.

4.14.2 El producto en \mathbb{Q}

Queremos definir el producto en \mathbb{Q} , mediante la composición de funciones, para esto debemos escoger representantes adecuados.

$\frac{a}{b}$ es la clase de equivalencia de $b\mathbb{Z} \xrightarrow{f_{a,b}} a\mathbb{Z}$, por otra parte $\frac{c}{d}$ es la clase de equivalencia de $d\mathbb{Z} \xrightarrow{f_{c,d}} c\mathbb{Z}$. Ahora, $\frac{a}{b} = \frac{ac}{bc}$ (si $c \neq 0$) y $\frac{c}{d} = \frac{cb}{db}$ y las funciones $f_{bc,ac}, f_{db,cb}$ se pueden componer:

$$\begin{array}{ccc} db\mathbb{Z} & \xrightarrow{f_{db,cb}} & bc\mathbb{Z} & \xrightarrow{f_{bc,ac}} & ac\mathbb{Z} \\ db & \longmapsto & cb & \longmapsto & ac\mathbb{Z} \end{array} = \begin{array}{ccc} db\mathbb{Z} & \xrightarrow{f_{db,ac}} & ac\mathbb{Z} \\ db & \longmapsto & ac\mathbb{Z} \end{array},$$

cuya clase de equivalencia es $\frac{ac}{db}$.

Esto sugiere definir $\frac{a}{b} \frac{c}{d} = \frac{ac}{db}$.

Definición 74 . $\frac{a}{b} \frac{c}{d} = \frac{ac}{db}$ en \mathbb{Q} .

Ejercicio 209 . Demuestre que la definición anterior es buena, es decir, que no depende de los representantes.

Observemos que el producto es conmutativo y asociativo. $\frac{1}{1}$ es neutro para el producto ($\frac{a}{b} \frac{1}{1} = \frac{a}{b1} = \frac{a}{b}$). Notemos que $\frac{a}{b} = \frac{0}{1}$ si y sólo si $a = 0$. Por último, notemos que si $\frac{a}{b} \neq \frac{0}{1}$ entonces $\frac{a}{b} \frac{b}{a} = \frac{ab}{ba} = \frac{1}{1}$, es decir que $\frac{b}{a}$ es el inverso multiplicativo de $\frac{a}{b}$.

Teorema 79 . $\frac{a}{b} \left(\frac{c}{d} + \frac{m}{n} \right) = \frac{a}{b} \frac{c}{d} + \frac{a}{b} \frac{m}{n}, \forall \frac{a}{b}, \frac{c}{d}, \frac{m}{n} \in \mathbb{Q}$.

Demostración.

$$\frac{a}{b} \left(\frac{c}{d} + \frac{m}{n} \right) = \frac{a}{b} \left(\frac{cn + md}{dn} \right) = \frac{acn + amd}{bdn}.$$

Por otra parte, $\frac{a}{b} \frac{c}{d} + \frac{a}{b} \frac{m}{n} = \frac{acbn + ambd}{b^2dn}$.

Ahora $\frac{acn + amd}{bdn} = \frac{acn + amd}{bdn} \frac{b}{b} = \frac{acbn + ambd}{b^2dn}$. ■

Las propiedades algebraicas de \mathbb{Q} se pueden resumir de la manera siguiente.

Teorema 80 . \mathbb{Q} es un campo.

4.14.3 El orden en \mathbb{Q}

Definimos el orden en \mathbb{Q} mediante una clase positiva.

Definición 75 . $\mathbb{Q}^+ = \left\{ \frac{a}{b} \in \mathbb{Q} \mid ab \in \mathbb{Z}^+ \right\}$.

Donde \mathbb{Z}^+ de nota la clase de los enteros positivos.

Observemos que \mathbb{Q}^+ está bien definida, es decir que si $\frac{a}{b} = \frac{\alpha}{\beta}$ y $ab \in \mathbb{Z}^+$, entonces $\alpha\beta \in \mathbb{Z}^+$.

En efecto, se tiene que $a\beta = b\alpha$ y $ab > 0$. Multipicando por ab , obtenemos

$$ab\alpha\beta = b^2\alpha^2$$

que es positivo por ser el producto de dos cuadrados distintos de cero. Así que $(ab)(\alpha\beta) > 0$, y como $ab > 0$ entonces $(\alpha\beta) > 0$. (Ver la Proposición 17, en la página 205).

Teorema 81 . \mathbb{Q}^+ es una clase positiva.

Demostración. Tenemos que demostrar que \mathbb{Q}^+ es cerrada bajo la suma, bajo el producto y la propiedad de tricotomía.

+) Supongamos que $\frac{a}{b}, \frac{c}{d} \in \mathbb{Q}^+$. $\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}$, y queremos ver

que $(ad + bc)bd \in \mathbb{Z}^+$.

$(ad + bc)bd = (ab)d^2 + b^2(cd)$, que es una suma de dos positivos, ya que ab y cd son positivos por hipótesis y b^2, d^2 son positivos porque son cuadrados no nulos.

*) Si $\frac{a}{b}, \frac{c}{d} \in \mathbb{Q}^+$, entonces su producto $\frac{ac}{bd}$ es positivo pues $(ac)(bd) = (ab)(cd)$ que es el producto de dos enteros positivos.

Tricotomía) Si $\frac{a}{b} \in \mathbb{Q}$, entonces pasa una y sólo una de las siguientes posibilidades:

i) $ab \in \mathbb{Z}^+$, ii) $-ab \in \mathbb{Z}^+$ iii) $ab = 0$ (y por lo tanto $a = 0$) cada una de estas posibilidades se corresponde respectivamente con:

$$\text{i) } \frac{a}{b} \in \mathbb{Q}^+, \text{ ii) } -\frac{a}{b} \in \mathbb{Q}^+, \text{ iii) } \frac{a}{b} = \frac{0}{1}. \blacksquare$$

Definición 76 . $\frac{a}{b} < \frac{c}{d}$ si $\frac{c}{d} - \frac{a}{b} \in \mathbb{Q}^+$, es decir si $(cb - da)db \in \mathbb{Z}^+$, es decir, si $abd^2 < cdb^2$.

Como siempre se puede escoger el denominador positivo ($\frac{a}{b} = \frac{-a}{-b}$), basta considerar que en ese caso, $\frac{a}{b} < \frac{c}{d}$ si y sólo si $ad < bc$.

Como de costumbre, $\frac{a}{b} < \frac{c}{d}$ y $\frac{c}{d} > \frac{a}{b}$ son proposiciones equivalentes.

$\frac{a}{b} \leq \frac{c}{d}$ es equivalente a $\left(\frac{a}{b} < \frac{c}{d}\right) \vee \left(\frac{a}{b} = \frac{c}{d}\right)$.

Ejemplo 117 . $\frac{1}{3} < \frac{1}{2}, \frac{2}{3} < \frac{3}{4}$.

Observe que $\frac{a}{b} \in \mathbb{Q}^+$ significa lo mismo que $\frac{a}{b} > \frac{0}{1}$.

Observación 73 . Si $\frac{a}{b} < \frac{c}{d}$ entonces $\frac{a}{b} < \frac{1}{2} \left(\frac{a}{b} + \frac{c}{d}\right) < \frac{c}{d}$.

Demostración. Podemos suponer $b, d > 0$.

$$\text{Ahora, } \frac{1}{2} \left(\frac{a}{b} + \frac{c}{d} \right) - \frac{a}{b} = \frac{ad + bc}{2bd} - \frac{a}{b} = \frac{ad + bc}{2bd} - \frac{a \cdot 2d}{b \cdot 2d} =$$

$$= \frac{bc - ad}{2bd} > \frac{0}{1} \text{ pues } bc - ad > 0.$$

$$\text{También tenemos que } \frac{c}{d} - \frac{1}{2} \left(\frac{a}{b} + \frac{c}{d} \right) = \frac{c}{d} - \frac{ad + bc}{2bd} = \frac{c \cdot 2b}{d \cdot 2b} -$$

$$\frac{ad + bc}{2bd} = \frac{bc - ad}{2bd}. \blacksquare$$

Como la diferencia entre $\frac{1}{2} \left(\frac{a}{b} + \frac{c}{d} \right)$ y $\frac{a}{b}$ es la misma que entre $\frac{c}{d}$ y $\frac{1}{2} \left(\frac{a}{b} + \frac{c}{d} \right)$, a $\frac{1}{2} \left(\frac{a}{b} + \frac{c}{d} \right)$ se le llama el promedio (o punto medio) entre $\frac{a}{b}$ y $\frac{c}{d}$.

Observación 74 . Note que una consecuencia de la definición del orden mediante una clase positiva es que cualesquiera dos racionales son comparables. Es decir el orden \leq en \mathbb{Q} es total.

En contraste con el orden de \mathbb{N} , el orden en \mathbb{Q} no es un buen orden, como se muestra en la siguiente observación.

Observación 75 . La sucesión $\frac{1}{2} > \frac{1}{4} > \frac{1}{8} > \dots > \frac{1}{2n} > \dots$ es una sucesión que no tiene primer elemento.

Teorema 82 . Para todo $\varepsilon \in \mathbb{Q}^+$, para todo $N \in \mathbb{N}$, existe $M \in \mathbb{N}$ tal que $\frac{M}{1} \varepsilon > \frac{N}{1}$.

Demostración. Hagamos $\varepsilon = \frac{\alpha}{\beta}$, con $\alpha, \beta \in \mathbb{Z}^+$. Queremos encontrar M tal que $M\alpha > N\beta$. Como $\alpha \geq 1$, entonces $M\alpha \geq M$. Por lo que basta mostrar que existe M tal que $M > N\beta$.

En caso contrario, $m \leq N\beta$, para cada $m \in \mathbb{N}$. Podemos tomar β el menor natural positivo tal que $m \leq N\beta$ (se hace uso del principio del buen orden). Como no sucede que $N+1 \leq N$, entonces $\beta \neq 1$. Por la elección de β , existe $n \in \mathbb{N}$ tal que $n \not\leq N(\beta-1) = N\beta - N$, de aquí que

$$n + N \not\leq N\beta,$$

contradicción. ■

Teorema 83 . $\forall N \in \mathbb{N}, \forall \varepsilon \in \mathbb{Q}^+, \exists M \in \mathbb{N}^+ \text{ tal que } \frac{1}{M}N < \varepsilon$.

Demostración. Se quiere mostrar que existe M tal que $M\varepsilon > N$. Esto se sigue del teorema anterior. ■

Lema 17 . $\forall \varepsilon \in \mathbb{Q}^+, \exists n \in \mathbb{N} \text{ tal que } \frac{1}{2^n} < \varepsilon$.

Demostración. Se quiere ver que hay un natural n tal que $2^n\varepsilon > 1$.

Se puede ver por inducción, que $2^n > n, \forall n \in \mathbb{N}$. Escogiendo M tal que $M\varepsilon > 1$ (ver el teorema 82), tenemos que $2^M\varepsilon > M\varepsilon > 1$. ■

Observemos que todo racional tiene racionales cercanos. Ya que si $\frac{a}{b} \in \mathbb{Q}$ y $\varepsilon \in \mathbb{Q}^+$, como existe $m \in \mathbb{N}$ tal que $\frac{1}{m} < \varepsilon$, entonces la diferencia entre $\frac{a}{b} + \frac{1}{m}$ y $\frac{a}{b}$ es menor que ε .

Ejercicio 210 . Demuestre que todo racional positivo se puede representar de manera única en la forma $\frac{m}{n}$, con $(m; n) = 1$ y m, n enteros positivos.

4.14.4 Inmersión de \mathbb{Z} en \mathbb{Q}

Teorema 84 . La función $\begin{array}{ccc} \mathbb{Z} & \xrightarrow{f} & \mathbb{Q} \\ z & \mapsto & \frac{z}{1} \end{array}$ es inyectiva, respeta las operaciones, los neutros y el orden.

Demostración. Inyectividad) $\frac{z}{1} = \frac{w}{1} \iff z1 = w1 \iff z = w$.

$+)$ $f(a+b) = \frac{a+b}{1}$. Por otra parte, $f(a) + f(b) = \frac{a}{1} + \frac{b}{1} = \frac{a1+b1}{1 \cdot 1} = \frac{a+b}{1}$.

$$\cdot) f(ab) = \frac{ab}{1} = \frac{a}{1} \frac{b}{1} = f(a)f(b).$$

Orden) Supongamos que $w < z$ son enteros, queremos ver que entonces $\frac{w}{1} < \frac{z}{1}$. En efecto, $\frac{z}{1} - \frac{w}{1} = \frac{z-w}{1}$, con $(z-w) \cdot 1 \in \mathbb{Z}^+$. ■

En vista del teorema anterior podemos identificar los enteros con su imagen en \mathbb{Q} . Con esta identificación (n se identifica con $\frac{n}{1}$), podemos pensar que cada entero es un racional.

Ejemplo 118 . No existe un racional $\frac{a}{b}$ tal que $\frac{a}{b} \frac{a}{b} = 2$.

En caso de que $\frac{a}{b} \frac{a}{b} = 2$ entonces $a^2 = 2b^2$. Si 2^α es la mayor potencia de 2 que divide a a entonces $2^{2\alpha}$ es la mayor potencia de 2 que divide a a^2 (es decir que 2α es el número de veces en que 2 aparece en la factorización en primos de a^2). Pero por otro lado, si 2^β es la mayor potencia de 2 que divide a b entonces $2^{2\beta}$ es la mayor potencia de 2 que divide a b^2 , y por lo tanto $2^{2\beta} = 2^{2\beta+1}$ es la mayor potencia de 2 que divide $2b^2 = a^2$. Entonces $2\alpha = 2\beta + 1$ y así $0 \stackrel{2}{\equiv} 1$, absurdo.

Ejercicio 211 . Demuestre que un entero no puede ser par e impar. (Sugerencia: demuestre que 1 no es par).

Ejercicio 212 . Demuestre que no hay un racional r tal que $r^2 = 3$.

Ejercicio 213 . Demuestre que no hay un racional r tal que $r^2 = 6$.

Ejercicio 214 . Demuestre que no hay un racional r tal que $r^3 = 2$.

Ejercicio 215 . Demuestre que no hay un racional r tal que $r^3 + r^2 = 5$. (Sugerencia: suponga que $r = \frac{a}{b}$ con $(a; b) = 1$).

Capítulo 5

¿De cuántas maneras?

En este capítulo estaremos interesados en asuntos acerca de las cardinalidad de los conjuntos finitos. Determinar la cardinalidad de un conjunto finito es lo mismo que “contar los elementos de dicho conjunto”.

Así que empezaremos por aclarar la siguiente pregunta:

¿Cuándo dos conjuntos A, B tienen el mismo número de elementos?

Definición 77 . *Se dice que A y B tiene el mismo número de elementos cuando hay una función biyectiva*

$$f : A \longrightarrow B.$$

Recordemos que una función biyectiva $f : A \longrightarrow B$ es

1. Inyectiva (manda dos elementos diferentes en A a dos elementos diferentes en B) y
2. Suprayectiva (cada elemento de B proviene de uno de A).

Si nos hicieramos la imagen de que f “reparte los elementos de A a los elementos de B ” diríamos que f es inyectiva cuando elementos distintos de A se reparten a elementos distintos de B (es decir, a un elemento de B no le pueden tocar dos elementos de A) y suprayectiva cuando todo elemento de B recibe su elemento de A .

Cuando en un salón de clases todos los alumnos están sentados en su silla, y no hay sillas desocupadas, podemos decir que hay el mismo número de alumnos que de sillas.

Recordemos que un conjunto X es un conjunto infinito cuando existe una correspondencia biyectiva entre X y uno de sus subconjuntos propios.

Por ejemplo,

$$\begin{array}{ccc} \mathbb{N} & \xrightarrow{2 \cdot} & 2\mathbb{N} \\ k & \longmapsto & 2k \end{array}$$

es una función biyectiva entre \mathbb{N} y uno de sus subconjuntos propios, así que \mathbb{N} es infinito.

Definición 78 . *Un conjunto F es finito si toda función suprayectiva $g : F \rightarrow F$ es biyectiva.*

Es una consecuencia del axioma de Elección el hecho de que toda función suprayectiva tiene inverso por la derecha (que es por lo tanto inyectiva). Toda función inyectiva tiene inverso por la derecha, así que podemos demostrar la siguiente proposición.

Proposición 21 . *Son equivalentes para un conjunto F :*

1. F es finito.
2. Toda función inyectiva $f : F \rightarrow F$ es una biyección.

Demostración. 1) \Rightarrow 2)

Supongamos que $f : F \rightarrow F$ es una función inyectiva. Tomemos g un inverso izquierdo de f . Como g tiene inverso derecho, entonces g es suprayectiva, entonces, por 1) tenemos que g es una biyección y como tal es invertible. Así que de la ecuación

$$g \circ f = Id_F = g \circ g^{-1}$$

tenemos, aplicando g^{-1} del lado izquierdo, que

$$f = Id_F \circ f = g^{-1} \circ g \circ f = g^{-1} \circ g \circ g^{-1} = Id_F \circ g^{-1} = g^{-1}.$$

Entonces tenemos que f es una biyección.

2) \Rightarrow 1) Tenemos que demostrar que toda función suprayectiva $g : F \rightarrow F$ es una biyección. Tómese $h : F \rightarrow F$ un inverso derecho de g . h es inyectiva puesto que tiene a g como inverso izquierdo. Por 2) tenemos que h es biyectiva. Como arriba, es fácil ver que $g = h^{-1}$ (Dedúzcase de $g \circ h = Id = h^{-1} \circ h$, aplicando a h^{-1}). Así, $g = h^{-1}$ es una biyección. ■

Recordemos la siguiente consecuencia del Teorema de Recursión:

Proposición 22 . *X es un conjunto infinito si y sólo si $\exists f : \mathbb{N} \rightarrow X$ inyectiva*

Demostración. \Leftarrow) Escribamos $X = f(\mathbb{N}) \cup (X \setminus (f(\mathbb{N})))$. Tomemos una biyección entre \mathbb{N} y un subconjunto propio de \mathbb{N} , digamos $\lambda : \mathbb{N} \rightarrow \mathbb{B}$, componiendo con la función inclusión $\mathbb{B} \hookrightarrow \mathbb{N}$, obtenemos $\mathbb{N} \xrightarrow{\lambda} \mathbb{B} \hookrightarrow \mathbb{N}$ una función inyectiva que no es suprayectiva.

Ahora, del diagrama

$$\begin{array}{ccccc}
 \mathbb{N} & \xrightarrow{\lambda} & \mathbb{B} & \hookrightarrow & \mathbb{N} \\
 \downarrow f & & \downarrow f_{|\mathbb{B}} & \neq & \downarrow f \\
 f(\mathbb{N}) & \xrightarrow{f_{|\mathbb{B}} \circ \lambda \circ f_{|f(\mathbb{N})}^{-1}} & f(\mathbb{B}) & \hookrightarrow & f(\mathbb{N}) \\
 & & \neq & &
 \end{array}$$

tenemos que

$$\gamma =: f(\mathbb{N}) \xrightarrow{f_{|\mathbb{B}} \circ \lambda \circ f_{|f(\mathbb{N})}^{-1}} f(\mathbb{B}) \xrightarrow{\neq} f(\mathbb{N})$$

es una función inyectiva que no es suprayectiva. Podemos usarla para definir una función inyectiva no suprayectiva entre X y X :

$$\begin{array}{ccc}
 \rho : & X & \rightarrow X \\
 & a & \longmapsto \begin{cases} \gamma(a) & \text{si } a \in f(\mathbb{N}) \\ a & \text{si } a \notin f(\mathbb{N}) \end{cases}
 \end{array}$$

ρ nos proporciona una biyección entre X y un subconjunto propio de X .

\Rightarrow) Si X es un conjunto infinito, entonces $\exists f : X \rightarrow X$ inyectiva pero no suprayectiva. Tomemos $a \in X \setminus f(X)$ y usemos el Teorema de recursión

$$\begin{array}{ccc}
 \mathbb{N} & \xrightarrow{\sigma} & \mathbb{N} \\
 \{0\} & \xrightarrow{\stackrel{\hat{0}}{\nearrow}} & \downarrow \mu \qquad \downarrow \mu \\
 & \xrightarrow{\hat{a}} & \\
 X & \xrightarrow{f} & X
 \end{array}$$

para deducir la existencia de una función μ que hace commutativo el diagrama anterior. Resta comprobar que μ es una función inyectiva.

Tomemos, si no lo fuera, $k \neq m \in \mathbb{N}$, tales que $\mu(k) = \mu(m)$, con k mínima posible.

Si $k = 0$ entonces $0 < m$, por lo que $m = \sigma(m - 1)$. Entonces

$$a = \mu(0) = \mu(m) = \mu(\sigma(m - 1)) = f(\mu(m - 1)) \stackrel{\nabla}{\circ}$$

Esta sería una contradicción a la elección de a , que no pertenece a $f(X)$.

Entonces $k > 0$, $m > k > 0$ y

$$\begin{aligned} \mu(k) &= \mu(m) = \mu(\sigma(m - 1)) = f(\mu(m - 1)) = \\ &= f(\mu(k - 1)) \end{aligned}$$

Como f es inyectiva entonces $\mu(k - 1) = \mu(m - 1) \stackrel{\nabla}{\circ}$. Se contradice la elección de k .

Por lo tanto μ es inyectiva, y hemos terminado ■

Como consecuencia de la afirmación anterior tenemos la siguiente observación, aparentemente trivial, pero útil..

Observación 76 . *Si F es un conjunto finito y $\mathbb{N} \xrightarrow{f} F$ es una función, entonces f no es inyectiva y por lo tanto*

$$\exists k \neq m \in \mathbb{N} \text{ tales que } f(k) = f(m).$$

Observación 77 . *Si $F \subseteq \mathbb{N}$ es finito y transitivo, entonces $F \in \mathbb{N}$.*

Demostración. Podemos suponer que $F \neq \emptyset = 0$. $k \in F \Rightarrow k \in \mathbb{N}$. Entonces $k = 0$ ó $0 \in k$.

En cualquier caso,

$$0 \in F$$

$(0 \in k \in F \Rightarrow 0 \in F$, pues F es transitivo).

Si $k \in F$, entonces $k \subseteq F$. Ahora, $\{k\} \in F \Rightarrow k + 1 \in F$.

Por el principio de inducción, y como F es finito, existe $k \in F$ tal que $k + 1 \notin F$ (si no fuera así, $F = \mathbb{N} \stackrel{\nabla}{\circ}$). Escogiendo k mínima con la propiedad anterior, tenemos que $\{0, \dots, k - 1\} \subseteq F$, pero $k \notin F$. Por transitividad, ningún elemento de \mathbb{N} , mayor o igual que k pertenece a F ($k < m \in F \Rightarrow k \in m \in F \Rightarrow k \in F \Rightarrow k + 1 \in F, \stackrel{\nabla}{\circ}$). Por lo tanto $F = \{0, \dots, k - 1\} = k \in \mathbb{N}$.

■

Observación 78 . *Si $F \subseteq \mathbb{N}$ es infinito y transitivo, entonces $F = \mathbb{N}$.*

Demostración. $F \neq \emptyset$, pues \emptyset es finito.

$$(k \in F \Rightarrow k + 1 \in F, \forall k \in F) \Rightarrow F = \mathbb{N}.$$

Si hubiera $k \in F$, tal que $k + 1 \notin F$, entonces por la transitividad de k , tendríamos que $k \subseteq F$. También por la transitividad de k tenemos que $m \notin F$, para cada $m \geq k + 1$. Por lo tanto $k + 1 = \{0, \dots, k\} = F^\nabla$. Pues todos los naturales son conjuntos finitos. ■

Lema 18 . Si $F \subseteq \mathbb{N}$ es finito entonces $\exists k \in \mathbb{N}$, y una biyección entre k y F .

Demostración. Si F no tiene elementos, entonces $F = \emptyset$ y $\emptyset \xrightarrow{Id_\emptyset} F$ es una biyección entre 0 y F .

Podemos suponer que $F \neq \emptyset$.

Por el principio del buen orden,

$$\exists a_0 = \min(F),$$

así tenemos una función inyectiva

$$\begin{array}{ccc} 1 = \{0\} & \xrightarrow{f_1} & F \\ 0 & \mapsto & a_0 \end{array}.$$

Si f_1 es suprayectiva, hemos encontrado una biyección entre 1 y F . En caso contrario, $F \setminus \{a_0\} \neq \emptyset$, es un subconjunto no vacío de naturales, así que, otra vez invocando el principio del buen orden,

$$\exists a_1 = \min(F \setminus \{a_0\}),$$

con lo que tenemos una función inyectiva

$$\begin{array}{ccc} 2 = \{0, 1\} & \xrightarrow{f_2} & F \\ 0 & \mapsto & a_0 \\ 1 & \mapsto & a_1 \end{array}.$$

Y nuevamente, f_2 es biyectiva o $F \setminus \{a_0, a_1\} \neq \emptyset$.

Podemos continua este proceso.

Si termina, es porque encontramos una biyección

$$k = \{0, \dots, k-1\} \xrightarrow{f_k} F.$$

Si no termina, tenemos una función

$$\begin{array}{ccc} \mathbb{N} & \xrightarrow{T} & F \\ j & \mapsto & f_k(j) \quad , k \geq j+1 \end{array}$$

T está bien definida pues en la sucesión

$$f_0, f_1, \dots, f_j, \dots$$

cada función extiende a la anterior.

$$\begin{array}{ccc} j & \mapsto & f_l(j) \\ l & \xrightarrow{f_l} & F \\ \uparrow & \uparrow & \parallel \\ k & \xrightarrow{f_k} & F \\ j & \mapsto & f_k(j) \end{array}$$

T es inyectiva, pues si $j < k$, tomemos $m > k$. Entonces $T(j) = f_m(j) \neq f_m(k) = T(k)$.

En resumen, o bien hay una $k \in \mathbb{N}$ tal que $k \xrightarrow{f_k} F$ es una biyección, o $\exists T : \mathbb{N} \rightarrow F$, en cuyo caso F sería infinito, lo que no sucede. ■

¿Qué quiere decir “contar”?

Contar un conjunto finito F quiere decir establecer una correspondencia biyectiva entre F y algún número natural k . Para esto haremos uso de la siguiente observación.

Observación 79. *Si F es un conjunto finito entonces*

$$\exists k \in \mathbb{N} \text{ y } f : k \longrightarrow F$$

biyectiva.

Demostración. Supongamos lo contrario.

Primero veremos que existen funciones inyectivas

$$f_k : k \rightarrowtail F,$$

para cada natural k .

Notemos que $F \neq \emptyset$, ya que en otro caso $F = 0$.

Tomando un elemento $a_0 \in F$, definimos la inyección

$$\begin{array}{ccc} \{0\} & \xrightarrow{f_1} & F \\ 0 & \mapsto & a_0 \end{array}.$$

Supongamos que tenemos una inyección

$$\{0, 1, \dots, k-1\} \xrightarrow{f_k} F,$$

como esta función no es suprayectiva, entonces $\exists b \in F \setminus f_k(\{0, 1, \dots, k-1\})$.

Podemos definir ahora una función inyectiva que extienda la anterior:

$$\begin{array}{ccc} \{0, 1, \dots, k-1, k\} & \xrightarrow{f_{k+1}} & F \\ j & \mapsto & f_k, \text{ si } j \leq k-1 \\ k & \mapsto & b \end{array}.$$

Ahora es fácil ver, como en la demostración anterior, que

$$\begin{array}{ccc} \mathbb{N} & \xrightarrow{T} & F \\ j & \mapsto & f_k(j), \text{ si } j \geq k+1 \end{array},$$

es una función inyectiva.

Entonces F es infinito ∇ . La contradicción viene de suponer que no hay una biyección entre F y un natural. ■

Así, contar los elementos de un conjunto finito F es determinar el número natural que está en correspondencia biyectiva con F .

Las páginas de un periódico normalmente están numeradas $1, 2, 3, \dots, k$.

El conjunto de vértices de un triángulo está en correspondencia biyectiva con

$$3 = \{0, 1, 2\} :$$

5.1 ¿Cuántos subconjuntos tiene un conjunto con n elementos?

Tomemos

$$A = \{a_1, a_2, \dots, a_n\}.$$

Podemos notar que no es necesario cargar con la letra a y podemos tomar el conjunto

$$n = \{0, 1, 2, \dots, n - 1\}$$

que también tiene n elementos.

Esto se debe a lo siguiente:

Proposición 23 . *Si dos conjuntos tienen el mismo número de elementos, entonces también tienen el mismo número de subconjuntos.*

Demostración. Supongamos que $A \xrightarrow{f} B$ es una biyección entre A y B . Denotemos, como de costumbre por $\wp(A)$ el conjunto de los subconjuntos de A . Daremos una biyección entre $\wp(A)$ y $\wp(B)$:

$$\begin{array}{ccc} \wp(A) & \xrightarrow{f^*} & \wp(B) \\ X & \longmapsto & f(X). \end{array}$$

f^* es inyectiva:

Si $X \neq Y$ son subconjuntos de A , entonces hay un elemento z , tal que

$$z \in (X \cup Y) \setminus (X \cap Y),$$

por lo tanto

$$f(z) \notin (f(X) \cap f(Y))$$

($(f(z) = f(x) = f(y)) \Rightarrow (x = y = z)$, porque f es inyectiva). Pero $f(z) \in f(X) \cup f(Y)$.

Por lo tanto

$$f(z) \in (f(X) \cup f(Y)) \setminus (f(X) \cap f(Y)),$$

Por lo tanto $f(X) \neq f(Y)$.¹

f^* es suprayectiva:

Si $W \subseteq B$, entonces $W = f(f^{-1}(W))$, ya que f es una función suprayectiva.

Por lo tanto,

$$|\wp(A)| = |\wp(B)|.$$

¹Recuérdese, del capítulo de conjuntos, que la diferencia simétrica de un conjunto consigo mismo es el conjunto vacío.

■ Así que al preguntarnos cuántos subconjuntos tiene un conjunto con n elementos, podemos contar en cualquier conjunto con n elementos, por ejemplo,

$$n = \{0, 1, \dots, n - 1\}$$

Consideremos ahora las n -adas ordenadas de 0 y 1 tal como

$$(0, 1, 0, 1, 0, 0, \dots, 0),$$

¿cuántas hay?

Para la primer coordenada tenemos dos opciones: 0 y 1.

Para la segunda coordenada también tenemos dos opciones: 0 y 1, así que las dos primeras coordenadas pueden ser

$$\left\{ \begin{array}{l} 0, 0, \dots \\ 0, 1, \dots \\ 1, 0, \dots \\ 1, 1, \dots \end{array} \right.$$

$4 = 2 \times 2$ posibilidades para las dos primeras coordenadas.

Para cada una de las posibilidades anteriores, tenemos dos posibilidades para escoger la tercera coordenada, así que tendremos 2^3 posibilidades para las tres primeras coordenadas.

Repetiendo el argumento, tenemos 2^n n -adas de 0 y 1.

Escribamos el conjunto de las n -adas de 0 y 1 como \mathbb{Z}_2^n . Veremos que tantos como subconjuntos de

$$\{0, 1, \dots, n - 1\}.$$

Sea

$$g : \begin{array}{ccc} \mathbb{Z}_2^n & \rightarrow & \wp(n) \\ (s_0, s_1, \dots, s_{n-1}) & \longmapsto & X_s = \{k \mid s_k = 1\} \end{array}$$

Ejemplo 119 . Si $n = 4$, entonces

$$g : \begin{array}{ccc} \mathbb{Z}_2^4 & \rightarrow & \wp(4) \\ (0, 1, 1, 0) = \\ (s_0, s_1, s_2, s_3) & \longmapsto & \{1, 2\} \\ (0, 1, 0, 1) & \longmapsto & \{1, 3\} \\ (0, 0, 0, 0) & \longmapsto & \emptyset \\ (1, 1, 1, 1) & \longmapsto & \{0, 1, 2, 3\} \end{array}.$$

Proposición 24 . $g : \mathbb{Z}_2^n \rightarrow \wp(n)$ es una función biyectiva. ²

Demostración. Inyectividad) Supongamos que

$$s = (s_0, s_1, \dots, s_{n-1}) \neq t = (t_0, t_1, \dots, t_{n-1}),$$

entonces $s_i \neq t_i$, para alguna i . Por ejemplo, supongamos que $s_i = 1$ y $t_i = 0$, entonces $i \in g(s)$ pero $i \notin g(t)$, por lo que $g(s) \neq g(t)$.

Por lo tanto g es una función inyectiva.

Suprayectividad) Tomemos ahora un subconjunto X de n , consideremos la función

$$\chi_X : n \rightarrow 2 \text{ tal que } \chi_X(j) = \begin{cases} 1 & \text{si } j \in X \\ 0 & \text{si } j \notin X \end{cases}$$

(χ_X se llama la función característica de X .) Es claro que $X = g(\chi_X)$, pues

$$i \in X \iff \chi_X(i) = 1 \iff i \in g(\chi_X).$$

Por lo tanto g es suprayectiva. ■

Por lo tanto

$$|\mathbb{Z}_2^n| = |\wp(n)|$$

Así que

Teorema 85

$$|\wp(n)| = |\mathbb{Z}_2^n| = 2^n. \quad (5.1)$$

Reflexionemos un poco sobre lo anterior:

\mathbb{Z}_2^n es el conjunto de n -adas de ceros y unos, que es lo mismo que el número de funciones entre $n = \{0, 1, \dots, n-1\}$ y $2 = \{0, 1\}$. Si denotamos el conjunto de las funciones de n a 2 por $\{0, 1\}^n$, tendremos que

$$|\{0, 1\}^n| = 2^n!$$

Denotemos ahora

$$B^A = \{A \xrightarrow{f} B \mid f \text{ es una función}\}.$$

Podemos preguntarnos, igual que arriba, si vale la ecuación siguiente:

$$\mathcal{Z}|B^A| = |B|^{|A|}?$$

La respuesta es sí, como veremos enseguida,

² $(k \in g(s) \iff s(k) = 1)$.

Teorema 86 . Si $|B| = n$ y $|A| = m$, entonces $|B^A| = |B|^{|A|} = n^m$.

Demostración. Por inducción sobre m .

Base.

Si $m = 0$ entonces A es el conjunto vacío y así,

$$B^A = \{A \xrightarrow{f} B \mid f \text{ es función}\} = \{\emptyset \xrightarrow{\emptyset} B\}$$

es un conjunto con un único elemento.

Por lo tanto

$$|B^\emptyset| = |\{\emptyset \xrightarrow{\emptyset} B\}| = 1 = n^0.$$

Así que la base de la inducción es válida.

Paso inductivo.

Supongamos que la afirmación vale si A tiene m elementos.

Sea ahora A' con $m + 1$ elementos. Un conjunto así se puede formar al agregar a un conjunto con m elementos un elemento nuevo.

Supongamos que

$$A' = A \cup \{x\}.$$

Cada función $A \xrightarrow{f} B$ da lugar a n funciones diferentes de A' a B , que extienden f , una función por cada una de las n maneras de escoger la imagen de x .

Como hay n^m funciones de A a B ³, entonces hay $n^m \cdot n$ funciones de A' a B .

Así que

$$|B^{A \cup \{x\}}| = n^m \cdot n = n^{m+1} = |B|^{|A \cup \{x\}|}.$$

■

5.1.1 El principio de la pichoneras

Este principio no es otra cosa que una manera de decir que si existe

$$A \xrightarrow{f} B$$

³(por hipótesis de inducción)

inyectiva, entonces

$$|A| \leq |B|.$$

Esto resulta muy natural, pues si a cada elemento de A se le puede asignar un elemento de B , sin repetir elementos, es claro que B debe tener por lo menos tantos elementos como A .

En otras palabras,

$$\left(\exists A \xrightarrow{f} B, \text{ inyectiva} \right) \Rightarrow (|A| \leq |B|).$$

Si tomamos la contrapuesta de esta proposición (y tomando en cuenta que una consecuencia del axioma de Elección es que cualesquiera dos cardinales son comparables), tenemos que

$$(|A| > |B|) \Rightarrow \left(A \xrightarrow{f} B \text{ no es inyectiva, } \forall f \in B^A \right).$$

Es decir, si A tiene más elementos que B , entonces una función $A \xrightarrow{f} B$ le asigna el mismo elemento de B a dos elementos de A .

Hay frases que resumen esta situación:

“Si hay más pichones que pichoneras, habrá una pichonera con más de un pichón (si los pichones están dentro de sus pichoneras)”.

“Si se reparten 20 monedas entre 15 personas, habrá alguna persona a la que se le dá más de una moneda”.

5.2 Subconjuntos con k elementos de un conjunto con n elementos

El número de subconjuntos con k elementos de un conjunto con n elementos se denota $\binom{n}{k}$ o por C_n^k en algunos textos.

Notemos los siguientes hechos evidentes:

Observación 80

1. Si $k > n$, entonces $\binom{n}{k} = 0$.

2. Si $k = n$, entonces $\binom{n}{k} = 1$.

3. Si $k = 0$, entonces $\binom{n}{k} = 1$.

4. Si $k = 1$, entonces $\binom{n}{k} = n$.

Notación 10 . Denotemos

$$\wp_k(B) = \{X \subseteq B \mid |X| = k\}.$$

Es decir, $\wp_k(B)$ es el conjunto de los subconjuntos de B que tienen k elementos.

Vamos a realizar la cuenta de $\wp_k(B)$, para un conjunto B , tal que $|B| = n$, de dos maneras distintas, cada una de las cuales nos dará información interesante.

Primera forma.

Como hemos notado arriba, basta analizar la situación en que $k < n$.

Queremos determinar $\binom{n}{k}$.

Escojamos un elemento x de B (como $n > k \geq 0$, entonces B no es vacío).

¿Cuántos subconjuntos de B con k elementos contienen a x ?

Es claro para formar un conjunto A con k elementos, uno de los cuales es x , basta tomar los $k - 1$ elementos de B que no son x .

¿De cuántas maneras se puede escoger un subconjunto con $k - 1$ elementos de los $n - 1$ elementos de B que no son x ?

Es claro que de

$$\binom{n-1}{k-1}$$

maneras.

Fijémonos ahora en un subconjunto A en particular, digamos que

$$A = \{x_1, \dots, x_k\}$$

Pensemos en las maneras de obtener A de la manera descrita en el párrafo de arriba:

A se obtiene empezando con $x = x_1$ y después agregando $\{x_2, \dots, x_k\}$;

A se obtiene empezando con $x = x_2$ y después agregando $\{x_1, x_3, \dots, x_k\}$;

A se obtiene empezando con $x = x_3$ y después agregando $\{x_1, x_2, x_4, \dots, x_k\}$;

\vdots

A se obtiene empezando con $x = x_k$ y después agregando $\{x_1, x_2, \dots, x_{k-1}\}$;

Así que la misma A se obtiene de k maneras distintas, según se empiece escogiendo x como x_1 ó bien como x_2, \dots , o bien, como x_k .

Así que si contamos los subconjuntos con k elementos, contando los subconjuntos con k elementos que contienen a x y haciendo esto para cada x de A , (hay n elementos en A), cada subconjunto con k elementos estará tomado en cuenta k veces, así que

$$\binom{n}{k} = \frac{n \binom{n-1}{k-1}}{k} = n/k \binom{n-1}{k-1}.$$

Estamos en condiciones de seguir aplicando la fórmula para obtener

$$\begin{aligned} \binom{n}{k} &= n/k \binom{n-1}{k-1} = (n/k)(n-1/k-1) \binom{n-2}{k-2} = \\ &= \frac{n \cdot (n-1) \cdots (n-(k-1))}{k \cdot (k-1) \cdots (k-(k-1))} \binom{n-k}{0} = \\ &= \frac{n \cdot (n-1) \cdots (n-(k-1))}{k!} (n-k) = \\ &= \frac{n!}{k!(n-k)!}. \end{aligned}$$

Desde luego, $n! = n \cdot (n-1) \cdot \dots \cdot 2 \cdot 1$.

En resumen, tenemos que

$$\boxed{\binom{n}{k} = \frac{n!}{k!(n-k)!}}$$

Segunda forma.

Escojamos un elemento x de B , vamos a partir $\wp_k(B)$ en dos conjuntos;

$$\{A \in \wp_k(B) \mid x \notin A\}$$

y en

$$\{A \in \wp_k(B) \mid x \in A\}.$$

(Es claro que un subconjunto A de B o bien contiene a x o bien no lo contiene).

Entonces

$$\wp_k(B) = \{A \in \wp_k(B) \mid x \notin A\} \stackrel{\circ}{\cup} \{A \in \wp_k(B) \mid x \in A\}.$$

⁴De donde vemos que

$$\binom{n}{k} = |\{A \in \wp_k(B) \mid x \notin A\}| + |\{A \in \wp_k(B) \mid x \in A\}|.$$

Hagámonos ahora la pregunta siguiente: ¿cuántos subconjuntos $A \in \wp_k(B)$ no contienen a x ?

Es claro que este número es

$$\binom{B \setminus \{x\}}{k} = \binom{n-1}{k}.$$

(Aquí, $\binom{B \setminus \{x\}}{k}$ denota el número de subconjuntos con k elementos que tiene $B \setminus \{x\}$).

Ahora nos preguntémonos por la cardinalidad del conjunto

$$\{A \in \wp_k(B) \mid x \in A\}.$$

Otra vez, un conjunto con k elementos que contiene a x se forma escogiendo los $k-1$ elementos de A que no son x . Hay

$$\binom{n-1}{k-1} = \binom{A \setminus \{x\}}{k-1},$$

de éstos.

Por lo tanto,

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1} \quad (5.2)$$

La fórmula anterior es conocida como la fórmula del triángulo de Pascal, un fragmento de él es:

⁴ $X \stackrel{\circ}{\cup} Y$ denota la unión de X y Y y además dice que X y Y son ajenos.

$$\begin{array}{ccccccc}
 & \binom{0}{0} & & \binom{1}{1} & & & \\
 & \binom{1}{0} & & \binom{2}{1} & & & \\
 \binom{2}{0} & \binom{2}{1} & & & \binom{2}{2} & & \\
 \binom{3}{1} & & \binom{3}{2} & & & & \\
 & \vdots & & \ddots & & & \\
 & & \binom{n-1}{k-1} & & \binom{n-1}{k} & & \\
 \binom{n}{k-1} & & & & \binom{n}{k} & &
 \end{array}$$

Si uno observa el renglón n -ésimo del triángulo arriba, uno obtiene

$$\binom{n}{0}, \binom{n}{1}, \dots, \binom{n}{k-1}, \binom{n}{k}, \dots, \binom{n}{n-1}, \binom{n}{n}$$

que se llaman los coeficientes binomiales, pues cuando uno calcula

$$(a+b)^n,$$

se obtiene

$$\begin{aligned}
 & \binom{n}{0} a^n b^0 + \binom{n}{1} a^{n-1} b + \dots + \binom{n}{k-1} a^{n-k+1} b^{k-1} + \\
 & + \binom{n}{k} a^{n-k} b^k + \dots + \binom{n}{n-1} a b^{n-1} + \binom{n}{n} a^0 b^n.
 \end{aligned}$$

Esto no es una casualidad, observemos

$$\underbrace{(a+b)_1(a+b)_2 \cdots (a+b)_n}_{n \text{ factores } (a+b)},$$

Este producto se puede realizar escogiendo uno de los dos sumandos dentro de cada paréntesis. Por ejemplo a en el primero, b en el segundo, a en el

tercero, b en todos los demás. Se obtiene un término a^2b^{n-2} . ¿De cuántas maneras podemos obtener este mismo término?

Para responder esto, notemos que lo que se hace es escoger dos a de los n paréntesis (escogiendo b en los demás). Esto se puede hacer de $\binom{n}{2}$ maneras, por lo que el coeficiente de a^2b^{n-2} debe ser

$$\binom{n}{2} = \binom{n}{n-2}.$$

El mismo razonamiento nos dice que el coeficiente de $a^k b^{n-k}$, al calcular el producto es

$$\binom{n}{k} = \binom{n}{n-k}.$$

Así, hemos visto que

$$(a+b)^n = \binom{n}{0} a^n b^0 + \binom{n}{1} a^{n-1} b + \dots + \binom{n}{k} a^{n-k} b^k + \dots + \binom{n}{n} a^0 b^n.$$

Notemos, como un caso particular, que

$$\begin{aligned} 2^n &= (1+1)^n = \\ &= \binom{n}{0} 1^n 1^0 + \binom{n}{1} 1^{n-1} 1^1 + \dots + \binom{n}{k} 1^{n-k} 1^k + \dots + \binom{n}{n} 1^0 1^n = \\ &= \binom{n}{0} + \binom{n}{1} + \dots + \binom{n}{k} + \dots + \binom{n}{n}. \end{aligned}$$

Que se puede interpretar de esta manera: 2^n es el número de subconjuntos de un conjunto con n elementos, que también se pueden contar así:

subconjuntos con 0 elementos: $\binom{n}{0} +$

subconjuntos con 1 elemento: $\binom{n}{1} +$

subconjuntos con 2 elementos: $\binom{n}{2} + \dots$

$\dots +$ subconjuntos con n elementos: $\binom{n}{n},$

lo que dice otra vez que

$$2^n = \binom{n}{0} + \binom{n}{1} + \dots + \binom{n}{k} + \dots + \binom{n}{n}.$$

5.3 Permutaciones

¿Cuántas funciones biyectivas hay de A a A ?

Una función biyectiva de A en A se llama una permutación en A .

Proposición 25 . *Si $|A| = |B|$, entonces hay el mismo número de biyecciones entre A y A que entre B y B .*

Demostración. Denotemos por

$$Bi(A) = \{A \xrightarrow{f} A \mid f \text{ es biyectiva}\}.$$

Tomemos $A \xrightarrow{\varphi} B$ una biyección entre A y B ($|A| = |B|$). Observando el siguiente diagrama

$$\begin{array}{ccc} A & \xleftarrow{\varphi^{-1}} & B \\ f \downarrow & = & \downarrow \varphi \circ f \circ \varphi^{-1} \\ A & \xrightarrow{\varphi} & B \end{array},$$

podemos notar que $\varphi \circ f \circ \varphi^{-1}$ es una biyección, ya que es una composición de biyecciones.

Esto sugiere la siguiente función

$$\begin{array}{ccc} Bi(A) & \xrightarrow{\varphi \circ (\) \circ \varphi^{-1}} & Bi(B) \\ f & \mapsto & \varphi \circ f \circ \varphi^{-1} \end{array}.$$

Esta función es biyectiva, pues su inversa es $\varphi^{-1} \circ (\) \circ \varphi$. Por lo tanto,

$$|A| = |B| \Rightarrow |Bi(A)| = |Bi(B)|.$$

■

Una pequeña modificación al argumento anterior nos da lo siguiente:

Proposición 26 *Si $|A| = |B| = n$, entonces $|Bi(A)| = |Bi(A, B)|$.*

Demostración. Sea $A \xrightarrow{\varphi} B$ una biyección. Si $A \xrightarrow{f} A$ es una biyección, entonces $A \xrightarrow{\varphi \circ f} B$, es una biyección de A en B .

Así, la correspondencia

$$\begin{array}{ccc} Biy(A) & \xrightarrow{\varphi(\)} & Biy(A, B) \\ f & \longmapsto & \varphi \circ f \end{array}$$

es una función cuyo inverso es $\varphi^{-1}(\)$.

Por lo tanto, $|Biy(A)| = |Biy(A, B)|$. ■

Proposición 27 . $|A| = n \Rightarrow |Biy(A)| = n! = n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot 2 \cdot 1$.

Demostración. Por la proposición anterior, podemos tomar

$$A = n = \{0, \dots, n-1\}.$$

Sea ahora

$$n \xrightarrow{f} n$$

una biyección notemos que

$$\begin{array}{ccc} n & \xrightarrow{f} & n \\ incl \uparrow & & \uparrow incl \\ n-1 & \xrightarrow{f|_{n-1}} & n \setminus \{f(n-1)\} \end{array}$$

commuta.

Nótese que una biyección de n en n se consigue escogiendo la imagen del último elemento de n , (que es $n-1$) lo que se puede hacer de n maneras, y después tomando una biyección

$$n-1 \longrightarrow n \setminus \{f(n-1)\}$$

(lo que se puede hacer de $|Biy(n-1)|$ maneras).

Por lo tanto,

$$\begin{aligned} |Biy(n)| &= n|Biy(n-1)| = n(n-1)|Biy(n-2)| = \dots \\ &\dots = n(n-1)\dots 2|Biy(1)| = n!. \end{aligned}$$

Ya que hay una única biyección entre 1 y $1 = \{0\}$. (La que manda 0 a 0). ■

5.3.1 Ordenaciones

Sea A un conjunto con n elementos y B un conjunto con m elementos, denotemos

$$Iny(A, B) = \{A \xrightarrow{f} B \mid f \text{ inyectiva}\}.$$

Como ya es usual, demostraremos la siguiente:

Proposición 28 . Si $|A| = |C|$ y $|B| = |D|$, entonces $Iny(A, B) = Iny(C, D)$.

Demostración. Tomemos biyecciones $A \xrightarrow{f} C$ y $B \xrightarrow{g} D$.

La asignación

$$\begin{array}{ccc} A & & C \\ \downarrow \Psi & \mapsto & \downarrow g \circ \Psi \circ f^{-1} \\ B & & D \end{array}$$

produce una biyección entre $Iny(A, B)$ y $Iny(C, D)$ cuya inversa es $f \circ (\) \circ g^{-1}$.

■

Corolario 6 . Si $|A| = k$ y $|B| = n$, entonces $|Iny(A, B)| = |Iny(k, n)|$.

$$|Iny(k, n)| = \begin{cases} 0 & \text{si } k > n, \text{ por el principio de las pichoneras} \\ n! & \text{si } k = n, \text{ pues en este caso una inyección es una biyección} \\ \frac{n!}{(n-k)!} & \text{si } k < n. \end{cases}$$

Demostración. Lo único que tenemos que demostrar es el tercer caso de arriba. Para ello, analicemos las maneras de construir una función inyectiva $f : k \rightarrow n$.

Esto se puede hacer así:

primero escojamos la imagen de f , que es un subconjunto con k elementos del conjunto n . Esto se puede hacer de $\binom{n}{k}$ maneras.

Una vez que hemos escogido $F = Im f$, respondamos lo siguiente: ¿cuántas funciones inyectivas hay de k a F ?

Como ambos conjuntos tienen k elementos, entonces una inyección entre ellos resulta una biyección, por lo que hay $k!$ inyecciones de k a F .

Por lo tanto hay

$$\binom{n}{k} k!$$

funciones inyectivas de k a n . Es decir que

$$|Iny(k, n)| = \frac{n!}{(n - k)!}.$$

■ Vamos a matizar un poco nuestro lenguaje.

Definición 79

1. Llamaremos alfabeto a un conjunto finito F , y desde luego, sus elementos se llamarán letras.
2. Una sucesión finita de letras,

$$a_1 a_2 \dots a_k$$

se llamará “palabra” con k letras (también se llama “ordenación con repetición”).

3. Si la palabra no tiene letras repetidas, la palabra se llamará “ordenación”.

Notemos que una palabra con k letras corresponde a una función

$$\{1, 2, \dots, k\} \rightarrow F.$$

(Que manda i a a_i). Si la “palabra no tiene letras repetidas se llamará una “ordenación”.

4. Denotemos con O_k^n el número de ordenaciones (sin repetición) con k letras tomadas de un alfabeto con n letras. Como ya hemos notado, este número es

$$O_k^n = |Iny(k, n)| = \frac{n!}{(n - k)!}.$$

5. Denotemos con OR_k^n el número de ordenaciones con repetición con k letras tomadas de un alfabeto con n letras. Como cada una de estas ordenaciones puede pensarse como una función de un conjunto con k elementos a un conjunto con n elementos, este número es

$$|\mathbf{n}^k| = n^k.$$

(Ver el teorema 86, en la página 291).

Para completar este bosquejo acerca de algunos números que aparecen cuando uno hace cuentas, notemos que:

Observación 81

1. Si A y B son conjuntos ajenos (es decir, de intersección vacía) entonces

$$|A \cup B| = |A| + |B|.$$

2. Como consecuencia de lo anterior, tenemos que

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

Como se puede ver si uno escribe

$$(A \cup B) = (A \setminus (A \cap B)) \stackrel{\circ}{\cup} (B \setminus (B \cap A)) \stackrel{\circ}{\cup} (A \cap B),$$

pues entonces

$$\begin{aligned} |A \cup B| &= |A \setminus (A \cap B)| + |B \setminus (B \cap A)| + |A \cap B| = \\ &= |A \setminus (A \cap B)| + |B \setminus (B \cap A)| + |A \cap B| + |A \cap B| - |A \cap B| = \\ &= (|A \setminus (A \cap B)| + |A \cap B|) + (|B \setminus (B \cap A)| + |A \cap B|) - |A \cap B| = \\ &= |A| + |B| - |A \cap B|. \end{aligned}$$

- 3.

$$|A \times B| = |A| |B|.$$

Aquí hay que notar dos cosas:

- La primera, es que si A tiene el mismo número de elementos que A' y B tiene el mismo número de elementos que B' , entonces

$$|A \times B| = |A'| |B'|.$$

Pues si $A \xrightarrow{f} A'$ y $B \xrightarrow{g} B'$ son biyecciones, entonces

$$\begin{array}{ccc} A \times B & \xrightarrow{(f,g)} & A' \times B' \\ (a, b) & \mapsto & (f(a), g(b)) \end{array}$$

es una biyección, cuyo inverso es $A' \times B' \xrightarrow{(f^{-1},g^{-1})} A \times B$.

- La segunda cosa que hay que notar, es que si alguno de los dos conjuntos es infinito, A , o B , entonces

$$|A \times B| = |A| |B|$$

es la definición de lo que significa multiplicar el cardinal de A por el cardinal de B .

En el primer caso, (A y B finitos) notemos que

$$A \times B = \bigcup_{a \in A}^{\circ} \{\{a\} \times B\},$$

es una unión de conjuntos ajenos dos a dos, cada uniendo con $|B|$ elementos.

Por lo tanto

$$|A \times B| = \sum_{a \in A} \{\{a\} \times B\} = |A| |B|.$$

Las siguientes proposiciones, aunque no sean precisamente una herramienta para el conteo, caben aquí.

Proposición 29

1. Si A, B, C son conjuntos entonces

$$|(A^B)^C| = |A^{C \times B}|.$$

2. $(n^m)^p = n^{pm}$.

Demostración. En virtud de la proposiciones anteriormente vistas, basta demostrar 1), pues 2) se sigue tomando A con n elementos, B con m elementos y C con p elementos.

Entonces debemos dar una función

$$(A^B)^C \rightarrow A^{C \times B}$$

que tenga inverso.

Tomemos

$$\begin{array}{ccc} (A^B)^C & \xrightarrow{\Psi} & A^{C \times B} \\ f & \longmapsto & \hat{f} \\ & & \hat{f}(c, b) =: f(c, b)^* \end{array} \quad 5 \text{ (ver nota pág. sig.)}$$

Ahora definamos la función Γ definida por

$$\begin{array}{ccc} A^{C \times B} & \xrightarrow{\Gamma} & (A^B)^C \\ g & \longmapsto & \vec{g} \\ & & (\vec{g}(c))(b) =: g(c, b) \end{array},$$

y veamos que en efecto es el inverso de Ψ ⁶.

Tomemos una f en $(A^B)^C$,

$$(\Gamma \circ \Psi)(f) = \Gamma((\Psi)(f)) = \overrightarrow{(\Psi)(f)},$$

queremos ver que esta función coincide con f :

Queremos demostrar que

$$\forall c \in C, \overrightarrow{(\Psi)(f)}(c) = f(c),$$

que es una función de B en A .

Para esto, queremos demostrar que

$$\forall b \in B, ((\overrightarrow{(\Psi)(f)}(c))(b)) = f(c)(b),$$

pero

$$((\overrightarrow{(\Psi)(f)}(c))(b)) =: \Psi(f)((c, b))$$

y

$$\Psi(f)((c, b)) = \hat{f}(c, b) =: f(c)(b), \quad \forall b \in B.$$

Por lo tanto,

$$((\overrightarrow{(\Psi)(f)}(c))(b)) = f(c)(b), \quad \forall b \in B.$$

Por lo tanto,

$$((\overrightarrow{(\Psi)(f)}(c))) = f(c) \quad (\forall c \in C)$$

⁵Notemos que f es una función de C en A^B , es decir que a cada elemento c de C , f le asocia una función $f(c)$ de B en A , es decir que $f(c)(b)$ es un elemento de A .

⁶ \hat{f} tiene que ser una función de $C \times B$ en A , así que debe asociarle a un elemento (c, b) de $C \times B$, un elemento de A . ¿Qué podría ser más natural que definir $\hat{f}(c, b)$ como $f(c)(b)$?

⁶Esta función también es muy natural, pero tenemos que comprobar que es en efecto el inverso de Ψ .

Por lo tanto,

$$\overrightarrow{(\Psi)(f)} = f.$$

Por lo tanto, $\Gamma \circ \Psi$ es la identidad.

Además $\Psi \circ \Gamma$ es la identidad:

$$(\Psi \circ \Gamma)(g) = \Psi((\Gamma)(g)) = \widehat{\Gamma(g)},$$

queremos ver que esta función es g .

Sea $(c, b) \in C \times B$, entonces

$$\widehat{\Gamma(g)}((c, b)) =: ((\Gamma(g))(c))(b) = (\vec{g}(c))(b) =: g(c, b), \quad \forall (c, b) \in C \times B.$$

Por lo tanto $\widehat{\Gamma(g)} = g$. Así, tenemos que Ψ y Γ , son cada una la inversa de la otra. ■

Proposición 30 . *Si B y C son conjuntos ajenos, entonces*

1. $|A^B \times A^C| = |A^{B \dot{\cup} C}|$.
2. $nm \cdot np = n(m + p)$.

Demostración. Como en la proposición anterior, basta demostrar el inciso 1).

Tomemos B y C ajenos.

Entonces la función

$$\begin{aligned} A^B \times A^C &\rightarrow A^{B \dot{\cup} C} \\ (f, g) &\mapsto f \vee g \end{aligned} \quad \text{definida por} \quad (f \vee g)(x) = \begin{cases} f(x) & \text{si } x \in B \\ g(x) & \text{si } x \in C \end{cases}$$

Tiene inversa, que es

$$\begin{aligned} A^{B \dot{\cup} C} &\rightarrow A^B \times A^C \\ h &\mapsto (h|_B, h|_C) \end{aligned}$$

La comprobación de que en efecto son funciones inversas, se deja al lector.

■

Ejercicio 216 *Demuestre que en efecto las dos funciones de la Proposición 30 son inversas una de la otra.*

1. Se puede pensar que la realización de un evento es “escoger un elemento dentro del conjunto finito de sucesos posibles”. Por ejemplo el evento de que un dado caiga mostrando dos puntos, lo podemos pensar como el evento de escoger el 2 dentro del conjunto $\{1, 2, 3, 4, 5, 6\}$.
2. El evento de escoger un placa de automóvil con tres números y tres letras equivale al evento de escoger una palabra con tres letras en el conjunto de las palabras con tres letras

$$OR_3^{a,b,c,\dots,z} = \{aaa, aab, aac, \dots, aaz, aba, abb, abc, abd, \dots, abz, \dots, zzy, zzz\}$$

y escoger un número con tres cifras dentro del conjunto Lo anterior equivale a escoger un elemento del producto cartesiano, lo que se puede hacer de $27^3 \cdot 10^3 = 19683000$ maneras, más que suficientes para las placas de cada automóvil mexicano.

3. Se tiene un librero con n niveles, cada uno con un libro de Álgebra, otro de biología y otro de Cálculo. Si se escoge un libro por cada nivel del librero, ¿de cuántas maneras se pueden escoger i libros de Álgebra, j de Biología y el resto de Cálculo?

Si quisiéramos hacer una elección como se pide, podríamos empezar por escoger los i niveles del librero de donde tomaremos un libro de Álgebra ($\binom{n}{i}$ maneras), luego, de los $n-i$ niveles de donde no se ha escogido un libro, podemos escoger los j niveles de donde tomaremos libros de biología ($\binom{n-i}{j}$ maneras). De los niveles que quedan, tomamos los libros de Cálculo.

En resumen hay

$$\binom{n}{k} \cdot \binom{n-i}{j}$$

maneras.

4. Supongamos que queremos encontrar el coeficiente de $a^i b^j c^k$ en $(a+b+c)^n$. Éste, aunque no lo parezca, es el mismo problema que el anterior. Escribamos

$$(a+b+c)^n = \underbrace{(a+b+c) \bullet (a+b+c) \bullet \dots \bullet (a+b+c)}_{n \text{ factores}}$$

Al hacer el producto, se escoge un sumando por cada paréntesis. Determinar el coeficiente de $a^i b^j c^k$ es exactamente lo mismo que contar el número de maneras en que se pueden escoger i áes en el producto anterior, y después j bes entre los $n - i$ factores restantes. Por lo tanto la solución es otra vez:

$$\binom{n}{k} \binom{n-i}{j}.$$

5. Queremos encontrar el coeficiente de $a_1^{l_1} a_2^{l_2} \cdots a_k^{l_k}$ en

$$(a_1 + a_2 + \dots + a_k)^n.$$

Es claro que la respuesta es

$$\binom{n}{i_1} \binom{n-i_1}{i_2} \binom{n-i_1-i_2}{i_3} \cdots \binom{n-(i_1+i_2+\dots+i_{k-2})}{i_{k-1}}$$

6. $(1 + 2 + 3 + 4)^2 = 10^2 = 100$, y se puede escribir como

$$\sum_{(i,j,k)} \binom{2}{i} \binom{2-i}{j} \binom{2-i-j}{k} 1^i 2^j 3^k 4^{(2-i-j)}$$

Como se ilustra en la siguiente tabla

i	j	k	$\binom{2}{i} \cdot \binom{2-i}{j} \cdot \binom{k}{2-i-j}$	$1^i 2^j 3^k 4^{(2-i-j-k)}$	$\binom{2}{i} \cdot \binom{2-i}{j} \cdot \binom{k}{2-i-j} \cdot 1^i 2^j 3^k 4^{(2-i-j-k)}$
0	0	0	111	16	16
0	0	1	112	12	+24
0	0	2	111	9	+9
0	1	0	121	8	+16
0	1	1	121	6	+12
0	2	0	111	4	+4
1	0	0	211	4	+8
1	0	1	211	3	+6
1	1	0	211	2	+4
2	0	0	111	1	+1
100					

Ejercicio 217 . Calcular el coeficiente de $x^2y^3z^4w$ en $(x+2y-z+w)^{10}$.

5.4 ¿Cuántas funciones suprayectivas hay de A a B ?

Empezaremos con la siguiente proposición.

Proposición 31 Si $|A| = |A'|$ y $|B| = |B'|$, entonces el número de funciones suprayectivas de A a B es el mismo que el de funciones suprayectivas de A' a B' .

Demostración. Denotemos

$$Supr(A, B) = \{f : B \rightarrow A \mid f \text{ es suprayectiva}\}.$$

Tomemos $A \xrightarrow{\varphi} A'$ y $B \xrightarrow{\psi} B'$, biyecciones. Definamos

$$\begin{array}{ccc} \text{Supr}(A, B) & \xrightarrow{\psi(\)\varphi^{-1}} & \text{Supr}(A', B) \\ g & \longmapsto & \psi \circ g \circ \varphi^{-1} \end{array} .$$

Vea el siguiente diagrama

$$\begin{array}{ccc} A & \xleftarrow{\varphi^{-1}} & A' \\ \downarrow g & & \\ B & \xrightarrow{\psi} & B' \end{array}$$

es decir que si g es una función suprayectiva de A a B , entonces $\psi \circ g \circ \varphi^{-1}$ lo es de A' a B' (se está usando el hecho de que una composición de funciones suprayectivas es suprayectiva). La función

$$\text{Supr}(A, B) \xrightarrow{\psi(\)\varphi^{-1}} \text{Supr}(A', B)$$

es biyectiva (porque su inverso es $\psi^{-1}(\)\varphi$).

Por lo tanto $|\text{Supr}(A, B)| = |\text{Supr}(A', B)|$. ■

Volvamos a hacernos la pregunta:

¿cuántas funciones suprayectivas hay de n a m ?

Denotemos este número por S_m^n . Es claro que si $n < m$, entonces $S_m^n = 0$ (si hubiera una función suprayectiva de n a m , entonces habría una función inyectiva de m a n , y así $m < n$, ∇).

Notemos ahora que si $n \geq m$, y

$$n \xrightarrow{f} m$$

es una función suprayectiva, entonces, dado que

$$m = \{0, 1, \dots, m-1\},$$

podemos considerar los siguientes m subconjuntos de n :

$$\begin{aligned} f^{-1}(0) &= \{i \in n \mid f(i) = 0\} \\ f^{-1}(1) &= \{i \in n \mid f(i) = 1\} \\ f^{-1}(2) &= \{i \in n \mid f(i) = 2\} \\ &\vdots \\ f^{-1}(m-1) &= \{i \in n \mid f(i) = m-1\} \end{aligned}$$

que parten a n en m partes no vacías⁷, ajenas dos a dos.

Es decir que el conjunto

$$\{f^{-1}(0), f^{-1}(1), \dots, f^{-1}(m-1)\}$$

es una partición de n , en el sentido usual.

Podemos replantear lo que hace la función f de la siguiente manera: f envía todos los elementos de $f^{-1}(0)$ a 0, todos los elementos de $f^{-1}(1)$ a 1, etcétera.

Renombremos los elementos de la partición:

$$\{f^{-1}(0), f^{-1}(1), \dots, f^{-1}(m-1)\},$$

denotando

$$X_j = f^{-1}(j).$$

Así la partición se denota ahora como $\{X_0, X_1, \dots, X_{m-1}\}$.

¿Habrá otras funciones además de f que produzcan la misma partición?

Es claro que sí, pues lo que se necesita es que todos los elementos de una misma parte vayan a dar al mismo elemento pero que elementos en partes distintas vayan a dar a elementos distintos.

Notemos que en este momento estamos contando el número de biyecciones entre el conjunto $\{X_0, X_1, \dots, X_{m-1}\}$ y m .

Así pues, hay $m!$ funciones de n a m que dan lugar a la misma partición.

Así que si conociéramos el número P_m^n de particiones de n en m partes, tendríamos resuelto el problema, pues debe ser claro de lo anterior que

$$S_m^n = m! P_m^n.$$

Es claro que

$$P_m^n = 0 \text{ si } m > n,$$

y que

$$P_n^n = 1$$

⁸, así que lo más interesante es decir algo de P_m^n cuando $n > m$.

⁷Como f es suprayectiva, cada elemento de m proviene de un elemento de n , así que cada

$$f^{-1}(j) = \{i \in n \mid f(i) = j\}$$

no es vacío.

⁸Un conjunto con n elementos sólo se puede partir en n partes no vacías de una sola manera: tomando n partes cada una con un elemento.

5.4.1 Relación de recurrencia para P_m^n

Escojamos un elemento de

$$n = \{0, 1, 2, \dots, n-1\},$$

por ejemplo $n-1$. Podemos separar las particiones de n en m partes en dos clases:

1. Las particiones en las que una parte es $\{n-1\}$ (es decir, $n-1$ está solitario en su parte). Es claro que las $m-1$ partes restantes forman una partición de $\{0, 1, 2, \dots, n-2\} = n-1$.

Entonces el número de particiones que estamos contando aquí, es

$$P_{m-1}^{n-1}.$$

2. Las particiones en las que $n-1$ no está solitario en su parte. Por ejemplo, supongamos que la partición es $\{Y_0, Y_1, \dots, Y_{m-1}\}$, y que $n-1 \in Y_j$, con $|Y_j| > 1$.

¿Qué obtenemos si quitamos $n-1$ de Y_j ?

Es claro que obtenemos una partición de $n-1 = \{0, 1, 2, \dots, n-2\}$, en m partes:

$$\{Y_0, Y_1, \dots, Y_j \setminus \{n-1\}, \dots, Y_{m-1}\}.$$

Si ahora empezáramos con una partición en m partes de $n-1$, por ejemplo

$$\{Z_0, Z_1, \dots, Z_j, \dots, Z_{m-1}\}.$$

En esta partición $n-1$ no aparece en ninguna de las partes, la queremos agregar a alguna parte, para tener una partición de n , con la propiedad de que $n-1$ no quede solitaria.

¿Cuántas elecciones podemos hacer? Es claro que m . Entonces, es claro que por cada una de las P_m^{n-1} maneras de partir $n-1$ en m partes, tenemos m particiones de n en las que $n-1$ no queda sola. El resultado de nuestra cuenta es

$$P_m^{n-1} \cdot m.$$

Sumando ahora los números en 1) y en 2) tenemos que

$$P_m^n = P_{m-1}^{n-1} + P_m^{n-1} \cdot m.$$

Ejemplo 120 . ¿Cuántas funciones suprayectivas hay de $\{0, 1, 2, 3\}$ a $\{a, b\}$?

Por lo que vimos arriba, hay

$$\begin{aligned} S_2^4 &= 2!P_2^4 = 2[P_1^3 + P_2^32] = 2[1 + 2(P_1^2 + P_2^22)] = \\ &= 2[1 + 2(1 + 1 \cdot 2)] = 2 \cdot 7 = 14. \end{aligned}$$

¿Cuáles son las funciones?

1) Contemos primero las que mandan únicamente a 3 en su imagen (es decir, las imágenes de 2, de 1 y de 0 son distintas de la de 3). La imagen de 3 se puede escoger de dos maneras: se escoge a ó b . Como sólo 3 va a su imagen, los demás elementos van a dar al otro elemento de $\{a, b\}$, así que la cuenta va en dos:

$$\begin{array}{ll} 3 \longmapsto a & 3 \longmapsto b \\ 2 \longmapsto b & 2 \longmapsto a \\ 1 \longmapsto b & 1 \longmapsto a \\ 0 \longmapsto b & 0 \longmapsto a \end{array}.$$

2) Contemos ahora las funciones que mandan 3 y otro elemento de $\{0, 1, 2, 3\}$ al mismo elemento de $\{a, b\}$. Por ejemplo pensemos en que $3 \longmapsto a$. Quitando 3 del dominio, contamos ahora el número de funciones suprayectivas de $\{0, 1, 2\}$ a $\{a, b\}$ (hay tantas como $2!P_2^3$). Pero ¿de cuántas formas se puede partir un conjunto con tres elementos en dos partes? En una parte deben haber dos elementos y en la otra uno solo. ¿De cuántas maneras se puede escoger el que queda solo?, de tres. Por lo tanto

$$2!P_2^3 = 6.$$

Así que hay 6 funciones suprayectivas que mandan 3 a a , y algún otro elemento también a a .

Como también hay 6 funciones suprayectivas que mandan 3 a b y a algún otro elemento a b tenemos 12 funciones suprayectivas en las que 3 y otro elemento van a dar al mismo elemento de $\{a, b\}$.

Seamos un poco más explícitos en la descripción de las funciones suprayectivas de $\{0, 1, 2\}$ en $\{a, b\}$:

Particiones de $\{0, 1, 2\}$ en dos partes:

$$\{\{0, 1\}, \{2\}\}, \{\{0, 2\}, \{1\}\}, \{\{1, 2\}, \{0\}\}$$

De la partición $\{\{0, 1\}, \{2\}\}$ obtenemos las dos funciones:

$$\begin{array}{l} 0 \mapsto a \\ 1 \mapsto a \\ 2 \mapsto b \end{array} \quad \begin{array}{l} 0 \mapsto b \\ 1 \mapsto b \\ 2 \mapsto a \end{array}$$

De las otras dos particiones obtenemos otras 4 funciones (así que contamos 6 en total).

Veamos ahora como usar las funciones de arriba, para construir funciones suprayectivas de $\{0, 1, 2, 3\}$ en $\{a, b\}$, para las que 3 comparte su imagen con otro elemento de $\{0, 1, 2, 3\}$:

$3 \mapsto a$	$3 \mapsto b$	$3 \mapsto a$	$3 \mapsto b$
$0 \mapsto a$	$0 \mapsto a$	$0 \mapsto b$	$0 \mapsto b$
$1 \mapsto a$	$1 \mapsto a$	$1 \mapsto b$	$1 \mapsto b$
$2 \mapsto b$	$2 \mapsto b$	$2 \mapsto a$	$2 \mapsto a$

Así es como se construyen las 12 funciones del inciso 2.

5.5 Ejercicios

Ejercicio 218 . *En una fiesta, había n personas que se saludaron de mano un número impar de veces. Demostrar que n es par.*⁹

Ejercicio 219 . *Si n es un cuadrado (demuéstrese que) n tiene un número impar de divisores. Si n no es un cuadrado entonces tiene un número par de divisores.*¹⁰

Ejercicio 220 . *Un dominó puede cubrir exactamente dos cuadrados adyacentes de un tablero de ajedrez.*

⁹ A cada persona pregúntese cuántas manos estrechó. Sumando todos estos números obtenemos el doble del número de apretones de manos (es decir, un número par, así que no puede ser impar el número de personas que saludaron un número impar de veces).

¹⁰ Sea

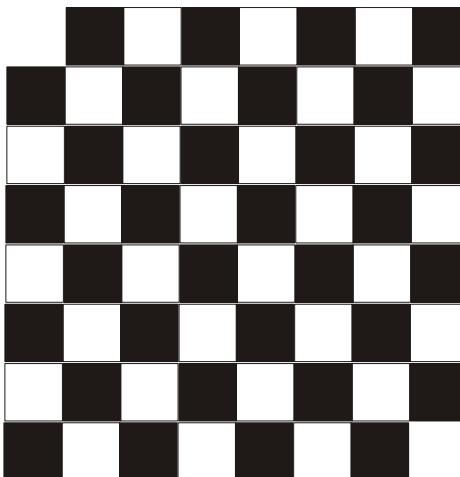
$$n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$$

la factorización en primos de n (Teorema fundamental de la Aritmética).

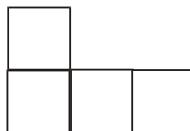
Si n es un cuadrado, entonces cada exponente a_i es par. ¿Cómo escogemos un divisor de n ? Tomando una sucesión de exponentes b_1, b_2, \dots, b_k tales que $b_i \leq a_i$.

Esto se puede hacer de $|a_1 + 1||a_2 + 1| \cdots |a_k + 1|$ maneras, pues, por ejemplo, b_1 puede escogerse en el conjunto $\{0, 1, 2, \dots, a_1\}$.

1. Muestre que 32 dominós pueden cubrir exactamente el tablero completo de ajedrez (que es de 8×8).
2. Ahora recorte dos cuadrados en esquinas opuestas del tablero. Demuestre que lo que queda no se puede cubrir con 31 dominós.



Ejercicio 221 . Los ladrillos con forma de L del tipo



que cubren cuatro cuadrados cada uno pueden cubrir un rectángulo de 5×8 como se muestra abajo

1. Demuestre que un rectángulo de 5×4 o de 6×6 no se puede cubrir con estos ladrillos.¹¹
2. Demuestre que si un rectángulo se puede cubrir, entonces el enladrillado usa un número par de ladrillos.¹²

¹¹Es claro que esto se sigue del inciso siguiente, pues las cantidades de ladrillos necesarias serían 5 y 9 respectivamente.

¹²Observemos un ladrillo:

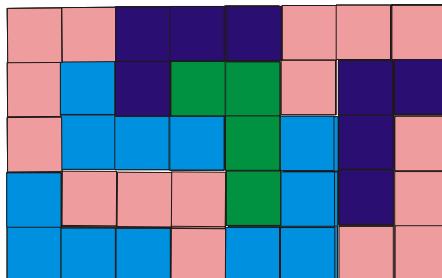
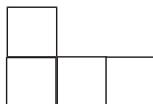


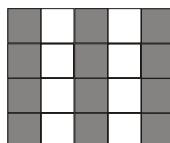
Figura 5.1:

3. Demuestre que si 8 divide mn y tanto m como n son mayores que tres, entonces un rectángulo de $m \times n$ se puede enladrillar.

Ejercicio 222 . *Un Senado tiene 100 miembros, así que tiene 2^{100-1} subcomités posibles (quitando el subcomité vacío). ¿Cuál es el mayor número de subcomités que se pueden formar, sujetos a la condición de que cualesquiera dos subcomités tengan por lo menos un elemento en común?*



Supongamos que tenemos un rectángulo que se puede cubrir con los ladrillos dados. Como cada ladrillo consta de 4 cuadrados “unitarios”, entonces, si el rectángulo mide nm unidades entonces nm es un múltiplo de 4, y n ó m tiene que ser par. Supongamos que n es par, y veamos el rectángulo de manera que la altura mida n y la base mida m :



Ahora, sombreamos las columnas alternadamente, como en el dibujo anterior. Notemos que hay un número par de cuadros grises y un número par de cuadros blancos. Ahora, coloquemos un ladrillo y notemos que como quiera que se ponga, cubre un número impar de cuadros grises: 1 ó 3. Así que si el rectángulo se puede enladrillar debemos usar un número par de ladrillos para cubrir el número par de cuadros grises.

Ejercicio 223 . *En el club X cada miembro pertenece a dos comités y cualesquiera dos comités tienen exactamente un miembro en común. Hay cinco comités ¿Cuántos miembros tiene el club?*

Ejercicio 224 . *Se tienen n pelotas etiquetadas 1, 2, 3, ..., ¿De cuántas maneras pueden disponerse en un círculo de tal forma que los números de dos bolas adyacentes cualesquiera difieran en 1 ó en 2?*

Ejercicio 225 . *Una ficha de dominó (o un dominó) es un par de números cada uno en $\{0, 1, \dots, 6\}$.*

1. ¿Cuántas fichas diferentes de dominó hay si consideramos (a, b) y (b, a) la misma ficha. (éste no es un problema para los que conocen el dominó).
2. ¿De cuántas maneras podemos escoger un par de dominóes que cacen, es decir que comparten un número?

Ejercicio 226 . *n un motor de seis cilindros los cilindros pares están a la izquierda y los impares a la derecha. Un buen orden de explosiones es una permutación de los números del 1 al 6 en que se alternen los lados del motor.*

1. ¿Cuántos buenos órdenes para explosiones hay?
2. Haga lo mismo para un motor con $2n$ cilindros.

Ejercicio 227 . *Hay nueve libros diferentes en un librero. Cuatro de ellos son rojos y cinco de ellos son negros. ¿Cuántos arreglos se pueden hacer si*

1. no hay restricciones?
2. los libros negros tienen que ir juntos?
3. los libros negros tienen que ir juntos y también los libros rojos deben ir juntos?
4. los colores deben alternarse?

Ejercicio 228 . *Hay 24 tomos de una enciclopedia en un librero. ¿De cuántas maneras se pueden escoger 5 tomos sin escoger tomos consecutivos?*

Ejercicio 229 . *Se escriben los números desde 1 hasta n , hasta que el número total de dígitos que se escribe es 1890? ¿Quién es n ?*

Ejercicio 230 . *Consideremos equivalentes dos números de 10 dígitos si uno se puede obtener del otro permutando sus dígitos. ¿Cuántos números de 10 dígitos no equivalentes hay?*

Ejercicio 231 . *Demostrar que el número de apareamientos diferentes para la primera ronda de un torneo de tenis con $2n$ participantes es¹³*

$$(1)(3)(5) \dots (2n-1).$$

Ejercicio 232 . *Demuestre que el número de maneras diferentes en que m números distintos del conjunto $\{1, 2, 3 \dots n\}$ se pueden colocar en un círculo es*

$$\frac{n!}{m(n-m)!}$$

donde arreglos que difieren solamente por una rotación se consideran iguales. Por ejemplo,

$$\begin{array}{ccccc} & 2 & & 12 & \\ 12 & & 3 & = & 9 & 2 \\ & 9 & & & & 3 \end{array}.$$

1. *¿De cuántas maneras se pueden escoger tres números del conjunto $\{1, 2, \dots, 99\}$ tal que su suma sea múltiplo de 3?*
2. *Generaliza lo anterior para selecciones de tres números del conjunto*

$$\{1, \dots, 3n\}.$$

Ejercicio 233 . *Si se escribieran los números desde 1 hasta 1000000, ¿cuántas veces aparecería el número 0?*

¹³Si se etiquetan los jugadores 1, 2, 3, ..., $2n$, escojamos el rival de 1, esto se puede hacer de $2n-1$ maneras. Enseguida hay que hacer la misma cuenta pero con los $2n-2$ jugadores restantes (quitando a 1 y su rival).

Ejercicio 234 . Demuestre que si n dados idénticos se lanzan, hay $\binom{n+5}{5}$ posibles resultados.¹⁴

Ejercicio 235 . En el parlamento de cierto país hay 201 asientos, y tres partidos políticos. ¿De cuántas maneras se pueden dividir estos asientos de tal manera que ningún partido tenga asegurada la mayoría (es decir, que ninguno tenga más de la mitad de los asientos).

Ejercicio 236 . Muestre que n pelotas idénticas se pueden colocar en r cajas etiquetadas ($n \geq r$) de tal manera que ninguna caja quede vacía, de $\binom{n-1}{r-1}$ maneras.

Ejercicio 237

1. ¿Cuántos triángulos se pueden dibujar de tal manera que todos sus vértices sean vértices de un polígono con n lados y tales que todos sus lados sean diagonales del polígono?
2. Demuestre que el número de k -gonos que se pueden dibujar de esta manera es

$$\frac{n}{k} \binom{n-k-1}{k-1}.$$

Ejercicio 238 . Cada arreglo de equis en cuatro cajas codifica una cadena creciente de 1, 2, 3, y 4. Por ejemplo:



codifica a 122244 (escriba una vez 1, escriba tres veces 2, escriba cero veces 3, escriba dos veces 4).

¹⁴Imagínese que se toman k_1 dados con el número 1, después agréguese un dado marcado con una x , para separar los dados con 1 de los dados con 2, después de los k_2 dados con dos, agréguese otro dado marcado con x , para separar los dados con 2 de los dados con 3, y así sucesivamente. Tenemos ahora los n dados originales más cinco dados separadores. Un resultado posible al lanzar los n dados se toma escogiendo las posiciones de los cinco dados separadores entre los $n+5$ dados. Es decir, $\binom{n+5}{5}$.

Demuestre que el número de palabras crecientes de longitud n formadas de un alfabeto con m letras es

$$\binom{m+n-1}{n}.$$

(Una palabra es creciente si sus letras, excepto por repeticiones, aparecen en orden alfabético, por ejemplo *abbbbccdeeff*).

Ejercicio 239 . Muestre que el número de maneras de colocar x 1s y y 0s en línea sin que haya 1s adyacentes es

$$\binom{y+1}{x}.$$

Ejercicio 240 . Muestre que el número de r -subconjuntos de $\{1, 2, \dots, n\}$ que no contiene ningún par de enteros consecutivos es

$$\binom{n-r+1}{r}.$$

(Vea el ejercicio anterior).

Ejercicio 241 . Dados n conjuntos, el primero con a_1 elementos, el segundo con a_2 y así sucesivamente. Todos los elementos en todos los conjuntos son distintos. Demuestre que el número total de maneras de escoger una muestra de estos conjuntos que no tome más de un elemento de cada conjunto (puede que no se tome ningún elemento en algún conjunto) es:

$$(a_1 + 1)(a_2 + 1) \dots (a_n + 1).$$

Ejercicio 242 . Si n se factoriza en primos p_1, p_2, \dots en la forma

$$n = p_1^{\alpha_1} p_2^{\alpha_2} p_3^{\alpha_3} \dots p_k^{\alpha_k}$$

demuestre que el número de divisores de n es

$$(a_1 + 1)(a_2 + 1) \dots (a_k + 1)$$

Usando esta fórmula demuestre que un entero es un cuadrado si y sólo si tiene un número impar de divisores.

Ejercicio 243 . En mecánica estadística, se necesita contar el número de maneras en que r partículas se pueden meter dentro de n cajas bajo tres tipos de hipótesis:

- (Maxwell-Boltzmann) Las partículas son diferentes y cualquier número de ellas se puede meter dentro de una caja.
- (Bose-Einstein) Las partículas son idénticas y se puede meter cualquier número de ellas dentro de una caja.
- (Fermi-Dirac) Las partículas son idénticas pero no puede haber más de una dentro de una caja.

Encuentre el número de arreglos para cada caso.

Ejercicio 244 . Se tiene un sistema con cuatro partículas que satisface la hipótesis de Bose-Einstein. Cada partícula puede tener nivel de energía 0, E , $2E$, $3E$ o $4E$ pero la energía total del sistema es $4E$. Una partícula con energía nE puede ocupar cualquiera de $(n^2 + 1)$ estados diferentes de energía. ¿Cuántas configuraciones de estados de energía hay?

Ejercicio 245 . Repita el problema anterior con la hipótesis de Fermi-Dirac excepto porque ahora una partícula de energía nE puede ocupar cualquiera de $2(n^2 + 1)$ estados de energía diferentes, y no hay dos partículas que tengan el mismo estado de energía simultáneamente.

Ejercicio 246 . ¿De cuántas maneras se pueden escoger k subconjuntos de un conjunto con n elementos tales que sean ajenos dos a dos. Ilustre esto con un conjunto con tres elementos de donde se tomen parejas de subconjuntos ajenos.

Ejercicio 247 . Supongamos que tenemos una provisión ilimitada de pelotas con n colores y un número primo p . Demuéstrese que el número de arreglos diferentes de p pelotas en un círculo es

$$\frac{n^p - n}{p},$$

si quitamos los arreglos en los que todas las pelotas son del mismo color. Esto demuestra que p divide a $n^p - n$ (Pequeño Teorema de Fermat).

Ejercicio 248 . *En una tienda hay k clases de tarjetas postales. Queremos enviarlas a n amigos. ¿De cuántas maneras se puede hacer esto? ¿Qué sucede si queremos enviarles diferentes postales? ¿Qué sucede si queremos enviar dos postales diferentes a cada amigo, pero diferentes amigos pueden tener la misma postal?*

Ejercicio 249 . *Tenemos k tarjetas postales diferentes. Queremos enlárselas a n amigos (un amigo puede recibir cualquier número de tarjetas postales incluyendo como posibilidad a 0). ¿De cuántas maneras se puede hacer? ¿Qué sucede si queremos enviar por lo menos una postal a cada amigo?*

Ejercicio 250 . *¿Cuántos anagramas se puede formar de la palabra characterization? (Un anagrama de una palabra es otra palabra que tiene las mismas letras que la primera, apareciendo el mismo número de veces. No hace falta que tenga un significado, es decir que podría no estar en el diccionario).*

Ejercicio 251

1. ¿De cuántas maneras se pueden distribuir k monedas entre n personas de manera que cada persona reciba por lo menos una?
2. Si retiramos la condición de que cada persona reciba alguna moneda, ¿de cuántas maneras se puede hacer el reparto?

Ejercicio 252 . *Hay k clases de postales, pero solamente un número limitado de cada clase, hay a_i copias de la i -ésima. ¿De cuántas maneras se les pueden enviar a n amigos? (Podemos mandar más de una copia de la misma postal a la misma persona).*

Ejercicio 253 . *Mostrar que el número de particiones de un número n en exactamente m sumandos es igual al número de particiones de $n - m$ en no más de m sumandos.*

Ejercicio 254 . *Mostrar que el número de particiones de un número n en cualquier número de sumandos distintos es igual al número de particiones de n en sumandos impares.*

Ejercicio 255 . Tenemos n monedas. Cada día compramos exactamente uno de los siguientes productos: taco (1 moneda), tamal (2 monedas) refresco (2 monedas). ¿Cuál es el número B_n de maneras posibles de gastar todo el dinero?

Ejercicio 256 . ¿Cuál es el número A_n de maneras posibles de subir n escalones, si podemos subir un escalón o dos en cada paso?

Ejercicio 257 . ¿Cuántas sucesiones de longitud n se pueden hacer con a, b, c, d de tal manera que a y b nunca queden juntos?

Ejercicio 258 . ¿Cuál es el número de k -adas que se pueden escoger de

$$\{1, 2, \dots, n\}$$

que no contengan enteros consecutivos?

Ejercicio 259 . Queremos romper un palo de longitud n en n piezas de longitud 1. ¿Cuál es el número de maneras para hacer esto? si:

1. en cada paso, rompemos una de las piezas con longitud mayor que 1 en dos.
2. En cada paso rompemos cada una de las piezas de tamaño mayor que 1 en dos.

Ejercicio 260 . ¿De cuántas maneras se le pueden poner paréntesis al producto

$$x_1 x_2 \dots x_r$$

(de manera que cualquier paréntesis incluya un producto de dos factores).

Ejercicio 261 . ¿Cuál es el número D_n de triangulaciones de un n -gono convexo? (Una triangulación es un conjunto de $n - 3$ diagonales que no se cruzan y que por lo tanto dividen al n -gono en $n - 2$ triángulos).

Ejercicio 262 . ¿De cuántas maneras se puede dividir un n -gono convexo en triángulos usando $n - 1$ diagonales que no se crucen, de tal manera que cada triángulo tenga una arista en común con el n -gono convexo?

Ejercicio 263 . En el salón de clase de una preparatoria hay 30 estudiantes, a 12 de ellos les gustan las Matemáticas, a 14 les gusta la física y 13 la química, a 5 alumnos les gustan tanto la física como las Matemáticas, a 7 les gustan la física y la química a la vez y a 4 les gusta las Matemáticas y la química. Hay 3 a los que les gustan las tres materias. ¿A cuántos alumnos no les gusta ninguna materia?

Ejercicio 264 . Tenemos n alcancías, con llaves diferentes. Alguien cierra las alcancías, revuelve las llaves y deposita una llave dentro de cada alcancía. Rompemos k alcancías. ¿Cuál es la probabilidad de que podamos abrir las restantes, con las llaves que obtuvimos?

Ejercicio 265 . A lo largo de un circuito de carreras de autos hay gasolineras. La cantidad total de gasolina disponible es la misma que la que nuestro auto necesita para el recorrido. Demuestre que existe una gasolinera en donde se puede empezar con el tanque vacío, de tal forma que podemos completar el recorrido del circuito. Por ejemplo si sólo hay una gasolinera, entonces ésta contiene gasolina para completar todo el recorrido. Si hay dos gasolineras, es seguro que en una de ellas hay suficiente gasolina para llegar a la segunda (¿por qué?).

Ejercicio 266 . En el primer grupo de la clase «A» de un campeonato de futbol participan 17 equipos. Los premios son medallas de oro, de plata y de bronce. ¿De cuántas formas éstas pueden ser distribuidas?

Ejercicio 267 . Hay una sociedad científica formada por 25 personas. Es necesario elegir al presidente de la sociedad, al vicepresidente, al secretario científico y al tesorero. ¿De cuántas formas se puede efectuar esta elección, si cada miembro de la sociedad puede ocupar sólo un cargo?

Ejercicio 268 . ¿De cuántas formas se pueden colocar en el tablero de ajedrez 8 torres de modo que no se puedan comer una a la otra?

Ejercicio 269 . Un domador de fieras quiere sacar a la arena del circo 5 leones y 4 tigres. Un tigre no puede ir detrás de otro. ¿De cuántas maneras se pueden distribuir las fieras?

Ejercicio 270 . Se dan n objetos diferentes y k cajones. Hay que colocar n_1 objetos en el primer cajón, n_2 en el segundo, . . ., n_k en el k -ésimo, siendo

$$n_1 + n_2 + \dots + n_k = n.$$

¿De cuántas maneras se puede efectuar dicha distribución?

Ejercicio 271 . Tres niños juntaron 40 manzanas de un árbol. ¿De cuántas maneras pueden repartirlas, si todas las manzanas se ven iguales (es decir, si sólo nos interesa cuántas manzanas obtiene cada uno, y no cuáles manzanas le tocan).

Ejercicio 272 . ¿De cuántas maneras se pueden repartir 10 hongos blancos, 15 setas y 8 trufas entre 4 niños?

Ejercicio 273 . Quiero enviar a mi amiga 8 fotos distintas. ¿De cuántas maneras puedo hacerlo, utilizando 5 sobres diferentes?

Ejercicio 274

1. Por el envío de un paquete hay que pagar 18 monedas. ¿De cuántas formas se puede pagar con estampillas de valor de 4, 6 y 10 monedas, si dos formas que se distingan en el orden de las estampillas se consideran diferentes? La reserva de estampillas de distinto valor se considera ilimitada. La manera 4, 4, 10 se considera distinta de la manera 10, 4, 4.
2. Se dispone de estampillas de valores de n_1, n_2, \dots, n_h monedas. ¿De cuántas maneras se puede pagar con ellas una suma de N monedas, si dos formas que se distingan en el orden se consideran distintas?.

Ejercicio 275 . ¿Cuántos son los pares diferentes de números enteros x e y desde 1 hasta 1000, para los cuales

$$(x^2 + y^2)/49$$

es un número entero?. (Los pares $\{z, y\}$ y $\{y, z\}$ son iguales.)

Ejercicio 276 . Partiendo de los números desde 1 a n se han compuesto toda clase de productos, que constan de k factores diferentes (k fijo). ¿Cuántos de los productos obtenidos son divisibles por un número primo p , $p \leq n$?

Ejercicio 277 . En el campeonato de futbol de un país participan 20 equipos. ¿Cuál es el número mínimo de partidos que deben jugarse, para que entre cualesquiera tres equipos haya dos que ya hayan jugado entre sí?

Ejercicio 278 . Si se colocan en un tablero de ajedrez dos torres (negra y blanca). ¿Qué es lo más probable, que dichas torres pueden comerse una a la otra o que no puedan?

Ejercicio 279 . Un tablero de ajedrez de 6×6 de dimensión está cubierto con 18 fichas de dominó de 2×1 de dimensión (de tal modo que cada ficha cubre dos casillas). Demuéstrese que, cualquiera que sea la forma en que las fichas cubren el tablero, éste puede ser cortado en dos partes, horizontal o verticalmente sin perjudicar ninguna ficha de dominó.

Ejercicio 280 . Las casillas (escaques) de un tablero de ajedrez se han enumerado de la siguiente manera: la primera fila horizontal con los números del 1 al 8 de izquierda a derecha; la segunda fila horizontal, con los números del 9 al 16 de izquierda a derecha, etc. En el tablero están colocadas 8 torres de tal modo que no se ataquen una a la otra. ¿Qué valor puede tomar la suma de números de los escaques en los cuales están colocadas las torres?

Ejercicio 281 . En cada escaque de un tablero de ajedrez de dimensión $n \times n$ se ha puesto un número que indica la cantidad de rectángulos que contienen dicho escaque. ¿Cuánto es la suma de todos los números escritos?

Ejercicio 282 . Demuéstrese que de cualesquiera cinco hongos que crecen en un bosque y que no están dispuestos en una recta, siempre hay cuatro hongos tales que sirven de vértices de un cuadrilátero convexo.

Ejercicio 283 . Cierta comisión se reunió 40 veces. Cada vez había en las sesiones 10 miembros. Además, cualesquiera dos miembros de la comisión presenciaron juntos una sesión a lo más una vez. Demuéstrese que el número de miembros de la comisión es superior a 60.

Ejercicio 284 . En una oficina hay 25 empleados. Demuéstrese que de dichos empleados no pueden formarse más de 30 comisiones con 5 personas en cada una, de tal modo que ningún par de comisiones tenga más de un miembro común.

Ejercicio 285 . *Para pintar una cara de un cubo hacen falta 5 segundos. ¿Cuál es el tiempo mínimo en el transcurso del cual 3 hombres pueden pintar 188 cubos? (Se supone que dos hombres no pueden pintar simultáneamente un cubo).*

Ejercicio 286 . *En un campeonato gimnástico dos equipos contaban con un número igual de participantes. En total la suma general de tantos obtenidos por todos los participantes era igual a 156. ¿Cuál fue el número de participantes, si cada uno de ellos obtuvo notas sólo de 8 y 9 tantos?.*

Ejercicio 287 . *Se tienen 9 palos de diferente longitud desde 1 hasta 9 cm. ¿De cuántos métodos y con qué lados pueden formarse cuadrados; haciendo uso de dichos palos? (No es obligatorio que se usen todos los palos; los métodos de formación de un cuadrado se consideran diferentes, si se emplean palos distintos.)*

Ejercicio 288 . *¿Qué hay más entre el primer millón de números: aquellos en cuya notación se encuentra 1, o aquellos en cuya notación 1 no está contenido?*

Ejercicio 289 . *Demuéstrese que si una partida de ajedrez se desarrollara con un número infinito de jugadas, entonces*

1. existiría por lo menos una posición se repetiría un número infinito de veces.
2. Demuéstrese que existiría una sucesión de jugadas de longitud tan grande como se quiera, que se repetiría un número infinito de veces.

