Carlos Ivorra Castillo

LÓGICA Y TEORÍA DE CONJUNTOS

No puedes encontrar la verdad con la lógica si no la has encontrado ya sin ella.

G.K. Chesterton

Índice General

Primera parte

Lógica de primer orden

Introducción a la lógica matemática

La lógica y su historia Tradicionalmente se ha dicho que la lógica se ocupa del estudio del razonamiento. Esto hoy en día puede considerarse desbordado por la enorme extensión y diversidad que ha alcanzado esta disciplina, pero puede servirnos como primera aproximación a su contenido.

Un matemático competente distingue sin dificultad una demostración correcta de una incorrecta, o mejor dicho, una demostración de otra cosa que aparenta serlo pero que no lo es. Sin embargo, no le preguntéis qué es lo que entiende por demostración, pues —a menos que además sepa lógica— no os sabrá responder, ni falta que le hace. El matemático se las arregla para reconocer la validez de un argumento o sus defectos posibles de una forma improvisada pero, al menos en principio, de total fiabilidad. No necesita para su tarea contar con un concepto preciso de demostración. Eso es en cambio lo que ocupa al lógico: El matemático demuestra, el lógico estudia lo que hace el matemático cuando demuestra.

Aquí se vuelve obligada la pregunta de hasta qué punto tiene esto interés y hasta qué punto es una pérdida de tiempo. Hemos dicho que el matemático se las arregla solo sin necesidad de que nadie le vigile los pasos, pero entonces, ¿qué hace ahí el lógico? Posiblemente la mejor forma de justificar el estudio de la lógica sea dar una visión, aunque breve, de las causas históricas que han dado a la lógica actual tal grado de prosperidad.

En el sentido más general de la palabra, el estudio de la lógica se remonta al siglo IV a.C., cuando Aristóteles la puso a la cabeza de su sistema filosófico como materia indispensable para cualquier otra ciencia. La lógica aristotélica era bastante rígida y estrecha de miras, pero con todo pervivió casi inalterada, paralelamente al resto de su doctrina, hasta el siglo XVI. A partir de aquí, mientras su física fue sustituida por la nueva física de Galileo y Newton, la lógica simplemente fue ignorada. Se mantuvo, pero en manos de filósofos y en parte de los matemáticos con inclinaciones filosóficas, aunque sin jugar ningún papel relevante en el desarrollo de las ciencias. Leibniz le dio cierto impulso, pero sin abandonar una postura conservadora. A principios del siglo XIX, los trabajos de Boole y algunos otros empezaron a relacionarla más directamente con la matemática, pero sin obtener nada que la hiciera especialmente relevante

(aunque los trabajos de Boole cobraran importancia más tarde por motivos quizá distintos de los que él mismo tenía *in mente).*

Así pues, tenemos que, hasta mediados del siglo XIX, la lógica era poco más que una curiosidad que interesaba a quienes sentían alguna inquietud por la filosofía de la matemática o del pensamiento en general. La lógica como hoy la entendemos surgió básicamente con los trabajos de Frege y Peano. En principio éstos eran, al igual que los anteriores, nuevos ensayos sobre el razonamiento, si bien más complejos y ambiciosos. Lo que les dio importancia fue que no aparecieron como productos de mentes inquietas, sino como culminación del proceso de formalización que la matemática venía experimentando desde los tiempos de Newton y Leibniz.

En efecto, el cálculo infinitesimal que éstos trazaron con tanta imaginación y que después desarrollaron Cauchy, Gauss y otros, tuvo que ser precisado a medida que se manejaban conceptos más generales y abstractos. Dedekind, Riemann, Weierstrass, fueron sistematizando la matemática hasta el punto de dejarla construida esencialmente a partir de los números naturales y de las propiedades elementales sobre los conjuntos. La obra de Frege y de Peano pretendía ser el último eslabón de esta cadena. Trataron de dar reglas precisas que determinaran completamente la labor del matemático, explicitando los puntos de partida que había que suponer así como los métodos usados para deducir nuevos resultados a partir de ellos.