Ejercicio 290 . *Demuéstrese que en cada fracción decimal infinita existe una sucesión de signos decimales de longitud arbitraria, que en el desarrollo de la fracción se encuentra un número infinito de veces.*

Ejercicio 291 . *¿Cuántos ceros tiene al final el número $((3!)!)!$ en notación decimal?*

Ejercicio 292 . *Los incisos siguientes se refieren a los números enteros del 5 al 200, incluyendo ambos.*

1. ¿Cuántos hay?

2. ¿Cuántos números pares hay?
3. ¿Cuántos números impares hay?
4. ¿Cuántos son divisibles entre 5?
5. ¿Cuántos son mayores que 72?
6. ¿Cuántos contienen dígitos diferentes?
7. ¿Cuántos contienen el dígito 7?
8. ¿Cuántos no contienen el dígito 0?
9. ¿Cuántos son mayores que 101 y no contienen el dígito 6?
10. ¿Cuántos contienen dígitos en orden estrictamente creciente? (Algunos ejemplos son 147, 8.)
11. ¿Cuántos son de la forma xyz , en donde $0 \neq x < y$ y $y > z$?

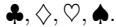
Ejercicio 293

1. ¿De cuántas maneras pueden cinco personas cumplir años en meses diferentes?
2. ¿Cuántas posibilidades hay para los meses de cumpleaños de cinco personas?
3. ¿De cuántos modos pueden al menos dos de las cinco personas cumplir años en el mismo mes?

Ejercicio 294

1. ¿Cuántas manos de bridge hay que tengan exactamente dos palos?¹⁵

¹⁵Una mano de bridge consta de 13 cartas. La baraja consta de 52 cartas repartidas en cuatro palos o bazas:



Estos palos se llaman tréboles, diamantes, corazones y espadas, respectivamente. Por ejemplo hay 13 cartas de tréboles: $A, 2, \dots, 10, J, Q, K$. Nótese que J funciona como el 11, Q como el 12 y K como el 13. Para algunos propósitos, por ejemplo en el póker, A funciona como 14 (“el as mata al rey”) o como 1. Una “corrida” puede tener los números $A, 2, 3, 4, 5$, pero también $10, J, Q, K, A$ es una corrida.

2. ¿Cuántas manos de bridge hay que tengan los cuatro ases?
3. ¿Cuántas manos de bridge hay que tengan cinco naipes de espadas, cuatro de corazones, tres de tréboles y uno de diamantes?
4. ¿Cuántas manos de bridge hay que tengan cinco cartas del mismo palo, cuatro de otro, tres del tercero y una del cuarto?
5. ¿Cuántas manos de bridge hay que tengan cuatro cartas de cada una de tres figuras diferentes y una de la otra?
6. ¿Cuántas manos de bridge hay que no tengan cartas de cara? (Una carta de cara es una con la denominación J, Q, K).

Ejercicio 295 . ¿De cuántas maneras pueden colocarse seis llaves diferentes en una argolla? (Voltear la argolla no cuenta como un arreglo diferente.)

Ejercicio 296

1. ¿De cuántas maneras pueden formar una fila cinco rusos y cinco chinos? (Se considera que todos son diferentes, es decir distinguimos un ruso de otro).
2. ¿De cuántos modos pueden formar una fila cinco rusos y cinco chinos si dos rusos no pueden estar juntos?
3. ¿De cuántos modos cinco rusos y cinco chinos pueden sentarse alrededor de una mesa circular?
4. ¿De cuántas maneras cinco rusos y cinco chinos pueden sentarse alrededor de una mesa circular si dos rusos no pueden estar juntos?
5. ¿De cuántas maneras pueden formar una fila cinco rusos y ocho chinos si dos rusos no pueden estar juntos?
6. ¿De cuántos modos cinco rusos y ocho chinos pueden sentarse alrededor de una mesa circular si dos rusos no pueden estar juntos?

Ejercicio 297 . En los incisos siguientes se supone que hay 20 bolas: 6 rojas, 6 verdes, 8 moradas.

1. ¿De cuántas maneras pueden seleccionarse cinco bolas si todas las bolas se consideran distintas? (por ejemplo si estuvieran numeradas del 1 al 20).
2. ¿De cuántas maneras pueden seleccionarse cinco bolas si las del mismo color se consideran idénticas?
3. ¿De cuántas maneras pueden seleccionarse dos rojas, tres verdes y dos moradas, si las bolas del mismo color son distintas? (por ejemplo las rojas están numeradas del 1 al 6, las verdes lo mismo y las moradas del 1 al 8).
4. Se sacan cinco bolas, se las repone y se vuelven a sacar cinco. ¿De cuántas maneras es posible hacer lo anterior si las bolas se consideran distintas?
5. Se sacan cinco pelotas y, sin reponerlas, se vuelven a sacar cinco de ellas. ¿De cuántas maneras es posible hacer lo anterior si las bolas se consideran distintas?
6. Se sacan cinco bolas, con al menos una roja y se las repone; se vuelven a sacar otras cinco con a lo sumo una verde. ¿De cuántas maneras es posible hacer la anterior si las bolas se consideran distintas?
7. Se sacan cinco bolas con al menos una roja, y sin reponerlas, se vuelven a sacar otras cinco con a lo sumo una verde. ¿De cuántas maneras es posible hacer lo anterior si las bolas se consideran distintas?

Ejercicio 298 . ¿De cuántas maneras se pueden repartir 15 ejemplares idénticos de un libro de Matemáticas entre seis estudiantes?

Ejercicio 299 . ¿De cuántas maneras se pueden repartir 10 libros diferentes entre tres estudiantes si el primer estudiante debe tener cinco libros, el segundo tres y el tercero dos?

Ejercicio 300 . Los incisos siguientes se refieren a tres pilas de pelotas iguales rojas, azules y verdes, donde cada pila contiene al menos diez bolas.

1. ¿De cuántas maneras se pueden seleccionar diez pelotas?

2. ¿De cuántas maneras se pueden seleccionar diez pelotas, si debe tenerse, al menos, una bola roja?
3. ¿De cuántos modos se pueden elegir diez bolas, si debe tenerse, al menos, una pelota roja, dos azules y tres verdes?
4. ¿De cuántos modos se pueden elegir diez bolas, si debe tenerse, exactamente, una pelota roja?
5. ¿De cuántas maneras pueden seleccionarse diez pelotas, si debe tenerse, exactamente, una pelota roja y, al menos, una azul?
6. ¿De cuántas maneras pueden seleccionarse diez pelotas, si debe tenerse, a lo sumo, una roja?
7. ¿De cuántas maneras se pueden seleccionar diez pelotas, si el número de bolas rojas debe ser el doble del número de pelotas verdes?

Capítulo 6

El campo de los números reales

6.1 Consideraciones generales

No parece exagerado decir que el conjunto \mathbb{R} de los números reales es una de las columnas principales en que se apoya el conocimiento matemático moderno, cuyas raíces se remontan hasta la Grecia antigua. En efecto, una parte fundamental del desarrollo teórico de la Matemática fue el descubrimiento, atribuido a los pitagóricos, de la incommensurabilidad de segmentos en la recta, lo que evidenció que la colección de magnitudes de estos, es más numerosa que el conjunto \mathbb{Q} de los números racionales.

Quizá la primera de tales incommensurabilidades fue la relación entre la magnitud de la diagonal de un cuadrado y la longitud de cualquiera de sus lados, hallazgo que planteó graves contradicciones en la concepción pitagórica de la recta, y aunque Eudoxio (408, 355 a.C.) desarrolló su “teoría sobre los segmentos de recta incommensurables”, que es el análogo geométrico de la construcción de las cortaduras de Dedekind -a la que precedió por más de 2000 años-, para los griegos, los conceptos de longitud y área eran puramente geométricos. No olvidar que ellos separaban tajantemente la Geometría de la Aritmética que “contaminaba al idealismo platónico de los objetos geométricos”. El concepto de número real no fue un concepto griego sino que apareció mucho después -en el Renacimiento- como resultado de la notación decimal- más o menos 1600-. Esta notación, que se utiliza actualmente en los cursos elementales, y que resulta ventajosa para un buen número de aplicaciones de la Matemática, tanto prácticas como teóricas, tiene sin embargo el grave problema de que la construcción rigurosa de \mathbb{R} en términos de sucesiones

decimales infinitas conduce a complicaciones técnicas muy inconvenientes. Considere por ejemplo la definición de suma o de producto aunado al hecho de no contar con representaciones explícitas -ni siquiera explícitamente definidas-.

La falta de claridad de los conceptos básicos del Análisis no impidió el desarrollo adecuado de las teorías matemáticas que sin embargo, no podían explicar razonablemente la validez de las operaciones que realizaban. Tanto Newton como Leibniz -considerados por muchos como los padres del Cálculo- fracasaron en su intento de explicar sus resultados. Ellos y sus discípulos al tratar de justificar sus procedimientos, sólo aumentaron la confusión, lo que era lógico que sucediera, dado que sus argumentos se apoyaban en unos misteriosos “infinitésimos”, que -como veremos después- son inexistentes en el campo ordenado de los números reales.

Un avance importante en la conceptualización del Análisis fue la hipótesis sobre la existencia de un isomorfismo entre los puntos de la recta y el conjunto de los números reales. Idea que utilizó provechosamente Descartes (1596-1650) para relacionar el Álgebra con las curvas geométricas.

La conveniencia de contar con un campo que permitiera fundamentar razonablemente todo el edificio que la Matemática había construido apoyándose en las ideas intuitivas -vagas- de lo que podía ser el conjunto de los números reales, obligó a los matemáticos de la época -último cuarto del siglo XIX- a proponer modelos de extensiones de \mathbb{Q} que resultaran adecuados para sus propósitos, y así surgieron varias construcciones, cada una de las cuales, con sus ventajas y desventajas, condujo a la misma estructura abstracta del “continuo de los números reales” esto es, cada una proporcionó un modelo de un *campo arquimedianamente ordenado y completo* a pesar de que en cada caso, los números reales resultaron de una naturaleza diferente. Queremos enfatizar aquí, 2 cosas:

1. Para el usuario común de la Matemática, lo que en definitiva importa, es el conjunto de propiedades que tienen las operaciones y el orden de \mathbb{R} y no lo que cada número real pueda ser, esto justifica plenamente el uso de los modelos axiomáticos para \mathbb{R} en los que tanto el conjunto como sus elementos resultan conceptos primitivos -en el sentido de que no se definen- y a los que se asignan las propiedades deseadas a través de los postulados de la teoría, (que son definiciones implícitas de los números y de sus relaciones). Recuérdese que -como ya se dijo- la Matemática se desarrolló considerablemente a pesar de no contar con una base

lógicamente correcta en que apoyarse. (La Teoría de las ecuaciones, la Geometría analítica y el Cálculo; la Geometría diferencial, las Ecuaciones diferenciales ordinarias y en derivadas parciales no necesitaron saber a ciencia cierta qué eran en realidad los números sobre los que trabajaban para desarrollarse en la forma en que lo hicieron).

2. Se puede demostrar que los diferentes modelos que se obtienen de las construcciones que hemos mencionado, son *isomorfos*, es decir, que se puede definir entre cualesquiera dos de ellos una biyección bien comportada con respecto a la estructura - operaciones y orden - de modo que, desde el punto de vista de sus propiedades algebraicas, cada modelo es indistinguible de cualquiera de los otros, por lo que desde esta perspectiva puede considerarse que existe un único campo arquimediano ordenado y completo, al que llamaremos el campo \mathbb{R} de los números reales y así afirmaremos que la definición axiomática de \mathbb{R} es *categórica*¹.

6.2 Construcción de \mathbb{R} a partir de las cortaduras en \mathbb{Q}

Procedamos ahora a ordenar, de alguna manera las consideraciones anteriores.

Supongamos que en una recta ℓ hemos marcado los puntos P y Q que corresponden a los racionales p y q respectivamente, (lo que, según se verá más adelante, siempre es posible). Entonces a

$$r_1 = \frac{p+q}{2}$$

le corresponde precisamente el punto medio del segmento \overline{PQ}

Podemos considerar ahora los números,

$$r_2 = \frac{p+r_1}{2}, \quad r_3 = \frac{r_1+q}{2},$$

marcarlos en la recta y proseguir así, indefinidamente el proceso de “tomar mitades”. Es obvio que en esta forma podemos intercalar -teóricamente- en

¹Una teoría axiomática T es *categórica* si tiene esencialmente un solo modelo excepto por isomorfismos, e.d., si T es *categórica*, dos modelos cualesquiera para T , resultan algebraicamente indistinguibles.

\overline{PQ} tantos puntos correspondientes a números racionales, como deseemos, y por lo tanto demostrar que en cada segmento de extremos racionales, sin importar que tan pequeña pueda ser su longitud, existe un número infinito, de estos números. Esta observación y la propiedad arquimediana del orden natural en \mathbb{Q} , aseguran que los racionales se extienden densamente a lo largo de toda la recta ℓ .

Es fácil demostrar que si p es un punto racional y ε es un “pequeño número” mayor que cero, existe una infinidad de racionales cuya distancia a p es menor que ε . Es decir podemos afirmar que la distancia de p a su complemento en \mathbb{Q} , es cero. Esta sobre población de racionales podría hacernos creer que en la recta ordenada ℓ , una vez colocados todos ellos, no queda lugar para punto adicional alguno. Sin embargo, la construcción geométrica de la diagonal de un cuadrado de lados de medida 1, que de acuerdo con el teorema de Pitágoras debe tener una longitud cuyo cuadrado es 2, muestra que si la recta “no tiene agujeros”, entonces, después de haber asignado lugares para cada $p \in \mathbb{Q}$ debe quedar alguna vacante, y, en un sentido muy preciso, puede demostrarse que en ℓ , la cardinalidad de las “vacantes” es mucho mayor que la de los “lugares ocupados”. De hecho puede demostrarse que entre cualesquiera 2 números racionales, existe una infinidad no numerable de puntos no ocupados.

Cuando Euclides construyó su Geometría, utilizó, -sin mención explícita- la propiedad de la recta de “ser completa” pensando que esto era “intuitivamente claro”. Para corregir esta omisión, Hilbert, en su axiomática para la geometría del plano, postula los *Axiomas de continuidad (grupo V)*, que garantizan que:

1. “Los puntos de una recta constituyen un sistema tal que no puede asignarse ningún punto nuevo a ella, sin que se viole al menos uno de nueve postulados anteriores”.
2. Después de definir, para los puntos de un segmento \overline{AB} A “el origen” y B “el extremo”, la relación “ P precede a Q ” ó “ Q sigue de P ”, y decidir que el segmento \overline{AB} con esa relación se llama un “segmento ordenado”, el postulado asegura que si los puntos de un segmento ordenado de origen A y extremo B se separan en dos clases de manera que:
 - (a) Cada punto de \overline{AB} pertenezca a una y sólo una de las clases,
 - (b) Los puntos A y B pertenezcan a clases distintas (que llamó respectivamente, la primera y la segunda clase) y

- (c) Cada punto de la primera clase precede a cada punto de la segunda.

Entonces existe un punto C sobre \overline{AB} tal que todo punto de \overline{AB} que preceda a C estará en la primera clase y todo punto de \overline{AB} que siga de C estará en la segunda.

Apuntaremos que con estos axiomas, puede probarse con todo rigor, que la recta está completa, es decir, que en toda recta orientada ℓ , a la medida de cualquier segmento, le corresponde un punto (y sólo uno) de ℓ y que no puede “transitarse continuamente” de un lado a otro de ella sin pasar por alguno de sus puntos. Esta última observación -que precisaremos más adelante- justifica la existencia de los puntos que Euclides supuso en sus construcciones:

1. Con centro O y radio r trácese con el compás una circunferencia,...y
2. ...sea P el punto de intersección de ésta con el rayo \overrightarrow{OQ} ,...

Observaciones:

1. Existe una infinidad de racionales positivos, cuya raíz cuadrada no está en \mathbb{Q} .
2. Suponiendo que ℓ es completa, en ella se puede construir geométricamente el punto que corresponde a la raíz cuadrada (positiva) de todo número racional $a > 0$.

Corolario 7 . *El conjunto \mathbb{Q} de los números racionales “no llena” completamente a la recta.*

Justificaremos nuestra primera observación demostrando que los únicos enteros positivos que tienen raíces n -ésimas racionales, son los que corresponden a las n -ésimas potencias de los naturales: De este modo si n es igual a 2, se puede ver que sólo tienen raíz cuadrada racional los enteros que son elementos de

$$\beta = \{1, 4, 9, \dots, m^2, \dots\}$$

y que por lo tanto, basta considerar

$$\mathbb{Z}^+ \subset \mathbb{Q}^+$$

para encontrar un conjunto infinito de racionales positivos cuya raíz no está en \mathbb{Q} .

Teorema 87 . *Si a, b, c y n son enteros positivos tales que $(a, b) = 1$, $(a/b)^n = c$, entonces $b = 1$.*

Demostración. Sean a, b, c , y n enteros positivos que satisfacen la hipótesis del teorema. Por lo tanto,

$$a^n = cb^n.$$

Ahora bien, si b es distinto de 1, entonces existe p un número primo tal que $p \mid b$ es decir $p \mid a^n$ luego $p \mid a$. Y por lo tanto, $p \mid (a, b) \nmid 1$. Este absurdo se obtuvo de suponer $b \neq 1$, luego $b = 1$, y $a/b \in \mathbb{Z}$. ■

Procedamos ahora a justificar la observación 2.

Algoritmo 1 . *Como dibujar la raíz cuadrada (positiva) de un número racional $a > 0$.*

1. Sea ℓ una recta cualquiera, en la que asignamos un punto 0 al cero y marcamos una escala con un punto $U \neq 0$ -que orienta a ℓ - y convenimos que en adelante la longitud OU será nuestra unidad de medida ($OU = 1$). Debe quedar claro que con el uso iterado del compás podemos construir en ℓ , segmentos de longitud m para cada entero positivo m y por lo tanto, cambiando la orientación cuando sea el caso, también podemos encontrar los puntos correspondientes, en ℓ , a cada entero, positivo o negativo. Estamos ahora en posibilidad de dibujar el punto que corresponde a cualquier número racional

$$a = p/q,$$

(obviamente bastará considerar el caso en que tanto p como q son enteros positivos).

En efecto, designemos por P al punto de ℓ que corresponde a p ($OP = p$) y tracemos por O cualquier rayo \overrightarrow{OQ} que no esté contenido en ℓ . Marquemos ahora en \overrightarrow{OQ} los puntos Q_q y Q_1 tales que $OQ_q = q$, $OQ_1 = 1$. ($\overrightarrow{OQ_1} \sim \overrightarrow{OU}$).

Tracemos ahora la recta $\overline{Q_qP}$ y por Q_1 una paralela a ella. Si denotamos por P_1 al punto de intersección con esta última paralela, es inmediato que

$$OP_1 = p/q$$

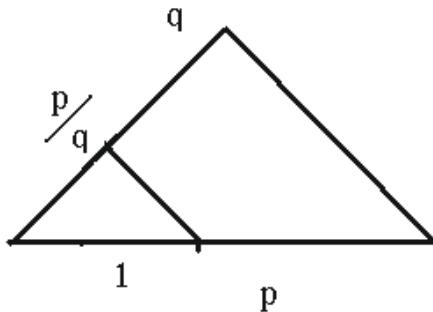


Figura 6.1:

y que por lo tanto P_1 es la representación geométrica de p/q . (ver fig). La discusión anterior muestra que es posible asignar a cada $a \in \mathbb{Q}$ un punto $P_a \in \ell$ y como $OP_1 = |a|$, resulta que la asignación es inyectiva (a números racionales distintos, corresponden puntos de ℓ distintos).

2. Sea $a \in \mathbb{Q}^+$. En la recta orientada ℓ constrúyase el segmento \overline{OP} , de longitud $a + 1$ y márquese el punto Q que corresponde a a .

Sea M el punto medio de \overline{OP} y trácese la circunferencia con centro en M y radio $\frac{a+1}{2}$.

Por el punto Q , constrúyase la perpendicular a ℓ y sea R el punto de intersección de ésta con la circunferencia C . Observe que \widehat{ORP} es un ángulo recto (subtiende un semicírculo) y que, por lo tanto

$$\overset{\triangle}{OQR} \sim \overset{\triangle}{RQP}$$

(En cada uno de ellos, tanto \widehat{QOR} como $\widehat{QR P}$ son complemento de \widehat{OPR} por lo tanto son congruentes ... etc.)

Entonces

$$\frac{OQ}{QR} = \frac{QR}{QP} \therefore (QR)^2 = OQ \cdot QP = a \quad (QP = 1)$$

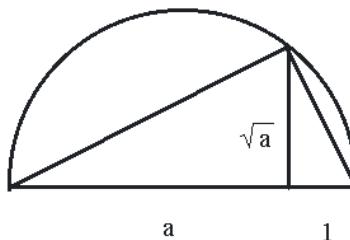


Figura 6.2:

Es decir, la longitud del segmento \overline{QR} es la raíz cuadrada de a . Es inmediato que con el compás podemos construir, en ℓ , el segmento \overline{OA} congruente con \overline{QR} . $A \in \ell$ es el punto que corresponde a \sqrt{a} y dado que este último número fue escogido arbitrariamente, podemos ahora asegurar que en ℓ están todos los puntos que corresponden a cada raíz cuadrada de cualquier número racional positivo, independientemente de que tal raíz pueda ser racional o no.

Nótese que en todo el argumento anterior hemos supuesto -como Euclides- que la intersección de cada arco trazado con el compás con cada recta que el arco corta, es no vacío (la recta no tiene agujeros), pero que esta hipótesis puede justificarse plenamente utilizando los axiomas de Hilbert para el plano, como mencionamos previamente.

La discusión anterior reafirma que \mathbb{Q} es insuficiente para medir los diferentes segmentos de la recta.

Dedekind, utilizando como guía la concepción geométrica de Descartes que postula una biyección entre los puntos de ℓ y los números reales, utilizó al conjunto \mathbb{Q} de los números racionales para construir su modelo para \mathbb{R} , en el que cada número real resulta ser el resultado de “cortar” la recta racional.

Ejercicio 301 . ¿Para cuáles de los siguientes valores de k se puede construir un rectángulo con lados de longitud entera y cuya diagonal mida \sqrt{k} : $k = 2, 3, 4, 5, 6, 7, 8, 9, 10$?

6.3 Cortaduras de Dedekind

Daremos a continuación una descripción del modelo de Dedekind, y de la manera en que se definen en él las operaciones de suma y producto, así como la relación de orden canónica.

Definición 80 . *Una cortadura en \mathbb{Q} es una pareja de subconjuntos de \mathbb{Q} , (A, B) tal que:*

1. $A \neq \emptyset, B \neq \emptyset$.
2. $A \cup B = \mathbb{Q}, A \cap B = \emptyset$.
3. $\forall a \in A, \forall b \in B, a < b$.

Ejemplo 121 . *Si $A = \{p \in \mathbb{Q} \mid p < 0\}$ y $B = \{q \in \mathbb{Q} \mid q \geq 0\}$, entonces (A, B) es una cortadura.*

(Nótese que en este caso, A no tiene máximo, y B tiene mínimo 0)

Ejemplo 122 . $A = \{p \in \mathbb{Q} \mid p \leq 2\}, B = \{q \in \mathbb{Q} \mid p > 2\}$.

En este ejemplo A tiene máximo, pero B no tiene mínimo. Obviamente no puede darse el caso de que teniendo A máximo p , B tuviera mínimo q , ya que entonces

$$r = \frac{p+q}{2}$$

satisface

$$p < r < q$$

y por tanto no puede estar ni en A ni en B , violando la condición 2 de la definición. Pero sí puede suceder -sucede- que ni A tenga máximo ni B mínimo. En este caso se dirá que (A, B) define una “hendidura” (“hueco”, “rajada”, “hiato”) en \mathbb{Q} .

Ejemplo 123 . *Un ejemplo de esta situación es el siguiente.*

$$B = \{p \in \mathbb{Q}^+ \mid p^2 > 2\}, A = \mathbb{Q} \setminus B.$$

Nótese que A consta de todos los racionales negativos más el cero y de los positivos cuyo cuadrado es menor que 2. (el hecho de que para ningún racional r se da el caso de que $r^2 = 2$, justifica el uso del “menor estricto” en la caracterización de A).

Teorema 88 . En el ejemplo anterior, B no tiene mínimo, ni A máximo.

Demostración. En efecto, $\forall p \in B, \exists q \in B$ tal que $q < p$, y $\forall r \in A, \exists t \in A$ tal que $r < t$:

1) Sea $p \in B$ ($\therefore p \in \mathbb{Q}^+ \wedge p^2 > 2$, luego

$$\frac{p^2 - 2}{2p}$$

es un racional positivo.

Tómese ahora $h \in \mathbb{Q}$ tal que

$$0 < h < \min \left\{ \frac{p^2 - 2}{2p}, p \right\}$$

(por ejemplo, la mitad del que sea menor de ellos).

Entonces:

$$h < \frac{p^2 - 2}{2p}, \text{ luego } 2ph < p^2 - 2$$

$$\therefore 2 < p^2 - 2ph < p^2 + 2ph = (p - h)^2$$

y como $p - h \in \mathbb{Q}^+$ (ya que $h < p$), entonces

$$(p - h \in B) \wedge (p - h < p).$$

Demostraremos ahora que A no tiene máximo.

En efecto, sea $p \in A$, Observe que si p fuera $p \leq 0$, entonces $\frac{1}{2} \in A$ y $p < \frac{1}{2}$

Sea pues $p \in \mathbb{Q}^+, p^2 < 2, \therefore \frac{2 - p^2}{2p + 1} \in \mathbb{Q}^+$.

Nuevamente consideramos

$$0 < h < \min \left\{ \frac{2 - p^2}{2p + 1}, 1 \right\}$$

Y entonces: $2 - p^2 > (2p + 1)h = 2ph + h$.

$$\therefore 2 > p^2 + 2ph + h > p^2 + 2ph + h^2 = (p + h)^2$$

pues $h < 1 \Rightarrow h^2 < h$.

$$\therefore (p + h) \in \mathbb{Q}$$

y obviamente $p < p + h$. ■

Definidas las cortaduras se vio que conservar la pareja (A, B) era redundante. Pues en vista de que $A \cup B = \mathbb{Q}$, cada conjunto define bien a su complemento. Por lo que se optó por conservar solamente uno de los conjuntos, quedando la elección subordinada al gusto del elector. Algunos prefieren trabajar con las “cortaduras inferiores” que tienen la ventaja de asimilar el orden con la contención, sin embargo nosotros usaremos al conjunto B o “cortadura superior”, que aun cuando tiene el inconveniente de invertir “orden-contención”, lo que no es trascendente $-(\alpha \leq \beta \Leftrightarrow \beta \subset \alpha)$, permite definir la multiplicación de una manera más natural.

Operativamente se vio la conveniencia de eliminar al mínimo de B , en el caso de que lo hubiera, adjuntándolo a su complemento.

Con todo esto en mente procederemos a definir nuestras cortaduras.

Definición 81 . *Un conjunto $\alpha \subseteq \mathbb{Q}$ es una cortadura de Dedekind (en adelante sólo “cortadura”) si y sólo si:*

1. $\alpha \neq \emptyset$, $\alpha \neq \mathbb{Q}$.
2. α no tiene elemento mínimo.
3. $\forall p, q \in \mathbb{Q}$, $((p \in \alpha \wedge p < q) \Rightarrow q \in \alpha)$.

Llamaremos “número real” a cada cortadura y denotaremos con \mathbb{R} al conjunto de todas ellas.

Ejemplo 124 . *Son ejemplos de cortaduras:*

1. Para $r \in \mathbb{Q}$, se tiene que $r^* = \{p \in \mathbb{Q} \mid p > r\}$ es una cortadura. Esta cortadura se identificará posteriormente con r . De modo que si para cada racional p asignamos el real p^* , tendremos -como se verá después- una inmersión de \mathbb{Q} en \mathbb{R} , con lo que -como se ha hecho con anterioridad- podremos considerar a \mathbb{R} como una extensión de \mathbb{Q} .

2.

$$\sqrt{2} = \{p \in \mathbb{Q}^+ \mid p^2 > 2\}$$

Nótese que en este caso “ $\sqrt{2}$ ” se está usando para designar una cortadura que tiene la propiedad asociada con su nombre o sea:

$$(\sqrt{2})(\sqrt{2}) = 2.$$

(Ver ejercicio 312, en la página 356).

Sea \mathbb{R} el conjunto de las cortaduras superiores de \mathbb{Q} , entonces \mathbb{R} es “el campo arquimedianamente ordenado y completo de los números reales”, y justificaremos (parcialmente) lo correcto del nombre.

Motivada por la observación geométrica obvia, damos la primera definición:

Definición 82

1. Si α y β están en \mathbb{R} , se dirá que “ α es menor que β ” ($\alpha < \beta$) si y sólo si $\beta \subset \alpha$, y $\beta \neq \alpha$.
2. Agregamos: $\alpha \leq \beta$ si ($\alpha < \beta$) \vee ($\alpha = \beta$)).
3. $\alpha > \beta$ si $\beta < \alpha$.
4. $\alpha \geq \beta$ si $\beta \leq \alpha$.

Nótese que -como ya se apuntó- la relación \leq que hemos definido invierte la relación de contención. Además, siendo esta última una relación de orden, fuerza a que \leq también lo sea.

En efecto, $\forall \alpha, \beta, \gamma \in \mathbb{R}$, se cumple:

1. $\alpha \subset \alpha \therefore \alpha \leq \alpha$ (reflexividad).
2. $((\alpha \subset \beta) \wedge (\beta \subset \alpha)) \Rightarrow \alpha = \beta \therefore (\alpha \leq \beta) \wedge (\beta \leq \alpha) \Rightarrow \alpha = \beta$ (antisimetría).
3. $((\alpha \supset \beta) \wedge (\beta \supset \gamma)) \Rightarrow (\alpha \supset \gamma) \therefore (\alpha \leq \beta) \wedge (\beta \leq \gamma) \Rightarrow \alpha \leq \gamma$ (transitividad).

Además, no puede suceder que $\alpha \cap \beta = \emptyset$, ya que α y β no son vacíos y si $p \in \alpha$ y $q \in \beta$, entonces $\max\{p, q\}$ está necesariamente en ambos. Ahora bien, $\alpha = \beta$ ó existe $p \in \mathbb{Q}$, tal que p está en alguno de ellos y no en el otro, y en ese caso la cortadura en la que está p es menor que la otra. En efecto, supongamos que $p \in \alpha$ y $p \notin \beta$ y por lo tanto, $\forall q \in \beta$, $p < q$. Entonces $q \in \alpha$, lo que prueba que $\alpha \supset \beta$ o sea que $\alpha < \beta$.

En resumen, podemos decir que \leq “ordena linealmente a \mathbb{R} ”.

Para demostrar otras propiedades de \leq , (que es arquimédiano, compatible con $+$ y \cdot y completo en el sentido de Dedekind), introducimos en una estructura de *campo*, definiendo:

Definición 83 . Si $\alpha, \beta \in \mathbb{R}$,

$$\alpha + \beta = \{p + q \mid p \in \alpha, q \in \beta\}.$$

Definición 84 . Si $r \in \mathbb{Q}$,

$$r^* = \{p \in \mathbb{Q} \mid r < p\}$$

y señalamos los casos particulares:

$$0^* = \mathbb{Q}^+,$$

$$1^* = \{p \in \mathbb{Q} \mid 1 < p\}$$

Ejercicio 302 . Demuestre que Si $r, s \in \mathbb{Q}$, entonces $(r \neq s \Rightarrow r^* \neq s^*)$.

Definición 85 . $\mathbb{R}^+ = \{\alpha \in \mathbb{R} \mid 0^* < \alpha\}$, es decir: $\alpha \in \mathbb{R}^+ \Leftrightarrow \alpha \subset \mathbb{Q}^+$.

Para definir el inverso aditivo de una cortadura α , debemos encontrar la que está caracterizada por la propiedad de que $-\alpha + \alpha = 0^*$ y por lo tanto, si $p \in -\alpha$, debe suceder que

$$\forall q \in \alpha, p + q > 0$$

e. d. p debe ser mayor que $-q \forall q \in \alpha$. Pero nótese que esta condición no basta.

Si $\alpha = r^* = \{p \in \mathbb{Q} \mid p > r\}$ y si se define

$$\beta = \{q \in \mathbb{Q} \mid q > -p, \forall p \in \alpha\}.$$

entonces $q \in \beta$ si $q \geq -r$. Por lo tanto $-r \in \beta$ y es evidentemente su primer elemento (es decir es el menor). Luego β no es cortadura. Una manera de evitar esta complicación -eliminar el menor elemento de β - es definir:

Definición 86 . Si $\alpha \in \mathbb{R}$, entonces

$$-\alpha =: \{x \in \mathbb{Q} \mid \exists \delta_x \in \mathbb{Q}^+ \text{ tal que } x + q > \delta_x, \forall q \in \alpha\}.$$

Ejercicio 303 . $\forall r \in \mathbb{Q}, r^* \in \mathbb{R}$, y si $r, s \in \mathbb{Q}$, entonces $r^* + s^* = (r + s)^*$.

Notación 11 . Si (A, \leq) es un conjunto parcialmente ordenado y $B \subseteq A$, denotaremos por B^\uparrow el conjunto de cotas superiores de B .

Denotaremos por B_\downarrow el conjunto de las cotas inferiores de B .

Ejercicio 304 . Demostrar que $\forall r \in \mathbb{Q}$, el conjunto

$$\{r\}^* = \{s \in \mathbb{Q} \mid s > r\}$$

es una cortadura.

Definición 87 . Si α y β son cortaduras, entonces $\alpha \leq \beta \Leftrightarrow \beta \subseteq \alpha$.

Lema 19 . Si α es una cortadura, entonces para toda $\varepsilon > 0$, $\exists d \in \alpha$, $\exists x \notin \alpha$ tales que $d - x < \varepsilon$.

Demostración. Podemos tomar un elemento $d_1 \in \alpha$, y $x_1 \notin \alpha$, ya que como α es una cortadura, ni es vacía ni es todo \mathbb{Q} .

Consideremos ahora su promedio aritmético, es decir

$$\frac{d_1 + x_1}{2}.$$

Pasa una de las dos posibilidades:

$$\text{o bien } \frac{d_1 + x_1}{2} \in \alpha \text{ ó } \frac{d_1 + x_1}{2} \notin \alpha.$$

En el primer caso, hacemos $d_2 = \frac{d_1 + x_1}{2}$, $x_2 = x_1$,

$$\text{en el segundo caso, hacemos } x_2 = \frac{d_1 + x_1}{2}, d_2 = d_1.$$

Todavía tenemos que $d_2 \in \alpha$ y $x_2 \notin \alpha$, pero ahora

$$d_2 - x_2 = \frac{d_1 - x_1}{2}.$$

De la misma manera, podemos definir $d_3 \in \alpha$, $x_3 \notin \alpha$ de tal manera que

$$d_3 - x_3 = \frac{d_1 - x_1}{2^2},$$

inductivamente, podemos encontrar dos sucesiones:

$$d_1, \dots, d_n \in \alpha$$

$$x_1, \dots, x_n \notin \alpha$$

de tal manera que $d_n - x_n < \frac{d_1 - x_1}{2^{n-1}}$.

Es claro que existe un natural n tal que $\frac{d_1 - x_1}{2^{n-1}} < \varepsilon$. ■

Ejercicio 305 . Sean p y r dos racionales positivos, demostrar que existe $n \in \mathbb{N}$ tal que $\frac{p}{2^n} < r$.

Teorema 89 . Sean $\alpha, \beta, \gamma \in \mathbb{R}$, las siguientes afirmaciones se cumplen.

1. $\alpha + \beta \in \mathbb{R}$.
2. $\alpha + \beta = \beta + \alpha$.
3. $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$.
4. $(\alpha + 0^*) = \alpha$.
5. $\alpha \in \mathbb{R} \Rightarrow (-\alpha \in \mathbb{R} \wedge (\alpha + -\alpha = 0^*))$.
6. $(\mathbb{R}, +, 0^*)$ es un grupo abeliano² y note que

$$\alpha < 0^* \Leftrightarrow -\alpha > 0^*.$$

Demostración. 1) Necesitamos demostrar que $\alpha + \beta$ es una cortadura, si α y β lo son.

Como α es una cortadura, entonces existe $x \in \alpha$. Análogamente, existe $y \in \beta$. Por lo tanto $x + y \in \alpha + \beta$. Así tenemos que $\alpha + \beta \neq \emptyset$.

Por otra parte, si $a \notin \alpha$ y $b \notin \beta$, entonces a es una cota inferior para α (y como no es un elemento de α , a es estrictamente menor que cualquier elemento de α) y b es una cota inferior para β . De esta manera, $a + b$ es una cota inferior para $\alpha + \beta$. Como tenemos que $a + b$ es estrictamente menor que cualquier elemento de $\alpha + \beta$, concluimos que $\alpha + \beta \notin \alpha + \beta$.

Por lo tanto tenemos que $\alpha + \beta \neq \mathbb{Q}$.

Veamos ahora que $\alpha + \beta$ no puede tener un elemento menor. Si lo tuviera, digamos que $a \in \alpha$, $b \in \beta$ y que $a + b$ es el elemento menor de $\alpha + \beta$.

Sea $z \in \beta$, entonces $a + b \leq a + z$, pero de aquí tendríamos que $b \leq z$. Como esto sucedería para cualquier $z \in \beta$, entonces b sería el menor elemento de β . Contradicciendo que la cortadura β no tiene un elemento menor.

Por último, supongamos que $a + b < z$, con $a \in \alpha$ y $b \in \beta$. En este caso podemos escribir $z = a + (z - a)$. Es inmediato que $b < z - a$. Pero como β es una cortadura, tenemos que $z - a \in \beta$.

² $(G, +, 0)$ es un grupo abeliano si $+$ es una operación asociativa y commutativa con neutro 0 y donde cada elemnto tiene un inverso.

Por lo tanto $z = a + (z - a) \in \alpha + \beta$.

Hemos demostrado que $\alpha + \beta$ es una cortadura.

2 y 3) Que la suma de cortaduras es conmutativa y asociativa, se sigue de la definición y del hecho de que la suma de racionales es conmutativa y asociativa.

4) Demostraremos que $\{0\}^\dagger = \mathbb{Q}^+$ es el neutro para la suma de cortaduras.

Sea α una cortadura; por definición

$$\alpha + \mathbb{Q}^+ = \{d + r \mid d \in \alpha, r \in \mathbb{Q}^+\}.$$

$d < d + r \implies d + r \in \alpha$, ya que α es una cortadura. Por lo tanto $\alpha + \mathbb{Q}^+ \subseteq \alpha$.

Recíprocamente, si $d \in \alpha$, existe $d' \in \alpha$ tal que $d' < d$, ya que las cortaduras no tienen elemento menor. Entonces $d = d' + (d - d') \in \alpha + \mathbb{Q}^+$. Por lo tanto $\alpha + \mathbb{Q}^+ = \alpha$. Esto significa que $\mathbb{Q}^+ = \mathbb{O}$, es el cero real.

5) Afirmamos que

$$-\alpha = \{x \in \mathbb{Q} \mid \exists \delta_x > 0, \forall d \in \alpha, x + d > \delta_x\}$$

es el inverso aditivo de α .

Primero veremos que el conjunto anterior es una cortadura, y luego veremos que efectivamente tiene la propiedad de que sumado con α da \mathbb{O} .

El conjunto descrito no es vacío:

Sea $d \in \alpha$, $x \notin \alpha$. Como d no es el menor elemento de α , entonces existe $d' < d$ tal que $d' \in \alpha$. Entonces $x - (d - d') < x < d' < d$.

$\forall z \in \alpha, z > x$. Entonces

$$(d - d') - x + z > (d - d') - x + x = d - d' > 0, \quad \forall z \in \alpha$$

Por lo tanto $(d - d') - x \in -\alpha$.

El conjunto propuesto no es todo \mathbb{Q} :

Es muy fácil notar que $d \in \alpha \implies -d \notin -\alpha$:

$$d - d = 0 \not> 0, \forall \delta > 0.$$

Consecuentemente $-\alpha \neq \mathbb{Q}$.

Veamos ahora que $-\alpha$ no tiene elemento menor: supongamos que $x \in -\alpha$, entonces $\exists \delta > 0$ tal que

$$x + d > \delta, \forall d \in \alpha.$$

Pero entonces $x - \frac{\delta}{2} + d > \frac{\delta}{2}, \forall d \in \alpha$. Por lo tanto $x - \frac{\delta}{2} \in -\alpha$, y $x - \frac{\delta}{2} < x$.

Por último si $-\alpha \ni x < y$, entonces $\exists \delta > 0$ tal que

$$x + d > \delta, \forall d \in \alpha.$$

Pero también $y + d > x + d > \delta, \forall d \in \alpha$. Por lo tanto $y \in -\alpha$.

De lo anterior concluimos que el conjunto propuesto para ser el inverso aditivo de α es en efecto una cortadura. Veamos ahora que en efecto es el inverso aditivo de α .

Por la definición de $-\alpha$, es claro que $\alpha - \alpha \subseteq \mathbb{Q}^+$. (Sumar un elemento de $-\alpha$ con un elemento de α produce un racional positivo).

Recíprocamente, sea $r \in \mathbb{Q}^+$. Por el Lema 19, $\exists d \in \alpha$ y $x \notin \alpha$ tales que $d - x < r$. Como d no es el primer elemento de α , existe $d' \in \alpha$ tal que $d' < d$.

$x - (d - d') < x < z, \forall z \in \alpha$ (todo elemento de α es mayor que todo racional que no pertenezca a α).

Entonces $(d - d') - x + z > d - d', \forall z \in \alpha$. Por lo que $(d - d') - x \in -\alpha$.

Además $[(d - d') - x] + d' = d - x \in -\alpha + \alpha$.

Como $-\alpha + \alpha$ es una cortadura, $d - x < r \implies r \in \alpha - \alpha$. Por lo tanto $\mathbb{Q}^+ \subseteq \alpha - \alpha$.

De esta manera tenemos que $\alpha - \alpha = \mathbb{Q}^+ = \mathbb{O}$.

Es decir que $\{x \in \mathbb{Q} \mid \exists \delta_x > 0, \forall d \in \alpha, x + d > \delta_x\}$ es exactamente el inverso aditivo de α .

6) Por último, demostraremos que $\alpha > \mathbb{O} \Leftrightarrow -\alpha < \mathbb{O}$.

En general si uno tiene que λ, μ, ρ son cortaduras y $\mu \subseteq \rho$, (o lo que es lo mismo, $\rho \leq \mu$) entonces

$$\lambda + \mu \subseteq \lambda + \rho :$$

pues si $x \in \lambda$ y $y \in \mu$, entonces $x + y \in \lambda + \rho$, ($y \in \rho$). Por lo tanto $\rho \leq \mu \implies \lambda + \rho \leq \lambda + \mu$. (La implicación es válida para desigualdades estrictas: $\rho < \mu \implies \lambda + \rho < \lambda + \mu$).

Entonces $\alpha > \mathbb{O} \implies \mathbb{O} = \alpha - \alpha > \mathbb{O} - \alpha = -\alpha$. También sucede que $-\alpha < \mathbb{O} \implies \mathbb{O} < \alpha$. (Sólo hay que sumar α). ■

Podemos resumir la proposición anterior diciendo que $(\mathbb{R}, +, \mathbb{O})$ es un grupo abeliano.

Observación 82

$$1. \alpha > \mathbb{O} \Leftrightarrow \mathbb{Q}^+ = \mathbb{O} \supsetneq \alpha \Leftrightarrow \exists r \in \mathbb{Q}^+ \setminus \alpha.$$

2. $\alpha < \mathbb{O} \Leftrightarrow \mathbb{Q}^+ = \mathbb{O} \subsetneq \alpha \Leftrightarrow \Leftrightarrow \exists d \in \alpha \setminus \mathbb{Q}^+ \Leftrightarrow \exists d \in \alpha, d \geq 0 \Leftrightarrow 0 > d' \in \alpha.$
3. $\alpha > \mathbb{O} \Leftrightarrow -\alpha < \mathbb{O}.$

Ejercicio 306 . *Demostrar la observación anterior.*

6.4 El producto en \mathbb{R}

Definición 88 . *Definimos $* : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$ de la manera siguiente:*

1. Si α_1, α_2 son $\geq \mathbb{O}$, entonces

$$\alpha_1 * \alpha_2 = \{xy \mid x \in \alpha_1, y \in \alpha_2\}.$$

2. Si $\alpha_1 \geq \mathbb{O}, \alpha_2 < \mathbb{O}$

$$\alpha_1 * \alpha_2 = -(\alpha_1 * (-\alpha_2)),$$

3. Si $\alpha_1 < \mathbb{O}, \alpha_2 \geq \mathbb{O}$

$$\alpha_1 * \alpha_2 = -((- \alpha_1) * \alpha_2),$$

4. Si $\alpha_1 < \mathbb{O}, \alpha_2 < \mathbb{O}$

$$\alpha_1 * \alpha_2 = (-\alpha_1) * (-\alpha_2).$$

Ejercicio 307 . *Use la observación 82 para demostrar que un producto de reales positivos es positivo y que el producto de un real positivo por un real negativo es negativo.*

Ejercicio 308 . *Demuestre que $(-\alpha) * \beta = -(\alpha * \beta)$. Recuerde que no se puede usar la propiedad distributiva, que no se ha probado todavía. Sin embargo sí se puede usar que $-(-\alpha) = \alpha$. Por ejemplo, si $\alpha > \mathbb{O}$ y $\beta > \mathbb{O}$, entonces $(-\alpha) * \beta = -[(-(-\alpha)) * \beta] = -(\alpha * \beta)$.*

Lema 20 . $\mathbb{O} * \alpha = \mathbb{O}, \forall \alpha \in \mathbb{R}$.

Demostración. Comencemos suponiendo que $\alpha \geq \mathbb{O}$. (Por lo tanto $\alpha \subseteq \mathbb{O} = \mathbb{Q}^+$.)

En este caso $\mathbb{O} * \alpha = \{xy \mid x \in \mathbb{Q}^+, y \in \alpha\} \subseteq \mathbb{Q}^+ * \mathbb{Q}^+ \subseteq \mathbb{Q}^+$.

Recíprocamente, si $r \in \mathbb{Q}^+$ y $d \in \alpha$, entonces

$$r = \frac{r}{d}d \in \mathbb{Q}^+ * \alpha.$$

Por lo tanto $\mathbb{Q}^+ \subseteq \mathbb{Q}^+ * \alpha$. Así que $\mathbb{O} * \alpha = \mathbb{O}$.

Supongamos ahora que $\alpha < \mathbb{O}$. entonces $\mathbb{O} * \alpha = -(\mathbb{O} * (-\alpha)) = -\mathbb{O} = \mathbb{O}$. ■

Teorema 90 . *El producto está bien definido.*

Demostración. Lo que queremos demostrar es que el producto dos cortaduras es una cortadura. En vista de las definiciones, bastará demostrar que el producto de dos cortaduras mayores o iguales que \mathbb{O} es una cortadura.

Supongamos que $\alpha \geq \mathbb{O}$ y que $\beta \geq \mathbb{O}$ son cortaduras. Queremos demostrar que el conjunto

$$\{xy \mid x \in \alpha, y \in \beta\}$$

es una cortadura.

Es claro que el conjunto de arriba no es vacío dado que α no lo es, ni β tampoco.

Es claro que tanto α como β constan de racionales positivos por lo tanto

$$\{xy \mid x \in \alpha, y \in \beta\}$$

consta de sólo elementos positivos. Por esta razón

$$\{xy \mid x \in \alpha, y \in \beta\} \neq \mathbb{Q}.$$

Demostraremos ahora que $\{xy \mid x \in \alpha, y \in \beta\}$ no tiene primer elemento.

Sean $a \in \alpha$, $b \in \beta$, como α y β no tienen primer elemento, podemos tomar $a' < a$, $b' < b$ con $a' \in \alpha$ y $b' \in \beta$. Entonces $a'b' < ab$ y $a'b' \in \alpha\beta$.

Supongamos ahora que $ab < x$ con $a \in \alpha$, $b \in \beta$. Entonces $x = a\left(\frac{x}{a}\right)$ (note que $a > 0$), ahora $ab < a\left(\frac{x}{a}\right)$ implica que $b < \frac{x}{a}$. Como β es una cortadura, entonces $\frac{x}{a} \in \beta$. Por lo tanto $x = a\left(\frac{x}{a}\right) \in \alpha\beta$.

■

Teorema 91 . *El producto es conmutativo.*

Demostración. Si α_1 y α_2 son ambas $\geq \mathbb{O}$, entonces

$$\begin{aligned}\alpha_1 * \alpha_2 &= \{xy \mid x \in \alpha_1, y \in \alpha_2\} = \\ &= \{yx \mid x \in \alpha_1, y \in \alpha_2\} = \\ &= \alpha_2 * \alpha_1.\end{aligned}$$

Si $\alpha_1 \geq \mathbb{O}$ y $\alpha_2 < \mathbb{O}$, entonces

$$\alpha_1 * \alpha_2 = -(\alpha_1 * (-\alpha_2)).$$

Pero

$$\alpha_2 * \alpha_1 = -((- \alpha_2) * (\alpha_1)).$$

Del caso anterior se sigue que $\alpha_1 * (-\alpha_2) = (-\alpha_2) * (\alpha_1)$, de aquí que también valga la igualdad entre $\alpha_1 * \alpha_2$ y $\alpha_2 * \alpha_1$.

El caso $\alpha_2 \geq \mathbb{O}$ y $\alpha_1 < \mathbb{O}$ se sigue del caso anterior.

Por último, si $\alpha_1 < \mathbb{O}$ y $\alpha_2 < \mathbb{O}$, tenemos que

$$\alpha_1 * \alpha_2 = (-\alpha_1) * (-\alpha_2) = (-\alpha_2) * (-\alpha_1) = \alpha_2 * \alpha_1.$$

Lo anterior termina de establecer la conmutatividad del producto. ■

Teorema 92 . *El producto es asociativo.*

Demostración. Es fácil ver que si alguno de los tres factores es el \mathbb{O} , entonces el producto dará \mathbb{O} , independientemente de como se colocan los paréntesis.

En vista de la conmutatividad, basta verificar los casos siguientes: los tres números reales positivos, dos positivos y uno negativo, uno positivo y dos negativos y por último los tres negativos.

Supongamos para empezar, que α_1, α_2 y α_3 son cortaduras $\geq \mathbb{O}$.

$$\begin{aligned}\alpha_1 * (\alpha_2 * \alpha_3) &= \{xz \mid x \in \alpha_1, z \in \alpha_2 * \alpha_3\} = \\ &= \{x(yz) \mid x \in \alpha_1, y \in \alpha_2, z \in \alpha_3\} = \\ &= \{(xy)z \mid x \in \alpha_1, y \in \alpha_2, z \in \alpha_3\} = \\ &= (\alpha_1 * \alpha_2) * \alpha_3.\end{aligned}$$

Supongamos ahora que $\alpha_1 > \mathbb{O}, \alpha_2 < \mathbb{O}, \alpha_3 > \mathbb{O}$.

Entonces $\alpha_1 * (\alpha_2 * \alpha_3) = -[\alpha_1 * (-(\alpha_2 * \alpha_3))]$ ya que $\alpha_2 * \alpha_3 < \mathbb{O}$.

Así que $\alpha_1 * (\alpha_2 * \alpha_3) = -[\alpha_1 * (-(\alpha_2 * \alpha_3))] =$

$$= -[\alpha_1 * (-(\alpha_2 * \alpha_3))] = -[\alpha_1 * ((-\alpha_2) * \alpha_3)]$$

Por otra parte, $(\alpha_1 * \alpha_2) * \alpha_3 =$
 $[-(\alpha_1 * (-\alpha_2))] * \alpha_3 = [(\alpha_1 * (-\alpha_2)) * \alpha_3].$

(Recuerde que $(-\lambda) * \mu = -(\lambda * \mu)$).

Para terminar este caso, simplemente notemos que

$$\alpha_1 * ((-\alpha_2) * \alpha_3) = (\alpha_1 * (-\alpha_2)) * \alpha_3,$$

ya que $\alpha_1, (-\alpha_2)$ y α_3 son positivos.

Consideremos ahora el caso en que $\alpha_1 > \mathbb{O}, \alpha_2 < \mathbb{O}, \alpha_3 < \mathbb{O}$.
 $\alpha_1 * (\alpha_2 * \alpha_3) = \alpha_1 * [(-\alpha_2) * (-\alpha_3)].$

Por otra parte,

$$\begin{aligned} (\alpha_1 * \alpha_2) * \alpha_3 &= -[(\alpha_1 * \alpha_2) * (-\alpha_3)] = \\ &= -[-(\alpha_1 * (-\alpha_2)) * (-\alpha_3)] = \\ &= --[(\alpha_1 * (-\alpha_2)) * (-\alpha_3)] = \\ &= [(\alpha_1 * (-\alpha_2)) * (-\alpha_3)]. \end{aligned}$$

El caso en que los tres números son negativos se deja como ejercicio al lector. ■

Ejercicio 309 . Demuestre que $\alpha_1 * (\alpha_2 * \alpha_3) = (\alpha_1 * \alpha_2) * \alpha_3$, si $\alpha_1, \alpha_2, \alpha_3$ son $< \mathbb{O}$.

Teorema 93 . El producto tiene neutro.

Demostración. Demostraremos que el neutro para el producto es

$$\{1\}^* = \{x \in \mathbb{Q} \mid x > 1\}.$$

Sea $\alpha \geq \mathbb{O}$, entonces

$$\alpha * \{x \in \mathbb{Q} \mid x > 1\} \subseteq \alpha,$$

ya que si $d \in \alpha$ y $x > 1$ entonces $dx > d \in \alpha$ y como α es una cortadura, se tiene que $dx \in \alpha$.

Recíprocamente, sea $d \in \alpha$. Como las cortaduras no tienen un elemento menor, $\exists d' \in \alpha$ tal que $d' < d$.

Entonces

$$d = d' \left(\frac{d}{d'} \right)$$

con $d' \in \alpha$ y $\left(\frac{d}{d'}\right) > 1$.

Por lo tanto $\alpha \subseteq \alpha * \{x \in \mathbb{Q} \mid x > 1\}$.

$$\begin{aligned} \text{Si } \alpha > \mathbb{O}, \text{ entonces } \alpha * \{1\}^* &= -((- \alpha) * \{1\}^*) = \\ &= -((- \alpha)) = \alpha. \blacksquare \end{aligned}$$

Lema 21 . Si $\alpha > \mathbb{O}$ es una cortadura, entonces $\forall r > 1$, existe $d \in \alpha$ tal que $\frac{d}{r} \notin \alpha$.

Demostración. En caso contrario, dada $d \in \alpha$, todos los elementos de la sucesión

$$d, \frac{d}{r}, \frac{\left(\frac{d}{r}\right)}{r} = \frac{d}{r^2}, \frac{d}{r^3}, \dots$$

pertenecerían a α .

Recuerde que como $\alpha > \mathbb{O}$, entonces existe algún racional positivo c , tal que $c \notin \alpha$.

Como todos los elementos de α son mayores que todos los elementos de su complemento, tendríamos que

$$\frac{d}{r^n} > c > 0, \forall n \in \mathbb{N}.$$

Pero $\frac{d}{r^n} > c \Leftrightarrow \frac{d}{c} > r^n$. Como $r > 1$, entonces $r = 1 + h$ con $h > 0$. Así que $r^n = (1 + h)^n > 1 + nh$ (truncando el desarrollo de $(1 + h)^n$, según el Teorema del binomio). Así que si $1 + nh > \frac{d}{c} \Leftrightarrow n > \frac{\frac{d}{c} - 1}{h}$ entonces

$$r^n > \frac{d}{c},$$

contradicción. ■

Teorema 94 . Todo real distinto de \mathbb{O} tiene inverso multiplicativo.

Demostración. Nuevamente, basta demostrar que las cortaduras mayores que \mathbb{O} tienen inverso multiplicativo.

Sea $\alpha > \mathbb{O}$. Demostraremos que el conjunto

$$\beta = \left\{ x \in \mathbb{Q}^+ \mid \exists t_x > 1 \text{ tal que } \frac{1}{x}d > t_x, \forall d \in \alpha \right\}$$

es el inverso multiplicativo de α . Primero demostraremos que β es una cortadura.

β no es vacío:

Sea $b > 1$ y $d \in \alpha$ tal que $\frac{d}{b} \notin \alpha$ (Lema anterior). Como d no es el primer elemento de α (las cortaduras no tienen primer elemento) entonces existe $d' \in \alpha$ tal que $d' < d$.

Entonces $\frac{d'}{b} < \frac{d}{b}$, $\frac{d}{b} \notin \alpha$. $\forall x \in \alpha$ se tiene que $\frac{d}{b} < x$ (todo elemento en una cortadura es mayor que todo elemento en el complemento de la cortadura).

Por lo tanto,

$$\frac{b}{d'}x > \frac{b}{d'}\frac{d}{b} = \frac{d}{d'} > 1,$$

esto coloca a $\frac{b}{d'}$ en el conjunto β . Por lo tanto $\beta \neq \emptyset$.

$\beta \neq \mathbb{Q}$:

Es claro que si $d \in \alpha$, entonces $\frac{1}{d} \notin \beta$. ($d\frac{1}{d} = 1 \not> c, \forall c > 1$).

Veamos ahora que β no tiene primer elemento:

sean $x \in \beta$ y $c_x > 1$ tales que

$$xd > c_x = 1 + h, \forall d \in \alpha.$$

Tomemos $r = 1 + \frac{h}{2}$ y $s = \frac{1+h}{1+\frac{h}{2}}$. Así, $r > 1$, $s > 1$ y $rs = c_x$. Entonces

$$\frac{x}{r}d > s, \forall d \in \alpha,$$

por lo que $\frac{x}{r} \in \beta$, pero $\frac{x}{r} < x$.

Por último si $\beta \ni x < y$, entonces $\exists c_x > 1$ tal que

$$yd > xd > c_x, \forall d \in \alpha.$$

Esto coloca a y en β .

Lo anterior termina la demostración de que β es una cortadura.

Veremos ahora que $\beta\alpha = \{1\}^*$.

Por la definición del conjunto β , tenemos que todo producto xd con $x \in \beta$ y $d \in \alpha$ es mayor que 1. Es decir, $\beta\alpha \subseteq \{1\}^*$.

Recíprocamente, si $c > 1$ entonces podemos encontrar $d' \in \alpha$ tal que $\frac{c}{d'} \in \beta$ (véase la demostración de que $\beta \neq \emptyset$, unas líneas arriba).

Entonces $\frac{c}{d'} d' = c \in \beta\alpha$. Por lo tanto $\{1\}^* \subseteq \beta\alpha$. ■

Teorema 95 . *El producto se distribuye sobre la suma.*

Demostración. Sean $\alpha_1, \alpha_2, \alpha_3$ tres cortaduras, queremos demostrar que

$$\alpha_1 * (\alpha_2 + \alpha_3) = (\alpha_1 * \alpha_2) + (\alpha_1 * \alpha_3).$$

En vista de que $\mathbb{O} * \alpha = \mathbb{O}$, para cualquier cortadura α , es claro que la igualdad de arriba es verdadera si alguna de las tres cortaduras es \mathbb{O} .

Podemos suponer que las tres cortaduras son distintas de \mathbb{O} .

Supondremos que $\alpha_1 > \mathbb{O}$ y dejaremos el caso $\alpha_1 < \mathbb{O}$ como ejercicio.

Consideraremos los siguientes casos:

- i) $\alpha_2, \alpha_3 > \mathbb{O}$.
- ii) $\alpha_2 > \mathbb{O}, \alpha_3 < \mathbb{O}, (\alpha_2 + \alpha_3) > \mathbb{O}$.
- iii) $\alpha_2 > \mathbb{O}, \alpha_3 < \mathbb{O}, (\alpha_2 + \alpha_3) = \mathbb{O}$.
- iv) $\alpha_2 > \mathbb{O}, \alpha_3 < \mathbb{O}, (\alpha_2 + \alpha_3) < \mathbb{O}$.
- v) $\alpha_2, \alpha_3 < \mathbb{O}$.

Caso i) $\alpha_1 * (\alpha_2 + \alpha_3) = \{x(y+z) \mid x \in \alpha_1, y \in \alpha_2, z \in \alpha_3\} = \{xy + xz \mid x \in \alpha_1, y \in \alpha_2, z \in \alpha_3\}$.

Por otra parte,

$$\alpha_1 * \alpha_2 + \alpha_1 * \alpha_3 = \{xy + x'z \mid x, x' \in \alpha_1, y \in \alpha_2, z \in \alpha_3\}.$$

Así es claro que $\alpha_1 * (\alpha_2 + \alpha_3) \subseteq \alpha_1 * \alpha_2 + \alpha_1 * \alpha_3$.

Recíprocamente, si $xy + x'z$ con $x, x' \in \alpha_1, y \in \alpha_2, z \in \alpha_3$ y $x < x'$ (el caso $x' < x$ es análogo) entonces $xy + x'z = xy + (x + (x' - x))z = xy + xz + (x' - x)z > xy + xz \in \alpha_1 * (\alpha_2 + \alpha_3)$. Como $\alpha_1 * (\alpha_2 + \alpha_3)$ es una cortadura, entonces $xy + x'z \in \alpha_1 * (\alpha_2 + \alpha_3)$.

Caso ii) $\alpha_2 > \mathbb{O}, \alpha_3 < \mathbb{O}, (\alpha_2 + \alpha_3) > \mathbb{O}$.

$$\alpha_1 * (\alpha_2 + \alpha_3) = \alpha_1 * \alpha_2 + \alpha_1 * \alpha_3 \Leftrightarrow$$

$$\Leftrightarrow -\alpha_1 * \alpha_3 + \alpha_1 * (\alpha_2 + \alpha_3) = \alpha_1 * \alpha_2.$$

Ahora $\alpha_1 * \alpha_3 = -[\alpha_1(-\alpha_3)]$, por lo que $-\alpha_1 * \alpha_3 = [\alpha_1(-\alpha_3)]$.

Hagamos $X = -\alpha_3$, $Y = \alpha_2 + \alpha_3$, entonces $X + Y = \alpha_2$, por el caso i) tenemos que

$$\alpha_1 * (X + Y) = \alpha_1 * X + \alpha_1 * Y,$$

es decir que

$$\begin{aligned}\alpha_1 * \alpha_2 &= \alpha_1 * (-\alpha_3) + \alpha_1 * (\alpha_2 + \alpha_3) = \\ &= -\alpha_1 * \alpha_3 + \alpha_1 * (\alpha_2 + \alpha_3),\end{aligned}$$

es decir que

$$\alpha_1 * \alpha_2 + \alpha_1 * \alpha_3 = \alpha_1 * (\alpha_2 + \alpha_3).$$

Caso iii) $\alpha_2 > \mathbb{O}$, $\alpha_3 < \mathbb{O}$, $(\alpha_2 + \alpha_3) = \mathbb{O}$.

$$\alpha_1 * (\alpha_2 + \alpha_3) = \alpha_1 * \mathbb{O} = \mathbb{O}.$$

$$\begin{aligned}\text{Por otra parte } \alpha_1 * (-\alpha_2) &= \alpha_1 * \alpha_3 = -[\alpha_1 * (-\alpha_3)] = \\ &= -[\alpha_1 * (-(-\alpha_2))] = -[\alpha_1 * \alpha_2].\end{aligned}$$

$$\text{Por lo tanto } \alpha_1 * \alpha_2 + \alpha_1 * \alpha_3 = \alpha_1 * \alpha_2 - [\alpha_1 * \alpha_2] = \mathbb{O}.$$

Caso iv) $\alpha_2 > \mathbb{O}$, $\alpha_3 < \mathbb{O}$, $(\alpha_2 + \alpha_3) < \mathbb{O}$.

$$\alpha_1 * (\alpha_2 + \alpha_3) = \alpha_1 * \alpha_2 + \alpha_1 * \alpha_3 \Leftrightarrow$$

$-\alpha_1 * \alpha_2 - \alpha_1 * \alpha_3 = -\alpha_1 * (\alpha_2 + \alpha_3)$. Hagamos $\beta = -\alpha_2$, $\gamma = -\alpha_3$, entonces

$$\alpha_1 * \beta + \alpha_1 * \gamma = \alpha_1 * (\beta + \gamma),$$

así que

$$-\alpha_1 * \beta - \alpha_1 * \gamma = -\alpha_1 * (\beta + \gamma),$$

es decir que

$$\begin{aligned}-\alpha_1 * (-\alpha_2) - \alpha_1 * (-\alpha_3) &= -\alpha_1 * (-\alpha_2 - \alpha_3) = \\ &= -\alpha_1 * (-(\alpha_2 + \alpha_3)),\end{aligned}$$

por lo que $\alpha_1 * \alpha_2 + \alpha_1 * \alpha_3 = \alpha_1 * (\alpha_2 + \alpha_3)$.

Caso v) Si $\alpha_2 < \mathbb{O}$ y $\alpha_3 < \mathbb{O}$, entonces

$$\begin{aligned}\alpha_1 * (\alpha_2 + \alpha_3) &= -[\alpha_1 * (-(\alpha_2 + \alpha_3))] = \\ &= -[\alpha_1 * (-\alpha_2 - \alpha_3)] = \\ &= -[\alpha_1 * (-\alpha_2) + (\alpha_1 * (-\alpha_3))] = \\ &= -[\alpha_1 * (-\alpha_2)] - [\alpha_1 * (-\alpha_3)] = \\ &= \alpha_1 * \alpha_2 + \alpha_1 * \alpha_3.\end{aligned}$$



Ejercicio 310 *Demostrar que*

$$\alpha_1 * (\alpha_2 + \alpha_3) = (\alpha_1 * \alpha_2) + (\alpha_1 * \alpha_3)$$

si alguna de las tres cortaduras es \mathbb{O} .

Ejercicio 311 *. Demostrar que*

$$\alpha_1 * (\alpha_2 + \alpha_3) = (\alpha_1 * \alpha_2) + (\alpha_1 * \alpha_3)$$

si $\alpha_1 < \mathbb{O}$.

Ejercicio 312 *. Demostrar que*

$$\sqrt{2} =: \{p \in \mathbb{Q}^+ \mid p^2 > 2\}$$

*tiene la propiedad de que $\sqrt{2} * \sqrt{2} = 2^* =: \{r \in \mathbb{Q}^+ \mid r > 2\}$.*

6.5 Supremos e ínfimos

Teorema 96 *. Si $\{\alpha_i\}_{i \in X}, X \neq \emptyset$, es un conjunto de cortaduras acotado por debajo por γ , entonces $\{\alpha_i\}_{i \in X}$ tiene un ínfimo en \mathbb{R} .*

Demostración. $\gamma \leq \alpha_i$ es equivalente a decir que $\alpha_i \subseteq \gamma$. Demostraremos que $\cup \{\alpha_i\}_{i \in X}$ es una cortadura:

Sea $j \in X$, entonces $\alpha_j \neq \emptyset$, por lo que $\cup \{\alpha_i\}_{i \in X} \neq \emptyset$.

Por otra parte, $\cup \{\alpha_i\}_{i \in X} \subseteq \gamma$, así que $\emptyset \neq (\mathbb{Q} \setminus \gamma) \subseteq \mathbb{Q} \setminus (\cup \{\alpha_i\}_{i \in X})$. Es decir que $\cup \{\alpha_i\}_{i \in X} \neq \mathbb{Q}$.

$\cup \{\alpha_i\}_{i \in X}$ no tiene elemento menor, ya que si $x \in \cup \{\alpha_i\}_{i \in X}$, entonces $x \in \alpha_j$, para alguna $j \in X$. Como α_j no tiene elemento menor, entonces $\exists x' \in \alpha_j$ tal que $x' < x$. Notemos que $x' \in \cup \{\alpha_i\}_{i \in X}$.

Por último si $\cup \{\alpha_i\}_{i \in X} \ni x < y$, entonces $x \in \alpha_j$, para alguna $j \in X$, y como α_j es una cortadura, entonces $\alpha_j \ni x < y \implies y \in \alpha_j \implies y \in \cup \{\alpha_i\}_{i \in X}$.

Tenemos que $\alpha_j \subseteq \cup \{\alpha_i\}_{i \in X}, \forall j \in X$. Por lo tanto $\cup \{\alpha_i\}_{i \in X} \leq \alpha_j, \forall j \in X$. Es decir que $\cup \{\alpha_i\}_{i \in X}$ es una cota inferior para $\{\alpha_i\}_{i \in X}$.

Si γ es otra cota inferior, entonces $\gamma \leq \alpha_j, \forall j \in X$, que equivale a $\alpha_j \subseteq \gamma, \forall j \in X$. Por lo tanto $\cup \{\alpha_i\}_{i \in X} \subseteq \gamma$, es decir que $\gamma \leq \cup \{\alpha_i\}_{i \in X}$. Concluimos que $\cup \{\alpha_i\}_{i \in X}$ es la mayor de las cotas inferiores para $\{\alpha_i\}_{i \in X}$.

En resumen, $\cup \{\alpha_i\}_{i \in X} = \inf (\{\alpha_i\}_{i \in X})$. ■

6.5.1 El principio del supremo

Corolario 8 . *Como corolario del teorema anterior, demostraremos que todo conjunto de reales acotado por arriba tiene un supremo en \mathbb{R} .*

Demostración. Sea $\{\alpha_i\}_{i \in X}$ un conjunto de cortaduras acotado por arriba por γ . Es fácil ver que en este caso $-\gamma$ es una cota inferior para $\{-\alpha_i\}_{i \in X}$. Si llamamos β al ínfimo de $\{-\alpha_i\}_{i \in X}$, es inmediato que $-\beta$ es el supremo de $\{\alpha_i\}_{i \in X}$. ■

Ejercicio 313 . *Verificar con detalle las afirmaciones en la demostración anterior.*

Ejercicio 314 . *Si $r, q \in \mathbb{Q}$, entonces*

$$r < q \Rightarrow r^* < q^*.$$

Ejercicio 315 . *Si $r, s \in \mathbb{Q}$ entonces $r^*s^* = (rs)^*$.*

Ejercicio 316 . *(Recuerde la definición de “clase positiva” en un dominio entero). Demuestre que \mathbb{R}^+ es una clase positiva y que la definición de “ $<$ ” coincide con la que se obtiene de*

$$\alpha < \beta \Leftrightarrow \beta - \alpha \in \mathbb{R}^+$$

que -como se vio- produce una relación de orden (\leq) compatible con las operaciones, es decir, $\forall \alpha, \beta, \gamma \in \mathbb{R}$,

i) $\alpha \leq \beta \Rightarrow \alpha + \gamma \leq \beta + \gamma$.

ii) $\alpha \leq \beta \wedge \gamma \in \mathbb{R}^+ \Rightarrow \alpha\gamma \leq \beta\gamma$.

*Concluya que $(\mathbb{R}, +, *, 0^*, 1^*)$ es un campo (ver la definición 103 y la nota al pie de la página 197) ordenado linealmente con un orden compatible con las operaciones (el orden que se deriva de una clase positiva).*

Definición 89 . *Se define $h : \mathbb{Q} \rightarrow \mathbb{R}$ de modo que $h(r) = r^*$.*

Entonces tenemos que:

1. h es inyectiva.
2. $h(r + s) = h(r) + h(s)$.

$$3. \ h(rs) = h(r)h(s).$$

$$4. \ r < s \Rightarrow h(r) < h(s).$$

Lo que demuestra que hf es una inmersión de \mathbb{Q} en \mathbb{R} es decir: \mathbb{Q} es isomorfo a un subcampo $h(\mathbb{Q})$ de \mathbb{R} , en donde, por supuesto,

$$h(\mathbb{Q}) = \{r^* \in \mathbb{R} \mid r \in \mathbb{Q}\}$$

y con esta identificación ($r \longleftrightarrow r^*$) puede asegurarse que \mathbb{R} es una extensión (de campo) de \mathbb{Q} .

Mencionemos tres aplicaciones más del teorema de Dedekind (Principio del Supremo):

Observación 83 . *En \mathbb{R} no existen los “infinitésimos”.*

Observación 84 . *El orden canónico de \mathbb{R} es arquimediano. (Vease el Corolario 10, en la página 359)..*

Observación 85 . *La recta está completa.*

Para justificar nuestras conclusiones, comenzaremos con una observación que no es más que otra manera -equivalente- de definir supremo de un conjunto.

Observación 86 . *Dado $A \subset \mathbb{R}$, $A \neq \emptyset$, una cota superior s de A es tal que*

$$s = \sup A$$

si y sólo si ningún número menor que s es cota superior de A y por lo tanto debe suceder que si ε es positivo, entonces $s - \varepsilon$ no es cota superior y por lo tanto debe existir

$$a \in A \text{ tal que } s - \varepsilon < a.$$

Teorema 97 . \mathbb{N} no está acotado superiormente en \mathbb{R} .

Demostración. Supongamos lo contrario. Es decir que existe

$$s = \sup \mathbb{N},$$

luego $s - \frac{1}{2}$ no es cota superior de \mathbb{N} y por lo tanto, $\exists n_0 \in \mathbb{N}_0$ tal que $s - \frac{1}{2} < n_0$, de donde resulta que $s + \frac{1}{2} < n_0 + 1 < s < s + \frac{1}{2}$, es decir que

$$s + \frac{1}{2} < s + \frac{1}{2} \circ$$

El absurdo se obtuvo de suponer que \mathbb{N} está acotado por arriba, luego debe ser cierto lo contrario, y con esto termina la demostración del teorema. ■

Corolario 9 . *Dado que ningún real es cota superior de \mathbb{N} , $\forall a \in \mathbb{R}$, $\exists n_0 \in \mathbb{N}$ tal que $a < n_0$.*

Corolario 10 . *(Propiedad arquimediana del orden en \mathbb{R}) Dados $a, b \in \mathbb{R}^+$, $\exists n_0 \in \mathbb{N}$ tal que $n_0 a > b$.*

Demostración. En efecto, $a \in \mathbb{R}^+ \Rightarrow a \neq 0$ y $\therefore b/a$ es un real. De acuerdo con el Corolario precedente, existe $n_0 \in \mathbb{N}$ tal que $b/a < n_0$, y multiplicando por a ($a > 0$), se obtiene $b < n_0 a$. ■

Observación 87 . *Si se define “infinitésmo positivo” como un número positivo ε tal que $\varepsilon < 1/n, \forall n \in \mathbb{N}$ (equivalentemente $n\varepsilon < 1, \forall n \in \mathbb{N}$), el Corolario anterior asegura que en \mathbb{R} los infinitésmos no existen³.*

³Cabe mencionar aquí que, como un resultado del “Teorema de compacidad” de la Lógica matemática, deben existir campos ordenados, que contengan a \mathbb{R} como subcampo, en los que existen los infinitésmos –y por lo tanto, también existen sus inversos: “infinitos”-. Cada uno de estos campos se llama “un campo no estándar”. Las teorías matemáticas desarrolladas a partir de estos super-campos corresponden a lo que se conoce genéricamente como “análisis no-estándar” y apuntamos que este análisis ha tenido un crecimiento considerable en los últimos años.

El teorema de compacidad afirma que si $\{A_i\}_{i \in \mathbb{N}}$ es una colección numerable de axiomas tal que cualquier subconjunto finito de ella es consistente, entonces debe existir un modelo en el que valga la colección completa (Gödel demostró que un sistema axiomático es consistente si y sólo si tiene un modelo, es decir si los conceptos y relaciones del sistema se pueden interpretar como elementos de ciertos conjuntos, y esta interpretación es tal que en ella, los axiomas del sistema resultan ser proposiciones ciertas). Consideréngase los axiomas,

$\{A_i\}_{i \in \mathbb{N}}$ para un campo ordenado K , que contiene \mathbb{R} .

$$A_i : \left(\exists \varepsilon \in K, \text{ tal que } 0 < \varepsilon < \frac{1}{i} \right).$$

Y obsérvese que es un modelo en el que vale cualquier colección finita de tales axiomas. Por lo tanto, debe existir un campo K_0 ordenado, tal que contiene a \mathbb{R} y en el que valen todos los axiomas A_i .

e.d. $\exists \sigma \in K_0^+, 0 < \sigma < 1/n, \forall n \in \mathbb{N}$.

6.5.2 La recta está completa

Uno de los conceptos fundamentales del Cálculo, es el de “continuidad” para una función, que esencialmente dice que si una función $f : [a, b] \rightarrow \mathbb{R}$ es continua en algún punto x_0 de su dominio, entonces, para cualquier condición que se fije arbitrariamente que permita considerar que dos puntos de \mathbb{R} están “suficientemente cerca”, se puede encontrar una distancia $\delta > 0$ que garantice que cualquier punto del dominio de la función, cuya distancia a x_0 sea menor que δ , tiene, bajo f , una imagen que está suficientemente cerca de $f(x_0)$. Si como es costumbre, se define la distancia entre dos números reales a y b como $d(a, b) = |a - b|$, y además se denota con $\varepsilon > 0$, la condición impuesta para aceptar la “cercanía suficiente”, la definición usual de continuidad para una función f en un punto x_0 de su dominio es:

Definición 90 . *Sea $f : [a, b] \rightarrow \mathbb{R}$ una función, $x_0 \in [a, b]$. f es continua en x_0 si para cada $\varepsilon > 0$, $\exists \delta(\varepsilon) > 0$, tal que $\forall x \in [a, b]$, $|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon$.*

Se extiende la definición de continuidad en un punto a la de continuidad en un intervalo, de la manera natural:

Definición 91 . *$f : [a, b] \rightarrow \mathbb{R}$ es continua en $[a, b]$ si y sólo si f es continua en x , $\forall x \in [a, b]$.*

Una consecuencia inmediata de la definición es que si f es continua en $x_0 \in [a, b]$ y $f(x_0) > 0$, entonces $\exists I$ un intervalo con centro en x_0 tal que

$$x \in I \cap [a, b] \Rightarrow f(x) > 0.$$

(Análogamente si $f(x) < 0$, entonces $f(x) < 0$, $\forall x \in I \cap [a, b]$).

Aceptadas estas consecuencias, podemos ahora garantizar que los puntos de la recta la llenan completamente, es decir, que no se puede pasar en forma continua de un lado a otro de ella sin cortarla en alguno de sus puntos, (resultado que justifica la existencia de los puntos que se requieren en las construcciones geométricas). Explícitamente:

Teorema 98 . *Sea $f : [a, b] \rightarrow \mathbb{R}$. una función continua tal que*

$$f(a)f(b) < 0$$

(f cambia de signo de signo de un extremo del intervalo al otro). Entonces $\exists c \in (a, b)$ tal que $f(c) = 0$.

Demostración. Sin pérdida de generalidad podemos suponer

$$f(a) < 0$$

y

$$f(b) > 0$$

(si $f(a)$ fuera tal que $f(a) > 0$ y $f(b) < 0$, cámbiese, en la forma obvia, el sentido de las desigualdades que aparecen en la demostración).

Definimos:

$$A = \{x \in [a, b] \mid f(x) < 0\}.$$

$f(a) < 0 \Rightarrow a \in A$, luego A no es vacío, y como $\forall x \in A, x \in [a, b], b$ es una cota superior, entonces (Teorema de Dedekind) A tiene una cota superior mínima c . Entonces: $a \in A$, b cota superior de A y $c = \sup A$ implican que

$$a \leq c \leq b$$

y por tanto, $f(c) \in \mathbb{R}$.

Ahora bien, $f(c)$ no puede ser positivo. En efecto, si $f(c)$ lo fuera, la continuidad de f asegura que existe un intervalo I con centro en c tal que

$$x \in I \cap [a, b] \Rightarrow f(x) > 0$$

y puesto que $f(a) < 0$, $c \in (a, b]$ y entonces $\exists g, a \leq g < c$ tal que $(g, c] \subset (I \cap [a, b])$ y como en todo ese intervalo $f(x)$ es positiva (no existe ningún elemento de A en $(g, c]$, $(g+c)/2$ resulta ser una cota superior de A y entonces:

$$(g+c)/2 < c \leq (g+c)/2^{\nabla}$$

Este absurdo se obtuvo de suponer $f(c) > 0$ luego $f(c) \leq 0$.

De manera análoga se prueba que suponer $f(c) < 0$ conduce a un absurdo, por lo tanto, concluimos que $f(c) = 0$ y entonces c es diferente de a y de b o sea:

$$f(c) = 0, \quad c \in (a, b).$$

■

Entre las importantes consecuencias que tiene este teorema, -además de las geométricas que ya apuntamos- mencionaremos las siguientes:

Teorema 99 . Si $f : [a, b] \rightarrow \mathbb{R}$ es continua y para x_1 y x_2 en $[a, b]$ y M en \mathbb{R} es tal que

$$f(x_1) < M < f(x_2),$$

entonces en el intervalo I de extremos x_1 y x_2 existe un c tal que $f(c) = M$.

Demostración. En efecto, considérese la función $h : I \rightarrow \mathbb{R}$ definida por:

$$h(x) = f(x) - M$$

que cambia de signo en x_1 y x_2 y es continua y por lo tanto

$$\exists c \in I \text{ tal que } h(c) = 0 = f(c) - M, \text{ e. d., } f(c) = M.$$

■

Teorema 100 . $\forall a \in \mathbb{R}^+, \forall n \in \mathbb{Z}^+, \exists! r \in \mathbb{R}^+$ tal que $r^n = a$. (Todo real positivo tiene una única raíz n -ésima positiva para cada entero positivo n).

Demostración. (Existencia) Sea $f(x) = x^n - a$.

Entonces $f(0) = -a < 0$ y como

$$(1+a)^n = 1 + na + \frac{n(n-1)}{2}a^2 + \dots + a^n > a$$

$$f(1+a) = (1+a)^n - a \text{ es tal que } f(1+a) > 0.$$

Es decir f cambia de signo en los extremos del intervalo $[0, 1+a]$ y es continua (todo polinomio lo es). Entonces, de acuerdo con el teorema 98

$$\exists r \in (0, 1+a) \text{ tal que } f(r) = 0$$

pero $f(r) = r^n - a$. Luego $r^n = a$.

(Unicidad) si $t \in \mathbb{R}^+$ fuera tal que $t^n = a$,

entonces

$$r < t \Rightarrow r^n < t^n \therefore a < a^{\nabla}$$

$$r > t \Rightarrow r^n > t^n \therefore a > a^{\nabla}$$

$$\therefore r = t.$$

r se llama “la raíz n -ésima de a y se denota:

$$r = \sqrt[n]{a}.$$

■

6.6 Representación decimal de un número real

En esta sección demostramos que hay una correspondencia entre los números reales y sus expresiones decimales.

En primer lugar demostraremos que todo número real tiene una expresión decimal. Para este propósito basta demostrar que los reales positivos la tienen. Pero además, basta demostrar que los reales positivos menores que uno la tienen. Esto se debe a que ya sabemos que todo entero positivo tiene expresión decimal y a que todo real positivo se encuentra en algún intervalo $[N, N + 1)$, donde N es natural.

Teorema 101 . *Si r es un número real positivo, entonces existe un natural N tal que $r \in [N, N + 1)$.*

Demostración. Hemos visto que \mathbb{N} no está acotado en \mathbb{R} (principio arquimediano del orden en \mathbb{R}). En particular, r no es una cota superior para \mathbb{N} , por lo que $\exists M \in \mathbb{N}$ tal que $r \not\geq M$. De esta manera tenemos que $r < M$. Por el principio del buen orden en \mathbb{N} , podemos escoger la menor M con esta propiedad. Note que como M es natural positivo (porque $M > r > 0$) entonces $M = N + 1$, para algún natural N . Por la manera en que escogimos M entonces $N \not> r$. Por la tricotomía del orden de los reales, entonces $N \leq r$.

Concluimos que $N \leq r < N + 1$, que es lo mismo que decir que $r \in [N, N + 1)$. ■

En vista del teorema anterior un número real positivo r siempre satisface $N \leq r < N + 1$, p. a. $N \in \mathbb{N}$, o lo que lo mismo,

$$0 \leq r - N < 1.$$

Como $r = N + (r - N)$, para demostrar que r tiene una expresión decimal, basta demostrar que $r - N$ la tiene.

Teorema 102 . *Todo número real positivo menor que uno tiene expresión decimal.*

Demostración. Sea $0 < r < 1$ un número real. Queremos encontrar un dígito a_1 tal que $r \in \left[\frac{a_1}{10}, \frac{a_1 + 1}{10}\right)$. Es decir queremos que $\frac{a_1}{10} \leq r < \frac{a_1 + 1}{10}$. Esto sucede si y sólo si

$$a_1 \leq 10r < a_1 + 1.$$

Esto sugiere la manera de determinar a_1 :

tómese $10r$, es claro que $0 < 10r < 10$ (simplemente multiplique por 10 la desigualdad $0 < r < 1$).

Los intervalos $(0, 1), [1, 2), \dots, [9, 10)$ forman una partición del intervalo $(0, 10)$. Por esta razón, $10r$ pertenece a uno solo de los intervalos. Llámemos $\left[\frac{a_1}{10}, \frac{a_1 + 1}{10}\right)$ a ese intervalo.

Hemos encontrado la primera cifra adicional de r , a saber, a_1 .

Como $r \in \left[\frac{a_1}{10}, \frac{a_1 + 1}{10}\right)$, es claro que entonces

$$r - \frac{a_1}{10} \in \left[0, \frac{1}{10}\right).$$

Ahora, los intervalos $\left(0, \frac{1}{100}\right), \left[\frac{1}{100}, \frac{2}{100}\right), \dots, \left[\frac{9}{100}, \frac{10}{100}\right)$ forman una partición del intervalo $\left(0, \frac{1}{10}\right)$; así que $r - \frac{a_1}{10}$ pertenece uno solo de ellos, digamos que a

$$\left[\frac{a_2}{100}, \frac{a_2 + 1}{100}\right),$$

con lo que hemos encontrado la segunda cifra decimal de r .

Notemos que la segunda cifra decimal de r es la primera cifra decimal de

$$10r - a_1,$$

ya que

$$r - \frac{a_1}{10} \in \left[\frac{a_2}{100}, \frac{a_2 + 1}{100}\right) \Leftrightarrow 10r - a_1 \in \left[\frac{a_2}{10}, \frac{a_2 + 1}{10}\right).$$

En general, si ya hemos encontrado las primeras k cifras decimales de r , es decir que

$$\frac{a_1}{10} + \frac{a_2}{100} + \dots + \frac{a_k}{10^k} \leq r < \frac{a_1}{10} + \frac{a_2}{100} + \dots + \frac{a_k + 1}{10^k}$$

o bien

$$.a_1a_2\dots a_k \leq r < .a_1a_2\dots a_k + \frac{1}{10^k}$$

que equivale a

$$0 \leq r - .a_1a_2\dots a_k < \frac{1}{10^k},$$

multiplicando por 10^k obtenemos

$$0 \leq 10^k r - a_1 a_2 \dots a_k < 1.$$

La siguiente cifra decimal de r es la primera cifra decimal de $10^k r - a_1 a_2 \dots a_k$.

De esta manera, encontramos las cifras decimales de r . ■

Hemos visto que a todo número real se le puede asociar una expresión decimal.

Ejemplo 125 . Calcularemos algunas cifras de la expresión decimal de $\sqrt{2}$: $1 < \sqrt{2} < 2$, pues $1^2 < 2$ pero $2^2 = 4 < 2$.

Así, $\sqrt{2} - 1 \in (0, 1)$, por lo que $\sqrt{2} - 1 \in \left[\frac{a_1}{10}, \frac{a_1 + 1}{10} \right) \Leftrightarrow$

$\sqrt{2} \in \left[1 + \frac{a_1}{10}, 1 + \frac{a_1 + 1}{10} \right) \Leftrightarrow$

$2 \leq \left(1 + \frac{a_1}{10} \right)^2 \wedge \left(1 + \frac{a_1 + 1}{10} \right)^2 > 2$.

Ahora, $\left(1 + \frac{3}{10} \right)^2 = 1.69$, $\left(1 + \frac{4}{10} \right)^2 = 1.96$, $\left(1 + \frac{5}{10} \right)^2 = 2.25$.

De lo anterior vemos que la primera cifra decimal de $\sqrt{2}$ es 1.

Así que $1.4 < \sqrt{2} < 1.5$.

La segunda cifra decimal de $\sqrt{2}$ es la primera cifra decimal de $10(\sqrt{2} - 1.4) = 10\sqrt{2} - 14$.

Ahora, $10\sqrt{2} - 14 \in \left[\frac{a_2}{10}, \frac{a_2 + 1}{10} \right) \Leftrightarrow$

$10\sqrt{2} \in \left[14 + \frac{a_2}{10}, 14 + \frac{a_2 + 1}{10} \right) \Leftrightarrow$

$100\sqrt{2} \in [140 + a_2, 140(a_2 + 1)) \Leftrightarrow$

$20000 \in [(140 + a_2)_2, (140(a_2 + 1))^2)$.

Pero

$$(140 + 0)^2 = 19600, \quad (140 + 1)^2 = 19881, \quad (140 + 2)^2 = 20164.$$

De lo anterior vemos que la siguiente cifra decimal es 1, por lo tanto

$$\overline{1.41} < \sqrt{2} < 1.42.$$

Podríamos seguir de esta manera para tener tantas cifras decimales de $\sqrt{2}$ como quisieramos.

De manera recíproca, podemos ver que a cualquier expresión decimal

$$.a_1a_2a_3\dots$$

le corresponde un número real. ¿A qué real corresponde?

Si consideramos el conjunto $\left\{ .a_1 = \frac{a_1}{10}, .a_1a_2 = \frac{a_1}{10} + \frac{a_2}{100}, .a_1a_2a_3, \dots \right\}$ es claro que este conjunto está acotado por 1. Por el Principio del Supremo, este conjunto tiene supremo, digamos r . Es fácil ver que a este número le corresponden expresión decimal $.a_1a_2a_3\dots$

Ejercicio 317 . *Respecto al último párrafo, demuestre que la expresión decimal de r es precisamente $.a_1a_2a_3\dots$*

Hemos visto que todo número real tiene una expresión decimal y que cada expresión decimal corresponde a un real.

Observación 88 . *Un número real r es racional si sólo si tiene una expresión decimal periódica.*

Empecemos por considerar un racional positivo de la forma $\frac{m}{n}$ con m, n enteros positivos. Por el algoritmo de la división, $m = nq + r$, donde $0 \leq r < n$, r entero. Es claro que $\frac{m}{n} = q + \frac{r}{n}$, con q un entero. En vista de esto, basta demostrar que $\frac{r}{n}$ tiene una expresión decimal periódica, en el caso $r < n$.

La primera cifra decimal de $\frac{r}{n}$ se obtiene multiplicando r por 10 y encontrando la cifra decimal correspondiente, según la división

$$q \overline{)r * 10} \quad .$$

a_1
 r_1

Para obtener a_2 se multiplica r_1 por 10 y se divide entre q :

$$q \overline{)r * 10} \quad .$$

$a_1 \quad a_2$
 $r_1 \quad *10$
 r_2

las demás cifras decimales se obtienen repitiendo el procedimiento. Note que como los residuos son menores que q en el argumento, tiene que ocurrir una repetición de residuo. Pero si $r_i = r_j$ entonces $a_{i+1} = a_{j+1}$, además $r_{i+1} = r_{j+1}$, por lo que $a_{i+2} = a_{j+2}$. En general, $a_{i+k} = a_{j+k}$. De esta manera, la sucesión de dígitos

$$a_i, a_{i+1}, \dots, a_{j-1}$$

se repite periódicamente en la expresión decimal de $\frac{r}{q}$.

Recíprocamente, si se tiene que

$$r = .a_1 \dots a_{i-1} \overline{a_i \dots a_{j-1}},$$

es decir si el periodo es $\overline{a_i \dots a_{j-1}}$, entonces

$$r * 10^{i-1} = a_1 \dots a_{i-1} \overline{a_i \dots a_{j-1}}.$$

Pero entonces $\overline{a_i \dots a_{j-1}} * 10^{j-i} = a_i \dots a_{j-1} \overline{a_i \dots a_{j-1}}$, así que

$$\overline{a_i \dots a_{j-1}} * 10^{j-i} - \overline{a_i \dots a_{j-1}} = a_i \dots a_{j-1},$$

de aquí que $(10^{j-i} - 1) * \overline{a_i \dots a_{j-1}} = a_i \dots a_{j-1}$, es decir que

$$\overline{a_i \dots a_{j-1}} = \frac{a_i \dots a_{j-1}}{10^{j-i} - 1}.$$

En resumen, $r * 10^{i-1} = a_1 \dots a_{i-1} + \frac{a_i \dots a_{j-1}}{10^{j-i} - 1}$ de aquí que $r = \frac{a_1 \dots a_{i-1}}{10^{i-1}} + \frac{a_i \dots a_{j-1}}{(10^{j-i} - 1) 10^{i-1}}$, un número racional.

Ejemplo 126 . Consideremos el número $.123456456\overline{456}$ veamos que es un racional. Llamémosle r , entonces $1000r = 123.\overline{456}$, por el argumento de arriba, $\overline{456} = \frac{456}{999}$.

en efecto,

$$\begin{array}{r} & 4 & 5 & 6 \\ 999 & \overline{4560} \\ & 3996 \\ \hline & 564 & 0 \\ & 499 & 5 \\ \hline & 64 & 5 & 0 \\ & 59 & 9 & 4 \\ \hline & 4 & 5 & 6 & 0 \end{array}.$$

Note que en la división que estábamos haciendo, la siguiente división es la misma con la que empezamos, así que sin hacerla, ya sabemos que el siguiente dígito en el cociente es 4, luego 5, luego 6 y así sucesivamente.

Por lo tanto $\overline{.456} = \frac{456}{999}$, así que $1000r = 123 + \frac{456}{999}$, por lo que $r = \frac{123}{1000} + \frac{456}{999000}$.

Ejemplo 127 . El número $.010010001\dots$ (un 0, un 1, dos 0, un 1, tres 0, un 1, ..., n 0, un 1, $n+1$ ceros, un 1, ...) no es racional pues no tiene expresión decimal periódica.

Ejercicio 318 . Encuentre una expresión $\frac{a}{b}$, con $a, b \in \mathbb{Z}$ para los racionales con las expresiones decimales siguientes:

1. $21.0\bar{1}$.
2. $4.\overline{0001012}$.
3. $.0\overline{0011}$.

Ejercicio 319 Encuentre la expresión decimal de:

1. $\frac{3}{7}$.
2. $\frac{301}{999}$.
3. $\frac{24}{23}$.

Ejercicio 320 . Digamos que se escribe un punto, y después se escriben los naturales uno tras otro en su representación decimal, de tal manera que las primeras cifras son las siguientes

$$.01234567891011121314151617181920\dots$$

¿el número anterior será racional? Demuestre su afirmación.

Ejercicio 321 . Demuestre que $\sum_{k=1}^{\infty} 10^{-(k!)}$ no es racional.

Capítulo 7

El campo C de los números complejos

Una característica importante del conjunto \mathbb{R} de los números reales, que hemos estudiado en capítulos anteriores, es que tiene una clase positiva \mathbb{R}^+ , que define al orden canónico que por ser compatible con las operaciones (ver el Capítulo 6), tiene entre otras, la propiedad de que para toda a diferente de cero, a^2 es un positivo y por lo tanto la ecuación

$$x^2 + 1 = 0 \quad (7.1)$$

no puede tener solución en \mathbb{R} . Como en los casos anteriores- construcciones de \mathbb{Z} y de \mathbb{Q} - se pensó extender a un campo más grande, en el que la mencionada ecuación pudiera resolverse.

Era necesario construir un campo en el que existiera un número -imaginario- “ i ” que satisficiera la ecuación $x^2 + 1 = 0$, que fuera una extensión de \mathbb{R} y, por supuesto, que resultara el más “económico” -en el sentido de la contención- con esas dos propiedades.

En la época en que surgió este problema no se conocía el teorema que asegura que para todo campo K y todo polinomio $f(x)$ no constante, con coeficientes en K ; existe una extensión de K en la que el polinomio tiene al menos una raíz. Teorema que valida la construcción, que resulta más natural y que la hubiera librado de las objeciones -injustificadas- que en su momento se hicieron y que se referían al invento de los números imaginarios.

Es pertinente observar que en el campo cuya construcción se deseaba, no puede haber una relación de orden compatible con las operaciones -que es la única que interesa al Álgebra- ya que en ese caso, como en el de los reales,

los cuadrados tendrían que ser no negativos. Por esta razón, algunos autores que enfatizan la propiedad, dicen que \mathbb{C} es el “desordenado” campo de los números complejos, a pesar de que una consecuencia del axioma de selección asegura que en todo conjunto se puede definir un buen orden. Lo que no puede asegurarse es que ese orden resulte necesariamente compatible con las operaciones.