Si sólo fuera por esto, probablemente este trabajo habría acabado como una curiosidad de presencia obligada en las primeras páginas de cada libro introductorio a la matemática y que continuaría interesando tan sólo a los matemáticos con inclinaciones filosóficas. Pero sucedieron hechos que confirmaron la necesidad de la lógica como herramienta matemática. A finales del siglo XIX, Georg Cantor creó y desarrolló la parte más general y más abstracta de la matemática moderna: la teoría de conjuntos. No pasó mucho tiempo sin que el propio Cantor, junto con otros muchos, descubriera descaradas contradicciones en la teoría, es decir, se obtenían demostraciones de ciertos hechos y de sus contrarios, pero de tal forma que burlaban el ojo crítico del matemático, tan de fiar hasta entonces. Se obtenían pares de pruebas de forma que cada una por separado parecía irreprochable pero que ambas juntas eran inadmisibles.

El ejemplo más simple de estos resultados fue descubierto por Bertrand Russell al despojar de contenido matemático a otro debido a Cantor: En la teoría cantoriana se puede hablar de cualquier conjunto de objetos con tal de que se especifiquen sus elementos sin ambigüedad alguna. En particular podemos considerar el conjunto R cuyos elementos son exactamente aquellos conjuntos que no son elementos de sí mismos. Es fácil ver que si R es un elemento de sí mismo, entonces por definición no debería serlo, y viceversa. En definitiva resulta que R no puede ni pertenecerse como elemento ni no hacerlo. Esto contradice a la lógica más elemental.

El lector puede pensar que esto es una tontería y que basta no preocuparse de estas cosas para librarnos de tales problemas, sin embargo sucede que contradicciones similares surgen continuamente en la teoría pero afectando a conjuntos no tan artificiales y rebuscados como pueda parecer el conjunto R, sino a otros

que aparecen de forma natural al trabajar en la materia. En cualquier caso estos hechos mostraban que el criterio que confiadamente han venido usando desde siempre los matemáticos no es inmune a errores difíciles —por no decir imposibles— de detectar, al menos al enfrentarse a la teoría de conjuntos.

La primera muestra de la importancia de la lógica fue un estrepitoso fracaso. Frege había creado (tras mucho tiempo de cuidadosa reflexión) un sistema que pretendía regular todo el razonamiento matemático, de manera que cualquier resultado que un matemático pudiera demostrar, debería poder demostrarse siguiendo las reglas que con tanto detalle había descrito. Russell observó que la paradoja antes citada podía probarse en el sistema de Frege y que, a consecuencia de esto, cualquier afirmación, fuera la que fuera, podía ser demostrada según estas reglas, que se volvían, por tanto, completamente inútiles.

Este desastre, no obstante, mostraba que la laboriosa tarea de Frege no era en modo alguno trivial, y urgía encontrar una sustituta a su fallida teoría. Con el tiempo surgieron varias opciones. La primera fueron los Principia Mathematica de Whitehead y Russell, de una terrible complejidad lógica, a la que siguieron muchas teorías bastante más simples aunque quizá menos naturales. Destacan entre ellas las teorías de conjuntos de Zermelo-Fraenkel (ZF) y de von Neumann-Bernays-Gödel (NBG). Ambas constan de unos principios básicos (axiomas) y unas reglas precisas de demostración que permiten deducir de ellos todos los teoremas matemáticos y —hasta donde hoy se sabe— ninguna contradicción.

De esta forma la lógica ha probado ser indispensable a la hora de trabajar en teoría de conjuntos, hasta el punto de que es inconcebible el estudio de ésta sin un buen conocimiento de aquélla.

El contenido de la lógica matemática En el apartado anterior hemos mostrado una de las funciones principales de la lógica matemática: servir de fundamento al razonamiento matemático, evitando ambigüedades y contradicciones mediante la determinación absolutamente precisa y rigurosa de lo que es un razonamiento matemático válido. Pero cuando la necesidad obliga al estudio de un determinado campo, el esfuerzo pronto es premiado con nuevos resultados inesperados:

Si uno tiene paciencia o un libro de geometría a mano, puede coger una regla y un compás y dibujar un pentágono regular. Si ahora prueba suerte con un heptágono no encontrará ningún libro de ayuda y la paciencia servirá de muy poco. Puede probarse que es imposible construir un heptágono regular sin más ayuda que una regla (no graduada) y un compás, pero, para demostrarlo no basta con coger una regla y un compás y terminar no construyéndolo. Es necesario reflexionar sobre qué es construir con regla y compás, dar una definición precisa, comprobar que ésta se corresponde con lo que usualmente se entiende por construir con regla y compás y, finalmente, ver que eso es imposible para el caso del heptágono regular.