Puestos a estudiar ese hipotético campo en el que figura esa misteriosa i , se vio que tenían que estar también todas sus potencias (i^2, i^3, \dots) productos de éstas por números reales y sumas de tales productos, es decir que debían estar consideradas todas las expresiones de la forma

$$a_0 + a_1i + a_2i^2 + \dots + a_ni^n, a_j \in \mathbb{R}, j \in \{0, \dots, n\}, \quad (*)$$

además de sus inversos multiplicativos. Se notó que como:

$$i^2 = -1, \quad (7.2)$$

$$i^3 = i^2i = -i, \quad (7.3)$$

$$i^4 = i^2i^2 = (-1)(-1) = 1, \quad (7.4)$$

si n es un número natural tal que

$$n = 4q + r, \quad 0 \leq r < 4, \quad (7.5)$$

entonces

$$i^n = (i^4)^q \cdot i^r = i^r, \quad (7.6)$$

o sea i^n es $1, i, -1$ ó $-i$. Observación que permite simplificar las expresiones (*) que pueden reducirse a “binomios” de la forma $a + bi$, con $a, b \in \mathbb{R}$.

1. $3 + 2i - 7i^2 + 2i^3 - i^4 + 7i^5 = 3 + 2i - 7(-1) + 2(-i) - 1 + 7i = 9 + 7i.$
2. $1 + i^3 + i^{37} - i^{204} = 1 - i + i - 1 = 0.$

Ejercicio 322 . *Expresé en la forma $a + bi$ las siguientes expresiones:*

$$1. \quad 2 - 8i + 7i^3 - 3i^7 + 16i^{20}$$

$$2. \quad (2i^3)^5$$

$$3. \quad \frac{(2 - 7i)^2}{1 + 3i - 2i^2 + i^3 + 2i^7}$$

Tomando en cuenta lo anterior, se procedió a estudiar al subconjunto formado por los elementos del nuevo campo que pueden expresarse como binomios:

$$\beta = \{a + bi \in \mathbb{C} \mid a, b \in \mathbb{R}, i^2 = -1\}. \quad (7.7)$$

Como siempre que se define un conjunto nombrando a sus elementos conviene aportar un criterio que permita decidir cuando dos nombres corresponden al mismo individuo, hacemos notar que, puesto que β es un campo, entonces

$$(a + bi = c + di) \Rightarrow ((a - c) = (d - b)i).$$

Por lo tanto, $(a - c)^2 = -(d - b)^2$, que es una igualdad en \mathbb{R} , lo que implica que cada uno de los cuadrados debe ser necesariamente 0. Por lo tanto $a = c$ y $b = d$. Es decir que cada elemento tiene una representación única. Así por ejemplo si supiéramos que $i = u + vi$, sabríamos que $u = 0$ y que $v = 1$.

Como β es un subconjunto de un campo,

$$(a + bi) + (c + di) = (a + c) + (b + d)i, \quad (7.8)$$

de donde resulta que la suma -en \mathbb{C} - de elementos de β , produce un elemento de β . Así, β es cerrado bajo la suma de \mathbb{C} . Por lo tanto, la restricción de esta suma a $\beta \times \beta$ es una operación binaria en β , que por herencia, resulta asociativa y commutativa. También se tiene que $0 = 0 + 0i$ pertenece a β y que para cada $a + bi \in \beta$, $-a + (-b)i$ (el inverso aditivo de $a + bi$) es un elemento de β . Luego $(\beta, +, 0)$ es un grupo abeliano.

$$1. (3 + 4i) + (6 + 5i) = (3 + 6) + (4 + 5)i = 9 + 9i.$$

$$2. (6 - 14i) - (-10 + 17i) = 16 - 31i.$$

Ejercicio 323 . *Expresa el resultado las operaciones siguientes en la forma $a + bi$:*

$$1. (5 + 7i) + (8 + 2i).$$

$$2. (5 + 7i) - (8 + 2i).$$

$$3. (11 + 2i) + (3 - 14i).$$

$$4. (19 - i) + (2 - 13i).$$

$$5. (15 + 2i) - (6 + 4i).$$

6. $(\sqrt{2} + i\sqrt{3}) - (6\sqrt{2} - 7i\sqrt{3})$.
7. $(19 - i) + (2 - 12i)$.
8. $[(2 + 6i) + (6 - 5i)] - (4 - 11i)$.
9. $(-6 + 4i) - [(18 - 6i) - (-7 - 2i)]$.
10. $(1 + i\sqrt{3}) - (7 + 2\sqrt{3}i) + (7 - \sqrt{3}i)$.

El producto -en \mathbb{C} - de dos elementos de β , está en β . En efecto,

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i, \quad (7.9)$$

y por lo tanto, β tiene también una multiplicación que por ser la restricción de la de un campo, es asociativa y commutativa. Además esta multiplicación tiene neutro ($1 = 1 + 0i$) y se distribuye sobre la suma por ambos lados.

1. $(6 + 3i)(2 + 4i) = (12 - 12) + (6 + 24)i = 30i$.
2. $(2 + i\sqrt{3})(5 - 6\sqrt{3}i) = (10 + 18) + (5\sqrt{3} - 12\sqrt{3})i = 28 - 7\sqrt{3}i$.

Finalmente, si $a + bi \in \beta$ es $\neq 0 + 0i$, ($a \neq 0$ ó $b \neq 0$ y por lo tanto $a^2 + b^2 > 0$). Así que

$$(a + bi)^{-1} = \frac{a}{a^2 + b^2} - i\frac{b}{a^2 + b^2}, \quad (7.10)$$

como puede comprobarse efectuando el producto $(a + bi) \cdot \left(\frac{a}{a^2 + b^2} - i\frac{b}{a^2 + b^2} \right)$.

Por lo que resulta que β es un campo, donde cada número real $a = a + 0i$ está en β , así como $i = 0 + 1i$. Obviamente, β es el menor campo, en el sentido de la contención, con esos dos propiedades. Luego, $\beta = \mathbb{C}$ es el campo que se deseaba construir.

Para ver como se obtiene el inverso multiplicativo de $a + bi$, consideremos un número $x + yi$ tal que multiplicado por $a + bi$ nos dé $1 + 0i$, es decir:

$$(x + yi)(a + bi) = 1 + 0i, \quad (7.11)$$

efectuando la multiplicación:

$$(ax - by) + (ay + bx)i = 1 + 0i, \quad (7.12)$$

de donde

$$ax - by = 1 \quad (7.13)$$

$$bx + ay = 0. \quad (7.14)$$

Como $a+bi \neq 0$, $a \neq 0$ ó $b \neq 0$, y por lo tanto $a^2+b^2 > 0$, por lo que aplicando la regla de Cramer (ver el teorema 151 en la página 531) se obtiene

$$x = \frac{\begin{vmatrix} 1 & -b \\ 0 & a \end{vmatrix}}{\begin{vmatrix} a & -b \\ b & a \end{vmatrix}} = \frac{a}{a^2+b^2} \text{ y} \quad (7.15)$$

$$y = \frac{\begin{vmatrix} a & 1 \\ b & 0 \end{vmatrix}}{\begin{vmatrix} a & -b \\ b & a \end{vmatrix}} = \frac{-b}{a^2+b^2}, \quad (7.16)$$

es decir que

$$(a+bi)^{-1} = \frac{a}{a^2+b^2} + \frac{-b}{a^2+b^2}i. \quad (7.17)$$

Consideremos el ejemplo siguiente:

Ejemplo 128 . Se desea encontrar $(3+4i)^{-1}$.

De acuerdo con lo anterior, el resultado es $\frac{3-4i}{25}$. Comprobamos:

$$(3+4i) \left(\frac{3-4i}{25} \right) = \frac{9+16}{25} = 1. \quad (7.18)$$

Definamos el conjugado del número complejo $z = a+bi$ como $\bar{z} = a-bi$.

Así si $z = 3+4i$, su conjugado es $3-4i$, y entonces, recordando que la expresión $\frac{a}{b}$ representa el producto de a por el inverso de b , ($\frac{a}{b} = ab^{-1}$), encontramos que para efectuar la división de z entre w , basta multiplicar el cociente $\frac{z}{w}$ por $\frac{\bar{w}}{\bar{w}}$, con lo que se tiene el resultado deseado: $\frac{z}{w} = \frac{z\bar{w}}{w\bar{w}}$.

Ejemplo 129 . Dividiremos $(2-i)$ entre $(1+i)$:

$$\frac{2-i}{1+i} = \frac{2-i}{1+i} \cdot \frac{1-i}{1-i} = \frac{1-3i}{2} = \frac{1}{2} - \frac{3}{2}i.$$

$$\text{En efecto, } (1+i) \left(\frac{1}{2} - \frac{3}{2}i \right) = \left(\frac{1}{2} + \frac{3}{2} \right) + \left(\frac{1}{2} - \frac{3}{2} \right) i = (2-i).$$

$$\text{Ejemplo 130 . } \frac{6+7i}{4-3i} = \frac{6+7i}{4-3i} \cdot \frac{4+3i}{4+3i} = \frac{(24-21)+(28+18)i}{25} = \frac{3+46i}{25}.$$

$$\begin{aligned} \text{Ejemplo 131 . } \frac{\sqrt{7}+i\sqrt{5}}{5+i\sqrt{7}} &= \left(\frac{\sqrt{7}+i\sqrt{5}}{5+i\sqrt{7}} \right) \left(\frac{5-i\sqrt{7}}{5-i\sqrt{7}} \right) = \\ &= \frac{(\sqrt{35}+\sqrt{35})+(5-7)i}{5+7} = \frac{2\sqrt{35}-2i}{12} = \frac{1}{6}(\sqrt{35}-i). \end{aligned}$$

Ejercicio 324 . Exprese en la forma $a+bi$ el resultado de las siguientes operaciones:

1. $(9+8i)(7-6i)$.
2. $(6+3i)(2+5i)$.
3. $(6-3i)(2+5i)$.
4. $(6-3i)(2-5i)$.
5. $(-6+3i)(2-5i)$.
6. $(5-3i)(7-2i)$.
7. $(7\sqrt{2}-6\sqrt{5}i)(2\sqrt{2}-7\sqrt{5}i)$.
8. $(3+2i)[(7-8i)(-2+9i)]$.
9. $(5+3i)^2$.
10. $(3-2i)^3$.
11. $(8-3i)/(9+4i)$.

12. $(2 - 11i) / (3 - 2i)$.
13. $5i / (6 - 7i)$.
14. $(\sqrt{3} - i\sqrt{7}) / (2\sqrt{3} + 3\sqrt{7}i)$.
15. $(8\sqrt{5} + 21\sqrt{3}i) / (2\sqrt{5} - 7\sqrt{3}i)$.
16. $(\sqrt{6} + 6\sqrt{10}i) / (2\sqrt{6} - 3\sqrt{10}i)$.
17. $(2 + 3i)(3 - 4i) / (4 + 5i)$.
18. $(5 + 6i)^2 / (9 - 2i)$.
19. $(-5 + 8i) / (3 + 2i)^2$.
20. $\frac{8 + 7i}{(5 + 6i)(5 - 2i)}$.
21. $\frac{(10i - 3)(4i + 3)}{9i - 8}$.
22. $\frac{(1 + 2i)(2 + 3i)}{(3 + 4i)(4 + 5i)}$.
23. $\frac{(3 - 5i)(7 + 4i)}{(5 + 3i)(6 - i)}$.

7.1 La inmersión de \mathbb{R} en \mathbb{C}

Con objeto de contestar las objeciones que se hicieron a la construcción anterior -la existencia de números imaginarios-, Gauss, tomando como base los resultados anteriores, propuso el siguiente modelo:

7.1.1 Modelo

Sea

$$\mathbb{C} = \{(a, b) \mid a, b \in \mathbb{R}\}, \quad (7.19)$$

junto con las siguientes propiedades:

1. $(a, b) = (c, d) \iff a = c$. (Representación única).

2. $(a, b) \oplus (c, d) =: (a + c, b + d)$. (La suma dentro del último paréntesis es la de \mathbb{R}).
3. $(a, b) \odot (c, d) =: (ac - bd, ad + bc)$.

El neutro para \oplus es $\vec{0} = (0, 0)$, $-(a, b) = (-a, -b)$, el neutro para el producto es $e = (1, 0)$ y si $(a, b) = z \in \mathbb{C} \setminus \{\vec{0}\}$, entonces $z^{-1} = \left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right)$ ¹.

En este modelo, se acostumbra llamar a la primera componente de cada pareja, la parte real y a la segunda la imaginaria. Nótese que ambas partes son números reales, así si $z = (a, b)$ es un complejo, entonces $\operatorname{Re}(z) = a$, $\operatorname{Im}(z) = b$ y esta costumbre queda justificada con la inmersión

$$\begin{aligned} f : \mathbb{R} &\rightarrow \mathbb{C} \\ a &\mapsto (a, 0) \end{aligned} \tag{7.20}$$

(Recuerde que una inmersión de una estructura en otra es una función inyectiva que “respeta” las relaciones de ambas. Explícitamente, f debe ser inyectiva y $\forall a, b \in \mathbb{R}, f(a + b) = f(a) + f(b)$ y $f(ab) = f(a)f(b)$ en donde - por supuesto- las operaciones a la izquierda de las igualdades son operaciones en \mathbb{R} y las de la derecha son operaciones en \mathbb{C}).

En vista de que las operaciones de las parejas -suma y producto- se definieron tomando como modelo las de los binomios, estas nuevas operaciones tienen -necesariamente- las propiedades de las anteriores, y por lo tanto,

$$\mathbb{C} = \{(a, b); a, b \in \mathbb{R}\}, \oplus, \odot, \vec{0}, e \tag{7.21}$$

resulta un campo, como puede comprobarse fácilmente. La inmersión $f : \mathbb{R} \rightarrow \mathbb{C}$ definida por $f(a) = (a, 0)$ muestra que puede considerarse \mathbb{C} como una extensión de campo \mathbb{R} y definiendo i como $i = (0, 1)$ e identificando a con $f(a) = (a, 0)$, puede verse que:

1. $i^2 = (0, 1)(0, 1) = (-1, 0) = -1$.
2. $(a, b) = (a, 0) + (0, b) = (a, 0) + (b, 0)(0, 1) = a + bi$.

Con lo que se recuperan los binomios de los que se partió en la construcción de Gauss.

¹ $z \neq \vec{0} \Rightarrow a^2 + b^2 \neq 0$.

La construcción del inverso multiplicativo de un complejo (a, b) no cero, $(a, b)^{-1} = (a/(a^2 + b^2), -b/(a^2 + b^2))$ muestra la conveniencia de definir dos funciones $f : \mathbb{C} \rightarrow \mathbb{C}$, $g : \mathbb{C} \rightarrow \mathbb{R}$ por medio de las fórmulas:

$$f((a, b)) = (a, -b) \quad (7.22)$$

-la conjugación- y

$$g((a, b)) = \sqrt{(a^2 + b^2)} \quad (7.23)$$

-el módulo ó tamaño- .

7.2 La conjugación

Si $z = (a, b)$, denotaremos

$$\bar{z} = f(z) = (a, -b) \quad (7.24)$$

(Cuando se identifican los complejos como puntos del plano coordenado \mathbb{R}^2 , $(z = (a, b)$ es el punto de abscisa a y de ordenada b), la conjugación puede interpretarse geométricamente como la reflexión sobre el eje X).

En adelante usaremos los símbolos $+$ y \cdot para las operaciones en \mathbb{C} .

La conjugación tiene, entre otras, las propiedades que expresa el siguiente:

Teorema 103 . $\forall z, w \in \mathbb{C}$,

1. $\overline{z + w} = \bar{z} + \bar{w}$ “El conjugado de la suma es la suma de los conjugados”.
2. $\overline{z \cdot w} = \bar{z} \cdot \bar{w}$ “El conjugado de un producto es el producto de los conjugados”.
3. $\bar{\bar{z}} = z \Leftrightarrow z \in \mathbb{R}$.
4. $\overline{\overline{z}} = z$.
5. $\bar{z} = \bar{w} \Rightarrow z = w$.
6. $z + \bar{z} = 2 \operatorname{Re}(z)$, $z - \bar{z} = 2i \operatorname{Im}(z)$.

Demostración. Sean $z = (a, b)$, $w = (c, d)$ Entonces:

$$1) \overline{z+w} = \overline{((a+c), (b+d))} = ((a+c), -(b+d)) = (a, -b) + (c, -d) = \bar{z} + \bar{w}.$$

$$2) \overline{z \cdot w} = (ac - bd, -(ad + bc)) = (a, -b)(c, -d) = \bar{z} \cdot \bar{w}.$$

3) $z \in \mathbb{R} \Rightarrow (b = 0) \therefore \bar{z} = (a, -0) = (a, 0) = z$. Recíprocamente, si $\bar{z} = z$ entonces $(a, -b) = (a, b)$. Por lo tanto $-b = b \in \mathbb{R}$, de aquí que $b = 0$. Por lo tanto $z = (a, 0)$.

$$4) \overline{\bar{z}} = \overline{\overline{(a, b)}} = \overline{(a, -b)} = (a, -(-b)) = (a, b) = z.$$

5) Es claro.

$$6) z + \bar{z} = (a, b) + (a, -b) = (2a, 0) = 2Re(z).$$

$$z - \bar{z} = (a, b) - (a, -b) = (0, -2b) = 2bi = 2i \operatorname{Im} z. \blacksquare$$

Corolario 11 . () : $\mathbb{C} \rightarrow \mathbb{C}$ es una biyección, pues es su propia inversa.

Demostración. Es 4) en el teorema anterior. ■

Corolario 12 . Si $w \in \mathbb{C} \setminus \{0\}$, entonces $\overline{\left(\frac{z}{w}\right)} = \left(\frac{\bar{z}}{\bar{w}}\right)$.

Demostración. En efecto, $z = \frac{z}{w} \cdot w \therefore \bar{z} = \overline{\left(\frac{z}{w}\right)} \cdot \bar{w}$ por lo tanto

$$\overline{\left(\frac{z}{w}\right)} = \left(\frac{\bar{z}}{\bar{w}}\right). \quad (7.25)$$

Nótese que $w \neq 0 \Rightarrow \bar{w} \neq 0$. ■

Cuando se interpreta la conjugación como una función $f : \mathbb{C} \rightarrow \mathbb{C}$, el teorema anterior prueba que f es una función biyectiva que “va bien” con las operaciones de \mathbb{C} y que “deja fijo” a \mathbb{R} - en el sentido de la parte (3) del teorema 103².

Teorema 104 . Si $\eta : \mathbb{C} \rightarrow \mathbb{C}$ es un automorfismo que deja fijo a \mathbb{R} , ($\eta(a) = a, \forall a \in \mathbb{R}$), entonces η es la conjugación ó la identidad en \mathbb{C} .

²Las funciones biyectivas que respetan las operaciones (y cuyo inverso también las respeta), se llaman *isomorfismos* y cuando “van” de un campo en él mismo, se conocen como *automorfismos*.

Demostración. Sea $z = (a, b) = a + bi$.

Entonces

$$\eta(z) = \eta(a + bi) = \eta(a) + \eta(b)i \quad (7.26)$$

$$(a, b \in \mathbb{R} \Rightarrow \eta(a) = a; \eta(b) = b).$$

Además, $-1 = \eta(-1) = \eta(ii) = (\eta(i))^2$ y por lo tanto $\eta(i)$ tiene que ser una raíz cuadrada de -1 ó sea $\eta(i) \in \{i, -i\}$.

Si $\eta(i) = i$, la ecuación 7.26 muestra que η es la identidad en \mathbb{C} y si $\eta(i) = -i$, entonces η es la conjugación. ■

Ejemplo 132 . Se desea calcular z si $iz + (2 - i)\bar{z} = 10 + 6i$.

Entonces sí $z = x + iy$, $\bar{z} = x - iy$, por lo tanto $iz + (2 - i)\bar{z} = i(x + iy) + (2 - i)(x - iy) = (2x - 2y) - 2iy$.

Ahora $(2x - 2y) - 2iy = 10 + 6i \Rightarrow -2y = 6$, $2x - 2y = 10$ y por lo tanto, $x = 2$, $y = -3$.

Ejercicio 325 . Resuelva

$$1. (1 + i)z + (1 - i)\bar{z} = 4$$

$$2. z\bar{z} + 3(z + \bar{z}) = 7$$

$$3. \begin{cases} iz + (1 + i) = 3 + i \\ (1 + i)\bar{z} - (6 + i)\bar{w} = 4 \end{cases}$$

7.3 La norma

Si $z = (a, b)$, hacemos $\|z\| =: g(z) = (a^2 + b^2)^{1/2}$.

Si se identifican los complejos como puntos del plano, la norma - o tamaño - de z puede interpretarse como la distancia euclíadiana de (a, b) al origen.

La función distancia tiene, entre otras, las propiedades que están enumeradas en el siguiente:

Teorema 105 . $\forall z, w \in \mathbb{C}$,

$$1. \|z\| = (z\bar{z})^{1/2}. \text{ (O bien, } \|z\|^2 = (z\bar{z})).$$

$$2. \|z\| \geq 0; \|z\| = 0 \iff z = 0.$$

3. $\|zw\| = \|z\| \|w\|$.
4. $\|z + w\| \leq \|z\| + \|w\|$.

Demostración. 1) $(z\bar{z})^{1/2} = ((a, b)(a, -b))^{1/2} = (a^2 + b^2)^{1/2} = \|z\|$.

2) Obvia, ya que el tamaño de z es la raíz cuadrada de un número no negativo, y ésta sólo es cero si el radicando $(a^2 + b^2)$ lo es.

3) $\|zw\|^2 = (zw)(\bar{zw}) = z\bar{z}w\bar{w} = \|z\|^2 \|w\|^2$ y como los tamaños son números reales, “se vale” extraer raíces cuadradas. Luego $\|zw\| = \|z\| \|w\|$.

4) Nótese que $\forall z, \|z\| = \|\bar{z}\|$ y que $|\operatorname{Re}(z)| \leq \|z\|$; $|\operatorname{Im}(z)| \leq \|z\|$.

$$\begin{aligned} \|z + w\|^2 &= (z + w)\overline{(z + w)} = z\bar{z} + z\bar{w} + \bar{z}w + w\bar{w} = \\ &= \|z^2\| + 2\operatorname{Re}(z\bar{w}) + \|w\|^2 \leq \|z^2\| + 2\|(z\bar{w})\| + \|w\|^2 = (\|z\| + \|w\|)^2. \end{aligned}$$

Es decir $\|z + w\|^2 \leq (\|z\| + \|w\|)^2$ que es una desigualdad de números reales no negativos. Por lo tanto $\|z + w\| \leq \|z\| + \|w\|$. ■

Corolario 13 . De 3) se sigue que

1. si $t \in \mathbb{R}$, entonces, $\|t\| = (t\bar{t})^{1/2} = (t^2)^{1/2} = |t|$,
- $$\therefore \|t\bar{z}\| = |t| \|\bar{z}\| = |t| \|z\|. \quad (7.27)$$
2. En particular, si $t = -1$, entonces $\|-z\| = \|z\|$.

Una función de $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ en con las propiedades 2), 3) y 4) del teorema anterior, se llama una *norma*, y permite definir la distancia entre dos puntos de \mathbb{R}^2 -en este caso dos complejos z, w - como sigue:

Definición 92 . $\forall z, w \in \mathbb{C}$, $d(z, w) = \|z - w\|$.

De la definición y de las propiedades de la norma, se deducen las propiedades siguientes que muestran que “ d ” es una *métrica* (*o distancia*).

Teorema 106 . $\forall z, y, w \in \mathbb{C}$,

1. $d(z, w) \geq 0$; $d(z, w) = 0 \Leftrightarrow z = w$.
2. $d(z, w) = d(w, z)$.

$$3. \ d(z, w) \leq d(z, y) + d(y, z).$$

Demostración. En efecto:

$$(1) \|z - w\| \geq 0; \|z - w\| = 0 \Leftrightarrow z - w = 0 \Leftrightarrow z = w.$$

$$(2) \|z - w\| = \|w - z\|.$$

$$(3) \|z - w\| = \|z - y + y - w\| \leq \|z - y\| + \|y - w\|.$$

(De las propiedades consignadas en el teorema 105 y sus Corolarios). ■

7.4 La ecuación general de segundo grado

Un teorema cuya importancia le valió el nombre de *Teorema fundamental del Álgebra* -cuya demostración se sale del nivel de estas notas- asegura que \mathbb{C} es algebraicamente cerrado, es decir que todo polinomio $f(x) \in \mathbb{C}[x]$ de grado n , tiene n raíces -bien contadas- en \mathbb{C} ³. Sin embargo vale la pena demostrar la existencia de algunas de éstas. En particular las raíces cuadradas que deben calcularse cuando se usa la fórmula para resolver la ecuación general de segundo grado. Veamos:

Se desea resolver la ecuación

$$ax^2 + bx + c = 0,$$

en donde $a, b, c \in \mathbb{C}$, $a \neq 0$.

- Dividimos la ecuación entre a ($a \neq 0$) y restamos c/a de cada lado:

$$x^2 + (b/a)x = -c/a.$$

- Completamos el trinomio cuadrado perfecto por la izquierda, sumando a cada lado $\frac{b^2}{4a^2}$:

$$x^2 + \frac{b}{a}x + \frac{b^2}{4a^2} = \frac{b^2}{4a^2} - \frac{c}{a}$$

o sea:

$$\left(x + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2}.$$

³La demostración de este teorema dentro del Análisis complejo es tan elegante -corta- que justifica plenamente que por ahora, se posponga tal demostración, que a este nivel resulta complicada y larga, y que, por descansar -necesariamente- en la construcción de \mathbb{R} , no puede ser algebraica pura.

- Suponiendo que se puede sacar raíz cuadrada a $b^2 - 4ac$ y que la representamos como: $\sqrt{b^2 - 4ac}$, entonces $x + \frac{b}{2a} = \pm \frac{\sqrt{b^2 - 4ac}}{2a}$, en donde el doble signo expresa el hecho de que para el caso, sirve tanto la raíz cuadrada de $b^2 - 4ac$ cuya existencia supusimos, como su inverso.
- Finalmente,

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Queda por justificar la antes mencionada existencia de las raíces cuadradas de $b^2 - 4ac$, que es lo que afirma el teorema siguiente:

Teorema 107 . *Para cada complejo $z = a + bi$ diferente de cero, existen (exactamente) dos raíces cuadradas complejas, (una inversa aditiva de la otra).*

Demostración. Supongamos que $w = x + yi$ es tal que $w^2 = z$. Entonces:

$$(x + yi)(x + yi) = x^2 - y^2 + 2xyi = a + bi$$

$$\therefore x^2 - y^2 = a, \quad 2xy = b.$$

Elevando al cuadrado cada ecuación y sumando:

$$x^4 - 2x^2y^2 + y^4 = a^2,$$

$$4x^2y^2 = b^2$$

$$x^4 + 2x^2y^2 + y^4 = a^2 + b^2$$

$$\therefore (x^2 + y^2)^2 = a^2 + b^2.$$

Entonces:

$$x^2 + y^2 = \sqrt{a^2 + b^2} \quad (= \|z\|)$$

$$x^2 - y^2 = a \quad (= \operatorname{Re}(z)).$$

$$\therefore x^2 = \frac{\sqrt{a^2 + b^2} + a}{2} \quad y^2 = \frac{\sqrt{a^2 + b^2} - a}{2}$$

Y como $\sqrt{a^2 + b^2} \geq |a|$, cada una de las expresiones de la derecha en estas igualdades tiene raíces cuadradas (reales) de donde resulta que

$$x = \pm \sqrt{\frac{\sqrt{a^2 + b^2} + a}{2}} \quad y = \pm \sqrt{\frac{\sqrt{a^2 + b^2} - a}{2}},$$

Para seleccionar la pareja de raíces que satisface nuestro problema -producir una raíz cuadrada de z - debe tenerse en cuenta que como $2xy = b$, si $b > 0$ deben escogerse x e y con signos iguales (ambos positivos o ambos negativos) y si $b < 0$, x e y deben tener signos diferentes. ■

Ejemplo 133 . Se desea encontrar las raíces cuadradas de $5 - 12i$. Entonces si $w = x + yi$ es una de ellas:

$$x^2 + y^2 = \sqrt{5^2 + 12^2} = 13.$$

$$x^2 - y^2 = 5.$$

Entonces:

$$x^2 = 9, \quad x = \pm 3,$$

$$y^2 = 4, \quad y = \pm 2,$$

y como b es menor que cero, x e y deben escogerse con signos diferentes. Entonces

$$w_1 = 3 - 2i, \quad w_2 = -3 + 2i$$

son las raíces cuadradas de z .

Ejemplo 134 . Se desea obtener las raíces de la siguiente ecuación: $z^2 - 3z + 3 - i = 0$.

Aquí, $a = 1, b = -3, c = 3 - i$ y por lo tanto $b^2 - 4ac = 9 - 4(3 - i) = -3 + 4i$.

Las raíces son: $z_1 = \frac{3 + \sqrt{-3 + 4i}}{2}$ y $z_2 = \frac{3 - \sqrt{-3 + 4i}}{2}$.

Calculemos $\sqrt{-3 + 4i}$. Si $(x + yi)^2 = -3 + 4i$, entonces $\begin{cases} x^2 + y^2 = 5 \\ x^2 - y^2 = -3 \end{cases} \therefore$

$x = \pm 1, y = \pm 2$ y como $2xy = 4 > 0$, las raíces resultan $1 + 2i$ y $-1 - 2i$.

Por lo tanto $z_1 = \frac{3 + 1 + 2i}{2}$ y $z_2 = \frac{3 - 1 - 2i}{2}$ es decir, $2 + i$ y $1 - i$.

En efecto, $z_1^2 - 3z_1 + 3 - i = (2 + i)^2 - 3(2 + i) + 3 - i = 4 + 4i - 1 - 6 - 3i + 3 - i = 0$ y $z_2^2 - 3z_2 + 3 - i = (1 - i)^2 - 3(1 - i) + 3 - i = 1 - 2i - 1 - 3 + 3i + 3 - i = 0$.

Ejemplo 135 . Encontremos las raíces de $z^2 - 2iz - 9 - 6i = 0$:

$$\text{una raíz es } z_1 = \frac{2i + \sqrt{(-2i)^2 - 4(-9 - 6i)}}{2} = \frac{2i + \sqrt{32 + 24i}}{2}.$$

Resolvamos ahora $(x + iy)^2 = 32 + 24i$:

$$\text{Entonces } \begin{aligned} x^2 + y^2 &= 40 = \sqrt{32^2 + 24^2} \\ x^2 - y^2 &= 32 \end{aligned}, \therefore x = \pm 6, y = \pm 2. \text{ Y como } 2xy =$$

$$24 > 0 \text{ entonces las raíces resultan } 6+2i \text{ y } -6-2i. \text{ Entonces } z_1 = \frac{2i + 6 + 2i}{2} =$$

$$3 + 2i, \text{ la otra raíz, } z_2 = \frac{2i - 6 - 2i}{2} = -3.$$

$$\text{En efecto, } z_1^2 - 2iz_1 - 9 - 6i = 5 + 12i + 4 - 6i - 9 - 6i = 0, \text{ y}$$

$$z_2^2 - 2iz_2 - 9 - 6i = 9 + 6i - 9 - 6i = 0.$$

Ejercicio 326 . Encuentre las raíces cuadradas de

1. $1 + \sqrt{3}i$.

2. $-1 - \sqrt{3}i$.

3. $2i$.

4. -16 .

5. $-2i$.

6. -1 .

7. $24 - 10i$.

8. $2 + 2\sqrt{3}i$.

9. $-8i$.

10. $5 - 12i$.

11. $3 - 4i$.

12. $3 + 4i$.

13. $-3 + 4i$.

14. $-3 - 4i$.

15. $12 + 5i$.

16. $15 + 8i$.

17. $-40 + 42i$.

Ejercicio 327 . Resuelva las ecuaciones siguientes:

1. $z^2 - 3z + 3 - i = 0$.

2. $z^2 - 2iz - 9 - 6i = 0$.

3. $z^2 - 3(1 + i)z + 5i = 0$.

4. $(1 + i)z^2 + (1 + 2i)z - 2 = 0$.

5. $z^2 - 2iz + 1 = 0$.

6. $-5x^2 + \sqrt{2}x - 1 = 0$.

7. $z^2 - (4 + i)z + 2 - 6i = 0$.

8. $z^2 - (5 - 3i)z - (5 + 5i) = 0$.

9. $z^2 + (2 - 3i)z - (5 + 5i) = 0$.

10. $z^6 + z^3 + 1 = 0$.

7.4.1 Sistemas de ecuaciones

Ejemplo 136 . Se desea resolver el sistema

$$\begin{aligned} iz + (1 + i)w &= 3i \\ (1 + i)\bar{z} - (6 + i)\bar{w} &= -12 + 3i. \end{aligned}$$

Observemos que las incógnitas en la segunda ecuación son las conjugadas de la primera, así que si la conjugamos el sistema queda:

$$\begin{aligned} iz + (1 + i)w &= 3i \\ (1 - i)z - (6 - i)w &= -12 - 3i. \end{aligned}$$

Multiplicando la primera ecuación por $-i$ y la segunda ecuación por $1+i$ obtenemos

$$\begin{aligned} z + (1-i)w &= 3 \\ 2z - (1+i)(6-i)w &= (1+i)(-12-3i), \end{aligned}$$

es decir

$$\begin{aligned} z + (1-i)w &= 3 \\ 2z - (7+5i)w &= -9-15i. \end{aligned}$$

Restando el doble de la primera ecuación a la segunda tenemos:

$$(-9-3i)w = -15-15i,$$

así que $w = \frac{-15-15i}{-9-3i} = 2+i$. Entonces $z + (1-i)(2+i) = 3$, por lo que $z = 3 - (1-i)(2+i) = i$.

Ejemplo 137 . Se desea resolver el sistema:

$$iz + 2w = 3 + 4i \quad (7.28)$$

$$2z - iw = 5 - 3i. \quad (7.29)$$

Calcúlese

$$\Delta = \begin{vmatrix} i & 2 \\ 2 & -i \end{vmatrix} = -3 \quad (7.30)$$

Por lo tanto el sistema tiene solución única, y

$$\Delta_z = \begin{vmatrix} 3+4i & 2 \\ 5-3i & -i \end{vmatrix} = -6+3i \quad \therefore z = \frac{-6+3i}{-3} = 2-i \quad (7.31)$$

$$\Delta_w = \begin{vmatrix} i & 3+4i \\ 2 & 5-3i \end{vmatrix} = -3-3i \quad \therefore w = \frac{-3-3i}{-3} = 1+i. \quad (7.32)$$

Comprobación:

$$i(2-i) + 2(1+i) = 3+4i \quad (7.33)$$

$$2(2-i) - i(1+i) = 5-3i \quad (7.34)$$

Ejemplo 138 . Se desea resolver el sistema

$$x + y + z = 4 + 2i \quad (7.35)$$

$$x + 2y - 2z = -4 - 2i \quad (7.36)$$

$$2x - 2y - z = -5 + 5i \quad (7.37)$$

Indicaremos las operaciones usando: R_2' (léase el nuevo renglón 2)

$R_2 - R_1$ (léase renglón 2 menos renglón 1)

La matriz aumentada es:

$$\begin{pmatrix} 1 & 1 & 1 & 4 + 2i \\ 1 & 2 & -2 & -4 - 2i \\ 2 & -2 & -1 & -5 + 5i \end{pmatrix} \quad (7.38)$$

ahora, si $R'_2 = R_2 - R_1$ y $R'_3 = R_3 - 2R_1$, obtenemos

$$\begin{pmatrix} 1 & 1 & 1 & 4 + 2i \\ 0 & 1 & -3 & -8 - 4i \\ 0 & -4 & -3 & -13 + i \end{pmatrix}, \quad (7.39)$$

ahora, si $R'_1 = R_1 - R_2$ y $R'_3 = \frac{R_3 + 4R_2}{-15}$, obtenemos

$$\begin{pmatrix} 1 & 0 & 4 & 12 + 6i \\ 0 & 1 & -3 & -8 - 4i \\ 0 & 0 & 1 & 3 + i \end{pmatrix}. \quad (7.40)$$

Por último, si $R'_1 = R_1 - 4R_3$ y $R'_2 = R_2 - 3R_3$, obtenemos

$$\begin{pmatrix} 1 & 0 & 0 & 2i \\ 0 & 1 & 0 & 1 - i \\ 0 & 0 & 1 & 3 + i \end{pmatrix}. \quad (7.41)$$

Por lo tanto $x = 2i$, $y = 1 - i$, $z = 3 + i$.

Ejercicio 328 . Resuelva:

$$1. z\bar{z} - 3(z + \bar{z}) = i.$$

$$2. 2z - \bar{z} = 4 - 2i.$$

3.
$$\begin{cases} 2z + 3w = 10 - 5i \\ \bar{z} - 6\bar{w} = -\left(\frac{37}{2}\right) + i. \end{cases}$$

4.
$$\begin{cases} 3z + w = 4 + 2i \\ 2z - iw = 3 + 2i. \end{cases}$$

5.
$$\begin{cases} (1+i)z + (1-i)w = 0 \\ (1-i)z + (1+i)w = 4. \end{cases}$$

6.
$$\begin{cases} z_1 + z_2 + z_3 = 6 + 4i \\ iz_1 + (1+i)z_2 + (1-i)z_3 = 7 + 4i \\ z_1 + iz_2 - z_3 = 2i. \end{cases}$$

7.
$$\begin{cases} z_1 + 2z_2 + 3z_3 = 1 - 2i \\ 4z_1 + 5z_2 + 6z_3 = 2 + i \\ 7z_1 + 8z_2 + 9z_3 = 3 + 4i. \end{cases}$$

7.5 Representación geométrica de los números complejos

Tal como se ha mencionado en párrafos anteriores, todo número complejo $a + bi$ puede hacerse corresponder con el punto del plano cuyas coordenadas son a y b . Cuando se usa esta representación, el eje X se conoce como el eje real y el Y es el imaginario. Cada número complejo z puede considerarse como la suma $a + bi$, la pareja (a, b) , el punto del plano $P = (a, b)$, o el vector apoyado en el origen de extremo P , y si la longitud del segmento es r y el ángulo que forma con el eje X es θ , el complejo z puede expresarse también en coordenadas polares (r, θ) en donde, por supuesto, $a = r \cos(\theta)$, $b = r \sin(\theta)$. El cambio inverso -rectangulares a polares- está dado por las relaciones:

$$r = \sqrt{a^2 + b^2},$$

$$\theta = \begin{cases} \arctan(b/a) & \text{si } a \neq 0 \\ \frac{\pi}{2} & \text{si } a = 0, b > 0 \\ -\frac{\pi}{2} & \text{si } a = 0, b < 0 \\ \text{no está definido} & \text{si } a = b = 0 \end{cases}$$

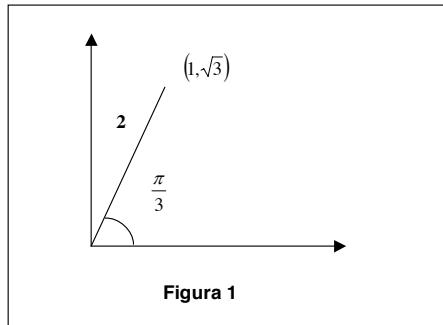


Figura 1

Figura 7.1:

Equivalentemente, se define θ como el ángulo cuyo coseno es $\frac{a}{\sqrt{a^2 + b^2}}$ y cuyo seno es $\frac{b}{\sqrt{a^2 + b^2}}$. θ no está definido si $a^2 + b^2 = 0$, es decir si $z = \vec{0}$.

Puede verse que esta manera de definir el argumento de z , permite la posibilidad de que dos ángulos θ_1 y θ_2 sean argumentos de z si θ_1 y θ_2 difieren en un múltiplo entero de 2π y sólo entonces, (ver “Argumento de un número complejo” más adelante). Observación que resultará conveniente en lo que sigue. ($\cos(\theta_1) = \cos(\theta_2)$ y $\sin(\theta_1) = \sin(\theta_2) \Rightarrow \theta_2 = \theta_1 + 2k\pi$, $k \in \mathbb{Z}$).

7.5.1 Pasar de coordenadas rectangulares a forma polar

Notación 12 . $cis(\theta)$ es abreviatura de $\cos(\theta) + i\sin(\theta)$.

Así $cis(60^\circ) = \cos(60^\circ) + i\sin(60^\circ)$.

Ejemplo 139 . Considérese el vector (en el plano) de módulo 2 y argumento $\pi/3$, y calculemos sus formas rectangular y polar.

Entonces sus coordenadas rectangulares son 1 y $\sqrt{3}$. Por lo tanto su forma rectangular es $1 + \sqrt{3}i$ y su forma polar es $2cis(\pi/3)$ ó $2cis(60^\circ)$.

Ejemplo 140 . Calcular las formas rectangular y polar de z de tamaño 4 y argumento $\pi/4$ (ó 45°). Entonces la forma rectangular es $2\sqrt{2} + 2\sqrt{2}i$ y la forma polar es $4cis(\pi/4)$.

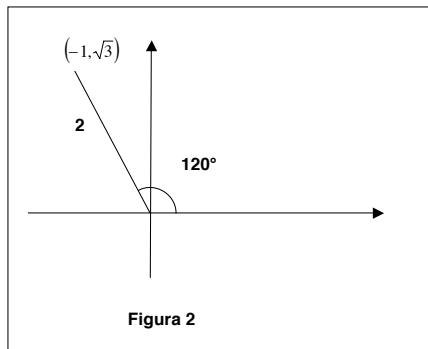


Figura 7.2:

Ejemplo 141 . Pasar de $4\sqrt{3} + 4i$ a forma polar. (También se sugiere empezar con un dibujo). Como se puede ver, el tamaño resulta $\sqrt{(4\sqrt{3})^2 + 4^2} = \sqrt{64} = 8$ y el argumento $\pi/6$ (ó 30°) y entonces, la forma polar queda: $z = 8\text{cis } 30^\circ = 8(\cos 30^\circ + i \sin 30^\circ)$.

Ejemplo 142 . Pasar de $2\text{cis } 120^\circ$ a forma rectangular. Entonces, como puede verse, las coordenadas de z son $(-1, \sqrt{3})$ y por lo tanto $z = -1 + \sqrt{3}i$.

Ejercicio 329 . Calcular las coordenadas de los puntos del plano cuyo módulo es r y cuyo argumento es s :

1. $r = 3\sqrt{2}$, $s = 225^\circ$.
2. $r = 2$, $s = 30^\circ$.
3. $r = 3$, $s = 90^\circ$.
4. $r = 2$, $s = 270^\circ$.
5. $r = \sqrt{2}$, $s = 45^\circ$.
6. $r = 4$, $s = 120^\circ$.
7. $r = 2$, $s = 300$.

Ejercicio 330 . Expresar en coordenadas polares:

1. $7 + 7i$.
2. $\sqrt{3} - i$.
3. $-5 + 5i$.
4. $1 - i\sqrt{3}$.
5. $4 + 2i$.
6. $-6 - 6\sqrt{3}i$.
7. -4 .
8. $8i$.

Ejercicio 331 *Calcule las potencias siguientes:*

1. $(1 + i)^8$.
2. $(3 + 4i)^3$.
3. $(5 - 12i)^2$.

Una de las ventajas que tiene la representación geométrica de los números complejos es que permite estudiar la geometría del plano a través del álgebra de \mathbb{C} .

La suma de dos complejos z_1, z_2 como la diagonal del paralelogramo construido sobre los sumandos. Además la multiplicación se interpreta en la forma que expresa el siguiente

Teorema 108 . *Sean $z_j = r_j(\cos\theta_j + i\sin\theta_j)$, $j \in \{1, 2\}$ dos números complejos (que por comodidad denotaremos (algunas veces) en forma abreviada como*

$$z_j = r_j \operatorname{cis}(\theta_j)$$

Entonces

$$z_1 z_2 = r_1 r_2 \operatorname{cis}(\theta_1 + \theta_2).$$

(Este teorema dice que para multiplicar dos complejos, se multiplican sus tamaños y se suman sus argumentos).

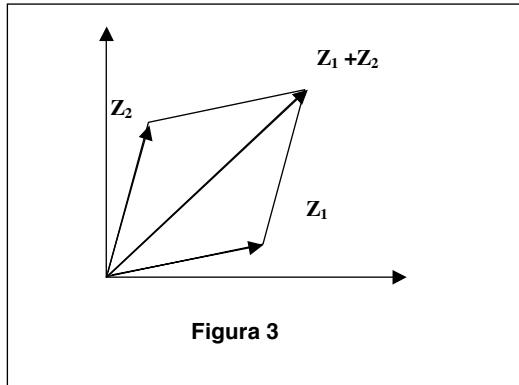


Figura 7.3:

Demostración.

$$\begin{aligned}
 & (r_1 \cos \theta_1 + i r_1 \sin \theta_1)(r_2 \cos \theta_2 + i r_2 \sin \theta_2) = \\
 & = r_1 r_2 \cos \theta_1 \cos \theta_2 - \sin \theta_1 \sin \theta_2 + i(\sin \theta_1 \cos \theta_2 + \sin \theta_2 \cos \theta_1) = \\
 & = r_1 r_2 (\cos(\theta_1 + \theta_2) + i \sin(\theta_1 + \theta_2)) = r_1 r_2 \operatorname{cis}(\theta_1 + \theta_2).
 \end{aligned}$$

■

Ejercicio 332 . Muestre que si $z \in \mathbb{C} \setminus \{0\}$, entonces $(z^n)^{-1} = z^{-n}$, $\forall n \in \mathbb{Z}$.

Corolario 14 (teorema de De Moivre) . Si $z = r \operatorname{cis} \theta$ es un número complejo, entonces $\forall n \in \mathbb{Z}$, $z^n = r^n \operatorname{cis}(n\theta)$.

Demostración. (Inducción sobre n).

La base, $n = 1$, es obvia.

Paso inductivo: $T(k) \Rightarrow T(k + 1)$.

Supóngase que el teorema vale para $n = k$ e. d. $z^k = r^k \operatorname{cis}(k\theta)$ y aplíquese el teorema anterior. Entonces $z^{k+1} = z^k z = r^k \operatorname{cis}(k\theta) \cdot r \operatorname{cis}(\theta) = r^{k+1} \operatorname{cis}((k+1)\theta)$.

Paso inductivo aumentado: $T(n) \Rightarrow T(-n)$:

$$\begin{aligned}
 z^{-n} &= (z^n)^{-1} = \overline{z^n} / \|z^n\|^2 = \\
 &= \frac{r^n(\cos(n\theta) - i\sin(n\theta))}{r^{2n}} = \\
 &= r^{-n}(\cos(-n\theta) + i\sin(-n\theta)) = \\
 &= r^{-n} \text{cis}(-n\theta).
 \end{aligned}$$

■

Ejemplo 143 . $2(\cos 25^\circ + i \sin 25^\circ) \cdot 5(\cos 75^\circ + i \sin 75^\circ) = 10(\cos 100^\circ + i \sin 100^\circ) = 10 \text{ cis } 100^\circ$.

Ejemplo 144 . $(15 \text{cis } 140^\circ) \cdot (4 \text{cis } 280^\circ) = 60 \cdot \text{cis } (420^\circ) = 60 \text{cis } (60^\circ)$.

Ejemplo 145 . $(1+i)^5 = (\sqrt{2} \text{cis } 45^\circ)^5 = 2^{5/2} \text{cis } (5 \cdot (45^\circ)) = 4\sqrt{2} \text{cis } (225^\circ)$.

Ahora lo hacemos desarrollando el binomio:

$$(1+i)^5 = 1 + 5i + 10i^2 + 10i^3 + 5i^4 + i^5 = \quad (7.42)$$

$$= 1 + 5i - 10 - 10i + 5 + i = \quad (7.43)$$

$$= -4 - 4i. \quad (7.44)$$

Pasando a forma polar este resultado:

$$\|z\| = \sqrt{16+16} = 4\sqrt{2} \text{ y } \theta = 225^\circ, \quad (7.45)$$

es decir que $(1+i)^5 = 4\sqrt{2} \text{cis } (225^\circ)$.

En este ejemplo se nota que es mucho más corto sacar potencias de un número complejo en forma polar usando el teorema de De Moivre que desarrollando el binomio en forma rectangular (trate de calcular $(1+i)^{200}$ en forma rectangular).

Ejercicio 333 . Realice las operaciones indicadas. En el caso de los ejercicios que están en forma rectangular, primero habrá que reducirlos a forma polar.

1. $2(\cos 25^\circ + i \sin 25^\circ) 5(\cos 75^\circ + i \sin 75^\circ)$.

2. $8(\cos 18^\circ + i \sin 18^\circ)6(\cos 24^\circ + i \sin 24^\circ)$.
3. $15(\cos 140^\circ + i \sin 140^\circ)4(\cos 280^\circ + i \sin 280^\circ)$.
4. $(1 + i\sqrt{3})(4 + 4i)$.
5. $(6 - 6i)(-7 + 7\sqrt{3}i)$.
6. $2(\cos 15^\circ + i \sin 15^\circ)3(\cos 70^\circ + i \sin 70^\circ)4(\cos 65^\circ + i \sin 65^\circ)$.

Ejercicio 334 . Use el teorema de De Moivre para elevar a la potencia indicada

1. $[4(\cos 12^\circ + i \sin 12^\circ)]^3$.
2. $[3(\cos 18^\circ + i \sin 18^\circ)]^4$.
3. $[\sqrt{3}(\cos 25^\circ + i \sin 25^\circ)]^6$.
4. $[2(\cos 20^\circ + i \sin 20^\circ)]^8$.
5. $(1 - i)^{10}$.
6. $(1 + i\sqrt{3})^9$.
7. $(-2\sqrt{3} + 2i)^8$.
8. $(\sqrt{3} - i)^{12}$.
9. $[5(\cos 16^\circ + i \sin 16^\circ)]^{-4}$.
10. $[\sqrt{7}(\cos 20^\circ + i \sin 20^\circ)]^8$.
11. $[\cos(-10^\circ) + i \sin(-10^\circ)]^6$.
12. $(\cos 10^\circ + i \sin 10^\circ)^{-6}$.
13. $(\sqrt{3} + i)^{-10}$.
14. $(-1 - i)^{-12}$.

7.6 Raíces n -ésimas de un número complejo

Supóngase que se desea encontrar todos los complejos w tales que $w^n = z$, y que $z = \rho \cdot cis(\theta)$. Supongamos además que $w_0 = rcis(\varphi)$ es uno de ellos. Entonces, por el teorema de De Moivre,

$$w_0^n = r^n cis(n\varphi) = \rho cis(\theta), \therefore r = \sqrt[n]{\rho} \text{ (la raíz real).}$$

y

$$\varphi = \frac{\theta}{n}.$$

Es decir

$$w_0 = \sqrt[n]{\rho} cis\left(\frac{\theta}{n}\right).$$

Considerando que si θ es un argumento de z , los números $\theta + 2\pi k, k \in \mathbb{Z}$, también lo son, se encuentran las otras $n - 1$ raíces de z , sumando a φ , los números $\frac{2\pi k}{n}, k = 1, \dots, n - 1$.

Así por ejemplo si $z = 1$ ($= 1cis(0)$), sus raíces cúbicas (que son tres), son:

$$\begin{aligned} r_1 &= \sqrt[3]{1} cis\left(\frac{0}{3}\right) = 1 \\ r_2 &= \sqrt[3]{1} cis\left(0 + \frac{2\pi}{3}\right) = -\frac{1}{2} + \frac{\sqrt{3}}{2}i = \omega \\ r_3 &= 1 cis\left(0 + 2 \cdot 2\pi/3\right) = -\frac{1}{2} - \frac{\sqrt{3}}{2}i = \omega^2. \end{aligned}$$

Por ejemplo $z = 32cis150^\circ$ sus raíces quintas son:

$$r_1 = 2cis\frac{150^\circ}{5} = 2cis30^\circ$$

$$r_2 = 2cis\left(30^\circ + \frac{2\pi}{5}\right) = 2cis(30^\circ + 72^\circ) = 2cis102^\circ$$

$$r_3 = 2cis[30^\circ + 2(72^\circ)] = 2cis174^\circ$$

$$r_4 = 2cis[30^\circ + 3(72^\circ)] = 2cis246^\circ$$

$$r_5 = 2cis[30^\circ + 4(72^\circ)] = 2cis318^\circ$$

Justificamos las afirmaciones anteriores con el siguiente:

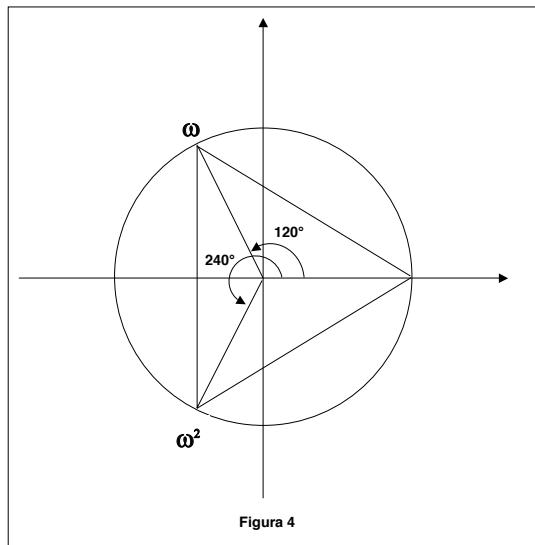


Figura 7.4:

Teorema 109 . Sea $z = \rho cis(\theta)$ un número complejo diferente de cero. Entonces para cada n entero positivo, z tiene exactamente n raíces n -ésimas, que están dadas por la fórmula:

$$w_r = \sqrt[n]{\rho} cis(\varphi_k) \quad \varphi_k = \frac{\theta + 2\pi k}{n}, \quad k \in \{0, 1, \dots, n-1\}.$$

(Para fines prácticos, los ángulos pueden expresarse en grados, y en este caso, $\varphi_k = \frac{\theta + 360k}{n}$, $k \in \{0, 1, \dots, n-1\}$).

Demostración. Sea

$$\beta = \{w_0, \dots, w_{n-1}\}$$

el conjunto formado por los n valores de la fórmula que corresponden a $k = 0, 1, \dots, n-1$. Entonces:

- 1) Cada w_i es raíz n -ésima de z y
- 2) Si $i \neq j$, entonces $w_i \neq w_j$ (β tiene exactamente n elementos).

En efecto:

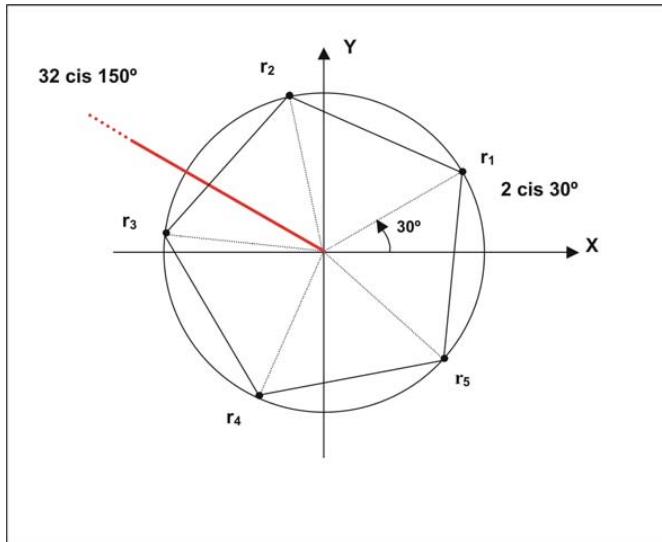


Figura 7.5:

$$1) (w_k)^n = (\rho)^n cis \left(n \frac{\theta + 2\pi k}{n} \right) = \rho cis(\theta + 2k) = z.$$

2) Supóngase $w_i = w_j$ $0 \leq j \leq i < n$.

Entonces $\sqrt[n]{\rho} cis(\varphi_i) = \sqrt[n]{\rho} cis(\varphi_j)$ y por lo tanto φ_i y φ_j difieren, necesariamente, en un múltiplo entero de 2π es decir:

$$\varphi_i - \varphi_j = 2\pi k = \frac{\theta + 2\pi i}{n} - \frac{\theta + 2\pi j}{n} = 2\pi \frac{i - j}{n},$$

Por lo tanto $k = \frac{i-j}{n}$, que debe ser entero, dice que $n \mid i - j$ pero $0 \leq i - j < n$ $\therefore i - j = 0, i = j$.

En resumen, $w_i = w_j \Rightarrow i = j$ y, por contrapuesta,

$$i \neq j \Rightarrow w_i \neq w_j.$$

Luego la cardinalidad de β es n y dado que la ecuación $x^n - z = 0$ no puede tener más de n raíces en \mathbb{C} , entonces β consta de todas las raíces n -ésimas de z . ■

Ejercicio 335 . Obtenga las raíces que se indican.

1. Raíces cúbicas de $8(\cos 72^\circ + i \sin 72^\circ)$.
2. Raíces cúbicas de $216(\cos 27^\circ + i \sin 27^\circ)$.
3. Raíces cuadradas de $(1 - i\sqrt{3})$.
4. Raíces cúbicas de $(1 - i\sqrt{3})$.
5. Raíces cuadradas de $(-1 - i\sqrt{3})$.
6. Raíces cúbicas de 1.
7. Raíces cúbicas de 8.
8. Raíces cúbicas de -1 .
9. Raíces cúbicas de i .
10. Raíces cúbicas de $(3 + 4i)$.
11. Raíces cuartas de $16\sqrt{2}(-1 + i)$.
12. Raíces quintas de 1.
13. Raíces sextas de $-27i$.
14. Raíces quintas de $16\sqrt{2}(-1 + i)$.

7.7 El argumento de un número complejo

Cuando se cambian las coordenadas polares a rectangulares en un complejo $z = (r, \theta)$, estas últimas quedan bien determinadas por las expresiones $x = r \cos(\theta)$ y $y = r \sin(\theta)$ pero el cambio inverso, de rectangulares a polares (en el que si $z = (x, y)$, entonces r resulta ser $r = (x^2 + y^2)^{\frac{1}{2}}$), sólo determina θ módulo 2π . Es decir: si θ satisface $\sin(\theta) = \frac{y}{r}$, $\cos \theta = \frac{x}{r}$, ($r \neq 0$, por supuesto), entonces ϕ satisfará las mismas relaciones si y sólo si $\phi = \theta + 2\pi k$, $k \in \mathbb{Z}$. Es decir si y solamente si θ y ϕ difieren en un múltiplo entero de 2π . Para evitar la ambigüedad que esta situación ocasiona, se *conviene* en definir $\theta = \operatorname{Arg} z$ como el único real con las propiedades:

$$x = r \cos \theta, \quad y = r \sin \theta, \quad -\pi < \theta \leq \pi \quad (z \neq 0). \quad (7.46)$$

Se sabe (Teorema de De Moivre) que para multiplicar 2 complejos, se multiplican (en \mathbb{R}) sus tamaños y se “suman sus argumentos”, pero para ser congruentes con la convención anterior, a esta suma se le debe aplicar una corrección para el caso en que no caiga dentro del rango acordado $(-\pi, \pi]$. De modo que si :

$$z_j = r_j cis \theta_j \quad (\operatorname{Arg}(z_j) = \theta_j) \quad j = 1, 2, \quad (7.47)$$

se tiene que :

$$\operatorname{Arg}(z_1 z_2) = \theta_1 + \theta_2 + 2\pi c(z_1, z_2), \quad (7.48)$$

en donde :

$$c(z_1, z_2) = \begin{cases} 1 & \text{si } \theta_1 + \theta_2 \leq -\pi \\ 0 & \text{si } -\pi < \theta_1 + \theta_2 \leq \pi \\ -1 & \text{si } \theta_1 + \theta_2 > \pi \end{cases} \quad (7.49)$$

Así por ejemplo si

$$z_1 = z_2 = -1; \quad -1 = 1 cis \pi, \quad (7.50)$$

entonces $c(z_1, z_2) = -1$ y

$$(-1)(-1) = 1 = 1 cis(2\pi - 2\pi) = 1 cis(0). \quad (7.51)$$

7.8 Algunas transformaciones del plano

Sea $z \in \mathbb{C}$ debe ser claro que la función $\begin{array}{ccc} z + & : \mathbb{C} & \longrightarrow \mathbb{C} \\ w & \longmapsto & z + w \end{array}$ corresponde con una traslación (la que envía w a $z + w$).

7.8.1 Contracciones y expansiones

Es claro que multiplicar por un número real produce una expansión si este número es mayor que uno y una contracción si el número real es menor que uno.

7.8.2 Rotaciones

Por lo que vimos, multiplicar por el complejo $\text{cis}(\theta)$, produce una rotación por un ángulo θ alrededor del origen.

Ejercicio 336 . *Demuestre que una rotación por un ángulo θ alrededor del complejo w está dada por*

$$(w + _) \circ (\text{cis}(\theta) \cdot _) \circ (-w + _).$$

⁴De tal manera que el efecto sobre el complejo z es

$$z \longmapsto w + (\text{cis}(\theta) \cdot (-w + z)).$$

7.8.3 Reflexión sobre el eje X

Es claro que tal reflexión se tiene mediante la conjugación. Ahora supongamos que queremos reflejar sobre una línea ℓ que pasa por el origen y que hace un ángulo α con el eje X . Para producir el efecto de reflejar sobre esta línea ℓ , podemos primero aplicar una rotación por un ángulo $-\alpha$ (es decir, multiplicando por $\text{cis}(-\alpha)$, luego conjugamos (es decir, aplicamos $\overline{()}$) y por último, multiplicamos por $\text{cis}(\alpha)$.

Para entender mejor lo que hemos hecho, apliquemos las transformaciones anteriores al complejo z :

$$\begin{aligned} \text{cis}\alpha \cdot \overline{(\text{cis}(-\alpha) \cdot z)} &= \text{cis}\alpha \cdot \left(\overline{\text{cis}(-\alpha)} \cdot \overline{z} \right) = \\ &= \text{cis}\alpha \cdot (\text{cis}(\alpha) \cdot \overline{z}) = (\text{cis}2\alpha) \cdot \overline{z}. \end{aligned}$$

Por lo que $\sigma_\ell = (\text{cis}(2\alpha)) \cdot \overline{()}$. Es decir, σ_ℓ la operación de reflejar sobre la línea ℓ , es equivalente a: primero conjugar seguida de rotar por un ángulo de 2α .

Ejercicio 337 . *Describa el efecto de reflejar sobre una línea que pasa por el punto w ($w \neq (0,0)$), usando traslaciones, reflexiones sobre líneas que pasan sobre el origen y rotaciones.*

Ejercicio 338 . *Demuestre que la composición de dos reflexiones sobre líneas que pasan por el origen es una rotación. Describa el ángulo de la rotación en términos de los ángulos que hacen las líneas que definen las reflexiones, con el eje Y .*

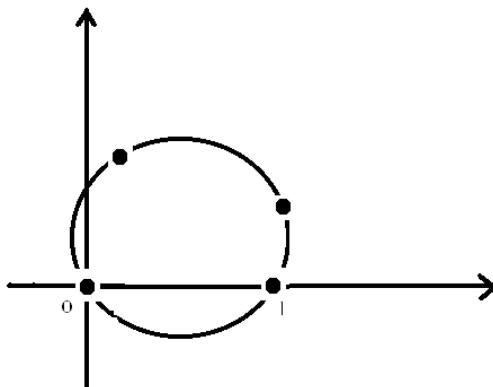
⁴ $\text{cis}(\theta) \cdot _$ denota la multiplicación por $\text{cis}(\theta)$.

7.8.4 Reflexión respecto al origen

Es claro que la función que manda un complejo z a su inverso aditivo, se puede interpretar como una reflexión respecto al origen.

Ejercicio 339 . *Demuestre que multiplicar por un complejo $z \neq (0,0)$ es una operación que manda círculos en círculos. Diga explícitamente en qué círculo se convierte el círculo que tiene centro en w y tiene radio s .*

Ejercicio 340 . *Encuentre la condición que describe cuando dos complejos z y w están en el mismo círculo junto con 0 y 1.*



Ejercicio 341 . *Encuentre una condición para que cuatro puntos estén en el mismo círculo.*

7.9 La función exponencial compleja

Una observación importante:

Observación 89 *Sea $z : [a, b] \rightarrow \mathbb{C}$ la descripción de una trayectoria en \mathbb{R}^2 . Entonces, interpretando la variable t como “el tiempo”, $z(t) = x(t) + iy(t)$ muestra que tanto la parte real como la imaginaria de z son funciones que dependen de t , y entonces, la derivada de z con respecto a t , debe definirse como:*

$$z' = x'(t) + iy'(t). \quad (7.52)$$

Así por ejemplo si

$$z(t) = \cos(t) + i\sin(t) \quad (7.53)$$

(z describe una rotación alrededor del origen, con radio uno), la velocidad del movimiento -la derivada de z con respecto a t - es

$$z'(t) = -\sin(t) + i\cos(t), \quad (7.54)$$

y del mismo modo, la aceleración resulta:

$$z''(t) = -\cos(t) + i\sin(t). \quad (7.55)$$

Se desea extender $\exp : \mathbb{R} \rightarrow \mathbb{R}$ a $E : \mathbb{C} \rightarrow \mathbb{C}$ de manera que se conserven las propiedades básicas de la exponencial. Explícitamente se desea que E tenga las propiedades siguientes: (Ver apéndice al final).

1. $E(x) = \exp(x), \forall x \in \mathbb{R}.$
2. $\forall z, w \in \mathbb{C}, E(z + w) = E(z)E(w).$
3. $E'(z) = E(z) \cdot z'.$

En vista de esto, si $z = x + iy, x, y \in \mathbb{R}$, debe suceder que

$$E(z) = E(x + iy) = E(x)E(iy) = e^x E(iy) \quad (7.56)$$

y por lo tanto, nuestro problema -encontrar la forma correcta de definir $E(z)$ - se reduce a decidir la manera en que debe interpretarse $E(iy)$ que obviamente es un complejo cuyas partes real e imaginaria dependen de y . Es decir $E(iy) = U(y) + iV(y)$ y se debe encontrar funciones U y V que resulten adecuadas para nuestro propósito (conseguir que E tenga las propiedades deseadas 1), 2) y 3)).

Entonces:

1. $E(iy) = U(y) + iV(y)$ por definición.
Derivando, en el supuesto de que E satisface 3,
2. $E'(iy) = iE(iy) = U'(y) + iV'(y)$ (en donde U' y V' son derivadas con respecto a su variable y).
Derivando una vez más:

3. $E''(iy) = -E(iy) = U''(y) + iV''(y) = -U(y) - iV(y)$, resultado que muestra que tanto U como V son funciones que satisfacen la ecuación $f'' + f = 0$.

Haciendo $y = 0$ en (1) y (2), se obtiene:

$$4. \quad 1 = U(0) + iV(0) \therefore U(0) = 1; V(0) = 0.$$

$$5. \quad i = U'(0) + iV'(0) \therefore U'(0) = 0; V'(0) = 1,$$

y como la solución de cada problema de valores iniciales es única, entonces U y V tienen que ser, necesariamente, $\cos(y)$ y $\operatorname{sen}(y)$ respectivamente. Luego:

$$E(x + iy) = e^x(\cos(y) + i \cdot \operatorname{sen}(y)) = e^x \cdot \operatorname{cis}(y). \quad (7.57)$$

Teorema 110 . La función $E : \mathbb{C} \rightarrow \mathbb{C}$ así definida tiene las propiedades (1) (2) y (3).

Demostración. 1) $E(x) = \exp(x) \ \forall x \in \mathbb{R}$:

Si $z \in \mathbb{R}$, entonces $z = x + 0i$, y $E(z) = \exp(x)\operatorname{cis}(0)$, pero $\operatorname{cis}(0) = 1$

$$\therefore E(z) = E(x) = \exp(x).$$

2) $E(z_1 + z_2) = E(z_1) + E(z_2)$:

Sea $z_j = x_j + iy_j$, $j \in \{1, 2\}$ Entonces:

$$z_1 + z_2 = (x_1 + x_2) + i(y_1 + y_2)$$

$$\therefore E(z_1 + z_2) = \exp(x_1 + x_2)(\operatorname{cis}(y_1 + y_2)) =$$

$$= e^{x_1}e^{x_2}\operatorname{cis}(y_1 + y_2) = (e^{x_1}\operatorname{cis}(y_1))(e^{x_2}\operatorname{cis}(y_2)) = E(z_1)E(z_2).$$

3) $E'(z) = E(z)z'$:

Sea $z(t) = x(t) + iy(t)$

$$\therefore E(z) = e^x(\cos(y) + i\operatorname{sen}(y)),$$

$$E'(z) = x'e^x(\cos(y) + i\operatorname{sen}(y)) + y'e^x(-\operatorname{sen}(y) + i\cos(y)).$$

Haciendo $-\operatorname{sen}(y) = i^2\operatorname{sen}(y)$,

$$E'(z) = e^x(\cos + i\operatorname{sen})(x' + iy') = E(z) \cdot z'.$$

■

Teorema 111 . *La función exponencial es periódica, y cada período es de la forma: $2k\pi$, $k \in \mathbb{Z}$.*

Demostración. Sea w un período de E , luego, $\forall z \in \mathbb{C}$, $E(z) = E(z+w)$.

Haciendo $z = 0$, entonces $1 = E(w)$.

Si $w = x + yi$, entonces $E(w) = e^x \operatorname{cis}(y)$, así que

$$e^x = 1 \quad (\because x = 0), \quad \operatorname{cis}(y) = 1,$$

es decir $\operatorname{sen}(y) = 0$, entonces $y = 2\pi k$, $k \in \mathbb{Z}$, luego $w = 2\pi ki$, $k \in \mathbb{Z}$. ■

Una consecuencia de la demostración es que

$$E(z) = 1 \Leftrightarrow z = 2\pi ki, \quad k \in \mathbb{Z}.$$

7.9.1 Representación geométrica de algunas rectas bajo la transformación E

Consideremos la transformación: $E : \mathbb{C} \rightarrow \mathbb{C}$

Obsérvese que el origen de \mathbb{C} , va a dar al punto $(1, 0)$.

A medida que la variable recorre el eje X alejándose del origen en el sentido positivo, e^x aumenta -exponencialmente-, pero y se mantiene igual a cero, luego la imagen de $[0, \infty)$ es $[1, \infty)$, mientras que $E(-\infty, 0)$ es $(0, 1)$. De la misma manera puede analizarse la imagen bajo la transformación E de cualquier recta horizontal ($y = cte$). Se encuentra que la imagen de la semirrecta cuyos puntos tienen abscisa positiva, es la parte que queda afuera del círculo unitario, del rayo que parte del origen y cuyo argumento es la y de la recta. La semirrecta de puntos con x negativa tiene por imagen la porción del rayo que queda dentro del círculo.

Nótese que como $E(z) \cdot E(-z) = 1$, $\forall z \in \mathbb{C}$, el origen del plano \mathbb{C} no es imagen de ningún complejo z bajo la transformación “exponencial”, y por lo tanto ($E(z) \neq 0 \quad \forall z \in \mathbb{C}$).

Las rectas verticales $x = cte$ (tamaño fijo y argumento de $-\infty$ a ∞), van a dar a circunferencias de radio e^x y las rectas que parten del origen se “retratan” como espirales que “arrancan” de $(1, 0)$. (Las imágenes van aumentando tanto de tamaño como de argumento, a medida que la variable se va alejando del origen).

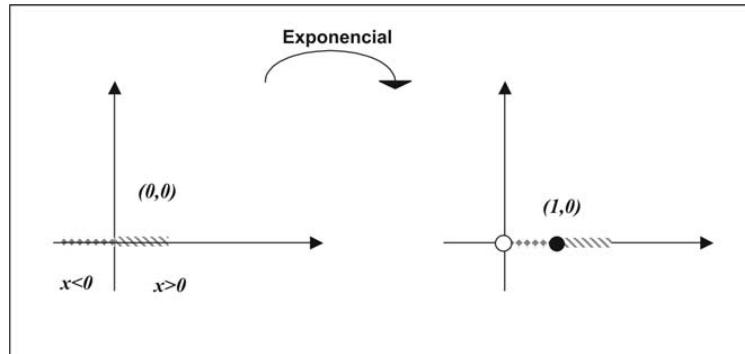


Figura 7.6:

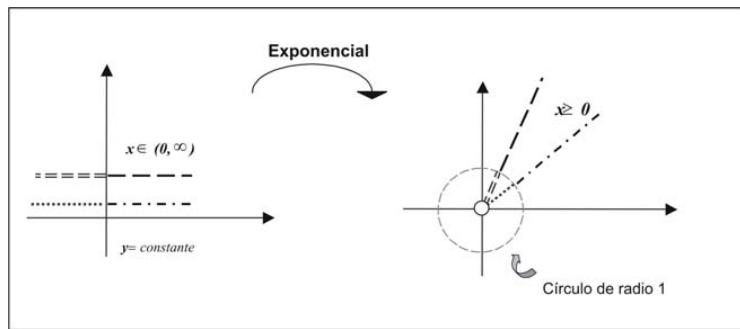
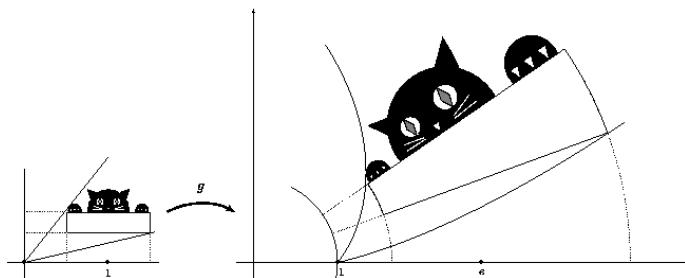
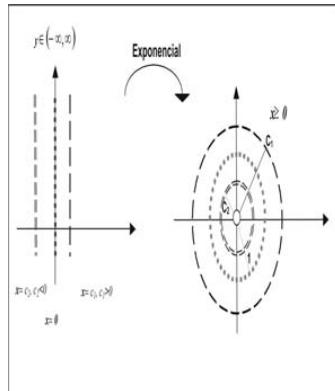


Figura 7.7:

Ilustración de la manera en que el plano complejo es transformado por la función $z \mapsto e^z$.

7.9.2 La función logaritmo

Con el deseo de definir la función inversa de la exponencial, -el logaritmo complejo- observamos que si $z = x + yi$, $E(z) = e^x cis(y)$, (por supuesto, $\|E(z)\| = e^x$, $\operatorname{Arg}(E(z)) = y$). Por lo tanto debe definirse:

$$L(z) = \ln \|z\| + i(\operatorname{Arg}(z)).$$

Nótese que si $z = x$, $x \in \mathbb{R}^+$, $L(z) = \ln|x| + 0i = \ln(x)$ e. d., L es una extensión de \ln . (si $z = -x$, $L(z) = \ln|x| + \pi i$.

Nótese también que siendo la exponencial una función periódica, debe escogerse una “banda” del plano de ancho 2π para seleccionar en ella los argumentos de los logaritmos complejos.

En efecto, si, por ejemplo escogemos la región

$$\{(x, y) \in \mathbb{C} \mid -\pi < y \leq \pi\}, \quad (7.58)$$

entonces, la función exponencial definida en ella, la “mapea” biyectivamente en todo el plano menos el origen.

Otra vez: Si se define

$$\alpha = \{x + yi \in \mathbb{C} \mid -\pi < y \leq \pi\},$$

entonces

$$E : \alpha \rightarrow \mathbb{C} \setminus \{0\}$$

es biyectiva y su inversa es

$$L : \mathbb{C} \setminus \{0\} \rightarrow \alpha.$$

En efecto, $\forall z \in \alpha$, $z = x + yi$, $E(z) = e^x cis(y)$;

$$L(e^x cis(y)) = \ln(e^x) + iy = x + yi$$

y si $z = x + yi$ es un complejo no cero de tamaño r y argumento θ ($x = r \cos \theta$, $y = r \sin \theta$) entonces $L(z) = \ln r + i\theta$ y por lo tanto $E(L(z)) = e^{\ln(r)} cis(\theta) = r \cos(\theta) + i r \sin(\theta) = x + yi = z$.

Obsérvese que la función $L : \mathbb{C} \setminus \{0\} \rightarrow \alpha$ está bien definida (en el sentido de que $\forall z \neq 0$, $L(z) \in \alpha$ ya que, según convinimos, el argumento de $L(z)$ debe pertenecer al intervalo $(-\pi, \pi]$).

Ejemplo 146 . $L(i) = \ln\|i\| + i \operatorname{Arg}(i) = (\pi/2)i$.
 $(-1) = \ln\|-1\| + i \operatorname{Arg}(-1) = \pi i$

Teorema 112 . Si $z_1 = r_1 cis \theta_1$, $z_2 = r_2 cis \theta_2$, entonces

$$L(z_1 z_2) = L(z_1) + L(z_2) + 2\pi i c(z_1, z_2) \quad (7.59)$$

Demostración.

$$L(z_1 z_2) = L(r_1 r_2 \operatorname{cis}(\theta_1 + \theta_2 + 2\pi c(z_1, z_2))) = \quad (7.60)$$

$$= \ln(r_1 r_2) + i((\theta_1 + \theta_2 + 2\pi c(z_1, z_2))) = \quad (7.61)$$

$$= \ln r_1 + i\theta_1 + \ln r_2 + i\theta_2 + 2\pi i c(z_1, z_2) = \quad (7.62)$$

$$= L(z_1) + L(z_2) + 2\pi i c(z_1, z_2). \quad (7.63)$$

■

Definición 93. Si $z \neq 0, w \in \mathbb{C}, z^w = E(wL(z))$.

En el caso de que $z = x \in \mathbb{R}^+$ y $w = y \in \mathbb{R}$, entonces $x^y = E(yL(x)) = E(y \ln(x)) = \exp(y \ln(x))$ por lo que, como puede verse, la definición anterior extiende a la que se tenía para \mathbb{R} .

1. $\forall z \neq \vec{0}, w_1, w_2 \in \mathbb{C}, z^{w_1+w_2} = z^{w_1} \cdot z^{w_2}$.
2. $(z^{w_1})^{w_2} = z^{w_1 w_2}$
3. Si z_1 y z_2 son distintos de cero, $w \in \mathbb{C}, (z_1 z_2)^w = z_1^w z_2^w E(2\pi w i c(z_1, z_2))$:

Demostración. 1) $z^{w_1+w_2} = E((w_1 + w_2)L(z)) = E(w_1 L(z) + w_2 L(z)) = E(w_1 L(z))E(w_2 L(z)) = z^{w_1} z^{w_2}$
 2)

$$(z^{w_1})^{w_2} = E(w_2 L(z^{w_1})) = E(w_2 L(E(w_1 L(z)))) = E(w_2 \cdot w_1 L(z)) = z^{w_1 w_2}. \quad (7.64)$$

3)

$$(z_1 z_2)^w = E(wL(z_1 z_2)) = E[w(L(z_1) + L(z_2) + 2\pi i c(z_1, z_2))] = \quad (7.65)$$

$$= E[wL(z_1) + wL(z_2) + 2\pi w i c(z_1, z_2)] = \quad (7.66)$$

$$= E(wL(z_1))E(wL(z_2))E(2\pi w i c(z_1, z_2)) = \quad (7.67)$$

$$= z_1^w z_2^w E(2\pi w i c(z_1, z_2)) \quad (7.68)$$

■

1. $(-1)^{1/2} = E((1/2)L(-1)) = E((1/2)\pi i) = e^{(\pi/2)i} = \cos \pi/2 + i \operatorname{sen} \pi/2 = i$.

2.

$$1 = [(-1)(-1)]^{1/2} = (-1)^{1/2}(-1)^{1/2}E((1/2)2\pi i c(-1, -1)) = \quad (7.69)$$

$$= i^2 E(-\pi i) = -1 cis(-\pi) = (-1)(\cos(\pi) - i \sin(\pi)) = (-1)(-1) = 1. \quad (7.70)$$

3.

$$\sqrt{-4} = (-4)^{1/2} = (-1 \cdot 4)^{1/2} = (-1)^{1/2} 4^{1/2} E(2\pi \frac{1}{2} i c(-1, 4)) \quad (7.71)$$

pero $c(-1, 4) = 0, \therefore \sqrt{-4} = 2i$.

$$4. \quad L(z^w) = L(E(wL(z))) = wL(z).$$

7.10 Las funciones trigonométricas

Basados en la observación de que si $\theta \in \mathbb{R}$, entonces

$$e^{\theta i} + e^{-\theta i} = 2 \cos(\theta). \quad (7.72)$$

$$e^{i\theta} - e^{-i\theta} = 2i \sin(\theta) \quad (7.73)$$

y que, por lo tanto, $\cos(\theta) = \frac{e^{\theta i} + e^{-\theta i}}{2}$; $\sin(\theta) = \frac{e^{i\theta} - e^{-i\theta}}{2i}$.

Definimos:

$$\cos(z) = \frac{e^{zi} + e^{-zi}}{2}, \quad \sin(z) = \frac{e^{iz} - e^{-iz}}{2i} \quad (7.74)$$

de donde resulta que las funciones $\cos, \sin : \mathbb{C} \Rightarrow \mathbb{C}$, -que extienden a las correspondientes funciones de variable real -tienen entre sus propiedades básicas, las siguientes, cuya demostración se hace simplemente aplicando las definiciones.

Ejercicio 342 . *Demostrar que $\forall z, w \in \mathbb{C}$,*

1. $\sin^2(z) + \cos^2(z) = 1$.
2. $\cos(z + w) = \cos(z)\cos(w) - \sin(z)\sin(w)$,
3. $\sin(z + w) = \sin(z)\cos(w) + \cos(z)\sin(w)$.

Capítulo 8

Espacios vectoriales

El estudio del Álgebra lineal, que ha tenido un extraordinario desarrollo en los últimos años, puede considerarse como una teoría de las transformaciones lineales, funciones que para definirse requieren del conocimiento previo de los conjuntos sobre los que actúan y, por supuesto, de la estructura de éstos.

El propósito de este capítulo es dar una explicación lo más clara posible, del significado de conceptos tales como “espacio vectorial”, “vector” y “transformación lineal”. Así como desarrollar, en forma elemental, una herramienta que permita al estudiante entender algunos resultados, teoremas y aplicaciones de la teoría.

Se enfatiza el estudio de los sistemas de ecuaciones lineales y se incluye una breve introducción al estudio de los determinantes.

8.1 Conceptos preliminares

Recordemos del capítulo 2 los siguientes conceptos acerca de las operaciones.

Definición 94 . *Una operación en un conjunto X es una función*

$$f : X \times X \longrightarrow X.$$

A veces, en lugar de una letra, usaremos signos como $+, *, -, \div$, para denotar una operación, y también escribiremos $a * b$ en lugar de $*(a, b)$.

Definición 95 . *Una operación $X \times X \xrightarrow{*} X$ es asociativa si*

$$a * (b * c) = (a * b) * c, \quad \forall a, b, c \in X.$$

Ejemplos 147 . *Algunas operaciones asociativas que se definen en el conjunto $\{0, 1\}$*

$\begin{array}{ c c c } \hline & 0 & 1 \\ \hline 0 & 0 & 0 \\ \hline 1 & 0 & 0 \\ \hline \end{array}$	$\begin{array}{ c c c } \hline & 0 & 1 \\ \hline 0 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array}$	$\begin{array}{ c c c } \hline & 0 & 1 \\ \hline 0 & 0 & 1 \\ \hline 1 & 0 & 1 \\ \hline \end{array}$	$\begin{array}{ c c c } \hline & 0 & 1 \\ \hline 0 & 0 & 0 \\ \hline 1 & 1 & 1 \\ \hline \end{array}$
$\begin{array}{ c c c } \hline + & 0 & 1 \\ \hline 0 & 0 & 1 \\ \hline 1 & 0 & 0 \\ \hline \end{array}$	$\begin{array}{ c c c } \hline \wedge & 0 & 1 \\ \hline 0 & 0 & 0 \\ \hline 1 & 0 & 1 \\ \hline \end{array}$	$\begin{array}{ c c c } \hline \vee & 0 & 1 \\ \hline 0 & 0 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array}$	

Ejercicio 343 . *Demuestre que las siete operaciones definidas arriba son asociativas.*

Ejercicio 344 . *Proporcione otro ejemplo de una operación asociativa que se puede definir en el conjunto $\{0, 1\}$.*

Ejercicio 345 . *Demuestre que las demás operaciones que se pueden definir en $\{0, 1\}$ no son asociativas.*

Definición 96 . *Una operación $X \times X \xrightarrow{*} X$ es conmutativa si*

$$a * b = b * a \quad \forall a, b \in X.$$

Definición 97 . *Sea $X \times X \xrightarrow{*} X$ una operación en el conjunto X . Un elemento $e \in X$ es:*

1. neutro derecho para $*$, si $a * e = a \quad \forall a \in X$.
2. neutro izquierdo para $*$, si $e * a = a \quad \forall a \in X$.
3. neutro, si es neutro izquierdo y neutro derecho.

Definición 98 . *Un semigrupo es una pareja ordenada $(X, *)$ formada por un conjunto y una operación asociativa definida en X .*

Definición 99 . *Un monoide es una terna ordenada $(X, *, e)$ tal que $(X, *)$ es un semigrupo y $e \in X$ es un neutro para la operación $*$.*

Definición 100 . Si $(X, *, e)$ es un monoide y $a \in X, b \in X$ son tales que $a * b = e$, diremos que b es un inverso derecho de a y que a es un inverso izquierdo de b .

Si además pasara que $b * a = e$, entonces b es un inverso de a , y a es inverso de b .

Observación 90 . En un monoide, un neutro izquierdo y un neutro derecho coinciden.

Demostración. Si e es un neutro izquierdo y f es un neutro derecho entonces

$$e = e * f = f.$$

La primera igualdad vale porque f es neutro derecho, la segunda se da porque e es neutro izquierdo. ■

Observación 91 . El neutro en un monoide es único.

Demostración. Un neutro es neutro izquierdo y derecho. ■

Observación 92 . En un monoide un inverso izquierdo de un elemento coincide con un inverso derecho del mismo elemento (si ambos existen).

Demostración. Supongamos que $a * b = e = b * c$. Entonces

$$c = e * c = (a * b) * c = a * (b * c) = a * e = a.$$

■

Notación 13 . En vista de la observación anterior, podemos denotar con a^{-1} al inverso de a (cuando existe).

Observación 93 . Si un elemento en un monoide tiene inverso, entonces este inverso es único.

Demostración. Se sigue de la observación anterior notando que un inverso es un inverso izquierdo y derecho. ■

Ejemplo 148 . Note que $\left(X^X = \left\{X \xrightarrow{f} X \mid f \text{ es función}\right\}, \circ, \text{Id}_X\right)$ es un monoide. La composición de funciones es asociativa y $\text{Id}_X : X \longrightarrow X$, es neutro. Note que los elementos de este monoide con inverso izquierdo son las funciones inyectivas, y que las funciones suprayectivas son los elementos del monoide que tienen inverso derecho.

Ejemplo 149 . En el monoide $\left(\mathbb{N}^{\mathbb{N}} = \left\{ \mathbb{N} \xrightarrow{f} \mathbb{N} \mid f \text{ es función} \right\}, \circ, Id_{\mathbb{N}} \right)$, la función multiplicar por dos: $\mathbb{N} \xrightarrow{2 \cdot} \mathbb{N}$ tiene inverso izquierdo pero no tiene inverso derecho (es decir, es una función inyectiva que no es suprayectiva).

Ejercicio 346 . En el ejemplo anterior, encuentre un elemento que tenga inverso derecho pero que no tenga inverso izquierdo.

Ejercicio 347 . Muestre que si X es un conjunto finito entonces son equivalentes para $f \in X^X$:

1. f tiene inverso.
2. f tiene inverso derecho.
3. f tiene inverso izquierdo.

Definición 101 . Un grupo $(X, *, e)$ es un monoide en el que cada elemento tiene inverso.

Definición 102 . Un grupo es abeliano o commutativo si su operación es commutativa.

1. Los enteros con la suma y con 0.
2. El conjunto de los subconjuntos de un conjunto X , con la diferencia simétrica (el neutro es el conjunto vacío)
3. Los racionales distintos de 0 con el producto (el neutro es el 1).

Observación 94 . En un grupo $(G, *, e)$ las siguientes afirmaciones son equivalentes:

1. $b = a^{-1}$.
2. $a * b = e$.
3. $b * a = e$.

Demostración. Es claro que 1) \Rightarrow 2) y que 1) \Rightarrow 3), ya que un inverso es un inverso derecho e izquierdo.

2) \Rightarrow 1) El inverso de a es un inverso izquierdo de a . La condición 2) dice que b es inverso derecho de a . Como un inverso izquierdo de a coincide con un inverso derecho de a , concluimos que $b = a^{-1}$.

3) \Rightarrow 1) Se sigue de manera análoga que en el argumento anterior. ■

Observación 95 . En un grupo $(G, *, e)$:

$$1. (a * b)^{-1} = b^{-1} * a^{-1}.$$

$$2. (a^{-1})^{-1} = a.$$

Demostración. 1) Como

$$\begin{aligned} (a * b) * (b^{-1} * a^{-1}) &= ((a * b) * b^{-1}) * a^{-1} = (a * (b * b^{-1})) * a^{-1} = \\ &= (a * e) * a^{-1} = a * a^{-1} = e. \end{aligned}$$

Tenemos que $(b^{-1} * a^{-1})$ es un inverso derecho de $(a * b)$. Por la observación anterior, un inverso derecho es lo mismo que un inverso, en un grupo.

2) Como $a * a^{-1} = e$, tenemos que a es un inverso izquierdo de a^{-1} . Por la observación anterior tenemos que a es el inverso de a^{-1} . ■

Ejemplo 150 . Si X es un conjunto, entonces

$$Biye(X) = \{f \in X^X \mid f \text{ es biyectiva}\}$$

es un grupo. Esto se debe a que el inverso de una función biyectiva de X a X es también una función biyectiva de X a X (ya sabemos que una composición de funciones biyectivas es biyectiva, que la composición de funciones es asociativa y que Id_X es neutro para la composición de funciones de X a X).

Ejemplo 151 . Consideremos $Biye(\{1, 2, 3\})$ el conjunto de las permutaciones de $\{1, 2, 3\}$. Sus seis elementos son:

$$\begin{array}{ccc} \{1, 2, 3\} & \xrightarrow{R} & \{1, 2, 3\} \\ Id, & \begin{array}{c} 1 \mapsto 1 \\ 2 \mapsto 2 \\ 3 \mapsto 3 \end{array} & , \quad \begin{array}{c} 1 \mapsto 2 \\ 2 \mapsto 3 \\ 3 \mapsto 1 \end{array} \\ & & \begin{array}{c} 1 \mapsto 3 \\ 2 \mapsto 1 \\ 3 \mapsto 2 \end{array} \end{array}$$

$$\xrightarrow{R^{-1}} \{1, 2, 3\}$$

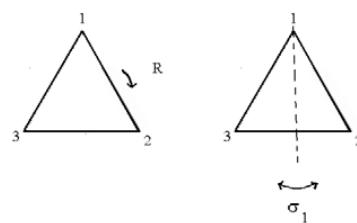


Figura 8.1:

$$\begin{array}{ccc}
 \{1, 2, 3\} & \xrightarrow{\sigma_1} & \{1, 2, 3\} \\
 1 \mapsto & 1 & \\
 2 \mapsto & 3 & , \\
 3 \mapsto & 2 & \\
 & & \{1, 2, 3\} \xrightarrow{\sigma_2} \{1, 2, 3\} \\
 & & 1 \mapsto 3 \\
 & & 2 \mapsto 2 \\
 & & 3 \mapsto 1 \\
 & & \{1, 2, 3\} \xrightarrow{\sigma_3} \{1, 2, 3\} \\
 & & 1 \mapsto 2 \\
 & & 2 \mapsto 1 \\
 & & 3 \mapsto 3
 \end{array}$$

Notemos que $R \circ \sigma_1 : 1 \xrightarrow{\sigma_1} 1 \xrightarrow{R} 2$, mientras que $\sigma_1 \circ R : 1 \xrightarrow{R} 2 \xrightarrow{\sigma_1} 3$, por lo que $R \circ \sigma_1 \neq \sigma_1 \circ R$. Esto muestra que el grupo $Biye(\{1, 2, 3\})$ no es conmutativo.

Ejercicio 348 . *Haga la tabla de la composición en $Biye(\{1, 2, 3\})$.*

Observación 96 . *Si $(G, *, e)$ es un grupo entonces:*

1. Las funciones

$$\begin{array}{l}
 G \xrightarrow{e*} G \text{ y} \\
 G \xrightarrow{*_e} G
 \end{array}$$

(multiplicar por e por la izquierda y multiplicar por e por la derecha, respectivamente, son ambas la función identidad).

2. Dada $g \in G$, la función

$$G \xrightarrow{g*} G$$

es biyectiva, pues su inverso es

$$G \xrightarrow{g^{-1}*} G.$$

3. Dada $g \in G$, la función

$$G \xrightarrow{\circ g} G$$

es biyectiva, pues su inverso es

$$G \xrightarrow{-*g^{-1}} G.$$

4. Recuérdese la definición de tabla de multiplicar. Note que si G es un grupo finito, en cada renglón de su tabla de multiplicar, todos los elementos son distintos (esto es porque $g * \underline{}$ es una función inyectiva). Note también que todos los elementos de G aparecen en cada renglón (esto se debe a que $g * \underline{}$ es una función suprayectiva).

5. Observe que si G es un grupo finito, cada elemento de G aparece una sola vez en cada columna de la tabla de multiplicar de G .

Ejercicio 349 . *Construya la tabla de multiplicar para un grupo con tres elementos.*

Ejemplo 152 . *Si G es un grupo abeliano, entonces G^X también es un grupo abeliano.*

Demostración. (Bosquejo) $G^X = \{f : X \longrightarrow G \mid f \text{ es una función}\}$. Llamémosle suma a las operaciones. La suma en G^X se define de la manera usual:

$$\begin{aligned} \text{Si } f, g &\in G^X, \text{ entonces} \\ (f + g)(x) &= : f(x) + g(x). \end{aligned}$$

Esta operación es commutativa, y asociativa. Si e es el neutro de G , entonces la función constante

$$\begin{aligned} \widehat{e} : X &\longrightarrow G \\ x &\longmapsto e, \end{aligned}$$

es el neutro para la operación en G^X .

El inverso de f es la función

$$\begin{aligned} X &\longrightarrow G \\ x &\longmapsto f(x)^{-1}. \end{aligned}$$

■

Ejercicio 350 . *Demostrar las afirmaciones hechas en el bosquejo de demostración de arriba.*

Ejemplo 153 . *Como los reales \mathbb{R} con su suma y el 0 son un grupo, entonces $\mathbb{R}^{\mathbb{R}}$, el conjunto de las funciones reales de variable real, junto con la suma usual de funciones y la función constante $\widehat{0}$, forman un grupo.*

1. $(\mathbb{N}, +, 0)$ no es un grupo abeliano (¿por qué?).
2. $(\mathbb{Z}, \cdot, 1)$ no es un grupo abeliano (¿por qué?).
3. $(\mathbb{R}, \cdot, 1)$ no es un grupo abeliano (¿por qué?).

1. Los racionales sin el 0, con el producto, con el uno.
2. Los reales sin el 0, con el producto, con el 1.

Definición 103 . *Un campo es una quinteta ordenada $(F, +, 0, \cdot, 1)$ tal que*

1. $(F, +, 0)$ Es un grupo abeliano.
2. $(F \setminus \{0\}, \cdot, 1)$ es un grupo abeliano.
3. $\forall a, b, c \in F, \quad a \cdot (b + c) = a \cdot b + a \cdot c.$

Observación 97 . *Aunque el conjunto en un semigrupo puede ser vacío, el conjunto en un monoide es distinto del vacío. ¿Por qué?*

Observación 98 . $0 \neq 1$ en un campo. (¿Por qué?, observe la definición).

Observación 99 . *En vista de la observación anterior, en un campo debe haber por lo menos dos elementos. Por otra parte, hay un campo con exactamente dos elementos:*

$$(\{0, 1\}, +, 0, \cdot, 1),$$

donde las operaciones están dadas por las tablas

$+$	0	1
0	0	1
1	1	0

 \cdot

	0	1
0	0	0
1	0	1

.

Para entender la tabla de la multiplicación, hagamos la siguiente observación.

1. $a \cdot 0 = 0, \forall a \in F, F$ un campo.
2. $(b + c) \cdot a = b \cdot a + c \cdot a.$
3. $-1 \cdot a = -a.$
4. $-(-a) = a.$
5. $(-a) \cdot b = -(a \cdot b).$
6. $(-a) \cdot (-b) = ab.$

Demostración. 1) $a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$. Sumando de cada lado el inverso aditivo de $a \cdot 0$, obtenemos

$$0 = a \cdot 0.$$

$$2) (b + c) \cdot a = a \cdot (b + c) = a \cdot b + a \cdot c = b \cdot a + c \cdot a.$$

3) $a + (-1) \cdot a = 1 \cdot a + (-1) \cdot a = (1 + (-1)) \cdot a = 0 \cdot a = 0$. Así que $(-1) \cdot a = -a$.

4) Es claro que de $a + x = 0$ se sigue que $x = -a$ (sume $-a$ de cada lado). Como $-a + a = 0$, entonces $a = -(-a)$.

5) $a \cdot b + (-a) \cdot b = (a + (-a)) \cdot b = 0 \cdot b = 0$. Por lo tanto $(-a) \cdot b$ es el inverso (aditivo) de $a \cdot b$.

$$6) (-a) \cdot (-b) = - (a \cdot (-b)) = - ((-b) \cdot a) = - (- (b \cdot a)) = b \cdot a = a \cdot b.$$

■

Definición 104 (*La restricción de una operación*). *Sea $* : X \times X \rightarrow X$ una operación en un conjunto X . Supongamos que $A \subseteq X$, entonces $A \times A \subseteq X \times X$, así que podemos considerar la situación*

$$\begin{array}{ccc} X \times X & \xrightarrow{*} & X \\ \uparrow & \nearrow & \\ A \times A & & \end{array}.$$

De tal forma que tenemos una función $*_{|A \times A} : A \times A \rightarrow X$. (dos elementos de A se multiplican como elementos de X que son). Desde luego ésta no es una operación en A , porque el codominio no es necesariamente A .

Si se puede escoger A como el codominio de $*_{|A \times A}$, diremos que A es cerrado bajo la operación $*$. (Debe ser claro que esto pasa si y sólo si $a, b \in A \Rightarrow a * b \in A$).

Observación 100 . Si $*$ es una operación asociativa en X (o conmutativa) y A es un subconjunto de X cerrado bajo la operación, entonces la restricción $*_{|A \times A}$ también es asociativa (conmutativa).

8.2 Espacios vectoriales

Definición 105 . Un espacio vectorial es una quinteta ordenada

$$(V, +, \vec{0}, F, \cdot : F \times V \rightarrow V)$$

tal que

1. $(V, +, \vec{0})$ es un grupo abeliano.
2. F es un campo.
3. La función $\cdot : F \times V \longrightarrow V$ tiene las siguientes cuatro propiedades:
 - (a) $1\vec{v} = \vec{v}$, $\forall \vec{v} \in V$.
 - (b) $(ab)\vec{v} = a(b\vec{v})$, $\forall a, b \in F, \forall \vec{v} \in V$.
 - (c) $(a+b)\vec{v} = a\vec{v} + b\vec{v}$, $\forall a, b \in F, \forall \vec{v} \in V$.
 - (d) $a(\vec{v} + \vec{w}) = a\vec{v} + a\vec{w}$, $\forall a \in F, \forall \vec{v}, \vec{w} \in V$.

Observación 101 . Si

$$(V, +, \vec{0}, F, \cdot : F \times V \longrightarrow V)$$

es un espacio vectorial, entonces

1. $0\vec{v} = \vec{0}$.
2. $c\vec{0} = \vec{0}$.
3. $(-1)\vec{v} = -\vec{v}$.
4. $c\vec{v} = \vec{0} \implies ((c=0) \vee (\vec{v} = \vec{0}))$.

Demostración. 1) $0\vec{v} = (0+0)\vec{v} = 0\vec{v} + 0\vec{v}$, sumando $-0\vec{v}$:
 $\vec{0} = 0\vec{v}$.
 2) $c\vec{0} = c(\vec{0} + \vec{0}) = c\vec{0} + c\vec{0}$, sumando $-c\vec{0}$: $\vec{0} = c\vec{0}$.
 3) $\vec{v} + (-1)\vec{v} = 1\vec{v} + (-1)\vec{v} = (1 + (-1))\vec{v} = 0\vec{v} = \vec{0} \therefore (-1)\vec{v} = -\vec{v}$.
 4) Si $c \neq 0$, $c\vec{v} = \vec{0}$, tenemos

$$\vec{v} = 1\vec{v} = \frac{1}{c}c\vec{v} = \frac{1}{c}(c\vec{v}) = \frac{1}{c}\vec{0} = \vec{0}.$$

■
Notación 14 . Si $(V, +, \vec{0}, F, \cdot : F \times V \longrightarrow V)$ es un espacio vectorial, llamaremos vectores a los elementos de V y escalares a los elementos de F .

Estrictamente, un espacio vectorial sobre un campo, es una quinteta ordenada (véase la definición). Sin embargo es mucho más cómodo usar la notación ${}_F V$, en donde V denota el grupo abeliano y F denota el campo. Los elementos de V se llamarán vectores, y los elementos de F se llamarán escalares.

Cuando F es el campo \mathbb{R} de los números reales, se dice que V es un espacio vectorial real, (si F es \mathbb{C} , V es un espacio vectorial complejo, etcétera).

Ejemplos 154

- Si $G = \{0\}$ con $0 + 0 = 0$, y $\forall a \in F$, $a0 = 0$, entonces G es un espacio vectorial sobre cualquier campo F , y se llama “el espacio vectorial cero” o “trivial”.
- Todo campo F es un espacio vectorial sobre cualquiera de sus subcampos. (\mathbb{C} es un espacio vectorial real, o un espacio vectorial sobre \mathbb{Q}, \dots).¹
- Denotamos por \mathbb{R}^2 el conjunto de parejas ordenadas de números reales:

$$\mathbb{R}^2 = \{(a, b) \mid a, b \in \mathbb{R}\}.$$

Por otra parte, $2 = \{0, 1\}$, así que

$$\mathbb{R}^{\{0,1\}} = \left\{ \{0, 1\} \xrightarrow{f} \mathbb{R} \mid f \text{ es función} \right\}.$$

Notemos que \mathbb{R}^2 y $\mathbb{R}^{\{0,1\}}$ se pueden identificar mediante la función biyectiva que envía la pareja (a, b) a la función

$\{0, 1\}$	\longrightarrow	\mathbb{R}
0	\mapsto	a
1	\mapsto	b

Como sabemos que \mathbb{R} es un grupo abeliano con la suma de funciones acostumbrada y con la función constante $\widehat{0}$.

Es fácil ver que la suma en \mathbb{R}^2 debe estar dada por la siguiente regla (inducida por la manera en que se suman dos funciones)

$$(a, b) + (c, d) = (a + c, b + d).$$

¹Si K, F son campos y $K \subseteq F$, se dice que K es un subcampo de F cuando las operaciones de K son las restricciones de las operaciones de F , y los neutros de las operaciones de F pertenecen a K .

De esta manera, se suman y se multiplican dos elementos de K tal como se suman y multiplican en F .

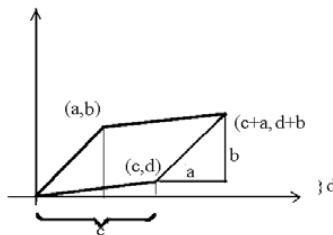


Figura 8.2:

Por otra parte, una manera de definir el producto de un real por una función es la siguiente:

$$(r \cdot f)(x) = \underbrace{r(f(x))}_{\substack{\text{producto de dos} \\ \text{números reales}}}.$$

Lo anterior sugiere que el producto de un real por una pareja ordenada de reales se debe definir de la manera siguiente:

$$r \cdot (a, b) = (ra, rb).$$

Ejercicio 351 . *Demuestre que \mathbb{R}^2 es un espacio vectorial sobre el campo de los números reales, con las operaciones como las definimos arriba.*

Ejemplo 155 . Sea F un campo, y $F^n = F \times F \times \dots \times F$ (n factores). Entonces F^n resulta un espacio vectorial sobre F , si se define $\forall a \in F, \forall v = (v_1, \dots, v_n), w = (w_1, \dots, w_n) \in F^n$,

$$v + w = (v_1 + w_1, \dots, v_n + w_n)$$

$$av = (av_1, \dots, av_n)$$

2

Ejemplo 156 . Si $V = F[x]$ denota el anillo de los polinomios en una indeterminada x con coeficientes en un campo F , con la suma y la multiplicación por un escalar definidas en la forma usual, V es un espacio vectorial sobre el campo F . (Ver la definición de polinomio en la página 540).

Ejemplo 157 . Sea W un espacio vectorial sobre un campo F y sea S un conjunto no vacío. Se define el conjunto de las funciones de S en W (W^S), y se le da estructura de grupo definiendo

$$(f + g)(s) = f(s) + g(s), \quad f, g \in W^S, \quad s \in S,$$

Si además se define

$$(af)(s) = a(f(s)), \quad a \in F, f \in W^S, \quad s \in S.$$

Entonces W^S resulta un espacio vectorial sobre F .

Ejercicio 352 . Demuestre que \mathbb{R}^n es un espacio vectorial sobre el campo de los números reales si se definen las operaciones de la manera siguiente:

$$(a_1, a_2, \dots, a_n) + (b_1, b_2, \dots, b_n) = (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n)$$

$$r \cdot (a_1, a_2, \dots, a_n) = (ra_1, ra_2, \dots, ra_n).$$

Ejercicio 353 . Demuestre que un campo F puede considerarse como espacio vectorial sobre sí mismo si se toman las operaciones como las operaciones usuales del campo.

²En los cursos de Álgebra lineal se demuestra que todo espacio vectorial de dimensión finita n (ver definición más adelante) sobre un campo F es isomorfo a F^n (indistinguible desde el punto de vista de sus propiedades algebraicas) y en ese sentido se justifica la afirmación de que "de alguna manera", los F^n son los espacios vectoriales más importantes, entre los de dimensión finita.

8.3 Subespacios

Definición 106. Sea $W \subseteq {}_F V$. Diremos que W es un subespacio de V (y escribiremos $W \leq V$) si

$$\left(W, +|_{W \times W}, \vec{0}, F, \cdot|_{F \times W} : F \times W \longrightarrow W \right)$$

es un espacio vectorial.

Notemos que para que W pueda ser un subespacio vectorial de V se necesita que la restricción de la suma a $W \times W$, sea una operación en W , es decir, se necesita que W sea cerrado bajo la suma.

Notemos también que el producto de una escalar por un elementos de W , debe ser un elemento de W .

Teorema 113. Supongamos que $W \subseteq {}_F V$. Son equivalentes:

1. W es un subespacio de V . (Escribiremos $W \leq V$).
2. (a) W es cerrado bajo la suma.
(b) El vector $\vec{0}$ pertenece a W .
(c) W es cerrado bajo la multiplicación por escalares.

Demostración. 1) \Rightarrow 2) Ya hemos notado que si W es un subespacio de V , W es cerrado bajo la suma y bajo la multiplicación por escalares. Por otra parte, como $(W, +, \vec{0})$ es un grupo, entonces $\vec{0} \in W$.

2) \Rightarrow 1) Por hipótesis, $(W, +, \vec{0})$ es un monoide (véase la observación 100). Falta demostrar que el inverso de cada elemento de W está también en W . Sea $\vec{w} \in W$, bastará notar que por la observación 101

$$(-1) \cdot \vec{w} = -\vec{w}.$$

Por lo tanto, $(W, +, \vec{0})$ es un grupo abeliano.

Por otra parte, tenemos lo siguiente:

$$1\vec{v} = \vec{v} \quad \forall \vec{v} \in V \therefore 1\vec{w} = \vec{w} \quad \forall \vec{w} \in W.$$

$$(ab)\vec{v} = a(b\vec{v}) \quad \forall a, b \in F, \quad \forall \vec{v} \in V \therefore ab\vec{w} = a(b\vec{w}) \quad \forall a, b \in F, \quad \forall \vec{w} \in W.$$

$(a + b)\vec{v} = a\vec{v} + b\vec{v} \quad \forall a, b \in F, \quad \forall \vec{v} \in V \therefore (a + b)\vec{w} = a\vec{w} + b\vec{w} \quad \forall a, b \in F, \quad \forall \vec{w} \in W.$

$a(\vec{v}_1 + \vec{v}_2) = a\vec{v}_1 + a\vec{v}_2 \quad \forall a \in F, \quad \forall \vec{v}_1, \vec{v}_2 \in V \therefore (\vec{w}_1 + \vec{w}_2) = a\vec{w}_1 + a\vec{w}_2 \quad \forall a \in F, \quad \forall \vec{w}_1, \vec{w}_2 \in W. \blacksquare$

Ejemplo 158 . $\left\{ \vec{0} \right\} \leq V :$

Es claro que $\left\{ \vec{0} \right\}$ contiene a $\vec{0}$, y que es cerrado bajo la suma. Demostrar que es cerrado bajo multiplicación por escalares, es lo mismo que demostrar que

$$c \cdot \vec{0} = \vec{0} \quad \forall c \in F.$$

En efecto,

$$c \cdot \vec{0} = c \cdot \left(\vec{0} + \vec{0} \right) = c \cdot \vec{0} + c \cdot \vec{0}$$

sumando $- \left(c \cdot \vec{0} \right)$, obtenemos

$$\vec{0} = c \cdot \vec{0}.$$

Ejemplo 159 . ${}_F V \leq_F V$.

Lema 22 . Si $\vec{u} \in V$, entonces el conjunto de todos los múltiplos de \vec{u} : $\{c\vec{u} \mid c \in F\} =: F\vec{u}$ es un subespacio de V :

Demostración. a) $\vec{0} = 0\vec{u} \in F\vec{u}$.

b) $c\vec{u} + d\vec{u} = (c + d)\vec{u}$.

c) $k(d\vec{u}) = (kd)\vec{u}$. \blacksquare

Lema 23 . Si W_1, W_2 son subespacios de V , entonces

$$\{\vec{w}_1 + \vec{w}_2 \mid \vec{w}_1 \in W_1, \vec{w}_2 \in W_2\}$$

es un subespacio de ${}_F V$:

Demostración. a) $\vec{0} = \vec{0} + \vec{0}, \quad \vec{0} \in W_1, \vec{0} \in W_2$.

b) Si $\vec{w}_1, \vec{u}_1 \in W_1, \vec{w}_2, \vec{u}_2 \in W_2$, entonces

$$(\vec{w}_1 + \vec{w}_2) + (\vec{u}_1 + \vec{u}_2) = (\vec{w}_1 + \vec{u}_1) + (\vec{w}_2 + \vec{u}_2)$$

con $\vec{w}_1 + \vec{u}_1 \in W_1$ y $\vec{w}_2 + \vec{u}_2 \in W_2$.

c) Si $c \in F, \vec{w}_1 \in W_1, \vec{w}_2 \in W_2$, entonces $c(\vec{w}_1 +) = c\vec{w}_1 + c\vec{w}_2$ con $c\vec{w}_1 \in W_1$ y $c\vec{w}_2 \in W_2$. \blacksquare

Ejemplos 160

1. El conjunto de los números reales es un subespacio del conjunto de los polinomios con coeficientes reales;
2. El conjunto de los polinomios con coeficientes reales es un subespacio del conjunto de las funciones reales de variable real que tienen todas las derivadas;
3. El conjunto de las funciones reales de variable real que tienen todas las derivadas es un subespacio del conjunto de las funciones derivables;
4. El conjunto de las funciones reales derivables es un subespacio del conjunto de las funciones continuas;
5. El conjunto de las funciones reales de variable real continuas es un subespacio de las funciones reales de variable real.
6. $\mathbb{R}^{\mathbb{R}}$ es un espacio vectorial, como ya hemos visto. (Es de la forma F^X).

Teorema 114 . *La intersección de una familia de subespacios vectoriales de FV , es un subespacio.*

Demostración. Supongamos que $\{W_i\}_{i \in I}$ es una familia de subespacios de FV . Para demostrar que $\cap \{W_i\}_{i \in I}$ es un subespacio de FV , necesitamos verificar que contiene al vector $\vec{0}$, que es cerrado bajo la suma, y que es cerrado bajo la multiplicación por escalares. En efecto,

$\vec{0} \in W_i \quad \forall i \in I$, pues cada W_i es un subespacio. Por lo tanto $\vec{0} \in \cap \{W_i\}_{i \in I}$.

Si $\vec{u}, \vec{v} \in \cap \{W_i\}_{i \in I}$, entonces $\vec{u}, \vec{v} \in W_i, \forall i \in I$, como cada W_i es cerrado bajo la suma, entonces $\vec{u} + \vec{v} \in W_i, \forall i \in I$. Por lo tanto $\vec{u} + \vec{v} \in \cap \{W_i\}_{i \in I}$.

Si $c \in F$ y $\vec{u} \in \cap \{W_i\}_{i \in I}$, entonces $c \in F$ y $\vec{u} \in W_i, \forall i \in I$, por lo tanto $c\vec{u} \in W_i, \forall i \in I$. Esto quiere decir que

$$c\vec{u} \in \cap \{W_i\}_{i \in I}.$$

■

Teorema 115 . *Si $X \subseteq FV$ entonces $\cap \{W \leq V \mid X \subseteq W\}$ es el menor subespacio de FV , que contiene a X .*

Demostración. Deben ser claras dos cosas: la primera es que

$$\cap \{W \leq V \mid X \subseteq W\}$$

es un subespacio de V (Teorema anterior).

Por otra parte, $X \subseteq \cap \{W \leq V \mid X \subseteq W\}$, claramente.

Ahora, si $Z \leq V$ y $X \subseteq Z$, entonces Z es uno de los intersectandos, por lo que

$$\cap \{W \leq V \mid X \subseteq W\} \leq Z.$$

(Note que los elementos de $\cap \{W \leq V \mid X \subseteq W\}$ son los elementos que pertenecen a cada uno de los intersectandos). ■

Notación 15 . Denotaremos por $\langle X \rangle$ o por $\mathcal{S}(X)$ al menor subespacio de FV que contiene a X . Este subespacio se llama el subespacio (de FV) generado por X .

Ejemplo 161 . El subespacio generado por \emptyset .

Por lo visto anteriormente,

$$\mathcal{S}(\emptyset) = \cap \{W \mid \emptyset \subseteq W \leq V\} = \cap \{W \mid W \leq V\} \subseteq \left\{ \overrightarrow{0} \right\},$$

puesto que $\left\{ \overrightarrow{0} \right\}$ es uno de los subespacios de V .

Por otra parte, y en vista de que $\mathcal{S}(\emptyset)$ es un subespacio, entonces $\overrightarrow{0} \in \mathcal{S}(\emptyset)$, que es lo mismo que decir que $\left\{ \overrightarrow{0} \right\} \subseteq \mathcal{S}(\emptyset)$.

Por lo tanto,

$$\mathcal{S}(\emptyset) = \left\{ \overrightarrow{0} \right\}.$$

Ejercicio 354 . Demuestre que $\mathcal{S}\left(\left\{ \overrightarrow{0} \right\}\right) = \left\{ \overrightarrow{0} \right\}$.

Definición 107 . Si W_1, W_2 son subespacios de V , definimos $W_1 + W_2$ como $\mathcal{S}(W_1 \cup W_2)$.

Observación 102 . $X \subseteq Y \implies \mathcal{S}(X) \leq \mathcal{S}(Y)$.

Demostración. Por definición, $Y \subseteq \mathcal{S}(Y)$. Por lo tanto

$$X \subseteq Y \subseteq \mathcal{S}(Y).$$

Entonces $\mathcal{S}(Y)$ es un subespacio de V que contiene a X , por lo tanto debe contener también al subespacio generado por X :

$$\mathcal{S}(X) \leq \mathcal{S}(Y).$$

■

Teorema 116

1. Si W_1, W_2 son subespacios de V , entonces

$$W_1 + W_2 = \{ \vec{w}_1 + \vec{w}_2 \mid \vec{w}_1 \in W_1, \vec{w}_2 \in W_2 \}.$$

2. Para X, Y subconjuntos de V :

$$\mathcal{S}(X \cup Y) = \mathcal{S}(X) + \mathcal{S}(Y).$$

Demostración. 1) Hemos visto en un ejemplo anterior que

$$\{ \vec{w}_1 + \vec{w}_2 \mid \vec{w}_1 \in W_1, \vec{w}_2 \in W_2 \}$$

es un subespacio de V . Debe ser claro que contiene tanto a W_1 como a W_2 . (Por ejemplo, $\vec{w}_1 = \vec{w}_1 + \vec{0}$). Entonces $\{ \vec{w}_1 + \vec{w}_2 \mid \vec{w}_1 \in W_1, \vec{w}_2 \in W_2 \}$ es un subespacio de V que contiene a $W_1 \cup W_2$. Por lo tanto debe contener también al subespacio generado por $W_1 \cup W_2$:

$$\mathcal{S}(W_1 \cup W_2) \subseteq \{ \vec{w}_1 + \vec{w}_2 \mid \vec{w}_1 \in W_1, \vec{w}_2 \in W_2 \}.$$

Recíprocamente, como $\mathcal{S}(W_1 \cup W_2)$ es un subespacio que contiene tanto a W_1 como a W_2 , entonces debe contener la suma de cada elemento de W_1 con cada elemento de W_2 .

- 2) Como $X \subseteq X \cup Y$, entonces $\mathcal{S}(X) \leq \mathcal{S}(X \cup Y)$. Análogamente, $\mathcal{S}(Y) \leq \mathcal{S}(X \cup Y)$. Por lo tanto,

$$\mathcal{S}(X) \cup \mathcal{S}(Y) \leq \mathcal{S}(X \cup Y)$$

Por lo tanto

$$\mathcal{S}(\mathcal{S}(X) \cup \mathcal{S}(Y)) \leq \mathcal{S}(X \cup Y).$$

Pero por definición de suma de subespacios, $\mathcal{S}(\mathcal{S}(X) \cup \mathcal{S}(Y)) = \mathcal{S}(X) + \mathcal{S}(Y)$, luego entonces

$$\mathcal{S}(X) + \mathcal{S}(Y) \leq \mathcal{S}(X \cup Y).$$

Recíprocamente, $X \cup Y \subseteq \mathcal{S}(X) \cup \mathcal{S}(Y) \subseteq \mathcal{S}(\mathcal{S}(X) \cup \mathcal{S}(Y)) = \mathcal{S}(X) + \mathcal{S}(Y)$.

Por lo tanto

$$\mathcal{S}(X \cup Y) \leq \mathcal{S}(X) + \mathcal{S}(Y).$$

■

Teorema 117 . *Sea $X \subseteq {}_F V$, entonces*

$$\mathcal{S}(X) = \{a_1 \vec{x}_1 + a_2 \vec{x}_2 + \dots + a_n \vec{x}_n \mid n \in \mathbb{N}, a_i \in F, \vec{x}_i \in X\}.$$

Demostración. Denotemos por $\mathcal{C}(X)$ el conjunto de la derecha. Necesitamos demostrar que $\mathcal{C}(X)$ es el menor subespacio de V que contiene a X .

Notemos que si $X = \emptyset$, entonces $\mathcal{S}(X) = \{\vec{0}\}$. Por otra parte, $\mathcal{C}(X) = \{\vec{0}\}$, pues siguiendo la convención usual, una suma vacía (es decir, una suma sin sumandos) debe producir el neutro, dado que la suma es asociativa.

Si $X \neq \emptyset$, veamos que $\mathcal{C}(X)$ es un subespacio de V .

- $\vec{0} = 0 \cdot \vec{x} \in \mathcal{C}(X)$, si $\vec{x} \in X$.
- $(a_1 \vec{x}_1 + a_2 \vec{x}_2 + \dots + a_n \vec{x}_n) + (b_1 \vec{y}_1 + b_2 \vec{y}_2 + \dots + b_m \vec{y}_m)$ sigue siendo de la forma de los elementos de $\mathcal{C}(X)$.
- Si $c \in F$, entonces

$$c(a_1 \vec{x}_1 + a_2 \vec{x}_2 + \dots + a_n \vec{x}_n) = (ca_1) \vec{x}_1 + (ca_2) \vec{x}_2 + \dots + (ca_n) \vec{x}_n.$$

Con esto hemos demostrado que $\mathcal{C}(X) \leq V$.

Además, $X \subseteq \mathcal{C}(X)$: pues $\forall \vec{x} \in X$, se tiene que $\vec{x} = 1 \cdot \vec{x}$.

Como $\mathcal{C}(X)$ es un subespacio de V que contiene a X , entonces también debe contener al subespacio generado por X :

$$\mathcal{S}(X) \subseteq \mathcal{C}(X).$$

Recíprocamente, si tomamos un elemento $a_1 \vec{x}_1 + a_2 \vec{x}_2 + \dots + a_n \vec{x}_n \in \mathcal{C}(X)$, donde cada $\vec{x}_i \in X$ y cada $a_i \in F$, entonces cada $a_i \vec{x}_i \in \mathcal{S}(X)$, ya que $\mathcal{S}(X)$ es un subespacio que contiene a X . Como $\mathcal{S}(X)$ es cerrado bajo la suma entonces

$$a_1 \vec{x}_1 + a_2 \vec{x}_2 + \dots + a_n \vec{x}_n \in \mathcal{S}(X).$$

Por lo tanto,

$$\mathcal{C}(X) \subseteq \mathcal{S}(X).$$

■

Definición 108 . Un vector de la forma $a_1 \vec{x}_1 + a_2 \vec{x}_2 + \dots + a_n \vec{x}_n \in \mathcal{S}(X)$, con $x_i \in X$, $i \in \{1, \dots, n\}$, se llama combinación lineal (de elementos) de X .

Ejemplo 162 . Como ya hemos visto, el subespacio generado por $\{\vec{u}\}$ tiene por elementos los múltiplos de \vec{u} . En particular, si $\vec{u} \in \mathbb{R}^2$ estos múltiplos se corresponden con los puntos sobre la recta que pasa por \vec{u} y por el origen: $(0, 0)$.

Ejemplo 163 . El espacio generado por dos vectores \vec{u} y \vec{v} es:

$$\begin{aligned} \mathcal{S}(\{\vec{u}, \vec{v}\}) &= \mathcal{S}(\{\vec{u}\} \cup \{\vec{v}\}) = \\ \mathcal{S}(\{\vec{u}\}) + \mathcal{S}(\{\vec{v}\}) &= \{c\vec{u} + d\vec{v} \mid c, d \in F\}. \end{aligned}$$

Ejemplo 164 . Considerando el ejemplo anterior, si tomamos dos vectores no colineales en \mathbb{R}^3 , el subespacio generado por ellos es el conjunto de puntos en el plano que pasa por el origen y por ellos dos.

8.3.1 Dependencia lineal

Definición 109

1. Se dice que $X \subseteq {}_F V$ es linealmente dependiente si $\exists \vec{x} \in X$ tal que $\vec{x} \in \mathcal{S}(X \setminus \{\vec{x}\})$.
2. Se dice que $X \subseteq {}_F V$ es linealmente independiente si no es linealmente dependiente.

Ejemplo 165 . *El conjunto $\{\vec{0}\}$ es linealmente dependiente porque*

$$\vec{0} \in \mathcal{S}(\{\vec{0}\} \setminus \{\vec{0}\}) = \mathcal{S}(\emptyset) = \{\vec{0}\}. \quad (8.1)$$

Ejemplo 166 . *Un conjunto con un único vector es linealmente independiente si y sólo si su vector es distinto del vector $\vec{0}$. Una parte está dada por el ejemplo anterior.*

Ahora, si $\vec{u} \neq \vec{0}$ entonces $\{\vec{u}\}$ es linealmente independiente pues en caso contrario,

$$\vec{u} \in \mathcal{S}(\{\vec{u}\} \setminus \{\vec{u}\}) = \mathcal{S}(\emptyset) = \{\vec{0}\},$$

contradicciendo que $\vec{u} \neq \vec{0}$.

Teorema 118 . *Si $X \subseteq Y \subseteq_F V$ y X es un conjunto linealmente dependiente, entonces también Y es un conjunto linealmente dependiente.*

Demostración. Si X linealmente dependiente es porque existe un elemento \vec{x} en X tal que es combinación lineal de los demás elementos de X :

$$\vec{x} \in \mathcal{S}(X \setminus \{\vec{x}\}),$$

como es claro que $\mathcal{S}(X \setminus \{\vec{x}\}) \subseteq \mathcal{S}(Y \setminus \{\vec{x}\})$, entonces tenemos que

$$\vec{x} \in \mathcal{S}(Y \setminus \{\vec{x}\}),$$

por lo que Y es un conjunto linealmente dependiente. ■

Observación 103 . *Si $\vec{0} \in X \subseteq V$, entonces X linealmente dependiente. Dicho de otra manera: un conjunto linealmente independiente no puede contener al vector $\vec{0}$.*

Corolario 15 . *Si $X \subseteq Y \subseteq_F V$ y Y es linealmente independiente, entonces X también es linealmente independiente.*

Demostración. Se sigue inmediatamente del teorema anterior. ■

Lema 24 . *Un conjunto finito es linealmente dependiente si y sólo si uno de sus vectores es combinación lineal de los anteriores.*

Demostración. \implies) Supongamos que $X = \{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n\}$ es un conjunto linealmente dependiente de vectores. $\exists i \in \{1, 2, \dots, n\}$ tal que

$$\vec{x}_i \in \mathcal{S}(X \setminus \{\vec{x}_i\}).$$

Entonces podemos escribir $\vec{x}_i = \sum_{j \neq i} a_j \vec{x}_j$, o bien

$$\vec{0} = \sum_{j < i} a_j \vec{x}_j - \vec{x}_i + \sum_{j > i} a_j \vec{x}_j$$

Hagamos $k = \max \{l \mid a_l \neq 0\}$, notemos que $k \geq i$ ($a_i = -1 \neq 0$).

Entonces

$$\vec{0} = \sum_{j=1}^k a_j \vec{x}_j,$$

con $a_k \neq 0$.

Podemos reescribir esto de manera siguiente:

$$\sum_{j=1}^{k-1} (-a_j) \vec{x}_j = a_k \vec{x}_k,$$

multiplicando por el recíproco de a_k :

$$\sum_{j=1}^{k-1} \left(\frac{-a_j}{a_k} \right) \vec{x}_j = \vec{x}_k,$$

de donde vemos que \vec{x}_k es una combinación lineal de los vectores anteriores.

Si $|X| = 1$ entonces estaríamos diciendo que el único elemento $\vec{x} \in X$ es combinación lineal de los elementos de

$$X \setminus \{\vec{x}\} = \{\vec{x}\} \setminus \{\vec{x}\} = \emptyset.$$

Como hemos visto, esto pasa sólo si $\vec{x} = \vec{0}$. Por lo tanto $X = \{\vec{0}\}$, que como ya vimos, es linealmente dependiente.

\iff) Supongamos que en el conjunto $X = \{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n\}$ hay un vector que es combinación lineal de los anteriores. Si \vec{x}_1 es el vector que es combinación lineal de los anteriores, es porque $\vec{x}_1 = \vec{0}$, y en ese caso el conjunto X es linealmente dependiente, como ya vimos.

Si $i > 1$, y $\vec{x}_i \in \mathcal{S}(\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_{i-1}\})$, entonces, por definición, el conjunto $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_i\}$ es linealmente dependiente. Consecuentemente el conjunto X también es linealmente dependiente. ■

Corolario 16 . *Un conjunto finito $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n\}$ es linealmente independiente si y sólo si ninguno de sus elementos es combinación lineal de los anteriores.*

Ejemplo 167 . *Consideremos el siguiente subconjunto de \mathbb{R}^3*

$$\{(1, 2, 2), (3, -1, 0), (1, 1, 1)\},$$

veamos que es linealmente independiente:

$(1, 2, 2)$ no es combinación lineal de los anteriores, porque no es el vector $\vec{0}$;

$(3, -1, 0)$ no es combinación lineal de los anteriores, porque el único múltiplo de $(1, 2, 2)$ que tiene 0 en su última coordenada es $(0, 0, 0)$. Por último, si

$$x(1, 2, 2) + y(3, -1, 0) = (1, 1, 1),$$

entonces se podría resolver el sistema de ecuaciones siguiente:

$$\begin{array}{rcl} x & + 3y & = 1 \\ 2x & - y & = 1, \\ 2x & & = 1 \end{array}$$

que tiene las mismas soluciones que el sistema

$$\begin{array}{rcl} x & + 3y & = 1 \\ -7y & = 1, \\ -6y & = 1 \end{array}$$

pero este sistema no se puede resolver porque la segunda ecuación nos dice que $y = -\frac{1}{7}$, mientras que la tercera ecuación nos dice que $y = -\frac{1}{6}$, es una contradicción.

Teorema 119 . *Un conjunto $X \subseteq {}_F V$ es linealmente dependiente si y sólo si X contiene un subconjunto finito F que es linealmente dependiente.*

Demostración. \Leftarrow) Si F es linealmente dependiente y $F \subseteq X$, entonces X es linealmente dependiente.

\Rightarrow) Si X es linealmente dependiente, entonces existe un elemento \vec{x} que es combinación lineal de los demás elementos de X . Si $\vec{x} = \vec{0}$, entonces $\{\vec{0}\}$ es un subconjunto finito de X , que es linealmente dependiente.

Si $\vec{x} \neq \vec{0}$, entonces podemos escribir $\vec{x} = \sum_{i=1}^n a_i \vec{x}_i$, con $\vec{x}_i \in X \setminus \{\vec{x}\}$ y $a_i \neq 0$. Pero entonces debe ser claro que el conjunto

$$\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n, \vec{x}\}$$

es linealmente dependiente y es finito. ■

Corolario 17 . *Un conjunto $X \subseteq {}_F V$ es linealmente independiente si y sólo si todos sus subconjuntos finitos son linealmente independientes.*

Teorema 120 . *Un conjunto X es linealmente independiente si y sólo si*

$$\left(\sum_{i=1}^n a_i \vec{x}_i = \vec{0}, \text{ con } \{\vec{x}_i\}_{i=1}^n \subseteq X \right) \implies a_i = 0, \forall i \in \{1, \dots, n\}.$$

Demostración. \implies) Si $\sum_{i=1}^n a_i \vec{x}_i = \vec{0}$, pero $a_j \neq 0$, entonces

$$\sum_{i \neq j}^n a_i \vec{x}_i = -a_j \vec{x}_j,$$

de donde tendríamos que

$$\vec{x}_j = \sum_{i \neq j}^n \frac{a_i}{-a_j} \vec{x}_i,$$

por lo que el conjunto $\{\vec{x}_i\}_{i=1}^n$ sería linealmente dependiente, y consecuentemente también lo sería X .

\Leftarrow) Si X no es linealmente independiente entonces existe en él un vector que es combinación lineal de los demás, así que podríamos escribir

$$\vec{x} = \sum_{i=1}^n a_i \vec{x}_i, \vec{x}_i \in X \setminus \{\vec{x}\},$$

de donde tenemos

$$\vec{0} = (-1) \vec{x} + \sum_{i=1}^n a_i \vec{x}_i,$$

una combinación lineal de elementos de X con por lo menos un coeficiente distinto de 0. ■

Corolario 18 . *Un conjunto X es linealmente dependiente si y sólo si $\vec{0}$ se puede escribir como combinación lineal de elementos de X , con coeficientes distintos de 0.*

Demostración. Se sigue directamente del teorema anterior. ■

Ejercicio 355 . *Demuestre que el conjunto de los polinomios*

$$\{1, x, x^2, \dots, x^n, \dots\}$$

es linealmente independiente.

Ejercicio 356 . *Demuestre que el conjunto $\{\sin(x), \cos(x)\}$ es un subconjunto linealmente independiente del espacio de las funciones reales de variable real.*

Ejercicio 357 . *Demuestre que el conjunto $\{e^{(x)}, e^{(2x)}\}$ es un subconjunto linealmente independiente del espacio de las funciones reales de variable real.*

8.4 Bases

Definición 110 . *Se dice que X genera $\mathcal{S}(X)$. En particular, se dice que el conjunto X genera el espacio vectorial ${}_F V$, cuando $\mathcal{S}(X) = {}_F V$.*

Definición 111 . *Una base para el espacio vectorial ${}_F V$ es un subconjunto β tal que*

1. β es linealmente independiente y
2. β genera ${}_F V$.

Ejemplo 168 . *Por ejemplo,*

$$\{(1, 0), (0, 1)\}$$

es una base para \mathbb{R}^2 ;

$$\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$$

es una base para \mathbb{R}^3 ; y en general,

$$\{\vec{e}_1, \dots, \vec{e}_i, \dots, \vec{e}_n\}$$

es una base para \mathbb{R}^n , \vec{e}_i es el vector que tiene 1 en su coordenada i -ésima y tiene 0 en sus demás coordenadas.

Ejemplo 169 . $\{1, x, x^2, x^3, x^4 \dots\}$ es la base canónica de $F[x]$.

Definición 112 . Un espacio vectorial es *finitamente generado* si alguno de sus subconjuntos finitos lo genera.

Ejemplo 170 . Los espacios \mathbb{R}^n son finitamente generados puesto que tienen bases finitas.

Teorema 121 . Si $x \in X \subseteq {}_F V$ es tal que $x \in \mathcal{S}(X \setminus \{x\})$, entonces

$$\mathcal{S}(X) = \mathcal{S}(X \setminus \{x\}).$$

Demostración. Como $X \setminus \{x\} \subseteq X$ entonces $\mathcal{S}(X \setminus \{x\}) \subseteq \mathcal{S}(X)$.

Por otra parte $X \setminus \{x\} \subseteq \mathcal{S}(X \setminus \{x\})$, pero como también $x \in \mathcal{S}(X \setminus \{x\})$, entonces $X = (X \setminus \{x\}) \cup \{x\} \subseteq \mathcal{S}(X \setminus \{x\})$, por lo tanto

$$\mathcal{S}(X) \subseteq \mathcal{S}(X \setminus \{x\}).$$

■

Teorema 122 . Un subconjunto linealmente independiente tiene a lo más tantos elementos como un conjunto generador finito.

Demostración. Sea $G = \{y_1, y_2, \dots, y_n\}$ un subconjunto generador finito del espacio vectorial ${}_F V$, y sea X un conjunto linealmente independiente.

Notemos que si X es el conjunto vacío, entonces no tenemos nada que demostrar ($0 \leq n, \forall n \in \mathbb{N}$).

Podemos suponer, que $G \cap X = \{y_1, y_2, \dots, y_k\}$, y que

$$y_1 = x_1, y_2 = x_2, \dots, y_k = x_k,$$

reenumerando si hiciera falta.

Notemos que si $X \subseteq G$, ya tendríamos que $|X| \leq |G|$, por lo tanto podemos suponer que $X \not\subseteq G$. Así que podríamos tomar un elemento $x_{k+1} \in X \setminus G$.

Si consideramos el conjunto

$$\{x_{k+1}, y_1, y_2, \dots, y_n\}$$

notaremos que es linealmente dependiente puesto que cualquier vector es combinación lineal de G .

Pero entonces en el conjunto

$$\{x_{k+1}, y_1, y_2, \dots, y_n\}$$

existe un elemento que es combinación lineal de los anteriores, y que no es ninguno de los elementos de $\{x_{k+1}, y_1, y_2, \dots, y_k\} = \{x_{k+1}, x_1, x_2, \dots, x_k\}$ pues este es linealmente independiente siendo un subconjunto de X . Llamemos y_{i_1} al vector que es combinación lineal de los anteriores. Por el teorema precedente, $\{x_{k+1}, y_1, y_2, \dots, y_n\} \setminus \{y_{i_1}\}$ genera el mismo subespacio que $\{x_{k+1}, y_1, y_2, \dots, y_n\}$, que genera V .

Podemos repetir el argumento con $G_1 = \{x_{k+1}, y_1, y_2, \dots, y_n\} \setminus \{y_{i_1}\}$ (que sigue siendo un conjunto generador con n elementos) y X .

Si $X \not\subseteq G_1$, podemos repetir el argumento tomando un elemento $x_{k+2} \in X \setminus G_1$, y encontrando un elemento y_{i_2} tal que

$$\{x_{k+2}, x_{k+1}, y_1, y_2, \dots, y_n\} \setminus \{y_{i_1}, y_{i_2}\}$$

es un conjunto generador con n elementos.

Este proceso tiene que terminar porque en cada paso quitamos un nuevo elemento de G que es finito. El proceso termina cuando todos los elementos de X han sido sustituidos por elementos de G . Pero esto quiere decir que $|X| \leq n$. ■

Ejemplo 171 . Consideremos el conjunto de vectores en \mathbb{R}^3 ,

$$\{(1, 2, 3), (1, -1, 1), (0, 2, 3)\}.$$

Notemos que el primer vector no es combinación lineal de los anteriores porque no es el vector $\vec{0}$. Notemos también que el segundo vector no es un múltiplo del primero (¿por qué?). Si el tercer vector fuera combinación lineal de los dos anteriores, entonces podríamos resolver la ecuación

$$x(1, 2, 3) + y(1, -1, 1) = (0, 2, 3)$$

que se puede transformar en el sistema de ecuaciones

$$\begin{array}{rcl} x & + y & = 0 \\ 2x & - y & = 2 \\ 3x & + y & = 3 \end{array},$$

eliminando x de las ecuaciones 2 y 3, obtenemos

$$\begin{array}{rcl} x & +y & = 0 \\ -3y & = 2, \\ -2y & = 3 \end{array}$$

que claramente no tiene solución porque implicaría que

$$\frac{-2}{3} = y = \frac{-3}{2}.$$

Esto muestra que el conjunto $\{(1, 2, 3), (1, -1, 1), (0, 2, 3)\}$ es linealmente independiente. Por otra parte, sabemos que la base canónica de \mathbb{R}^3 genera \mathbb{R}^3 . Si tomamos el conjunto $\{(1, 2, 3), (1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ podemos estar seguros de que es linealmente dependiente. Entonces en él debe haber un vector que es combinación lineal de los anteriores, note además que el vector que es combinación lineal de los anteriores no puede ser ninguno de los tres primeros (convéñzase usted mismo). Por lo tanto el vector que es combinación lineal los anteriores es el último. Consecuentemente

$$\{(1, 2, 3), (1, 0, 0), (0, 1, 0)\}$$

es un conjunto generador.

Ahora,

$$\{(1, -1, 1), (1, 2, 3), (1, 0, 0), (0, 1, 0)\}$$

es un conjunto linealmente dependiente, así que debe contener un elemento que es combinación lineal de los anteriores; pero éste no puede ser el primero ni el segundo.

Si $(1, 0, 0) = x(1, -1, 1) + y(1, 2, 3)$, poniendo atención en las coordenadas segunda y tercera, tendríamos que

$$\begin{array}{rcl} -x & +2y & = 0 \\ x & +3y & = 0 \end{array}$$

por lo que

$$\begin{array}{rcl} -x & +2y & = 0 \\ & +5y & = 0 \end{array}$$

lo que implica que $y = 0 \wedge x = 0$. Lo que es imposible.

Por lo tanto $(1, 0, 0)$ no puede ser combinación lineal de los dos vectores

anteriores. Así que otra vez es el último vector el que es combinación lineal de los anteriores. Consecuentemente el conjunto

$$\{(1, -1, 1), (1, 2, 3), (1, 0, 0)\}$$

es un conjunto generador.

Por último, el conjunto

$$\{(0, 2, 3), (1, -1, 1), (1, 2, 3), (1, 0, 0)\}$$

es linealmente dependiente y como en los otros, el vector que es combinación lineal de los anteriores es el último, por lo que el conjunto

$$\{(0, 2, 3), (1, -1, 1), (1, 2, 3)\}$$

es un conjunto generador de \mathbb{R}^3 . Como ya habíamos observado que también es un conjunto linealmente independiente, entonces es una base para \mathbb{R}^3 .

Corolario 19 . Cualquier subconjunto de \mathbb{R}^n con más de n vectores es linealmente dependiente. Por ejemplo, un conjunto de cuatro vectores en \mathbb{R}^3 es linealmente dependiente.

Teorema 123 . Si un espacio vectorial es finitamente generado, cualesquiera dos bases tienen el mismo número de elementos.

Demostración. Sea X un conjunto finito que genera FV . Si β y γ son dos bases de FV , entonces por ser conjuntos linealmente independientes ambas bases tienen a lo más tantos elementos como $|X|$. En particular, tanto β como γ son conjuntos generadores finitos.

Como β es un conjunto linealmente independiente y como γ es un conjunto generador finito, entonces $|\beta| \leq |\gamma|$.

Por simetría, $|\gamma| \leq |\beta|$. ■

Definición 113 . La dimensión de un espacio vectorial finitamente generado es el número de elementos en cualquiera de sus bases.

Teorema 124 . Un espacio vectorial finitamente generado tiene base.

Demostración. Tomemos un conjunto generador X del espacio V . Si X es linealmente independiente, entonces X ya es una base.

En caso contrario, $\exists \vec{x} \in X$ tal que $\mathcal{S}(X \setminus \{\vec{x}\}) = \mathcal{S}(X) = V$.

Es decir, se podría quitar un elemento a X y seguir teniendo un conjunto generador.

Podemos repetir el argumento con $X \setminus \{\vec{x}\}$, que tiene un elemento menos que X .

El argumento se repite tantas veces como se requiera, el proceso tiene que terminar, dado que X es finito. El proceso termina cuando encontramos un conjunto que genera y que es linealmente independiente, es decir, cuando obtenemos una base para V . ■

Teorema 125 . *En un espacio vectorial finitamente generado, cualquier subconjunto linealmente independiente puede extenderse a una base.*

Demostración. Tomemos un subconjunto finito I linealmente independiente del espacio V . Por ejemplo, podemos tomar $I = \emptyset$.

Si I genera V , entonces I ya es una base para V . En caso contrario, $\exists \vec{x} \in V \setminus \mathcal{S}(X)$, de aquí que $I \cup \{\vec{x}\}$ es linealmente independiente (considere que los elementos de I van antes que \vec{x} , entonces note que en $I \cup \{\vec{x}\}$ ningún vector es combinación lineal de las anteriores).

Podemos repetir el argumento con el conjunto $I \cup \{\vec{x}\}$.

Repetimos el argumento tantas veces como sea necesario. Este proceso tiene que terminar ya que un conjunto linealmente independiente no puede tener más elementos que un conjunto generador, y el espacio tiene un conjunto generador, por hipótesis.

Cuando el proceso termina, es porque hemos encontrado un conjunto linealmente independiente que ya genera V . ■

Ejemplo 172 . *$\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ es una base para \mathbb{R}^3 .*

Tomemos el conjunto $B = \{(1, 2, 3), (-1, -1, -1)\}$ que es linealmente independiente porque ninguno de sus elementos es combinación lineal de los anteriores. El conjunto B no puede ser una base de \mathbb{R}^3 , ya que sólo tiene dos elementos. Algún elemento de la base canónica no pertenece al espacio generado por B , ya que B no genera \mathbb{R}^3 .

Veamos si $(1, 0, 0)$ es combinación lineal de B , es decir, veamos si se puede resolver

$$(1, 0, 0) = x(1, 2, 3) + y(-1, -1, -1),$$

esta ecuación produce el sistema

$$\begin{aligned} x - y &= 1 \\ 2x - y &= 0 \\ 3x - y &= 0 \end{aligned}$$

podemos eliminar x de la segunda y de la tercera ecuación:

$$\begin{aligned} x - y &= 1 \\ y &= -2 \\ 2y &= -3 \end{aligned}$$

con lo que se ve que el sistema no tiene solución.

Por lo tanto, $B \cup \{(1, 0, 0)\}$ es linealmente independiente, y es una base para \mathbb{R}^3 . (Véase el siguiente teorema).

Observación 104 . Para un espacio vectorial ${}_F V$ es lo mismo decir que es finitamente generado, a decir que es de dimensión finita. Esto se debe a que un conjunto generador finito X de V , contiene una base. Por otra parte una base ya es un conjunto generador de V .

Teorema 126 . Son equivalentes para un subconjunto β de un espacio vectorial V , de dimensión finita n :

1. β es una base de V y $|\beta| = n$.
2. β es linealmente independiente y $|\beta| = n$.
3. β es un conjunto generador de V y $|\beta| = n$.

Demostración. Es claro que 1) \implies 2) y que 1) \implies 3).

Debería ser claro también que basta demostrar la equivalencia entre 2) y 3).

2) \implies 3) Supongamos que β es linealmente independiente y que tiene n elementos. Si no fuera un generador de V , existiría un elemento $\vec{x} \in V \setminus \mathcal{S}(\beta)$. Pero entonces $\beta \cup \{\vec{x}\}$ sería linealmente independiente (si se piensa que los elementos de β van antes que \vec{x} , entonces en $\beta \cup \{\vec{x}\}$, ningún vector es combinación lineal de los anteriores). Entonces hay un conjunto linealmente independiente con $n + 1$ elementos. Esto no puede ser porque como la dimensión de V es n , entonces debe haber un conjunto generador (una base)

con n elementos. Ya vimos que un conjunto linealmente independiente tiene a lo más tantos elementos como un conjunto que genera. La contradicción demuestra que β genera V .

3) \Rightarrow 2)

Supongamos ahora que β es un conjunto con n elementos, que genera V . Si β no fuera linealmente independiente, $\exists \vec{x} \in \beta$, tal que \vec{x} es combinación lineal de los demás elementos de β . Pero entonces $\beta \setminus \{\vec{x}\}$ genera lo mismo que lo que genera β . Es decir que $\beta \setminus \{\vec{x}\}$ es un conjunto generador con $n-1$ elementos. Como una base de V es un conjunto linealmente independiente con n elementos, y que un conjunto linealmente independiente tiene a lo más tantos elementos como un conjunto generador, obtenemos la siguiente contradicción:

$$n \leq n-1.$$

■

Teorema 127 . *Son equivalentes para $\beta \subseteq {}_F V \neq \{\vec{0}\}$:*

1. β es una base para ${}_F V$.
2. $\forall \vec{x} \in V \setminus \{\vec{0}\}$ existen únicos $\vec{x}_1, \dots, \vec{x}_n \in \beta$ y $c_1, c_2, \dots, c_n \in F \setminus \{0\}$ tales que

$$\vec{x} = c_1 \vec{x}_1 + \dots + c_n \vec{x}_n.$$

Demostración. 1) \Rightarrow 2) Si $\vec{x} \in V \setminus \{\vec{0}\}$ entonces existen $\vec{x}_1, \dots, \vec{x}_n \in \beta$ y $c_1, c_2, \dots, c_n \in F \setminus \{0\}$ reales que

$$\vec{x} = c_1 \vec{x}_1 + \dots + c_n \vec{x}_n,$$

pues β genera V . Resta ver que esta expresión es única, supongamos, reordenando los términos si es necesario, que

$$\vec{x} = d_1 \vec{x}_1 + \dots + d_k \vec{x}_k + d_{k+1} \vec{y}_{k+1} + \dots + d_m \vec{y}_m,$$

donde $d_j \neq 0$ y $\{\vec{y}_{k+1}, \dots, \vec{y}_m\} \subseteq \beta \setminus \{\vec{x}_1, \dots, \vec{x}_n\}$.
entonces

$$\begin{aligned} \vec{0} &= \vec{x} - \vec{x} = \\ &= (c_1 - d_1) \vec{x}_1 + \dots + (c_k - d_k) \vec{x}_k + c_{k+1} \vec{x}_{k+1} + \dots + c_n \vec{x}_n - d_{k+1} \vec{y}_{k+1} - \dots - d_m \vec{y}_m, \end{aligned}$$

como β es linealmente independiente, entonces $n = k = m$ (si $n > k$, entonces $c_n = 0$, contra las hipótesis acerca de los coeficientes). Además tenemos que $c_i - d_i = 0$ para cada $i \in \{1, \dots, k\}$.

2) \Rightarrow 1) Si vale 2) es claro que β genera V . Si β fuera linealmente dependiente existiría un elemento de β que es combinación lineal de otros elementos de β , digamos

$$\vec{x} = \alpha_1 \vec{y}_1 + \dots + \alpha_n \vec{y}_n$$

que serían dos maneras distintas de expresar \vec{x} como combinación lineal de elementos de β . ■

8.4.1 Intersección de subespacios y suma de subespacios

Como ya hemos visto, la intersección de la familia de subespacios de FV , sigue siendo un subespacio.

Teorema 128 . *Sean W_1, W_2 dos subespacios del espacio de dimensión finita V . Entonces: $\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2)$.*

Demostración. Sea γ una base para $W_1 \cap W_2$. γ se puede extender a una base $\gamma \cup \lambda_1$ de W_1 y también se puede extender a una base $\gamma \cup \lambda_2$ de W_2 . Demostraremos que $\gamma \cup \lambda_1 \cup \lambda_2$ es una base para $W_1 + W_2$.

Como $\gamma \cup \lambda_1 \cup \lambda_2 = (\gamma \cup \lambda_1) \cup (\gamma \cup \lambda_2)$, entonces

$$\begin{aligned} \mathcal{S}(\gamma \cup \lambda_1 \cup \lambda_2) &= \mathcal{S}((\gamma \cup \lambda_1) \cup (\gamma \cup \lambda_2)) = \\ &= \mathcal{S}(\gamma \cup \lambda_1) + \mathcal{S}(\gamma \cup \lambda_2) = \\ &= W_1 + W_2. \end{aligned}$$

Por lo tanto $\gamma \cup \lambda_1 \cup \lambda_2$ genera $W_1 + W_2$.

Como $\gamma \cup \lambda_1$ es una base para W_1 , en particular es linealmente independiente. Veamos ahora que ningún elemento en $\gamma \cup \lambda_1 \cup \lambda_2$, es combinación lineal de los anteriores. En caso de haberlo, sería un elemento $\vec{x} \in \lambda_2$, entonces

$$\vec{x} = \sum_{\vec{y}_j \in \gamma} c_j \vec{y}_j + \sum_{\vec{u}_k \in \lambda_1} d_k \vec{u}_k + \sum_{\vec{v}_l \in \lambda_2} h_l \vec{v}_l,$$

por lo que

$$\vec{x} - \sum_{\vec{v}_l \in \lambda_2} h_l \vec{v}_l = \sum_{\vec{y}_j \in \gamma} c_j \vec{y}_j + \sum_{\vec{u}_k \in \lambda_1} d_k \vec{u}_k \in W_1,$$

por otra parte, $\vec{x} - \sum_{\vec{v}_l \in \lambda_2} h_l \vec{v}_l \in \mathcal{S}(\lambda_2) \subseteq W_2$. Por lo que

$$\vec{x} - \sum_{\vec{v}_l \in \lambda_2} h_l \vec{v}_l \in W_1 \cap W_2 = \mathcal{S}(\gamma).$$

De lo anterior, tenemos que

$$\sum_{\vec{y}_j \in \gamma} c_i \vec{y}_j + \sum_{\vec{u}_k \in \lambda_1} d_k \vec{u}_k \in \mathcal{S}(\gamma),$$

de aquí que $\sum_{\vec{u}_k \in \lambda_1} d_k \vec{u}_k \in \mathcal{S}(\gamma)$. Digamos que $\sum_{\vec{u}_k \in \lambda_1} d_k \vec{u}_k = \sum_{\vec{y}_j \in \gamma} b_i \vec{y}_j$, de dónde obtenemos

$$\sum_{\vec{u}_k \in \lambda_1} d_k \vec{u}_k - \sum_{\vec{y}_j \in \gamma} b_i \vec{y}_j = \vec{0},$$

pero como $\gamma \cup \lambda_1$ es linealmente independiente, la ecuación anterior implica que todos los coeficientes son 0. En particular, cada $d_k = 0$. Por lo que de la ecuación original

$$\vec{x} = \sum_{\vec{y}_j \in \gamma} c_i \vec{y}_j + \sum_{\vec{u}_k \in \lambda_1} d_k \vec{u}_k + \sum_{\vec{v}_l \in \lambda_2} h_l \vec{v}_l,$$

tenemos que

$$\vec{x} = \sum_{\vec{y}_j \in \gamma} c_i \vec{y}_j + \sum_{\vec{v}_l \in \lambda_2} h_l \vec{v}_l,$$

pero esto contradice que $\gamma \cup \lambda_2$ es linealmente independiente.

Esta contradicción demuestra que $\gamma \cup \lambda_1 \cup \lambda_2$ es una base para $W_1 + W_2$.

Por lo tanto

$$\begin{aligned} \dim(W_1 + W_2) &= |\gamma \cup \lambda_1 \cup \lambda_2| = \\ &= |\gamma \cup \lambda_1| + |\lambda_2| = |\gamma \cup \lambda_1| + |\gamma| + |\lambda_2| - |\gamma| = \\ &= \dim(W_1) + \dim(W_2) - |\gamma| = \\ &= \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2). \blacksquare \end{aligned}$$

8.5 Producto punto

Definición 114 . *Se define el producto punto*

$$\cdot : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$$

de la manera siguiente: si

$$\vec{v} = (x_1, x_2, \dots, x_n) \quad y \quad \vec{w} = (y_1, y_2, \dots, y_n),$$

entonces

$$\vec{v} \cdot \vec{w} = x_1y_1 + x_2y_2 + \dots + x_ny_n.$$

Observación 105 . El producto punto así definido tiene las siguientes propiedades:

1. $\vec{v} \cdot \vec{w} = \vec{w} \cdot \vec{v}$. (Esto se sigue de la commutatividad de la suma y del producto en \mathbb{R}).
2. $\vec{v} \cdot (\vec{w}_1 + \vec{w}_2) = \vec{v} \cdot \vec{w}_1 + \vec{v} \cdot \vec{w}_2$. (Consecuencia de la distributividad en \mathbb{R}).
3. $(c\vec{v}) \cdot \vec{w} = c(\vec{v} \cdot \vec{w}) = \vec{v} \cdot (c\vec{w})$.
4. $\vec{v} \cdot \vec{v} \geq 0$ y $(\vec{v} \cdot \vec{v} = 0 \iff \vec{v} = \vec{0})$.

Nota 2 . Si $\cdot : V \times V \longrightarrow F$ satisface las cuatro propiedades anteriores, se dice que \cdot es un producto interior en V (el campo es F).

Definición 115 . \vec{v} y \vec{w} en \mathbb{R}^n son ortogonales si $\vec{v} \cdot \vec{w} = \vec{0}$. (Escribimos $\vec{v} \perp \vec{w}$).

Así, por ejemplo,

$(1, 0) \perp (0, 1)$ en \mathbb{R}^2 y $(1, 0, 0) \perp (0, 1, 0)$ en \mathbb{R}^3 .

Nota 3 . La definición de ortogonalidad puede extenderse a espacios que tengan asociado cualquier producto escalar no degenerado (uno que satisfaga que $\vec{v} \neq 0 \implies \vec{v} \cdot \vec{v} \neq 0$) pero en estas notas sólo nos interesan los espacios \mathbb{R}^n y la Geometría euclíadiana, por lo que no abundaremos en el tema.

Definición 116 . La norma (euclíadiana) de \mathbb{R}^n es:

$$\begin{array}{rcl} \|\quad\|: & \mathbb{R}^n & \longrightarrow & \mathbb{R}^+ \cup \{0\} \\ \vec{v} & \longmapsto & & \sqrt{\vec{v} \cdot \vec{v}} \end{array}.$$

Observación 106

$$\|\cdot\|: \mathbb{R}^n \longrightarrow \mathbb{R}^+ \cup \{0\}$$

satisface:

1. $\|\vec{v}\| \geq 0$ y ($\|\vec{v}\| = 0 \iff \vec{v} = \vec{0}$).
2. $\|c\vec{v}\| = |c| \|\vec{v}\|$, $c \in \mathbb{R}$, $\vec{v} \in \mathbb{R}^n$.
3. $\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$ (Desigualdad del triángulo).

Para comprobar 3), calculemos $(\vec{v} + \vec{w}) \cdot (\vec{v} + \vec{w})$:

$$(\vec{v} + \vec{w}) \cdot (\vec{v} + \vec{w}) = \vec{v} \cdot \vec{v} + 2(\vec{v} \cdot \vec{w}) + \vec{w} \cdot \vec{w},$$

así que

$$\|\vec{v} + \vec{w}\|^2 = \|\vec{v}\|^2 + 2(\vec{v} \cdot \vec{w}) + \|\vec{w}\|^2.$$

Basta ahora demostrar que $2(\vec{v} \cdot \vec{w}) \leq 2\|\vec{v}\|\|\vec{w}\|$, pues en este caso tendremos que

$$\|\vec{v} + \vec{w}\|^2 \leq (\|\vec{v}\| + \|\vec{w}\|)^2.$$

Esto es lo que afirma el Lema de Schwarz:

Lema 25 . $2(\vec{v} \cdot \vec{w}) \leq 2\|\vec{v}\|\|\vec{w}\|$.

Demostración. Notemos que la afirmación se cumple cuando $\vec{w} = \vec{0}$. Supongamos ahora que $\vec{w} \neq \vec{0}$.

Notemos que existe $\lambda \in \mathbb{R}$ tal que $(\vec{v} - \lambda\vec{w}) \perp \vec{w}$: simplemente de

$$(\vec{v} - \lambda\vec{w}) \cdot \vec{w} = 0,$$

despejemos λ :

$$(\vec{v} \cdot \vec{w} - \lambda\vec{w} \cdot \vec{w} = 0) \implies \left(\lambda = \frac{\vec{v} \cdot \vec{w}}{\|\vec{w}\|^2} \right).$$

Para esta λ tenemos que

$$\begin{aligned} 0 &\leq \|\vec{v} - \lambda\vec{w}\|^2 = (\vec{v} - \lambda\vec{w}) \cdot (\vec{v} - \lambda\vec{w}) = \\ &= (\vec{v} - \lambda\vec{w}) \cdot \vec{v} \\ &= \|\vec{v}\|^2 - \lambda(\vec{v} \cdot \vec{w}). \end{aligned}$$

Así que

$$0 \leq \|\vec{v}\|^2 - \lambda(\vec{v} \cdot \vec{w}),$$

de donde

$$\lambda(\vec{v} \cdot \vec{w}) \leq \|\vec{v}\|^2,$$

como

$$\lambda = \frac{\vec{v} \cdot \vec{w}}{\|\vec{w}\|^2}$$

tenemos que

$$(\vec{v} \cdot \vec{w})^2 \leq \|\vec{v}\|^2 \|\vec{w}\|^2,$$

de aquí que

$$\vec{v} \cdot \vec{w} \leq |\vec{v} \cdot \vec{w}| \leq \|\vec{v}\| \|\vec{w}\|.$$

■

Ejemplo 173 . Para el caso $n = 2$, la definición de ortogonalidad reproduce el Teorema de Pitágoras. (Si \vec{v}, \vec{w} son ortogonales, es decir, si $\vec{v} \cdot \vec{w} = 0$, entonces

$$\begin{aligned} \|\vec{v} + \vec{w}\|^2 &= (\vec{v} + \vec{w}) \cdot (\vec{v} + \vec{w}) = \vec{v} \cdot \vec{v} + 2\vec{v} \cdot \vec{w} + \vec{w} \cdot \vec{w} = \\ &= \|\vec{v}\|^2 + \|\vec{w}\|^2, \end{aligned}$$

$$\text{Es decir, } \|\vec{v} + \vec{w}\| = \sqrt{\|\vec{v}\|^2 + \|\vec{w}\|^2}$$

Nota 4 . Apuntamos aquí, que también existen otras normas, (en las que tampoco estamos interesados). Los espacios que tienen alguna norma se llaman espacios normados.

Definición 117 . En un espacio normado, un vector es unitario si su norma es 1.

Ejemplo 174 . En \mathbb{R}^n , para cada $i \in \{1, \dots, n\}$

$$\vec{e}_i = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \end{pmatrix} \quad \left. \right\} \begin{array}{l} i - \text{ésima} \\ \text{coordenada} \end{array}$$

es unitario.

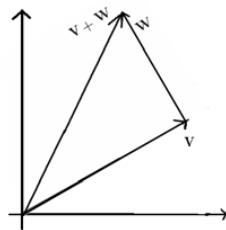


Figura 8.3:

Ejemplo 175 . En \mathbb{R}^3

$$\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right)$$

también es unitario.

Definición 118 . $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\} \subseteq \mathbb{R}^m$ es un conjunto ortonormal si

1. $\|\vec{v}_i\| = 1 \quad \forall i \in \{1, 2, \dots, n\}$ y
2. $\vec{v}_i \perp \vec{v}_j$ si $i \neq j$.

Nota 5 . Observe que todo conjunto finito ortonormal es linealmente independiente. En efecto, si

$$c_1 \vec{v}_1 + c_2 \vec{v}_2 + \dots + c_n \vec{v}_n = \vec{0}$$

entonces para cada $i = 1, \dots, n$ se tiene que

$$0 = \vec{v}_i \cdot \vec{0} = \vec{v}_i \cdot c_1 \vec{v}_1 + c_2 \vec{v}_2 + \dots + c_n \vec{v}_n = c_i \|\vec{v}_i\|^2 = c_i$$

Teorema 129 . *Todo espacio vectorial real V de dimensión finita n , tiene bases ortogonales y todo subconjunto ortogonal, linealmente independiente puede completarse a una base ortogonal.*

Demostración. Supongamos que $T = \{v_1, v_2, \dots, v_m\}$ es un conjunto ortogonal linealmente independiente, demostremos que T se puede extender a una base ortogonal de V , por inducción sobre $n - m$.

Notemos primero que $m \leq n$, por lo que $n - m \geq 0$.

Base. Si $n - m = 0$, entonces $n = m$ y el conjunto T ya es una base pues es un conjunto linealmente independiente con n elementos.

Supongamos que $n - m > 0$. Así que $m < n$ y por lo tanto T no puede generar V . Que la dimensión de V sea n significa que existe una base $\{w_1, w_2, \dots, w_n\}$ de V con n elementos. Si cada w_i perteneciera al subespacio generado por T , entonces T generaría V .

Por lo tanto, existe w_i tal que $w_i \notin \mathcal{S}(T)$. Esto significa que

$$T \cup \{w_i\}$$

es linealmente independiente, y $\dim(\mathcal{S}(T \cup \{w_i\})) = m + 1$.

Hagamos

$$z_{m+1} = w_i - \left(\sum_{j=1}^m \frac{w_i \cdot v_j}{\|v_j\|^2} v_j \right).$$

Usando la ortogonalidad del conjunto T , podemos notar que

$$z_{m+1} \perp v_j, \quad j \in \{1, 2, \dots, m\}.$$

Entonces

$$T \cup \{z_{m+1}\}$$

es un conjunto ortogonal, pero además z_{m+1} no es combinación lineal de los elementos de T , pues si lo fuera también w_i sería combinación lineal de los elementos de T .

Por lo tanto

$$T \cup \{z_{m+1}\}$$

es un conjunto ortogonal linealmente independiente con $m + 1$ vectores, como $n - (m + 1) < n - m$, podemos usar la hipótesis de inducción para concluir que

$$T \cup \{z_{m+1}\}$$

se puede completar a una base ortogonal de V . Esta base contiene a T . ■

Definición 119 . Sea $W \leq V$ un subespacio de un espacio vectorial real de dimensión finita. El complemento ortogonal de W , W^\perp es:

$$\{\vec{v} \in V \mid \vec{v} \perp \vec{w}, \forall \vec{w} \in W\}.$$

Ejemplo 176 . Si $W = \{(a, b, 0) \mid a, b \in \mathbb{R}\}$, entonces

$$W^\perp = \{(x, y, c) \in \mathbb{R}^3 \mid ax + by = 0, \forall a, b \in \mathbb{R}\}.$$

Ejemplo 177 . Si

$$W = \{(x, y, z) \in \mathbb{R}^3 \mid 2x + 3y - z = 0\},$$

entonces

$$W^\perp = \left\{ \vec{x} \in \mathbb{R}^3 \mid \vec{x} = t \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}, t \in \mathbb{R} \right\}.$$

Observación 107 . Si $W \leq V$ es un subespacio de V y β es un subconjunto de V tal que cada $\vec{w} \in W$ es combinación lineal de β (β genera W), entonces

$$W^\perp = \beta^\perp.$$

Demostración. Es claro que como $\beta \subseteq W$ entonces cualquier vector ortogonal a cualquier elemento de W es, en particular, ortogonal a cualquier elemento de β . Es decir $W^\perp \subseteq \beta^\perp$. Por otra parte, si $\vec{v} \in \beta^\perp$, y $\vec{w} \in W$, entonces $\vec{w} = c_1 \vec{x}_1 + \dots + c_k \vec{x}_k$ con $c_i \in F$ y $\vec{x}_i \in \beta$,

por lo tanto

$$\vec{v} \cdot \vec{w} = c_1 (\vec{v} \cdot \vec{x}_1) + \dots + c_k (\vec{v} \cdot \vec{x}_k) = 0 + \dots + 0 = 0,$$

por lo que $\vec{v} \in W^\perp$. ■

Teorema 130 . Si $W \leq V$ y V es un espacio de dimensión finita n con producto punto entonces W^\perp es un subespacio de V , y

$$\dim W + \dim W^\perp = \dim V.$$

Demuestra. $\vec{0} \cdot \vec{w} = (\vec{0} \vec{0}) \cdot \vec{w} = 0 (\vec{0} \cdot \vec{w}) = 0, \forall \vec{w} \in W$, por lo que $\vec{0} \in W^\perp$.

Si $\vec{v}_1, \vec{v}_2 \in W^\perp$, entonces $(\vec{v}_1 + \vec{v}_2) \cdot \vec{w} = \vec{v}_1 \cdot \vec{w} + \vec{v}_2 \cdot \vec{w} = 0 + 0 = 0$. Por lo tanto $(\vec{v}_1 + \vec{v}_2) \in W^\perp$.

Si $c \in F, \vec{v} \in W^\perp$, entonces $(c \vec{v}) \cdot \vec{w} = c (\vec{v} \cdot \vec{w}) = c0 = 0$.

Supongamos ahora que β es una base ortogonal de W , extendámosla a una base ortogonal γ de V . Digamos que

$$\gamma = \beta \cup \beta'.$$

Entonces $\beta' \subseteq \beta^\perp = W^\perp$, por la observación anterior. β' es linealmente independiente en V y por lo tanto también es un subconjunto linealmente independiente de W^\perp .

Si $\vec{w} \in W^\perp$, escribamos $\vec{w} = \vec{x} + \vec{y}$, donde \vec{x} es una combinación lineal de elementos de β y \vec{y} lo es de β' .

Entonces

$$0 = \vec{w} \cdot \vec{x} = (\vec{x} + \vec{y}) \cdot \vec{x} = \vec{x} \cdot \vec{x} = \|\vec{x}\|^2$$

ya que $\vec{x} \in W$ y $\vec{y} \in W^\perp$. De aquí se sigue que $\vec{x} = \vec{0}$. Por lo tanto, cualquier elemento de W^\perp es combinación lineal de β' . Por lo tanto β' es una base de W^\perp . Así que

$$\dim(V) = |\gamma| = |\beta \cup \beta'| = |\beta| + |\beta'| = \dim(W) + \dim(W^\perp).$$

■

8.6 Matrices

Definición 120 . Una función $A : \{1, \dots, m\} \times \{1, \dots, n\} \rightarrow \mathbb{R}$ se llama matriz de m renglones y n columnas con coeficientes reales. En lugar de escribir $A(i, j)$, escribiremos $A_{i,j}$.

Se suele describir una matriz A de m por n por medio del arreglo rectangular

$$\begin{pmatrix} A_{1,1} & A_{1,2} & \dots & A_{1,n} \\ A_{2,1} & A_{2,2} & \dots & A_{2,n} \\ \vdots & \vdots & & \vdots \\ A_{m,1} & A_{m,2} & \dots & A_{m,n} \end{pmatrix},$$

denotaremos A_i el i -ésimo renglón del arreglo, por A^j la j -ésima columna del arreglo.

Notación 16 . Denotaremos por $M_{m,n}(\mathbb{R})$ al conjunto de todas las matrices de m por n con coeficientes en \mathbb{R} .

Ejercicio 358 . Demuestre que $M_{m,n}(\mathbb{R})$ es un conjunto.

8.6.1 El rango de una matriz

Definición 121 . Sean

$$A = (a_{i,j}) \in M_{m,n}(\mathbb{R}),$$

$$\beta = \{A_1, \dots, A_m\},$$

el conjunto de los renglones de la matriz, sea W el subespacio de \mathbb{R}^n generado por β . El rango (de renglones) de A es $\dim(W)$ y se denota

$$\text{rango}(A).$$

Teorema 131 . Sea $(^*H)$ un sistema homogéneo de ecuaciones lineales, y sea

$$A = (a_{i,j})$$

la matriz de sus coeficientes. Definamos $\beta = \{A_1, \dots, A_m\}$, en donde $A_i = (a_{i,1}, \dots, a_{i,n})$ es el i -ésimo renglón de A y sea W el subespacio generado por β . Entonces el conjunto

$$S_0 \leq \mathbb{R}^n$$

de soluciones de $(^*H)$ es precisamente el complemento ortogonal de W , es decir:

$$S_0 = W^\perp.$$

(Ver la sección 8.9, más adelante).

Demostración. En efecto,

$$\vec{c} = (c_1, \dots, c_n) \in S_0 \iff \text{para cada } i, i \in \{1, 2, \dots, n\}$$

$$A_i \cdot \vec{c} = a_{i1}c_1 + \dots + a_{in}c_n = 0,$$

es decir

$$\vec{c} \in S_0 \iff \vec{c} \perp A_i \quad \forall A_i \in \beta.$$

De aquí se sigue que

$$S_0 = W^\perp.$$

■

Se tiene que

$$n = \dim(W^\perp) + \dim(W) = \dim S_0 + \text{rango}(A).$$

O sea:

$$\dim S_0 = n - \text{rango}(A)$$

y esta última igualdad permite decir que S_0 es un espacio vectorial (subespacio de \mathbb{R}^n) con $n - \text{rango}(A)$ parámetros (o grados de libertad).

La dimensión de W , coincide con el número máximo de vectores linealmente independientes que pueden formarse con los elementos de β , que es, a su vez, el orden del determinante diferente de cero de mayor orden, que puede obtenerse del conjunto de submatrices cuadradas de la matriz A .

La observación anterior tiene como consecuencia inmediata el hecho de que para una matriz A , el espacio generado por sus renglones tiene la misma dimensión que el espacio que generan sus columnas a pesar de que el primero es un subespacio de \mathbb{R}^n mientras que el segundo lo es de \mathbb{R}^m .

Esto se sigue de que si A es una matriz de m renglones y n columnas entonces considerando el sistema de ecuaciones

$$A \cdot \vec{x} = \vec{0}$$

donde

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

notamos que esto se puede reescribir de la siguiente manera:

$$x_1 A^1 + x_2 A^2 + \dots + x_n A^n = \vec{0} \tag{8.2}$$

donde $\{A^1, \dots, A^n\}$ es el conjunto de las columnas de la matriz A .

Supongamos que $\{A^{l_1}, \dots, A^{l_s}\}$ es una base para el subespacio de \mathbb{R}^m generado por las columnas de A .

Podríamos reescribir la ecuación 8.2 de la siguiente manera:

$$x_{l_1}A^{l_1} + x_{l_2}A^{l_2} + \dots + x_{l_s}A^{l_s} = -\sum_{j \notin \{l_1, \dots, l_s\}} x_j A^j. \quad (8.3)$$

Notemos ahora que para cualquier elección de las x_j en el campo, existen únicos elementos del campo b_{l_1}, \dots, b_{l_s} tales que satisfacen 8.3 al hacer las sustituciones. Esto es una consecuencia de que $\{A^{l_1}, \dots, A^{l_s}\}$ es una base para el espacio generado por las columnas de A .

Si perder generalidad, y para simplificar la notación, supongamos que las s primeras columnas de A generan el espacio de las columnas. (Si esto no fuera así, podríamos reordenar las incógnitas y las columnas).

Con esta simplificación podemos escribir

$$x_1A^1 + x_2A^2 + \dots + x_sA^s = -\sum_{j>s} x_j A^j. \quad (8.4)$$

Como habíamos hecho notar, hay una solución para cada $e_i \in \mathbb{R}^{n-s}$. Sean

$$\begin{aligned} \vec{u}_1 &= (t_{1,1}, \dots, t_{s,1}, 1, 0, \dots, 0), \\ \vec{u}_2 &= (t_{1,2}, \dots, t_{s,2}, 0, 1, \dots, 0), \\ &\vdots \\ \vec{u}_{n-s} &= (t_{1,n-s}, \dots, t_{s,n-s}, 0, 0, \dots, 1). \end{aligned}$$

Podemos notar que estas $n - s$ soluciones son linealmente independientes, pues si $\sum d_k \vec{u}_k = \vec{0}$, note que las coordenadas $s + l$ -ésimas son $d_l = 0$.

También es claro que si (b_1, \dots, b_n) es una solución es porque

$$b_1A^1 + b_2A^2 + \dots + b_sA^s = -\sum_{j>s} b_j A^j.$$

Como también

$$b_{s+1}\vec{u}_1 + b_{s+2}\vec{u}_2 + \dots + b_n\vec{u}_{n-s} = (\alpha_1, \dots, \alpha_s, b_{s+1}, \dots, b_n)$$

es una solución entonces

$$\alpha_1A^1 + \alpha_2A^2 + \dots + \alpha_sA^s = -\sum_{j>s} \alpha_j A^j.$$

Como la expresión de cada vector en términos de una base es única, tenemos que

$$(b_1, \dots, b_n) = (\alpha_1, \dots, \alpha_s, b_{s+1}, \dots, b_n) = b_{s+1} \overrightarrow{u_1} + b_{s+2} \overrightarrow{u_2} + \dots + b_n \overrightarrow{u_{n-s}}$$

así que $\{\overrightarrow{u_1}, \overrightarrow{u_2}, \dots, \overrightarrow{u_{n-s}}\}$ es una base para el conjunto de soluciones.

Por lo tanto, $\dim(S_0) = n - s$, donde s es la dimensión del espacio de columnas de A .

Por lo tanto, la dimensión del espacio generado por las columnas de A es

$$n - \dim(S_0),$$

que como ya vimos también es la dimensión del espacio generado por los renglones de A .

Si llamamos rango de columna de la matriz A a la dimensión del espacio generado por las columnas de A y llamamos rango de renglón la dimensión del espacio generado por los renglones de A podemos hacer la siguiente conclusión.

Proposición 32 . *El rango de renglón de una matriz A coincide con su rango de columna.*

Notación 17 . *En vista de la proposición anterior hablaremos simplemente del rango de una matriz.*

8.7 Funciones lineales

Definición 122 . *Una función $T : {}_F V \longrightarrow {}_F W$ es lineal si:*

1. T respeta la suma, es decir que

$$T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v}), \quad \forall \vec{u}, \vec{v} \in V.$$

2. T respeta multiplicación por escalares, es decir que

$$T(c\vec{v}) = cT(\vec{v}), \quad \forall c \in F, \forall \vec{v} \in V.$$

Ejemplos 178

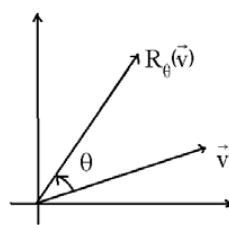


Figura 8.4:

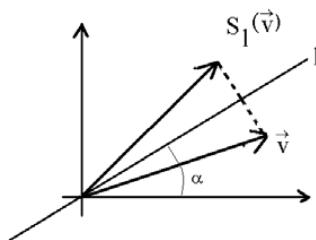


Figura 8.5:

1. Rotaciones. Consideremos la función

$$R_\theta : \mathbb{R}^2 \longrightarrow \mathbb{R}^2,$$

que rota cada vector por un ángulo θ alrededor del origen. Una rotación en \mathbb{R}^2 es una función lineal.

2. Reflexiones. Tomemos una línea recta l que pase por $(0,0)$ en \mathbb{R}^2 . La función S_l que envía un vector \vec{v} a su imagen reflejada sobre la línea l es una función lineal ($S_l(\vec{v})$ tiene el mismo tamaño que \vec{v} y el ángulo entre $S_l(\vec{v})$ y l es el mismo que el ángulo entre l y \vec{v}).

3. Proyecciones. Consideremos la función $p_i : \mathbb{R}^n \longrightarrow \mathbb{R}$ que envía el

vector

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

a su i -ésima coordenada x_i , es una función lineal.

4. Homotecias. Multiplicar por un escalar

$$\begin{aligned} c \cdot \underline{} : V &\longrightarrow V \\ \vec{v} &\longmapsto c\vec{v} \end{aligned}$$

es una función lineal

Lema 26 . Si β es una base del espacio vectorial FV entonces cada vector de V se puede expresar de manera única como combinación lineal (con coeficientes no nulos) de elementos de β

Demostración. Como la bases generan, cada vector de V es combinación lineal de elementos de β .

Si

$$c_1\vec{x}_1 + \dots + c_n\vec{x}_n = d_1\vec{y}_1 + \dots + d_m\vec{y}_m$$

con $c_i, d_j \in F \setminus \{0\}$, $\vec{x}_i, \vec{y}_j \in \beta$, entonces

$$c_1\vec{x}_1 + \dots + c_n\vec{x}_n - d_1\vec{y}_1 - \dots - d_m\vec{y}_m = \vec{0}. \quad (8.5)$$

Como β es linealmente independiente y los coeficientes en la ecuación anterior son distintos de 0, entonces \vec{y}_m debe coincidir con alguna \vec{x}_i , pues en caso contrario $\{\vec{x}_1, \dots, \vec{x}_n, \vec{y}_1, \dots, \vec{y}_m\}$ sería un subconjunto linealmente dependiente del conjunto linealmente independiente β , lo que no es posible. De la misma manera, cada \vec{y}_j debe pertenecer al conjunto $\{\vec{x}_1, \dots, \vec{x}_n\}$.

Por simetría, también tenemos que $\{\vec{x}_1, \dots, \vec{x}_n\} \subseteq \{\vec{y}_1, \dots, \vec{y}_m\}$. De esta manera tenemos que $\{\vec{x}_1, \dots, \vec{x}_n\} = \{\vec{y}_1, \dots, \vec{y}_m\}$, por lo que podemos suponer, reenumerando, que $\vec{x}_i = \vec{y}_i$ y que $n = m$. Entonces podemos escribir 8.5 como

$$(c_1 - d_1)\vec{x}_1 + \dots + (c_n - d_n)\vec{x}_n = \vec{0},$$

de donde tenemos que $c_1 = d_1, \dots, c_n = d_n$.

Es decir que cada combinación lineal de elementos de β con coeficientes distintos de 0 tiene expresión única, excepto por el orden de los sumandos. ■

Teorema 132 . *Para definir una función lineal basta definirla en una base del dominio.*

Demostración. Sea β una base del espacio vectorial FV . Si $f : \beta \rightarrow W$ es una función, en vista del lema precedente, podemos definir $\hat{f} : V \rightarrow W$ mediante la regla:

$$\hat{f}(c_1\vec{x}_1 + \dots + c_n\vec{x}_n) = c_1\hat{f}(\vec{x}_1) + \dots + c_n\hat{f}(\vec{x}_n), \vec{x}_i \in \beta, c_i \in F.$$

Que el dominio de \hat{f} es FV es consecuencia de que β genera FV . Que la definición es una buena definición es consecuencia de que β es linealmente independiente.

Notemos que $\hat{f}(\vec{0}) = \vec{0}$, ya que una suma sin sumandos es igual a $\vec{0}$, por convención (ver la página 151). En efecto,

$$\hat{f}(\vec{0}) = \hat{f}\left(\sum_{i \in \emptyset} c_i x_i\right) = \sum_{i \in \emptyset} c_i f(x_i) = \vec{0}, c_i \in F, x_i \in \beta,$$

donde las dos sumas son sumas “vacías” (sin sumandos).

Resta notar que \hat{f} es lineal, pero esto es una consecuencia inmediata de la definición, que el lector puede comprobar como un ejercicio. ■

Ejercicio 359 . *Compruebe que la función \hat{f} del teorema anterior es lineal.*

Para poder hacer el ejercicio siguiente se supone que todo subconjunto linealmente independiente de un espacio vectorial se puede completar a una base. Este resultado no se demuestra en este curso para espacios de dimensión infinita, pero se puede encontrar en casi cualquier texto de Álgebra Lineal.

Ejercicio 360 . *Suponga que $\beta \subseteq FV$ tiene la propiedad de que para cada función $f : \beta \rightarrow FW$ existe una única función lineal $\hat{f} : FV \rightarrow FW$ tal que su restricción a β es f . Demuestre que β tiene que ser una base de FV . (Sugerencia: demuestre que β tiene que ser linealmente independiente y generador de FV . Suponga, por reducción al absurdo, que β no es linealmente independiente, tome una expresión de $\vec{0}$ como combinación lineal de elementos de β con coeficientes distintos de 0, digamos que $\vec{0} = c_1\vec{x}_1 + \dots + c_n\vec{x}_n$ con $\vec{x}_i \in \beta$, tome $f : \vec{x}_i \mapsto (0, \dots, 1, \dots, 0) = e_i \in \mathbb{R}^n$, y contradiga que $\hat{f}(\vec{0}) = \vec{0}$. Con esto se tendrá que β es linealmente independiente tome ahora una base*

γ de V tal que $\beta \subseteq \gamma$. Defina $f : \gamma \rightarrow V$ por $\begin{cases} f(\vec{x}) = \vec{x} \text{ si } \vec{x} \in \beta \\ f(\vec{x}) = \vec{0} \text{ si } x \in \gamma \setminus \beta \end{cases}$, como γ es una base entonces existe $\hat{f} : V \rightarrow V$ lineal que es una extensión de f . Ahora, tanto \hat{f} como Id_V son funciones lineales **distintas** (¿por qué?) que extienden a $\beta \xrightarrow{\text{inclusión}} V$. Obtenga una contradicción.)

Ejercicio 361 . Demuestre que una función lineal f es la función cero, si y sólo si f se anula en una base de su dominio.

Ejercicio 362 . Demuestre que dos funciones lineales $f, g : V \rightarrow W$ son iguales si y sólo si f y g coinciden en una base de V .

Ejercicio 363 . ¿Cuál es la función lineal $\mathbb{R}^3 \xrightarrow{T} \mathbb{R}^3$ tal que $(1, 0, 0) \xrightarrow{T} (1, 1, 1)$, $(0, 1, 0) \xrightarrow{T} (0, 1, 2)$ y $(0, 0, 1) \xrightarrow{T} (2, -1, -2)$?

Ejercicio 364 . ¿Hay una función lineal $\mathbb{R}^3 \xrightarrow{T} \mathbb{R}^3$ tal que $(1, 0, 0) \xrightarrow{T} (1, 1, 1)$, $(0, 1, 0) \xrightarrow{T} (0, 1, 2)$ y $(0, 0, 1) \xrightarrow{T} (2, -1, -2)$ y $(1, 2, 3) \xrightarrow{T} (0, 0, 1)$?

Ejercicio 365 . Encuentre dos funciones lineales $\mathbb{R}^3 \xrightarrow{T,U} \mathbb{R}^3$ que manden $(1, 0, 0)$ a $(2, 3, 5)$ y $(1, 1, 0)$ a $(-1, 1, -1)$.

Ejercicio 366 . Demuestre que una composición de dos funciones lineales también es lineal.

8.8 La matriz de una función lineal entre $F^n \xrightarrow{T} F^m$

Consideremos la base canónica de F^n , $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ y tomemos el conjunto de sus imágenes bajo T ,

$$\{T(\vec{e}_1), T(\vec{e}_2), \dots, T(\vec{e}_n)\}$$

este es un conjunto de n vectores cada uno con m coordenadas. Formamos el arreglo rectangular con las coordenadas de los vectores en

$$\{T(\vec{e}_1), T(\vec{e}_2), \dots, T(\vec{e}_n)\}.$$

De tal manera que obtenemos un arreglo con nm entradas (coeficientes). El coeficiente en el i -ésimo renglón y en la j -ésima columna del arreglo es la i -ésima coordenada del vector $T(\vec{e}_j)$. El arreglo de nm números se pone entre paréntesis y se llama la matriz de T .

Si llamamos A a la matriz de T , entonces $A_{i,j}$ es la i -ésima coordenada de $T(\vec{e}_j)$.

En resumen, para construir la matriz $A \in M_{m \times n}(\mathbb{R})$ que corresponde a una transformación lineal $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, dadas las bases canónicas de \mathbb{R}^n y de \mathbb{R}^m , tómese para cada columna A^j de A la imagen $T(\vec{e}_j)$ de \vec{e}_j bajo la transformación T .

Así por ejemplo si $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ es la transformación definida por $T(x,y,z) = (x+2y-z, 3x-7y)$, entonces $T(1,0,0) = (1,3)$, $T(0,1,0) = (2,-7)$ y $T(0,0,1) = (-1,0)$. Por lo tanto $A = \begin{pmatrix} 1 & 2 & -1 \\ 3 & -7 & 0 \end{pmatrix}$ es la matriz asociada. Nótese ahora que para $\vec{x} \in \mathbb{R}^3$, $T(\vec{x})$ es $A\vec{x}$.

Ejemplos 179

1. La matriz de una rotación. Si consideramos

$$R_\theta : \mathbb{R}^2 \longrightarrow \mathbb{R}^2,$$

notando que

$$R_\theta \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}$$

y que

$$R_\theta \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix},$$

tenemos que

$$[R_\theta] = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix},$$

2. La matriz de una reflexión sobre una línea que pasa por el origen de \mathbb{R}^2

Si reflejamos el vector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ sobre una línea ℓ que pasa por $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ y que hace un ángulo α con el eje X , obtenemos un vector que hace un ángulo 2α con el eje X . Por lo tanto, sus coordenadas son $\begin{pmatrix} \cos(2\alpha) \\ \sin(2\alpha) \end{pmatrix}$.

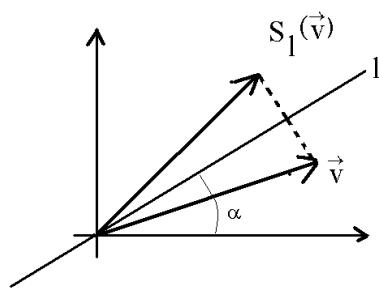


Figura 8.6:

Si reflejamos el vector $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ sobre la línea ℓ , obtenemos el vector que hace un ángulo

$$\frac{\pi}{2} - 2\left(\frac{\pi}{2} - \alpha\right) = 2\alpha - \frac{\pi}{2}$$

con el eje X . Por lo tanto, sus coordenadas son $\begin{pmatrix} \cos\left(2\alpha - \frac{\pi}{2}\right) \\ \sin\left(2\alpha - \frac{\pi}{2}\right) \end{pmatrix}$. De la siguiente figura, vemos que

$$\cos(\theta) = \sin\left(\frac{\pi}{2} - \theta\right) = -\left(\sin\left(\theta - \frac{\pi}{2}\right)\right),$$

de aquí que

$$\sin\left(2\alpha - \frac{\pi}{2}\right) = -\cos(2\alpha).$$

Además

$$\sin(\theta) = \cos\left(\frac{\pi}{2} - \theta\right) = \cos\left(\theta - \frac{\pi}{2}\right),$$

por lo que

$$\cos\left(2\alpha - \frac{\pi}{2}\right) = \sin(2\alpha).$$

3

Así que la matriz de la reflexión es

$$\begin{pmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{pmatrix}.$$

Ejercicio 367 . Calcule la matriz la proyección $p_2 : \mathbb{R}^3 \longrightarrow \mathbb{R}$, que envía un vector en \mathbb{R}^3 a su segunda coordenada.

Ejercicio 368 . Calcule la matriz de la homotecia $\mathbb{R}^3 \xrightarrow{\lambda} \mathbb{R}^3$, (multiplicar por λ).

Definición 123 . Sea $A \in M_{m,n}(\mathbb{R})$, queremos definir una función lineal

$$\mathbb{R}^n \xrightarrow{A} \mathbb{R}^m$$

³Hemos usado que $\sin(-\theta) = -\sin(\theta)$ y que $\cos(-\theta) = \cos(\theta)$. Recuerde que \cos es una función par mientras que \sin es impar.

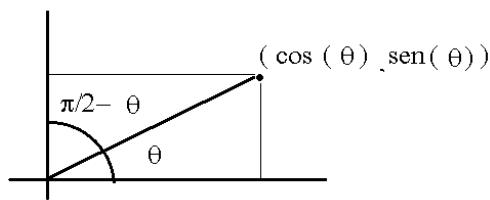


Figura 8.7:

(multiplicar por A). Quisiéramos hacer esto de tal manera que la matriz de $A \cdot \underline{}$ sea A misma. Entonces la j -ésima columna de la matriz de $A \cdot \underline{}$ debe ser

$$A^j = (A \cdot \underline{})(\vec{e}_j) = A \cdot \vec{e}_j.$$

En este caso

$$\begin{aligned} A \cdot \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} &= A \cdot (c_1 \vec{e}_1 + \dots + c_n \vec{e}_n) = c_1 (A \cdot \vec{e}_1) + \dots + c_n (A \cdot \vec{e}_n) = \\ &= c_1 A^1 + \dots + c_n A^n. \end{aligned}$$

Para tener linealidad.

Definición 124 (Producto de matrices). Sean $A \in M_{m,n}(\mathbb{R})$ y $B \in M_{n,r}(\mathbb{R})$ entonces

$$\mathbb{R}^r \xrightarrow{B \cdot \underline{}} \mathbb{R}^n, \quad \mathbb{R}^n \xrightarrow{A \cdot \underline{}} \mathbb{R}^m$$

son funciones lineales de la manera que se definió arriba. Por lo tanto su composición es una función lineal

$$\mathbb{R}^r \xrightarrow{(A \cdot \underline{}) \circ (B \cdot \underline{})} \mathbb{R}^m,$$

como tal, tiene matriz y llamamos AB a su matriz.

Observación 108 . Por definición la j -ésima columna de AB es

$$\begin{aligned} (AB)^j &= (A \cdot \underline{}) \circ (B \cdot \underline{})(\vec{e}_j) = A((B \cdot \underline{})(\vec{e}_j)) = A \cdot B^j = \\ &= B_{1,j} A^1 + \dots + B_{n,j} A^n \in \mathbb{R}^m. \end{aligned}$$

Si nos fijamos en la i -ésima coordenada de este vector obtenemos

$$\begin{aligned} (AB)_{i,j} &= B_{1,j} A_{i,1} + \dots + B_{n,j} A_{i,n} = A_{i,1} B_{1,j} + \dots + A_{i,n} B_{n,j} = \\ &= A_i \cdot B^j. \end{aligned}$$

Lo que nos puede servir como una guía para definir el producto de dos matrices.

Observación 109 . El producto de matrices es asociativo.

Demostración. Supongamos que A es una matriz de m por n , que B es una matriz de n por r y que C es una matriz de r por s .