Igualmente, el tener una noción precisa de demostración nos permite comprender y resolver problemas que de otro modo serían inabordables: cuando un matemático hace una conjetura, puede meditar sobre ella y, si tiene suerte, la demostrará o la refutará. Pero también puede ser que no tenga suerte y no

consiga ni lo uno ni lo otro. Esto último puede significar dos cosas: que no es lo suficientemente buen matemático o que pretendía un imposible. Cantor llegó a la locura en gran parte por la frustración que le producía el no lograr decidir la verdad o falsedad de una de sus conjeturas, la llamada hipótesis del continuo. Con ayuda de la nueva lógica se ha probado que ésta no puede probarse ni refutarse, y no se trata de un caso aislado. Sucede que estas afirmaciones no surgen sólo en teoría de conjuntos, donde son el pan de cada día, sino que son también abundantes en el análisis y la topología, incluso hay casos en álgebra. Por ello el matemático necesita en ocasiones de la lógica para determinar sus propias posibilidades y limitaciones. El establecer este tipo de resultados de independencia es una de las partes más importantes de la lógica aplicada a la teoría de conjuntos.

Por otra parte, toda teoría suficientemente rica contiene resultados de interés interno, en sí mismo. La lógica moderna, principalmente de la mano de Gödel, ha obtenido resultados sorprendentes e interesantísimos que nos permiten comprender mejor la capacidad y las limitaciones del razonamiento humano, resultados que justifican por sí solos el estudio de la lógica. Por ejemplo: ¿Puede un matemático probar que $2+2=5$? El lector que responda: "Claramente no", o "No, porque es mentira", o "No, porque $2+2=4$", o similares, no tiene claros ciertos conceptos lógicos. Está claro que un matemático puede demostrar que $2+2=4$, más aún, está claro que $2+2=4$, pero el problema es que la existencia de una demostración de que $2+2\neq 5$ o incluso de la falsedad de que $2+2=5$ no aportan la menor garantía de que no pueda traer alguien unos cuantos folios escritos según las "costumbres" de razonamiento de los matemáticos, aun cumpliendo todas las condiciones que estipulan los lógicos, pero que termine con la conclusión $2+2=5$. ¿Por qué no puede ser? No es un problema evidente, hasta el punto de que puede probarse —como consecuencia del llamado segundo teorema de incompletitud de Gödel— que es imposible garantizar que no exista tal catastrófica prueba. Lo demostraremos en su momento.

Sin ánimo de ser exhaustivos, podríamos decir que la lógica moderna se divide en cuatro áreas:

a) Teoría de la demostración.

b) Teoría de modelos.

c) Teoría de la recursión.

d) Teoría de conjuntos.

En esta primera parte haremos especial hincapié en la teoría de la demostración, que es la parte más clásica de la lógica, y usaremos la teoría de modelos y la teoría de la recursión como auxiliares para el estudio de la primera. Finalmente aplicaremos los resultados que obtendremos a la teoría de conjuntos como ejemplo más significativo. Vamos a probar la mayoría de los resultados clásicos de la teoría de la demostración, mientras que la teoría de modelos y la teoría de la recursión serán tocadas muy superficialmente, con la suficiente profundidad como para obtener resultados importantes que nos serán necesarios, pero no

como para formarnos una idea del trabajo que se lleva a cabo en estos campos. Este planteamiento es el más conveniente para los objetivos que perseguimos, que son dos: por una parte dotar al lector de un bagaje lógico más que suficiente para abordar con comodidad el estudio de la teoría de conjuntos, y por otra, tratar de explicar a través de estos resultados la naturaleza del trabajo del matemático.

Matemática y metamatemática Una gran parte de la lógica moderna constituye una rama más de la matemática, como pueda serlo el álgebra o el análisis, pero hay otra parte que no puede ser considerada del mismo modo, y es precisamente la que más nos va a interesar. Se trata de la parte que se ocupa de los fundamentos de la matemática. Para que un argumento matemático sea aceptable es necesario que satisfaga unas condiciones de rigor, condiciones que los matemáticos aplican inconscientemente y que ahora nos proponemos establecer explícitamente, pero precisamente por eso sería absurdo pretender que los razonamientos y discusiones que nos lleven a establecer el canon de rigor matemático deban someterse a dicho canon, del que —en nuestra peculiar situación— no disponemos a priori. Esto plantea el problema de cómo ha de concebirse todo cuanto digamos hasta que dispongamos de la noción de rigor matemático.