Entonces para ver que

$$A(BC) = (AB)C,$$

basta ver que para cada i, l

$$(A(BC))_{i,l} = ((AB)C)_{i,l},$$

en efecto,

$$\begin{aligned} (A(BC))_{i,l} &= A_i \cdot (BC)^l = \sum_k A_{i,k} (BC)_{k,l} = \sum_k A_{i,k} (B_k \cdot C^l) = \\ &= \sum_k A_{i,k} \left(\sum_j B_{k,j} C_{j,l} \right) = \sum_k \left(\sum_j A_{i,k} (B_{k,j} C_{j,l}) \right) = \\ &= \sum_k \sum_j A_{i,k} (B_{k,j} C_{j,l}) = \sum_j \sum_k A_{i,k} (B_{k,j} C_{j,l}) = \\ &= \sum_j \left(\sum_k (A_{i,k} B_{k,j}) C_{j,l} \right) = \sum_j ((AB)_{i,j} C_{j,l}) \\ &= ((AB)C)_{i,l}. \end{aligned}$$

■

Observación 110 . Si $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ y $U : \mathbb{R}^m \rightarrow \mathbb{R}^s$ son funciones lineales, entonces

$$U \circ T : \mathbb{R}^n \rightarrow \mathbb{R}^s$$

es una función lineal y su matriz es el producto de las matrices respectivas, es decir

$$[U \circ T] = [U] [T]$$

Demostración. Denotemos $A = [U]$ y $B = [T]$, entonces

$$\begin{aligned} [U \circ T]^j &= [U \circ T] \vec{e}_j = U(T(e_j)) = U(B^j) = \\ &= U \left(\sum_k B_{k,j} \vec{e}_k \right) = \sum_k B_{k,j} U(\vec{e}_k) = \sum_k B_{k,j} A^k. \end{aligned}$$

La i -ésima coordenada de este vector es:

$$[U \circ T]_{i,j} = \sum_k B_{k,j} A_{i,k} \sum_k A_{i,k} B_{k,j} = A_i \cdot B^j = (AB)_{i,j}$$

Por lo tanto $[U \circ T] = AB = [U] [T]$. ■

Ejemplo 180 (*La matriz de una composición de rotaciones*). *Consideremos rotaciones en el plano \mathbb{R}^2 por ángulos θ y ϕ . Denotemos R_θ y R_ϕ dichas rotaciones entonces*

$$R_\theta \circ R_\phi = R_{\phi+\theta},$$

por lo tanto

$$[R_{\phi+\theta}] = [R_\theta] [R_\phi],$$

así que

$$\begin{pmatrix} \cos(\theta+\phi) - \sin(\theta+\phi) \\ \sin(\theta+\phi) \cos(\theta+\phi) \end{pmatrix} = \begin{pmatrix} \cos(\theta) - \sin(\theta) \\ \sin(\theta) \cos(\theta) \end{pmatrix} \begin{pmatrix} \cos(\phi) - \sin(\phi) \\ \sin(\phi) \cos(\phi) \end{pmatrix} =$$

$$= \begin{pmatrix} \cos(\theta) \cos(\phi) - \sin(\theta) \sin(\phi) & -\cos(\theta) \sin(\phi) - \sin(\theta) \cos(\phi) \\ \sin(\theta) \cos(\phi) + \cos(\theta) \sin(\phi) & \cos(\theta) \cos(\phi) - \sin(\theta) \sin(\phi) \end{pmatrix}$$

De donde tenemos

$$\cos(\theta + \phi) = \cos(\theta) \cos(\phi) - \sin(\theta) \sin(\phi)$$

y

$$\sin(\theta + \phi) = \sin(\theta) \cos(\phi) + \cos(\theta) \sin(\phi).$$

Definición 125 . *La matriz identidad de n por n es la matriz de*

$$Id : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ 0 & 0 & \dots & 1 \end{pmatrix} = I_n.$$

Definición 126 . *Se dice que una matriz cuadrada A tiene inverso se existe una matriz cuadrada B tal que*

$$AB = I_n = BA.$$

Ejemplo 181 . *Es claro que la operación de reflejar sobre una línea ℓ que pasa por el origen de \mathbb{R}^2 es autoinversa, por lo tanto*

$$\begin{aligned} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} &= [Id] = [\sigma_\ell \circ \sigma_\ell] = \\ &= \begin{pmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{pmatrix} \begin{pmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{pmatrix}, \end{aligned}$$

de aquí que

$$\cos^2(2\alpha) + \sin^2(2\alpha) = 1,$$

o bien

$$\cos^2(\beta) + \sin^2(\beta) = 1.$$

8.9 Sistemas de ecuaciones lineales

Cuando utilizamos el lenguaje matemático para modelar situaciones (que pueden referirse a la Matemáticas mismas, pero que también puede salir de la física, de la química, o de la biología, por ejemplo), se presentan con frecuencia los *Sistemas de ecuaciones lineales*, cuyas soluciones, en general, admiten interpretaciones importantes en el problema original. En estos casos, resulta de interés el tener *criterios efectivos* que permitan decidir *cuando* tales sistemas tienen solución, y en este caso, *cuantas y como* encontrarlas.

Consideremos un sistema de m ecuaciones con n incógnitas, al que etiquetaremos (*) y que deseamos a resolver:

$$\begin{array}{rcl} a_{1,1}x_1 + \dots + a_{1,n}x_1 & = & k_1 \\ \vdots & \vdots & \vdots \\ a_{1,1}x_1 + \dots + a_{m,n}x_n & = & k_n \end{array} \quad ((*))$$

Encaminados a este propósito, nuestra primera pregunta podría ser:

- ¿Qué debemos entender por *resolver un sistema*?

Aceptemos que un sistema queda resuelto cuando hemos encontrado todas sus soluciones.

Aquí surgen de inmediato las siguientes cuatro preguntas:

- ¿Qué es una solución de (*)?
- ¿Cuando un sistema tiene solución?
- ¿Cuántas soluciones hay?
- ¿Cómo podemos encontrarlas, ya sea explícitamente o bien caracterizándolas por medio de expresiones adecuadas?

En esta parte se describe una manera de responder a estas preguntas y al mismo tiempo se pretende discutir brevemente los fundamentos teóricos en los que se basa el algoritmo.

Como se verá más adelante, en el caso de los sistemas ecuaciones lineales, las preguntas 3, 4 y 5 se responde simultáneamente, lo que no es común en problemas de las Matemáticas. En efecto, la existencia de soluciones, la unicidad de estas, y la manera de encontrarlas son en general problemas independientes.

8.9.1 Algunas definiciones

Definición 127 . *Se llama sistema de ecuaciones lineales reales de $m \times n$ con coeficientes reales o bien sistema de m ecuaciones lineales con n incógnitas a una expresión (conjunto de igualdades) de la forma:*

$$\begin{array}{rcl} a_{1,1}x_1 + \dots + a_{1,n}x_n & = & k_1 \\ \vdots & & \vdots \quad \vdots \\ a_{m,1}x_1 + \dots + a_{m,n}x_n & = & k_m \end{array} \quad ((*))$$

en donde cada $a_{i,j}$ (coeficiente) y cada k_i (término independiente), son números reales. Las incógnitas (las x_j) son variables que deberán representar también números reales.

Nos referiremos a un sistema como el anterior simplemente como a un sistema de ecuaciones y sólo seremos más explícitos cuando la situación particular lo requiera.

Definición 128 . En el caso en que cada k_i sea cero, ser el sistema es homogéneo. A cada sistema $(*)$ se le puede asociar de manera natural un sistema de ecuaciones homogéneo que se conoce como **el sistema homogéneo asociado $(*_H)$** .

Definición 129 . Para cada i , la expresión

$$a_{i,1}x_1 + \dots + a_{i,n}x_n = k_i \quad ((i))$$

se llama la **i-ésima ecuación del sistema**.

Definición 130 . Una n -ada ordenada $\vec{c} = (c_1, c_2, \dots, c_n) \in \mathbb{R}^n$ es una solución de la ecuación i -ésima de $(*)$ si

$$a_{i,1}c_1 + \dots + a_{i,n}c_n = k_i ,$$

lo que también se expresa diciendo que \vec{c} satisface (i) .

Una vez convenido que los coeficientes y los términos independientes son números reales, y que las variables quedan restringidas a tomar como valores también números reales, a cada ecuación se le asocia el conjunto

$$S_i = \{(c_1, \dots, c_n) \in \mathbb{R}^n \mid a_{i,1}c_1 + \dots + a_{i,n}c_n = k_i\}$$

de sus soluciones, que en los casos en que $n \leq 3$ tiene la interpretación geométrica usual.

Así por ejemplo si la i -ésima ecuación es: $2x - y = 4$, y $n = 2$, entonces S_i resulta ser el lugar geométrico de todos los puntos del plano cuyas coordenadas satisfacen la ecuación $(*)$, que en este caso es la recta de pendiente 2 que tiene ordenada al origen -4 .

Aquí resulta conveniente considerar la situación en que para alguna ecuación dada, todos los coeficientes sean 0 y entonces distinguir los dos casos extremos siguientes:

$$1. \ k_i = 0 \text{ y}$$

$$2. \ k_i \neq 0.$$

Explícitamente, si $n = 3$, en el primer caso, la ecuación queda:

$$0x + 0y + 0z = 0,$$

cuyo conjunto de soluciones es, evidentemente, todo \mathbb{R}^3 .

Para el segundo, la situación es la siguiente:

$$0x + 0y + 0z = k, \quad k \neq 0,$$

que no es satisfactible por ninguna terna en \mathbb{R}^3 y por lo tanto su conjunto solución es vacío.

Observación 111 . *En vista de que algunos coeficientes $a_{i,j}$ pueden ser cero, y en ese caso puede omitirse la escritura del término correspondiente, debe recordarse cuál es el número de incógnitas con el que está trabajando, y que se fija al definir el sistema. Si, por ejemplo, la i -ésima ecuación de algún sistema fuera: $x = 1$, entonces las soluciones serían:*

si $n = 1$, $\{1\}$, un solo número,

si $n = 2$, $\{(1, t) \mid t \in \mathbb{R}\}$, una recta vertical;

y si $n = 3$, $\{(1, s, t) \mid s, t \in \mathbb{R}\}$, un plano paralelo al yz ,

lo que simplemente reitera el hecho de que para describir un conjunto, (en ese caso es de las soluciones de la ecuación), no basta dar una condición, (la ecuación misma), sino que es preciso señalar el conjunto del cual la condición escoge elementos (\mathbb{R}^n).

Definición 131 . $\vec{c} \in \mathbb{R}^n$ es una solución de $(*)$ si satisface cada una de sus ecuaciones, es decir, si $\vec{c} \in S_i \forall i \in \{1, \dots, m\}$. Al conjunto se le llama el conjunto solución de $(*)$ o el conjunto de todas sus soluciones, dependiendo de que se desee enfatizar: al conjunto mismo, por sus elementos. Remarcamos aquí que entonces el conjunto solución de $(*)$ es

$$S_* = \bigcap_{i=1}^m \{S_i\}.$$

En efecto, desde el punto de vista de la lógica, el sistema $(*)$ es un predicado compuesto por la conjunción de sus operaciones, que cuando se interpreta en la teoría de conjuntos, corresponde a la intersección.

Por esa razón, si se da alguno de los casos extremos analizados con anterioridad:

1. $0x_1 + 0x_2 + \dots + 0x_n = 0$,

la ecuación puede omitirse ya que el conjunto S_i , que es el universo, es el idéntico con respecto a la intersección.

2. $0x_1 + 0x_2 + \dots + 0x_n \neq 0$,

en este caso se dice el sistema no tiene solución, lo que corresponde el hecho de que S_* está contenido en cada S_i , uno de los cuales es vacío.

Cuando $S_* = \emptyset$ ((*) no tiene soluciones) se dice que el sistema es inconsistente.

Si $S_* \neq \emptyset$ (existe al menos una solución), el sistema es consistente.

Cuando S_* consta de un solo elemento, se llama determinado.

Así por ejemplo, si (*) es

1.

$$\begin{array}{rcl} x & + & y = 2 \\ x & + & y = 3 \end{array}, \text{ entonces } S_* = \emptyset.$$

2.

$$\begin{array}{rcl} x & + & y = 2 \\ x & + & y = 3 \end{array}, \text{ entonces } S_* = \{(6, 4)\}.$$

3.

$$\begin{array}{rcl} x & + & y = 10 \\ 2x & + & 2y = 20 \end{array}, \text{ entonces } S_* \text{ es infinito.}$$

(En este caso suele decirse que el sistema es indeterminado).

Observación 112 . Aunque con frecuencia se dice que “resolver una ecuación” (o resolver un sistema de ecuaciones) es encontrar todas sus soluciones, no quisimos emplear esta expresión debido a que las “soluciones” de todo el sistema (y en particular las de cada ecuación) forman un conjunto (el conjunto de sus soluciones) y como tal puede describirse o “encontrarse” de muchas maneras. ¡De hecho, los sistemas mismos son descripciones implícitas de sus soluciones!

Observación 113 . *Como se verá más adelante, siempre que S_* no sea vacío, resultará ser el trasladado de un subespacio, (una variedad lineal), y por lo tanto se podrá entender que S_* está definido en forma explícita cuando se diga que es vacío (y lo sea), o cuando no siéndolo se dé una representación*

paramétrica de él, o bien se describa en términos de base y punto de apoyo; forma a la que también llamaremos “expresión vectorial paramétrica de la solución”; es decir, cuando se dé la solución general del sistema homogéneo asociado por medio de una base, más una solución particular del sistema completo (punto de apoyo).

Consideremos los siguientes ejemplos.

Ejemplos 182

1. Las soluciones de la ecuación $x^2 - 5x + 6 = 0$ son los elementos del conjunto $\{2, 3\}$ que puede describirse igualmente como:
 - $\{x \in \mathbb{Z}^+ \mid x \text{ es primo, y } x < 5\}$ o bien como
 - $\{x \in \mathbb{R} \mid x^2 - 5x + 6 = 0\}$,
 entre muchas otras formas.
2. La solución general S_* del sistema

$$\begin{array}{rcl} x & +y & +z = 3 \\ y & +z & = 2 \end{array} \quad (*)$$

es $S_* = \{(c_1, c_2, c_3) \in \mathbb{R}^3 \mid c_1 = 1, \quad c_2 + c_3 = 2\}$ que corresponde también a las siguientes descripciones:

- $S_* = \left\{ \vec{x} \in \mathbb{R}^3 \mid \vec{x} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, t \in \mathbb{R} \right\}.$
- $S_* = \{(1, 2 - t, t) \in \mathbb{R}^3 \mid t \in \mathbb{R}\}.$
- $S_* = \{(x, y, z) \in \mathbb{R}^3 \mid x = 1, \quad y = 2 - t, \quad z = t, \quad t \in \mathbb{R}\}.$
- S_* es $x = 1, y = 2 - t, z = t, t \in \mathbb{R}.$

Que son distintas (?) versiones de la llamada *descripción paramétrica*. En este caso, suele decirse que “las incógnitas están despejadas”. De hecho lo están, pero en función de sus parámetros (la t), si esto justifica el que en muchas ocasiones se afirme que resolver un sistema de ecuaciones es “despejar sus incógnitas”.

Otras formas de describir S_* podrían ser:

3 S_* es la intersección de los planos cuyas ecuaciones son:

$$x + y + z = 3 \text{ y } y + z = 2$$

$$4 \quad S_* = \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 3 \text{ y } y + z = 2\} .$$

Definición 132 . *Resolver un sistema de ecuaciones es escribir su solución general en forma explícita.*

Lo que se trata de conseguir aquí, es una descripción de S_* efectiva, en el sentido de que permita encontrar soluciones particulares fácilmente.

8.9.2 Un método para resolver sistemas de ecuaciones lineales

Uno de los métodos que aprendimos en la escuela secundaria para resolver sistemas ecuaciones lineales es el de suma y resta que consiste esencialmente en ir eliminando en cada ecuación a todas las incógnitas menos una, de manera que se termine con un sistema de la forma:

$$\begin{aligned} x_1 &= c_1 \\ x_2 &= c_2 \\ &\vdots \\ x_n &= c_n \end{aligned} \tag{8.6}$$

cuya solución general salta la vista. (Ver la observación 111).

La eliminación de las incógnitas se hace por medio de las **operaciones elementales** que se efectúan en las ecuaciones del sistema y que son:

1. Cambiar el orden en que aparecen las ecuaciones.
2. Multiplicar cada término de una ecuación por cualquier constante no 0.
3. Sustituir una ecuación por resultado de sumarle otra.

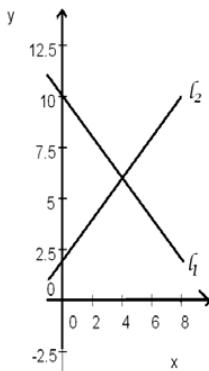
El resultado final, aunque no siempre queda de la forma 8.6, sí es totalmente satisfactorio en cuanto que contestan las últimas preguntas de la página 470.

Para fijar ideas comenzaremos suponiendo que debemos resolver el sistema

$$x + y = 10 \quad (l_1)$$

$$-x + y = 2, \quad (l_2)$$

que geométricamente corresponde encontrar a la intersección de las rectas l_1 y l_2 que se ven en la figura siguiente:



Transformamos el sistema sumando la primera ecuación a la segunda y luego dividiendo cada término de la suma así obtenidas entre 2

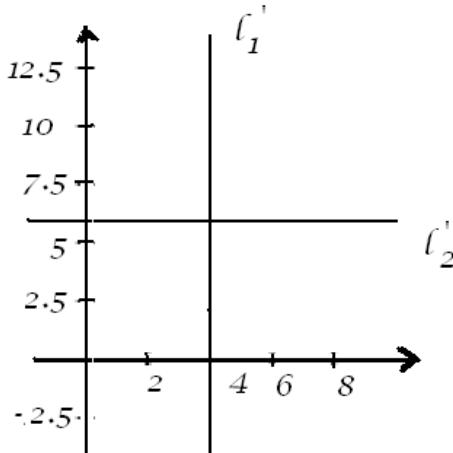
$$x + y = 10 \quad (l_1)$$

$$y = 6 \quad (l'_2)$$

Finalmente, restando la segunda ecuación de la primera, obtenemos:

$$x = 4$$

$$y = 6$$



Terminaremos diciendo que la solución del sistema es $(4, 6)$ y para estar seguros comprobamos:

$$4 + 6 = 10 \quad (8.7)$$

$$-4 + 6 = 2 \quad (8.8)$$

Seguramente lector habrá notado que en cada paso de la construcción anterior, hemos ido cambiando el problema.

En el ejemplo comenzamos con rectas inclinadas, una de las cuales 8.7 se transforma en la recta horizontal $y = 6$. Sin embargo, el conjunto solución de cada sistema es el mismo los tres casos.

Esta situación de cambio de problema, se repite cada vez que usamos el método de eliminación de las incógnitas por suma y resta. Por supuesto que tendrá que demostrarse que a pesar de estos cambios, el conjunto solución al que se llegue al concluir el proceso, es el mismo que el que nos interesa conocer desde el principio. Al final de este capítulo, daremos la justificación requerida.

Podríamos esquematizar un poco lo que hemos hecho, comenzando con una matriz en la que pondremos sólo los coeficientes y los términos independientes. Al ir cambiando el sistema por medio de las operaciones elementales

que se requieren en el proceso eliminación, la matriz que lo representa también va transformándose de igual manera, de modo que se puede representar todo el proceso utilizando exclusivamente las matrices y agregando una pequeña instrucción en cada caso como una cortesía para el lector. Así por ejemplo en el caso que nos ocupa, la situación podría ser la siguiente:

1. Comenzamos con la matriz aumentada

$$\begin{pmatrix} 1 & 1 & 10 \\ -1 & 1 & 2 \end{pmatrix} \quad (8.9)$$

que representa el sistema original.

2. Pasamos ahora a la matriz

$$\begin{pmatrix} 1 & 1 & 10 \\ 0 & 2 & 12 \end{pmatrix} \quad (8.10)$$

que describe el resultado de sustituir el segundo renglón de la matriz anterior, por la suma de los renglones, lo que abreviaremos: $\mathbf{R}_2 + \mathbf{R}_1 = \mathbf{R}'_2$.

El resto del proceso puede describirse, de manera análoga por:

$$\frac{1}{2} \mathbf{R}_2 = \mathbf{R}'_2 \quad \begin{pmatrix} 1 & 1 & 10 \\ 0 & 1 & 6 \end{pmatrix} \xrightarrow{\mathbf{R}_1 - \mathbf{R}_2 = \mathbf{R}'_1} \begin{pmatrix} 1 & 0 & 4 \\ 0 & 1 & 6 \end{pmatrix} \quad (8.11)$$

Finalmente recuperamos la información subyacente escribiendo $x = 4$, $y = 6$.

Nótese que en este caso la solución es única, y corresponde en el final del proceso esquematizado, a una matriz en la que el número de renglones diferentes de 0 (r) es igual al número de incógnitas (n); ($r = n$).

Por supuesto, un renglón cero es $(0, 0, 0)$.

Probemos ahora con el sistema

$$\begin{aligned} x + 2y + 3z &= 6 \\ 4x + 5y + 6z &= 15 \\ 7x + 8y + 9z &= 24 \end{aligned} \quad (8.12)$$

que geométricamente consiste en buscar de intersección de tres planos y que corresponde en el método esquemático a

$$\left(\begin{array}{cccc} 1 & 2 & 3 & 6 \\ 4 & 5 & 6 & 15 \\ 7 & 8 & 9 & 24 \end{array} \right) \xrightarrow{\substack{R_2 \leftarrow R_2 - 4R_1 \\ R_3 \leftarrow R_3 - 7R_1}} \left(\begin{array}{cccc} 1 & 2 & 3 & 6 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 2 & 3 \end{array} \right) \xrightarrow{\substack{R_1 \leftarrow R_1 - 2R_2 \\ R_3 \leftarrow R_3 - R_2}} \quad (8.13)$$

$$\left(\begin{array}{cccc} 1 & 2 & 3 & 6 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{array} \right) \xrightarrow{R_1 \leftarrow 2R_2} \left(\begin{array}{cccc} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{array} \right) \quad (8.14)$$

Esta última matriz representa el sistema

$$\begin{aligned} x - z &= 0 \\ y + 2z &= 3 \\ 0x + 0y + 0z &= 0 \end{aligned} \quad (8.15)$$

Hagamos las siguientes observaciones:

Observación 114 . *No es posible llevar más adelante el proceso de eliminación. Por supuesto se puede eliminar z de las ecuaciones primera y segunda, pero a costa de introducir y en la primera o x en la segunda.*

Observación 115 . *Notese que en esta situación (imposibilidad de eliminar “bien” las incógnitas) corresponde al caso en que la última matriz tiene menos renglones distintos de 0, que el número de incógnitas ($r < n$).*

Observación 116 . *La tercera ecuación dice que*

$$0x + 0y + 0z = 0 \quad (8.16)$$

y por lo tanto, como se dijo antes, se cumple para cualquier terna ordenada de números reales es decir, su solución parcial es todo \mathbb{R}^3 , así que al intersectarla con las soluciones de la primera y de la segunda ecuaciones, simplemente quedan las soluciones de éstas, por lo que la tercera ecuación se puede desechar.

Las dos primeras ecuaciones se pueden reescribir como sigue:

$$x = z \quad (8.17)$$

$$y = 3 - 2z \quad (8.18)$$

y puestas así, permiten observar que la z es *libre* de tomar cualquier valor real que se ocurra y que cada vez que se seleccione alguno para ella, la x y la y quedan *fijas*. Esto nos permite expresar la solución general de 8.12 como

$$\begin{aligned} x &= t \\ y &= 3 - 2t \quad (t \in \mathbb{R}) . \\ z &= t \end{aligned} \tag{8.19}$$

Cada valor (arbitrario) que la z tome permite encontrar una solución particular. z se denotará por “ t ”, (el parámetro), y se acostumbra entonces decir que se trata de un sistema con un grado de libertad. la solución general de 8.12 también puede describirse la manera siguiente:

$$S_* = \left\{ \vec{x} \in \mathbb{R}^3 \mid \vec{x} = \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, t \in \mathbb{R} \right\}, \tag{8.20}$$

que es la recta en \mathbb{R}^3 que pasa por $\begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix}$ y está generada por $\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$.⁴

8.9.3 Algoritmo para la solución de sistemas de ecuaciones lineales

Para resumir, concluimos con un algoritmo (receta) para resolver sistemas de ecuaciones lineales por el método matricial y damos algunos ejemplos.

Algoritmo 2 . *Para resolver sistemas de ecuaciones lineales.*

1. **Escriba el sistema cuidando de respetar el orden en el que aparecen las incógnitas en cada ecuación (x debajo de x, y debajo de y, etc.).**

(Si en alguna ecuación falta alguna incógnita, incorpórela con un coeficiente 0).

⁴Como se apuntó en el comienzo de este capítulo, la solución que un sistema de ecuaciones admite diferentes interpretaciones. En este caso, dentro de la cinemática, S_* también puede interpretarse como un modelo del movimiento (rectilíneo) de un móvil que se mueve con velocidad constante $(1, -2, 1)$ y que parte del punto $(0, 3, 0)$ ($t = 0$). En este caso, t representa el "tiempo", por supuesto.

Así por ejemplo, si el sistema es

$$\begin{array}{rcl} 3x & -2y & +z = 4 \\ y & +3z & -x = 0 \\ x & +z & = 2 \end{array},$$

se debe cambiar por

$$\begin{array}{rcl} 3x & -2y & +z = 4 \\ -x & +y & +3z = 0 \\ x & +0y & +z = 2 \end{array}.$$

2. Escriba la matriz de coeficientes aumentada con los términos independientes.

(Cuando convenga, cambie el orden de las ecuaciones. En el ejemplo:

$$\left(\begin{array}{rrrr} 3 & -2 & 1 & 4 \\ -1 & 1 & 3 & 6 \\ 1 & 0 & 1 & 2 \end{array} \right) \text{ se puede cambiar a } \left(\begin{array}{rrrr} 1 & 0 & 1 & 2 \\ -1 & 1 & 3 & 6 \\ 3 & -2 & 1 & 4 \end{array} \right).$$

3. Diagonalice por medio de operaciones elementales en los renglones (llévela a su forma escalonada reducida).

(Si al estar ejecutando este paso aparece algún renglón en el que la única componente diferente de cero es la última, que por lo tanto representa una ecuación de la forma $0x_1 + 0x_2 + \dots + 0x_n = k$, $k \neq 0$, que obviamente no puede ser satisfecha, suspenda el proceso y declare que el sistema es inconsistente (el conjunto de las soluciones es vacío)).

4. Terminado el proceso de diagonalización, identifique las incógnitas “fijas” (también llamadas “principales”) -que son las que corresponden a los pivotes que encabezan cada renglón distinto de 0 (que debe ser 1 en cada caso)- y llame “parámetros” o “incógnitas libres” a las demás y desígnelas t_1, \dots, t_r
(Si por ejemplo

$$A = \left(\begin{array}{rrrrrrr} 1 & 3 & 0 & 2 & 1 & 0 & 3 \\ 0 & 0 & 1 & 3 & -1 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 & 1 & -6 \end{array} \right),$$

y las incógnitas son x_1, \dots, x_6 , entonces los parámetros son x_2, x_4 y x_5 . Hágase $x_2 = t_1$, $x_4 = t_2$ y $x_5 = t_3$).

5. Reinterprete cada renglón de la matriz como ecuación y despeje cada incógnita fija. Escriba la “solución paramétrica”.

$$\begin{aligned}x_1 &= 3 - 3t_1 - 2t_2 - t_3 \\x_2 &= t_1 \\x_3 &= 4 - 3t_2 + t_3 \\x_4 &= t_2 \\x_5 &= t_3 \\x_6 &= -6.\end{aligned}$$

(Remarque la condición de ser parámetro, especificando que cada uno puede tomar libremente cualquier valor real: $t_1, t_2, t_3 \in \mathbb{R}$).

6. Escriba la solución en forma “vectorial paramétrica”.

En el ejemplo anterior,

$$\vec{x} = \begin{pmatrix} 3 \\ 0 \\ 4 \\ 0 \\ 0 \\ -6 \end{pmatrix} + t_1 \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} -2 \\ 0 \\ -3 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t_3 \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

Nota 6 Cuando se escribe la solución en forma vectorial paramétrica, el vector de los términos independientes, que es el que no está multiplicado por ningún parámetro, es una solución particular del sistema original (completo). Los vectores que corresponden a los parámetros, son soluciones del sistema homogéneo asociado y además, toda solución del sistema homogéneo es una combinación lineal de estos últimos. En el último ejemplo, si llamáramos

$$\vec{C}_p = \begin{pmatrix} 3 \\ 0 \\ 4 \\ 0 \\ 0 \\ -6 \end{pmatrix}, \vec{C}_1 = \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \vec{C}_2 = \begin{pmatrix} -2 \\ 0 \\ -3 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \vec{C}_3 = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix},$$

entonces \vec{C}_i resuelve $A\vec{y} = \vec{0}$ ($i \in \{1, 2, 3\}$) y si \vec{C} es una solución del sistema, entonces existen $\alpha, \beta, \gamma \in \mathbb{R}$ tales que

$$\vec{C} = \vec{C}_p + \alpha\vec{C}_1 + \beta\vec{C}_2 + \gamma\vec{C}_3,$$

y en este sentido se dice que el conjunto de soluciones del sistema completo, es un trasladado del espacio de soluciones del sistema homogéneo, una de cuyas bases es $\{\vec{C}_1, \vec{C}_2, \vec{C}_3\}$. Recuérdese que en este caso también se dice que el conjunto solución tiene tres grados de libertad, tres parámetros, o que es una **variedad lineal de “dimensión” 3**.

Ejemplos 183

1. Consideremos el sistema

$$\begin{aligned} x - 2y + 3z &= -6 \\ 2x + y - z &= 5 \\ x + 3y + z &= 6 \end{aligned}$$

cuya matriz aumentada es

$$\left(\begin{array}{cccc} 1 & -2 & 3 & -6 \\ 2 & 1 & -1 & 5 \\ 1 & 3 & 1 & 6 \end{array} \right).$$

Se efectúan las operaciones que se indican y resultan las matrices de la derecha:

$$\begin{array}{l} R'_2 = R_2 - 2R_1 \\ R'_3 = R_3 - R_1 \end{array} \quad \left(\begin{array}{cccc} 1 & -2 & 3 & 6 \\ 0 & 5 & -7 & 17 \\ 0 & 5 & -2 & 12 \end{array} \right)$$

$$\begin{array}{l} R'_1 = R_1 + \frac{2}{5}R_2 \\ R'_3 = R_3 - R_2 \end{array} \quad \left(\begin{array}{cccc} 1 & 0 & \frac{1}{5} & \frac{4}{5} \\ 0 & 5 & -7 & 17 \\ 0 & 0 & 5 & -5 \end{array} \right)$$

$$\begin{array}{l} R'_1 = R_1 - \frac{1}{5}R_3 \\ R'_3 = \frac{1}{5}R_3 \end{array} \quad \left(\begin{array}{cccc} 1 & 0 & 0 & 1 \\ 0 & 5 & -7 & 17 \\ 0 & 0 & 1 & -1 \end{array} \right)$$

$$R'_2 = R_2 \quad \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 5 & 0 & 10 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

$$R'_2 = \frac{1}{5}R_2 + 7R_3 \quad \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

Finalmente $x = 1, y = 2, z = -1$ y $S_0 = \left\{ \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \right\}$.

2.

$$\begin{aligned} x + 2y - z &= 4 \\ 2x + 4y - 2z &= 5 \\ x - 2y + z &= 1 \end{aligned}$$

La matriz asociada es

$$\begin{pmatrix} 1 & 2 & -1 & 4 \\ 2 & 4 & -2 & 5 \\ 1 & -2 & 1 & 1 \end{pmatrix}$$

Nuevamente, anotamos las operaciones realizadas a la izquierda y las matrices obtenidas a la derecha:

$$R'_2 = R_2 - 2R1 \quad \begin{pmatrix} 1 & 2 & -1 & 4 \\ 0 & 0 & 0 & -3 \\ 1 & -2 & 1 & 1 \end{pmatrix}.$$

El proceso termina. El sistema es **inconsistente**.

3.

$$\begin{aligned} x + 2y + 3z &= 6 \\ 4x + 5y + 6z &= 15 \\ 7x + 8y + 9z &= 24 \end{aligned}$$

que tiene la matriz asociada

$$\begin{pmatrix} 1 & 2 & 3 & 6 \\ 4 & 5 & 6 & 15 \\ 7 & 8 & 9 & 24 \end{pmatrix}.$$

$$\begin{aligned}
 R'_2 &= R_2 - 4R_1 & \begin{pmatrix} 1 & 2 & 3 & 6 \\ 0 & -3 & -6 & -9 \\ 0 & -6 & -12 & -18 \end{pmatrix} \\
 R'_3 &= R_3 - 7R_1 & \\
 R'_2 &= -\frac{1}{3}R_2 & \begin{pmatrix} 1 & 2 & 3 & 6 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 2 & 3 \end{pmatrix} \\
 R'_3 &= -\frac{1}{6}R_3 & \\
 R'_1 &= R_1 - 2R_2 & \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\
 R'_3 &= R_3 - R_2 &
 \end{aligned}$$

Las incógnitas fijas son x, y el parámetro es $z = t$.

Entonces:

$$\begin{aligned}
 x &= t \\
 y &= 3 - 2t \quad \text{o bien, } \vec{x} = \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, t \in \mathbb{R} \\
 z &= t
 \end{aligned}$$

Para cada valor de t , se obtiene una solución particular, así si:

$$t = 0, \vec{x} = \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix}; \text{ si } t = 1, \vec{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \dots \text{ etc.}$$

$$\begin{aligned}
 3x_1 + 9x_2 + x_3 + x_4 + 3x_5 + 2x_6 &= 7 \\
 2x_1 + 6x_2 + x_3 - x_4 + 3x_5 &= 2 \\
 x_1 + 3x_2 + 2x_4 + 5x_5 - 2x_6 &= 3 \\
 x_1 + 3x_2 + 2x_4 + x_5 + 2x_6 &= 1.
 \end{aligned}$$

Entonces la matriz asociada es

$$\begin{pmatrix} 3 & 9 & 1 & 1 & 3 & 2 & 7 \\ 2 & 6 & 1 & -1 & 3 & 0 & 2 \\ 1 & 3 & 0 & 2 & 5 & -2 & 3 \\ 1 & 3 & 0 & 2 & 1 & 2 & 1 \end{pmatrix}.$$

$$\begin{array}{l}
 R'_1 = R_4 \\
 \xrightarrow{R'_4 = R_1} \begin{pmatrix} 1 & 3 & 0 & 2 & 1 & 2 & 1 \\ 2 & 6 & 1 & -1 & 3 & 0 & 2 \\ 1 & 3 & 0 & 2 & 5 & -2 & 3 \\ 3 & 9 & 1 & 1 & 3 & 2 & 7 \end{pmatrix}
 \end{array}$$

$$R'_2 = R_2 - 2R_1$$

$$R'_3 = R_3 - R_1$$

$$\begin{array}{l} R'_4 = R_4 - 3R_1 \\ \hline \end{array} \quad \left(\begin{array}{ccccccc} 1 & 3 & 0 & 2 & 1 & 2 & 1 \\ 0 & 0 & 1 & -5 & 1 & -4 & 0 \\ 0 & 0 & 0 & 0 & 4 & -4 & 2 \\ 0 & 0 & 1 & -5 & 0 & -4 & 1 \end{array} \right)$$

$$R'_3 = \frac{1}{4}R_4$$

$$\begin{array}{l} R'_4 = R_4 - R_2 \\ \hline \end{array}$$

$$\left(\begin{array}{ccccccc} 1 & 3 & 0 & 2 & 1 & 2 & 1 \\ 0 & 0 & 1 & -5 & 1 & -4 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 \end{array} \right)$$

$$R'_1 = R_1 + R_4$$

$$R'_2 = R_2 + R_4$$

$$R'_3 = R_3 + R_4$$

$$\begin{array}{l} R'_4 = -R_4 \\ \hline \end{array}$$

$$\left(\begin{array}{ccccccc} 1 & 3 & 0 & 2 & 0 & 2 & 2 \\ 0 & 0 & 1 & -5 & 0 & -4 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 & \frac{3}{2} \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 \end{array} \right)$$

$$\begin{array}{l} R'_3 = R_4 \\ \hline \end{array} \quad \cdot \quad \left(\begin{array}{ccccccc} 1 & 3 & 0 & 2 & 0 & 2 & 2 \\ 0 & 0 & 1 & -5 & 0 & -4 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -\frac{3}{2} \end{array} \right)$$

$$\begin{array}{l} R'_1 = R_1 - 2R_4 \\ \hline \end{array} \quad \left(\begin{array}{ccccccc} 1 & 3 & 0 & 2 & 0 & 0 & 5 \\ 0 & 0 & 1 & -5 & 0 & 0 & -5 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -\frac{3}{2} \end{array} \right).$$

Por lo tanto $x_2 = t_1$; $x_4 = t_2$, y así,

$$\begin{array}{l} x_1 = 5 - 3t_1 - 2t_2 \\ x_2 = t_1 \\ x_3 = -5 + 5t_2 \\ x_4 = t_2 \\ x_5 = -1 \\ x_6 = -\frac{3}{2} \end{array} \quad \text{o bien } \vec{x} = \left(\begin{array}{c} 5 \\ 0 \\ -5 \\ 0 \\ -1 \\ -\frac{3}{2} \end{array} \right) + t_1 \left(\begin{array}{c} -3 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{array} \right) + t_2 \left(\begin{array}{c} -2 \\ 0 \\ 5 \\ 1 \\ 0 \\ 0 \end{array} \right), \quad t_1, t_2 \in \mathbb{R}.$$

En los siguientes ejercicios, resuelva los sistemas cuyas matrices están dadas.

Ejercicio 369 . $\begin{pmatrix} 1 & -1 & 2 & 2 \\ 2 & 1 & -1 & 2 \\ 3 & 0 & 2 & 5 \end{pmatrix}; n = 3.$

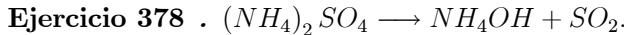
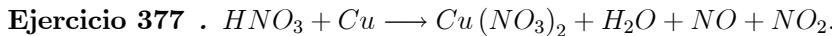
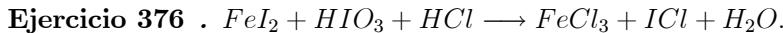
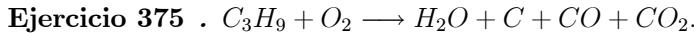
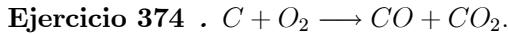
Ejercicio 370 . $\begin{pmatrix} 1 & -2 & 3 & -4 & 0 \\ -2 & 5 & 1 & 2 & 1 \\ -3 & 2 & 5 & 1 & 3 \\ -1 & -3 & 4 & -2 & 1 \end{pmatrix}; n = 4.$

Ejercicio 371 . $\begin{pmatrix} 1 & -1 & 2 & 0 & 4 & 0 & 0 & 3 \\ 0 & 0 & 0 & 1 & 2 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -7 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}; n = 7.$

Ejercicio 372 $\begin{pmatrix} 1 & 2 & -1 & 6 & 5 & 13 \\ 1 & 3 & 1 & -4 & 2 & 2 \\ 1 & 2 & 1 & -4 & 1 & 1 \end{pmatrix}; n = 5.$

Ejercicio 373 $\begin{pmatrix} 3 & 2 & 0 & -2 & 3 \\ 2 & -1 & 1 & 6 & 7 \\ 5 & 1 & 1 & 4 & 9 \end{pmatrix}; n = 4.$

En los siguientes ejercicios balancee las siguientes ecuaciones que representan reacciones químicas:



Ejercicio 380 . $Zn + HNO_3 \longrightarrow Zn(NO_3)_2 + H_2 + NH_4NO_3 + Zn(NH_3)_2 + NO + NO_2$.

Ejercicio 381 . Un móvil parte de $(-3, 5, -3)$ con una velocidad $(2, -1, 1)$ y al mismo tiempo otro sale de $(-5, -3, -1)$ con una velocidad $(3, 3, 0)$. ¿Se cruzan sus trayectorias? ¿Chocan?

Ejercicio 382 . Igual que el anterior pero el segundo móvil con velocidad $(1, 1, 0)$.

Ejercicio 383 . En un espectáculo los hombres pagan \$2.00, las mujeres \$5.00 y los niños \$0.10. Si hay 100 personas y la recaudación fue de \$100.00 ¿Cuántos espectadores hay de cada clase?

8.10 Matrices reducidas y escalonadas

Definición 133 . Si (c_1, \dots, c_n) es un vector distinto de $\vec{0}$, diremos que su coeficiente principal (o pivote) es su primer coeficiente distinto de 0: si c_i es el coeficiente principal de (c_1, \dots, c_n) entonces

$$i = \min \{j \mid c_j \neq 0\}.$$

Definición 134 . Una matriz A con m renglones y n columnas está reducida y escalonada si

1. El coeficiente principal de cada renglón distinto de $\vec{0}$ es 1.
2. Si el coeficiente principal del i -ésimo renglón es $c_{i,j} (= 1)$ entonces la j -ésima columna de la matriz es \vec{e}_j .
3. Si A_i, A_j son renglones distintos de $\vec{0}$ y $i < j$ entonces el coeficiente principal de A_j está a la derecha del coeficiente principal de A_i .
4. Los renglones de puros ceros van debajo de los renglones $\neq \vec{0}$.

Definición 135 . Una matriz A con m renglones y n columnas está escalonada si satisface 2) y 4) de la definición anterior.

1. $\begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ está reducida y escalonada.

2. $\begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ está escalonada pero no reducida.

3. $\begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ está escalonada pero no reducida.

4. $\begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ no está escalonada.

5. $\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ no es reducida y escalonada.

6. $\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ está reducida y escalonada.

Definición 136 . Una operación elemental de renglón es una operación de los siguientes tres tipos:

1. Intercambiar dos renglones de una matriz, que es una operación del primer tipo.
2. Multiplicar un renglón de una matriz por un escalar $\neq 0$, que es una operación del segundo tipo.
3. Sumar a un renglón un múltiplo de otro, que es una operación elemental del tercer tipo.

Teorema 133 . Toda matriz se puede reducir y escalonar mediante un número finito de operaciones elementales de renglón.

Demostración. Notemos primero que una matriz de ceros ya está reducida y escalonada. Demostraremos la afirmación para matrices A distintas de 0, por inducción sobre el número de renglones de la matriz.

Base. Si la matriz sólo tiene un renglón, entonces

$$A = (A_{1,1}, \dots, A_{1,n}) \neq (0, \dots, 0).$$

Supongamos que el coeficiente principal es $A_{1,j}$. Si este coeficiente principal es 1 entonces la matriz ya está reducida y escalonada, si no, multiplicando por $\frac{1}{A_{1,j}}$ obtenemos una matriz reducida y escalonada.

Paso Inductivo. Supongamos que la matriz A tiene m renglones, con $m > 1$, y la afirmación cierta para matrices con $m - 1$ renglones.

Entre los renglones de la matriz, escoja un renglón A_i que tenga su coeficiente principal lo más a la izquierda que sea posible (si $A_{i,j}$ es el coeficiente principal del renglón i -ésimo y $A_{k,l}$ es el coeficiente principal del renglón k -ésimo, entonces $j \leq l$).

Multipliquemos el i -ésimo renglón por $\frac{1}{A_{i,j}}$ para obtener una matriz cuyo renglón i -ésimo tiene coeficiente principal 1.

Intercambiemos los renglones 1 e i , si $i \neq 1$.

Hemos hecho a lo más 2 operaciones elementales y tenemos una matriz B de la forma

$$\begin{pmatrix} 0 & \dots & 0 & 1 & B_{1,j+1} & \dots & B_{1,n} \\ 0 & \dots & 0 & B_{2,j} & B_{2,j+1} & \dots & B_{2,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & B_{m,j} & B_{m,j+1} & \dots & B_{m,n} \end{pmatrix}.$$

Si al segundo renglón le sumamos $-(B_{2,j})$ veces el primer renglón, al tercer renglón le sumamos $-(B_{3,j})$ veces el primer renglón,..., al m -ésimo renglón le sumamos $-(B_{m,j})$ veces el primer renglón obtenemos una matriz

$$\begin{pmatrix} 0 & \dots & 0 & 1 & B_{1,j+1} & \dots & B_{1,n} \\ 0 & \dots & 0 & 0 & B_{2,j+1} & \dots & B_{2,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & B_{m,j+1} & \dots & B_{m,n} \end{pmatrix}. \quad (8.21)$$

Denotemos $C = \begin{pmatrix} B_{2,j+1} & \dots & B_{2,n} \\ \vdots & & \\ B_{m,j+1} & \dots & B_{m,n} \end{pmatrix}$, como esta matriz tiene $m - 1$ ren-

glones, se puede reducir y escalar por medio de un número finito de operaciones elementales renglón, obteniéndose una matriz α . Aplicando las operaciones correspondientes a la matriz en 8.21, obtenemos una matriz

$$G = \begin{pmatrix} 0 & \dots & 0 & 1 & B_{1,j+1} & \dots & B_{1,n} \\ 0 & \dots & 0 & 0 & \alpha_{1,1} & \dots & \alpha_{2,n-j} \\ \vdots & & \vdots & \vdots & \vdots & & \\ 0 & \dots & 0 & 0 & \alpha_{m-1,1} & \dots & \alpha_{m-1,n-j} \end{pmatrix}$$

que está escalonadas, y donde la submatriz $\begin{pmatrix} \alpha_{1,1} & \dots & \alpha_{2,n-j} \\ \vdots & & \\ \alpha_{m-1,1} & \dots & \alpha_{m-1,1} \end{pmatrix}$ está

reducida y escalonada. Si G no estuviera reducida sería porque la columna

de un coeficiente principal de α es $\begin{pmatrix} B_{1,k} \\ 0 \\ \vdots \\ 1 = \alpha_{l,k-j} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$. Si éste fuera

el caso, sumando $-(B_{1,k})$ veces el renglón de $\alpha_{l,k-j}$ al primer renglón, y haciendo lo correspondiente para cada coeficiente principal de α , obtenemos una matriz reducida y escalonada. ■

Ejercicio 384 . Demuestre el teorema anterior por inducción sobre el número de columnas.

Ejemplo 184 . Para la matriz

$$\begin{pmatrix} 0 & 0 & 1 & 2 & 3 \\ 0 & 2 & 0 & -1 & 2 \\ 3 & 0 & 6 & 0 & -3 \end{pmatrix}$$

el renglón que tiene su coeficiente principal más a la izquierda es el tercero.

Multiplicando este renglón por $\frac{1}{3}$, obtenemos

$$\begin{pmatrix} 0 & 0 & 1 & 2 & 3 \\ 0 & 2 & 0 & -1 & 2 \\ 1 & 0 & 2 & 0 & -1 \end{pmatrix}.$$

Intercambiando los renglones primero y tercero obtenemos

$$\begin{pmatrix} 1 & 0 & 2 & 0 & -1 \\ 0 & 2 & 0 & -1 & 2 \\ 0 & 0 & 1 & 2 & 3 \end{pmatrix}.$$

Ahora tenemos que obtener una matriz reducida y escalonada a partir de la submatriz

$$\begin{pmatrix} 2 & 0 & -1 & 2 \\ 0 & 1 & 2 & 3 \end{pmatrix}.$$

El proceso es

$$\begin{pmatrix} 2 & 0 & -1 & 2 \\ 0 & 1 & 2 & 3 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & \frac{-1}{2} & 1 \\ 0 & 1 & 2 & 3 \end{pmatrix}.$$

Así

$$\begin{pmatrix} 1 & 0 & 2 & 0 & -1 \\ 0 & 2 & 0 & -1 & 2 \\ 0 & 0 & 1 & 2 & 3 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & \mathbf{2} & 0 & -1 \\ 0 & 1 & \mathbf{0} & \frac{-1}{2} & 1 \\ 0 & 0 & \mathbf{1} & 2 & 3 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & -4 & -7 \\ 0 & 1 & 0 & \frac{-1}{2} & 1 \\ 0 & 0 & 1 & 2 & 3 \end{pmatrix}.$$

Observación 117 . Recordemos que un sistema de m ecuaciones con n incógnitas se puede representar como

$$A\vec{x} = \vec{b},$$

donde A es la matriz de coeficientes del sistema, es decir la ecuación anterior es una forma sucinta de escribir el sistema de ecuaciones

$$\begin{aligned} A_{11}x_1 + \dots + A_{1n}x_n &= b_1 \\ A_{21}x_1 + \dots + A_{2n}x_n &= b_2 \\ &\vdots \\ A_{m1}x_1 + \dots + A_{mn}x_n &= b_m. \end{aligned}$$

Definición 137 . Si a la matriz del sistema le agregamos al final la columna

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \text{ obtenemos la matriz aumentada del sistema de ecuaciones } (A \mid \vec{b}).$$

Observación 118 . *Podemos hacer operaciones elementales en un sistema de ecuaciones, haciendo las operaciones a la matriz aumentada del sistema. Por ejemplo, intercambiar dos renglones de la matriz aumentada, equivale a intercambiar las ecuaciones correspondientes del sistema.*

Observación 119 . *Consideremos el sistema de m ecuaciones con n incógnitas*

$$A\vec{x} = \vec{0}.$$

Recordemos que las soluciones de este sistema son los elementos

$$\{A_1, \dots, A_m\}^\perp,$$

donde A_i denota el i -ésimo renglón de la matriz A . (Ver el teorema 131)

Observación 120 . *Las operaciones elementales no alteran las soluciones del sistema. $A\vec{x} = \vec{0}$.*

Demostración. 1) Intercambiar dos renglones no altera las soluciones pues es claro que

$$\{A_1, \dots, A_m\}^\perp$$

no cambia si se cambia el orden en que se enlistan los elementos.

2) Notemos que si c es una escalar distinto de 0, entonces

$$\vec{w} \perp A_i \iff \vec{w} \perp cA_i$$

$(\vec{w} \cdot A_i = 0 \iff \vec{w} \cdot cA_i = c(\vec{w} \cdot A_i) = 0)$. Por lo tanto multiplicar un renglón por un escalar distinto de 0, no altera las soluciones.

3) Notemos que

$$\{A_i, A_j\}^\perp = \{A_i, cA_i + A_j\}^\perp :$$

es claro que

$$\vec{w} \in \{A_i, A_j\}^\perp \iff \vec{w} \perp A_i, \vec{w} \perp A_j \text{ y } \therefore \vec{w} \perp (cA_i + A_j).$$

Con el mismo argumento,

$$\vec{w} \in \{A_i, (cA_i + A_j)\}^\perp \implies \vec{w} \perp ((-c)A_i + (cA_i + A_j)) \implies \vec{w} \perp A_j.$$

Por lo tanto una operación elemental del tercer tipo no cambia las soluciones.

■

Teorema 134 . *Las operaciones elementales de renglón, no alteran el rango de una matriz.*

Demostración. Hemos visto que para una matriz A con m renglones y n columnas se tienen que

$$\text{rango}(A) + \dim(S_0) = n,$$

donde S_0 es el conjunto de soluciones de $A\vec{x} = \vec{0}$. Si aplicáramos una operación elemental \mathcal{R} , obtendríamos una nueva matriz $\mathcal{R}(A)$ pero el sistema

$$\mathcal{R}(A)\vec{x} = \vec{0}$$

tendría las mismas soluciones, como vimos en la observación anterior. Entonces

$$\text{rango}(\mathcal{R}(A)) + \dim(S_0) = n.$$

Por lo que $\text{rango}(A) = \text{rango}(\mathcal{R}(A))$. ■

Teorema 135 (De existencia de soluciones de un sistema de ecuaciones). *El sistema de m ecuaciones con n incógnitas*

$$A\vec{x} = \vec{b}$$

(A una matriz de m renglones con n columnas y coeficientes en un campo) tiene solución si y sólo si

$$\text{rango}(A) = \text{rango}\left(A \mid \vec{b}\right).$$

Demostración. Como ya hemos notado antes, el sistema $A\vec{x} = \vec{b}$ se puede reescribir como:

$$x_1A^1 + \dots + x_nA^n = \vec{b}, \quad (A^j \text{ es la } j\text{-ésima columna de } A)$$

así que el sistema tiene solución si y sólo si \vec{b} es una combinación lineal de las columnas de A . Esto último sucede si y sólo si el subespacio de \mathbb{R}^m generado por

$$\{A^1, \dots, A^n\}$$

coincide con el espacio generado por $\{A^1, \dots, A^n, \vec{b}\}$. Como el rango de una matriz es también la dimensión del espacio generado por sus columnas entonces nuestra condición necesaria y suficiente para que $A\vec{x} = \vec{b}$ tenga solución es que

$$\begin{aligned} rango(A) &= \\ &= \dim(\mathcal{S}(\{A^1, \dots, A^n\})) = \dim(\mathcal{S}(\{A^1, \dots, A^n, \vec{b}\})) = \\ &= rango(A \mid \vec{b}). \end{aligned}$$

■

Teorema 136 . Consideremos el sistema $A\vec{x} = \vec{b}$ como en el teorema anterior, denotemos por S el conjunto de sus soluciones, y denotemos S_0 el conjunto de soluciones de $A\vec{x} = \vec{0}$ entonces

$$S = \vec{u} + S_0 = \{\vec{u} + \vec{s} \mid \vec{s} \in S_0\},$$

donde \vec{u} es una solución de $A\vec{x} = \vec{b}$.

Demuestra&on. Si $\vec{s} = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix} \in S_0$ entonces

$$s_1A^1 + \dots + s_nA^n = \vec{0},$$

como además

$$u_1A^1 + \dots + u_nA^n = \vec{b},$$

donde $\vec{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$ entonces

$$\begin{aligned} \vec{b} &= \vec{0} + \vec{b} = (s_1A^1 + \dots + s_nA^n) + (u_1A^1 + \dots + u_nA^n) = \\ &= (s_1 + u_1)A^1 + \dots + (s_n + u_n)A^n. \end{aligned}$$

Por lo que $\begin{pmatrix} s_1 + u_1 \\ s_2 + u_2 \\ \vdots \\ s_n + u_n \end{pmatrix} = \vec{s} + \vec{u} \in S$.

Por lo tanto,

$$\vec{u} + S_0 \subseteq S.$$

Recíprocamente, si $\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \vec{w} \in S$ entonces $\vec{w} = \vec{u} + (\vec{w} - \vec{u})$. como queremos ver que $\vec{w} \in \vec{u} + S_0$, basta demostrar que $\vec{w} - \vec{u} \in S_0$ pero

$$u_1 A^1 + \dots + u_n A^n = \vec{b},$$

y también

$$w_1 A^1 + \dots + w_n A^n = \vec{b},$$

así que

$$\begin{aligned} \vec{0} &= \vec{b} - \vec{b} = (w_1 A^1 + \dots + w_n A^n) - (u_1 A^1 + \dots + u_n A^n) = \\ &= (w_1 - u_1) A^1 + \dots + (w_n - u_n) A^n. \end{aligned}$$

Es decir, $\vec{w} - \vec{u} = \begin{pmatrix} w_1 - u_1 \\ w_n - u_n \end{pmatrix} \in S_0$. ■

En resumen, para resolver un sistema

$$A\vec{x} = \vec{b},$$

hay que saber hacer dos cosas:

1. Reducir y escalar matrices.
2. Resolver sistemas

$$A\vec{x} = \vec{b},$$

cuya matriz aumentada ya esté reducida y escalarizada.

La demostración de los siguientes teoremas se deja al lector como un ejercicio.

Teorema 137 . $A \cdot \underline{\quad}$ es inyectiva \Leftrightarrow las columnas de A son linealmente independientes $\Leftrightarrow A\vec{x} = \vec{0}$ tiene únicamente la solución $\vec{0}$.

Ejercicio 385 . Demuestre el teorema anterior.

Teorema 138 . $F^n \xrightarrow{A} F^m$ es suprayectiva \Leftrightarrow las columnas de A generan $F^m \Leftrightarrow A\vec{x} = \vec{b}$ tiene solución para cada $\vec{b} \in F^m$.

Ejercicio 386 . Demuestre el teorema anterior.

Teorema 139 . $F^n \xrightarrow{A} F^m$ es biyectiva \Leftrightarrow las columnas de A son una base para $F^m \Leftrightarrow A\vec{x} = \vec{b}$ tiene solución única para cada $\vec{b} \in F^m$. (Note que en este caso, $n =$ número de columnas de $A = \dim(F^m) = m$, por lo que la matriz A es cuadrada).

Ejercicio 387 . Demuestre el teorema anterior.

8.11 Determinantes

8.11.1 Notaciones para permutaciones

Una permutación σ de un conjunto finito X es una función biyectiva $X \xrightarrow{\sigma} X$. Denotamos con S_X al conjunto de todas las permutaciones de X . Como ya notamos en el capítulo 5, página 298, si X tiene n elementos, entonces S_X tiene $n!$ elementos. Cuando $X = \{1, \dots, n\}$ escribiremos S_n en lugar de $S_{\{1, \dots, n\}}$.

Para denotar una permutación en S_n podríamos usar la notación

$$\alpha = \begin{bmatrix} 1 & 2 & \dots & n \\ \alpha(1) & \alpha(2) & \dots & \alpha(n) \end{bmatrix}$$

que incluye toda la información acerca de α . Sin embargo sería mucho más económico escribir simplemente

$$\alpha = [\alpha(1), \alpha(2), \dots, \alpha(n)]$$

(nótese el uso de corchetes).

Según lo anterior, $[2, 4, 3, 6, 1, 5]$ denota la permutación en S_6 tal que $1 \mapsto 2, 2 \mapsto 4, 3 \mapsto 3, 4 \mapsto 6, 5 \mapsto 1, 6 \mapsto 5$.

Definición 138

1. Una transposición $\tau \in S_n$ es una permutación que intercambia los elementos $i, j \in \{1, 2, \dots, n\}$ y que deja fijos todos los demás elementos. Explícitamente, $\tau(i) = j, \tau(j) = i$ y $\tau(k) = k$ si $k \notin \{i, j\}$. Denotaremos $\tau = (i, j)$ (note el uso de paréntesis, en lugar de corchetes).

2. Llamaremos *inversión* a una transposición que intercambia i con $i + 1$, $i \in \{1, \dots, n - 1\}$.

Por ejemplo $[2, 1, 3, 4, 5]$ es una inversión en S_5 , ya que es la transposición que intercambia 1 y 2.

Observación 121 . *Toda transposición coincide con su inversa.*

Lema 27 . *Toda permutación es una composición de inversiones.*

Demostración. Consideremos la composición $\alpha \circ \tau$ de

$$\alpha = [\alpha(1), \alpha(2), \dots, \alpha(n)]$$

con la inversión

$$\tau = [1, 2, \dots, i + 1, i, \dots, n]$$

intercambia i con $i + 1$.

Tomando en cuenta que se aplica primero τ y después α , tenemos que

$$1 \xrightarrow{\alpha \circ \tau} \alpha(1), \dots, i \xrightarrow{\alpha \circ \tau} \alpha(i + 1), i + 1 \xrightarrow{\alpha \circ \tau} \alpha(i), \dots, n \xrightarrow{\alpha \circ \tau} n,$$

es decir que

$$\alpha \circ \tau = [\alpha(1), \alpha(2), \dots, \alpha(i + 1), \alpha(i), \dots, \alpha(n)].$$

Donde se han intercambiado $\alpha(i + 1)$ con $\alpha(i)$ respecto de la notación para α .

Notemos ahora, que si $\alpha(1) \neq 1$, es decir si $1 = \alpha(k)$ con $k > 1$, entonces

$$\alpha_1 = \alpha \circ (k - 1, k) \circ \dots \circ (1, 2) = [1, \dots]$$

es una permutación que manda 1 en 1. Si α_1 no fijara 2 entonces (pre)componiendo con algunas cuantas inversiones, podemos obtener una permutación α_2 que manda 2 en 2. Repitiendo el argumento, es claro que podemos encontrar inversiones $\gamma_1, \dots, \gamma_m$ tales que

$$\alpha \circ \gamma_1 \circ \dots \circ \gamma_m = [1, 2, \dots, n] = Id_{\{1, 2, \dots, n\}}.$$

entonces debe ser claro que

$$\alpha = Id_{\{1, 2, \dots, n\}} \circ \gamma_m \circ \dots \circ \gamma_1 = \gamma_m \circ \dots \circ \gamma_1.$$

ya que $\gamma_l = \gamma_l^{-1}$. ■

Teorema 140 . *Toda permutación se puede expresar como un producto de transposiciones.*

Demostración. Es inmediato del Lema anterior. ■

Definición 139 . *Sean $i_1, i_2, \dots, i_m \in \{1, \dots, n\}$ diremos que la permutación α tal que*

$i_1 \mapsto i_2 \mapsto \dots \mapsto i_m \mapsto i_1$ y tal que $\alpha(k) = k, \forall k \in \{1, \dots, n\} \setminus \{i_1, i_2, \dots, i_m\}$.

es un ciclo y la denotaremos $\alpha = (i_1, i_2, \dots, i_m)$ (Note el uso de paréntesis).

Notación 18 . *Si $\alpha \in S_n$, denotaremos*

$$\text{Mov}(\alpha) = \{i \mid \alpha(i) \neq i\} \text{ y } \text{Fij}(\alpha) = \{i \mid \alpha(i) = i\}.$$

Notemos que $\text{Fij}(\alpha) = \{1, \dots, n\} \setminus \text{Mov}(\alpha)$.

Notación 19 . *Por comodidad, denotaremos (1) a la permutación identidad.*

Definición 140 . *Sean α y $\beta \in S_n$, diremos que α es ajena con β si $\text{Mov}(\alpha) \subseteq \text{Fij}(\beta)$.*

Notemos que la definición anterior es simétrica pues: α es ajena con β es equivalente a $i \in \text{Mov}(\alpha) \implies i \in \text{Fij}(\beta)$, cuya contrapuesta es $i \in \text{Mov}(\beta) \implies i \in \text{Fij}(\alpha)$, es decir, β es ajena con α .

Observación 122 . *Si α es ajena con β , entonces $\alpha \circ \beta = \beta \circ \alpha$.*

Demostración. Comprobaremos que $(\alpha \circ \beta)(i) = (\beta \circ \alpha)(i)$, $\forall i \in \{1, \dots, n\}$.

1) Si $i \in \text{Mov}(\alpha)$, entonces $i \in \text{Fij}(\beta)$ así que

$$(\alpha \circ \beta)(i) = \alpha(\beta(i)) = \alpha(i),$$

Notemos ahora que en este caso $\alpha(i) \in \text{Mov}(\alpha)$, (ya que no puede pasar que $i \mapsto \alpha(i)$ y $\alpha(i) \mapsto \alpha(i)$, dado que α es inyectiva y $\alpha(i) \neq i$). Entonces $\alpha(i) \in \text{Fij}(\beta)$, por lo que

$$(\beta \circ \alpha)(i) = \beta(\alpha(i)) = \alpha(i).$$

- 2) El caso $i \in \text{Mov}(\beta)$, es simétrico del anterior.
 3) Si $i \in \text{Fij}(\alpha) \cap \text{Fij}(\beta)$, entonces

$$(\alpha \circ \beta)(i) = \alpha(\beta(i)) = \alpha(i) = i = \beta(i) = \beta(\alpha(i)) = (\beta \circ \alpha)(i).$$

■

Teorema 141 . *Toda permutación α es un producto de ciclos ajenos dos a dos, de manera única excepto por el orden de los factores.*

Demostración. Existencia)

Por inducción sobre $|\text{Mov}(\alpha)|$.

Base.

Si $|\text{Mov}(\alpha)| = 0$ entonces α es la permutación identidad y la única manera de expresarla como producto de ciclos, es como el producto vacío.

Supongamos ahora que $|\text{Mov}(\alpha)| > 0$ y que lo que se afirma vale para permutaciones que mueven menos elementos que los que mueve α . Supongamos que $\alpha(i) \neq i$, hagamos $i_2 = \alpha(i)$, $i_3 = \alpha\alpha(i)$, etc. Tomemos $k = \text{menor } \{k \mid i_k = i\}$, que existe ya que en la lista

$$i, i_2, i_3, \dots$$

hay a lo más n elementos distintos (note que $i_k = i_j$ con $k > j \implies \alpha^{k-1}(i) = \alpha^{j-1}(i) \implies \alpha^{k-j}(i) = i$), entonces

$$i, i_2, i_3, \dots, i_k$$

son elementos distintos y el ciclo

$$\gamma = (i, i_2, i_3, \dots, i_k)$$

coincide con α en los elementos i, i_2, i_3, \dots, i_k .

Consideremos ahora $\gamma^{-1} \circ \alpha \in S_n$.

Notemos que si $j \in \text{Mov}(\gamma^{-1} \circ \alpha)$, entonces

$$(\gamma^{-1} \circ \alpha)(j) \neq j$$

por lo que $j \in \text{Mov}(\alpha)$ o $j \in \text{Mov}(\gamma^{-1})$. Si $j \in \text{Mov}(\gamma^{-1})$ entonces $j \in \{i, i_2, i_3, \dots, i_k\}$, pero entonces $(\gamma^{-1} \circ \alpha)(j) = (\gamma^{-1} \circ (\gamma^{-1} \circ \gamma)(j)) = j$, contradicción.

Por lo tanto $\text{Mov}(\gamma^{-1} \circ \alpha) \subseteq \text{Mov}(\alpha)$. Además

$$j \in \{i, i_2, i_3, \dots, i_k\} \implies j \in \text{Mov}(\alpha) \setminus \text{Mov}(\gamma^{-1} \circ \alpha),$$

pues como vimos para una j así tenemos que $(\gamma^{-1} \circ \alpha)(j) = (\gamma^{-1})((\alpha)(j)) = (\gamma^{-1})((\gamma)(j)) = j$. Esto quiere decir que γ es ajena con $\gamma^{-1} \circ \alpha$.

Aplicando la hipótesis de inducción a $\gamma^{-1} \circ \alpha$, tenemos que

$$\gamma^{-1} \circ \alpha = \underbrace{\gamma_2 \circ \dots \circ \gamma_s}_{\text{ciclos mutuamente ajenos}} \dots$$

Además, $\text{Mov}(\gamma_t) \subseteq \text{Mov}(\gamma^{-1} \circ \alpha) \subseteq \text{Fij}(\gamma)$, por lo que γ es ajena con cada γ_i entonces $\alpha = \underbrace{\gamma \circ \gamma_2 \circ \dots \circ \gamma_s}_{\text{ciclos mutuamente ajenos}}$.

Unicidad) Por inducción sobre $|\text{Mov}(\alpha)|$.

Base: si $|\text{Mov}(\alpha)| = 0$, entonces $\alpha = \text{Id}$ y ya notamos que la única manera de expresarla como producto de ciclos ajenos es con la factorización vacía.

Paso inductivo. Si $|\text{Mov}(\alpha)| > 0$, y

$$\alpha = \underbrace{\gamma_1 \circ \gamma_2 \circ \dots \circ \gamma_s}_{\text{ciclos ajenos}} = \underbrace{\lambda_1 \circ \lambda_2 \circ \dots \circ \lambda_t}_{\text{ciclos ajenos}}$$

supongamos sin perder generalidad que γ_1 y λ_1 mueven ambos a i_1 . Entonces

$\gamma_1(i_1) = \alpha(i_1) = \lambda_1(i_1)$, $\gamma_1(\gamma_1(i_1)) = \alpha(\alpha(i_1)) = \lambda_1(\lambda_1(i_1))$, etcétera.

Entonces

$$\gamma_1 = (i_1, \gamma_1(i_1), \dots) = (i_1, \alpha(i_1), \dots) = (i_1, \lambda_1(i_1), \dots) = \lambda_1.$$

Entonces tenemos que $\alpha \gamma_1^{-1} = \gamma_2 \circ \dots \circ \gamma_s = \lambda_2 \circ \dots \circ \lambda_t$ mueve menos elementos que α (ya que fija los elementos movidos por $\gamma_1 = \lambda_1$). Por hipótesis de inducción tenemos que $\{\gamma_2, \dots, \gamma_s\}_\neq = \{\lambda_2, \dots, \lambda_t\}_\neq$ ⁵ por lo que tenemos que $s = t$ y como además tenemos que $\gamma_1 = \lambda_1$, hemos establecido la unicidad de la factorización. ■

⁵ $\{\gamma_2, \dots, \gamma_s\}_\neq$ significa que los elementos son distintos.

8.11.2 La paridad de una permutación

Definición 141

1. Si γ es un ciclo tal que $|\text{Mov}(\gamma)| = m$, diremos que $\text{sig}(\gamma) = (-1)^{m+1}$.
2. Si $\alpha = \underbrace{\gamma_1 \circ \dots \circ \gamma_s}_{\text{ciclos ajenos}}$ entonces $\text{sig}(\alpha) = \prod_{i=1}^s \text{sig}(\gamma_i)$.
3. Diremos que una permutación es par si su signo es 1, diremos que es impar si su signo es -1 .

Notemos que la definición anterior es buena, debido al Teorema de factorización única. Notemos también que $\text{sig}((i, j)) = (-1)^{2+1} = -1$, por lo que una transposición es impar.

También notemos que $\text{sig}((1)) = \text{sig}(\Pi \emptyset) = \prod_{\gamma \in \emptyset} \text{sig}(\gamma) = 1$, dado que la identidad es un producto vacío de ciclos y dada la convención acerca de productos vacíos. Consecuentemente, (1) es una permutación par.

En el siguiente teorema mostraremos que componer con una transposición cambia el signo de una permutación. En términos más precisos, tenemos el siguiente teorema.

Teorema 142 . $\text{sig}(\alpha \circ \tau) = -\text{sig}(\alpha)$, si τ es una transposición.

Demostración. Supongamos que $\tau = (i, j)$ y que la factorización de α en ciclos ajenos es

$$\alpha = \gamma_1 \circ \dots \circ \gamma_m.$$

Consideramos los siguientes casos:

1) τ es ajena con α . En este caso, τ es ajena con cada γ_i , por lo que la factorización de $\alpha \circ \tau$ en ciclos ajenos es precisamente

$$\alpha \circ \tau = \underbrace{\gamma_1 \circ \dots \circ \gamma_m \circ \tau}_{\text{ciclos ajenos}}$$

entonces, es claro que $\text{sig}(\alpha \circ \tau) = \text{sig}(\alpha) \text{sig}(\tau) = -\text{sig}(\alpha)$.

2) $\text{Mov}(\alpha) \cap \text{Mov}(\tau)$ consta de un solo elemento.

Si no perder generalidad, supongamos que i es un elemento movido por α y por τ y que además el ciclo de α que mueve a i es $\gamma_k = (i, \alpha(i), \dots, \alpha^r(i))$. Es fácil ver que $\gamma_k \circ \tau = (i, j, \alpha(i), \dots, \alpha^r(i))$ ya que

$$i \xrightarrow{\tau} j \xrightarrow{\gamma_k} j, \quad j \xrightarrow{\tau} i \xrightarrow{\gamma_k} \alpha(i), \dots, \alpha^r(i) \xrightarrow{\tau} \alpha^r(i) \xrightarrow{\gamma_k} i,$$

por lo tanto $\text{sig}(\gamma_k \circ \tau) = (-1)^{r+2+1} = -(-1)^{r+2} = -\text{sig}(\gamma_k)$. Por lo tanto $\text{sig}(\alpha \circ \tau) = -\text{sig}(\alpha)$ en este caso.

3) $\text{Mov}(\alpha) \cap \text{Mov}(\tau) = \{i, j\}$.

Distinguimos dos subcasos:

a) i, j están movidos por el mismo ciclo de α , que sin perder generalidad supondremos es $\gamma_k = (i, \dots, \alpha^r(i), j, \dots, \alpha^l(j))$. Entonces $\text{sig}(\gamma_k) = (-1)^{r+1+l+1+1} = (-1)^{r+l+1}$.

Por otra parte,

$$\gamma_k \circ \tau = (i, \alpha(j), \dots, \alpha^l(j)) \circ (j, \alpha(i), \dots, \alpha^r(i))$$

como se comprueba fácilmente. Entonces

$$\begin{aligned} \text{sig}(\gamma_k \circ \tau) &= \text{sig}((i, \alpha(j), \dots, \alpha^l(j))) \text{sig}((j, \alpha(i), \dots, \alpha^r(i))) = \\ &= (-1)^{l+1+1} (-1)^{r+1+1} = (-1)^{l+r}. \end{aligned} \quad (8.22)$$

De donde tenemos que $\text{sig}(\gamma_k \circ \tau) = (-1)^{l+r} = -\text{sig}(\gamma_k) = -(-1)^{r+l+1}$.

b) Supongamos que i y j son movidos por ciclos distintos de α . Sin perder generalidad, supongamos que

$$\gamma_{k-1} = (i, \dots, \alpha^r(i)), \quad \gamma_k = (j, \dots, \alpha^l(j)),$$

entonces $\text{sig}(\gamma_{k-1}) = (-1)^{r+2}$ y $\text{sig}(\gamma_k) = (-1)^{l+2}$, por lo que

$$\text{sig}(\gamma_{k-1} \circ \gamma_k) = (-1)^{r+l}.$$

Ahora,

$$\gamma_{k-1} \circ \gamma_k \circ \tau = (i, \alpha(j), \dots, \alpha^l(j), j, \alpha(i), \dots, \alpha^r(i)),$$

de donde tenemos que

$$\text{sig}(\gamma_{k-1} \circ \gamma_k \circ \tau) = (-1)^{l+1+r+1+1} = (-1)^{l+r+1} = -\text{sig}(\gamma_{k-1} \circ \gamma_k).$$

En cualquier caso $\text{sig}(\alpha \circ \tau) = -\text{sig}(\alpha)$. ■

Ahora, como cualquier permutación es un producto de transposiciones,

$$\alpha = \underbrace{\tau_1 \circ \dots \circ \tau_k}_{\text{transposiciones}} = (1) \circ \tau_1 \circ \dots \circ \tau_k$$

entonces $\text{sig}(\alpha) = (-1)^k = \begin{cases} 1 & \text{si } k \text{ es par} \\ -1 & \text{si } k \text{ es impar.} \end{cases}$

Así tenemos el siguiente teorema.

Teorema 143 . *Si $\alpha = \underbrace{\tau_1 \circ \dots \circ \tau_k}_{\text{transposiciones}} \in S_n$ entonces $\text{paridad}(\alpha) = \text{paridad}(k)$.*

Observación 123 . *Si $n \geq 2$ entonces el número de permutaciones pares en S_n coincide con el número de permutaciones impares.*

Demostración. Denotemos $A_n = \{\alpha \in S_n \mid \alpha \text{ es par}\}$ sea $\tau = (1, 2)$. Entonces la función

$$\begin{aligned} A_n &\xrightarrow{\circ\tau} S_n \setminus A_n \\ \alpha &\mapsto \alpha \circ \tau \end{aligned}$$

es una biyección cuyo inverso es

$$\begin{aligned} S_n \setminus A_n &\xrightarrow{\circ\tau} A_n \\ \beta &\mapsto \beta \circ \tau \end{aligned} .$$

■

Ejercicio 388. Demuestre que si α, β son permutaciones, entonces $\text{sig}(\alpha \beta) = \text{sig}(\alpha) \text{sig}(\beta)$.

8.11.3 Determinantes

En la siguiente definición F denota un campo, por ejemplo \mathbb{R} , pero también tiene sentido si F es un anillo comunitativo.

Definición 142 . *Si $A \in M_{n \times n}(F)$ definimos*

$$\det(A) = \sum_{\sigma \in S_n} \left(\text{sig}(\sigma) \left(\prod_{i=1}^n A_{i, \sigma(i)} \right) \right).$$

⁶Usamos también la notación

$$\det(A) = |A|. \quad (8.23)$$

Ejemplos 185

1. Por ejemplo si $n = 1$, entonces $S_1 = \{(1)\}$, así que

$$\det((A_{1,1})) = \text{sig}((1)) A_{1,1} = A_{1,1}.$$

2. Si $n = 2$, entonces $S_2 = \{(1), (1, 2)\}$ y si $A = \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix}$ entonces

$$\begin{aligned} \det(A) &= \text{sig}((1)) A_{1,1} A_{2,2} + \text{sig}((1, 2)) A_{1,2} A_{2,1} = \\ &= A_{1,1} A_{2,2} - A_{1,2} A_{2,1}. \end{aligned}$$

3. $S_3 = \{(1), (1, 2, 3), (1, 3, 2), (1, 2), (1, 3), (2, 3)\}$. Aquí el conjunto de la permutaciones pares es $A_n = \{(1), (1, 2, 3), (1, 3, 2)\}$ por lo que

$$\begin{aligned} \det \left(\begin{pmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{2,1} & A_{2,2} & A_{2,3} \\ A_{3,1} & A_{3,2} & A_{3,3} \end{pmatrix} \right) &= \\ &= \text{sig}((1)) A_{1,1} A_{2,2} A_{3,3} + \text{sig}((1, 2, 3)) A_{1,2} A_{2,3} A_{3,1} + \\ &\quad + \text{sig}((1, 3, 2)) A_{1,3} A_{2,1} A_{3,2} + \\ &\quad + \text{sig}((1, 2)) A_{1,2} A_{2,1} A_{3,3} + \\ &\quad + \text{sig}((1, 3)) A_{1,3} A_{2,2} A_{3,1} + \\ &\quad + \text{sig}((2, 3)) A_{1,1} A_{2,3} A_{3,2} \\ &= A_{1,1} A_{2,2} A_{3,3} + A_{1,2} A_{2,3} A_{3,1} + A_{1,3} A_{2,1} A_{3,2} \\ &\quad - A_{1,2} A_{2,1} A_{3,3} - A_{1,3} A_{2,2} A_{3,1} - A_{1,1} A_{2,3} A_{3,2}. \end{aligned}$$

Recordemos que una matriz B es una submatriz menor de A si se obtuvo tachando (no todos los) renglones y (no todas las) columnas de A . La matriz $\widehat{A_{\{i_1, \dots, i_s\}, \{j_1, \dots, j_t\}}}$ es la matriz que se obtiene al tachar de A los renglones i_1, \dots, i_s y las columnas j_1, \dots, j_t . Escribiremos $\widehat{A_{i,j}}$ en lugar de $\widehat{A_{\{i\}, \{j\}}}$.

Ejemplo 186 . Si $A = \begin{pmatrix} 4 & -5 & -4 \\ -1 & -9 & -7 \\ 2 & 6 & 8 \end{pmatrix}$ entonces $\widehat{A_{2,1}} = \begin{pmatrix} -5 & -4 \\ 6 & 8 \end{pmatrix}$ y $\widehat{A_{\{1,2\}, \{3\}}} = \begin{pmatrix} 2 & 6 \end{pmatrix}$.

8.11.4 El desarrollo del determinante respecto a un renglón

Obtendremos otra descripción del determinante, que se llama “desarrollo del determinante respecto al primer renglón”.

Sea $\sigma \in S_n$ tal que $\sigma(1) = j$ a esta permutación le asociaremos otra permutación en S_{n-1} de la manera siguiente:

como $\{1, \dots, n\} \xrightarrow{\sigma} \{1, \dots, n\} \in S_n$ entonces

$$\{2, \dots, n\} \xrightarrow{\sigma|_{\{2, \dots, n\}}} \{1, \dots, n\} \setminus \{j\} \text{ es una biyección,}$$

considerando las biyecciones

$$\begin{array}{ccc} \{2, \dots, n\} & \xrightarrow{\sigma|_{\{2, \dots, n\}}} & \{1, \dots, n\} \setminus \{j\} \\ \downarrow m & & \downarrow f \\ \{1, \dots, n-1\} & & \{1, \dots, n-1\} \end{array}$$

donde:

$$m(k) = k-1 \text{ y } f(k) = \begin{cases} k \text{ si } k < j \\ k-1 \text{ si } k > j. \end{cases},$$

podemos definir $\sigma' \in S_{n-1}$ por

$$\sigma' = f \circ \sigma|_{\{2, \dots, n\}} \circ m^{-1}.$$

Entonces el diagrama

$$\begin{array}{ccc} \{2, \dots, n\} & \xrightarrow{\sigma|_{\{2, \dots, n\}}} & \{1, \dots, n\} \setminus \{j\} \\ \downarrow m & & \downarrow f \\ \{1, \dots, n-1\} & \xrightarrow{\sigma'} & \{1, \dots, n-1\} \end{array}$$

commuta.

Note que por definición

$$\sigma'(i) = f(\sigma|_{\{2, \dots, n\}}(i+1)) = \begin{cases} \sigma|_{\{2, \dots, n\}}(i+1) & \text{si } \sigma|_{\{2, \dots, n\}}(i+1) < j \\ \sigma|_{\{2, \dots, n\}}(i+1)-1 & \text{si } \sigma|_{\{2, \dots, n\}}(i+1) > j. \end{cases}$$

Haremos ahora algunas observaciones importantes:

Observación 124 . *La función*

$$\begin{aligned} \{\sigma \in S_n \mid \sigma(1) = j\} &\rightarrow S_{n-1} \\ \sigma &\mapsto \sigma' = f \circ \sigma_{\{2, \dots, n\}} \circ m^{-1} \end{aligned}$$

es una biyección.

Pues note que

$$\begin{aligned} S_{n-1} &\rightarrow \text{Biyecc}(\{2, \dots, n\}, \{1, \dots, n\} \setminus \{j\}) \\ \gamma &\mapsto f^{-1} \circ \gamma \circ m \end{aligned}$$

$f^{-1} \circ \gamma \circ m$ se puede extender de manera única a una permutación $\overline{f^{-1} \circ \gamma \circ m} \in S_n$, simplemente definiendo $\overline{f^{-1} \circ \gamma \circ m}(1) = j$.

Ahora la función

$$\begin{aligned} S_{n-1} &\rightarrow \{\sigma \in S_n \mid \sigma(1) = j\} \\ \gamma &\mapsto \overline{f^{-1} \circ \gamma \circ m} \end{aligned}$$

es la inversa. de la función $\{1, \dots, n-1\} \xrightarrow{\sigma'} \{1, \dots, n-1\}$.

Ejercicio 389 . *Verifique que*

$$\begin{aligned} S_{n-1} &\rightarrow \{\sigma \in S_n \mid \sigma(1) = j\} \\ \gamma &\mapsto \overline{f^{-1} \circ \gamma \circ m} \end{aligned}$$

es la inversa de la función

$$\begin{aligned} \{1, \dots, n-1\} &\xrightarrow[\sigma]{f^{-1} \circ (\cdot) \circ m} \{1, \dots, n-1\} \\ &\mapsto \sigma' = f^{-1} \circ (\sigma) \circ m \end{aligned}.$$

Observación 125 . *Respecto a la notación de la observación anterior, tenemos que* $\text{sig}(\sigma) = (-1)^{j+1} \text{sig}(\sigma')$.

Demostración. $\sigma' \in S_{n-1}$ se puede extender a $\sigma'' \in S_n$ mandando n a n , entonces $\text{sig}(\sigma') = \text{sig}(\sigma'')$ (simplemente note que sus factorizaciones en ciclos ajenos son idénticas. De la definición de σ'

$$\begin{array}{ccc} \{2, \dots, n\} & \xrightarrow{\sigma_{\{2, \dots, n\}}} & \{1, \dots, n\} \setminus \{j\} \\ \downarrow _ - 1 & & \downarrow f \\ \{1, \dots, n-1\} & \xrightarrow{\sigma'} & \{1, \dots, n-1\} \end{array}$$

tenemos que $\{2, \dots, n\} \xrightarrow{-1} \{1, \dots, n-1\}$ ⁷ se puede extender a una permutación γ en S_n enviando 1 a n . También

$$\{1, \dots, n\} \setminus \{j\} \xrightarrow{f} \{1, \dots, n-1\}$$

se puede extender a una permutación $\varphi \in S_n$ definiendo $\varphi(j) = n$. Entonces

$$\begin{array}{ccccc}
 \{1, 2, \dots, n\} & & \xrightarrow{-\sigma} & & \{1, 2, \dots, n\} \\
 \nwarrow & & & & \nearrow \\
 & \{2, \dots, n\} & \xrightarrow{\sigma|_{\{2, \dots, n\}}} & \{1, \dots, n\} \setminus \{j\} & \\
 \downarrow \gamma & \downarrow -1 & & \downarrow f & \downarrow \varphi \\
 \{1, \dots, n-1\} & \xrightarrow{\sigma'} & \{1, \dots, n-1\} & & \\
 \swarrow & & \xrightarrow{-\sigma''} & & \searrow \\
 \{1, 2, \dots, n\} & & & & \{1, 2, \dots, n\}
 \end{array}$$

de aquí que $\sigma'' \circ \gamma = \varphi \circ \sigma$, es decir que $\sigma'' = \varphi \circ \sigma \circ \gamma^{-1}$. Por lo tanto $\text{sig}(\sigma'') = \text{sig}(\varphi) \text{sig}(\sigma) \text{sig}(\gamma)$.

Calculemos $\text{sig}(\varphi)$, φ expresada como producto de ciclos ajenos es: $(n, n-1, \dots, j)$ por lo que $\text{sig}(\varphi) = (-1)^{n-j+1+1} = (-1)^{n-j}$.

Por otra parte, $\gamma = (n, n-1, \dots, 1)$, por lo que $\text{sig}(\gamma) = (-1)^{n+1}$.

Entonces

$$\text{sig}(\sigma'') = \text{sig}(\sigma) (-1)^{n-j} (-1)^{n+1} = \text{sig}(\sigma) (-1)^{2n-j+1} = \text{sig}(\sigma) (-1)^{j+1} \quad (8.24)$$

Así, $\text{sig}(\sigma) = (-1)^{j+1} \text{sig}(\sigma')$. ■

Lema 28 . $\det(A) = \sum_{j=1}^n (-1)^{j+1} A_{i,j} \det(\widehat{A_{i,j}})$.

Demostración. Por definición, $\det(A) = \sum_{\sigma \in S_n} \left(\text{sig}(\sigma) \left(\prod_{i=1}^n A_{i, \sigma(i)} \right) \right)$.

⁷Desde luego $\{2, \dots, n\} \xrightarrow{-1} \{1, \dots, n-1\}$ es la función "restar 1".

Separaremos esta suma en varios sumandos, según la imagen de 1 :

$$\begin{aligned}\det(A) &= \sum_{\sigma \in S_n} \left(\text{sig}(\sigma) \left(\prod_{i=1}^n A_{i,\sigma(i)} \right) \right) = \\ &= \sum_{j=1}^n \left(\sum_{\substack{\sigma \\ \sigma(1)=j}} \left(\text{sig}(\sigma) \left(\prod_{i=1}^n A_{i,\sigma(i)} \right) \right) \right).\end{aligned}$$

Consideremos ahora uno de los sumandos:

$$\begin{aligned}&\sum_{\substack{\sigma \\ \sigma(1)=j}} \left(\text{sig}(\sigma) \left(\prod_{i=1}^n A_{i,\sigma(i)} \right) \right) \\ &= \sum_{\substack{\sigma \\ \sigma(1)=j}} A_{1,j} \left(\text{sig}(\sigma) \left(\prod_{i=2}^n A_{i,\sigma(i)} \right) \right).\end{aligned}$$

Recordando que

$$\begin{array}{ccc} \{2, \dots, n\} & \xrightarrow{\sigma|_{\{2, \dots, n\}}} & \{1, \dots, n\} \setminus \{j\} \\ \downarrow _ - 1 & & \downarrow f \\ \{1, \dots, n-1\} & \xrightarrow{\sigma'} & \{1, \dots, n-1\} \end{array} \quad (8.25)$$

donde $f(k) = \begin{cases} k & \text{si } k < j \\ k-1 & \text{si } k > j \end{cases}$, tenemos que $\sigma'(x) = (f \circ \sigma)(x+1)$ y $\sigma' \in S_{n-1}$.

Así tenemos que $A_{i,\sigma(i)} = \begin{cases} \widehat{A_{1,j}}_{i-1,\sigma(i)} & \text{si } i \geq 2, \sigma(i) < j \\ \widehat{A_{1,j}}_{i-1,\sigma(i)-1} & \text{si } i \geq 2, \sigma(i) > j \end{cases}$, es decir

$$\text{que } A_{i,\sigma(i)} = \widehat{A_{1,j}}_{i-1,f \circ \sigma(i)} = \widehat{A_{1,j}}_{i-1,\sigma'(i-1)} \text{ si } i \geq 2.$$

Entonces podemos escribir

$$\prod_{i=2}^n A_{i,\sigma(i)} = \prod_{i=1}^{n-1} \widehat{(A_{i,j})}_{i,\sigma'(i)}.$$

Ahora usaremos las relaciones entre $sig(\sigma)$ y $sig(\sigma')$, dadas en las dos Observaciones anteriores.

Regresando a 8.11.4,

$$\begin{aligned} A_{1,j} \sum_{\sigma} & \left(sig(\sigma) \left(\prod_{i=2}^n A_{i,\sigma(i)} \right) \right) \\ \sigma(1) = j \\ & = (-1)^{j+1} A_{1,j} \sum_{\sigma' \in S_{n-1}} \left(sig(\sigma') \left(\prod_{i=1}^{n-1} \widehat{(A_{1,j})}_{i,\sigma'(i)} \right) \right) \\ & = (-1)^{j+1} A_{1,j} \det \left(\widehat{(A_{1,j})} \right). \end{aligned}$$

Donde hemos usado que

$$\begin{aligned} \{\sigma \in S_n \mid \sigma(1) = j\} & \rightarrow S_{n-1} \\ \sigma & \mapsto \sigma' \end{aligned}$$

es una biyección.

$$\text{Entonces } \det(A) = \sum_{j=1}^n (-1)^{j+1} A_{1,j} \det \left(\widehat{(A_{1,j})} \right). \blacksquare$$

Recordemos que denotamos por $\mathcal{I}_{i,j}$ la operación elemental que intercambia el renglón i con el renglón j en una matriz A .

Lema 29 . $\det(\mathcal{I}_{i,j}(A)) = -\det(A)$.

Demostración. Sea $\tau = (i, j)$ entonces

$$\det(\mathcal{I}_{i,j}(A)) = \sum_{\sigma \in S_n} sig(\sigma) \prod_{k=1}^n (\mathcal{I}_{i,j}(A))_{k,\sigma(k)}.$$

Ahora,

$$(\mathcal{I}_{i,j}(A))_{i,\sigma(i)} = A_{j,\sigma(i)} = A_{j,\sigma(\tau(j))}$$

en virtud del cambio de renglón y de que $\tau(j) = i$.

También

$$(\mathcal{I}_{i,j}(A))_{j,\sigma(j)} = A_{i,\sigma(j)} = A_{i,\sigma(\tau(i))}$$

y si $k \notin \{i, j\}$

$$(\mathcal{I}_{i,j}(A))_{k,\sigma(k)} = A_{k,\sigma(k)} = A_{k,\sigma(\tau(k))},$$

por lo tanto

$$\det(\mathcal{I}_{i,j}(A)) = \sum_{\sigma \in S_n} \text{sig}(\sigma) \prod_{k=1}^n (A)_{k,\sigma(\tau(k))}.$$

Notemos ahora que $\text{sig}(\sigma) = -\text{sig}(\sigma \circ \tau)$ y que como $S_n \xrightarrow{\circ \tau} S_n$ es una biyección, entonces $(\sigma \circ \tau)$ “corre sobre S_n , cuando σ corre sobre S_n ”, por lo que haciendo $\gamma = \sigma \circ \tau$, tenemos que

$$\begin{aligned} \det(\mathcal{I}_{i,j} A) &= \\ \sum_{\sigma \in S_n} \text{sig}(\sigma) \prod_{k=1}^n (A)_{k,\sigma(\tau(k))} &= \\ \sum_{\sigma \in S_n} -\text{sig}(\sigma \circ \tau) \prod_{k=1}^n (A)_{k,\sigma(\tau(k))} &= \\ - \left(\sum_{\gamma \in S_n} \text{sig}(\gamma) \prod_{k=1}^n (A)_{k,\gamma(k)} \right) &= \\ = -\det(A). \end{aligned}$$

■

Ahora podemos encontrar una fórmula para desarrollar el determinante respecto de cualquier renglón.

Teorema 144 . *Si A es una matriz en $M_{n \times n}(F)$ entonces*

$$\det(A) = \sum_j (-1)^{i+j} A_{i,j} \det(\widehat{A_{i,j}}).$$

Demostración. Recordemos que $A_{\underline{i}}$ denota el renglón i de la matriz A . Mediante $i - 1$ intercambios de renglón podemos llevar el renglón i -ésimo de

A a ocupar el primer lugar de los renglones. Obtenemos una matriz B tal que

$$B_{\underline{1}} = A_{\underline{i}}, B_{\underline{2}} = A_{\underline{1}}, \dots, B_{\underline{i}} = A_{\underline{i-1}}, B_{\underline{i+1}} = A_{\underline{i+1}}, \dots, B_{\underline{n}} = A_{\underline{n}}.$$

Entonces $\det(B) = (-1)^{i-1} \det(A)$ y también

$$\begin{aligned} \det(B) &= \sum_j (-1)^{j+1} B_{1,j} \det(\widehat{B_{1,j}}) = \\ &= \sum_j (-1)^{j+1} A_{i,j} \det(\widehat{A_{i,j}}). \end{aligned}$$

Pues es inmediato que $\widehat{B_{1,j}} = \widehat{A_{i,j}}$.

Entonces $(-1)^{i-1} \det(A) = \det(B) = \sum_j (-1)^{j+1} A_{i,j} \det(\widehat{A_{i,j}})$, es decir que

$$\det(A) = \sum_j (-1)^{i+j+1-1} A_{i,j} \det(\widehat{A_{i,j}}).$$

■

Como consecuencia de este resultado, podemos demostrar lo siguiente.

Lema 30 . *Si $A \in M_{n \times n}(F)$ es una matriz con dos renglones iguales, entonces $\det(A) = 0$.*

Demostración. Por inducción sobre n .

Base. Si $n = 1$ no hay nada que demostrar.

$$\text{Si } n = 2 \text{ entonces } \det \begin{pmatrix} a & b \\ a & b \end{pmatrix} = ab - ab = 0.$$

Si $n \geq 3$ y los renglones iguales son $A_{\underline{i}}$ y $A_{\underline{j}}$, desarrollando el determinante respecto del k -ésimo renglón, con $k \notin \{i, j\}$ entonces

$$\det(A) = \sum_l (-1)^{k+l} A_{k,l} \det(\widehat{A_{k,l}}) = 0,$$

pues $\widehat{A_{k,l}} \in M_{n-1 \times n-1}(F)$ tiene dos renglones iguales, así que su determinante es 0, por hipótesis de inducción. ■

Recordemos que la matriz transpuesta de A , A^t es la matriz definida por $(A^t)_{i,j} = A_{j,i}$.

Teorema 145 . $\det(A^t) = \det(A)$, $A \in M_{n \times n}(F)$.

Demostración. $\det(A^t) = \sum_{\sigma} \text{sig}(\sigma) \prod_{i} (A^t)_{i, \sigma(i)} = \sum_{\sigma} \text{sig}(\sigma) \prod_{i} A_{\sigma(i), i}$. Como $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ es una biyección, cuando i corre sobre $\{1, \dots, n\}$, entonces $\sigma(i)$ también. Si hacemos $k = \sigma(i)$, entonces $i = \sigma^{-1}(k)$ y

$$\prod_{i \in \{1, \dots, n\}} A_{\sigma(i), i} = \prod_{k \in \{1, \dots, n\}} A_{k, \sigma^{-1}(k)}$$

puesto que ambos productos tienen los mismos factores (reordenados).

Entonces $\det(A^t) = \sum_{\sigma} \text{sig}(\sigma) \left(\prod_{k \in \{1, \dots, n\}} A_{k, \sigma^{-1}(k)} \right)$. Además tenemos

que $\text{sig}(\sigma) = \text{sig}(\sigma^{-1})$ y que

$$\begin{aligned} S_n &\rightarrow S_n \\ \sigma &\mapsto \sigma^{-1} \end{aligned}$$

es una biyección, por lo que

$$\begin{aligned} \det(A^t) &= \sum_{\sigma^{-1}} \text{sig}(\sigma^{-1}) \left(\prod_{k \in \{1, \dots, n\}} A_{k, \sigma^{-1}(k)} \right) = \\ &= \det(A). \end{aligned}$$

■

Teorema 146 . $\det(A) = \sum_i (-1)^{i+j} A_{i,j} \det(\widehat{A_{i,j}})$.

Demostración. Se deja como ejercicio. ■

En vista de que $\det(A^t) = \det(A)$, entonces para cada teorema sobre los determinantes formulado respecto de los renglones de una matriz, tenemos un teorema correspondiente respecto de las columnas de una matriz.

Por ejemplo, el determinante de una matriz que tenga dos columnas iguales vale 0, el efecto de intercambiar dos columnas en el determinante es un cambio de signo, etc.

8.11.5 El determinante de un producto de matrices I

El propósito de esta sección es demostrar que el determinante de un producto de matrices es el producto de sus determinantes. Esta demostración

puede parecer difícil de seguir, pero en la sección 8.11.7, presentaremos otro argumento.

El argumento que sigue tiene el mérito de que vale para matrices con coeficientes en un anillo comunitativo y no necesariamente en un campo, lo que podría ser importante en algunas situaciones. Por ejemplo cuando se consideran polinomios característicos de una matriz, que son determinantes de matrices con coeficientes en un anillo de polinomios. Pero el resultado vale también, por ejemplo, para matrices con coeficientes enteros.

Pero si el lector está interesado por el momento en matrices con coeficientes sobre un campo, puede ver la demostración de la sección 8.11.7.

Teorema 147 . $\det(AB) = \det(A)\det(B)$.