Esto nos lleva a la distinción entre *matemática* y *metamatemática.* Matemática es lo que hacen los matemáticos. Cuando hayamos alcanzado nuestro objetivo, podremos decir qué es exactamente hacer matemáticas. De momento podemos describirlo *grosso modo:* Hacer matemáticas consiste en demostrar afirmaciones, en un sentido de la palabra "afirmación" que hemos de precisar y en un sentido de la palabra "demostrar" que hemos de precisar, a partir de unas afirmaciones fijas que llamaremos axiomas y que también hemos de precisar.[1] Por otra parte, hacer metamatemáticas es razonar sobre afirmaciones, demostraciones, axiomas y, en general, sobre todo aquello que necesitemos razonar para establecer qué es la matemática y cuáles son sus posibilidades y sus límites.

Por ejemplo, una afirmación matemática es "los poliedros regulares son cinco", mientras que una afirmación metamatemática es "los axiomas de Peano son cinco". Pese a su similitud formal, es crucial reconocer que son esencialmente distintas. Cuando hayamos "capturado" la noción de razonamiento matemático, podremos entender la primera de ellas como un teorema, una afirmación cuya verdad se funda en que puede ser demostrada matemáticamente, mediante un razonamiento que satisfará todas las exigencias de rigor que habremos impuesto. En cambio, la segunda no es un teorema demostrable a partir de ningunos axiomas. Simplemente expresa que cuando escribimos en un papel los axiomas de Peano, escribimos cinco afirmaciones. Cuando contamos los axiomas de Peano hacemos lo mismo que cuando le contamos los pies a un gato. Podrá discutirse sobre qué es lo que hacemos, pero, ciertamente, no estamos demostrando un teorema formal.

[1] Ciertamente, esta concepción radicalmente formalista de las matemáticas es más que cuestionable. En realidad no afirmo que las matemáticas sean sólo esto, sino tan sólo que éste es exactamente el significado que tendrá el término "matemático" a lo largo de este libro.

Antes de continuar debo hacer una advertencia al lector: Los resultados que vamos a estudiar son todos hechos conocidos sobre la lógica de primer orden, que merecen el respeto y la consideración habituales para con los resultados matemáticos, sin embargo, entre éstos, hay interpretaciones subjetivas con las que unos lógicos y matemáticos estarán de acuerdo mientras que otros podrán discrepar. Mi intención no ha sido la de exponer imparcialmente todos los puntos de vista posibles, sino la de decantarme en cada momento por lo que me parece más adecuado, de modo que el lector es libre de estar de acuerdo o discrepar de lo que lea. Si el lector opta por lo segundo, debería tener presente que hay dos formas de discrepar: una destructiva y estéril, consistente únicamente en discrepar, y otra constructiva y enriquecedora, consistente en proponer una alternativa. Tengo la convicción de que el lector que trate de discrepar constructivamente no discrepará mucho.

La diferencia esencial entre una afirmación o un razonamiento matemático y una afirmación o un razonamiento metamatemático es que los primeros se apoyan esencialmente en una teoría axiomática, y los segundos no. Cuando afirmamos que "los poliedros regulares son cinco", aunque literalmente esto es una afirmación en castellano, si la consideramos como una afirmación matemática correcta es porque podríamos enunciarla en el lenguaje de la teoría de conjuntos y demostrarla según la lógica de la teoría de conjuntos. Por el contrario, la afirmación "los axiomas de Peano son cinco" es una afirmación en castellano, que podríamos traducir al inglés o al francés, pero no tiene sentido considerarla como un teorema integrante de un sistema axiomático.[2] Todo matemático, tanto si conoce explícitamente la teoría axiomática en la que trabaja como si no, entiende perfectamente qué es razonar formalmente en el seno de una teoría y, aunque no sepa —conscientemente— mucha lógica, entiende que eso es precisamente lo que hace y lo que da rigor a su trabajo. El problema es, pues, explicar cómo puede razonarse de forma rigurosa fuera de toda teoría axiomática. Dedicaremos a este problema las secciones siguientes. Para acabar ésta añadiremos únicamente la siguiente advertencia:

Un matemático puede encontrar esotéricos e incomprensibles o naturales y simples los resultados de los capítulos siguientes, no en función de su inteligencia o de su capacidad como matemático, sino exclusivamente en función de su capacidad de librarse de los prejuicios o de la "deformación profesional" que le impidan asumir que no está leyendo un libro de matemáticas. Si decide prescindir de las indicaciones que acompañan a los resultados, más cercanas a la filosofía que a la matemática en sí, corre el riesgo de entender todos los pasos intermedios pero no entender ninguna de las conclusiones.

El formalismo radical Antes de esbozar una concepción razonable para la metamatemática, será conveniente que descartemos de antemano la alternativa a la que es proclive una buena parte de los matemáticos no familiarizados con

[2]En realidad la metamatemática sí puede formalizarse, como cualquier teoría razonable, pero lo cierto es que en nuestro contexto no podemos hacerlo, por lo que es más aproximado a la verdad decir que no tiene sentido considerar a sus afirmaciones como teoremas de ninguna teoría formal.

la lógica: el formalismo radical. Ya hemos comentado que las contradicciones que achacaban a la matemática de finales del siglo XIX fueron desterradas estipulando unos axiomas y unas reglas de razonamiento lógico cuidadosamente seleccionadas para este fin. Más allá de cubrir esta necesidad elemental de consistencia, el método axiomático proporciona al matemático una seguridad sin precedentes: decidir si un razonamiento es válido o no cuando la teoría a la que pretende integrarse está debidamente axiomatizada es mera cuestión de cálculo, una tarea mecánica que, al menos en teoría, puede realizar incluso un ordenador debidamente programado.

Esto ha hecho que algunos matemáticos, convencidos de que el método axiomático es todo lo que necesitan para su trabajo, no reconozcan otra forma de razonamiento legítimo. Un formalista radical es alguien que no acepta un razonamiento a no ser que venga precedido de una enumeración de los conceptos que va a involucrar y de los axiomas que se van a aceptar sin prueba, y de modo que todo cuanto siga sean consecuencias lógicas formales de los axiomas dados (sin perjuicio de que, en la mayoría de los casos, estos principios se omitan por consabidos).

Es importante destacar el significado del adjetivo "formal" en la expresión "consecuencias lógicas formales". Una deducción formal es una deducción que no tiene en cuenta para nada el posible significado de las afirmaciones que involucra. Por ejemplo, de "todo H es M" y "S es H" se deduce formalmente que "S es M", sin que importe lo más mínimo a qué hagan referencia las letras H, M y S. Si uno quiere ver ahí el silogismo "Todos los hombres son mortales", "Sócrates es un hombre", luego "Sócrates es mortal", es libre de pensarlo así, pero la validez del razonamiento no depende de esa interpretación ni de ninguna otra.[3]

Hilbert fue el primero en concebir la posibilidad de reducir la totalidad de la matemática a una teoría axiomática formal, idea extremadamente fructífera y poderosa. La falacia del formalista radical —en la que, desde luego, Hilbert no cayó— consiste en creer que no hay nada más. En las secciones siguientes veremos qué más hay, pero en ésta hemos de convencernos de que algo más tiene que haber.

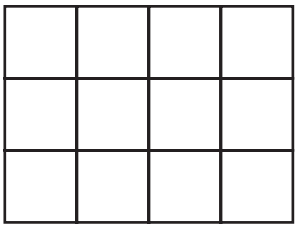
No es cierto que el formalismo radical baste para fundamentar la matemática. El problema es que establecer un lenguaje, unos axiomas y unas reglas de razonamiento requiere ciertos razonamientos: hay que discutir cuáles son los signos del lenguaje, cuáles son las combinaciones aceptables de esos signos, cuáles de ellas se toman concretamente como axiomas, hay que demostrar algunos hechos generales sobre demostrabilidad, etc. ¿Cómo podrían entenderse esos razonamientos si no admitiéramos razonamientos que no provengan de unos axiomas prefijados?, ¿hemos de presentar axiomáticamente la metamatemática?, ¿y cómo presentamos los axiomas necesarios para axiomatizar la metamatemática?, ¿Hemos de construir una metametamatemática?