Demostración.

$$\begin{aligned}\det(AB) &= \sum_{\sigma \in S_n} \text{sig}(\sigma) \left[\prod_{i=1}^n (AB)_{i,\sigma(i)} \right] = \\ &= \sum_{\sigma \in S_n} \text{sig}(\sigma) \left[\prod_{i=1}^n \left(\sum_{k=1}^n A_{i,k} B_{k,\sigma(i)} \right) \right].\end{aligned}$$

Ahora para una σ fija

$$\begin{aligned}\prod_{i=1}^n \left(\sum_{k=1}^n A_{i,k} B_{k,\sigma(i)} \right) &= \\ &= \sum_{\{f: \{1, \dots, n\} \rightarrow \{1, \dots, n\} \mid f \text{ es función}\}} A_{1,f(1)} B_{f(1),\sigma(1)} A_{2,f(2)} B_{f(2),\sigma(2)} \dots A_{n,f(n)} B_{f(n),\sigma(n)}.\end{aligned}$$

Para convencerse de lo anterior, hagamos una lista de los factores en

$$\prod_{i=1}^n \left(\sum_{k=1}^n A_{i,k} B_{k,\sigma(i)} \right) :$$

$$\begin{aligned}A_{1,1} B_{1,\sigma(1)} + A_{1,2} B_{2,\sigma(1)} + \dots + A_{1,n} B_{n,\sigma(1)} \\ A_{2,1} B_{1,\sigma(2)} + A_{2,2} B_{2,\sigma(2)} + \dots + A_{2,n} B_{n,\sigma(2)}\end{aligned}$$

$$A_{n,1} B_{1,\sigma(n)} + A_{n,2} B_{2,\sigma(n)} + \dots + A_{n,n} B_{n,\sigma(n)}$$

el producto de las sumas anteriores es una suma de productos. Es decir, consta de los productos que se pueden formar escogiendo un sumando en cada una de las listas anteriores y después multiplicándolos. Por ejemplo podríamos escoger el sumando $A_{1,2}B_{2,\sigma(1)}$ en la primera suma, el sumando $A_{2,1}B_{1,\sigma(2)}$ en la segunda hilera, el sumando $A_{3,1}B_{1,\sigma(3)}$ en la tercera hilera,..., el sumando $A_{n,n}B_{n,\sigma(n)}$, y multiplicándolos obtendríamos

$$A_{1,2}B_{2,\sigma(1)}A_{2,1}B_{1,\sigma(2)}A_{3,1}B_{1,\sigma(3)}\dots A_{n,n}B_{n,\sigma(n)}.$$

La elección anterior corresponde a la función $1 \xrightarrow{f} 2$ (se escogió el segundo sumando de la primera hilera), $2 \xrightarrow{f} 1$ (se escogió el primer sumando de la segunda hilera), $3 \xrightarrow{f} 1$ (se escogió el primer sumando de la tercera hilera),..., $n \xrightarrow{f} n$ (se escogió el enésimo sumando de la enésima hilera). Para esta f particular, podemos reescribir 8.11.5 como

$$A_{1,f(1)}B_{f(1),\sigma(1)}A_{2,f(2)}B_{f(2),\sigma(2)}A_{3,f(3)}B_{f(3),\sigma(3)}\dots A_{n,f(n)}B_{f(n),\sigma(n)}.$$

Lo anterior debe dejar claro que

$$\prod_{i=1}^n \left(\sum_{k=1}^n A_{i,k}B_{k,\sigma(i)} \right) = \sum_{\{f:\{1,\dots,n\} \rightarrow \{1,\dots,n\} \mid f \text{ es función}\}} A_{1,f(1)}B_{f(1),\sigma(1)}A_{2,f(2)}B_{f(2),\sigma(2)}\dots A_{n,f(n)}B_{f(n),\sigma(n)}.$$

Entonces

$$\begin{aligned} & \det(\mathcal{A}B) \\ &= \sum_{\sigma \in S_n} \text{sig}(\sigma) \sum_{\{f:\{1,\dots,n\} \rightarrow \{1,\dots,n\} \mid f \text{ es función}\}} A_{1,f(1)}B_{f(1),\sigma(1)}A_{2,f(2)}B_{f(2),\sigma(2)}\dots A_{n,f(n)}B_{f(n),\sigma(n)} \\ &= \sum_{\sigma \in S_n} \sum_{\{f:\{1,\dots,n\} \rightarrow \{1,\dots,n\} \mid f \text{ es función}\}} \text{sig}(\sigma) A_{1,f(1)}A_{2,f(2)}\dots A_{n,f(n)}B_{f(1),\sigma(1)}B_{f(2),\sigma(2)}\dots B_{f(n),\sigma(n)} \\ &= \sum_{\sigma \in S_n} \sum_{f \text{ es biyectiva}} \text{sig}(\sigma) A_{1,f(1)}A_{2,f(2)}\dots A_{n,f(n)}B_{f(1),\sigma(1)}B_{f(2),\sigma(2)}\dots B_{f(n),\sigma(n)} + \\ & \quad + \sum_{\sigma \in S_n} \sum_{f \text{ no es biyectiva}} \text{sig}(\sigma) A_{1,f(1)}A_{2,f(2)}\dots A_{n,f(n)}B_{f(1),\sigma(1)}B_{f(2),\sigma(2)}\dots B_{f(n),\sigma(n)}. \end{aligned}$$

Mostraremos que

$$\sum_{\sigma \in S_n} \sum_{\{f: \{1, \dots, n\} \rightarrow \{1, \dots, n\} \mid f \text{ es biyectiva}\}} \text{sig}(\sigma) A_{1,f(1)} A_{2,f(2)} \dots A_{n,f(n)} B_{f(1),\sigma(1)} B_{f(2),\sigma(2)} \dots B_{f(n),\sigma(n)}$$

es $\det(A) \det(B)$. Mientras que

$$\sum_{\sigma \in S_n} \sum_{\{f: \{1, \dots, n\} \rightarrow \{1, \dots, n\} \mid f \text{ no es biyectiva}\}} \text{sig}(\sigma) A_{1,f(1)} A_{2,f(2)} \dots A_{n,f(n)} B_{f(1),\sigma(1)} B_{f(2),\sigma(2)} \dots B_{f(n),\sigma(n)}$$

es 0. Comencemos por esto último:

$$\begin{aligned} & \sum_{\sigma \in S_n} \left(\sum_{f \text{ no es biyectiva}} \text{sig}(\sigma) A_{1,f(1)} A_{2,f(2)} \dots A_{n,f(n)} B_{f(1),\sigma(1)} B_{f(2),\sigma(2)} \dots B_{f(n),\sigma(n)} \right) = \\ &= \sum_{f \text{ no es biyectiva}} \left(\sum_{\sigma \in S_n} \text{sig}(\sigma) A_{1,f(1)} A_{2,f(2)} \dots A_{n,f(n)} B_{f(1),\sigma(1)} B_{f(2),\sigma(2)} \dots B_{f(n),\sigma(n)} \right) \end{aligned}$$

Tomemos una función $f: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ que no sea biyectiva. Entonces f no es inyectiva, para fijar ideas, supongamos que $f(1) = f(2)$.

Definamos la matriz C tal que $C_{i,j} = A_{i,f(i)} B_{f(i),j}$ entonces

$$C = \begin{pmatrix} A_{1,f(1)} B_{f(1),1} & A_{1,f(1)} B_{f(1),2} & \dots & A_{1,f(1)} B_{f(1),n} \\ A_{2,f(2)} B_{f(2),1} & A_{2,f(2)} B_{f(2),2} & \dots & A_{2,f(2)} B_{f(2),n} \\ \vdots & \vdots & & \vdots \\ A_{n,f(n)} B_{f(n),1} & A_{n,f(n)} B_{f(n),2} & \dots & A_{n,f(n)} B_{f(n),n} \end{pmatrix},$$

observemos que determinante de esta matriz es precisamente

$$\left(\sum_{\sigma \in S_n} \text{sig}(\sigma) A_{1,f(1)} B_{f(1),\sigma(1)} A_{2,f(2)} B_{f(2),\sigma(2)} \dots A_{n,f(n)} B_{f(n),\sigma(n)} \right).$$

Por otro lado, como $f(1) = f(2)$ entonces

$$\det(C) = A_{1,f(1)} A_{2,f(2)} \det \begin{pmatrix} B_{f(1),1} & B_{f(1),2} & \dots & B_{f(1),n} \\ B_{f(2),1} & B_{f(2),2} & \dots & B_{f(2),n} \\ \vdots & & & \vdots \\ A_{n,f(n)} B_{f(n),1} & A_{n,f(n)} B_{f(n),2} & \dots & A_{n,f(n)} B_{f(n),n} \end{pmatrix} = 0$$

pues los dos primeros renglones de la matriz anterior son iguales.

Ahora tenemos que ver que

$$\begin{aligned} & \sum_{\sigma \in S_n} \left(\sum_{f \in S_n} \text{sig}(\sigma) A_{1,f(1)} A_{2,f(2)} \dots A_{n,f(n)} B_{f(1),\sigma(1)} B_{f(2),\sigma(2)} \dots B_{f(n),\sigma(n)} \right) \\ &= \det(A) \det(B). \end{aligned}$$

Notemos que

$$\begin{aligned} & \sum_{\sigma \in S_n} \left(\sum_{f \in S_n} \text{sig}(\sigma) A_{1,f(1)} A_{2,f(2)} \dots A_{n,f(n)} B_{f(1),\sigma(1)} B_{f(2),\sigma(2)} \dots B_{f(n),\sigma(n)} \right) = \\ &= \sum_{f \in S_n} \left(\sum_{\sigma \in S_n} \text{sig}(\sigma) A_{1,f(1)} A_{2,f(2)} \dots A_{n,f(n)} B_{f(1),\sigma(1)} B_{f(2),\sigma(2)} \dots B_{f(n),\sigma(n)} \right) = \\ & \sum_{f \in S_n} A_{1,f(1)} A_{2,f(2)} \dots A_{n,f(n)} \left(\sum_{\sigma \in S_n} \text{sig}(\sigma) B_{f(1),\sigma(1)} B_{f(2),\sigma(2)} \dots B_{f(n),\sigma(n)} \right). \end{aligned}$$

Ahora

$$B_{f(1),\sigma(1)} B_{f(2),\sigma(2)} \dots B_{f(n),\sigma(n)} = \prod_{i=1}^n B_{f(i),\sigma(i)} = \prod_{j=1}^n B_{j,\sigma(f^{-1}(j))}.$$

Además $\text{sig}(\sigma \circ f^{-1}) = \text{sig}(\sigma) \text{sig}(f)$, por lo que $\text{sig}(\sigma) = \text{sig}(\sigma \circ f^{-1}) \text{sig}(f)$, de esta manera,

$$\begin{aligned} & \sum_{\sigma \in S_n} \text{sig}(\sigma) B_{f(1),\sigma(1)} B_{f(2),\sigma(2)} \dots B_{f(n),\sigma(n)} \\ &= \sum_{\sigma \in S_n} \text{sig}(\sigma \circ f^{-1}) \text{sig}(f) B_{1,\sigma \circ f^{-1}(1)} B_{2,\sigma \circ f^{-1}(2)} \dots B_{n,\sigma \circ f^{-1}(n)}. \end{aligned}$$

Como $S_n \xrightarrow{\circ f^{-1}} S_n$ es una biyección podemos hacer la sustitución $\gamma = g \circ f^{-1}$ en la ecuación de arriba para escribir:

$$\begin{aligned} & \sum_{\sigma \in S_n} \text{sig}(\sigma) B_{f(1),\sigma(1)} B_{f(2),\sigma(2)} \dots B_{f(n),\sigma(n)} \\ &= \text{sig}(f) \sum_{\gamma \in S_n} \text{sig}(\gamma) B_{1,\gamma(1)} B_{2,\gamma(2)} \dots B_{n,\gamma(n)} \\ &= \text{sig}(f) \det(B). \end{aligned}$$

Entonces

$$\begin{aligned}
 & \sum_{f \in S_n} A_{1,f(1)} A_{2,f(2)} \dots A_{n,f(n)} \left(\sum_{\sigma \in S_n} \text{sig}(\sigma) B_{f(1),\sigma(1)} B_{f(2),\sigma(2)} \dots B_{f(n),\sigma(n)} \right) = \\
 & = \sum_{f \in S_n} (A_{1,f(1)} A_{2,f(2)} \dots A_{n,f(n)}) (\text{sig}(f) \det(B)) = \\
 & = \det(B) \left(\sum_{f \in S_n} \text{sig}(f) (A_{1,f(1)} A_{2,f(2)} \dots A_{n,f(n)}) \right) = \\
 & = \det(B) \det(A).
 \end{aligned}$$

■

8.11.6 Determinantes y rango

Recordemos que el rango de una matriz es la dimensión del espacio generado por sus renglones (= dimensión del espacio generado por sus columnas).

Lema 31. *$A \in M_{n \times n}(F)$ es invertible si y sólo $\det(A) \neq 0$.*

Demostración. Si A no es invertible, entonces el conjunto de sus renglones es linealmente dependiente, y alguno de sus renglones es combinación lineal de los anteriores. Si el primer renglón es $\vec{0}$ entonces ya sabemos que $\det(A) = 0$. Si $A_{\underline{i}} = \sum_{j < i} c_j A_{\underline{j}}$, entonces

$$A_{\underline{i}} - \sum_{j < i} c_j A_{\underline{j}} = \vec{0}$$

por lo que por medio de operaciones elementales de renglón del tercer tipo (que no alteran el determinante) podemos llevar la matriz A a una matriz que tiene su renglón i de ceros, como esta matriz tiene determinante 0, entonces también se tiene que $\det(A) = 0$.

Si A es invertible, podemos proceder por inducción sobre n .

Base. Si $n = 1$, la afirmación es claramente cierta: $(A_{1,1})$ es invertible $\iff A_{1,1} \neq 0 \iff \det(A_{1,1}) = A_{1,1} \neq 0$.

Paso inductivo. Como el rango de A es n entonces todas sus columnas son linealmente independientes, y en particular la primera columna de A

no es $\vec{0}$. Entonces $A_{i,1} \neq 0$ intercambiando los renglones 1 e i , (operación que cambia el signo al determinante, ver la página 522) podemos suponer que el coeficiente $A_{1,1}$ es distinto de 0. Multipliquemos ahora el primer renglón por $\frac{1}{A_{1,1}}$, es decir apliquemos la operación elemental $\mathcal{M}_{\frac{1}{A_{1,1}},1}$.⁸ Como $\det(\mathcal{M}_{\frac{1}{A_{1,1}},1}(A)) = \frac{1}{A_{1,1}} \det(A)$, para demostrar que $\det(A) \neq 0$ basta ver que $\det(\mathcal{M}_{\frac{1}{A_{1,1}},1}(A)) \neq 0$. Con la ventaja de que $(\mathcal{M}_{\frac{1}{A_{1,1}},1}(A))_{1,1} = 1$.

$$(\mathcal{M}_{\frac{1}{A_{1,1}},1}(A)) = \begin{pmatrix} 1 & * & \dots & * \\ * & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \dots & * \end{pmatrix}$$

Ahora, restando múltiplos adecuados del primer renglón a los demás (operaciones elementales del tercer tipo que no alteran el determinante ni el rango), podemos obtener que haya 0 en cada coeficiente de la matriz debajo del primer coeficiente (pivote) del primer renglón:

$$\begin{pmatrix} 1 & * & \dots & * \\ 0 & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & * & \dots & * \end{pmatrix} =: D,$$

el rango de D sigue siendo n , así que el rango de $\widehat{D}_{1,1} = n - 1$.

Por otra parte, si desarrollamos el determinante de D respecto de la primera columna, obtenemos

$$\det(D) = 1 \cdot \det(\widehat{D}_{1,1}) \neq 0,$$

por hipótesis de inducción. ■

Ahora podemos establecer un criterio para encontrar el rango de una matriz, usando el determinante de matrices menores.

Teorema 148 . *Sea $A \in M_{n \times n}(F) \setminus \{\mathbf{0}\}$, con F un campo. Entonces*

$$\text{rango}(A) = \max \{k \mid \det(B) \neq 0, B \text{ de } k \times k, B \text{ submatriz de } A\}.$$

⁸ $\mathcal{M}_{c,i}$ denota la operación elemental que consiste en multiplicar el renglón i -ésimo de una matriz por c ,

Demostración. Supongamos que B de $k \times k$ es una submatriz cuadrada de A con determinante distinto de 0. Entonces B es invertible y por lo tanto sus k renglones son linealmente independientes. Entonces los k renglones de A que corresponden a los renglones de B son linealmente independientes.⁹

En vista de lo anterior, $\text{rango}(A) \geq k$. Por lo tanto

$$\text{rango}(A) \geq \max \{k \mid \det(B) \neq 0, B \text{ de } k \times k, B \text{ submatriz de } A\}.$$

Por otra parte, supongamos que $r = \text{rango}(A)$, así que el espacio de renglones de A tiene una base B formada por r de sus renglones. Denotemos X al conjunto de los renglones de A que no están en B . La matriz $\widehat{A_{X,\emptyset}}$ tiene rango r pues sus r renglones son linealmente independientes. Reduciendo y escalonando esta matriz, obtenemos una matriz R con r renglones distintos de $\vec{0}$. Esta matriz contiene una submatriz identidad de $r \times r$ (simplemente tache las columnas que no contengan un pivote de un renglón), entonces

$$\widehat{R_{\emptyset,Y}} = I_r,$$

para algún conjunto Y de columnas de R .

Entonces existen matrices elementales E_1, \dots, E_m tales que

$$E_m \cdot \dots \cdot E_1 \cdot \widehat{A_{X,\emptyset}} = R.$$

Observemos ahora que

$$E_1 \cdot \widehat{A_{\emptyset,j}} = (\widehat{E_1 \cdot A})_{\emptyset,j}$$

$${}^{10} \text{(en general } E_1 \cdot \widehat{A_{\emptyset,Y}} = (\widehat{E_1 \cdot A})_{\emptyset,Y}).$$

⁹Suponga que $B = \widehat{A_{x,Y}}$ se obtuvo tachando en A los renglones en X y las columnas en Y . Lo que estamos afirmando es que $\widehat{A_{x,\emptyset}}$ tiene sus renglones linealmente independientes: si así no fuera, alguno de sus renglones sería combinación lineal de los renglones anteriores, pero entonces un renglón de B sería combinación lineal de los anteriores, contradiciendo la elección de B .

¹⁰ya que tachar una columna y hacer una operación elemental de renglón es lo mismo que primero hacer la operación elemental de renglón y después tachar la columna:

Entonces

$$\begin{aligned}
 E_m \cdot \dots \cdot E_1 \cdot \widehat{A_{X,Y}} &= \\
 E_m \cdot \dots \cdot E_1 \cdot \left(\widehat{\left(\widehat{A_{X,\emptyset}} \right)_{\emptyset,Y}} \right) &= \\
 \left(E_m \cdot \dots \cdot \widehat{E_1 \cdot \left(\widehat{A_{X,\emptyset}} \right)} \right)_{\emptyset,Y} &= \\
 \widehat{R_{\emptyset,Y}} &= \\
 &= I_r,
 \end{aligned}$$

esto quiere decir que $\widehat{A_{X,Y}} \in M_{r \times r}(F)$ es invertible y por lo tanto $\det(\widehat{A_{X,Y}}) \neq 0$.

Por lo tanto

$$\text{rango}(A) \leq \max \{k \mid \det(B) \neq 0, B \text{ de } k \times k, B \text{ submatriz de } A\}.$$

Como ya demostramos la desigualdad inversa, tenemos que

$$\text{rango}(A) = \max \{k \mid \det(B) \neq 0, B \text{ de } k \times k, B \text{ submatriz de } A\}.$$

■

Note que en la demostración del teorema anterior, $\text{rango}(A)$ denota el **rango de renglón** de A . Usando que el rango de columna de A es el rango de renglón de A^t , el teorema anterior nos dice que rango de columna de A coincide con el rango de renglón de A^t . Además

$$\text{rango de columna de } A =$$

$$\begin{aligned}
 &= \text{rango}(A^t) = \\
 &= \max \{k \mid \det(B) \neq 0, B \text{ de } k \times k, B \text{ submatriz de } A^t\} = \\
 &= \max \{k \mid \det(B^t) \neq 0, B \text{ de } k \times k, B \text{ submatriz de } A\} = \\
 &= \max \{k \mid \det(B) \neq 0, B \text{ de } k \times k, B \text{ submatriz de } A\} = \\
 &= \text{rango}(A)
 \end{aligned}$$

(donde hemos usado que una submatriz de A^t es la transpuesta de una submatriz de A y que el rango de una matriz coincide con el rango de su transpuesta) entonces tenemos una nueva demostración de que el rango de renglón de una matriz coincide con su rango de columna.

Corolario 20 . *El rango de renglón de una matriz es igual a su rango de columna.*

Demostración. Discusión anterior. ■

Ejercicio 390 . *Refiriéndonos a la demostración del teorema anterior.*

1. Demuestre que $E_1 \cdot \widehat{A_{\emptyset,j}} = (\widehat{E_1 \cdot A})_{\emptyset,j}$. (Sugerencia use que $(AB)^j = AB^j$).
2. Demuestre que $E_m \cdot \dots \cdot E_1 \cdot \widehat{A_{\emptyset,Y}} = (\widehat{E_1 \cdot A})_{\emptyset,Y}$, Y un subconjunto propio del conjunto de las columnas de A .

Ejercicio 391 . *Demuestre que un menor de la matriz A^t es la transpuesta de un menor de A . Explícitamente demuestre que $\widehat{A_{X,Y}^t} = (\widehat{A_{Y,X}})^t$.*

8.11.7 El determinante de un producto de matrices II

Hagamos las siguientes observaciones acerca de como cambia el determinante, al efectuar una operación elemental de renglón (ver página 489).

- Hemos visto que un intercambio de renglones cambia el signo de un determinante.
- Si uno multiplica un renglón por una constante, aplicando la definición de \det , se tiene que el determinante de la nueva matriz queda multiplicada por el escalar.
- Por último, si uno cambia el renglón $A_{\underline{i}}$ por $A_{\underline{i}} + cA_{\underline{j}}$ $j \neq i$, tenemos que $\det(S_{c,j,i}(A)) = \det(A)$.¹¹

Es decir, una operación elemental del tercer tipo no cambia el determinante, como lo muestra el siguiente cálculo:

$$\begin{aligned} \det(S_{c,j,i}(A)) &= \\ &\sum_k (-1)^{i+k} (A_{i,k} + cA_{j,k}) \det((\widehat{A_{i,k} + cA_{j,k}})_{i,k}). \end{aligned}$$

¹¹ $S_{c,j,i}$ denota la operación elemental que consiste en sumar c veces el renglón j -ésimo al i -ésimo.

Notemos ahora que $(A_{i,k} + \widehat{cA_{j,k}})_{i,k} = \widehat{(A_{i,k})_{i,k}}$, entonces

$$\begin{aligned} & \sum_k (-1)^{i+k} (A_{i,k} + cA_{j,k}) \det \left((A_{i,k} + \widehat{cA_{j,k}})_{i,k} \right) \\ &= \sum_k (-1)^{i+k} A_{i,k} \det \left((A_{i,k} + \widehat{cA_{j,k}})_{i,k} \right) + \\ & \quad + \sum_k (-1)^{i+k} cA_{j,k} \det \left((A_{i,k} + \widehat{cA_{j,k}})_{i,k} \right). \end{aligned}$$

Observemos que $\sum_k (-1)^{i+k} cA_{j,k} \det \left((A_{i,k} + \widehat{cA_{j,k}})_{i,k} \right)$ es el determinante de la matriz C que se obtiene al sustituir el renglón i de A por el renglón j . De tal manera que $C_{\underline{i}} = cA_{\underline{j}}$ y $C_{\underline{j}} = A_{\underline{j}}$ entonces $\det(C) = c \cdot 0 = 0$. (Recuerde que una matriz con dos renglones iguales tiene determinante 0, y que al multiplicar un renglón por c el determinante también se multiplica por c).

Entonces

$$\begin{aligned} \det(S_{c,j,i}(A)) &= \\ \sum_k (-1)^{i+k} A_{i,k} \det \left((A_{i,k} + \widehat{cA_{j,k}})_{i,k} \right) &= \\ \sum_k (-1)^{i+k} A_{i,k} \det \left(\widehat{(A_{i,k})_{i,k}} \right) &= \\ &= \det(A). \end{aligned}$$

Ejemplo 187 . Si una matriz tiene un renglón cero, entonces su determinante es 0. Esto se puede comprobar de varias maneras, por ejemplo, desarrollando el determinante respecto del renglón $\vec{0}$.

Ejercicio 392 . El determinante de una matriz triangular es el producto de los elementos de su diagonal.

Sugerencia; Inducción sobre el número de renglones y desarrollo respecto al primer renglón (o columna).

Definición 143

1. Se dice que A es una matriz elemental si se obtiene al aplicarle una operación elemental a la matriz identidad.

2. Se dice que la matriz elemental A es del mismo tipo que el de la operación elemental que se aplicó a la matriz identidad para producirla.

Realizar una operación elemental de renglón a una matriz A equivale a multiplicarla por una matriz elemental:

$$\mathcal{R}(A) = (\mathcal{R}(I_n))A,$$

donde estamos suponiendo que $n =$ número de renglones de A , I_n es la matriz identidad de $n \times n$ y \mathcal{R} es una operación elemental.

Ejercicio 393 . *Demuestre la afirmación anterior. (Sugerencia: proceda por casos, según el tipo).*

Observación 126 . *Toda matriz invertible es un producto de matrices elementales.*

Demostración. Como una matriz A de $n \times n$ invertible entonces tiene rango n . Al reducirla y escalarla por medio de operaciones elementales de renglón obtenemos la matriz I_n . Es decir:

$$I_n = \mathcal{R}_s \dots \mathcal{R}_1(A) = (\mathcal{R}_s(I_n)) \cdot \dots \cdot (\mathcal{R}_1(I_n)) \cdot (A),$$

entonces

$$A = ((\mathcal{R}_s(I_n)) \cdot \dots \cdot (\mathcal{R}_1(I_n)))^{-1} = (\mathcal{R}_1(I_n))^{-1} \cdot \dots \cdot (\mathcal{R}_s(I_n))^{-1}$$

que muestra que A , es un producto de matrices elementales. ■

Veamos ahora como es el determinante de una matriz elemental.

1. $\det((I_n)) = -1.$ ¹²
2. $\det(\mathcal{M}_{c,i}(I_n)) = c.$
3. $\det(\mathcal{S}_{c,i,j}(I_n)) = 1.$

¹² $\mathcal{I}_{i,j}$ es la operación elemental que consiste en intercambiar los renglones i -ésimo y j -ésimo de una matriz.

Demostración. 1) Nótese que $\det(I_n) = 1$ (usando la definición, o por inducción sobre n).

$\mathcal{I}_{i,j}(I_n)$ es la matriz que se obtiene intercambiando dos renglones en la matriz identidad. Por lo tanto su determinante es -1 .

2) Ya hemos observado que al multiplicar un renglón por un escalar c , el determinante también se multiplica por c .

3) Esto porque sumar a un renglón un múltiplo de otro renglón, el determinante no cambia. ■

Corolario 21. *Si E es una matriz elemental de $n \times n$, y A es una matriz de $n \times n$, entonces $\det(EA) = \det(E) \det(A)$.*

Demostración. Comprobaremos la igualdad anterior, según el tipo de E .

Si E es de tipo 1, entonces su determinante es -1 . Por otra parte EA es una matriz que se obtiene al intercambiar dos renglones de A , por lo que

$$\det(EA) = -\det(A) = \det(E) \det(A).$$

Si E es de tipo 2, entonces $E = \mathcal{M}_{c,i}(I_n)$, así que su determinante es c . Por otra parte, $\det \mathcal{M}_{c,i}(A) = c \det(A)$, pues $\mathcal{M}_{c,i}(A)$ se obtiene al multiplicar por c el renglón i de A .

Por último, si E es de tipo 3, entonces su determinante es 1, y por otra parte una operación elemental del tercer tipo no cambia el valor de un determinante. Por lo tanto,

$$\det(EA) = \det(A) = \det(E) \det(A).$$

■

Observación 127. *Si $A = E_1 \cdot \dots \cdot E_m$ es un producto de matrices elementales, entonces $\det(A) = \det(E_1) \cdot \dots \cdot \det(E_m)$.*

Demostración. Por inducción sobre m .

La base es trivial.

Si $m > 1$ podemos usar el corolario anterior para escribir

$$\begin{aligned} \det(A) &= \det(E_1 \cdot (E_2 \cdot \dots \cdot E_m)) \\ &= \det(E_1) \det(E_2 \cdot \dots \cdot E_m) \\ &= \det(E_1) \det(E_2) \cdot \dots \cdot \det(E_m), \end{aligned}$$

en donde hemos usado que

$$\det(E_2 \cdot \dots \cdot E_m) = \det(E_2) \cdot \dots \cdot \det(E_m),$$

por hipótesis de inducción. ■

Observación 128 . $(AB)_{\underline{i}} = A_{\underline{i}}B$. Es decir: el i -ésimo renglón de AB se obtiene multiplicando el i -ésimo renglón de A por B .

Demostración.

$$(AB)_{\underline{i}} = (A_{\underline{i}}B^1, A_{\underline{i}}B^2, \dots, A_{\underline{i}}B^n),$$

donde B^j es la columna j de B y

$$A_{\underline{i}}B^j = (A_{i,1}, \dots, A_{i,n}) \begin{pmatrix} B_{1,j} \\ \vdots \\ B_{n,j} \end{pmatrix} = \sum_k A_{i,k} B_{k,j}.$$

Pero también

$$A_{\underline{i}}B = (A_{\underline{i}}B^1, A_{\underline{i}}B^2, \dots, A_{\underline{i}}B^n).$$

■

Estamos en condiciones de demostrar que $\det(AB) = \det(A)\det(B)$.

Teorema 149 . $\det(AB) = \det(A)\det(B)$, si $A, B \in M_{n \times n}(F)$.

Demostración. Analizaremos dos casos:

1) Si A no es invertible, entonces su rango es menor que n . Así que uno de sus renglones es combinación lineal de los anteriores. Usando el resultado anterior, tenemos que

$$(AB)_{\underline{i}} = A_{\underline{i}}B,$$

así que si el primer renglón de A es $\vec{0}$, entonces también el primer renglón de AB también es $\vec{0}$, y si el renglón j ($j > 1$) de A es combinación lineal de los

anteriores, $A_{\underline{j}} = \sum_{k < j} A_{\underline{k}}$ entonces

$$\begin{aligned}(AB)_{\underline{j}} &= A_{\underline{j}}B \\ &= \left(\sum_{k < j} A_{\underline{k}} \right) B \\ &= \sum_{k < j} A_{\underline{k}}B \\ &= \sum_{k < j} (AB)_{\underline{k}}.\end{aligned}$$

Entonces un renglón de AB es combinación lineal de los anteriores, por lo que el rango de AB es menor que n . Por lo tanto

$$\det(AB) = 0 = 0 \cdot \det(B) = \det(A) \det(B).$$

2) Si A es invertible, podemos expresar A como producto de matrices elementales:

$$A = E_1 \cdot \dots \cdot E_s.$$

Entonces

$$\begin{aligned}\det(AB) &= \\ &= \det(E_1 \cdot \dots \cdot E_s \cdot B) \\ &= \det(E_1) \dots \det(E_s) \det(B) \\ &= \det(A) \det(B),\end{aligned}$$

en donde hemos usado el Corolario 21 y la observación 127. ■

Ejercicio 394 . *Demuestre que la columna j de AB es*

$$(AB)_{\underline{j}}^j = AB_{\underline{j}}^j.$$

Ejemplo 188 . *Una operación elemental de columna de tipo 3 no cambia el determinante.*

Demostración. Sea A una matriz de $n \times n$, denotemos $\mathcal{S}_{c,i,j}^t$ la operación que consiste en sumar a la columna j c veces la columna i . Entonces $\mathcal{S}_{c,i,j}^t(A) = (\mathcal{S}_{c,i,j}(A^t))^t$. Pues para efectuar la operación de columna, podemos hacer la operación de renglón correspondiente a A^t y después transponemos. Entonces

$$\begin{aligned}\det(\mathcal{S}_{c,i,j}^t(A)) &= \det((\mathcal{S}_{c,i,j}(A^t))^t) \\ &= \det(\mathcal{S}_{c,i,j}(A^t)) \\ &= \det((A^t)) = \det(A).\end{aligned}$$

donde se usó dos veces que el determinante de una matriz es igual al determinante de su transpuesta. ■

Ejercicio 395 . Demuestre que una operación elemental de columna de tipo 1, cambia el signo de un determinante.

Ejercicio 396 . Enuncie y demuestre la proposición correspondiente para operaciones elementales de columna del tipo 2.

8.11.8 Matrices invertibles y determinantes

Consideremos la fórmula para el desarrollo del determinante respecto del renglón i :

$$\det(A) = \sum_j (-1)^{i+j} A_{i,j} \det(\widehat{A_{i,j}})$$

Definamos ahora el cofactor de $A_{i,j}$,

$$C_{i,j} = (-1)^{i+j} \det(\widehat{A_{i,j}}),$$

y esto define la matriz C de cofactores de A .

Notemos que

$$\begin{aligned}\det(A) &= \sum_j (-1)^{i+j} A_{i,j} \det(\widehat{A_{i,j}}) \\ &= \sum_j A_{i,j} C_{i,j} \\ &= \sum_j A_{i,j} (C_{j,i})^t \\ &= (AC)_{i,i}.\end{aligned}$$

Así que

$$AC^t = \begin{pmatrix} \det(A) & * & \dots & * \\ * & \det(A) & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \dots & \det(A) \end{pmatrix}.$$

Veamos que todo coeficiente fuera de la diagonal en la matriz anterior es 0.

$$\begin{aligned} (AC^t)_{i,j} &= \sum_k A_{i,k} C_{k,j}^t = \sum_k A_{i,k} C_{j,k} = \\ &= \sum_k A_{i,k} (-1)^{j+k} \det(\widehat{A_{j,k}}). \end{aligned}$$

Esta última expresión es el desarrollo respecto a la j -ésimo renglón de la matriz B que se obtiene al cambiar el renglón j por A_i , así que

$$B = \left. \begin{array}{c} A_1 \\ \vdots \\ \underline{A_i} \\ \vdots \\ \underline{A_i} \\ \vdots \\ A_n \end{array} \right\}_i,$$

note también que $\widehat{B_{j,k}} = \widehat{A_{j,k}}$ y que B tiene dos renglones iguales.

Por lo tanto

$$(AC^t)_{i,j} = \det(B) = 0.$$

Resumimos nuestras observaciones en le siguiente teorema.

Teorema 150 . $AC^t = \det(A) \cdot I_n$.

Corolario 22 . $A \in M_{n \times n}(F)$ es invertible $\Rightarrow A^{-1} = \frac{1}{\det(A)} C^t$.

Demostración. Si A es invertible, entonces su determinante es distinto de 0, ahora podemos usar el teorema 150, para obtener

$$A \left(\frac{1}{\det(A)} C^t \right) = I_n.$$

Así que $\left(\frac{1}{\det(A)}C^t\right)$ es inverso derecho de A^t . Como A es invertible, entonces

$$AA^{-1} = I_n = A \left(\frac{1}{\det(A)}C^t \right) \Rightarrow A^{-1} = \left(\frac{1}{\det(A)}C^t \right).$$

■

Notemos también que

$$\det(A) \det \left(\left(\frac{1}{\det(A)}C^t \right) \right) = 1,$$

así que tenemos que

$$\det(C^t) \left(\frac{1}{\det(A)} \right)^n \det(A) = 1$$

y

$$\det(C) = \det(C^t) = (\det(A))^{n-1}.$$

Si denotamos D la matriz de cofactores C de A , (A invertible) entonces

$$C \left(\frac{1}{\det(C)} D^t \right) = I_n = C \left(\frac{1}{(\det(A))^{n-1}} D^t \right).$$

Entonces $C^{-1} = \frac{1}{(\det(A))^{n-1}} D^t$, tomando transpuestas tenemos que

$$(C^{-1})^t = \frac{1}{(\det(A))^{n-1}} D,$$

es decir que

$$D = (\det(A))^{n-1} (C^{-1})^t. \quad (8.26)$$

Por otra parte, $A^{-1} = \left(\frac{1}{\det(A)} C^t \right)$, tenemos que

$$A = \left(\frac{1}{\det(A)} C^t \right)^{-1} = \det(A) (C^t)^{-1} = \det(A) (C^{-1})^t \quad (8.27)$$

(simplemente realice el producto:

$$\frac{1}{\det(A)} C^t \det(A) (C^{-1})^t = C^t (C^{-1})^t = (C^{-1} C)^t = I_n.$$

De las expresiones 8.26 y 8.27 tenemos que

$$D = (\det(A))^{n-1} (C^{-1})^t = (\det(A))^{n-2} A.$$

que relaciona la matriz de cofactores de la matriz de cofactores de A con A .

8.11.9 La regla de Cramer

Supongamos que

$$A\vec{x} = \vec{b} \quad (8.28)$$

es un sistema de n ecuaciones con n incógnitas con coeficientes en el campo F . Si A es invertible, el sistema anterior tiene la solución única $A^{-1}\vec{b} = \vec{s} =$

$$\begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix} \in F^n.$$

Teorema 151 (Regla de Cramer). *Sea $\vec{s} = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix}$ la solución de 8.28.*

Entonces

$$s_i = \frac{\det \left(\underbrace{\begin{matrix} A^1 & \dots & \overset{i}{\vec{b}} & \dots & A^n \end{matrix}}_{\text{Matriz con } i\text{-ésima columna reemplazada por } \vec{b}} \right)}{\det(A)},$$

donde $\left(\underbrace{\begin{matrix} A^1 & \dots & \vec{b} & \dots & A^n \end{matrix}}_{\text{Matriz que se obtiene al sustituir la } i\text{-ésima columna de } A \text{ por } \vec{b}} \right)$ es la matriz que se obtiene al sustituir la i -ésima columna de A por \vec{b} .

Demostración. Que \vec{s} sea solución de 8.28, también se puede expresar de la manera siguiente:

$$\vec{b} = s_1 A^1 + \dots + s_i A^i + s_n A^n.$$

Entonces

$$\begin{aligned} & \det \left(\underbrace{\begin{matrix} A^1 & \dots & \overset{i}{s_1 A^1 + \dots + s_i A^i + s_n A^n} & \dots & A^n \end{matrix}}_{\text{Matriz con } i\text{-ésima columna reemplazada por } s_1 A^1 + \dots + s_i A^i + s_n A^n} \right) \\ &= \det \left(\underbrace{\begin{matrix} A^1 & \dots & \overset{i}{s_i A^i} & \dots & A^n \end{matrix}}_{\text{Matriz con } i\text{-ésima columna reemplazada por } s_i A^i} \right) \end{aligned}$$

ya que por medio de operaciones elementales de columna podemos cambiar $s_1A^1 + \dots + s_iA^i + s_nA^n$ por s_iA^i .

Entonces

$$\begin{aligned} & \det \left(\begin{array}{c|ccccc} & \overbrace{A^1 \ \dots \ A^n}^i & \vec{b} & \dots & A^n \end{array} \right) \\ &= \det \left(\begin{array}{c|ccccc} & \overbrace{A^1 \ \dots \ s_iA^i}^i & \dots & & A^n \end{array} \right) \\ &= s_i \det \left(\begin{array}{c|ccccc} & \overbrace{A^1 \ \dots \ A^i}^i & \dots & & A^n \end{array} \right) \\ &= s_i \det(A). \end{aligned}$$

Dividiendo entre $\det(A)$ obtenemos el resultado. ■

8.11.10 Determinantes y funciones multilineales

Definición 144

1. Sea $\{V_i\}_{i \in \{1, \dots, n\}}$ una familia de espacios vectoriales sobre un campo F . Una función $f : V_1 \times \dots \times V_n \rightarrow F$ es lineal respecto al i -ésimo factor si la composición

$$\begin{aligned} V_i &\hookrightarrow V_1 \times \dots \times V_n \xrightarrow{f} F \\ x_i &\mapsto (v_1, \dots, x_i, \dots, v_n) \mapsto f((v_1, \dots, x_i, \dots, v_n)) \end{aligned}$$

es lineal para cada $(v_1, \dots, v_i, \dots, v_n) \in V_1 \times \dots \times V_n$.

2. f es n -lineal si es lineal en cada uno de sus factores.

Definición 145 . Una función $f : V \times \dots \times V \rightarrow F$ n -lineal, es alternante si $f((v_1, \dots, v_i, \dots, v_n)) = 0$, para cada $(v_1, \dots, v_i, \dots, v_n)$ que tenga dos coordenadas consecutivas iguales.

Observación 129 . Si f es alternante, entonces

$$f((v_1, \dots, v_i, v_{i+1}, \dots, v_n)) = -f((v_1, \dots, v_{i+1}, v_i, \dots, v_n))$$

Demostración.

$$\begin{aligned}
0 &= f((v_1, \dots, v_i + v_{i+1}, v_i + v_{i+1}, \dots, v_n)) \\
&= f((v_1, \dots, v_i, v_i + v_{i+1}, \dots, v_n)) + f((v_1, \dots, v_{i+1}, v_i + v_{i+1}, \dots, v_n)) \\
&= f((v_1, \dots, v_i, v_i, \dots, v_n)) + f((v_1, \dots, v_i, v_{i+1}, \dots, v_n)) + \\
&\quad + f((v_1, \dots, v_{i+1}, v_i, \dots, v_n)) + f((v_1, \dots, v_{i+1}, v_{i+1}, \dots, v_n)) \\
&= 0 + f((v_1, \dots, v_i, v_{i+1}, \dots, v_n)) + f((v_1, \dots, v_{i+1}, v_i, \dots, v_n)) + 0.
\end{aligned}$$

Donde se ha usado la alternancia y la n -linealidad de f . ■

Corolario 23 . Si \vec{v}_i es paralelo a \vec{v}_j y f es alternante, entonces

$$f(\vec{v}_1, \dots, \vec{v}_i, \dots, \vec{v}_j, \dots, \vec{v}_n) = 0.$$

Corolario 24 . $f(\vec{v}_1, \dots, \vec{v}_i, \dots, \vec{v}_j, \dots, \vec{v}_n) = f(\vec{v}_1, \dots, \vec{v}_i, \dots, c\vec{v}_i + \vec{v}_j, \dots, \vec{v}_n)$.

Ejercicio 397. Demuestre los dos corolarios anteriores.

Teorema 152. Sean $f: V \times V \times \dots \times V \rightarrow F$ alternante y $\beta = \{\vec{v}_1, \dots, \vec{v}_i, \dots, \vec{v}_j, \dots, \vec{v}_n\}$ una base para V , (supongamos que $\vec{x}_i = \sum_k a_{i,k} \vec{v}_k$) entonces

$$\begin{aligned}
&f(\vec{x}_1, \dots, \vec{x}_i, \dots, \vec{x}_j, \dots, \vec{x}_n) \\
&= f\left(\sum_k a_{1,k} \vec{v}_k, \dots, \sum_k a_{i,k} \vec{v}_k, \dots, \sum_k a_{j,k} \vec{v}_k, \dots, \sum_k a_{n,k} \vec{v}_k\right) \\
&= \sum_{\sigma \in S_n} \text{sig}(\sigma) f(\vec{v}_1, \dots, \vec{v}_i, \dots, \vec{v}_j, \dots, \vec{v}_n) \prod a_{i,\sigma(i)}.
\end{aligned}$$

Demostración. Al desarrollar (usando n -linealidad)

$$f\left(\sum_k a_{1,k} \vec{v}_k, \dots, \sum_k a_{i,k} \vec{v}_k, \dots, \sum_k a_{n,k} \vec{v}_k\right)$$

anulando cada sumando que resulte con columnas paralelas (y que por lo tanto vale 0), se ve que sólo quedan $n!$ sumas, una por cada $\sigma \in S_n$:

$$\begin{aligned}
&f\left(\sum_k a_{1,k} \vec{v}_k, \dots, \sum_k a_{i,k} \vec{v}_k, \dots, \sum_k a_{n,k} \vec{v}_k\right) \\
&= \sum_{\sigma \in S_n} a_{1,\sigma(1)} \dots a_{n,\sigma(n)} f(\vec{v}_{\sigma(1)}, \dots, \vec{v}_{\sigma(n)}) \\
&= \sum_{\sigma \in S_n} a_{1,\sigma(1)} \dots a_{n,\sigma(n)} \text{sig}(\sigma) f(\vec{v}_1, \dots, \vec{v}_n).
\end{aligned}$$

■

Corolario 25 . Si $V \times V \times \dots \times V \xrightarrow{f} F$ es alternante y si $f(v_1, \dots, v_n) = 1$ entonces

$$\begin{aligned} & f \left(\sum_k a_{1,k}(\vec{v}_k), \dots, \sum_k a_{i,k}(\vec{v}_k), \dots, \sum_k a_{n,k}(\vec{v}_k) \right) \\ &= \sum_{\sigma \in S_n} a_{1,\sigma(1)} \dots a_{n,\sigma(n)} \operatorname{sig}(\sigma). \end{aligned}$$

Observemos que si $V = F^n$ entonces hay un isomorfismo

$$\begin{aligned} V \times V \times \dots \times V &\xrightarrow{\psi} M_{n \times n}(F), \\ (\vec{v}_1, \dots, \vec{v}_n) &\mapsto (v_1 \dots v_n) \end{aligned} \tag{8.29}$$

$\psi(\vec{v}_1, \dots, \vec{v}_n)$ es la matriz cuyas columnas son $\vec{v}_1, \dots, \vec{v}_n$.

El diagrama

$$\begin{array}{ccc} V \times V \times \dots \times V \rightarrow F & \xrightarrow{f} & F \\ \downarrow \psi & & \parallel \\ M_{n \times n}(F) & \rightarrow & F \end{array}$$

permite reinterpretar una función n -lineal f como una función $M_{n \times n}(F) \rightarrow F$.

Note que si $V \times V \times \dots \times V \rightarrow F \xrightarrow{f} F$ es n -lineal, alternante y si $f(e_1, e_2, \dots, e_n) = 1$, entonces el corolario anterior muestra que

$$\begin{aligned} f \begin{pmatrix} a_{1,1} & a_{2,1} & \dots & a_{n,1} \\ a_{1,2} & a_{2,2} & \dots & a_{n,2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1,n} & a_{2,n} & \dots & a_{n,n} \end{pmatrix} &= \\ &= \sum_{\sigma \in S_n} \operatorname{sig}(\sigma) (a_{1,\sigma(1)} \dots a_{n,\sigma(n)}) = \\ &= \det \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{pmatrix} = \\ &= \det \begin{pmatrix} a_{1,1} & a_{2,1} & \dots & a_{n,1} \\ a_{1,2} & a_{2,2} & \dots & a_{n,2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1,n} & a_{2,n} & \dots & a_{n,n} \end{pmatrix}. \end{aligned}$$

Por lo tanto, podemos enunciar el siguiente teorema:

Teorema 153 Una función $f : M_{n \times n}(F) \rightarrow F$ que interpretada como función de sus columnas sea n -lineal, alternante y tal que $f(I_n) = 1$ es el determinante.

Desde luego, también se tiene que \det es una función n -lineal alternante cuando se piensa como función de $F^n \times \dots \times F^n \rightarrow F$.

8.11.11 Resumen de las propiedades del determinante

- Son equivalentes para $f : M_{n \times n}(F) \rightarrow F$

1. $f = \det$.

2. $f(A) = \sum_{\sigma \in S_n} \left(\text{sig}(\sigma) \left(\prod_{i=1}^n A_{i,\sigma(i)} \right) \right), \forall A \in M_{n \times n}(F)$.

3. $f(A) = \sum_j (-1)^{i+j} A_{i,j} \det(\widehat{A_{i,j}}), \forall A \in M_{n \times n}(F)$.

4. $f(A) = \sum_i (-1)^{i+j} A_{i,j} \det(\widehat{A_{i,j}}), \forall A \in M_{n \times n}(F)$.

5. f pensada como $f : F^n \times \dots \times F^n \rightarrow F$ es n -lineal, alternante y $f \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 1 \end{pmatrix} = 1$.

Otras posibilidades que no hemos demostrado son:

- 6 (a) f es multiplicativa: $f(AB) = f(A)f(B), \forall A, B \in M_{n \times n}(F)$.
 (b) f es una función aditiva respecto de la primera columna de cualquier matriz.
 (c) $f(E) = 1$ para toda matriz elemental E del tercer tipo.
- 7 (a) f es multiplicativa.

(b) $f(A) = \prod_{i=1} A_{i,i}$, para cualquier matriz triangular (inferior o superior).

- El determinante de una matriz con dos renglones (o columnas) iguales es 0.
- Intercambiar dos renglones (o dos columnas) cambia el signo del determinante.
- Sumar a un renglón un múltiplo de otro no cambia el valor del determinante.
- Sumar a una columna un múltiplo de otra no cambia el valor del determinante.
- El determinante de una matriz y el determinante de su transpuesta son iguales.
- El determinante de un producto de matrices es el producto de los determinantes respectivos.
- Una matriz es invertible si y sólo si su determinante es distinto de 0 (en el caso de que los coeficientes se tomen en un campo).
- Una matriz es invertible si y sólo si su determinante tiene inverso multiplicativo en R (en el caso de que los coeficientes se tomen en un anillo comunitativo).
- $\det(A^{-1}) = (\det(A))^{-1}$.
- Si $A \in M_{n \times n}(F)$, es invertible y F es un campo, entonces $A^{-1} = \frac{1}{\det(A)} C^t$, donde C es la matriz de cofactores de A .
- El rango de una matriz no nula A es

$$\max \{k \mid \det(B) \neq 0 \text{ si } B \text{ es submatriz de } A\}.$$

- La regla de Cramer.

Ejercicio 398 Demuestre que las condiciones 6 y 7 anteriores caracterizan la función \det .

1. Demuestre que si R es un anillo commutativo, entonces $M_{n \times n}(R)$ definido de la manera natural es un anillo.
2. Defina $\det(A)$ de manera adecuada para $A \in M_{n \times n}(R)$.
3. Demuestre que $A \in M_{n \times n}(R)$ es invertible si y sólo si $\det(A)$ es invertible en R .

Capítulo 9

Polinomios con coeficientes en \mathbb{R}

El estudio de los polinomios comienza muy frecuentemente definiéndolos como expresiones de la forma $a_0 + a_1x + \dots + a_nx^n$. Así, el conjunto $\mathbb{R}[x] = \{a_0 + a_1x + \dots + a_nx^n \mid n \in \mathbb{N}, a_i \in \mathbb{R}\}$ en donde cada a_i es un “coeficiente” real y x es una “indeterminada”. Los elementos del conjunto anterior -los polinomios- se conocen como “sumas formales” porque su expresión tiene la forma de suma; sin embargo no queda claro lo que significa la x o cualquiera de sus potencias. ¿Cómo se multiplica una indeterminada -o una de sus potencias- por un elemento del campo de los coeficientes? ¿Cómo se suman los “productos” resultantes de esas multiplicaciones?

Con objeto de responder a estas preguntas y a otras que surgen en el estudio de los polinomios, se hace una serie de definiciones de éstos, de la igualdad entre ellos y de la forma en que se suman y multiplican. En esencia, lo que se hace es identificar cada suma formal $a_0 + a_1x + \dots + a_nx^n$ con la sucesión -casi nula- $(a_0, a_1, \dots, a_n, 0, \dots)$ y se construye una inmersión de los coeficientes en los polinomios $a \mapsto (a, 0, \dots)$ y luego se describe la manera de recuperar las expresiones originales, que ahora sí representan sumas auténticas.

Se enfatiza la diferencia entre polinomios y funciones polinomiales, y se hace notar que la forma en que se definieron las operaciones entre los primeros permite construir un homomorfismo de anillos entre ambos, que cuando están definidos en campos infinitos, se convierte en isomorfismo, justificando así el que en el Álgebra elemental, no se haga distinción alguna entre ellos y se transite libremente entre las dos estructuras. (Ver por ejemplo el teorema

del residuo).

Se justifica el nombre “anillo de los polinomios” haciendo notar que se trata en efecto de un dominio entero (anillo conmutativo sin divisores de propios de cero) al igual que los enteros, y se describen algunos resultados que son consecuencia de su estructura.

El capítulo termina con una discusión de la teoría de las ecuaciones y de sus temas centrales, como el teorema del residuo, del factor, el teorema de Sturm, las raíces múltiples, raíces racionales y derivadas entre otras, así como una breve incursión a las desigualdades.

9.1 Construcción y definiciones

Denotemos con \mathbb{R} el campo de los números reales, y con \mathbb{N} al conjunto de los naturales, definimos

$$\mathbb{R}[x] = \mathbb{R}^{(\mathbb{N})} = \{ f \in \mathbb{R}^{\mathbb{N}} \mid f(x) \neq 0 \text{ en un subconjunto finito de } \mathbb{N} \}.$$

¹

Así, los elementos de $\mathbb{R}[x]$ son sucesiones de casi puros 0, que los podemos escribir también en forma de lista:

$$f : f(0), f(1), f(2), \dots, f(n), 0, 0, \dots$$

A partir de algún momento ya sólo aparecerán 0 en la lista lo que podríamos indicar poniendo una barra sobre el primer 0 de la lista infinita de 0 :

$$f : f(0), f(1), f(2), \dots, f(n), \bar{0}, \dots$$

y si $f(n)$ lo denotamos como a_n , la expresión de cada polinomio puede ser $(a_0, a_1, \dots, a_n, \bar{0}, \dots)$.

Definición 146 . *Los elementos de $\mathbb{R}[x]$ se llaman polinomios con coeficientes reales.*

Definición 147 . *Si f es un elemento de $\mathbb{R}[x]$, denotaremos*

$$sop(f) =: \{x \in \mathbb{N} \mid f(x) \neq 0\}.$$

¹El uso de esta notación se aclarará después.

Ejemplo 189. $(-1, 2, 0, \sqrt{2}, 0, 1, \bar{0}, \dots)$ es un polinomio con soporte $\{0, 1, 3, 5\}$.

Definición 148. La suma de polinomios que se define de manera natural:

$$\mathbb{R}[x] \times \mathbb{R}[x] \mathbb{R}[x]$$

está dada por $(f \tilde{+} g)(n) = f(n) + g(n)$.

Observación 130. Esta definición de suma es buena porque

$$n \in \text{sop}(f \tilde{+} g) \Rightarrow f(n) + g(n) \neq 0,$$

$$f(n) + g(n) \neq 0 \Rightarrow (f(n) \neq 0 \vee g(n) \neq 0),$$

$$(f(n) \neq 0 \vee g(n) \neq 0) \Leftrightarrow n \in \text{sop}(f) \vee n \in \text{sop}(g).$$

Así tenemos que $\text{sop}(f \tilde{+} g) \subseteq \text{sop}(f) \cup \text{sop}(g)$. Por lo que $\text{sop}(f \tilde{+} g)$ es un conjunto finito al ser un subconjunto de la unión de dos conjuntos finitos.

Ejercicio 399. Sume los siguientes polinomios

$$(-85, 0, -37, -35, 97, 50, 0, \dots)$$

y :

$$(56, 49, 0, 57, \bar{0}, \dots)$$

Observación 131. $\tilde{+}$ es asociativa y commutativa, pues:

- $(f \tilde{+} (g \tilde{+} h))(n) = f(n) + (g(n) + h(n)) = (f(n) + g(n)) + h(n) = ((f \tilde{+} g) \tilde{+} h)(n).$
- $(f \tilde{+} g)(n) = f(n) + g(n) = g(n) + f(n) = (g \tilde{+} f)(n).$

Observación 132. $\tilde{+}$ tiene neutro:

$$\begin{aligned} \hat{0} : \mathbb{N} &\rightarrow \mathbb{R} \\ n &\mapsto 0 \quad \forall n \in \mathbb{N} \end{aligned}.$$

Observe que $\text{sop}(\hat{0}) = \emptyset$, que es finito.

Observación 133. Cada polinomio tiene inverso respecto a $\tilde{+}$:

$$[-(f)](n) = -f(n).$$

Note que $\text{sop}(-f) = \text{sop}(f)$.

Enseguida vamos a definir en $\mathbb{R}[x]$ una multiplicación.

Definición 149

$$\mathbb{R}[x] \times \mathbb{R}[x] \xrightarrow{*} \mathbb{R}[x]$$

se define por

$$(f * g)(n) = \sum_{i+j=n} f(i)g(j).$$

Proposición 33 . $*$ es una operación asociativa, con neutro, en $\mathbb{R}[x]$:

Demostración. 1. $n \in \text{sop}(f * g) \Rightarrow 0 \neq (f * g)(n) = \sum_{i+j=n} f(i)g(j) \Rightarrow i \in \text{sop}(f) \wedge j \in \text{sop}(g)$ para alguna i y alguna j tales que $i + j = n$. Como $\text{sop}(f)$ y $\text{sop}(g)$ son finitos, también es finito el conjunto de productos $f(i)g(j)$ que sean distintos de 0. Por lo tanto $\text{sop}(f * g)$ es finito.

2. $*$ es asociativa:

$$\begin{aligned} ((f * g) * h)(n) &= \sum_{i+j=n} (f * g)(i)h(j) \\ &= \sum_{i+j=n} [(f * g)(i)]h(j) = \sum_{i+j=n} \left[\sum_{k+l=i} f(k)g(l) \right] h(j) \\ &= \sum_{k+l+j=n} (f(k)g(l))h(j). \end{aligned}$$

Por otra parte si calculamos $(f * (g * h))(n)$ obtendremos

$$\sum_{k+l+j=n} f(k)(g(l)h(j)).$$

3. El neutro es la función $\tilde{1} : \mathbb{N} \rightarrow \mathbb{R}$ tal que $\tilde{1}(0) = 1, \tilde{1}(n) = 0$, para todo $n > 0$. En efecto

$$(\tilde{1} * g)(n) = \sum_{i+j=n} \tilde{1}(i)g(j) = \tilde{1}(0)g(n) = g(n) = (g * \tilde{1})(n).$$

■

Ejercicio 400 . *Multiplique los polinomios :*

$$(2, 0, 3, 0, 2, 1, \bar{0}, \dots)$$

y :

$$(0, 1, 0, 0, 2, 0, \dots)$$

Proposición 34 . ** se distribuye sobre la suma de $\mathbb{R}[x]$.*

Demostración.

$$\begin{aligned} (f * (g + h))(n) &= \sum_{i+j=n} f(i)(g+h)(j) \\ &= \sum_{i+j=n} f(i)(g(j) + h(j)) = \left[\sum_{i+j=n} f(i)g(j) + f(i)h(j) \right] = \\ &= \sum_{i+j=n} f(i)g(j) + \sum_{i+j=n} f(i)h(j) = \\ &= (f * g)(n) + (f * h)(n) = [f * g + f * h](n). \end{aligned}$$

Por lo tanto $(f * (g + h)) = f * g + f * h$. ■

Podemos resumir las propiedades de la suma y del producto de polinomios que hemos visto en la siguiente proposición.

Proposición 35 . $(\mathbb{R}[x], \tilde{+}, \hat{0}, *, \tilde{1})$ es un anillo, el anillo de polinomios con coeficientes en \mathbb{R} .

Definición 150 . $x =: (0, 1, \bar{0}, \dots)$.

Observación 134 . *Como consecuencia de la definición anterior, se tiene que*

$$x^2 =: (0, 0, 1, 0, \bar{0}, \dots), \dots, x^n =: (\underbrace{0, \dots, 0}_{n+1 \text{ lugares}}, 1, 0, \bar{0}, \dots).$$

Además $x^0 = \tilde{1} =: (1, 0, \bar{0}, \dots)$.

Observación 135 . *Si $f : f(0), f(1), \dots, f(n), 0, \bar{0}, \dots$ entonces*

$$f = f(0)\tilde{1} + f(1)x + \dots + f(n)x^n.$$

Es esta propiedad, la que hace que un polinomio se pueda escribir en la forma

$$f = a_0 + a_1 x + \cdots + a_n x^n.$$

Por costumbre, se escribe $f(x)$.

Definición 151 . $\text{grad}(f) =: \max \{k \mid f(k) \neq 0\}$. Notemos que esta definición no incluye al polinomio $\hat{0}$.

Lema 32 . Si $f, g \in \mathbb{R}[x] \setminus \{\hat{0}\}$, entonces $f * g \neq \hat{0}$.

Demostración. Basta ver que $\text{grad}(f * g) = \text{grad}(f) + \text{grad}(g)$:

Si $\text{grad}(f) = n$ y $\text{grad}(g) = m$, entonces $f(n) \neq 0$ y $f(j) = 0 \forall j > n$. Además $g(m) \neq 0$ y $g(j) = 0 \forall j > m$. Entonces

$$(f * g)(n+m) = \sum_{i+j=n+m} f(i)g(j) = f(n)g(m) \neq 0, \text{ ya que } i \geq n \text{ ó } j \geq m.$$

Además si $k \geq n+m$, entonces $(f * g)(k) = \sum_{i+j=k} f(i)g(j) = 0$, ya que $i > n$ ó $j > m$. ■

Teorema 154 . La función $\psi : \mathbb{R} \rightarrow \mathbb{R}[x]$ definida por $\psi(r) = \tilde{r}$, la función $(r, 0, \bar{0}, \dots)$ es una función que respeta la suma, el producto el uno, y además es inyectiva. Es decir:

1. $\psi(r+s) = \psi(r) + \psi(s)$ para cualesquiera $r, s \in \mathbb{R}$.
2. $\psi(r \cdot s) = \psi(r) * \psi(s)$ para cualesquiera $r, s \in \mathbb{R}$.
3. $\psi(1) = \tilde{1}$.
4. $\psi(r) = \psi(s) \Rightarrow r = s$.

Por lo que ψ es una inmersión, lo que permite considerar a \mathbb{R} como un subconjunto de \mathbb{C} .

Demostración. 1. Sean $r, s \in \mathbb{R}$, entonces

$$\begin{aligned} \psi(r+s) &= \widetilde{r+s} = (r+s, 0, \bar{0}, \dots) = (r, 0, \bar{0}, \dots) + (s, 0, \bar{0}, \dots) = \\ &= \tilde{r} + \tilde{s} = \psi(r) + \psi(s) \end{aligned}$$

2. Sean $r, s \in \mathbb{R}$, entonces

$$\begin{aligned}\psi(r * s) &= \widetilde{r * s} = (r \cdot s, 0, \bar{0}, \dots) = (r, 0, \bar{0}, \dots) * (s, 0, \bar{0}, \dots) = \\ &= \tilde{r} * \tilde{s} = \psi(r) * \psi(s)\end{aligned}$$

$$3. \psi(1) = \hat{1} = 1_{\mathbb{R}[x]}.$$

$$4. \psi(r) = \psi(s) \Rightarrow (r, 0, \bar{0}, \dots) = (s, 0, \bar{0}, \dots) \Rightarrow r = s. \blacksquare$$

Proposición 36 . $(\mathbb{R}[x], \tilde{+}, \hat{0}, *, \tilde{1})$ es un dominio entero.

Demostración. Basta ver que $(\mathbb{R}[x] \setminus \{\hat{0}\}, *, \tilde{1})$ es un monoide con cancelación.

Supongamos que

$$f * g = f * h, f, g, h \in \mathbb{R}[x] \setminus \{\hat{0}\}.$$

Entonces $f * (g - h) = \hat{0}$. Como $f \neq \hat{0}$, tenemos que $g - h = \hat{0}$, por el Lema 32. ■

Proposición 37 . En $(\mathbb{R}[x], \tilde{+}, \hat{0}, *, \tilde{1})$ hay algoritmo de la división:
 Si $f(x), g(x) \in \mathbb{R}[x]$ y $g(x) \neq \hat{0}$, entonces

$$\exists!q(x), !r(x) \in \mathbb{R}[x]$$

tales que

$$0 = r(x) \text{ ó } \text{grad}(r(x)) < \text{grad}(q), \text{ se } f(x) = bq + r.$$

Como la proposición anterior es un poco larga, es mejor escribir el diagrama siguiente:

$$g(x) \quad \frac{q(x)}{r(x)} \quad \hat{0} = r(x) \text{ ó } grad(r(x)) < grad(g(x))$$

Demostración. Podemos suponer que $f(x) \neq \hat{0}$, pues en caso contrario, $q(x) = 0 = r(x)$, sirven.

Ahora, podemos hacerlo por inducción sobre $\text{grad}(f(x))$:

Base.

Si $\text{grad}(f) = 0 = \text{grad}(g)$, entonces

$$g \left| \begin{array}{c} f/g \\ \hline f \\ \hat{0} \end{array} \right. \hat{0} = r(x)$$

Si $\text{grad}(f) < \text{grad}(g)$, entonces

$$g \left| \begin{array}{c} \hat{0} \\ \hline f \\ f \end{array} \right. f = r(x), \text{ grad}(f) < \text{grad}(g)$$

Paso inductivo:

Si $\text{grad}(f) < \text{grad}(g)$ entonces

$$g \left| \begin{array}{c} \hat{0} \\ \hline f \\ f \end{array} \right. f = r(x)$$

Si $\text{grad}(f) \geq \text{grad}(g)$, escribamos

$$f = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$$

y

$$g = b_0 + b_1x + b_2x^2 + \dots + b_mx^m$$

Multipliquemos g por x^{n-m} . Entonces $f - x^{n-m}g = \hat{0}$ ó $\text{grad}(f - x^{n-m}g) < \text{grad}(f)$. En el primer caso tenemos

$$g \left| \begin{array}{c} x^{n-m} \\ \hline f \\ \hat{0} \end{array} \right. \hat{0} = r(x)$$

y en el segundo tenemos que por hipótesis de inducción

$$g \left| \begin{array}{c} q \\ \hline f - x^{n-m}g \\ r(x) \end{array} \right. \hat{0} = r(x) \text{ ó } \text{grad}(r) < \text{grad}(g)$$

En este último caso, $f - x^{n-m}g = qg + r$, es decir que

$$g \left| \begin{array}{c} x^{n-m} + q \\ \hline f \\ r(x) \end{array} \right. 0 = r(x) \text{ ó } \text{grad}(r) < \text{grad}(g)$$

La demostración de la unicidad se deja como ejercicio. ■

Ejercicio 401 *vEfectúe las siguientes divisiones:*

1. $x^3 + 3x^2 + 3x + 1 \overline{)x^6 + 4x^5 + 2x^3 + 3x}$,
2. $3x^3 + x^2 + 3x + 4 \overline{)27x^6 + 54x^5 + 162x^4 + 54x^3 + 81x^2 + 81x + 54}$.

9.2 Evaluación

Observación 136 . *Dada $a \in \mathbb{R}$ existe una única función $Ev_a : \mathbb{R}[x] \rightarrow \mathbb{R}$, a la que llamaremos “evaluación en a ”, tal que:*

1. $Ev_a(\hat{r}) = r, \forall r \in \mathbb{R}$.
2. $Ev_a(x) = a$.
3. Ev_a respeta la suma, el producto y el uno.

Demostración. Se deja como ejercicio. ■

Observemos que

$$\begin{aligned} Ev_a(1 + 4x^2 + x^3) &= Ev_a(1) + Ev_a(4x^2) + Ev_a(x^3) = \\ &= 1 + Ev_a(4) \cdot Ev_a(x^2) + (Ev_a(x))^3 = \\ &= 1 + 4 \cdot a^2 + a^3. \end{aligned}$$

Notación 20 . *De aquí en adelante, escribiremos $f(a)$ en lugar de $Ev_a(f(x))$.*

Lo anterior permite asociar a un polinomio f una función polinomial

$$\begin{aligned} Ev_{(\)}(f) : \mathbb{R} &\longrightarrow \mathbb{R} \\ a &\longmapsto Ev_a(f(x)) \end{aligned}.$$

Por costumbre, se identifican los conceptos de polinomio y de función polinomial y para no contravenir esta costumbre, de aquí en adelante, a menos que indique otra cosa, pensaremos en $f(x)$ de la siguiente manera:

$$f : \mathbb{R} \longrightarrow \mathbb{R},$$

en lugar de pensarla como la sucesión de sus coeficientes, $f : \mathbb{N} \longrightarrow \mathbb{R}$.

Ejemplo 190 . Para aclarar la diferencia entre polinomio y función polinomial, imagínese que el campo es \mathbb{Z}_2 en lugar de ser \mathbb{R} , tomemos el polinomio

$$x^3 + x^2,$$

salta a la vista que este polinomio no es el polinomio $0(x)$, ya que tiene grado 3. Sin embargo la función polinomial asociada es

$$\begin{array}{rcl} \mathbb{Z}_2 & \longrightarrow & \mathbb{Z}_2 \\ 0 & \longmapsto & 0^3 + 0^2 = 0 \\ 1 & \longmapsto & 1^3 + 1^2 = 1 + 1 = 0 \end{array}$$

que sí es la función $\hat{0}$.

Proposición 38 . $\forall r, s \in \mathbb{R}, \forall f, g \in \mathbb{R}[x] :$

1. $1 \cdot f = f$.
2. $(rs) \cdot f = r \cdot (s \cdot f)$.
3. $r(f \cdot g) = (rf)g = f(rg)$.

Ejercicio 402 . Si \mathbb{R} está incluido en un anillo S (como subanillo, es decir que las operaciones de S extienden a las de \mathbb{R}) y $a \in S$, entonces $\exists! Ev_a : \mathbb{R}[x] \rightarrow S$ tal que el siguiente diagrama commuta

$$\begin{array}{ccccc} & \mathbb{R}[x] & & & \\ \nearrow & \circlearrowleft & \searrow & \nearrow & \\ R & \hookrightarrow & S & \xrightarrow{Ev_a} & a \end{array} \quad .$$

Es claro que $Ev_a(r) = r, \forall r \in \mathbb{R}[x]$, tal que $r = \hat{0}$, o $\text{grad}(r) = 0$. Además

$$Ev_a(rx^n) = Ev_a(r) \cdot Ev_a(x^n) = r \cdot Ev_a(x)^n = r \cdot a^n.$$

Así que

$$Ev_a(a_0 + a_1x + \dots + a_nx^n) = (a_0 + a_1a + \dots + a_na^n).$$

(Ver el ejercicio siguiente).

Ejercicio 403 . Demuestre que $\mathbb{R}[x] \xrightarrow{Ev_a} S$ respeta la suma, el producto y 1.

Sugerencia: para demostrar que Ev_a respeta productos, es decir, para ver que

$$Ev_a(f(x) \circ g(x)) = Ev_a(f(x)) \cdot Ev_a(g(x)),$$

mostrar primero que se cumple para $f(x) = \hat{0}$, luego para $f(x) = c \cdot x^n$. Y después hacerlo por inducción sobre el grado de $f \neq \hat{0}$.

Teorema 155 (del residuo). Sea $f \in \mathbb{R}[x]$. Entonces

$$f(x) = g(x)(x-a) + f(a).$$

Demostración.

$$x-a \quad \overline{f} \quad \begin{array}{l} g(x) \\ \hline f \\ r(x) \end{array} \quad \hat{0} = r(x) \text{ ó } \text{grad}(r) < \text{grad}(x-a).$$

Entonces $r(x)$ es una constante, y $f(x) = g(x)(x-a) + r$. Evaluando en a , obtenemos la conclusión deseada. ■

Definición 152 . Sean $f(x), g(x) \in \mathbb{R}[x]$, diremos que f divide a g ($f \mid g$) si $\exists h(x) \in \mathbb{R}[x]$ tal que $fh = g$.

Definición 153 . Si $f(r) = 0$, decimos que r es una raíz de $f(x)$.

Teorema 156 (del factor). Sea $f \in \mathbb{R}[x]$. Entonces

$$f(a) = 0 \Leftrightarrow (x-a) \mid f(x).$$

Demostración. Por el Corolario anterior, tenemos que $f(x) = g(x)(x-a) + f(a)$.

■

Recordando que los elementos de \mathbb{R} pueden identificarse con elementos de $\mathbb{R}[x]$, definiremos un producto de $\mathbb{R} \times \mathbb{R}[x]$ en $\mathbb{R}[x]$, de manera que coincida con el producto en $\mathbb{R}[x]$.

Definición 154

$$\begin{array}{ccc} \cdot : & \mathbb{R} \times \mathbb{R}[x] & \rightarrow \mathbb{R}[x] \\ & (r, f) & \mapsto \widehat{r} \cdot f \end{array}.$$

Ejemplo 191 . Si $f = (a_0, a_1, \dots, a_n, \bar{0}, \dots)$ y $r \in \mathbb{R}$, entonces

$$rf =: \hat{r}f = (ra_0, ra_1, \dots, ra_n, \bar{0}, \dots).$$

De las definiciones y proposiciones antes demostradas, concluimos que para $f \in \mathbb{R}[x]$ con $f = (a_0, a_1, \dots, a_n, \bar{0}, \dots)$ se tiene que $f(x) = a_0 1 + a_1 x + \dots + a_n x^n$.

Así que

$$\mathbb{R}[x] =: \{a_0 + a_1 x + \dots + a_n x^n \mid n \in \mathbb{N}, a_i \in \mathbb{R}\}.$$

(El anillo de los polinomios con coeficientes en \mathbb{R}).

9.3 Los ideales de $\mathbb{R}[x]$

Definición 155 . Un subconjunto I de $\mathbb{R}[x]$ se llama ideal de $\mathbb{R}[x]$ si

1. $0 \in I$.
2. $f, g \in I \Rightarrow f + g \in I$.
3. $f \in \mathbb{R}[x], g \in I \Rightarrow fg \in I$.
1. El coeficiente principal de un polinomio es a_n si $f(x) = a_0 + \dots + a_n x^n$ y $a_n \neq 0$.
2. Se dice que un polinomio $f(x)$, distinto de cero es mónico si su coeficiente principal es 1.

Teorema 157

$$\begin{aligned} I \leq \mathbb{R}[x] &\Leftrightarrow I = \mathbb{R}[x]g(x) =: \\ &= \{f(x)g(x) \mid f \in \mathbb{R}[x], \text{ para alguna } g(x) \in \mathbb{R}[x], g(x) \text{ mónico}\}. \end{aligned}$$

Demostración. \Leftarrow

$$0 = 0 \cdot g \in I.$$

$$fg, f'g \Rightarrow fg + f'g = (f + f')g \in I.$$

$$f \in \mathbb{R}[x], hg \in I \Rightarrow f(hg) = (fh)g \in I. \therefore I \leq \mathbb{R}[x].$$

\Rightarrow

$$\text{Si } I = \{0\} \text{ entonces } I = \mathbb{R}[x] \cdot 0.$$

Si $I \neq \{0\}$, sea

$$A = \{n \in \mathbb{N} \mid \exists f \in I \text{ tal que } \text{grad}(f) = n\}.$$

Notemos que $A \neq \emptyset$. Escojamos el menor elemento de A (haciendo uso del principio del buen orden) y llamémoslo m . Después tomemos una $g \in I$ tal que $\text{grad}(g) = m$. (Nótese que se puede escoger g con coeficiente principal 1, multiplicando por el recíproco del coeficiente principal, si fuera necesario).

Demostraremos que $I = (\mathbb{R}[x])g$.

\subseteq) Sea $f \in I$. Por el algoritmo de la división, $f = tg + r$ con $r = 0$ ó $\text{grad}(r) < \text{grad}(g)$. Si r fuera distinto de 0 entonces $r = f - tg \in I$ y $\text{grad}(r) < m = \min(A)$.

Por lo tanto $r = 0$ y así $f \in (\mathbb{R}[x])g$.

\supseteq) $g \in I \Rightarrow (\mathbb{R}[x])g \leqslant I$. ■

Escribiremos $I \leqslant \mathbb{R}[x]$, para indicar que I es un ideal de $\mathbb{R}[x]$.

Ejercicio 404. Sea $a \in \mathbb{R}$, demuestre que $\{f \in \mathbb{R}[x] \mid f(a) = 0\}$ es un ideal de $\mathbb{R}[x]$.

Teorema 158. Sean $I, J \leqslant \mathbb{R}[x]$, entonces:

1. $I \cap J \leqslant \mathbb{R}[x]$.
2. $I + J = \{i + j \mid i \in I, j \in J\} \leqslant \mathbb{R}[x]$.
3. $I + J$ es el menor ideal de $\mathbb{R}[x]$ que incluye a $I \cup J$.

Demostración. 1. a. Como I y J son ideales, entonces $0 \in I$, $0 \in J$, por lo que $0 \in I \cap J$.

1.b.

$$\begin{aligned} f, g \in I \cap J &\Rightarrow (f, g \in I) \wedge (f, g \in J) \Rightarrow \\ &\Rightarrow (f + g \in I) \wedge (f + g \in J) \Rightarrow f + g \in I \cap J. \end{aligned}$$

1.c.

$$\begin{aligned} f &\in \mathbb{R}[x], g \in I \cap J \Rightarrow (f \in \mathbb{R}[x], g \in I, g \in J) \Rightarrow \\ &\Rightarrow (fg \in I) \wedge (fg \in J) \Rightarrow fg \in I \cap J. \end{aligned}$$

2.a. $0 \in I, 0 \in J \Rightarrow 0 = 0 + 0 \in I + J$.

2.b. $i + j, i' + j' \in I + J \Rightarrow (i + j) + (i' + j') = (i + i') + (j + j') \in I + J$.

2.c. $h \in \mathbb{R}[x], i + j \in I + J \Rightarrow h(i + j) = hi + hj \in I + J$.

3.a. $I \subseteq I + J$ ya que $\forall i \in I, i = i + 0 \in I + J$. Análogamente, $J \subseteq I + J$.

Por lo tanto

$$I \cup J \subseteq I + J.$$

3.b. Veamos ahora que $I + J$ es el menor ideal que incluye a $I \cup J$:

Si $I \cup J \subseteq K \leq \mathbb{R}[x]$, entonces $\forall i \in I, \forall j \in J, i, j \in K$. Como K es un ideal, $i + j \in K$. Por lo tanto, $I + J \subseteq K$. ■

Definición 156 . $h \in \mathbb{R}[x]$ es el máximo común divisor de f y g si

1. $h | f, h | g$ (es decir, h es un divisor común).
2. $k | f, k | g \Rightarrow k | h$ (cualquier otro divisor común divide a h).
3. h es mónico.

Corolario 26 . Dados $f, g \in \mathbb{R}[x]$, $(\mathbb{R}[x])f + (\mathbb{R}[x])g = (\mathbb{R}[x])h$, para alguna $h \in \mathbb{R}[x]$, mónico. Además h es el máximo común divisor de f y g .

Demostración. A la vista de los resultados anteriores, lo único que requiere demostración es la afirmación de que h es el máximo común divisor de f y g .

1. $f \in \mathbb{R}[x]h \Rightarrow f = th$, para alguna h en $\mathbb{R}[x]$. Es decir, $h | f$. Análogamente $h | g$. Con esto vemos que h es un divisor común de f y de g .

2. Ahora, si k es otro divisor común, $k | f$ y $k | g$, digamos que

$$f = k\phi \text{ y que } g = k\gamma,$$

con ϕ y $\gamma \in \mathbb{R}[x]$. Como también tenemos que

$$h = \alpha f + \beta g,$$

entonces

$$h = \alpha k\phi + \beta k\gamma = k(\alpha\phi + \beta\gamma).$$

Así que $k | h$. Por lo tanto h es el máximo común divisor de f y g . ■

Notación 21 . Denotemos $(f; g)$ el máximo común divisor de f y de g .

Note el punto y coma en lugar de la coma.

Ejercicio 405 . Demuestre que el máximo común divisor de dos polinomios se puede encontrar mediante el Algoritmo de Euclides.

1. Primero demuestre que dados dos polinomios $f(x)$ y $g(x) \neq 0$, entonces dado

$$g \left| \begin{array}{c} q \\ f \\ r \end{array} \right. \quad r = 0 \vee (\text{grad}(r) < \text{grad}(g))$$

se tiene que $(f; g) = (g; r)$.

2. Termine la demostración imitando la demostración correspondiente para los enteros.

Ejemplo 192 . Encontraremos el máximo común divisor de

$$x^4 + 5x^3 - 4x^2 - 2x \text{ y } 3x^4 + x^3 - 4x^2 :$$

$$\begin{array}{r} 3 \\ x^4 + 5x^3 - 4x^2 - 2x \left| \begin{array}{c} 3x^4 + x^3 - 4x^2 \\ -(3x^4 + 15x^3 - 12x^2 - 6x) \\ \hline -14x^3 + 8x^2 + 6x \end{array} \right. \\ - \frac{1}{14}x - \frac{39}{98} \\ \hline -14x^3 + 8x^2 + 6x \left| \begin{array}{c} x^4 + 5x^3 - 4x^2 - 2x \\ - \left(x^4 - \frac{4}{7}x^3 - \frac{3}{7}x^2 \right) \\ \hline \frac{39}{7}x^3 - \frac{25}{7}x^2 - 2x \end{array} \right. \\ - \frac{39}{49}x^3 - \frac{156}{49}x^2 - \frac{117}{49}x \\ \hline -\frac{19}{49}x^2 + \frac{19}{49}x \end{array}$$

Multiplicando $-\frac{19}{49}x^2 + \frac{19}{49}x$ por $-\frac{49}{19} : x^2 - x$

$$\begin{array}{r} -14x - 6 \\ x^2 - x \left| \begin{array}{c} -14x^3 + 8x^2 + 6x \\ -(-14x^3 + 14x^2) \\ \hline -6x^2 + 6x \end{array} \right. \\ - \frac{(-6x^2 + 6x)}{0} \end{array}$$

Entonces el máximo común divisor es $x^2 - x$.

Ejemplo 193 . Usaremos el Algoritmo de Euclides para escribir el máximo común divisor de $x^4 - 2x^3 - x^2 + 2x$ y $x^3 - 3x^2 + 2x + 1$ como combinación de ellos.

$$\begin{array}{r}
 \begin{array}{c} x+1 & x^2 - 4x + 6 & x+1 \\ \hline x+1 & x^3 - 3x^2 + 2x + 1 & x^4 - 2x^3 - x^2 + 2x \\ 0 & -(x^3 + x^2) & -(x^4 - 3x^3 + 2x^2 + x) \\ \hline & -4x^2 + 2x + 1 & x^3 - 3x^2 + x \\ & -(-4x^2 - 4x) & -(x^3 - 3x^2 + 2x + 1) \\ \hline & 6x + 1 & -x - 1 \\ & -(6x + 6) & \\ \hline & -5 & \end{array}
 \end{array}$$

Entonces el máximo común divisor es 1 y

$$\begin{aligned}
 -5 &= (x^3 - 3x^2 + 2x + 1) - (x^2 - 4x + 6)(x + 1) = \\
 -5 &= (x^3 - 3x^2 + 2x + 1) + (x^2 - 4x + 6)(-(x + 1)) = \\
 &= (x^3 - 3x^2 + 2x + 1) + (x^2 - 4x + 6) \left(\begin{array}{c} (x^4 - 2x^3 - x^2 + 2x) - \\ -(x + 1)(x^3 - 3x^2 + 2x + 1) \end{array} \right) = \\
 &= (x^2 - 4x + 6)(x^4 - 2x^3 - x^2 + 2x) + \\
 &\quad + (-x^3 + 3x^2 - 2x - 5)(x^3 - 3x^2 + 2x + 1).
 \end{aligned}$$

No hace falta decir que entonces

$$\begin{aligned}
 1 &= \left(-\frac{1}{5} \right) (x^2 - 4x + 6)(x^4 - 2x^3 - x^2 + 2x) + \\
 &\quad + \left(-\frac{1}{5} \right) (-x^3 + 3x^2 - 2x - 5)(x^3 - 3x^2 + 2x + 1).
 \end{aligned}$$

Ejercicio 406 . Encuentre el máximo común divisor de los siguientes polinomios:

1. $2x^3 + 3x^2 - 8x + 3$, $3x^3 + 4x^2 - 13x + 6$.
2. $x^3 - x^2 - x - 2$, $x^4 - x^3 - 4x^2 + 4x$.

Observación 137 . $\mathbb{R}[x]h = \mathbb{R}[x](c \cdot h)$, $\forall c \in \mathbb{R} \setminus \{0\}$.

Demostración. $h = 1/c(c \cdot h) \Rightarrow h \in \mathbb{R}[x] c \cdot h \Rightarrow \mathbb{R}[x] h \subseteq \mathbb{R}[x] c \cdot h$.
 Análogamente, $(c \cdot h) \in \mathbb{R}[x] \Rightarrow \mathbb{R}[x] c \cdot h \subseteq \mathbb{R}[x] h$. ■

Observación 138 . *El máximo común divisor de dos polinomios es único.*

Demostración. Sean d, d' dos máximos divisores comunes de f y de g .
 d divisor común y d' máximo común divisor $\Rightarrow d \mid d'$.

Por simetría, $d' \mid d$.

$d \mid d' \Rightarrow d\alpha = d'$, para alguna $\alpha \in \mathbb{R}[x]$.

$d' \mid d \Rightarrow d\beta = d$, para alguna $\beta \in \mathbb{R}[x]$.

Entonces $d = d\beta = d\alpha\beta$. Luego $0 = \text{grad}(\alpha\beta) = \text{grad}(\alpha) + \text{grad}(\beta)$.

Por lo que α y β son polinomios constantes.

Como d y d' son mónicos y $d = d\beta$, tienen coeficiente principal 1 y β , tenemos que $\beta = 1$, así que $d = d'$. ■

Definición 157 . Decimos que dos polinomios $f, g \in \mathbb{R}[x]$ son primos relativos si su máximo común divisor es 1.

Ejercicio 407 . Demuestre que dos polinomios f, g son primos relativos \Leftrightarrow existen $\alpha(x), \beta(x) \in \mathbb{R}[x]$ tales que $1 = \alpha(x)f(x) + \beta(x)g(x)$.

Ejercicio 408 . Encuentre el máximo común divisor de

$$x^3 + x^2 + 2x + 3 \text{ y } (x^4 + x^3 + 2x^2 + 6),$$

escriba 19 como combinación de dichos polinomios.

Definición 158 . $m(x) \in \mathbb{R}[x]$ es el mínimo común múltiplo de f y de g (se escribe $m(x) = [f; g]$) si

1. $m(x)$ es un múltiplo común de f y de g : $m(x) \in \mathbb{R}[x]f, m(x) \in \mathbb{R}[x]g$.
2. Si $k(x)$ es otro múltiplo común entonces $k(x)$ es múltiplo de $m(x)$: $k(x) \in \mathbb{R}[x]f, k(x) \in \mathbb{R}[x]g \Rightarrow m(x) \mid k(x)$.
3. $m(x)$ es mónico.

Denotaremos $[f; g]$ el mínimo común múltiplo de f y g . **Note** el punto y coma.

Observación 139 . *Dados $f, g \in \mathbb{R}[x]$, entonces $(\mathbb{R}[x])f \cap (\mathbb{R}[x])g = (\mathbb{R}[x])h$, p. a. $h \in \mathbb{R}[x]$, h mónico. Además h es el mínimo común múltiplo de f y g .*

Demostración. Tenemos que $(\mathbb{R}[x])f \cap (\mathbb{R}[x])g$ es un ideal de $\mathbb{R}[x]$, por el teorema 158, ahora, por el teorema 157, tenemos que $(\mathbb{R}[x])f \cap (\mathbb{R}[x])g = (\mathbb{R}[x])h$ y podemos suponer que h es mónico.

Veamos que h es el mínimo común múltiplo de f y g :

$$h \in (\mathbb{R}[x])f \cap (\mathbb{R}[x])g \Rightarrow$$

$$h = \alpha f \wedge h = \beta g \therefore \begin{matrix} f \mid h \\ g \mid h \end{matrix}, \therefore h \text{ es un múltiplo común de } f \text{ y de } g.$$

Si $k(x)$ es un polinomio tal que

$$\begin{matrix} f \mid k \\ g \mid k \end{matrix} \text{ entonces } k \in (\mathbb{R}[x])f \wedge k \in (\mathbb{R}[x])g.$$

$$\therefore k \in (\mathbb{R}[x])f \cap (\mathbb{R}[x])g = (\mathbb{R}[x])h$$

$$\therefore h \mid k.$$

Por lo que h es el mínimo común múltiplo común de f y de g . ■

Proposición 39 . *El mínimo común múltiplo de dos polinomios mónicos f, g en $\mathbb{R}[x]$ es*

$$\frac{f \cdot g}{(f; g)}.$$

Demostración. Notemos primero que

$$\frac{f \cdot g}{(f; g)} = \frac{f}{(f; g)} \cdot g = f \cdot \frac{g}{(f; g)}$$

es un múltiplo común de f y de g , por lo que

$$[f; g] \mid \frac{f \cdot g}{(f; g)}.$$

Recíprocamente,

$$\begin{aligned} \frac{f \cdot g}{(f; g)} \mid [f; g] &\Leftrightarrow \frac{f \cdot g}{[f; g]} \mid (f; g) \Leftrightarrow \left[\left(\frac{f \cdot g}{[f; g]} \mid f \right) \wedge \left(\frac{f \cdot g}{[f; g]} \mid g \right) \right] \Leftrightarrow \\ &\Leftrightarrow [(f \cdot g \mid f [f; g]) \wedge (f \cdot g \mid g [f; g])] \Leftrightarrow \\ &\Leftrightarrow (g \mid [f; g]) \wedge (f \mid [f; g]), \end{aligned}$$

pero eso es exactamente lo que sucede, pues $[f; g]$ es un múltiplo común de f y de g .

Como

$$\left([f; g] \mid \frac{f \cdot g}{(f; g)} \right) \text{ y } \left(\frac{f \cdot g}{(f; g)} \mid [f; g] \right)$$

tenemos que

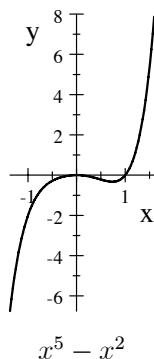
$$\frac{f \cdot g}{(f; g)} = [f; g].$$

pues tanto $[f; g]$ como $\frac{f \cdot g}{(f; g)}$ son polinomios mónicos. ■

Definición 159 . $f \in \mathbb{R}[x]$ es irreducible si:

1. $\text{grad}(f) > 0$.
2. $f = g \cdot h \Rightarrow (\text{grad}(g) = 0 \text{ o } \text{grad}(h) = 0)$.

Una consecuencia del Teorema del Valor intermedio del Cálculo, es que todo polinomio de grado impar tiene alguna raíz. Esto se debe a que los polinomios son funciones continuas (tienen derivada) y a que un polinomio de grado impar tiene valores de signo contrario si se evalúa “muy a la derecha” y “muy a la izquierda”.



$$x^5 - x^2$$

Por el teorema 156 si r es una raíz de $f(x)$ entonces $x - r \mid f(x)$.

De lo anterior, debe ser claro que ningún polinomio de grado impar es irreducible.

Un polinomio $f(x) \in \mathbb{R}[x]$ de grado par ≥ 4 tampoco es irreducible. Esto se debe al Teorema Fundamental del Álgebra y a que la operación de conjugación compleja preserva sumas, productos y reales. Veámoslo: si $z \in \mathbb{C}$ es una raíz compleja de

$$f(x) = a_0 + a_1x + \cdots + a_nx^n,$$

entonces

$$0 = f(z) = a_0 + a_1z + \cdots + a_nz^n,$$

conjugando:

$$\begin{aligned} 0 = \overline{0} &= \overline{f(z)} = \overline{a_0 + a_1z + \cdots + a_nz^n} = \overline{a_0} + \overline{a_1z} + \cdots + \overline{a_nz^n} = \\ &= \overline{a_0} + \overline{a_1}\overline{z} + \cdots + \overline{a_n}\overline{z^n} = a_0 + a_1\overline{z} + \cdots + a_n(\overline{z})^n = \\ &= f(\overline{z}). \end{aligned}$$

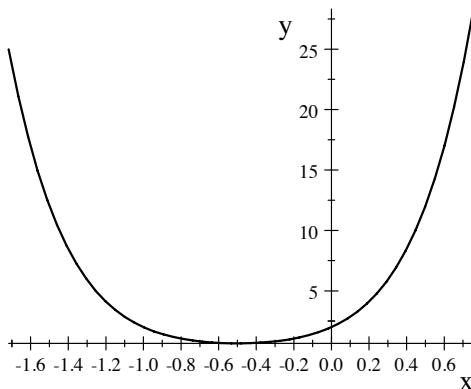
De donde \overline{z} también es raíz de $f(x)$, pero entonces

$$(x - z)(x - \overline{z}) \mid f(x)$$

$$^2 \text{y } (x - z)(x - \overline{z}) = x^2 - (z + \overline{z})x + \|z\|^2 \in \mathbb{R}[x].$$

Ejercicio 409 . Demuestre que si $z \in \mathbb{C}$. Entonces $x^2 - (z + \overline{z})x + \|z\|^2 \in \mathbb{R}[x]$.

²Notar que $x - z$ y $x - \overline{z}$, son polinomios primos relativos en $\mathbb{C}[x]$. Ya que $(x - \overline{z}) - (x - z) = z - \overline{z} \in \mathbb{R} \setminus \{0\}$, cuando $z \in \mathbb{C} \setminus \mathbb{R}$. Entonces $1 = \left(\frac{1}{z - \overline{z}}\right)(x - \overline{z}) - \left(\frac{1}{z - \overline{z}}\right)(x - z)$.



$$2x^6 + 6x^5 + 13x^4 + 16x^3 + 14x^2 + 7x + 2$$

Si $f(x) = 2x^6 + 6x^5 + 13x^4 + 16x^3 + 14x^2 + 7x + 2$, entonces una de sus raíces es

$$\frac{-1+i}{2}$$

y otra es

$$\frac{-1-i}{2}$$

por lo que

$$\left(x - \frac{(-1+i)}{2}\right) \left(x - \frac{(-1-i)}{2}\right) \mid f(x) :$$

en efecto:

$$\left(x - \frac{(-1+i)}{2}\right) \left(x - \frac{(-1-i)}{2}\right) = x^2 + x + \frac{1}{2}$$

y

$$\begin{aligned} (2x^6 + 6x^5 + 13x^4 + 16x^3 + 14x^2 + 7x + 2) &= \\ &= \left(x^2 + x + \frac{1}{2}\right) (2x^4 + 4x^3 + 8x^2 + 6x + 4). \end{aligned}$$

Observación 140 . Por lo anteriormente dicho, tenemos que los únicos polinomios irreducibles en $\mathbb{R}[x]$ son de grado 1 ó 2. Los polinomios irreducibles de grado 2 son de la forma

$$ax^2 + bx + c$$

con

$$b^2 - 4ac < 0.$$

Pues ya sabemos que si $b^2 - 4ac \geq 0$, entonces

$$x - \frac{(-b + \sqrt{b^2 - 4ac})}{2a}$$

sería un factor propio de

$$ax^2 + bx + c.$$

9.3.1 Traslación de la gráfica de un polinomio

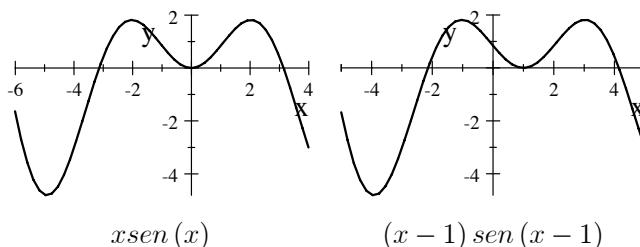
Sea $f : \mathbb{R} \rightarrow \mathbb{R}$, comparemos las gráficas de $f(x)$ y de

$$f(x - a) = (f \circ (\underline{} - a))(x).$$

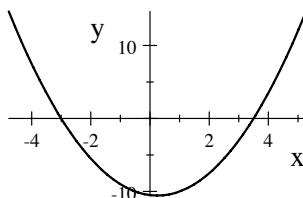
Desde luego,

$$\begin{array}{rccc} (\underline{} - a) : & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & x - a \end{array}.$$

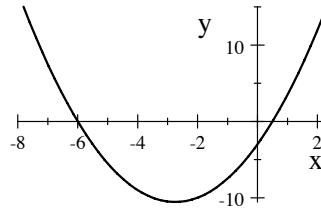
puede interpretarse como una traslación (a la derecha si $a > 0$).



Supongamos que un polinomio tiene una raíz entre 3 y 4, por ejemplo $f(x) = (x - 3.5)(x + 3)$, : $x^2 - .5x - 10.5$



entonces $f(x+3) = x^2 + 5.5x - 3.0$



tiene una raíz entre 0 y 1.

Para encontrar un cifra decimal para la raíz mayor de f uno tendría que calcular sucesivamente

$$f(3), f(3.1), f(3.2), \dots, f(3.9), f(4)$$

o lo que es lo mismo, calcular

$$f(3+x)(0), f(3+x)(0.1), f(3+x)(0.2), \dots, f(3+x)(0.9), f(3+x)(1)$$

y estar pendiente del cambio de signo.

Para cálculos con lápiz y papel es mucho mejor evaluar un polinomio alrededor de 0 que de otro número. Por ejemplo, es más fácil calcular $(0.2)^5$ que $(10.2)^5$.

Calcular

$$(x^4 + 3x^2 + x + 10)(7.2)$$

es lo mismo que calcular

$$((x+7)^4 + 3(x+7)^2 + (x+7) + 10)(0.2).$$

Tenemos que

$$\begin{aligned} ((x+7)^4 + 3(x+7)^2 + (x+7) + 10) &= \\ &= x^4 + 28x^3 + 297x^2 + 1415x + 2565. \end{aligned}$$

Ahora, calcular $(x^4 + 3x^2 + x + 10)(7.2)$ es equivalente a encontrar el residuo en la división

$$x - 7.2 \overline{)x^4 + 3x^2 + x + 10}.$$

Hagámoslo con el siguiente formato:

dividendo	divisor	cociente
$x^4 + 3x^2 + x + 10$	$x - 7.2$	$x^3 + 7.2x^2 + 54.84x + 395.85$
$-x^4 + 7.2x^3$		
$7.2x^3 + 3x^2 + x + 10$		
$-(7.2x^3 - 51.84x^2)$		
$54.84x^2 + x + 10$		
$-(54.84x^2 - 394.85x)$		
$395.85x + 10$		
$-(395.85x - 2850.1)$		
2860.1		

Comparemos con

dividendo	divisor	cociente
$x^4 + 28x^3 + 297x^2 + 1415x + 2565$	$x - .2$	$x^3 + 28x^2 + 302.6x + 1475.5$
$-(x^4 - .2)$		
$28x^3 + 297x^2 + 1415x + 2565.2$		
$-(28x^3 - 5.6x^2)$		
$302.6x^2 + 1415x + 2565.2$		
$-(302.6x^2 - 60.52x)$		
$1475.5x + 2565.2$		
$-(1475.5x - 295.1)$		
2860.1		

Notemos que en

dividendo	divisor	cociente
$a_nx^n + a_{n-1}x^{n-1} + \dots + a_0$	$x - b$	$a_nx^{n-1} + (a_{n-1} + a_n b)x^{n-2} + \dots$
$-a_nx^{n-1}(x - b)$		
$(a_{n-1} + a_n b)x^{n-1} + \dots$		
$-(a_{n-1} + a_n b)x^{n-2}(x - b)$		
$(a_{n-2} + a_n b^2)x^{n-2} + \dots$		
\vdots		

sería mucho más económico poner sólo los coeficientes, cambiar el signo de b en la columna del divisor para no tener que restar y además podemos hacer

la operaciones en una sola línea.

$$\begin{array}{cccccc|c}
 a_n & a_{n-1} & a_{n-2} & \cdots & a_0 & b \\
 a_n \cdot b & (a_{n-1} + (a_n \cdot b)) b & & \cdots & & \\
 \hline
 a_n & a_{n-1} + (a_n \cdot b) & a_{n-2} + (a_{n-1} + (a_n \cdot b)) b & \cdots & r & \\
 \text{coeficientes del cociente} & & & & &
 \end{array}$$

Volvamos a hacer la división

$$\begin{array}{c}
 x^4 + 28x^3 + 297x^2 + 1415x + 2565 \mid x - 0.2
 \end{array}$$

1	28	297	1415	2565	.2
	.2	28.2*.2=5.64	302.64*.2=60.528	1475.5*.2=295.1	
1	28.2	302.64	1475.5	2860.1	

Escribir un polinomio $f(x)$ en la forma $b_n(x-a)^n + \cdots + b_1(x-a) + b_0$ es la misma cosa que escribir un número en representación decimal, excepto porque aquí la base es $(x-a)$. Note que b_0 es el residuo de la división

$$x - a \mid \overline{f(x)},$$

digamos que

$$x - a \mid \frac{f_1(x)}{b_0},$$

es decir que

$$f(x) = (x-a)f_1(x) + b_0 = b_n(x-a)^n + \cdots + b_1(x-a) + b_0$$

por lo que

$$f_1(x) = b_n(x-a)^{n-1} + \cdots + b_1,$$

de donde se tiene que b_1 es el residuo en

$$x - a \mid \frac{f_2(x)}{b_1},$$

repitiendo el argumento (o mediante una sencilla inducción sobre el grado de $f(x)$), tenemos que los coeficientes

$$b_0, b_1, \dots$$

se obtienen al hacer las divisiones

$$x - a \left| \begin{array}{c} f_1(x) \\ f(x) \end{array} \right. , \quad x - a \left| \begin{array}{c} f_2(x) \\ f_1(x) \\ b_0 \end{array} \right. , \dots , \quad x - a \left| \begin{array}{c} f_{i+1}(x) \\ f_i(x) \\ b_i \end{array} \right. , \dots$$

Para este propósito conviene el formato

$$\begin{array}{c} f(x) \quad | \quad a \\ \hline f_1(x) \quad | \quad b_0 \\ f_2(x) \quad | \quad b_1 \\ f_3(x) \quad | \quad b_2 \\ \vdots \quad \quad \vdots \end{array}$$

que permite hacer las divisiones una tras otra.

Ejemplo 194 . Queremos expresar

$$x^3 + 5x^2 + x + 6,$$

en la forma

$$b_3(x+5)^3 + b_2(x+5)^2 + b_1(x+5)^1 + b_0(x+5)^0$$

Entonces

$$\begin{array}{ccccc} \text{dividendos} & & & & \text{divisor} \\ \hline 1 & 5 & 1 & 6 & \\ & -5 & 0 & -5 & \\ \hline 1 & 0 & 1 & | & 1 \\ & -5 & 25 & & \\ \hline 1 & -5 & | & 26 & \\ & -5 & & & \\ \hline 1 & | & -10 & & \end{array} \quad -5$$

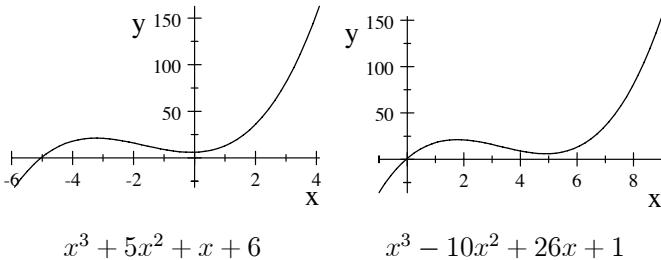
Así que

$$b_0 = 1, \quad b_1 = 26, \quad b_2 = -10, \quad b_3 = 1$$

y

$$x^3 + 5x^2 + x + 6 = (x+5)^3 - 10(x+5)^2 + 26(x+5) + 1$$

Notemos ahora que $f(x-5) = x^3 - 10x^2 + 26x + 1$ tiene la gráfica de $f(x)$, pero trasladada a la derecha:



Ejercicio 410 . $x^3 - 10x^2 + 26x + 1$ vale 1 en 0. Calcule cuánto vale en -1 y en 1 dividiendo entre $(x+1)$ y $(x-1)$. ¿En qué intervalo está la raíz de $x^3 - 10x^2 + 26x + 1$? ¿En $[-1, 0]$ o en $[0, 1]$?

9.3.2 El método de Horner

En el pasado reciente, se acostumbraba encontrar las raíces de un polinomio por el Método de Horner que consistía en lo siguiente:

Supongamos que queremos calcular la expresión decimal de la raíz

$$N.a_1a_2a_3\dots$$

de $f(x) \in \mathbb{R}[x]$. Para empezar, supongamos que ya sabemos que la raíz está entre N y $N+1$ (por ejemplo si hubiésemos encontrado que $f(N)$ y $f(N+1)$ tienen signos contrarios).

Entonces expresaríamos $f(x)$ en la forma

$$a_k(x-N)^k + \dots + a_1(x-N) + a_0$$

y consideraríamos $g(x) = a_kx^k + \dots + a_1x + a_0$. Notemos que

$$g(x-N) = f(x)$$

Por lo que $g(0) = g(N-N) = f(N)$, así que si f tiene una raíz entre N y $N+1$, $g(x)$ la tiene entre 0 y 1.

La siguiente cifra decimal de la raíz se calcula evaluando $g(x)$ en

$$0.1, 0.2, 0.3, \dots, 0.9$$

para detectar un cambio de signo. Digamos que la raíz está entre $.a_1$ y $.a_2$

Ahora, se expresa $g(x)$ en la forma

$$c_m(x - .a_1)^m + \cdots + c_1(x - .a_1) + c_0$$

y se considera $g_1(x) = c_m(x)^m + \cdots + c_1(x) + c_0$ que tiene la raíz entre $.0$ y $.10$, después se evalúa $g_1(x)$ en

$$0.01, 0.02, 0.03, \dots, 0.09$$

y el proceso se repite tantas veces como se quiera.

En el proceso anterior, se usaba el método de división abreviado (“sintética”) que se describió antes.

El método es eficaz, pues cuando $x \rightarrow 0$, los valores de x^m que cuentan más son los de m pequeña.

Por ejemplo, tomemos

$$x^4 + x^3 + x^2 + x + 1$$

Consideremos la siguiente tabla

x	x^4	x^3	x^2	x	1	$x^4 + x^3 + x^2 + x + 1$
1	1	1	1	1	1	5
0.1	.0001	.001	.01	.1	1	1.1111
.01	.00000001	.000001	.0001	.01	1	1.01010101

como podemos apreciar, la contribución relativa de las potencias grandes al valor del polinomio es cada vez menor, conforme vamos evaluando en valores menores de x .

El método de Horner ya no es de utilidad práctica dado que disponemos de computadoras y calculadoras, sin embargo tiene aspectos interesantes, por ejemplo, a pesar de su sencillez, nos proporciona uno tras otro, los dígitos de las raíces.

Ejemplo 195 . *Mostraremos como funciona el Método de Horner para calcular algunas cifras de $\sqrt{2}$.*

El polinomio a considerar es $x^2 - 2$

$$\begin{array}{r} 1 \ 0 \ -2 \\ \quad 1 \quad 1 \quad 1 \\ \hline 1 \ 1 \ \boxed{-1} \end{array}$$

$$\begin{array}{r}
 1 \ 0 \ -2 \\
 2 \ 4 \\
 \hline
 1 \ 2 \boxed{2}
 \end{array} \ .$$

Como $x^2 - 2$ cambia de signo entre 1 y 2 entonces tenemos una raíz en el intervalo $[1, 2]$.

Expresemos ahora $x^2 - 2$ como combinación de potencias de $x - 1$:

$$\begin{array}{r}
 1 \ 0 \ -2 \\
 1 \ 1 \\
 \hline
 1 \ 1 \boxed{-1} \ 1 \\
 1 \\
 \hline
 \boxed{1} \boxed{2} \boxed{}
 \end{array}$$

$$x^2 - 2 = (x - 1)^2 + 2(x - 1) - 1$$

Tomamos

$$g(x) = x^2 + 2x - 1,$$

evalúamos en 1, 2, ...:

$$\begin{array}{r}
 1 \ 2 \ -1 \ \boxed{.1} \ 1 \ 2 \ -1 \ \boxed{.2} \\
 .1 \ .21 \ \quad \quad \quad .2 \ .44 \\
 \hline
 1 \ 2.1 \ -.79 \ \quad \quad \quad 1 \ 2.2 \ -.56 \\
 \\
 1 \ 2 \ -1 \ \boxed{.3} \ 1 \ 2 \ -1 \ \boxed{.4} \\
 .3 \ .69 \ \quad \quad \quad .4 \ .96 \\
 \hline
 1 \ 2.3 \ -.31 \ \quad \quad \quad 1 \ 2.4 \ -.04 \\
 \\
 1 \ 2 \ -1 \ \boxed{.5} \\
 .5 \ 1.25 \\
 \hline
 1 \ 2.5 \ .25
 \end{array}$$

El cambio de signo nos dice que hay una raíz para $g(x)$ entre .4 y .5. Por lo que una aproximación a raíz de 2 es 1.4.

$$x^2 + 2x - 1 = (x - 0.4)^2 + c_1(x - 0.4) + c_0$$

se calcula así:

$$\begin{array}{r}
 1 \ 2 \ -1 \ \boxed{0.4} \\
 .4 \ .96 \\
 1 \ 2.4 \ \boxed{-0.4} \\
 .4 \\
 \hline
 1 \ \boxed{2.8}
 \end{array}$$

Así que

$$g_1(x) = x^2 + 2.8x - .04$$

que evaluaremos en

$$\begin{array}{rccccc} & & .01, & .02, & \dots \\ \begin{array}{rccccc} 1 & 2.8 & -0.04 & \boxed{.01} & 1 & 2.8 & -0.04 & \boxed{.02} \\ .01 & .0281 & & & .02 & .0564 & & \\ 1 & 2.81 & -0.0119 & & 1 & 2.82 & .0164 & \end{array} \end{array}$$

El cambio de signo nos dice que $g_1(x)$ tiene una raíz entre .01 y .02.

Por lo tanto 1.41 es una aproximación a $\sqrt{2}$.

Ejercicio 411 . Calcule $\sqrt[3]{2}$ con dos cifras decimales.

Observación 141 . Desde luego, también se puede expresar un polinomio $f(x)$ como

$$a_m(x)g(x)^m + \dots + a_1(x)g(x) + a_0$$

para cualquier polinomio g de grado > 0 , donde los coeficientes son polinomios en $\mathbb{R}[x]$ de grado menor que el grado de g .

Ejercicio 412 . Demuestre la existencia y la unicidad de los coeficientes en la observación anterior.

Ejemplo 196 . Expresaremos $x^4 + x^3 + x^2 + x + 1$ en términos de potencias de $x^2 + 2x + 1$:

$$\begin{array}{r} x^2 - x + 2 \\ x^2 + 2x + 1 \left| \begin{array}{r} x^4 + x^3 + x^2 + x + 1 \\ -x^4 - 2x^3 - x^2 \\ \hline -x^3 + x + 1 \\ x^3 + 2x^2 + x \\ \hline 2x^2 + 2x + 1 \\ -2(x^2 + 2x + 1) \\ \hline -2x - 1 \end{array} \right. \\ \hline \end{array}$$

$$\begin{array}{r} 1 \\ x^2 + 2x + 1 \left| \begin{array}{r} x^2 - x + 2 \\ -x^2 - 2x - 1 \\ \hline -3x + 1 \end{array} \right. \end{array}$$

Entonces

$$\begin{aligned} x^4 + x^3 + x^2 + x + 1 &= \\ &= (x^2 + 2x + 1)^2 + (-3x + 1)(x^2 + 2x + 1) + (-2x - 1). \end{aligned}$$

9.4 Un procedimiento gráfico para resolver algunas desigualdades

Se desea encontrar la solución del problema siguiente:

Problema. Dado un polinomio $f(x) \in \mathbb{R}[x]$, determinar

$$S_0 = \{x \in \mathbb{R} \mid f(x) < 0\}$$

$$(\delta f(x) \leq 0 \text{ ó } f(x) > 0 \text{ ó } f(x) \geq 0).$$

Para describir una manera de proceder, aceptaremos provisionalmente (ver “complejos”) que todo polinomio $f(x) \in \mathbb{R}[x]$ se puede expresar como producto de una constante k por el producto de una colección -posiblemente vacía- de polinomios mónicos irreducibles, y que esta factorización es única excepto por el orden en el que puedan aparecer los factores. (En $\mathbb{R}[x]$ son irreducibles todos los polinomios de grado uno así como aquellos que siendo de grado dos, tienen discriminante negativo y sólo esos).

Recordemos que si

$$f(x) = ax^2 + bx + cx,$$

su discriminante es el número $b^2 - 4ac$, y que $g(x) \in \mathbb{R}[x]$ es mónico si y sólo si su coeficiente principal es 1.

Obsérvese finalmente que si $g(x)$ es mónico e irreducible de 2° grado, entonces $\forall x \in \mathbb{R}, g(x) > 0$.

En efecto,

$$\begin{aligned} g(x) &= x^2 + bx + c = x^2 + bx + \frac{b^2}{4} - \frac{b^2}{4} + c = \\ &= \left(x + \frac{b}{2}\right)^2 + \frac{4c - b^2}{4}. \end{aligned}$$

y como $\Delta = b^2 - 4c$ es < 0 , entonces $\frac{4c - b^2}{4} > 0$, luego tiene raíz cuadrada positiva, entonces $g(x) = \left(x + \frac{b}{2}\right)^2 + \left(\sqrt{\frac{4c - b^2}{4}}\right)^2$, y por lo tanto, $g(x) > 0$.

Sea

$$f(x) = k(x - r_1)^{\alpha_1}(x - r_2)^{\alpha_2} \dots (x - r_n)^{\alpha_n} g_1^{\beta_1}(x) \dots g_m^{\beta_m}(x)$$

una factorización de $f(x)$, en donde, $\forall i \in \{1, \dots, m\}$, $g_i(x)$ es un mónico irreducible de 2^o grado,

$$\alpha_i \geq 1, i = 1, \dots, n,$$

$$\beta_j \geq 1, j = 1, \dots, m,$$

y $r_1 < r_2 < \dots < r_n$.

Supóngase que se desea resolver alguna de las desigualdades del problema, para $f(x)$.

Sin pérdida de generalidad podemos suponer $k > 0$ (si fuera $k < 0$, multiplicando por -1 todo el polinomio, nuestro problema consistiría en resolver la nueva desigualdad que resulta, y que, como es evidente, transforma el problema en otro del mismo tipo).

Hagamos las observaciones siguientes:

1. Cada $r_i, i \in \{1, \dots, n\}$, es raíz de $f(x)$, luego la gráfica de f toca el eje X en cada una de ellas. ($f(r_i) = 0, \forall i \in \{1, \dots, n\}$).
2. En vista de la observación que se hizo con anterioridad, cada factor $g_i^{\beta_i}(x)$ es positivo por lo que, en el análisis que haremos, podemos bautizar el producto de todos ellos como $t(x)$ que es mayor que $0, \forall x \in \mathbb{R}$. Luego

$$f(x) = kt(x) [(x - r_1)^{\alpha_1} \dots (x - r_n)^{\alpha_n}],$$

y entonces, el signo de $f(x)$ queda determinado exclusivamente por el signo del producto dentro del paréntesis rectangular.

3. Para cada $i, i \in \{1, \dots, n-1\}$, si $x \in (r_i, r_{i+1})$, todos los factores $(x - r_j)$ son negativos si $j > i$ y positivos si $j \leq i$. En efecto, de

$$r_i < x < r_{i+1}$$

se sigue que si $j > i$, entonces $r_j \geq r_{i+1} \therefore x < r_j \therefore x - r_j < 0$ y si $j \leq i$, $r_j \leq r_i \therefore r_j < x \Rightarrow 0 < x - r_j$ y por lo tanto cada vez que la x cambia de un intervalo al anterior, hay un único factor en

$$\{(x - r_1), \dots, (x - r_n)\}$$

que cambia de signo y por lo tanto $f(x)$ cambiará de signo también si y sólo si el exponente correspondiente al factor afectado es impar.

El resultado final de estas observaciones justifica el procedimiento que detallamos enseguida.

9.4.1 Procedimiento gráfico para resolver la desigualdad $f(x) < 0$

1. Factorice $f(x)$ como producto de mónicos irreducibles y agrupe los de 2^o grado: Sea

$$f(x) = k(x - r_1)^{\alpha_1} \dots (x - r_n)^{\alpha_n} t(x)$$

en donde $k > 0$, $t(x) > 0$, $\forall x \in \mathbb{R}$.

2. Obsérvese que si $x > r_n$, cada factor $(x - r_i)$ es positivo, por lo tanto $f(x) > 0$ (el dibujo de su gráfica va por arriba del eje X).
3. En cada r_i , $f(x) = 0$ y al pasar del intervalo que empieza con r_i , al que termina con r_i , $f(x)$ cambia de signo si y sólo si α_i es impar. (Si α_i es impar, $f(x)$ “cruza” el eje X si x pasa por r_i y “rebota” en r_i si α_i es par).
4. Recordando (o suponiendo) que los polinomios tienen gráficas “suaves” (continuas y sin “picos”), y aún cuando la información que se ha adquirido hasta aquí es incompleta (sólo afirma que $f(x)$ está arriba, abajo o sobre el eje X), basta para bosquejar su gráfica y proceder, a partir de ella, a definir el conjunto que resuelve la desigualdad.

Ejemplo 197 . Supóngase que se desea resolver: $f(x) \geq 0$ cuando

$$\begin{aligned} f(x) &= 3(x+4)^2(x+1)^3x^4(x-2)(x^2+x+1)^3(x^2+x+2)^2 \\ &= 3t(x)(x+4)^2(x+1)^3x^4(x-2), \end{aligned}$$

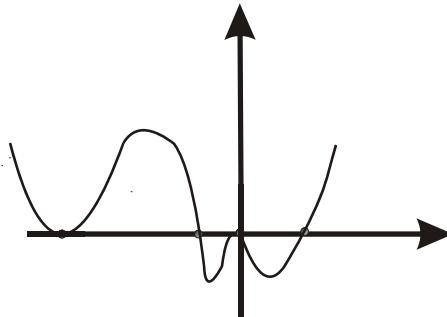
en donde $t(x) = (x^2+x+1)^3(x^2+x+2)^2$, que es positivo $\forall x \in \mathbb{R}$.

Las raíces de $f(x)$ son: $-4, -1, 0, 2$ por lo tanto:

- Si $x > 2$, entonces $f(x) > 0$.
- La observación de los exponentes de los factores lineales dice que

en $x = 2$ $f(x)$ cruza el eje X
 en $x = 0$ rebota
 en $x = -1$ cruza
 en $x = -4$ rebota

Luego la gráfica (cualitativa) es



- Se determina, observando la gráfica, que

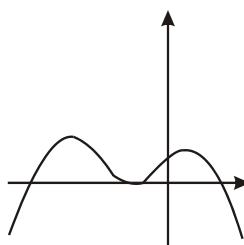
$$S_0 = \{x \in \mathbb{R} \mid f(x) \geq 0\} = (-\infty, -1] \cup \{0\} \cup [2, \infty).$$

Ejemplo 198 . Sea

$$f(x) = -7(x+6)^3(x+2)^2(x-3)^5(x^2+1)^7$$

y se desea resolver $f(x) < 0$.

- Multiplique por -1 , para hacer $k > 0$, y defínase $t(x) = (x^2 + 1)^7$ que es positiva $\forall x \in \mathbb{R}$.
- Nótese que las raíces reales -ordenadas- son; $-6, -2$ y 3 , y que en $x = 3$ y en $x = -6$ la gráfica cruza y que en -2 rebota. Luego la gráfica de $-f(x)$ es aproximadamente:



- *Por lo tanto*

$$S_0 = \{x \in \mathbb{R} \mid f(x) < 0\} = \{x \mid -f(x) > 0\} = (-\infty, -6) \cup (3, \infty)$$

9.4.2 Una aplicación

Es frecuente encontrar el problema de resolver desigualdades en donde aparecen fracciones polinomiales que, por supuesto no están definidas en los puntos x que anulan a los denominadores. En estos casos, luego de eliminar estos puntos prohibidos, se acostumbra proceder a resolver por partes, lo que, en general, complica el procedimiento a los estudiantes que enfrentan el problema de determinar signos y manejar uniones e intersecciones de intervalos.

Una manera alternativa, consiste en multiplicar arriba y abajo de cada fracción por su denominador -que es multiplicar por 1, y que produce nuevas fracciones equivalentes a las anteriores cuyos denominadores, por ser cuadrados de polinomios diferentes de cero, son necesariamente positivos, por lo que pueden “transitar” libremente de un lado a otro de las desigualdades, sin alterar el sentido de éstas, y transformar el problema en una desigualdad de polinomios para cuya resolución puede aplicarse el método que hemos descrito con anterioridad

Ejemplo 199 . Se desea resolver: $\frac{x}{x-1} \leq \frac{x+1}{x+2}$.

Eliminamos $x = 1$ y $x = -2$.

Se multiplica el lado izquierdo de la desigualdad por $\frac{x-1}{x-1}$, y el derecho por $\frac{x+2}{x+2}$, con lo que se obtiene

$$\frac{x(x-1)}{(x-1)^2} \leq \frac{(x+1)(x+2)}{(x+2)^2}$$

Se multiplica cada miembro de (9.4.2) por $(x-1)^2(x+2)^2$ y se obtiene:

$$(x+2)^2x(x-1) < (x+2)(x+1)(x-1)^2.$$

Se suma a cada lado de (9.4.2), $-(x+2)^2x(x-1)$ y resulta:

$$0 \leq (x+2)(x+1)(x-1)^2 - (x+2)^2x(x-1)$$

$$0 \leq (x+2)(x-1)[(x+1)(x-1) - (x+2)x]$$

$$0 \leq (x+2)(x-1)(x^2-1-x^2-2x) = (x+2)(x-1)(-2x-1).$$

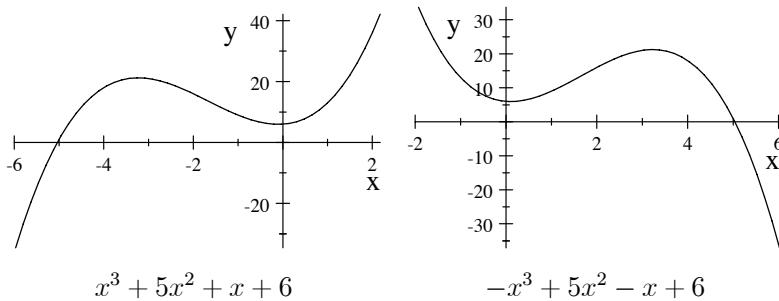
$0 \leq (-2)(x+2)(x+1/2)(x-1)$ se multiplica por -1 y se obtiene:

$0 \geq 2(x+2)(x+1/2)(x-1)$, cuya solución es $x \in (-\infty, -2] \cup [-1/2, 1]$ y como 2 y 1 estaban eliminados desde el principio, concluimos que la solución de (199) es $(-\infty, -2) \cup [-1/2, 1]$.

9.5 Reflexión sobre el eje Y

Observación 142 . Si f es un polinomio, entonces $f(-x) = f \circ (-(\))$ es un polinomio tal que su gráfica es la reflexión sobre el eje y de la gráfica de f . Es decir, $(a, f(a))$ pertenece a la gráfica de f mientras que $(-a, f(a)) = (-a, f(-(-a)))$ pertenece a la gráfica de $f(-x)$.

$$x^3 + 5x^2 + x + 6$$



9.6 Continuidad

Definición 160 . Diremos que $f(x)$ es creciente en (a, b) si $\forall x, y \in (a, b)$, $x < y \Rightarrow f(x) < f(y)$.

Teorema 159 . Si $f(x)$ es un polinomio tal que $f(0) > 0$ entonces $\exists a < 0 < b$ tales que $f((a, b)) \subseteq (0, \infty)$.

Demostración. Sea $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$, entonces $f(0) = a_0 > 0$.

Ahora,

$$|f(x) - a_0| = |a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x| \leq$$

$$\leq |a_n x^n| + |a_{n-1} x^{n-1}| + \cdots + |a_1 x|.$$

Ahora

$$|a_n x^n| < \frac{a_0}{n}$$

si

$$|x^n| < \frac{a_0}{n |a_n|}$$

es decir si

$$x \in \left(-\sqrt[n]{\frac{a_0}{n |a_n|}}, \sqrt[n]{\frac{a_0}{n |a_n|}} \right).$$

Análogamente

$$|a_i x^i| < \frac{a_0}{n}, \quad a_i \neq 0$$

si

$$x \in \left(-\sqrt[i]{\frac{a_0}{n |a_i|}}, \sqrt[i]{\frac{a_0}{n |a_i|}} \right).$$

Si $a_i = 0$, es claro que

$$|a_i x^i| < \frac{a_0}{n}.$$

Si $a_i \neq 0$, tomemos ahora la intersección de estos intervalos abiertos:

$$\cap \left\{ \left(-\sqrt[i]{\frac{a_0}{n |a_i|}}, \sqrt[i]{\frac{a_0}{n |a_i|}} \right) \right\}_{i \in \{1, \dots, n\}} = (a, b).$$

Ahora, es claro que $\forall x \in (a, b)$, se tiene que

$$|a_n x^n| + |a_{n-1} x^{n-1}| + \cdots + |a_1 x| < n \left(\frac{a_0}{n} \right) = a_0.$$

Entonces $|f(x) - a_0| < a_0$, $\forall x \in (a, b)$.

Si $f(x) < 0$ para alguna $x \in (a, b)$, entonces

$$a_0 - f(x) = |f(x) - a_0| < a_0.$$

Pero entonces $0 = a_0 - a_0 < f(x) \stackrel{\nabla}{<} 0$.

Por lo tanto, $f(x) > 0$, $\forall x \in (a, b)$. ■

Ejercicio 413 . Demuestre que la intersección de una familia finita no vacía de intervalos abiertos que comparten un elemento $c \in \mathbb{R}$ es un intervalo abierto.

Definición 161 . Sea $P(x) : \mathbb{R} \rightarrow \{0, 1\}$ una función proposicional. Interpretaremos $P(a) = 1$ como “ a tiene la propiedad P ” y $P(a) = 0$ como “ a no tiene la propiedad P ”.

Diremos que P es localmente cierta, alrededor de a , si

1. $P(a) = 1$ y
2. Existen $c < a < d$ tales $P(x) = 1$ para toda $x \in (c, d)$.

Ejemplo 200 . En el teorema 159, se demuestra que la propiedad “ $f(x) > 0$ ”, es localmente cierta, alrededor de 0, para cada polinomio $f(x)$ tal que $f(0) > 0$.

Teorema 160 . Si $f(x)$ es un polinomio tal que $f(0) < 0$ entonces f es menor que 0, localmente alrededor de 0.

Esto quiere decir que podemos encontrar $b, c \in \mathbb{R}$ tales que

$$b < 0 < d \text{ y } f((b, c)) \subseteq (-\infty, 0).$$

Demostración. Aplicemos el teorema anterior al polinomio $g(x) = -f(x)$.

Que $g(x)$ sea localmente positivo, alrededor de 0, es lo mismo que decir que $f(x)$ es localmente negativa alrededor de 0. ■

Corolario 27 . Sea $f(x)$ un polinomio tal que $f(a) > L$, para $a, L \in \mathbb{R}$. Entonces f es localmente mayor que L , alrededor de a .

Demostración. Queremos demostrar que existen $b, c \in \mathbb{R}$, $b < a < c$ tales que $f(b, c) \subseteq (L, \infty)$.

Consideremos el polinomio

$$g(x) = f(x + a) - L,$$

notemos que

$$g(0) = f(a) - L > 0.$$

Por el teorema 159, tenemos que g es localmente mayor que 0, alrededor de 0, por lo que existen $b' < 0 < c'$ tales que $g(b', c') \subseteq (0, \infty)$.

Así que

$$(f(x + a) - L)(b', c') \subseteq (0, \infty) \Leftrightarrow$$

$$\Leftrightarrow f(x + a)(b', c') \subseteq (L, \infty) \Leftrightarrow$$

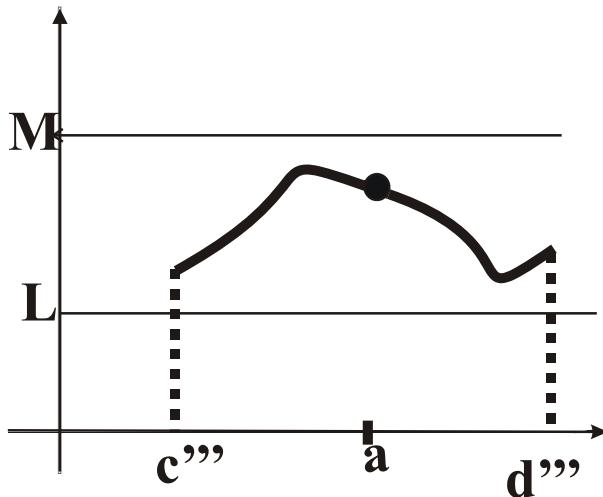


Figura 9.1:

$$\Leftrightarrow f((a+b', a+c')) \subseteq (L, \infty).$$

Tomemos $b = a + b', c = a + c'$ y notando que

$$b < a < c,$$

hemos terminado la demostración. ■

Ejercicio 414 . Demuestre que si $f(x)$ es un polinomio tal que $f(a) < M$ entonces f es localmente menor que M , alrededor de a .

Teorema 161 . Si $f(x)$ es un polinomio que tal que $L < f(a) < M$ entonces f es localmente (menor que M y mayor que L), alrededor de a .

Demostración. Por el ejercicio anterior f es localmente menor que M , alrededor de a , así que existen $b' < a < c'$ tales que $f((b', c')) \subseteq (-\infty, M)$.

Por el teorema anterior, f es localmente mayor que L , alrededor de a , es decir que existen $b'' < a < c''$ tales que $f((b'', c'')) \subseteq (L, \infty)$.

Tomemos $(b''', c''') = (b', c') \cap (b'', c'')$ y entonces en $f((b''', c''')) \subseteq (L, M)$.

■

El teorema anterior se puede expresar así: Si $f(x)$ es un polinomio tal que $f(a) \in (L, M)$ entonces $f(x)$ está localmente contenido en (L, M) , alrededor de a .

Lo anterior se puede expresar así:

$$f(a) \in (L, M) \Rightarrow \exists (b, c) \text{ tal que } a \in (b, c) \text{ y } (b, c) \subseteq f^{-1}(L, M).$$

Esta propiedad es la propiedad de que f es continua en a . Así que si quisieramos utilizar este concepto, diríamos que lo que se demostró arriba es que cualquier función polinomial es continua en cualquier punto, o más brevemente:

Toda función polinomial es continua.

9.7 Valores intermedios

Teorema 162 . *Los valores de una función polinomial están acotados por arriba en un intervalo cerrado $[a, b]$.*

Demostración. Supongamos que $a < b$.

Haremos uso del principio del supremo.

Si una función polinomial $f(x)$ no estuviera acotada por arriba en $[a, b]$, entonces

$$\mathcal{B} = \{z \in [a, b] \mid f \text{ no está acotada por arriba en } [a, z]\} \neq \emptyset,$$

pues $b \in \mathcal{B}$. Por otra parte $a \notin \mathcal{B}$ y es claro que a es una cota inferior para \mathcal{B} .

Como \mathcal{B} es un conjunto no vacío acotado por abajo, tiene que tener un ínfimo

$$s = \inf(\mathcal{B}).$$

Consideremos $f(s)$ y un valor $M > f(s)$. Entonces f es localmente menor que M alrededor de s . Tomemos $a \leq \alpha < s < \beta \leq b$ tales que $f((\alpha, \beta)) \subseteq (-\infty, M)$. Como s es la mayor de las cotas inferiores para \mathcal{B} y $\alpha < s$, entonces $\alpha \notin \mathcal{B}$ por lo que f está acotada por arriba en

$$[a, \alpha].$$

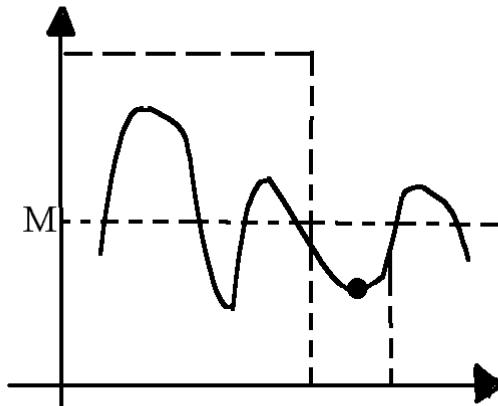


Figura 9.2:

Como f está acotada por arriba en $[a, \alpha]$ y también en (α, β) , entonces f está acotado por arriba en $[a, \beta]$, así, es claro que β es una cota inferior para \mathcal{B} , con $s < \beta \nabla$.

Esta contradicción muestra que \mathcal{B} es vacío, es decir que $\forall z \in [a, b]$, f está acotada por arriba en $[a, z]$.

En particular f está acotada por arriba en $[a, b]$. ■

Ejercicio 415 . *Demuestre que los valores de una función polinomial están acotados por debajo.*

Ejercicio 416 . *Demuestre que si un polinomio f está acotado en $[a, b]$ con $a < b$, entonces f está acotada en $[a, c]$, si $a \leq c \leq b$.*

Una vez que sabemos que un polinomio está acotado en un intervalo cerrado, podemos demostrar que asume su valor máximo en un intervalo cerrado.

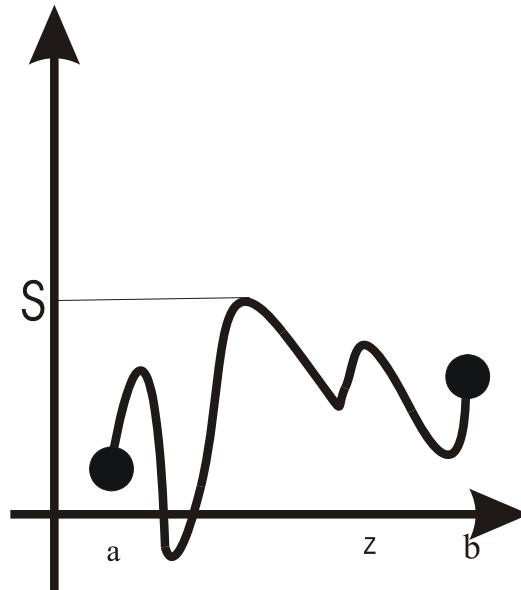


Figura 9.3:

Teorema 163 . Una función polinomial $f(x)$ asume su máximo en un intervalo cerrado.

Demostración. Supongamos que $a < b \in \mathbb{R}$. Ya sabemos que $f([a, b])$ está acotado por arriba, como además es un conjunto no vacío de reales, entonces tiene un supremo s que supondremos distinto de $f(a)$ y de $f(b)$, (porque en caso contrario no habría nada que demostrar).

Consideremos

$$\mathcal{B} = \{z \in [a, b] \mid s = \sup \{f(x) \mid x \in [a, z]\}\}.$$

Como $\sup \{f(x) \mid x \in [a, b]\} = s$ tenemos que $b \in \mathcal{B}$. Por otra parte, $a \notin \mathcal{B}$, pues $f(a) < s$. Es claro que a es una cota inferior para \mathcal{B} .

Podemos considerar

$$\gamma = \inf (\mathcal{B} \cap [a, b]).$$

Demostraremos que $f(\gamma) = s$.

Si $f(\gamma) < s$, escojamos un número $M \in (f(\gamma), s)$, entonces f es localmente menor que M alrededor de γ entonces existen $\alpha, \beta \in [a, b]$ tales que

$$\alpha < \gamma < \beta$$

y $f(\alpha, \beta) \subseteq (-\infty, M)$.

Como $\gamma \not\in \alpha$ entonces $\alpha \notin \mathcal{B}$, por lo que $\sup\{f(x) \mid x \in [a, \alpha]\} =: T$ es menor que s .

Entonces T es una cota superior para f en $[a, \alpha]$ y M es una cota superior para f en (α, β) por lo tanto $\max\{M, T\} < s$ es una cota superior para f en $[a, \beta]$. Por lo tanto $\beta \notin \mathcal{B}$ y ningún elemento de $[a, \beta]$ pertenece a \mathcal{B} . Esto significa que β es una cota inferior para \mathcal{B} , pues $\gamma < \beta$ era la mayor de las cotas inferiores para \mathcal{B} .

Por lo tanto $f(\gamma) \geq s$, pero $f(\gamma)$ no puede ser mayor que s , por la definición de s .

Por lo tanto $f(\gamma) = s$. ■

Ejercicio 417 . *Haga un esquema para ilustrar la demostración del teorema anterior.*

Ejercicio 418 . *Demuestre que un polinomio asume su ínfimo en un intervalo cerrado.*

Ejercicio 419 . *Dé un ejemplo de un polinomio que no asuma su máximo en un intervalo abierto.*

Teorema 164 . *Un polinomio $f(x) : \mathbb{R} \rightarrow \mathbb{R}$ tiene la siguiente propiedad:*

$$\forall a < b \in \mathbb{R}, \quad f([a, b]) = [c, d],$$

para algunos elementos $c, d \in \mathbb{R}$.

El teorema anterior dice sencillamente que un polinomio manda un intervalo cerrado en un intervalo cerrado. Este teorema, a pesar de su aparente sencillez, tiene consecuencias interesantes.

Demostración. Es claro que

$$d = \sup\{f(x) \mid x \in [a, b]\}$$

y que

$$c = \inf \{f(x) \mid x \in [a, b]\}.$$

Que son valores asumidos por f en el intervalo $[a, b]$.

Resta ver que cualquier valor intermedio también es asumido por f .

Dicho de otra manera, queremos demostrar que

$$[a, b] \xrightarrow{f|_{[a,b]}} [c, d]$$

es suprayectiva.

Sea $k \in (c, d)$ y consideremos $\mathcal{B} = \{x \in [a, b] \mid f(x) < k\}$.

Es claro que si $f(m) = c$, entonces $m \in \mathcal{B}$. Tomemos $\gamma = \sup(\mathcal{B})$.

Si $f(\gamma) < k$ tomemos ℓ tal que $f(\gamma) < \ell < k$, f sería localmente menor que ℓ alrededor de γ .

Entonces existen α y $\beta \in [a, b]$ tales que $\alpha < \gamma < \beta$ y $f((\alpha, \beta)) \subseteq (-\infty, \ell)$. Pero entonces $f(\beta) < k$ con $\gamma < \beta$, pero por definición, γ es cota superior para los elementos de \mathcal{B} , por lo que $\beta \leq \gamma$. ■

Por lo tanto $f(\gamma) \geq k$.

Si $f(\gamma) > k$, entonces f sería localmente mayor que k alrededor de γ . Entonces existen λ y μ en $[a, b]$ tales que $\gamma \in (\lambda, \mu)$ y $f((\lambda, \mu)) \subseteq (k, \infty)$ veamos que λ también es cota superior para \mathcal{B} : de no ser así, $\exists z \in \mathcal{B}$ tal que $\lambda < z$. Por definición de γ tendríamos que $z < \gamma$ así que

$$\lambda < z < \gamma,$$

pero entonces $f(z) < k$ porque $z \in \mathcal{B}$, pero $f(z) > k$ porque $z \in [\lambda, \mu]$. Esta contradicción muestra que λ es cota inferior para \mathcal{B} . Como $\gamma = \sup(\mathcal{B})$ tenemos que $\gamma \leq \lambda < \gamma$. Esta contradicción demuestra que $f(\gamma) = k$. ■

Corolario 28 . *Sea $f(x)$ un polinomio, si $a < b$ y $f(a), f(b)$ tienen signos contrarios, entonces f tiene una raíz en $[a, b]$.*

Ejercicio 420 . *Demuestre el Corolario anterior.*

9.8 Derivadas

Usaremos la palabra “polinomio”, en el sentido de función polinomial de \mathbb{R} a \mathbb{R} .

Definición 162 . *Definimos*

$$\mathbb{R}[x] \xrightarrow{\alpha} \mathbb{R}[x]$$

mediante:

1. α es aditiva, es decir que $\alpha(f + g) = \alpha(f) + \alpha(g)$.
2. $\alpha(c \cdot x^n) = cn \cdot x^{n-1}$, $\forall c \in \mathbb{R}$, $n \in \mathbb{N} \setminus \{0\}$.
3. $\alpha(\widehat{c}) = \widehat{0}$, donde \widehat{c} denota el polinomio constante c .

Usaremos también la notación $f' = \alpha(f)$.

Ejemplo 201 . $(x^3 + x^2 + 5x + 2)' = 3x^2 + 2x + 5$.

Ejercicio 421 . *Demuestre que si $c \in \mathbb{R}$ y $f(x)$ es un polinomio entonces $\alpha(c \cdot f) = c \cdot \alpha(f)$.*

Observación 143 . *Para un polinomio g se tiene que $(xg)' = g + xg'$.*

Demostración. :

Digamos que

$$g = c_m x^m + \cdots + c_1 x + c_0 \quad (9.1)$$

entonces

$$xg = c_m x^{m+1} + \cdots + c_1 x^2 + c_0 x,$$

así,

$$(xg)' = (m+1) c_m x^m + \cdots + c_1 x + c_0. \quad (9.2)$$

Por otra parte, $g' = m c_m x^{m-1} + \cdots + c_0$, así que

$$xg' = m c_m x^m + \cdots + c_0 x \quad (9.3)$$

sumando las ecuaciones 9.1 y 9.3 obtenemos 9.2 ■

Teorema 165 . *Si f, g son polinomios entonces $(fg)' = f'g + fg'$.*

Demostración. Si f es el polinomio cero, $\hat{0}$, es claro que ambas expresiones dan $\hat{0}$.

Supongamos pues, que f tiene grado y hagamos la demostración por inducción sobre el grado de f .

Base.

Si $\text{grad}(f) = 0$ entonces $f = \hat{c}$, por lo que $fg = cg$, así que

$$(cg)' = c(g') = fg' + \hat{0} = fg' + 0g = fg' + f'g.$$

Paso inductivo.

Supongamos que $\text{grad}(f) > 0$ y haciendo la división

$$x \overline{f} \overline{a_0}$$

podemos escribir

$$f = xq + \hat{a}_0$$

donde $\text{grad}(q) = \text{grad}(f) - 1$.

Por lo que $fg = xqg + a_0g$, y entonces

$$\begin{aligned} (fg)' &= \\ &= (q(xg))' + (a_0g)' = \\ &= \underbrace{q(xg)'}_{\text{hipótesis de i.}} + q'(xg) + a_0g'. \end{aligned}$$

Por la observación precedente,

$$(xg)' = g + xg'.$$

Así que

$$\begin{aligned} (fg)' &= q(xg)' + q'(xg) + a_0g' = \\ &= q(g + xg') + q'(xg) + a_0g' = \\ &= g'(qx + a_0) + g(q + q'x) = \\ &= g'f + gf'. \end{aligned}$$

La última igualdad se dá porque $f' = (qx + a_0)' = (qx)' = q + xq'$, usando nuevamente la observación precedente. ■

Teorema 166 . Para un polinomio f y un $a \in \mathbb{R}$ se tiene que

$$\alpha(f(x+a)) = (\alpha(f))(x+a).$$

Demostración. Si f es el polinomio $\hat{0}$, la afirmación es evidente.

Supongamos que $f \neq \hat{0}$, haremos la demostración por inducción sobre el grado de f .

Base.

Si el grado de f es cero, entonces se obtiene $\hat{0}$ en ambos lados de la ecuación

Paso inductivo.

Si $\text{grad}(f) > 0$ escribamos $f = xq + c$ donde $\text{grad}(q) < \text{grad}(f)$.

Entonces $f(x+a) = (x+a)q(x+a) + c$ por lo que

$$\begin{aligned} \alpha(f(x+a)) &= (x+a)\alpha(q(x+a)) + q(x+a) = \\ &= \underbrace{(x+a)(\alpha(q))(x+a) + q(x+a)}_{\text{hip. de induc.}}. \end{aligned}$$

Por otra parte,

$$\begin{aligned} \alpha(f)(x+a) &= (x\alpha q + q)(x+a) = \\ &= (x+a)(\alpha q)(x+a) + q(x+a). \end{aligned}$$

■

Ejemplo 202 . $(2x^3 + 4x^2 + 5x + 1)(x-1) = 2x^3 - 2x^2 + 3x - 2.$ (Note que el polinomio de la izquierda se evalúa en $x-1$, no se confunda con un producto).

Por lo que

$$\begin{aligned} \alpha((2x^3 + 4x^2 + 5x + 1)(x-1)) &= \\ &= \alpha(2x^3 - 2x^2 + 3x - 2) = \\ &= 6x^2 - 4x + 3. \end{aligned}$$

Por otra parte $\alpha(2x^3 + 4x^2 + 5x + 1) = 6x^2 + 8x + 5$ y

$$\begin{aligned} (6x^2 + 8x + 5)(x-1) &= \\ &= 6(x-1)^2 + 8(x-1) + 5 = \\ &= 6x^2 - 4x + 3. \end{aligned}$$

Definición 163 . Diremos que f tiene un *máximo local en a* si existe (α, β) que contenga a a y tal que $f(a) \geq f(x)$, para cada $x \in (\alpha, \beta)$.

Teorema 167 . Si f tiene un *máximo local en a* , entonces $f'(a) = 0$.

Demostración. Mediante el algoritmo de la división escribamos

$$f(x) = (x - a)q(x) + f(a).$$

Supongamos que $a \in (\alpha, \beta)$ y que $f(a)$ es el valor máximo de $f(x)$ en (α, β) .

Si $x \in (\alpha, a)$ entonces $f(x) \leq f(a)$, por lo que $(x - a)q(x) \leq f(a)$, con $(x - a) < 0$. Por lo tanto

$$q(x) \geq 0 \text{ en } (\alpha, a).$$

En cambio, para $x \in (a, \beta)$ tenemos que $(x - a)q(x) \leq 0$, con $(x - a) > 0$, por lo que

$$q(x) \leq 0 \text{ en } (a, \beta).$$

Vemos pues que $q(x)$ cruza el eje X al pasar por a . Si $q(a) > 0$ entonces q sería localmente mayor que 0 alrededor de a , así que q no cruzaría el eje X . Por lo tanto $q(a) \leq 0$ y de nuevo, no puede ser menor que 0. Por lo tanto

$$q(a) = 0.$$

Ahora,

$$\begin{aligned} f'(a) &= ((x - a)q(x) + f(a))'(a) = \\ &= (q(x) + (x - a)q'(x))(a) = \\ &= q(a) + (a - a)q'(a) = 0. \end{aligned}$$

■

Ejercicio 422 . Demuestre que si f tiene un *mínimo local en a* , entonces $f'(a) = 0$.

Teorema 168 . Si $f(a) = f(b)$ para $a < b$, entonces $\exists z \in (a, b)$ tal que $f'(z) = 0$.

Demostración. Si f es constante en $[a, b]$, entonces $f(x) = \widehat{f(a)}$ ($f(x) - f(a)$ tiene una infinidad de raíces). Entonces su derivada es 0.

Si no es constante, entonces o bien el máximo de f en $[a, b]$ es mayor que $f(a)$ o bien el mínimo de f en $[a, b]$ es menor que $f(a)$. Supongamos que se dá el primer caso (si no fuera así sustituimos f por $-f$ para seguir el argumento). Digamos que este valor máximo ocurre en γ , es claro que $f(\gamma)$ es un máximo local de f . Por el teorema anterior, $f'(\gamma) = 0$. ■

Ejercicio 423 . *Muestre que si un polinomio tiene una infinidad de raíces entonces es el polinomio 0. (Use el Teorema Fundamental de Álgebra).*

Teorema 169 . *Sea $h(x)$ un polinomio tal que $h(0) = 0$ y sea $b \in \mathbb{R}^+$ entonces existe $c \in (0, b)$ tal que*

$$f'(c) = \frac{f(b)}{b}.$$

Demostración. Si $h(b)$ también es 0, usando el teorema anterior tenemos que $\exists c \in (0, b)$ tal que $f'(c) = 0$. ■

Supongamos entonces que $h(b) \neq 0$ y supongamos además que $h(b) > 0$ (en caso contrario, podríamos seguir argumentando con $-h$).

Tomemos ahora el polinomio

$$g(x) = h(x) - \frac{h(b)}{b}x$$

este polinomio vale 0 tanto en 0 como en b , así que por el teorema anterior $\exists c \in (0, b)$ tal que

$$g'(c) = 0.$$

Es decir que

$$h'(c) - \frac{h(b)}{b} = 0.$$

■

Teorema 170 (del valor medio) . *Sean $a < b \in \mathbb{R}$ y $f(x)$ un polinomio, entonces existe $c \in (a, b)$ tal que*

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

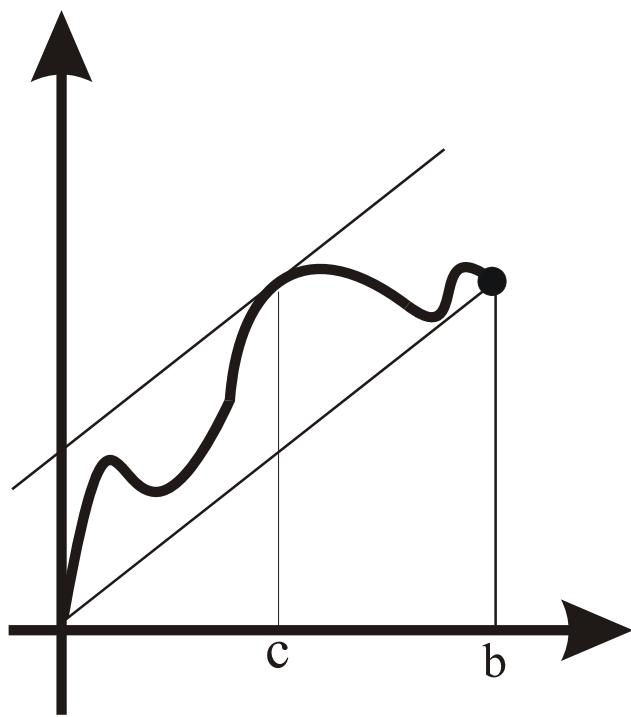


Figura 9.4:

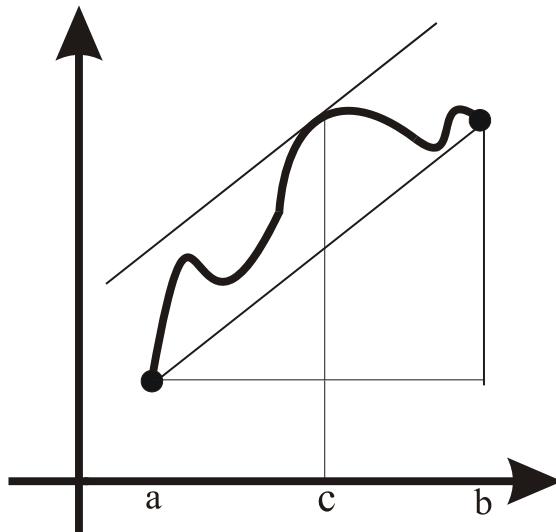


Figura 9.5:

Demostración. Consideremos el polinomio

$$h(x) = f(x + a) - f(a),$$

este polinomio tiene la propiedad de que $h(0) = 0$, así que aplicando el teorema anterior, tenemos que existe $d \in (0, b - a)$ tal que

$$h'(d) = \frac{h(b - a)}{b - a}$$

Pero entonces

$$f'(d + a) = \frac{h(b - a)}{b - a} = \frac{f(b) - f(a)}{b - a}.$$

Hagamos $c = (d + a)$ y notemos que

$$a < d + a < b.$$

■

Teorema 171 . *Sea f un polinomio tal que $f(0) = 0$, si $f'(0) > 0$, entonces f es creciente en un intervalo abierto alrededor de 0.*

Demostración. El polinomio f' es localmente mayor que 0 alrededor de 0. Sean $a < 0 < b$ tales que $f'(x) > 0$ en (a, b) . El teorema anterior nos permite asegurar que si $a \leq u < v \leq b$ entonces

$$f(v) - f(u) = f'(w)(v - u) > 0, \text{ para alguna } w \in (u, v).$$

Entonces $f(v) > f(u)$, es decir, f es creciente en (a, b) . ■

Teorema 172 . *Sea f un polinomio tal que $f'(a) > 0$, entonces f es creciente en un intervalo abierto alrededor de a .*

Demostración. Consideremos el polinomio $h(x) = f(x + a) - f(a)$, entonces $h(0) = 0$ y $h'(0) = f'(0 + a) = f'(a)$, por lo que podemos aplicar el teorema anterior para concluir que h es creciente en un intervalo abierto que contiene a 0. Esto es lo mismo que decir que f es creciente en un intervalo abierto que contiene a a . ■

Ejercicio 424 . *Sea f un polinomio tal que $f'(a) < 0$, entonces f es decreciente en un intervalo abierto alrededor de a .*

Ejercicio 425 . *Demuestre que son equivalentes:*

1. $h(x) = f(x + a) - f(a)$, es creciente en un intervalo abierto que contiene a 0.
2. f es creciente en un intervalo abierto que contiene a a .

Teorema 173 . *Si el polinomio f es tal que $f'(0) = 0$ y $f''(0) < 0$ entonces f tiene un máximo local en 0.*

Demostración. (Bosquejo, se deja al lector como ejercicio escribir los detalles). $f''(0) < 0$ implica que $f'(x)$ es decreciente en un intervalo abierto alrededor de 0, (a, b) digamos. Por lo tanto $f'(x)$ es positiva en $(a, 0)$ y negativa en $(0, b)$. Por lo tanto f crece en $(a, 0)$ y decrece en $(0, b)$, entonces $f(0)$ es un máximo local. ■

Ejercicio 426 . *Escribir los detalles en la demostración anterior.*

Ejercicio 427 . *Demostrar que si el polinomio f es tal que $f'(0) = 0$ y $f''(0) > 0$ entonces f tiene un mínimo local en 0.*

9.9 Derivadas y multiplicidad

1. Se dice que a es una raíz de multiplicidad m ($m \in \mathbb{N}$) del polinomio $f(x)$ si $(x - a)^m \mid f(x)$ pero $(x - a)^{m+1} \nmid f(x)$.
2. Se dice que a es una raíz múltiple de f si la multiplicidad de a es > 1 .
3. Se dice que una raíz es simple si su multiplicidad es 1.

Ejemplo 203 . La multiplicidad de 3 como raíz de $(x - 3)^2 (x - 1)$ es 2, la multiplicidad de 1 es 1.

Teorema 174 . a es una raíz múltiple de $f(x) \Leftrightarrow (x - a) \mid (f; f')$.

Demostración. \Rightarrow)

Supongamos que a es una raíz múltiple de $f(x)$, con multiplicidad m . Entonces podemos escribir

$$f(x) = (x - a)^m g(x)$$

con $g(a) \neq 0$ (que es lo mismo que decir que $(x - a) \nmid g(x)$).

Entonces por el teorema 165, tenemos que

$$f'(x) = ((x - a)^m)' g(x) + (x - a)^m g'(x).$$

Por otra parte, por el teorema 166, tenemos que $((x - a)^m)' = (x^m)' (x - a) = (mx^{m-1})(x - a) = m(x - a)^{m-1}$.

Por lo tanto

$$\begin{aligned} f'(x) &= m(x - a)^{m-1} g(x) + (x - a)^m g'(x) = \\ &= (x - a)^{m-1} (mg(x) + (x - a) g'(x)). \end{aligned}$$

Así tenemos que la multiplicidad de a como raíz de f' es $m - 1$. Note que $(x - a) \nmid (mg(x) + (x - a) g'(x))$ pues

$$mg(a) + (a - a) g'(a) = mg(a) \neq 0.$$

Entonces

$$(x - a) \mid (x - a)^{m-1} \mid (f; f').$$

\Leftarrow)

Supongamos que $(x - a) \mid (f; f')$ entonces a es raíz de f y de f' . Escribamos $f(x) = (x - a)^m g(x)$ donde $g(a) \neq 0$.

Como arriba,

$$f'(x) = (x - a)^{m-1} (mg(x) + (x - a)g'(x)).$$

Por hipótesis,

$$0 = f'(a) = (a - a)^{m-1} (mg(a) + (a - a)g'(a)),$$

por lo que $m > 1$ (si $m = 1$ entonces $f'(a) = mg(a) \neq 0$). ■

Observación 144 . $\frac{f}{(f; f')}$ tiene las mismas raíces que f pero no tiene raíces múltiples.

Demuestração. Como hemos visto en la demostración del teorema 174, si la multiplicidad de a como raíz de f es m entonces su multiplicidad como raíz de $f'(x)$ es $m - 1$. Por lo tanto la multiplicidad de a como raíz de $(f; f')$ es $m - 1$. Como

$$(f; f') = (x - a)^{m-1} h(x) \mid f(x) = (x - a)^m g(x),$$

con $h(a) \neq 0 \neq g(a)$, entonces

$$h(x) \mid (x - a)g(x)$$

por lo que

$$h(x) \mid g(x).$$

$(x - a)$ es irreducible y $x - a \nmid h(x)$, por lo que $(h(x); (x - a)) = 1$.

Entonces

$$\frac{f(x)}{(f; f')} = (x - a) \frac{g(x)}{h(x)}.$$

Pero además $\left(\frac{g(x)}{h(x)}\right)(a) \neq 0$, pues en caso contrario, $g(a) = 0$. $\left(\frac{g(x)}{h(x)}\right) \mid g(x)$, por lo que las raíces de $\frac{g(x)}{h(x)}$ son raíces de $g(x)$.

Por lo tanto si a es una raíz de f entonces a es raíz de $\frac{f(x)}{(f; f')}$ de multiplicidad 1.

Por otra parte, debemos notar que como $\frac{f(x)}{(f; f')}$ divide a $f(x)$, entonces todas las raíces de $\frac{f(x)}{(f; f')}$ son raíces de f . ■

Ejercicio 428 *Si a es una raíz de $f(x)$ de multiplicidad m y es también una raíz de $g(x)$ de multiplicidad n , demuestre que es una raíz de multiplicidad $\min\{n, m\}$ de $(f; g)$.*

Ejercicio 429 . *Encuentre la multiplicidad de 3 como raíz de $f(x) = x^4 - 8x^3 + 18x^2 - 27$ encontrando la multiplicidad de 3 como raíz de $(f; f')$.*

Ejercicio 430 . *Considere $f(x) = (x - a_1)^{m_1} \cdots (x - a_k)^{m_k}$, como*

$$\frac{f(x)}{(f; f')} = (x - a_1) \cdots (x - a_k)$$

demuestre que $f'(x) = c(x - a_1)^{m_1-1} \cdots (x - a_k)^{m_k-1} g(x)$, para alguna $c \in \mathbb{R}$ y $g(x)$ es un polinomio mónico de grado $k - 1$ que no tiene ninguna raíz en común con f . Encuentre c estimando el coeficiente principal de $f(x)$.

Ejercicio 431 . *Compruebe lo afirmado en el ejercicio anterior, con el polinomio*

$$(x - 1)^3 \cdot (x - 2) \cdot (x - 3)^3.$$

La observación 144 nos permite, al estimar las raíces de un polinomio, suponer que las raíces son simples. En particular, se puede suponer que $(f; f') = 1$.

Lema 33 . *Si $f(x)$ es un polinomio de grado mayor que 0 y de coeficiente principal positivo, entonces dada M existe N tal que*

$$f(x) > M$$

si $x \in (N, \infty)$.

Demostración. Por inducción sobre $\text{grad}(f)$.

Base.

Si $\text{grad}(f) = 1$, entonces $f(x) = ax + b$ con $a > 0$. Es claro que

$$ax + b > M \Leftrightarrow ax > M - b \Leftrightarrow x > \frac{M - b}{a}.$$

Se puede tomar $N = \frac{M-b}{a}$ en este caso.

Paso inductivo.

Supongamos que

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

con $n > 1$.

$$\begin{aligned} f(x) &> M \Leftrightarrow \\ a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x &> M - a_0 \Leftrightarrow \\ x(a_n x^{n-1} + a_{n-1} x^{n-2} + \cdots + a_1) &> M - a_0 \end{aligned}$$

Por hipótesis de Inducción $\exists N_1$ tal que

$$a_n x^{n-1} + a_{n-1} x^{n-2} + \cdots + a_1 > 1 \text{ si } x \in (N_1, \infty)$$

y $x > M - a_0$ si $x > |M - a_0|$.

Tomemos $N = \max \{N_1, |M - a_0|\}$ y es claro que

$$f(x) > M \text{ si } x > N.$$

■

Ejemplo 204 . Consideremos $2x^2 + 5x - 3$ queremos encontrar N tal que $2x^2 + 5x - 3 > 10$ si $x \geq N$.

$$2x^2 + 5x - 3 > 10 \Leftrightarrow 2x^2 + 5x > 13 \Leftrightarrow x(2x + 5) > 13.$$

$2x + 5 > 1 \Leftrightarrow 2x > -4 \Leftrightarrow x > -2$. Como $\max \{-2, 13\} = 13$, podemos tomar $N = 13$.

También podemos proceder resolviendo $x > 1$ y $(2x + 5) > 13$:

$$(2x + 5) > 13 \Leftrightarrow 2x > 8 \Leftrightarrow x > 4.$$

Lo que nos permite tomar $N = 4$.

9.10 El teorema de Sturm

Consideremos un polinomio $f(x) \in \mathbb{R}[x]$ sin raíces múltiples y con raíces ordenadas por tamaño

$$\rho_1, \rho_2, \dots, \rho_k.$$

Supongamos por simplificar, que el coeficiente principal de $f(x)$ es positivo.

Notemos que por la hipótesis de que no hay raíces múltiples entonces $(f; f')$ no tiene raíces así que debe ser 1).

Recordemos que al aplicar el algoritmo de Euclides, en la división

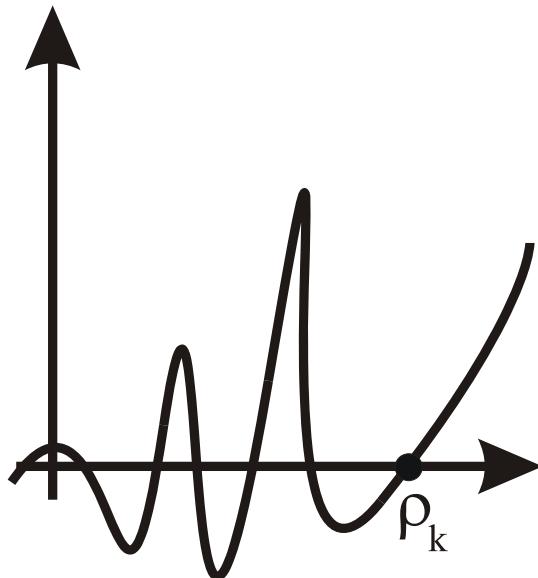
$$f'(x) \overline{) f(x)}^{\frac{q(x)}{r(x)}}$$

se tiene que $(f; f') = (f', r)$ por lo que tampoco f' y r tienen raíces comunes.

Notemos ahora que

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

es creciente en (ρ_k, ∞) :



Tenemos que $f(\rho_k) = 0$ y que ρ_k es la mayor raíz de f . Como f y $f'(x)$ no comparten raíces, entonces $f'(\rho_k) \neq 0$. Si $f'(\rho_k) < 0$, entonces

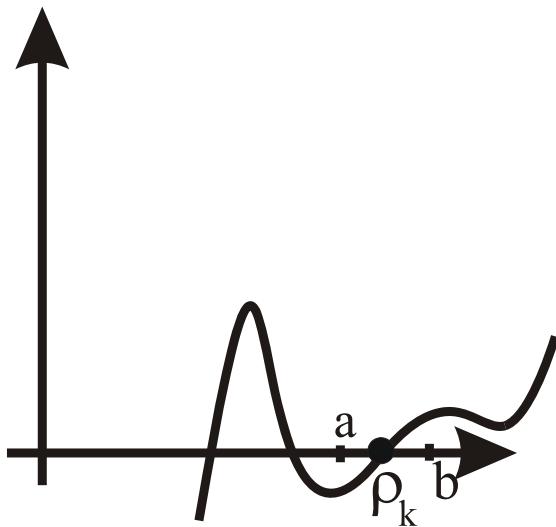


Figura 9.6:

f sería decreciente en ρ_k , por lo que tendría valores negativos a la derecha de ρ_k y como ρ_k es la raíz mayor de f , entonces $f(x) < 0$, $\forall x \in (\rho_k, \infty)$, contradiciendo el Lema 33.

Por lo tanto $f'(\rho_k) > 0$, y así $f'(\rho_k)$ sería localmente positiva alrededor de ρ_k . Así que existen $a < b$ tales que $\rho_{k-1} < a < \rho_k < b$, y tales que $f'(x) > 0$, para $x \in (a, b)$. Esto significa que f es creciente en (a, b) , por lo que es negativa en (a, ρ_k) y positiva en (ρ_k, b) .

Consideremos la siguiente tabla

	$a < x < \rho_k$	ρ_k	$\rho_k < x < b$
f	—	0	+
f'	+	+	+
no. de cambios de signo	1		0

Imagínese que no conocemos el valor de ρ_k pero que conocemos a .

Tomemos $c > a$ ¿está a la izquierda o a la derecha de ρ_k ? La respuesta depende de si hay cambio de signo o no en $\text{sig}(f(c)), \text{sig}(f'(c))$. (Claro que basta observar $\text{sig}(f(c))$, pero por favor, siga el argumento).

¿Podemos usar el argumento del cambio de signo para valores a la izquierda de a o a la derecha de b ?

Supongamos que γ_l es la mayor raíz de $f'(x)$ y, para fijar ideas, supongamos que $\gamma_l > b > \rho_k$, supongamos que es un mínimo de $f(x)$ (ver figura anterior). A la izquierda de γ_l hay valores para los que $f'(x)$ sería menor que 0 y a la derecha de γ_l hay valores con $f'(x)$ mayor que 0 así que repitiendo la tabla

	$\exists x \text{ tal que } \rho_k < x < \gamma_l \text{ y }$	γ_l	$\exists x \text{ tal que } \gamma_l < x < b \text{ y }$
f	+	+	+
f'	—	0	+
no. de cambios de signo	1		0

Observando la columna de la izquierda notamos que la raíz γ_l “nos ha metido ruido”.

Para que el valor de la derivada no altere la cuenta de los cambios de signo antes y después de la raíz ρ_k , agreguemos un renglón más a la tabla, consideremos un renglón para $-r_1(x)$ donde

$$f' \left| \begin{array}{c} q_1 \\ f \\ r_1 \end{array} \right. .$$

¿Por qué funciona esto? Note que $f = q_1 f' + r_1$ así que en una raíz de f' , f y r_1 tienen el mismo signo (recuerde que f y f' no tienen raíces en común). Por lo tanto, en una raíz de f' , f y $-r_1$ tienen signos contrarios así que f y $-r_1$ tienen signos contrarios localmente alrededor de las raíces de f' y la

tabla anterior se convierte en

	$\exists x$ tal que $\rho_k < x < \gamma_l$ y	γ_l	$\exists x$ tal que $\gamma_l < x < b$
f	+	+	+
f'	-	0	+
$-r_1$	-	-	-
cambios de signo	1	1	1

Como se vé, el número de cambios es 1 independientemente de la ubicación respecto a la raíz de f' .

¿Qué pasa a la izquierda de ρ_k ?

	$a < x < \rho_k$	ρ_k	$\rho_k < x < b$
f	-	0	+
f'	+	+	+
$-r_1$?	?	?
no. de cambios de signo	1 ó 2		0 ó 1

Ahora hay una ambigüedad producida por el comportamiento de r_1 (r_1 puede cambiar de signo), para eliminar el efecto de las raíces de r_1 agreguemos un renglón más a la tabla, incluyendo el signo de $-r_2$, donde

$$r_1 \left| \begin{array}{c} q_2 \\ f' \\ r_2 \end{array} \right. .$$

La introducción de este nuevo renglón a la tabla elimina el efecto de las raíces de r_1 , pero introduce el efecto producido por las raíces de r_2 . Para eliminarlo, introducimos el renglón correspondiente al signo de r_3 , donde

$$r_2 \left| \begin{array}{c} q_3 \\ r_1 \\ r_3 \end{array} \right. .$$

¿Por qué r_3 y no $-r_3$?

Pues porque en la tabla tomamos $-r_2$, así que si

$$r_1 = q_3 r_2 + r_3$$

entonces $-r_1 = q_3 (-r_2) - r_3$, así que en una raíz de r_2 , $-r_1$ y $-r_3$ tienen el mismo signo, y tomamos r_3 para que tenga signo contrario que $-r_1$ en una raíz de r_2 .

Notemos que este proceso de agregar renglones a la tabla termina, pues

$$\text{grad}(f) > \text{grad}(f') > \text{grad}(r_1) > \dots$$

termina (de la misma manera que termina el Algoritmo de Euclides).

Definición 164 . *Sea $f(x)$ un polinomio distinto de $\hat{0}$, sin raíces múltiples sean*

$$f, f', r_1, r_2, \dots, (f; f')$$

los polinomios que se obtienen aplicando el algoritmo de Euclides. Para $c \in \mathbb{R}$ definamos la sucesión

$$f(c), f'(c), -r_1(c), -r_2(c), r_3(c), r_4(c), -r_5(c), -r_6(c) \dots,$$

Note la colocación de los signos

$$+, +, -, -, +, +, -, -, \dots$$

Definamos ahora $V_f(c) =:$ número de cambio de signo en la sucesión (tome $\text{sig}(d) = +$ si $d > 0$, $\text{sig}(d) = -$ si $d < 0$, ignórense los ceros).

La demostración del siguiente teorema es la formalización de los párrafos precedentes.

Teorema 175 (de Sturm). *Sea $f(x)$ un polinomio no nulo sin raíces múltiples tomemos V_f como se definió arriba. Si $a < b$ entonces el número de raíces de f en (a, b) es $V(a) - V(b)$.*

Demostración. Numeremos las raíces de f

$$\rho_1 < \rho_2 < \dots < \rho_k.$$

Basta demostrar que $V(x) = V(z) + 1$ si $x \in (\rho_{i-1}, \rho_i)$ y $z \in (\rho_i, \rho_{i+1})$.

Como $f(\rho_i) = 0$ entonces $f'(\rho_i) \neq 0$, así que $f(x)$ cambia de signo al transponer ρ_i .

1) **Supongamos que no hay raíces de $f', r_1, r_2, r_3, \dots, (f, f')$ en el intervalo (ρ_{i-1}, ρ_{i+1}) .**

Entonces las funciones $f', r_1, r_2, r_3, \dots, (f, f')$ no cambiarían de signo en el intervalo (ρ_{i-1}, ρ_{i+1}) , así que las sucesiones

$$\operatorname{sig}(f(x)), \operatorname{sig}(f_0(x)), \operatorname{sig}(-r_1(x)), \operatorname{sig}(-r_2(x)), \operatorname{sig}(r_3(x)), \dots, x \in (\rho_{i-1}, \rho_i)$$

y

$$\operatorname{sig}(f(x)), \operatorname{sig}(f_0(x)), \operatorname{sig}(-r_1(x)), \operatorname{sig}(-r_2(x)), \operatorname{sig}(r_3(x)), \dots, x \in (\rho_i, \rho_{i+1})$$

coinciden, excepto por los primeros elementos que son opuestos.

1a) Si $\operatorname{sig}(f(x)) = +$ en $x \in (\rho_{i-1}, \rho_i)$ es porque f decrece para tomar el valor 0 en ρ_i , entonces $\operatorname{sig}(f'(x)) = -$ así que la sucesión de signos en (ρ_{i-1}, ρ_i) es

$$+, -, \operatorname{sig}(-r_1(x)), \dots$$

y en (ρ_i, ρ_{i+1}) es

$$-, -, \operatorname{sig}(-r_1(x)), \dots$$

Así que $V(c) = V(d) + 1$ si $c \in (\rho_{i-1}, \rho_i)$ y $d \in (\rho_i, \rho_{i+1})$

1b) Si $\operatorname{sig}(f(x)) = -$ en $x \in (\rho_{i-1}, \rho_i)$ es porque f crece para tomar el valor 0 en ρ_i , entonces $\operatorname{sig}(f'(x)) = +$, así que la sucesión de los signos en (ρ_{i-1}, ρ_i) es

$$-, +, \operatorname{sig}(-r_1(x)), \dots$$

y en (ρ_i, ρ_{i+1}) es

$$+, +, \operatorname{sig}(-r_1(x)), \dots$$

De nuevo note que $V(x)$ decrece en 1 al transponer ρ_i de izquierda a derecha.

2) Veamos que el valor de $V(x)$ en $(\rho_{i-1}, \rho_{i+1}) \setminus \{\rho_i\}$ no se altera por la presencia de raíces de f', r_1, r_2, \dots en (ρ_{i-1}, ρ_{i+1}) . Supongamos que $\gamma \in (\rho_{i-1}, \rho_{i+1}) \setminus \{\rho_i\}$ es raíz de algunas de las funciones f', r_1, r_2, \dots . γ no puede ser raíz de dos funciones consecutivas en la lista $f, f', -r_1, -r_2, \dots$ (porque f no tiene raíces múltiples). Lo que significa que al evaluar en γ , un 0 de la sucesión resultante está flanqueado por dos elementos de signo contrario. Esto se debe a que

$$\begin{array}{c} q_{i+2} \\ \hline r_{i+1} \mid \overline{r_i} \\ r_{i+2} \end{array}$$

y a que a r_{i+2} se le antepone el signo contrario que el que se le antepone a r_i , así que $r_i = q_{i+2}r_{i+1} - (-r_{i+2})$, por lo que si r_{i+1} se anula en γ , entonces r_i y $-r_{i+2}$ tienen signos opuestos.

Entonces se tendría una sucesión de la forma

$$\cdots, +, 0, -, \cdots -, 0, +, \cdots \quad (9.4)$$

Si $\xi < \gamma$ es tal que en (ξ, γ) no hay raíces de ninguna de las funciones, entonces para $x \in (\xi, \gamma)$ tendremos la misma sucesión 9.4, excepto porque los 0 serían sustituidos por + ó por -. Note que ninguna de estas posibilidades altera la cuenta de los cambios de signo, pues si originalmente teníamos un cambio en

$$+, 0, -$$

y el 0 se sustituye por + tendremos

$$+, +, -$$

(otra vez un cambio), y si 0 se cambia por -, tendremos

$$+, -, -,$$

de nuevo un cambio.

Como vemos, la presencia de raíces de las funciones auxiliares $f_i, -r_i, \dots$ no altera el valor de $V(x)$. Simplemente se pueden ignorar los 0.

El resultado se sigue ahora del hecho de que los valores de V disminuyen en 1 al transponer una raíz de izquierda a derecha. ■

Ejemplo 205 . Usaremos el teorema de Sturm para ubicar las raíces de $x^3 + 5x^2 + x + 6$:

$$\frac{d}{dx} (6 + x + 5x^2 + x^3) = 1 + 10x + 3x^2$$

$$\begin{array}{c} -\frac{3}{44}x - \frac{587}{1936} \quad \frac{1}{3}x + \frac{5}{9} \\ -44x + 49 \left[\begin{array}{c} 3x^2 + 10x + 1 \quad | \quad x^3 + 5x^2 + x + 6 \\ 30699 \quad | \quad -\frac{10}{3}x^2 - \frac{1}{3}x \\ 1936 \quad | \quad \frac{5}{3}x^2 + \frac{2}{3}x + 6 \\ \quad \quad \quad -\frac{44}{9}x + \frac{49}{9} \end{array} \right. \end{array} .$$

Entonces f y f' son primos relativos, la sucesión de polinomios auxiliares para calcular V es (excepto por factores positivos):

$$\begin{aligned} f(x) &= x^3 + 5x^2 + x + 6, \\ f_1(x) &= 3x^2 + 10x + 1, \\ -r_1(x) &= 44x - 49, \\ &\quad -1 \end{aligned}$$

Hagamos una tabla

	0	∞
f	$+ = \text{sig}(6)$	$+ = \text{sig}(1)$
f'	$+ = \text{sig}(1)$	$+ = \text{sig}(3)$
$-r_1 = 44x - 49$	$- = \text{sig}(-49)$	$+ = \text{sig}(44)$
$-r_2 = -1$	$- = \text{sig}(-1)$	$- = \text{sig}(-1)$
<i>cambios</i>	1	1

³Lo que dice que f no tiene raíces positivas. Agrandando la tabla a

	$-\infty$	0	∞
f	$-$	$+ = \text{sig}(6)$	$+ = \text{sig}(1)$
f'	$+$	$+ = \text{sig}(1)$	$+ = \text{sig}(3)$
$-r_1 = 44x - 49$	$-$	$- = \text{sig}(-49)$	$+ = \text{sig}(44)$
$-r_2 = -1$	$-$	$- = \text{sig}(2)$	$- = \text{sig}(-1)$
<i>cambios</i>	2	1	1

de la nueva columnnavemos que hay una raíz, y que esta es negativa.

Agreguemos más columnas:

	$-\infty$	-2	-1	0	∞
f	$-$	16	9	6	1
f'	$+$	-7	-6	1	3
$-r_1 = 44x - 49$	$-$	-137	-93	-49	$+ = \text{sig}(44)$
$-r_2 = -1$	$-$	-1	-	-1	$- = \text{sig}(-1)$
<i>cambios</i>	2	1	1	1	1

³Desde luego el símbolo de ∞ en la tabla no es un valor en el se evalúa el polinomio, corresponde al comportamiento eventual del polinomio y al signo del coeficiente principal, como sabemos si es positivo el coeficiente principal de un polinomio de grado mayor que 0, el polinomio eventualmente será positivo ($\exists N$ tal que $[(x > N) \Rightarrow (f(x) > 0)]$).

de donde vemos que la raíz negativa es menor que -2 .

Otros valores:

	$-\infty$	-10	-5	0	
f	—	-504	1	6	
f'	+	201	26	1	
$-r_1 = 44x - 49$	—	-391	-269	-49	
$-r_2 = 1$	—	-1	-1	-1	
<i>cambios</i>	2	2	1	1	

de lo anterior, se vé que la raíz negativa está en $(-10, -5)$.

Ejemplo 206 . Acorralemos la raíz negativa en el ejemplo anterior,

$$g(x) =: f(-x) = -x^3 + 5x^2 - x + 6.$$

$$g'(x) = -3x^2 + 10x - 1$$

$$-r_1(x) = -44x + 49$$

$$-r_2(x) = 1.$$

Como $g(5) = 1$ y $g(6) = -36$, hay una raíz entre 5 y 6 . Podemos proseguir, usando el Método de Horner.

Se puede trasladar los polinomios para que la raíz esté en $(0, 1)$, para simplificar las estimaciones de los signos:

$$g(x) = -x^3 + 5x^2 - x + 6 = h(x - 5)$$

$$\begin{array}{r}
 -1 \quad 5 \quad -1 \quad 6 \quad \boxed{+5} \\
 \quad -5 \quad 0 \quad -5 \\
 \hline
 -1 \quad 0 \quad -1 \quad \boxed{1} \\
 \quad -5 \quad -25 \\
 \hline
 -1 \quad -5 \quad \boxed{-26} \\
 \quad -5 \\
 \hline
 -1 \boxed{-10}
 \end{array}$$

Entonces $g(x + 5) = -x^3 - 10x^2 - 26x + 1 = h(x)$.

Evalúemos:

$$\begin{array}{r}
 -1 \quad -10 \quad -26 \quad 1 \quad \boxed{0} \quad -1 \quad -10 \quad -26 \quad 1 \quad \boxed{1} \\
 \quad 0 \quad 0 \quad 0 \\
 \hline
 -1 \quad -10 \quad -27 \quad \boxed{1} \quad \quad -1 \quad -11 \quad -37 \quad \boxed{-36}
 \end{array}$$

$$\begin{array}{r}
 \begin{array}{rrrrrrrrrr}
 -1 & -10 & -26 & 1 & \boxed{.5} & -1 & -10 & -26 & 1 & \boxed{.2} \\
 -0.5 & -5.5 & -15.75 & & & -0.2 & -2.04 & -5.68 & & \\
 \hline
 -1 & -10.5 & -31.5 & \boxed{-14.75} & & -1 & -10.2 & -28.4 & \boxed{-4.68} & \\
 & & & & & & & & & \\
 & -1 & -10 & -26 & 1 & \boxed{.1} & & & & \\
 & -0.1 & -1.01 & -2.701 & & & & & & \\
 \hline
 & -1 & -10.1 & -27.01 & \boxed{-1.701} & & & & & \\
 & & & & & & & & & \\
 & -1 & -10 & -26 & 1 & \boxed{.05} & & & & \\
 & -0.05 & -.5025 & -1.3252 & & & & & & \\
 \hline
 & -1 & -10.05 & -26.503 & \boxed{-.3252} & & & & & \\
 & & & & & & & & & \\
 & -1 & -10 & -26 & 1 & \boxed{.03} & & & & \\
 & -0.03 & -.3009 & -.78903 & & & & & & \\
 \hline
 & -1 & -10.03 & -26.301 & \boxed{.21097} & & & & & \\
 & & & & & & & & & \\
 & -1 & -10 & -26 & 1 & \boxed{.04} & & & & \\
 & -0.04 & -.4016 & -1.0561 & & & & & & \\
 \hline
 & -1 & -10.04 & -26.402 & \boxed{-.0561} & & & & & \\
 \end{array}
 \end{array}$$

La raíz es $5.03\cdots$

Corolario 29 . Sea $f(x)$ un polinomio sin raíces múltiples y sean

$$f, f', -r_1, -r_2, r_3, r_4, \dots$$

los polinomios que se usan para definir la función V .

El número de raíces del polinomio es $V(-\infty) - V(\infty)$.

1. $V(\infty)$ es el número de cambios de signo en los coeficientes principales de

$$f, f', -r_1, -r_2, r_3, r_4, \dots.$$

2. $V(-\infty)$ es el número de cambios de signo en los coeficientes principales de

$$f(-x), f'(-x), -r_1(-x), -r_2(-x), r_3(-x), r_4(-x), \dots.$$

3. $V(0)$ (si 0 no es raíz de f) es el número de cambios de signo en los términos independientes de

$$f, f', -r_1, -r_2, r_3, r_4, \dots.$$

que se obtiene con los signos, el número de raíces positivas, el número de raíces negativas.

4. $V(0) - V(\infty)$ es el número de raíces positivas.

5. $V(-\infty) - V(0)$ es el número de raíces negativas.

Ejercicio 432 . Use el teorema de Sturm para determinar el número de raíces del polinomio

$$x^4 + 12x^2 + 5x - 9.$$

9.11 Regla de los signos de Descartes

Definición 165 . Sea $f(x)$ un polinomio de grado n , consideremos la sucesión

$$f, f', f'', \dots, f^{(n)}.$$

En el punto $c \in \mathbb{R}$ definimos $W_f(c)$ como el número de cambios de signo en la sucesión

$$f(c), f'(c), f''(c), \dots, f^{(n)}(c).$$

Se ignoran los ceros que puedan aparecer.

Lema 34 . Sea $f(x)$ un polinomio de grado n y $a < b \in \mathbb{R}$, que no son raíces de f . Entonces

1. el número de raíces de f en (a, b) es $\stackrel{2}{\equiv} W_f(a) - W_f(b)$, además
2. el número de raíces de f en (a, b) es $\geq W_f(a) - W_f(b)$.

Lema 35 Una raíz de multiplicidad m cuenta como m raíces.

Demostración. Inducción sobre $\text{grad}(f)$.

Base.

Si $f = c$, entonces $W_f(x) = 0$, $\forall x$ por lo que número de raíces de f en (a, b) es $0 - 0 = 0 \stackrel{2}{\equiv} 0$.

Si $f = cx + b$, entonces $f'(x) = c$.

$cx + b > 0 \Leftrightarrow cx > -b$ supongamos que $c > 0$, entonces $cx > -b \Leftrightarrow x > \frac{-b}{c}$.

Por lo tanto $\text{sig}(cx + b) > 0$ si $x > \frac{-b}{c}$.

Por lo tanto

$$W_f(x) = 1 \text{ si } x < \frac{-b}{c},$$

$$W_f(x) = 0 \text{ si } x > \frac{-b}{c}.$$

De aquí se sigue la afirmación.

Si $c < 0$ entonces $cx + b > 0 \Leftrightarrow cx > -b \Leftrightarrow x < \frac{-b}{c}$

$$W_f(x) = 0 \text{ si } x > \frac{-b}{c},$$

$$W_f(x) = 1 \text{ si } x < \frac{-b}{c}.$$

Paso inductivo

Supongamos que $\text{grad}(f) > 1$.

Consideremos ρ una raíz de f . Tomemos un intervalo abierto (α, β) tal que la única raíz de f y de f' que contiene sea ρ .

Consideremos los dos casos: $(f')(\rho) \neq 0$. Entonces f' no tiene raíces en (α, β) , así que permanece con el mismo signo en (α, β) . Por lo tanto, f es siempre creciente o siempre decreciente en (α, β) . En cualquier caso, cambia de signo al transponer ρ de izquierda a derecha supongamos que

	(α, ρ)	ρ	(ρ, β)
$f(x)$	+	0	-
$f'(x)$	-	-	-
número de cambios	1		0

Como f' no tiene raíces en (α, β) podemos aplicar hipótesis de inducción a f' para concluir que $W_{f'}$ tiene la misma paridad en (α, β) . Además podemos suponer que $W_{f'}(\alpha, \rho) > W_{f'}(\rho, \beta)$. Por lo que podemos concluir de la tabla de arriba, que W_f cambia de paridad al transponer ρ en sentido creciente y que además

$$W_f(\alpha, \rho) > W_f(\rho, \beta).$$

Así que $W_f(\alpha, \rho) - W_f(\rho, \beta) \stackrel{2}{\equiv} 1$ que es la multiplicidad de ρ como raíz de f , al no ser raíz de f' .

Supongamos ahora que $f'(\rho) = 0$.

Consideremos dos posibilidades: f cambia de signo al cruzar ρ en sentido creciente y el otro caso, que es que no cambie de signo. Consideremos ambos casos en las dos tablas siguientes;

	(α, ρ)	ρ	(ρ, β)		(α, ρ)	ρ	(ρ, β)
$f(x)$	-	0	+	$f(x)$	-	0	-
$f'(x)$	+	0	+	$f'(x)$	+	0	-
número de cambios	1		0	número de cambios	1		0

Aplicando hipótesis de Inducción a f' concluimos que

$$W_{f'}(k) - W_{f'}(m) \stackrel{2}{\equiv} \text{multiplicidad de } \rho, \text{ como raíz de } f', \\ \text{si } k \in (\alpha, \rho), m \in (\rho, \beta) \text{ y} \\ W_{f'}(k) - W_{f'}(m) \geq 0.$$

Por lo tanto, de

$$W_f(x) = W_{f'}(x) + 1 \text{ para } x \in (\alpha, \rho) \\ W_f(x) = W_{f'}(x) \text{ para } x \in (\rho, \beta)$$

tenemos que

$$\text{Si } k \in (\alpha, \rho), m \in (\rho, \beta), \text{ entonces} \\ W_f(k) - W_f(m) = (W_{f'}(k) + 1) - W_{f'}(m) \stackrel{2}{\equiv} \\ \stackrel{2}{\equiv} (\text{multiplicidad de } \rho, \text{ como raíz de } f') + 1 \\ = (\text{multiplicidad de } \rho, \text{ como raíz de } f).$$

Además,

$$W_f(k) - W_f(m) = (W_{f'}(k) + 1) - W_{f'}(m) \geq 1.$$

Hemos demostrado que al transponer una raíz ρ de f , la paridad decrece en un número $\stackrel{2}{\equiv}$ (multiplicidad de ρ). ■

Corolario 30 (*Regla de los signos de Descartes*).

1. El número de raíces positivas de un polinomio $f(x)$ y tal que $f(0) \neq 0$ es $\stackrel{2}{\equiv} W(0) =$ número de cambios de signo en la sucesión de términos independientes de f, f', f'', \dots .

Además, el número de raíces positivas de un polinomio $f(x)$ y tal que $f(0) \neq 0$ es mayor o igual que

$$W(0) = \frac{\text{número de cambios de signo}}{\text{en la sucesión de términos}} = \frac{\text{número de cambios de signo}}{\text{en los coeficientes de}} \\ \text{independientes de} \quad f, f', f'', \dots$$

2. El número de raíces negativas de un polinomio $f(x)$ y tal que $f(0) \neq 0$ es $\overset{2}{\equiv} W(-\infty) - W(0)$, donde $W(-\infty)$ es el número de cambios de signo en la sucesión de los coeficientes principales de

$$(-1^n) f, (-1^{n-1}) f', (-1^{n-2}) f'', \dots$$

Es lo mismo que el número de cambios de signo en

$$\underbrace{+, -, +, -, \dots}_{n+1 \text{ símbolos}}$$

es decir n .

3. El número de raíces de un polinomio $f(x)$ es $\overset{2}{\equiv} W(-\infty) = \text{grad}(f)$.

Demostración. Todo lo que hay que notar es que si

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$$

entonces

$$f'(x) = n a_n x^{n-1} + (n-1) a_{n-1} x^{n-2} + \dots + a_1.$$

Así que eventualmente (para todos los valores mayores que una cierta $N \in \mathbb{R}$) f tendrá el mismo signo que a_n , $f'(x)$ tendrá el mismo signo que na_n , etc. Eventualmente $W(x) = 0$, y esto se expresa simbólicamente como $W(\infty) = 0$.

Claramente la sucesión

$$f(0), f'(0), (f''(0)), \dots$$

es

$$a_0, a_1, a_2, \dots$$

Por último para los valores menores que M , para alguna M suficientemente pequeña, $f(x)$ tendrá el mismo signo que $(-1)^n \text{signo}(a_n)$, $f'(x)$ tendrá el mismo signo que $(-1)^{n-1} \text{signo}(a_n)$, etc. ■

Ejemplo 207 . $x^3 - 7x^2 - 7$

$W(\infty) = 0, W(0) = 1, W(-\infty) = 3$, entonces

$$\begin{array}{ll} \text{número de} & \stackrel{2}{\equiv} 2 \stackrel{2}{\equiv} 0 \\ \text{raíces negativas} & \\ \text{número de} & \stackrel{2}{\equiv} 1 \\ \text{raíces positivas} & \\ \text{número de} & \stackrel{2}{\equiv} 3 \stackrel{2}{\equiv} 1 \\ \text{raíces} & \end{array}$$

posibilidades: $\begin{cases} 0 \text{ negativas y 1 positivas,} \\ 2 \text{ negativas, 1 positivas,} \\ 0 \text{ negativas, 3 positivas.} \end{cases}$

Con el teorema de Sturm, decidimos la cuestión:

$$\begin{array}{r} x^3 - 7x^2 - 7 \\ 3x^2 - 14x \\ \frac{98}{9}x + 7 \\ -1 \end{array}$$

Entonces

$$\begin{array}{l} V(\infty) = 1 \\ V(0) = 2 \\ V(-\infty) = 2 \end{array}$$

Es decir, $x^3 - 7x^2 - 7$ tiene una sola raíz, que es positiva.

Ejemplo 208 . $(x - 2)^2(x - 1)(x^2 + x + 1) = x^5 - 4x^4 + 4x^3 - x^2 + 4x - 4$.

Número de raíces positivas: $\stackrel{2}{\equiv} 5$

Número de raíces negativas: $\stackrel{2}{\equiv} 5 - 5 \stackrel{2}{\equiv} 0$.

Ejemplo 209 . $(x + 2)^2(x - 1)(x^2 + x + 1) = x^5 + 4x^4 + 4x^3 - x^2 - 4x - 4$.

Número de raíces positivas: $\stackrel{2}{\equiv} 1$

Número de raíces negativas: $\stackrel{2}{\equiv} 5 - 1 \stackrel{2}{\equiv} 4$.

Corolario 31 . Si $f(x)$ es un polinomio, entonces el número de raíces de f es congruente con $\text{grad}(f)$ módulo 2.

Ejercicio 433 . Encuentre las raíces de $x^3 + 6x^2 - 24x + 160$, que tiene $2 - 2\sqrt{3}i$ como raíz.

Ejercicio 434 . Encuentre las raíces de $x^3 + (1 - 2i)x^2 - (1 + 2i)x - 1$, que tiene una raíz doble.

Ejercicio 435 . Encuentre las raíces de $x^5 - 3x^4 + 4x^2 - 4x - 4$ que tiene $1 + i$ como raíz doble.

Ejercicio 436 . Encuentre las raíces de $x^3 - x^2 - 9x + 9$, que tiene una raíz que es el inverso aditivo de otra.

Ejercicio 437 . Encuentre las raíces de $x^3 + 2x^2 - x - 2$, si una raíz es el doble de la otra.

Ejercicio 438 . Encuentre las raíces de $3x^3 - 52x^2 + 107x - 30$, si el producto de dos de sus raíces es 5.

9.12 Raíces racionales

Supongamos que $f(x) = a_nx^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0$ es un polinomio con coeficientes racionales, es decir que

$$a_i = \frac{b_i}{c_i}, \quad b_i \in \mathbb{Z}, \quad c_i \in \mathbb{Z} \setminus \{0\}, \quad \text{para } i \in \{0, \dots, n\}.$$

Multipliquemos f por $[c_0; c_1; \dots; c_n]$ y hagamos

$$g = [c_0; c_1; \dots; c_n] f,$$

notemos ahora que los coeficientes de g son de la forma

$$\frac{b_i}{c_i} [c_0; c_1; \dots; c_n] = b_i \frac{[c_0; c_1; \dots; c_n]}{c_i} \in \mathbb{Z}.$$

Además es claro que $f(x)$ y $g(x)$ tienen las mismas raíces.

Entonces tenemos que para encontrar las raíces de un polinomio con coeficientes racionales, basta saber encontrar las raíces de polinomios con coeficientes enteros.