Por poner un ejemplo explícito: La teoría de conjuntos de Zermelo-Fraenkel es el sistema axiomático comúnmente aceptado como fundamento de la ma-

[3]Por eso una buena definición del formalista (radical) es la que lo caracteriza como alguien incapaz de entender algo a menos que carezca de significado.

temática. En efecto, a partir de sus axiomas se pueden demostrar todos los teoremas matemáticos, en particular de ellos se deducen las propiedades de los conjuntos infinitos. Un formalista radical sólo aceptará razonamientos que involucren el concepto de infinitud a partir del momento en que las propiedades de los conjuntos infinitos se hayan demostrado a partir de los axiomas, pero sucede que la teoría de conjuntos de Zermelo-Fraenkel tiene infinitos axiomas. Por consiguiente, el formalismo radical conduce a descalificar como falto de rigor a su propio canon de rigor. Por eso sólo son formalistas radicales quienes, con independencia de su capacidad como matemáticos, jamás han abordado con detalle —no a nivel teórico general, sino a nivel técnico— el problema de fundamentar rigurosamente la matemática.

El finitismo No toda la matemática necesita una fundamentación axiomática formal. Ésta es necesaria sólo porque la matemática trata con conjuntos infinitos. Si un matemático trabaja exclusivamente con conjuntos finitos, por ejemplo, grafos finitos, grupos finitos, etc., puede prescindir por completo de axiomas y reglas de razonamiento formal. Nadie ha encontrado jamás una paradoja que involucre exclusivamente conjuntos finitos[4] ni error de razonamiento sobre conjuntos finitos que no sea detectable sin más que prestar suficiente atención al discurso. Esto vuelve remilgados y vanos —en este contexto— muchos de los escrúpulos del formalista radical. Pongamos algunos ejemplos. Es fácil calcular $3 \times 4 = 12$ y $4 \times 3 = 12$, lo que nos convence de que $3 \times 4 = 4 \times 3$. Hay, sin embargo, una forma de razonarlo que es especialmente fructífera. Pensemos en el rectángulo siguiente:



Podemos considerarlo formado por 3 veces 4 cuadrados o por 4 veces 3 cuadrados, lo que muestra que, necesariamente $3 \times 4 = 4 \times 3$. Esto ya lo sabíamos, pero hay una diferencia: si calculamos $3 + 3 + 3 + 3$ y $4 + 4 + 4$ y vemos que da lo mismo, sabemos eso y nada más que eso, mientras que el argumento del rectángulo nos convence de que $m \times n = n \times m$ para cualquier par de números m y n (no nulos, en principio). En efecto, está claro que, sean quienes sean m y n, siempre podremos construir un rectángulo formado por m filas de n cuadrados o, equivalentemente, por n columnas de m cuadrados. Vemos así que —para desesperación de un formalista radical— la prueba de un caso particular contiene la prueba del caso general.

Quien considere que de un caso particular —o incluso de varios— nunca es lícito inferir el caso general, está generalizando ilícitamente a partir de uno o

[4]Podría objetarse que "el menor número natural no definible con menos de doce palabras" es contradictorio, pero es que aquí la noción de "definible" no está bien definida.

varios casos particulares. Por ejemplo, no es muy difícil probar que la ecuación $x^3 + y^3 = z^3$ no tiene soluciones enteras, pero la prueba no muestra más que eso, de modo que no es lícito deducir de ella que la ecuación $x^n + y^n = z^n$ no tiene soluciones enteras para $n > 2$. El hecho de que los primeros números de la forma $2^{2^n} + 1$ sean primos no nos permite asegurar que todos ellos lo sean. En ambos casos tenemos meras comprobaciones aisladas que no aportan nada sobre el caso general. Por el contrario, el argumento del rectángulo contiene un esquema uniforme de razonamiento, en el sentido de que cualquiera que comprenda el argumento se sabe capacitado para generar razonamientos concretos que prueben la conmutatividad de cualquier par de factores.[5]

El argumento del rectángulo es un ejemplo de razonamiento finitista que nos proporciona una verdad sobre los números naturales. El formalista radical preguntará qué debemos entender por "números naturales" y "producto" en dicho razonamiento. No podemos permitirnos el lujo de responderle como a él le gustaría: *necesitamos* los números naturales para fundamentar la matemática, es decir, mucho antes de estar en condiciones de responder a las exigencias del formalista. Eso no nos exime de responder:

Cójase a un niño que no sepa contar pero que esté en edad de aprender. Enséñesele a contar. Con ello, el niño habrá pasado de no saber contar a saber contar. Algo habrá aprendido. Lo que ha aprendido es lo que son los números naturales. Sería inútil que repitiera aquí lo que no sería ni más ni menos que lo que el lector aprendió en su infancia. Del mismo modo, "multiplicar" es eso que todos sabemos hacer cuando nos dan una expresión como "$12 \times 345 =$" y nos piden que la completemos. Es una operación que nos lleva de dos números a otro número de forma objetiva, en el sentido de que dos personas cualesquiera que sepan multiplicar llegarán siempre al mismo resultado y, de no ser así, será fácil sacar de su error a quien se haya equivocado.

Supongamos que hemos enseñado a contar a un niño de tal modo que éste es capaz de decidir cuál de dos números naturales dados (en forma decimal, por ejemplo) es mayor, así como de escribir el siguiente de un número dado. En cuanto tenga esto debidamente asimilado, pregúntesele cuál es el mayor de todos los números. Sin duda responderá que no hay tal número, pues él se sabe capaz de superar cualquier número que le sea dado. A poco que se le explique la diferencia entre lo finito y lo infinito, sabrá ver ahí la prueba de que el conjunto de los números naturales es infinito. Quizá no sepa si el conjunto de las estrellas es finito o infinito, pero sabrá que el conjunto de los dedos de su mano es finito y el conjunto de los números es infinito.

El punto crucial es que estos conocimientos no son precarios y basados en la credulidad de los niños, sino que son firmes y objetivos, en el sentido de que, en cuanto un niño ha comprendido adecuadamente el significado de los términos "número", "finito" e "infinito", tal vez podremos engañarle y hacerle

[5]La clave está en que se sabe capacitado a priori. En realidad, cualquiera está capacitado para ello aunque pueda no saberlo: basta calcular $m \times n$ y $n \times m$ y comprobar que da lo mismo. La diferencia es que quien conoce el argumento del rectángulo sabe de antemano que su argumento va a funcionar con factores cualesquiera, mientras que quien hace las operaciones no tiene la seguridad en cada caso hasta que no acaba los cálculos. Por eso no puede asegurar que la multiplicación es conmutativa.

creer cualquier cosa sobre el número de estrellas, pero jamás conseguiremos que crea que tiene infinitos dedos en su mano o que hay una cantidad finita de números naturales. Las afirmaciones estrictamente matemáticas sobre los números nunca han generado ni pueden generar polémica sobre si son verdaderas o falsas.[6]

Estos ejemplos pretenden mostrar que es posible razonar con objetividad, seguridad, precisión y, por consiguiente, rigurosamente, sobre algunos conceptos sin depender de sistemas axiomáticos. ¿De qué conceptos, concretamente? Es muy difícil, si no imposible, establecer fronteras precisas. El finitismo consiste en aceptar que el razonamiento humano no corre riesgos de extravío mientras se limite a considerar conceptos y procesos finitos. Así, Hilbert, en su programa de fundamentación de la matemática, propugnó la búsqueda de un sistema axiomático adecuado para este fin, de modo que, tanto la construcción del sistema como la comprobación de que satisfacía los requisitos necesarios para considerarlo aceptable, tenía que llevarse a cabo mediante argumentos finitistas que —por consiguiente— no requirieran la teoría buscada y no nos llevaran así al callejón sin salida al que conduce inexorablemente el formalismo radical.

En definitiva, la propuesta de Hilbert era fundamentar la matemática, no finitista en su mayor parte, con una metamatemática finitista, que carece de los problemas característicos de la matemática no finitista —que el formalista radical extrapola catastróficamente a toda la matemática— y por consiguiente no requiere de una fundamentación formal para justificar su solidez.