Observación 145 . Si $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ es un polinomio y $\frac{\alpha}{\beta}$ es una raíz racional de f con $(\alpha; \beta) = 1$, entonces $\beta \mid a_n$ y $\alpha \mid a_0$.

Demuestra. Evalúemos f en $\frac{\alpha}{\beta}$ y después multipliquemos por β^n :

$$a_n \left(\frac{\alpha}{\beta} \right)^n + a_{n-1} \left(\frac{\alpha}{\beta} \right)^{n-1} + \cdots + a_1 \left(\frac{\alpha}{\beta} \right) + a_0 = 0,$$

$$a_n \alpha^n + a_{n-1} \alpha^{n-1} \beta + \cdots + a_1 \alpha \beta^{n-1} = -a_0 \beta^n, \quad (9.5)$$

de lo anterior vemos que α es un factor de $-a_0 \beta^n$, pero entonces lo es de a_0 , ya que $(\alpha, \beta) = 1$. (Se han usado las siguientes proposiciones: $(\alpha, \beta) = 1 \Rightarrow (\alpha, \beta^n) = 1$; y $\alpha \mid cd, (\alpha; c) = 1 \Rightarrow \alpha \mid c$.)

Si expresamos la ecuación 9.5 en la forma

$$a_n \alpha^n = -a_{n-1} \alpha^{n-1} \beta - \cdots - a_1 \alpha \beta^{n-1} - a_0 \beta^n,$$

vemos que β es un factor de $a_n \alpha^n$ y por lo tanto lo es de a_n . ■

Ejemplo 210 . Encontremos las raíces racionales de $x^3 + \frac{17}{6}x^2 - \frac{2}{3}x - \frac{1}{2}$, multipliquemos por 6:

$$6 \left(x^3 + \frac{17}{6}x^2 - \frac{2}{3}x - \frac{1}{2} \right) = 6x^3 + 17x^2 - 4x - 3.$$

Las raíces racionales, si las hay, deben ser de la forma

$$\pm \frac{1}{6}, \pm \frac{1}{3}, \pm \frac{1}{2} \pm \frac{1}{1}, \frac{3}{6}, \pm \frac{3}{3} \pm \frac{3}{2} \pm \frac{3}{1},$$

eliminando redundancias:

$$\pm \frac{1}{6}, \pm \frac{1}{3}, \pm \frac{1}{2} \pm \frac{1}{1}, \pm \frac{3}{2} \pm \frac{3}{1}$$

La división

$$\begin{array}{r} 6 \quad 17 \quad -4 \quad -3 \parallel -3 \\ \quad -18 \quad 3 \quad 3 \\ \hline 6 \quad -1 \quad -1 \quad \boxed{0} \end{array}$$

muestra que -3 es una raíz, y además que $6x^2 - x - 1 = (x + 3)(6x^2 - x - 1)$, por lo que las otras raíces son

$$\frac{1 + \sqrt{1 + 24}}{12} = \frac{1}{2}, \quad \frac{1 - \sqrt{1 + 24}}{12} = -\frac{1}{3}.$$

Corolario 32 . Si un polinomio con coeficientes enteros es mónico, entonces sus raíces racionales son enteras.

Demostración. Si $\frac{\alpha}{\beta} \in \mathbb{Z}$, $\beta \in \mathbb{Z} \setminus \{0\}$ es una raíz racional, el teorema anterior nos dice que $\beta \mid 1$. ■

Corolario 33 . Sean $m, n \in \mathbb{N}$, entonces $\sqrt[n]{m} \in \mathbb{Q} \Rightarrow \sqrt[n]{m} \in \mathbb{Z}$.

Demostración. $\sqrt[n]{m}$ es solución de $x^n = m$, es decir es raíz de $x^n - m$. Concluimos del Corolario anterior que si $\sqrt[n]{m}$ es racional entonces es entero. ■

Ejemplos 211

1. No hay un entero cuyo cuadrado sea 2. Obsérvese la sucesión de los enteros cuadrados

$$0, 1, 4, 9, 16,$$

luego $\sqrt{2}$ no es racional.

Por la misma razón, $\sqrt{3}, \sqrt{5}, \sqrt{6}, \sqrt{7}, \sqrt{8}, \sqrt{10}, \dots$ son irracionales.

2. Obsérvese la sucesión de los enteros que son cubos (de enteros) positivos

$$0, 1, 8, 27, 64, 125, \dots$$

entonces $\sqrt[3]{2}, \sqrt[3]{3}, \sqrt[3]{4}, \sqrt[3]{60}, \sqrt[3]{100}$ son todos números irracionales.

9.13 Coeficientes y raíces

Sea $f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$ un polinomio con n raíces r_1, r_2, \dots, r_n entonces

$$(x - r_1)(x - r_2) \cdots (x - r_n) \mid f(x)$$

debe ser claro de la coincidencia de los grados y de los coeficientes que

$$a_n (x - r_1)(x - r_2) \cdots (x - r_n) = f(x).$$

Es claro que el coeficiente de x^i es

$$a_n \left(\sum_{\{B \subseteq \{1, 2, \dots, n\} \mid |B|=i\}} \left(\prod_{j \notin B} (-r_j) \right) \right). \quad (9.6)$$

En B están los índices de los factores de donde se toma x .

Por ejemplo, en

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

calculemos el coeficiente de grado 2, el número de subconjuntos $\{1, 2, 3, 4\}$ con dos elementos es $\binom{4}{2} = 6$, a saber:

$\{1, 2\}$, que contribuye con $3(-4 \cdot -5)$

$\{1, 3\}$, que contribuye con $3(-2 \cdot -5)$

$\{1, 4\}$, que contribuye con $3(-2 \cdot -4)$

$\{2, 3\}$, que contribuye con $3(-1 \cdot -5)$

$\{2, 4\}$, que contribuye con $3(-1 \cdot -4)$

$\{3, 4\}$, que contribuye con $3(-1 \cdot -2)$

Por lo tanto el coeficiente de x^2 en $3(x-1)(x-2)(x-4)(x-5)$:es

$$\left(3(-4 \cdot -5) + 3(-2 \cdot -5) + 3(-2 \cdot -4) + 3(-1 \cdot -5) + 3(-1 \cdot -4) + 3(-1 \cdot -2) \right) = 147.$$

Ejercicio 439 . Encuentre los coeficientes de x y de x^3 en el polinomio

$$3(x-1)(x-2)(x-4)(x-5),$$

usando 9.6.

Ejercicio 440 . Muestre que 9.6 es correcta aunque haya raíces repetidas, en particular compruebe que el coeficiente de x^i en $(x-a)^n$ es $\binom{n}{i} (-a)^{n-i}$.

Ejercicio 441 . Encuentre el coeficiente de x^i en $(x-a)^n(x-b)^m$, $a \neq b$.

De esta manera vemos que las n raíces de un polinomio de grado n determinan los coeficientes (salvo un factor constante, correspondiente al coeficiente principal).

Es muy natural el problema inverso ¿los coeficientes de un polinomio determinan las raíces? Desde luego, si los coeficientes determinan al polinomio, seguramente determinan sus raíces. Pero la pregunta debiera precisarse más.

Notemos que para obtener los coeficientes conociendo las raíces, las operaciones que hacemos son sumas, restas, productos, divisiones y potencias, además solamente se efectúan un número *finito de operaciones* que además puede expresarse en una fórmula tal como 9.6.

Con estas apreciaciones la pregunta que nos hacemos es ¿existe una fórmula que involucre un número finito de operaciones de sumas, productos, restas, divisiones y extracción de raíces, que aplicada a los coeficientes de un polinomio produzcan sus raíces?

Note que en el problema anterior es de crucial importancia la petición de finitud.

Este problema fascinó a la humanidad. Es un hecho sorprendente que una fórmula así existe para polinomios de grados 2, 3 y 4. Pero que no hay una fórmula que funcione para cualquier polinomio de grado n , para $n \geq 5$.

La fórmula para las raíces de un polinomio de grado 2 era conocida por los babilonios:

$$ax^2 + bx + c$$

tiene las raíces

$$\frac{-b + \sqrt{b^2 - 4ac}}{2a}, \frac{-b - \sqrt{b^2 - 4ac}}{2a}.$$

Las fórmulas para encontrar las raíces de los polinomios de grado 3 y 4 fueron encontradas hasta el Renacimiento italiano, por Scipio del Ferro, Tartaglia, Cardano, Ferrari.

Este hecho, fué verdaderamente un renacimiento para las Matemáticas, pues fue el primer descubrimiento importante desde la época de los griegos antiguos y de los árabes, quienes habían intentado la solución, obteniendo solamente resultados parciales, véase [44].

La no solubilidad del polinomio de grado 5 (y de los de grado mayor que 5) fué descubierta hasta el siglo XIX por Ruffini, Abel y Galois.

Hay exposiciones sencillas sobre como encontrar las fórmulas para las raíces de los polinomios de 3o.y 4o. grado véanse [43], [2], aquí seguimos la de Anglin y Lambek.

9.14 Polinomios de tercer grado

Tomemos el polinomio $f(x) = x^3 + ax^2 + bx + c$ con coeficientes en \mathbb{R} . Mediante una “traslación” podemos eliminar el coeficiente en x^2 :

$$f(x-d) = x^3 + (a-3d)x^2 + (-2ad+3d^2+b)x + (-bd+c+ad^2-d^3),$$

así que tomando d tal que $a-3d=0$, es decir, tomando $d=\frac{a}{3}$, tenemos

$$f\left(x-\frac{a}{3}\right) = x^3 + \left(b - \frac{1}{3}a^2\right)x + \left(-\frac{1}{3}ba + c + \frac{2}{27}a^3\right).$$

Que podemos escribir en la forma

$$g(x) = x^3 + \left(b - \frac{1}{3}a^2\right)x + \left(-\frac{1}{3}ba + c + \frac{2}{27}a^3\right).$$

$$\alpha = b - \frac{1}{3}a^2, \beta = \left(-\frac{1}{3}ba + c + \frac{2}{27}a^3\right).$$

Basta pues, saber encontrar las raíces de polinomios de la forma⁴ de g .

⁴Por otro lado, cualquier polinomio $g(x) = x^3 + \alpha x + \beta$ proviene de un polinomio de la forma de f , al eliminar el coeficiente de grado 2 mediante una traslación:

Si tomamos el sistema de ecuaciones

$$\begin{aligned} b - \frac{1}{3}a^2 &= \alpha \\ \left(-\frac{1}{3}ba + c + \frac{2}{27}a^3\right) &= \beta, \end{aligned}$$

$$b - \frac{1}{3}a^2 = \alpha \Rightarrow b = \frac{1}{3}a^2 + \alpha.$$

Sustituyendo $b = \frac{1}{3}a^2 + \alpha$ en

$$\left(-\frac{1}{3}\left(\frac{1}{3}a^2 + \alpha\right)a + c + \frac{2}{27}a^3\right) = \beta,$$

obtenemos

$$-\frac{1}{27}a^3 - \frac{1}{3}a\alpha + c = \beta,$$

es decir que

$$c = \frac{1}{27}a^3 + \frac{1}{3}a\alpha + \beta.$$

Por lo tanto los valores de b y de c están determinados por los de α , β y a , a se toma como un parámetro libre (para cada valor de a se obtienen los de b y de c).

Encontremos las raíces de

$$g(x) = x^3 + \alpha x + \beta.$$

Hagamos $x = y + z$, entonces

$$\begin{aligned} g(y + z) &= y^3 + 3y^2z + 3yz^2 + z^3 + \alpha y + \alpha z + \beta = \\ &= y^3 + 3yz(y + z) + \alpha(y + z) + z^3 + \beta = \\ &= y^3 + z^3 + (3yz + \alpha)(y + z) + \beta. \end{aligned}$$

De la última ecuación, vemos que si $3yz + \alpha = 0$, es decir si hacemos

$$y = -\frac{1}{3}\frac{\alpha}{z},$$

entonces la ecuación resulta muy sencilla:

$$g(y + z) = \left(-\frac{1}{3}\frac{\alpha}{z}\right)^3 + z^3 + \beta = -\frac{1}{27}\frac{\alpha^3}{z^3} + z^3 + \beta. \text{ Multiplicando por } z^3$$

obtenemos:

$$z^6 + \beta z^3 - \frac{1}{27}\alpha^3$$

que es una ecuación cuadrática en z^3 , de aquí que

$$\begin{aligned} z^3 &= \frac{-\beta \pm \sqrt{\beta^2 + \frac{4}{27}\alpha^3}}{2} = \frac{-\beta \pm \sqrt{\frac{4}{4}\beta^2 + \frac{4}{27}\alpha^3}}{2} = \\ &= \frac{-\beta \pm 2\sqrt{\frac{\beta^2}{4} + \frac{\alpha^3}{27}}}{2} = \frac{-\beta}{2} \pm \sqrt{\frac{\beta^2}{4} + \frac{\alpha^3}{27}}, \end{aligned}$$

de aquí obtenemos z , de $y = -\frac{1}{3}\frac{\alpha}{z}$, obtenemos y , y por último, $x = y + z$.

9.14.1 El discriminante y número de raíces reales

Llamaremos a $\frac{\beta^2}{4} + \frac{\alpha^3}{27}$ (o equivalentemente a $27\beta^2 + 4\alpha^3$), el discriminante de $g(x) = x^3 + \alpha x + \beta$.

Usaremos el teorema de Sturm, para ver que tiene que ver el discriminante con el número de raíces reales.

Si aplicamos el algoritmo de Euclides a $g(x)$ y $g'(x)$ y colocamos los signos como se indica en el teorema de Sturm, obtenemos los polinomios

$$x^3 + \alpha x + \beta, 3x^2 + \alpha, -\frac{2}{3}\alpha x - \beta, -4\alpha^3 - 27\beta^2,$$

donde puede ser que no aparezca $-4\alpha^3 - 27\beta^2$, si resulta cero.

Los casos discriminados son los siguientes:

1. $27\beta^2 + 4\alpha^3 < 0$ y $\alpha < 0$

	$-\infty$	∞
$x^3 + \alpha x + \beta$	-	+
$3x^2 + \alpha$	+	+
$-\frac{2}{3}\alpha x - \beta$	-	+
$-4\alpha^3 - 27\beta^2$	+	+
V	3	0.

Tres raíces reales.

2. $27\beta^2 + 4\alpha^3 < 0$ y $\alpha > 0$ entonces

	$-\infty$	∞
$x^3 + \alpha x + \beta$	-	+
$3x^2 + \alpha$	+	+
$-\frac{2}{3}\alpha x - \beta$	+	-
$-4\alpha^3 - 27\beta^2$	+	+
V	1	2.

No es posible.

3. $27\beta^2 + 4\alpha^3 > 0$ y $\alpha > 0$ entonces

	$-\infty$	∞
$x^3 + \alpha x + \beta$	-	+
$3x^2 + \alpha$	+	+
$-\frac{2}{3}\alpha x - \beta$	+	-
$-4\alpha^3 - 27\beta^2$	-	-
V	2	1.

Una raíz real.

4. Si $27\beta^2 + 4\alpha^3 > 0$ y $\alpha < 0$ entonces

	$-\infty$	∞
$x^3 + \alpha x + \beta$	-	+
$3x^2 + \alpha$	+	+
$-\frac{2}{3}\alpha x - \beta$	-	+
$-4\alpha^3 - 27\beta^2$	-	-
V	2	1.

Una raíz real.

5. $27\beta^2 + 4\alpha^3 = 0$, $\alpha \neq 0$.

Como en este caso $\frac{2}{3}\alpha x + \beta = (g(x); g'(x))$ tenemos que la raíz de $\frac{2}{3}\alpha x + \beta$, $x = -\frac{3\beta}{2\alpha}$ es una raíz de multiplicidad 2, la otra raíz la obtenemos de

$$\frac{g(x)}{\left(x + \frac{3\beta}{2\alpha}\right)^2}.$$

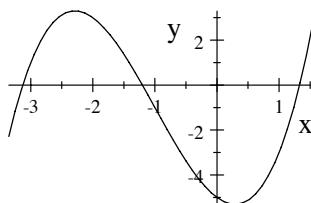
6. $\beta = 0 = \alpha$, en este caso $g(x) = x^3$, tiene 0 como raíz triple.

En resumen:

Proposición 40 . Sea $g(x) = x^3 + \alpha x + \beta \in \mathbb{R} [\curvearrowright]$, $\alpha \neq 0$, entonces

1. $27\beta^2 + 4\alpha^3 < 0 \Rightarrow g(x)$ tiene 3 raíces reales.
2. $27\beta^2 + 4\alpha^3 > 0 \Rightarrow g(x)$ tiene 1 raíz real (y dos complejas).
3. $27\beta^2 + 4\alpha^3 = 0 \Rightarrow g(x)$ tiene una raíz doble $(-\frac{3\beta}{2\alpha})$ y otra raíz real.

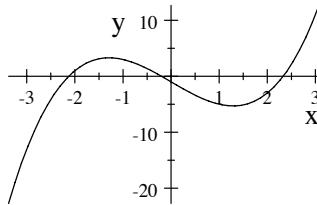
Ejemplo 212 . $h(x) = x^3 + 3x^2 - 2x - 5$



$$x^3 + 3x^2 - 2x - 5$$

$h(x+d) = x^3 + (3+3d)x^2 + (6d+3d^2-2)x - 2d - 5 + 3d^2 + d^3$,
así que tomamos $d = -1$,

$$h(x-1) = x^3 - 5x - 1.$$



$$x^3 - 5x - 1$$

El discriminante es $27\beta^2 + 4\alpha^3 = 27(-1)^2 + 4(-5)^3 = -473$, que es negativo,
por lo que hay 3 raíces reales

Hagamos $g(x) = x^3 - 5x - 1$ y hagamos $x = y + z$, entonces

$$\begin{aligned} g(y+z) &= y^3 + 3y^2z + 3yz^2 + z^3 - 5y - 5z - 1 = \\ &= y^3 + z^3 + (3yz - 5)(y+z) - 1. \end{aligned}$$

Tomamos $3yz - 5 = 0$, así que $y = \frac{5}{3z}$, sustituyendo:

$$\begin{aligned} \left(\frac{5}{3z}\right)^3 + z^3 + \left(3\left(\frac{5}{3z}\right)z - 5\right)\left(\left(\frac{5}{3z}\right) + z\right) - 1 &= \\ &= \frac{125}{27z^3} + z^3 - 1 \end{aligned}$$

Así que debemos resolver $z^6 - z^3 + \frac{125}{27} = 0$, de donde

$$z^3 = \frac{1 \pm \sqrt{1 - 4 * \frac{125}{27}}}{2}$$

Por lo tanto $z^3 = \frac{1}{2} \pm \frac{1}{18}i\sqrt{1419}$, una solución es:

$$z = \frac{1}{6} \sqrt[3]{\left(108 + 12i\sqrt{1419}\right)}$$

entonces

$$y = \frac{5}{3 * \frac{1}{6} \sqrt[3]{(108 + 12i\sqrt{1419})}} = \frac{10}{\sqrt[3]{(108 + 12i\sqrt{1419})}}$$

por lo que

$$\begin{aligned} x &= \frac{10}{\sqrt[3]{(108 + 12i\sqrt{1419})}} + \frac{1}{6} \sqrt[3]{(108 + 12i\sqrt{1419})} = \\ &= \frac{1}{6} \frac{60 + \left(\sqrt[3]{(108 + 12i\sqrt{1419})} \right)^2}{\sqrt[3]{(108 + 12i\sqrt{1419})}} \end{aligned}$$

es una raíz de $h(x - 1)$, de donde

$$\frac{1}{6} \frac{60 + \left(\sqrt[3]{(108 + 12i\sqrt{1419})} \right)^2}{\sqrt[3]{(108 + 12i\sqrt{1419})}} - 1$$

que es aproximadamente 1.3301 es una raíz de h . Teniendo esta raíz reducimos el polinomio a uno de grado 2.

Ejemplo 213 . $f(x) = x^3 - 7x - 7$.

El discriminante es $27(-7)^2 + 4(-7)^3 = -49$, así que debe haber tres raíces reales.

Hagamos $x = y + z$, entonces

$$\begin{aligned} f(y + z) &= y^3 + 3y^2z + 3yz^2 + z^3 - 7y - 7z - 7 = \\ &= y^3 + z^3 + (3yz - 7)(y + z) - 7 \end{aligned}$$

Hagamos $y = \frac{7}{3z}$

$$\begin{aligned} \left(\frac{7}{3z} \right)^3 + z^3 + \left(3 \left(\frac{7}{3z} \right) z - 7 \right) \left(\left(\frac{7}{3z} \right) + z \right) - 7 &= \\ &= \frac{343}{27z^3} + z^3 - 7 \end{aligned}$$

Así que hay que resolver $z^6 - 7z^3 + \frac{343}{27}$, de donde

$$z^3 = \frac{7 \pm \sqrt{49 - 4 * \frac{343}{27}}}{2} = \frac{7}{2} \pm \frac{7}{18}i\sqrt{3}$$

$z^3 = \frac{7}{2} + \frac{7}{18}i\sqrt{3}$, una solución es $\frac{1}{6}\sqrt[3]{(756 + 84i\sqrt{3})}$, luego

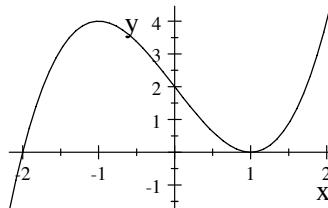
$$y = \frac{7}{3 * \frac{1}{6}\sqrt[3]{(756 + 84i\sqrt{3})}},$$

luego

$$\begin{aligned} \frac{1}{6}\sqrt[3]{(756 + 84i\sqrt{3})} + \frac{7}{3 * \frac{1}{6}\sqrt[3]{(756 + 84i\sqrt{3})}} &= \\ &= \frac{1}{6} \frac{\left(\sqrt[3]{(756 + 84i\sqrt{3})}\right)^2 + 84}{\sqrt[3]{(756 + 84i\sqrt{3})}} \end{aligned}$$

este valor, aproximadamente 3.0489 es muy cercano a una raíz de $f(x)$. De hecho, $f(3.0489)$ es $-.00036218$.

Ejemplo 214 . Si $x^3 - 3x + 2$, entonces $27\beta^2 + 4\alpha^3 = 27 \cdot 2^2 + 4 \cdot (-3)^3 = 0$,



$$x^3 - 3x + 2$$

y

$$-\frac{3\beta}{2\alpha} = -\frac{3}{2} \frac{2}{-3} = 1$$

es una raíz doble. La otra es la raíz de

$$\frac{x^3 - 3x + 2}{(x - 1)^2} = x + 2,$$

es decir, la otra raíz es -2 .

Ejemplo 215 . $f(x) = x^3 - 15x - 4$, hagamos

$$x = y + z,$$

$$\begin{aligned} f(y + z) &= y^3 + 3y^2z + 3yz^2 + z^3 - 15y - 15z - 4 = \\ &= y^3 + z^3 + (3yz - 15)(y + z) - 4. \end{aligned}$$

$$\text{Hagamos } y = \frac{5}{z},$$

$$\left(\frac{5}{z}\right)^3 + z^3 + \left(3\left(\frac{5}{z}\right)z - 15\right)\left(\left(\frac{5}{z}\right) + z\right) - 4 = \frac{125}{z^3} + z^3 - 4.$$

Obtenemos $z^6 - 4z^3 + 125$,

$$z^3 = \frac{4 \pm \sqrt{16 - 500}}{2} = 2 + 11i$$

una solución es $z = \sqrt[3]{2 + 11i}$, por lo que $y = \frac{5}{\sqrt[3]{2 + 11i}}$ y así $x = \frac{5}{\sqrt[3]{2 + 11i}} + \sqrt[3]{2 + 11i} = \frac{5 + \left(\sqrt[3]{(2 + 11i)}\right)^2}{\sqrt[3]{(2 + 11i)}}$ es una raíz de f .

Otra solución de $z^3 = 2 + 11i$ es $2 + i$ así que

$$\frac{5}{2+i} + 2 + i = 4$$

es una raíz de f . Por último, la tercera raíz de f es el conjugado de la primera que encontramos, es decir: $2 - 11i$.

9.15 Polinomios de grado cuatro

Consideremos

$$x^4 + ax^3 - bx^2 - cx - d.$$

Queremos resolver

$$x^4 + ax^3 - bx^2 - cx - d = 0$$

que equivale a resolver

$$x^4 + ax^3 = bx^2 + cx + d.$$

Completemos a un cuadrado en el lado izquierdo: $ax^3 = 2x^2z$, así que $z = \frac{1}{2}ax$, entonces sumamos de cada lado $(\frac{1}{2}ax)^2$, obtenemos

$$\begin{aligned} \left(x^2 + \frac{1}{2}ax\right)^2 &= bx^2 + \left(\frac{1}{2}ax\right)^2 + cx + d = \\ &= \left(b + \frac{1}{4}a^2\right)x^2 + cx + d. \end{aligned}$$

Sumemos $2\left(x^2 + \frac{1}{2}ax\right)t + t^2$ de cada lado para obtener

$$\begin{aligned} \left(x^2 + \frac{1}{2}ax + t\right)^2 &= bx^2 + \left(\frac{1}{2}ax\right)^2 + cx + d + 2\left(x^2 + \frac{1}{2}ax\right)t + t^2 = \\ &= \left(\frac{1}{4}a^2 + b + 2t\right)x^2 + (at + c)x + t^2 + d = \\ &= Ax^2 + Bx + C, \end{aligned}$$

$Ax^2 + Bx + C$ es un cuadrado si $B^2 - 4AC = 0$, es decir si

$$\begin{aligned} 0 &= (at + c)^2 - 4\left(\frac{1}{4}a^2 + b + 2t\right)t^2 + d = \\ &= -8t^3 - 4bt^2 + 2act + c^2 + d \Leftrightarrow \\ &= 8t^3 + 4bt^2 - 2act - c^2 - d \end{aligned}$$

que es una ecuación de grado 3 que ya sabemos resolver. Supongamos que τ es una raíz de la ecuación anterior entonces tenemos que

$$\begin{aligned} \left(x^2 + \frac{1}{2}ax + \tau\right)^2 &= A\left(x - \frac{B}{2A}\right)^2 = \\ &= \left(\frac{1}{4}a^2 + b + 2\tau\right)\left(x - \frac{(a\tau + c)}{2(\frac{1}{4}a^2 + b + 2\tau)}\right)^2 = \end{aligned}$$

Así que hay que resolver

$$\left(x^2 + \frac{1}{2}ax + \tau \right) = \sqrt{\left(\frac{1}{4}a^2 + b + 2\tau \right)} \left(x - \frac{(a\tau + c)}{2\left(\frac{1}{4}a^2 + b + 2\tau \right)} \right),$$

es decir, hay que resolver

$$x^2 + \left(-\frac{1}{2}\sqrt{(a^2 + 4b + 8\tau)} + \frac{1}{2}a \right) x + \tau + \frac{1}{2}\sqrt{(a^2 + 4b + 8\tau)} \frac{a\tau + c}{\frac{1}{2}a^2 + 2b + 4\tau},$$

que es una ecuación de segundo grado.

Ejemplo 216 . Resolver $x^4 + 2x^3 - 12x^2 - 10x + 3 = 0$, equivale a resolver

$$x^4 + 2x^3 = 12x^2 + 10x - 3.$$

Para completar un cuadrado del lado izquierdo, sumamos x^2 de cada lado :

$$x^4 + 2x^3 + x^2 = 13x^2 + 10x - 3.$$

Por lo tanto

$$(x^2 + x)^2 = 13x^2 + 10x - 3.$$

Sumamos $t^2 + 2t(x^2 + x)$ de cada lado:

$$\begin{aligned} (x^2 + x + t)^2 &= 13x^2 + 10x - 3 + t^2 + 2t(x^2 + x) = \\ &= (13 + 2t)x^2 + (2t + 10)x - 3 + t^2. \end{aligned}$$

Para que el lado derecho ($Ax^2 + Bx + C$) resulte un cuadrado, pedimos que $B^2 - 4AC = 0$, es decir que

$$0 = (2t + 10)^2 - 4(13 + 2t)(-3 + t^2)$$

así que

$$0 = 8t^3 + 48t^2 - 64t - 256$$

o bien

$$0 = t^3 + 6t^2 - 8t - 32.$$

Es fácil ver que -2 es una raíz del polinomio anterior (que ya sabemos resolver, por lo que proseguimos nuestro ejemplo). Sustituyamos -2 en

$$(x^2 + x + t)^2 = (13 + 2t)x^2 + (2t + 10)x - 3 + t^2$$

para obtener

$$\begin{aligned} (x^2 + x + (-2))^2 &= (13 + 2(-2))x^2 + (2(-2) + 10)x - 3 + (-2)^2 = \\ &= 9x^2 + 6x + 1 = (3x + 1)^2. \end{aligned}$$

Por lo tanto

$$(x^2 + x + (-2)) = (3x + 1),$$

así que

$$x^2 - 2x - 3 = (x + 1)(x - 3).$$

Cuyas raíces son -1 y 3 . De aquí que las otras dos raíces son las raíces de

$$\frac{x^4 + 2x^3 - 12x^2 - 10x + 3}{x^2 - 2x - 3} = x^2 + 4x - 1,$$

cuyas raíces son $-2 + \sqrt{5}$ y $-2 - \sqrt{5}$. El conjunto de raíces es

$$\{-1, 3, -2 + \sqrt{5}, -2 - \sqrt{5}\}.$$

Ejercicio 442 . Encontrar las raíces de $x^4 + x^3 - 2x^2 + 3x - 1$.

Ejercicio 443 . Encontrar las raíces de $x^4 + x^3 - 6x^2 - x + 1$.

9.16 Otra construcción de \mathbb{C}

Observación 146 . Sean $f, h \in \mathbb{R}[x]$, f irreducible y mónico. Entonces $(f; h) = f$ o $(f; h) = 1$.

Demuestra $\mathbf{r}.$ $f = (f; h)k$ para alguna $k \in \mathbb{R}[x]$. Entonces $\text{grad}(f; h) = 0$ o $\text{grad}(k) = 0$.

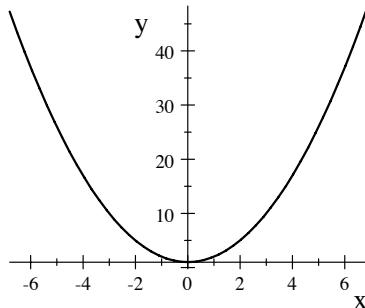
Si $\text{grad}(k) = 0$, entonces k es el coeficiente principal de $(f; h)k$, que es el coeficiente principal de f que es 1, así $k = 1$ y $f = (f; h)$.

Si $\text{grad}(f; h) = 0$, análogamente al argumento anterior, concluimos que $f = k$. ■

Corolario 34 . $x^2 + a$ es irreducible en $\mathbb{R}[x]$ si $a \in \mathbb{R}^+$.

Demuestra $\mathbf{r}.$ Supóngase que $x^2 + a = f \cdot g$ con $\text{grad}(f) = \text{grad}(g) = 1$ (supongamos que $f = x + r$). Entonces $x + r \mid x^2 + a \Rightarrow \text{Ev}_{(-r)}(x^2 + a) = 0 \Rightarrow r^2 + a = 0$ (la suma de dos reales positivos es un real positivo). ■

Corolario 35 . $x^2 + 1$ es irreducible en $\mathbb{R}[x]$.



$$x^2 + 1$$

Definición 166 . En $\mathbb{R}[x]$ definimos la relación de “congruencia módulo $x^2 + 1$ ” por:

$$f \stackrel{x^2+1}{\equiv} g \text{ si } (x^2 + 1) \mid (f - g)$$

$$(\Leftrightarrow (f - g) \in \mathbb{R}[x](x^2 + 1)).$$

Proposición 41 . La relación $\stackrel{x^2+1}{\equiv}$ es de equivalencia en $\mathbb{R}[x]$.

Demostración. Reflexividad)

$$f \stackrel{x^2+1}{\equiv} f, \text{ ya que } f - f = 0(x) \in \mathbb{R}[x](x^2 + 1).$$

Simetría)

$$f \stackrel{x^2+1}{\equiv} g \Rightarrow (f - g) \in \mathbb{R}[x](x^2 + 1) \Rightarrow -(f - g) \in \mathbb{R}[x](x^2 + 1).$$

$$-(f - g) \in \mathbb{R}[x](x^2 + 1) \Rightarrow (g - f) \in \mathbb{R}[x](x^2 + 1) \Rightarrow g \stackrel{x^2+1}{\equiv} f.$$

Transitividad) $f \stackrel{x^2+1}{\equiv} g \stackrel{x^2+1}{\equiv} h \Rightarrow (f - g) \in \mathbb{R}[x](x^2 + 1) \wedge (g - h) \in \mathbb{R}[x](x^2 + 1) \Rightarrow$

$$(f - g) + (g - h) \in \mathbb{R}[x](x^2 + 1) \Rightarrow (f - h) \in \mathbb{R}[x](x^2 + 1) \Rightarrow f \stackrel{x^2+1}{\equiv} h.$$

■

Definición 167

$$\mathbb{C} = \mathbb{R}[x] / \stackrel{x^2+1}{\equiv} = \left\{ \overline{f(x)} \mid f(x) \in \mathbb{R}[x] \right\}.$$

Donde $\overline{f(x)}$ denota la clase de congruencia de $f(x)$.

Definición 168 . Dotamos a \mathbb{C} de suma y de producto mediante las definiciones siguientes

$$\begin{aligned} \tilde{+} : \quad \mathbb{C} \times \mathbb{C} &\rightarrow \mathbb{C} \\ (\bar{f}, \bar{g}) &\longmapsto \overline{f+g} \\ \tilde{\cdot} : \quad \mathbb{C} \times \mathbb{C} &\rightarrow \mathbb{C} \\ (\bar{f}, \bar{g}) &\longmapsto \overline{f \cdot g} \end{aligned}$$

Ejercicio 444 . Demostrar que las operaciones recién definidas están bien definidas, es decir, no dependen de la elección de los representantes en las clases de congruencia.

Proposición 42 . $(\mathbb{C}, \tilde{+}, \bar{0}, \tilde{\cdot}, \bar{1})$ es un anillo commutativo (el anillo de los complejos).

- Demostración.**
1. $\bar{f} \tilde{+} \bar{0} = \overline{f+0} = \overline{f} = \bar{0} \tilde{+} \bar{f}, \forall f \in \mathbb{C}$.
 2. $\bar{f} \tilde{+} (\bar{g} \tilde{+} \bar{h}) = \overline{f \tilde{+} (g \tilde{+} h)} = \overline{f + (g + h)} = \overline{(f + g) + h} =$
 $= (\overline{f + g}) \tilde{+} \bar{h} = (\bar{f} \tilde{+} \bar{g}) \tilde{+} \bar{h}, \forall f, g, h \in \mathbb{C}$.
 3. $\bar{f} \tilde{+} \bar{g} = \overline{f+g} = \overline{g+f} = \bar{g} \tilde{+} \bar{f}, \forall f, g \in \mathbb{C}$.
 4. $\bar{f} \tilde{+} \overline{-f} = \overline{f+(-f)} = \bar{0}$, por lo tanto $\overline{-f} = -\bar{f}, \forall f \in \mathbb{C}$.
 5. $\bar{f} \tilde{\cdot} (\bar{g} \cdot \bar{h}) = \overline{f \cdot (g \cdot h)} = \overline{f \cdot (g \cdot h)} =$
 $= \overline{(f \cdot g) \cdot h} = (\overline{f \cdot g}) \tilde{\cdot} \bar{h} = (\bar{f} \tilde{\cdot} \bar{g}) \tilde{\cdot} \bar{h}, \forall f, g, h \in \mathbb{C}$.
 6. $\bar{f} \tilde{\cdot} \bar{1} = \overline{f \cdot 1} = \overline{f} = \bar{1} \tilde{\cdot} \bar{f}, \forall f \in \mathbb{C}$.
 7. $\bar{f} \tilde{\cdot} \bar{g} = \overline{f \cdot g} = \overline{g \cdot f} = \bar{g} \tilde{\cdot} \bar{f}, \forall f, g \in \mathbb{C}$.
 8. $\bar{f} \tilde{\cdot} (\bar{g} + \bar{h}) = \overline{f \cdot (g + h)} = \overline{f \cdot g + f \cdot h} =$
 $= \overline{f \cdot g} \tilde{+} \overline{f \cdot h} = \bar{f} \tilde{\cdot} \bar{g} \tilde{+} \bar{f} \tilde{\cdot} \bar{h}, \forall f, g, h \in \mathbb{C}$.

■

Proposición 43 . La función $\Psi : \mathbb{R} \rightarrow \mathbb{C}$, definida por $\Psi(r) = \bar{r}$, respeta la suma, el producto, el uno y es inyectiva.

Demostración. Sean $r, s \in \mathbb{R}$, entonces:

$$1. \Psi(r+s) = \overline{r+s} = \bar{r} + \bar{s} = \Psi(r) + \Psi(s).$$

$$2. \Psi(r \cdot s) = \overline{r \cdot s} = \bar{r} \bar{s} = \Psi(r) \bar{\Psi}(s).$$

$$3. \Psi(1_{\mathbb{R}}) = \overline{1_{\mathbb{R}}} = 1_{\mathbb{C}}.$$

$$4. \Psi(r) = \Psi(s) \Rightarrow \bar{r} = \bar{s} \Rightarrow r \stackrel{x^2+1}{\equiv} s \Rightarrow (x^2 + 1) \mid (r - s).$$

Pero como $r - s$ es 0 o su grado es cero, entonces $r - s = 0$. Por lo tanto $r = s$. ■

Observación 147 . En \mathbb{C} , $\bar{x}^2 = \overline{-1}$.

$$\text{Demostración. } x^2 + 1 \stackrel{x^2+1}{\equiv} 0 \Rightarrow \bar{0} = \overline{x^2 + 1} = \bar{1} + \overline{x^2} = \bar{1} + \bar{x}^2 \Rightarrow \bar{x}^2 = -\bar{1}$$

■

Proposición 44 . Denotando \bar{x} con i , tenemos que $\forall f \in \mathbb{C}$, $f = \bar{a} + \bar{b}i$, con $a, b \in \mathbb{R}$.

Demostración. Aplicando el algoritmo de la división a f y a $x^2 + 1$:

$$\begin{array}{c} q(x) \\ \hline x^2 + 1 \mid \overline{f(x)} \\ r(x) \quad 0 = r(x) \text{ ó } \text{grad}(r(x)) < \text{grad}(x^2 + 1) = 2 \end{array}$$

como $0 = r(x)$ o bien $\text{grad}(r(x)) \leq 1$, en cualquier caso podemos escribir $r(x) = a + bx$, con $a, b \in \mathbb{R}$. Entonces

$$\bar{f} = \bar{q} \cdot \overline{(x^2 + 1)} + \bar{a} + b\bar{x}$$

ahora, notando que $\overline{(x^2 + 1)} = \bar{0}$, que se puede identificar r con \bar{r} para $r \in \mathbb{R}$, podemos escribir

$$\bar{f} = \bar{a} + b\bar{x} = a + bi.$$

■

Teorema 176 . \mathbb{C} es un campo.

Demostración. Sea $\bar{f} \in \mathbb{C} \setminus \{\bar{0}\}$, entonces $x^2 + 1 \nmid f \Rightarrow (x^2 + 1; f) = 1 \Rightarrow 1 = \alpha(x)(x^2 + 1) + \beta(x)f \Rightarrow \overline{\beta(x)f} = \bar{1} \Rightarrow \bar{\beta} = \overline{f^{-1}}$. ■

Apéndice A

Una teoría axiomática para \mathbb{R}

El resultado final de la sección II y de sus proyectos, muestra que \mathbb{R} , con las operaciones de suma y producto que hemos definido y su clase positiva \mathbb{R}^+ , es un campo arquimediano ordenado en el que todo subconjunto no vacío acotado por arriba, tiene supremo, y extiende a \mathbb{Q} .

Estos son las propiedades que caracterizan -categóricamente- a \mathbb{R} y que sirven de base para construir los axiomas de la teoría cuando se escoge este camino, que evade los problemas de existencia.

En beneficio de los estudiantes que escogen tal manera de proceder (el método axiomático no constructivo para estudiar \mathbb{R}), presentamos una lista de resultados, que pueden considerarse, en ese caso, como los axiomas de la teoría, remarcando que basta suponerlos ciertos, para que, a partir de ellos pueda construirse rigurosamente la mayor parte del Análisis matemático.

A.1 Los axiomas

Sea \mathbb{R} un conjunto, cuyos elementos se llamarán números reales, en el que están definidas dos operaciones binarias $+$ y \cdot , tales que $\forall a, b, c \in \mathbb{R}$ se cumple:

A.1.1 Axiomas, Grupo I.

Axioma 21

$$\begin{aligned}(a + b) + c &= a + (b + c); \\ (ab)c &= a(bc).\end{aligned}$$

“la suma y el producto son operaciones asociativas”

Axioma 22

$$\begin{aligned} a + b &= b + a; \\ ab &= ba \end{aligned}$$

“la suma y el producto son operaciones conmutativas”.

Axioma 23

$$\begin{aligned} \exists 0 &\in \mathbb{R} . \exists . a + 0 = a; \\ \exists 1 &\in \mathbb{R} . 1 \neq 0, . \exists . a \cdot 1 = a \end{aligned}$$

“existencia de neutros para + y * .

Axioma 24

$$\begin{aligned} \forall a &\in \mathbb{R} , \exists b . \exists . a + b = 0; \forall a \in \mathbb{R} , \\ a &\neq 0, \exists b \in \mathbb{R} . \exists . ab = 1 \end{aligned}$$

“existencia de inverso aditivo y multiplicativo respectivamente”.

Axioma 25

$$a \cdot (b + c) = (a \cdot b) + (a \cdot c)$$

“La multiplicación se distribuye sobre la suma o la suma distribuye a la multiplicación o más brevemente: vale la ley distributiva”.

Este primer grupo de axiomas corresponde a la estructura de campo que tiene, y permite justificar plenamente las manipulaciones del álgebra elemental -y por supuesto muchas otras- entre las que seleccionaremos algunos a manera de ejemplo.

Aclaramos que aunque las operaciones + y · son binarias, se pueden generalizar cuando el número de sumandos -o de factores- es diferente de dos, de la manera siguiente:

1. $\sum_{i \in \emptyset} a_i = 0.$ ¹
2. $\sum_{i=1}^1 a_i = a_1.$
3. $\sum_{i=1}^n a_i = \left(\sum_{i=1}^{n-1} a_i \right) + a_n \text{ si } n \geq 2.$
4. $\prod_{i \in \emptyset} a_i = 1.$
5. $\prod_{i=1}^1 a_i = a_1.$
6. $\prod_{i=1}^n a_i = \left(\prod_{i=1}^{n-1} a_i \right) \cdot a_n \text{ si } n \geq 2.$

Ejemplo 217 . Son teoremas $\forall a, b, c \in \mathbb{R}$.

1. $a + b = a + c \Rightarrow b = c.$
“En la suma se vale cancelar” y por lo tanto el inverso aditivo de cada a es único. Se denota “ $-a$ ” y entonces se define:

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

2. Las ecuaciones $a + x = b$ tienen solución única, $x = b - a.$
“En \mathbb{R} vale restar”.
3. En una suma generalizada, el resultado es independiente de la forma en que se agrupen los sumandos. “El cambio en la forma de asociar sumandos, no altera la suma”.

¹Cuando se desea que en una operación generalizada valga la ley asociativa “generalizada” (a lo bestia), la operación vacía debe ser el neutro de la operación. Así la unión vacía debe ser \emptyset y la intersección vacía, el total.

En efecto

$$\begin{aligned} (a_1 \circ a_2 \circ \dots \circ a_n) &= (a_1 \circ a_2 \circ \dots \circ a_n) \circ () = \\ &= (a_1 \circ a_2 \circ \dots \circ a_n) \circ e \end{aligned}$$

E.d.

$$() =: e.$$

4. $ab = 0 \wedge a \neq 0 \Rightarrow b = 0.$

“Cero no tiene divisores propios en \mathbb{R} ”.

$$\therefore ab = ac \wedge a \neq 0 \Rightarrow b = c.$$

“En \mathbb{R} vale cancelar factores diferentes de cero”.

5. Para cada $a \in \mathbb{R}, a \neq 0$, el inverso multiplicativo es único. Se denota a^{-1} y se define

$$ba^{-1} = b/a.$$

6. Si a es distinta de 0, las ecuaciones $ax = b$ tienen solución única ($x = b/a$).

“En \mathbb{R} se vale dividir entre números diferentes de cero”.

7. $a \cdot 0 = 0.$

8. $a(-b) = -ab; (-a)(-b) = ab.$

“Vale la regla de los signos”.

9. $a \cdot \left(\sum_{i=1}^n b_i \right) = \sum_{i=1}^n ab_i.$

$$\left(\sum_{i=1}^n a_i \right) \cdot b = \sum_{i=1}^n a_i b.$$

“vale multiplicar polinomios en la forma usual”.

10. $a/b + c/d = (ad + bc)/bd.$

$$a/b \cdot c/d = ac/bd$$

y si $c \neq 0$, entonces $a/b : c/d = a/b \cdot d/c = ad/bc.$

“Los quebrados se suman, restan, multiplican y dividen, como en la primaria (cuando en la primaria no se equivoca uno)”.

Apéndice B

Las funciones trascendentes

Para redondear el desarrollo del tema de los números reales quisieramos hablar de algunas de las funciones más importantes del Análisis real entre las que se encuentran desde luego la función “logaritmo” (natural), la exponencial y las trigonométricas entre otras, y, como en el caso del conjunto \mathbb{R} , podemos postular su existencia y sus propiedades. Sin embargo, decidimos presentar al lector un pequeño mosaico deductivo basado en lo que llamaremos “un cúmulo de conocimientos previos” que no corresponden propiamente al Álgebra, a partir de los cuales se pueden construir las funciones antes mencionadas. El lector que no esté interesado en este desarrollo puede omitir tranquilamente la sección completa. Hacemos notar que las secciones B.1 y B.2 son dos diferentes maneras de definir las mismas funciones \log y \exp y que una tercera -que aquí omitimos- es la que utiliza las series de potencias.

B.1 “Un cúmulo de conocimientos previos”

Suponemos al lector familiarizado con algunos de los teoremas del Cálculo -que mencionaremos explícitamente- y así daremos por ciertos:

El teorema del valor medio y sus consecuencias elementales:

Teorema 177 (del valor medio). *Si $f : [a, b] \rightarrow \mathbb{R}$ es una función continua en $[a, b]$ y derivable en (a, b) , entonces*

$$\exists c \in (a, b) \text{ tal que } \frac{f(b) - f(a)}{b - a} = f'(c).$$

Observación 148 . $\forall x \in (a, b)$:

1. $f'(x) = 0$, entonces f es constante en $[a, b]$.
2. $f'(x) > 0$, entonces f es estrictamente creciente en $[a, b]$.
3. $f'(x) < 0$, entonces f es estrictamente decreciente en $[a, b]$.
4. $f'(x) = g'(x)$ entonces $\exists c \in \mathbb{R}$ tal que

$$f(x) = g(x) + c.$$

Teorema 178 . Si $f : [a, b] \rightarrow \mathbb{R}$ es continua, entonces es integrable en cualquier subintervalo $I \subset [a, b]$.

Y si, como es usual, se define:

$$\left(\int_a^a f \right) = 0, \quad \left(\int_b^a f \right) = - \left(\int_a^b f \right),$$

entonces $\forall a, b, c \in \mathbb{R}$, si existen dos de las tres integrales siguientes, entonces existe la tercera y además

$$\int_a^b f = \int_a^c f + \int_c^b f$$

independientemente del orden en el que aparezcan a, b y c sobre la recta real.

Teorema 179 . El primer Teorema fundamental del Cálculo:

Si $f : [a, b] \rightarrow \mathbb{R}$ es continua, y se define $F : [a, b] \rightarrow \mathbb{R}$ como $F(x) = \int_a^x f$, entonces F es derivable y $\forall x \in [a, b], F'(x) = f(x)$.

Teorema 180 . Regla de la cadena: Si f y g son funciones derivables tales que la composición $g \circ f$ está definida, entonces esta composición es derivable, y su derivada es el producto de las derivadas de g y de f

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

Definición 169 . Definimos la función $L : \mathbb{R}^+ \rightarrow \mathbb{R}$ como

$$L(x) = \int_1^x \frac{dt}{t}.$$

Que de acuerdo al teorema 178, está bien definida, ya que $\forall t \in \mathbb{R}^+, 1/t$ es continua y por lo tanto integrable, y notamos que de acuerdo a las definiciones previas, $L(x_0)$ puede interpretarse geométricamente como “el área bajo la curva” de la hipérbola $y = 1/x$ desde $x = 1$ hasta $x = x_0$ si $x_0 > 1$ y como el inverso de tal área si $0 < x_0 < 1$.

Entonces puede verse inmediatamente que L tiene las propiedades siguientes:

1. $L(1) = 0$.
2. $L(x) = 1/x, \forall x \in \mathbb{R}^+$.

Repetimos que cuando no se desea recurrir al multicitado cúmulo de conocimientos previos, la existencia de la función

$$L : \mathbb{R}^+ \rightarrow \mathbb{R}$$

junto con sus propiedades 1. y 2., puede agregarse como un axioma más para \mathbb{R} . También podríamos haber supuesto que “el área bajo la curva” de la hipérbola $y = 1/x$ es mensurable (puede medirse). Definir entonces $L(x_0)$ como la medida de tal área -o el inverso de la medida si $0 < x_0 < 1$ - y entonces deducir las propiedades 1. y 2. a partir de las consideraciones geométricas obvias.

Teorema 181 . *Sea $a \in \mathbb{R}^+$ y defínase: $h : \mathbb{R}^+ \rightarrow \mathbb{R}$ como sigue: $h(x) = L(ax)$. Entonces*

$$L(ax) = \frac{1}{ax}a = \frac{1}{x} = L(x).$$

Por lo tanto $L(ax)$ y $L(x)$ difieren en una constante. Es decir:

$$\exists c \in \mathbb{R} \ \exists \forall x \in \mathbb{R}^+, L(ax) = c + L(x).$$

Para calcular el valor de c , tomamos $x = 1$ y entonces

$$L(a1) = L(a) = c + L(1) = c.$$

En resumen:

3. $\forall a, b \in \mathbb{R}^+, L(ab) = L(a) + L(b)$, de donde, por inducción, resulta que

$$\forall n \in \mathbb{Z}^+, L(a^n) = nL(a).$$

Nótese que $a = (a/b)b \therefore L(a) = L((a/b)b) = L(a/b) + L(b)$ Luego

$$L(a/b) = L(a) - L(b),$$

de donde se sigue que

$$L(a^{-n}) = L(1/a^n) = L(1) - L(a^n) = 0 - nL(a)$$

y por lo tanto

$$L(a^n) = nL(a) \forall n \in \mathbb{Z}$$

(Se verá después que el teorema vale $\forall n \in \mathbb{R}$).

Teorema 182 . *En vista de que $L(x) = 1/x$, $L(x)$ es positiva para todo x en su dominio (\mathbb{R}^+) y por lo tanto $L(x)$ crece estrictamente. Ahora bien si $x \rightarrow \infty$, la gráfica de $L(x)$ se “acuesta” (su pendiente disminuye) a medida que se va alejando del eje Y , y aumenta muy rápidamente cuando x se acerca a cero, a pesar de lo cual, como probaremos, $L(\mathbb{R}^+)$ no está acotado ni por arriba ni por abajo, es decir: para cada $M \in \mathbb{R}^+$, existen*

$$x_0, y_0 \in \mathbb{R}^+ \ . \ L(x_0) > M, L(y_0) < -M$$

o sea que L toma valores arbitrariamente altos o bajos y de aquí, dada la continuidad de L , puede concluirse inmediatamente que

$$L(\mathbb{R}^+) = \mathbb{R}.$$

Teorema 183 . *$L(\mathbb{R}^+)$ no está acotado.*

Demostración. Sea $M \in \mathbb{R}$.

En vista de que $L(1) = 0$, y de que L crece estrictamente, entonces $L(2) > 0$ y por lo tanto -propiedad arquimediana-

$$\exists n_0 \in \mathbb{N} \ . \ . \ . \ M < n_0 \cdot L(2) = L(2^{n_0}); \ x_0 = 2^{n_0}$$

y

$$-M > -n_0 \cdot L(2) = L(2^{-n_0}); \ y_0 = 1/2^{n_0}.$$

■

Ahora bien, una función estrictamente creciente es inyectiva y por lo tanto, $L : \mathbb{R}^+ \rightarrow \mathbb{R}$, es biyectiva. Luego existe

$$L^{-1} : \mathbb{R} \rightarrow \mathbb{R}^+.$$

Si se bautiza

$$L^{-1} = E,$$

habremos demostrando la existencia de una función

$$E : \mathbb{R} \rightarrow \mathbb{R}^+$$

tal que

$$\forall x \in \mathbb{R}, L(E(x)) = x$$

y

$$\forall y \in \mathbb{R}^+, E(L(y)) = y.$$

Entonces, de las propiedades de L se obtienen las siguientes que corresponden a E :

1. $L(1) = 0 \Rightarrow 1 = E(0)$
2. $L(x) = \frac{1}{x} \Rightarrow E(x) = E(x).$
(Sea $y = E(x) \therefore L(y) = x, L(y) = \frac{1}{y} \Rightarrow 1 \therefore y = y$).
3. Sean

$$x, y \in \mathbb{R}, a = E(x), b = E(y)$$

Entonces

$$L(a) = x, L(b) = y$$

$$\therefore L(a) + L(b) = L(ab) = x + y$$

$$(\therefore E(L(ab)) = E(x + y))$$

O sea

$$E(x + y) = ab = E(x)E(y).$$

Por inducción:

$$E \left(\sum_{i=1}^n a_i \right) = \prod_{i=1}^n E(a_i), \forall n \in \mathbb{N}.$$

Si se define $e = E(1)$, resulta que $L(e) = 1$ lo que justifica -según se verá más adelante- la denominación de L como la función “logaritmo base e ” o “logaritmo natural”.

En el capítulo correspondiente a los números naturales, se definió lo que debe entenderse por a^n con a y n números naturales, y se vió que como consecuencia de la definición (recursiva) de estas expresiones, a^n se puede interpretar como el producto generalizado de n factores iguales a a . Consideración que permite extender la definición anterior al caso en que $a \in \mathbb{R}$. Pero ¿qué significan expresiones como 2^e ó $(n)^{\sqrt{2}}$?

Definamos:

Definición 170 . Sean $a, b \in \mathbb{R}$, $a > 0$, entonces

$$a^b = E(bL(a)),$$

$$\therefore L(a^b) = bL(a), \forall b \in \mathbb{R}.$$

Como consecuencia de esta definición, se tienen los siguientes resultados:

1. La definición 170 extiende a la que ya se tenía. En efecto si $a \in \mathbb{R}^+$ y $n \in \mathbb{Z}^+$, entonces

$$\begin{aligned} a^n &= E(nL(a)) = E(\underbrace{L(a) + \dots + L(a)}_{n \text{ sumandos}}) = \\ &= E(\underbrace{L(a)L(a)\dots L(a)}_{n \text{ factores}}) = \\ &= \underbrace{a \dots a}_{n \text{ factores}} \end{aligned}$$

2. “Valen las leyes de los exponentes”

En efecto, si $a \in \mathbb{R}^+$, $b, c \in \mathbb{R}$, entonces

$$\begin{aligned} (a) \quad a^b \cdot a^c &= E(bL(a)) \cdot E(cL(a)) = \\ &= E(bL(a) + cL(a)) = E((b+c)L(a)) = a^{b+c}. \end{aligned}$$

$$(b) \quad (a^b)^c = E(c(L(a^b))) = E(c(bL(a))) = E(bcL(a)) = a^{bc}$$

$$(c) \quad \forall x \in \mathbb{R}, e^x = E(xL(e)) = E(x).$$

Este último resultado justifica el extendido uso de la notación e^x para la función exponencial (recuerde que las funciones trigonométricas se representan como $\sin(x)$, $\cos(x)$, . . . lo que en rigor introduce una ambigüedad. Así por ejemplo $\sin(x)$ puede ser tanto la notación para la función “seno” como “el valor que la función seno asigna al número x ”, pero la extendida práctica -usos y costumbres- ha mostrado que el contexto en el que aparecen tales expresiones permite precisar la interpretación que deba tomarse)

B.2 Hipótesis. (Mosaico 1)

Supondremos también el siguiente resultado -que se demuestra en los cursos de ecuaciones diferenciales-:

Teorema 184 (*Teorema de existencia y unicidad*). *Dada una ecuación diferencial*

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_0 y^{(0)} = 0, \quad a_i \in \mathbb{R}, \quad i = 0, \dots, n, \quad (\text{b.1})$$

para cualquier colección de $n + 1$ números reales

$$x_0, A_0, A_1, \dots, A_{n-1},$$

$$\exists! f : \mathbb{R} \rightarrow \mathbb{R},$$

n-diferenciable tal que

.) $f^{(i)}(x_0) = A_i, \quad i = 0, 1, \dots, n - 1.$

..) $\forall x \in \mathbb{R}, \quad a_n f^{(n)}(x) + \dots + a_0 f(x) = 0.$

“Dadas condiciones iniciales

$$(x_0, \{A_i\}), \quad i \in \{0, \dots, n - 1\}, \quad \text{existe una única } f : \mathbb{R} \rightarrow \mathbb{R}$$

solución de b.1 que las satisface”.

Además si

$$S_0 = \{f : \mathbb{R} \rightarrow \mathbb{R} \mid f \text{ es solución de b.1}\},$$

entonces S_0 es un subespacio del espacio de las funciones n -diferenciables, de dimensión n y por lo tanto si

$$\beta = \{y_1, \dots, y_n\}$$

es un conjunto linealmente independiente de soluciones de b.1, toda $g \in S_0$ es combinación lineal de β .

Nos referiremos solamente a las dos ecuaciones particulares siguientes:

1.

$$y' - y = 0. \quad (b.2)$$

2.

$$y' + y = 0. \quad (b.3)$$

Primer caso:

$$y' - y = 0.$$

Una observación importante:

Observación 149 . *El teorema de existencia y unicidad cuya validez estamos suponiendo, garantiza que por cada punto del plano (x_0, y_0) , pasa una única curva integral que corresponde a la gráfica de la solución $f : \mathbb{R} \rightarrow \mathbb{R}$ tal que $f(x_0) = y_0$, lo que equivale a garantizar que la unión de todas estas curvas integrales llena completamente el plano y que cualesquiera, dos de ellas son ajenas. (Si f_1 y f_2 se cortarán en algún punto $(x_0, y_0) \in \mathbb{R}^2$, entonces ambas satisfarían el mismo problema de valores iniciales, luego no serían dos, e. d. $f_1 = f_2$).*

B.3 La función exponencial (2a. versión)

Consideremos, para la primera ecuación b.2, la (única) solución $E : \mathbb{R} \rightarrow \mathbb{R}$ tal que $E(0) = 1$. Entonces:

$$E(x) = E(x), \quad E(0) = 1.$$

Observando que la gráfica de la función constante cero corresponde al eje X , y que ésta también es solución de b.2, concluimos que $E(x)$ no puede ser cero para número real alguno, ya que dos soluciones distintas tienen gráficas ajenas -como ya se dijo- y por lo tanto,

$$\forall x \in \mathbb{R}, \quad E(x) > 0$$

(si para algún x_0 , $E(x_0)$ fuera negativa, entonces el teorema del valor intermedio del Cálculo asegura que E , necesariamente cortaría al eje X , lo que no puede ocurrir, por lo que:

Observación 150 . *E es estrictamente creciente, es de clase C^∞ y*

$$\forall k \in \mathbb{Z}^+, E^{(k)} = E.$$

Demostración. En efecto, $E(x) > 0$ y $E'(x) = E(x)$ demuestran la primera parte de la observación -E crece estrictamente- y la hipótesis induc-tiva

$$E^{(k)} = E$$

dice que siendo E derivable $E^{(k)}$ también lo es y por lo tanto $E^{(k+1)} = E' = E$

■

Observación 151

$$S_0 = \{f : \mathbb{R} \rightarrow \mathbb{R} \mid f \text{ es solución de } b.2\}$$

es un subespacio de las funciones derivables de \mathbb{R} en \mathbb{R} y su dimensión es 1.

Demostración. 1) La función 0 está en S_0 , obviamente.

2) Sean f y $g : \mathbb{R} \rightarrow \mathbb{R}$ tales que

$$f = f'$$

$$g = g'$$

Entonces, sumando miembro a miembro se obtiene:

$$f + g = f' + g' = (f + g)',$$

luego

$$f, g \in S_0 \Rightarrow (f + g) \in S_0.$$

3) Si $f = f'$ y $c \in \mathbb{R}$ entonces

$$cf = cf' = (cf)'$$

por lo tanto

$$f \in S_0 \text{ y } c \in \mathbb{R} \Rightarrow cf \in S_0.$$

Conclusión: S_0 es un subespacio vectorial.

2). $\beta = \{E\}$ es una base de S_0 .

.) $E \neq 0 \Rightarrow \{E\}$ es linealmente independiente.

..) Si $f : \mathbb{R} \rightarrow \mathbb{R}$ está en S_0 y $f(0) = k$, entonces $kE \in S_0$ satisface la misma condición inicial

$$(kE)(0) = k(E(0)) = k$$

y por lo tanto $f = kE$. Luego E genera S_0 , lo que demuestra que

$$\dim S_0 = 1.$$

■

(Versión alternativa)

Teorema 185 . *Toda función $f : \mathbb{R} \rightarrow \mathbb{R}$ que sea solución de b.2, es un múltiplo de E .*

Demuestra Recuerde que $E(x) \neq 0$, $\forall x \in \mathbb{R}$, $\therefore h(x) = \frac{f(x)}{E(x)}$ está bien definida en \mathbb{R} , es derivable (el cociente de funciones derivables lo es), y $h'(x) = \frac{E(x)f'(x) - E'(x)f(x)}{(E(x))^2} = 0$, $\therefore h(x) = k$ (constante), es decir, $f(x) = kE(x)$. ■

Teorema 186 . $\forall x, y \in \mathbb{R}$, $E(x+y) = E(x)E(y)$.

Demuestra Sea $x \in \mathbb{R}$. Defínase

$$h(y) = E(x+y).$$

Entonces $h' = h \therefore h \in S_0$ y por lo tanto existe una constante $k \in \mathbb{R}$

$$E(x+y) = kE(y),$$

haciendo $x = 0$ obtenemos: $E(x) = k$ e. d.

$$E(x+y) = E(x)E(y).$$

■

Corolario 36 . $\forall n \in \mathbb{Z}^+, E(na) = E(a + \dots + a) = E(a) \cdot \dots \cdot E(a) = (E(a))^n$.

Se probará -después de definir a^b - que

$$\forall a > 0, E(ra) = (E(a))^r \forall r \in \mathbb{R}.$$

Corolario 37 . $E : \mathbb{R} \rightarrow \mathbb{R}^+$ es biyectiva y por lo tanto tiene inversa

$$L : \mathbb{R}^+ \rightarrow \mathbb{R}$$

y entonces:

$$\forall x \in \mathbb{R}, L(E(x)) = x$$

$$y \in \mathbb{R}^+, E(L(y)) = y$$

Definición 171 . Se llama e al número real $E(1)$.

Entonces:

1. $L(e) = L(E(1)) = 1$ y $L(1) = L(E(0)) = 0$
2. Si $y = L(x)$, entonces $E(y) = x$ y derivando:

$$E'(y) = E(y)y' = 1; y' = 1/E(y)$$

o sea:

$$L'(x) = 1/x.$$

3. $L(ab) = L(a) + L(b)$

En efecto:

Sea $x = L(a)$, $y = L(b)$ Entonces $E(x) = a$, $E(y) = b$

$$\therefore ab = E(x)E(y) = E(x + y)$$

y tomando logaritmos,

$$L(ab) = L(E(x + y)) = x + y = L(a) + L(b).$$

Recuérdese que en su momento (ver N) se definió a^n para $a \neq 0$ inductivamente:

Definición 172 . $a^0 = 1$, $a^{n+1} = a^n \cdot a$.

Se vió que para estos casos, si $a \neq 0$, entonces

$$a^n = \underbrace{a \cdot \dots \cdot a}_{n \text{ factores}}$$

y que $\forall n, m \in \mathbb{N}$,

$$a^n a^m = a^{n+m}, (a^n)^m = a^{nm}.$$

Definición 173. Sean, $a, b \in \mathbb{R}$, $a > 0$, defínase $a^b = E(b \cdot L(a))$.

Se tienen ahora los teoremas siguientes:

Teorema 187. La Definición 173 extiende a la Definición 172 para $a > 0$.

Demostración. En efecto, si $b = n$ y $a \neq 0$,

$$\begin{aligned} a^n &= E(nL(a)) = E(\underbrace{L(a) + \dots + L(a)}_{n \text{ sumandos}}) = \\ &= \underbrace{E(L(a))E(L(a))\dots E(L(a))}_{n \text{ factores}} = \underbrace{a \cdot a \cdot \dots \cdot a}_{n \text{ factores}}. \end{aligned}$$

■

Teorema 188. $L(a^b) = L(E(bL(a))) = bL(a)$.

Ejercicio 445. Demuestre el teorema anterior.

Teorema 189. $a^0 = E(0L(a)) = E(0) = 1$.

Ejercicio 446. Demuestre el teorema anterior.

Teorema 190. (Leyes de los exponentes). Sea $a > 0$. Entonces $\forall b, c \in \mathbb{R}$,

$$1. a^b a^c = a^{b+c}.$$

$$2. (a^b)^c = a^{bc}.$$

Demostración.

$$\begin{aligned} a^{b+c} &= E((b+c)L(a)) = E(bL(a) + cL(a)) = \\ &= E(bL(a))E(cL(a)) = a^b a^c. \end{aligned}$$

y

$$(a^b)^c = E(cL(a^b)) = E(cbL(a)) = E(bcL(a)) = a^{bc}$$

■

Teorema 191. $e^x = E(x)$ pues $e^x = E(xL(e)) = E(x)$.

Este último resultado justifica el uso de la notación e^x como sinónimo de $E(x)$.

B.4 Funciones trigonométricas

Tomando ahora la ecuación $y'' + y = 0$, sabemos por el teorema 184 que:

para cada pareja (k_1, k_2) de números reales, existe una única función $f : \mathbb{R} \rightarrow \mathbb{R}$ que es solución de la ecuación diferencial

$$y' + y = 0, \quad (b.3)$$

y que es tal que

$$f(0) = k_1; f'(0) = k_2.$$

Tomemos ahora los siguientes acuerdos:

1. $S_0 = \{f : \mathbb{R} \rightarrow \mathbb{R} \mid f \text{ es solución de b.3}\}.$
2. $\mathbf{c} : \mathbb{R} \rightarrow \mathbb{R}$ es la (única) solución de (b.3) que satisface: $\mathbf{c}(0) = 1; \mathbf{c}'(0) = 0.$
3. $\mathbf{s} : \mathbb{R} \rightarrow \mathbb{R}$: es la (única) solución de (b.3) tal que $s(0) = 0; s'(0) = 1.$

Es importante observar que si $y \in S_0$, entonces $y' = -y$ y por tanto la función $-y$ es dos veces derivable y como cada vez que se derive $-y$ -que es igual a y' - se obtiene una derivada dos órdenes mayor para ella, resulta que todo elemento de S_0 (solución de b.3) es de clase C^∞ . Es decir, tiene derivadas continuas de todos los órdenes. En particular, \mathbf{c} y \mathbf{s} son de clase C^∞ .

Si hacemos $h(x) = s'(x)$, se obtiene, derivando, que:

$$h(x) = s''(x) = -s(x)$$

$$h''(x) = -s'(x) = -h(x)$$

es decir que $h(x) = s'(x) \in S_0$ y como

$$h(0) = s'(0) = 1, \quad h'(0) = s''(0) = -s(0) = 0,$$

entonces h satisface las mismas condiciones iniciales que la función \mathbf{c} . El teorema de unicidad de soluciones fuerza a concluir que:

$$s'(x) = \mathbf{c}(x).$$

Análogamente se obtiene:

$$\mathbf{c}'(x) = -\mathbf{s}(x).$$

De aquí resulta que si $g(x) = \mathbf{c}^2 + \mathbf{s}^2$, entonces g es una función C^∞ de \mathbb{R} en \mathbb{R} y como

$$g'(x) = 2\mathbf{c}(x)(-\mathbf{s}(x)) + 2\mathbf{s}(x)\mathbf{c}(x) = 0,$$

g es una función constante, que se puede evaluar calculándola cuando $x = 0$. En este caso:

$$g(0) = \mathbf{c}^2(0) + \mathbf{s}^2(0) = 1.$$

de lo que se concluye que:

$$\forall x \in \mathbb{R}, \mathbf{c}^2(x) + \mathbf{s}^2(x) = 1.$$

Obsérvese que si $h(x) = \mathbf{c}(-x)$, entonces:

$$h'(x) = -c'(-x); \quad h''(-x) = c''(-x) = -c(-x) = -h(x).$$

y, por lo tanto, $h(x) \in S_0$. Como $h(0) = 1$, y $h'(0) = 0$, entonces $h = c$, e. d.

$$\forall x \in \mathbb{R}, \mathbf{c}(x) = \mathbf{c}(-x)$$

En otras palabras, \mathbf{c} es una función “par”.

Análogamente se prueba

$$\mathbf{s}(-x) = -\mathbf{s}(x) \tag{B.1}$$

de lo que se concluye que \mathbf{s} es una función “ímpar”.

Demostremos ahora el siguiente teorema:

Teorema 192 . *S_0 es un espacio vectorial (subespacio de C^∞) de dimensión 2 y $\beta = \{c, \mathbf{s}\}$ es una de sus bases.*

Demostración. i) La función constante $\hat{0} : \mathbb{R} \rightarrow \mathbb{R}$: están en S_0 que, por lo tanto, no es vacío.

ii) Si $f, g \in S_0$ entonces $f + g \in S_0$:

En efecto, $f, g \in S_0$ implica que

$$f'' + f = 0$$

y

$$g'' + g = 0,$$

por lo que, sumando miembro a miembro,

$$(f + g)'' + (f + g) = 0,$$

lo que garantiza que $f + g \in S_0$.

iii) Si $\alpha \in \mathbb{R}$ y $f \in S_0$, de $f'' + f = 0$, se obtiene multiplicando por α ,

$$\alpha(f'' + f) = \alpha 0 = 0,$$

o sea: $(\alpha f)' + \alpha f = 0$, lo que dice que también $\alpha f \in S_0$ con lo que termina la demostración de que S_0 , es un subespacio de C^∞ .

Para probar que \mathbf{c}, s es una base de S_0 -que por lo tanto resultará de dimensión 2-, se debe demostrar que $\{\mathbf{c}, \mathbf{s}\}$ es linealmente independiente y que todo elemento de S_0 es una combinación lineal de \mathbf{c}, s .

En efecto, si $\alpha c(x) + \beta \mathbf{s}(x) = 0(x)$, derivando se obtiene:

$$\alpha c'(x) + \beta \mathbf{s}'(x) = 0(x)$$

por lo que, tomando para x el valor 0 en cada caso, resulta:

$$\alpha c(0) + \beta \mathbf{s}(0) = 0(0)$$

e. d..

$$\alpha \cdot 1 + \beta \cdot 0 = 0 \therefore \alpha = 0$$

y

$$\alpha c'(0) + \beta \mathbf{s}'(0) = 0(0)$$

e. d.

$$\alpha \cdot 0 + \beta \cdot 1 = 0 \therefore \beta = 0$$



Luego $\{\mathbf{c}, s\}$ es un conjunto linealmente independiente.

Sea ahora $f \in S_0$ y llámese $a = f(0), b = f'(0)$.

Si ahora definimos $u : \mathbb{R} \rightarrow \mathbb{R}$, por

$$u(x) = a \mathbf{c}(x) + b \mathbf{s}(x),$$

es fácil ver que $u \in S_0$, ya que es combinación lineal de elementos de S_0 que es un espacio vectorial. Por supuesto que se llega a la misma conclusión calculando u', u'' y notando que $u' + u = 0$,

Además $u(0) = a$, $u'(0) = b$ o sea que u satisface idénticas condiciones que f . Por lo tanto, apelando nuevamente al teorema de existencia y unicidad, concluimos que $u = f$ o sea que:

$$f(x) = a \mathbf{c}(x) + b \mathbf{s}(x),$$

con lo que termina la demostración del teorema que asegura, entre otras cosas, que toda solución de (b.3)(elemento de S_0) es una combinación lineal de $\{\mathbf{c}, \mathbf{s}\}$.

(Esto justifica la costumbre de describir al conjunto solución de (b.3) S_0 como:

$$\{y(x) = k_1 \mathbf{c}(x) + k_2 \mathbf{s}(x) \mid k_1, k_2 \in \mathbb{R}\}.$$

Es importante considerar ahora la función $\Phi_a : \mathbb{R} \rightarrow \mathbb{R}$, definida para cada a como:

$$\Phi_a(x) = \mathbf{c}(a + x)$$

que, como se comprueba directamente, es un elemento de S_0 y es por lo tanto, combinación lineal de $\{\mathbf{c}, \mathbf{s}\}$, es decir que existen constantes k_1 y k_2 tales que $\forall x \in \mathbb{R}$,

$$\mathbf{c}(a + x) = k_1 \mathbf{c}(x) + k_2 \mathbf{s}(x).$$

Derivando,

$$-\mathbf{s}(a + x) = -k_1 \mathbf{s}(x) + k_2 \mathbf{c}(x).$$

Ahora, si $x = 0$ se obtiene

$$k_1 = \mathbf{c}(a), k_2 = -\mathbf{s}(a)$$

y por esto

$$\forall x \in \mathbb{R}, \mathbf{c}(a + x) = \mathbf{c}(a) \mathbf{c}(x) - \mathbf{s}(a) \mathbf{s}(x),$$

que en particular, si $x = b$, corresponde a la conocida fórmula del coseno de la suma de a y b .

1.

$$\mathbf{c}(a + b) = \mathbf{c}(a) \mathbf{c}(b) - \mathbf{s}(a) \mathbf{s}(b).$$

Si $x = -b$, y recordando que

$$\mathbf{c}(-x) = \mathbf{c}(x), \mathbf{s}(-x) = -\mathbf{s}(x),$$

se concluye que:

2.

$$\mathbf{c}(a-b) = \mathbf{c}(a)\mathbf{c}(-b) - \mathbf{s}(a)\mathbf{s}(-b) = \mathbf{c}(a)\mathbf{c}(b) + \mathbf{s}(a)\mathbf{s}(b).$$

Procediendo de la misma manera, para

$h_a(x) = \mathbf{s}(a+b)$, se llega a:

3.

$$\mathbf{s}(a+b) = \mathbf{s}(a)\mathbf{c}(b) + \mathbf{c}(a)\mathbf{s}(b)$$

y

4.

$$\mathbf{s}(a-b) = \mathbf{s}(a)\mathbf{c}(b) - \mathbf{c}(a)\mathbf{s}(b).$$

Cuando $b = a$, (2) se transforma en :

5.

$$\mathbf{c}(2a) = \mathbf{c}^2(a) - \mathbf{s}^2(a) = \mathbf{c}^2(a) - (1 - \mathbf{c}^2(a)) = 2\mathbf{c}^2(a) - 1$$

y, por lo tanto:

6.

$$\mathbf{c}^2(a) = \frac{1 + \mathbf{c}(2a)}{2}.$$

Si en (5) cambiamos $\mathbf{c}^2(a)$ por $1 - \mathbf{s}^2(a)$, se obtiene:

7.

$$\mathbf{c}(2a) = 1 - 2\mathbf{s}^2(a)$$

y despejando $\mathbf{s}^2(a)$,

8.

$$\mathbf{s}^2(a) = \frac{1 - \mathbf{c}(2a)}{2}.$$

Demostraremos ahora el siguiente resultado, por reducción al absurdo y usando el Teorema del valor medio (Teorema 177).

Teorema 193 . *Existe $a \in \mathbb{R}^+$ tal que $\mathbf{c}(a) \leq 0$.*

Demostración. Supóngase que no. Por lo tanto, dado que $\mathbf{c}(0) = 1$, y \mathbf{c} es de clase C^∞ , debe suceder que $\forall x \in \mathbb{R}^+, \mathbf{c}(x) > 0$ (en caso contrario, se contradiría el teorema del valor intermedio).

Entonces \mathbf{s} , cuya derivada es \mathbf{c} , resulta estrictamente creciente en \mathbb{R}^+ y así, \mathbf{c} decrece estrictamente en todo \mathbb{R}^+ .

Si $h \in \mathbb{R}^+$, entonces

$$\mathbf{c}(0) - \mathbf{c}(h) = \varepsilon,$$

es un número positivo.

Por el teorema del valor medio,

$$\frac{\mathbf{c}(h) - \mathbf{c}(0)}{h} = \mathbf{c}'(\xi), \quad \xi \in (0, h).$$

Como $\mathbf{c}' = -\mathbf{s}$, y como $\mathbf{s}(\xi) > \mathbf{s}(0) = 0$ (por hipótesis, \mathbf{s} es creciente. Recordemos además que \mathbf{s} es impar, B.4), entonces

$$\frac{-\varepsilon}{h} = \frac{\mathbf{c}(h) - \mathbf{c}(0)}{h} = -\mathbf{s}(\xi), \quad \xi \in (0, h).$$

Análogamente,

$$\frac{\mathbf{c}(2h) - \mathbf{c}(h)}{h} = -\mathbf{s}(\varsigma) < \mathbf{s}(\xi) = \frac{-\varepsilon}{h} > 0, \quad \varsigma \in (h, 2h),$$

de donde tenemos que

$$\mathbf{c}(2h) - \mathbf{c}(h) < -\varepsilon$$

es decir que

$$\mathbf{c}(h) - \mathbf{c}(2h) > \varepsilon.$$

Por inducción, obtenemos que

$$\mathbf{c}(nh) - \mathbf{c}((n+1)h) > \varepsilon, \quad \forall n \in \mathbb{N}.$$

Con otra fácil inducción tenemos que

$$\forall n \in \mathbb{N}, \mathbf{c}(0) - \mathbf{c}(nh) > n\varepsilon.$$

Si $n_0 \in \mathbb{N}$ es tal que $n_0\varepsilon > 1$ (n_0 existe ya que el orden de \mathbb{R} es arquimediano), entonces

$$\mathbf{c}(0) - \mathbf{c}(n_0h) > 1$$

o sea que $\mathbf{c}(n_0h) < 0$ (absurdo).

El absurdo ($\mathbf{c}(n_0h) > 0$, $\mathbf{c}(n_0h) < 0$) se obtuvo de suponer que $\forall a \in \mathbb{R}^+, \mathbf{c}(a) > 0$.

Entonces lo que afirma el teorema es cierto. ■

Como resultado, sabemos que si

$$\beta = \{x \in \mathbb{R}^+ \mid \mathbf{c}(x) \leq 0\},$$

entonces β es no vacío y obviamente acotado por abajo. Luego tiene ínfimo, que conviene definir.

Definición 174 . Se llama π al doble del ínfimo de β . Entonces $\inf(\beta) = \pi/2$.

Observación 152 . Como \mathbf{c} es una función continua en $\pi/2$, suponer que $\mathbf{c}(\pi/2)$ es diferente de cero, contradice la definición de ínfimo:

- $\mathbf{c}(\pi/2) > 0 \implies \exists \varepsilon > 0. \exists \forall x \in [\pi/2, \pi/2 + \varepsilon], \mathbf{c}(x) > 0$, luego $\pi/2 + \varepsilon/2$ es cota inferior de β . Obviamente, $\pi/2 + \varepsilon/2 > \pi/2$.
- $\mathbf{c}(\pi/2) < 0 \implies \exists \varepsilon > 0. \exists \forall x \in (\pi/2 - \varepsilon, \pi/2), \mathbf{c}(x) < 0$
entonces $\pi/2 - \varepsilon/2 \in \beta$ y $\pi/2 - \varepsilon/2 < \pi/2$, ∇ .

Luego $\mathbf{c}(\pi/2) = 0$ y entonces $\mathbf{s}(\pi/2) = 1$ (la función \mathbf{s} , cuya derivada es \mathbf{c} , crece estrictamente a partir de 0 en el intervalo $[0, \pi/2]$, y como $\mathbf{s}^2 + \mathbf{c}^2 = 1$, $\mathbf{s}(\pi/2)$ no puede ser negativo).

Usando ahora la fórmula (6) se obtiene:

(13)...

$$\mathbf{c}(\pi) = \mathbf{c}(\pi/2 + \pi/2) = \mathbf{c}^2(\pi/2) - \mathbf{s}^2(\pi/2) = -1 \therefore \mathbf{s}(\pi) = 0.$$

(14)...

$$\mathbf{c}(2\pi) = \mathbf{c}(\pi + \pi) = \mathbf{c}^2\pi - \mathbf{s}^2\pi = 1; \therefore \mathbf{s}(2\pi) = 0$$

y finalmente

(15)...

$$\mathbf{c}(2\pi + x) = \mathbf{c}(2\pi) \mathbf{c}(x) - \mathbf{s}(2\pi) \mathbf{s}(x) = \mathbf{c}(x),$$

(16)...

$$\mathbf{s}(2\pi + x) = \mathbf{s}(2\pi) \mathbf{c}(x) + \mathbf{c}(2\pi) \mathbf{s}(x) = \mathbf{s}(x)$$

Las fórmulas (15) y (16), dicen que tanto \mathbf{c} como \mathbf{s} , son funciones periódicas de período 2π , y nótese que 2π es el menor real positivo para el que esto pasa, ya que si

$$\forall x \in \mathbb{R} \text{ y } a > 0, \mathbf{c}(a + x) = \mathbf{c}(x),$$

para $x = 0$ se obtiene

$$\mathbf{c}(a) = 1 \therefore a \in \{0, \pm 2\pi, \pm 4\pi, \dots\}.$$

Si $\mathbf{s}(a + x) = \mathbf{s}(x)$, entonces

$$\mathbf{s}(a) = \mathbf{s}(0) = 0 \therefore a \in \{\pm\pi, \pm 2\pi, \pm 3\pi, \dots\},$$

y como para $x = \pi$,

$$\mathbf{s}(a + \pi) = \mathbf{s}(a) \mathbf{c}(\pi) + \mathbf{s}(\pi) \mathbf{c}(a) = -\mathbf{s}(a),$$

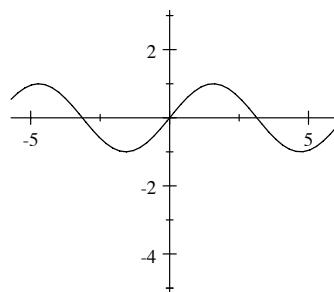
quedan descalificados los múltiplos impares de π como valores posibles para a .

A partir de las funciones \mathbf{s} y \mathbf{c} , se construyen las restantes funciones trigonométricas, definiendo:

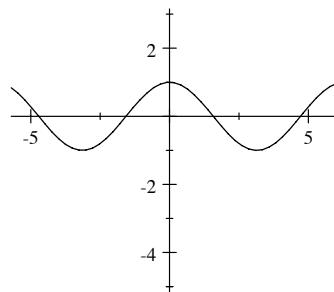
$$\tan = \mathbf{s} / \mathbf{c} \quad \operatorname{ctg} = \mathbf{c} / \mathbf{s}$$

$$\sec = 1 / \mathbf{c} \quad \csc = 1 / \mathbf{s}$$

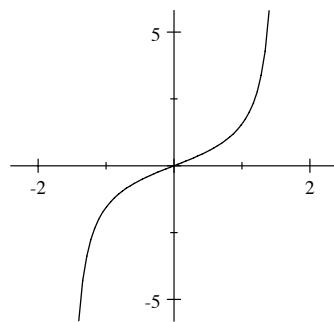
Cuyos dominios constan de todos los números reales que no hacen cero a los denominadores correspondientes, y con la información que se obtiene, tanto de las definiciones anteriores como de las propiedades de \mathbf{s} y \mathbf{c} , se pueden justificar las principales identidades trigonométricas y dibujar sus gráficas que -como se sabe- resultan de la forma que se ilustra.



seno



coseno



tangente

Bibliografía

- [1] Andrews G., “Number Theory”, Philadelphia, Saunders, 1971
- [2] Anglin W. S., Lambek J., “The heritage of Thales”, Springer, Undergraduate texts in Mathematics, New York, 1995.
- [3] Beaumont, Pierce, “The algebraic foundations of Mathematics”, Addison-Wesley, 1963.
- [4] Birkhoff-Mac Lane, “A survey of modern algebra”, New York : Macmillan, 1977.
- [5] Cárdenas H., Lluis E., Raggi F., Tomás F., “Álgebra Superior”, Trillas, México1973.
- [6] Cohen Daniel, “Basic techniques of combinatorial theory”, New york, J. Wiley, 1978.
- [7] Copi Irving, “Lógica simbólica”. CECSA, México, 1979.
- [8] Courant, John, “Introducción al Cálculo y al Análisis Matemático”, Vol. I. Editorial Limusa, México, 1974.
- [9] Devlin K. “The joy of sets : fundamentals of contemporary set theory” Springer, Undergraduate Texts in Mathematics, New York1993.
- [10] Dickson, L. “A new First course in the theory of equations”, New York : J. Wiley, 1922.
- [11] Dodge C., “Sets, Logic and Numbers”, Prindle, Weber & Schmidt, 1969.
- [12] Espinosa R. “Introducción al razonamiento matemático”, Grupo Editorial América, México 2002.

- [13] Faddeev D., Somiinski I., “Problemas de álgebra superior”, Moscú : Mir, 1980.
- [14] Friedberg, Insel, Spence, “Álgebra lineal”, México, Publicaciones Culturales, 1982.
- [15] Gentile Enzo, “Aritmética elemental”, Monografías científicas de la OEA, número 25, 1985.
- [16] Gentile Enzo, “Aritmética elemental en la formación matemática”, Vínculos matemáticos número 211, Facultad de Ciencias, UNAM.
- [17] Halmos P. “Teoría intuitiva de los Conjuntos”, CECSA, México, 1984.
- [18] Hamilton, “Logic for Mathematicians”, Cambridge University Press, Cambridge, 1991.
- [19] Hamilton, “Numbers, Sets and Axioms: the Apparatus of Mathematics”, Cambridge University Press, Cambridge, 1982.
- [20] Hernández F. “Teoría de Conjuntos”, México, Sociedad Matemática Mexicana, 1998.
- [21] Herstein I., “Topics in Algebra”, New York : J. Wiley, 1975.
- [22] Jacobson N. “Basic Algebra”, vol. I, San Francisco, Freeman, 1974.
- [23] Johnsonbaugh R., “Matemáticas discretas”, Grupo Ed. Iberoamericana, México, 1993.
- [24] Kurosch, “Curso de Álgebra Superior”, Limusa, México, 1994.
- [25] Landau E., “Foundations of analysis : The arithmetic of whole, rational, irrational and complex numbers”, New York : Chelsea, 1957.
- [26] Lara M. “Los matemáticos griegos”, Universidad Autónoma de Querétaro, Querétaro, 1991.
- [27] Lehmann C., “Álgebra”, Limusa, México, 2003.
- [28] Liang-Shin Hahn, “Complex numbers and Geometry”, Mathematical Association of America, USA, 1994.

- [29] Lovasz L., “Combinatorial problems and exercises”, Amsterdam, North-Holland, 1979.
- [30] Mac Lane-Birkhoff, “Algebra”, Chelsea, 1999.
- [31] Mendelsohn E. , “Introduction to Mathematical Logic”, Princeton., Van nostrand, 1965
- [32] Niven I., Zuckerman H., “Introducción a la teoría de los números”, México, Limusa-Wiley,1969.
- [33] Pineda M. “Aritmetica y teoria de grupos”, UAM, Unidad Iztapalapa, México, 1995.
- [34] Ribnikov K. “Análisis combinatorio”, Moscú, Mir, 1988.
- [35] Ribnikov K. (Director), “Análisis combinatorio. Problemas y ejercicios”, Moscú, Mir, 1989.
- [36] Rincón Mejía Hugo, “Álgebra Lineal”, UNAM, Facultad de Ciencias, México, 2001.
- [37] Rincón Mejía Hugo, “Cuando cuentas cuántos...”, Temas de matemáticas para Bachillerato #1., Instituto de Matemáticas, UNAM, México.
- [38] Rincón Orta César, (Coordinador) “Lógica Matemática”, Departamento de Matemáticas, Facultad de Química, UNAM., México, 2004.
- [39] Rincón Orta César, “El conjunto de los números naturales”, Departamento de Matemáticas, Facultad de Química, UNAM., México, 2005.
- [40] Rincón Orta César, “El campo de los números reales”, Departamento de Matemáticas, Facultad de Química, UNAM., México, 2005.
- [41] Rincón Orta César, “El campo de los números complejos”, Departamento de Matemáticas, Facultad de Química, UNAM., México, 2005.
- [42] Rotman J., “An introduction to the theory of groups”, Springer, Graduate texts in mathematics, New York, 1995.
- [43] Rotman J., “Galois theory”, Springer, Universitext, New York, 1998

- [44] Struik D., “Historia concisa de las Matemáticas”, Instituto Politécnico Nacional, México, 1994.
- [45] Suppes, Hill, “Introducción a la Lógica Matemática”, Editorial Reverté, México, 1976.
- [46] Uspensky JV, Teoria de ecuaciones, Mexico, Limusa, 1987.
- [47] Vilenkin N., “¿De cuántas formas?”, Moscu, Mir, 1972.
- [48] Vinogradov I., “Fundamentos de la teoria de los numeros”, Moscu, Mir, 1971.
- [49] Zubieta G. “Manual de Lógica para estudiantes de Bachillerato”, México, Editorial. Trillas.
- [50] Zubieta F., “Álgebra elemental”, México, Edición del autor, 1982.

Índice de materias

- absurdos, 9
- algoritmo
 - de Euclides, 231, 237, 257
 - de la división para polinomios, 545
 - de la división para \mathbb{Z} , 208
- anillo, 197, 234
- anticadena, 93
- argumento de un número
- complejo, 398
- asociatividad de la composición de funciones, 103
- axioma
 - de especificación, 65
 - de extensión, 62
 - de las partes, 78
 - de regularidad, 84
 - del infinito, 107
- axiomas de Peano, 144
- base de un espacio vectorial, 436
- bases ortogonales, 450
- cadena, 93
- campo, 197
- cardinal de un producto, 303
- cardinalidad, 107
 - de un conjunto finito, 281
- unión de, 303
- cis* (θ), 389
- coeficiente principal de un renglón, 488
- combinación lineal, 431
- C_n^k , 291
- comparables, 92
- complejo
- argumento, 403
- complementos, 67
- composición
 - de funciones, 100
 - de relaciones, 94
- congruencias, 210
 - módulo n , 233
- conjugación, 377
- conjunción, 6
- conjunto
 - finito, 93, 110
 - inductivo, 108, 146
 - infinito, 92, 161
 - potencia, 78, 189
 - transitivo, 86, 159
- conmutatividad, 207, 350, 414
- contención, 63
- contradominio, 95
- contradominio de una relación, 81
- contrapuesta, 10, 17

- cortadura, 341
- cota
 - inferior, 90
 - superior, 87
- cuantificador
 - existencial, 57
 - universal, 57
- dependencia lineal, 431
- derivada de un polinomio, 582
- desigualdad del triángulo, 447
- determinante
 - cofactor, 528
 - de la transpuesta, 512
 - desarrollo respecto al primer renglón, 506
- diferencia simétrica, 77
- dilema constructivo, 18
- dilema destructivo, 18
- disyunción, 6
- divisibilidad,
 - relación de, 214
- dominio, 95
- dominio de una relación, 82
- dominio entero, 197
- e*, 643
- ecuación diofantina, 255
- ecuación general de segundo grado, 381
- escalares, 422
- espacio vectorial, 411
 - finitamente generado, 437
- exportación, 18
- familias de conjuntos, 69, 74
- función, 96
- continua, 360
- lineal, 456
- suprayectiva, 101
- exponencial, 401, 640
- inyectiva, 98
- polinomial, 547
- grado
 - de un polinomio, 381, 546
- grupo, 414
- ideal, 212
- igualdad entre conjuntos, *véase*
 - Axioma de extensión
- imagen, 95
- imagen inversa, 112
- implicación, 7
- incomparables, 93
- independencia lineal, 434
- inducción, 147
 - segundo principio, 187
- ínfimo, 91
- intersección, 67
- inversión, 498
- inverso aditivo de un número real, 345
- Lema de Schwarz, 447
- leyes
 - de los exponentes, 638
 - de De Morgan para conjuntos, 73
 - distributivas para conjuntos, 76
- matriz, 452
 - elemental, 520, 523

- menor de una, 505
- reducida y escalonada, 488
- máximo común divisor de enteros, 216
- mayor elemento, 91
- menor, 87
 - de una matriz, 501
- método de Horner, 565
- mínimo común múltiplo, 218
- modus ponens, 17
- modus tollendo ponens, 17
- monoide, 412
-
- negación, 5
- neutro
 - derecho, 140, 412
 - izquierdo, 140, 412
- norma euclíadiana, 446
- numeración base b , 265, 520
- número
 - complejo, 369
 - argumento, 398
 - parte imaginaria, 376
 - parte real, 376
 - de funciones suprayectivas de n a m , 309
 - de particiones de un conjunto con n elementos en m partes, 311
 - de permutaciones de un conjunto finito, 301
 - entero, 199
 - primo, 226
- O^r_n , 301
- operación, 411
-
- asociativa, 137
- comutativa, 137
- elemental de renglón, 489
- operación en un conjunto, 133
- operaciones elementales, 475
- orden
 - lexicográfico, 92
 - total, 93
- orden en un producto, 92
- ordenación, 300
 - con repetición, 313
- O^r_m , 310
- ortogonalidad, 450
-
- pareja, 70
- pareja ordenada, 79
- partes de un conjunto, véase conjunto
- partición, 121
- permutación, 298
- paridad de una, 502
- signo, 502
- P^r_m , 310
- relación de recurrencia, 311
- polinomio
 - creciente, 574
 - derivada, 582
- polinomios, 539
- principio
 - de las pichoneras, 291
 - del Buen Orden, 170
 - de inducción, 147
- producto
 - cartesiano, 81
 - de matrices, 466
 - de números complejos, 388

- de números naturales, 183
- de números reales, 348
- de polinomios, 543
- de racionales, 270, 274
- interior, 446
- punto, 445
- proposición, 2
- quinto postulado de Peano, 147
- raíz
 - multiplicidad de una, 591
 - cuadrada de un segmento, 336
 - n -ésima de un número complejo, 395
- rango de una matriz, 453
- reales positivos, 348
- recursión, 171
 - generalizada, 150
- reducción al absurdo, 30
- reflexiones, 458
- regla
 - de inferencia, 15
 - de la cadena, 634
 - de la tautología, 21
 - de los signos de Descartes, 605
 - del reemplazo, 18
 - de Cramer, 531
- relación, 82
 - antisimétrica, 95
 - de equivalencia, 117
 - diagonal, 118
- representación decimal, 363
- restricción
 - de una operación, 135
- de una relación, 94
- retícula, 88
 - completa, 91
 - superior, 90
- rotación, 458
- semigrupo, 137, 412
- silogismo hipotético, 17
- sistema
 - completo de conectivos, 11
 - homogéneo asociado, 471
- sistemas de congruencias, 243, 247
- S^n_m , 309
- soporte, 237
 - de un polinomio, 539
- subespacio, 425
 - generado por un conjunto, 429
- suma
 - de números naturales, 176
 - de números racionales, 270
 - de números reales, 343
 - de polinomios, 541
 - de subespacios, 430
- supremo, 91
- tablas
 - de multiplicar, 138
 - de verdad, 5
- tautologías, 9
- teorema
 - Chino del residuo, 252
 - de Cantor-Bernstein-Schröeder, 108
 - de De Moivre, 392

- de existencia y unicidad de soluciones para una ecuación diferencial, 639
- de la deducción, 41
- de recursión, 171
- de Sturm, 594
- del factor, 549
- del residuo, 549
- del valor intermedio, 557
- del valor medio, 587
- fundamental de la Aritmética, 313, 226
- fundamental del Álgebra, 381, 558, 587
- fundamental del Cálculo, 634
- tollendo tollens, 17
- transposición, 497

- uniones, 69

- valuación, 53
- vector unitario, 448
- vectores, 421

- \mathbb{Z}_n , 234