Esto no significa que no se pueda especular sobre la fundamentación de la metamatemática, pero ésta ya no corresponde al ámbito de la matemática o de la lógica, sino de la teoría del conocimiento, y el matemático puede prescindir de tratar este problema ya que, en todo caso, la cuestión sería en qué se funda nuestra capacidad de razonamiento básico, no si dicha capacidad es o no sólida y fiable.[7]

Más allá del finitismo Aunque la mayor parte de la metamatemática puede desarrollarse en el marco finitista que exigía Hilbert, lo cierto es que algunos resultados valiosos, como el teorema de completitud semántica de Gödel, exigen nuestra confianza en argumentos algo más audaces. Por ello conviene cambiar la pregunta más tímida de ¿qué tipo de razonamientos necesitamos sostener sin el apoyo de una teoría axiomática? por la más ambiciosa de ¿qué tipo de razonamientos podemos sostener sin el apoyo de una teoría axiomática?

La tesis general que adoptaremos aquí es la siguiente: Para que un razonamiento sea aceptable metamatemáticamente ha de cumplir dos condiciones:

a) Ha de ser convincente, en el sentido de que nadie que lo comprenda pueda

[6]Otra cosa es polemizar sobre si podemos asegurar que cualquier afirmación sobre números es verdadera o falsa, especialmente cuando no sabemos cómo comprobarla, pero jamás —que yo sepa— ha habido dos personas que se creyeran con argumentos racionales que probaran tesis opuestas sobre una propiedad de los números naturales o de conjuntos finitos en general.

[7]Evidentemente, se puede dudar de la fiabilidad de nuestra capacidad de razonamiento finitista como se puede dudar de si existe o no el mundo, pero eso es escepticismo, un mal que sólo afecta a los que hablan por hablar y a los que piensan por pensar.

tener dudas serias[8] sobre la verdad de su tesis.

b) Todas las afirmaciones involucradas han de tener un significado preciso y objetivo independiente de los argumentos que las demuestren.

Nos encontramos aquí con un fenómeno omnipresente en la metamatemática: mientras el matemático está acostumbrado a ir de lo general a lo particular (así por ejemplo, sólo después de definir la noción general de continuidad de una función es cuando se plantea si una función dada es o no continua) esta actitud rara vez es posible en la metamatemática. De este modo, aunque no tenemos ninguna definición general, objetiva y precisa de qué es un razonamiento convincente —y por consiguiente el enunciado de la condición a) es obviamente ambiguo e impreciso—, afortunadamente, no necesitamos tenerla para reconocer un argumento objetivo y preciso (en particular convincente) cuando lo tenemos delante. Por ejemplo, el argumento del rectángulo demuestra la conmutatividad del producto de números naturales sin dejar lugar a dudas. Su poder de convicción es objetivo en el sentido de que no depende de la capacidad de sugestión o de dejarse engañar de quien lo escucha, sino que, por el contrario, nadie que lo conozca puede albergar ya el menor recelo de encontrarse con un par de números que al multiplicarlos en uno y otro orden produzcan resultados distintos.

La segunda condición está relacionada con la diferencia fundamental entre matemática y metamatemática: cuando un matemático trabaja en el seno de una teoría axiomática formal, no está legitimado a hablar de la verdad o falsedad de las afirmaciones que demuestra. Para él sólo hay afirmaciones demostrables y no demostrables o, si se quiere hilar más fino, afirmaciones demostrables, refutables e independientes de sus axiomas (las que no se pueden demostrar o refutar). En cambio, en metamatemática no podemos hacer esta distinción ya que no tenemos una noción precisa de lo que es ser (metamatemáticamente) demostrable. Nuestra única posibilidad, pues, de distinguir afirmaciones como $2+2=4$, $2+2=5$ y $2^{\aleph_0}=\aleph_1$ es decir que la primera es verdadera, la segunda es falsa y la tercera no tiene significado metamatemático porque no cumple la condición b). Una vez más, no tenemos una definición general de qué quiere decir que una afirmación sea verdadera, pero sí sabemos lo que quiere decir que algunas afirmaciones sean verdaderas, y esas afirmaciones son las únicas que podemos permitirnos el lujo de manejar metamatemáticamente. Pongamos algunos ejemplos.

Sabemos demostrar que el producto de números naturales es conmutativo, pero, independientemente de cualquier razonamiento que nos convenza de ello, sabemos lo que eso significa: significa que si tomamos dos números cualesquiera y hacemos lo que sabemos que hay que hacer para calcular su producto, el resultado es el mismo independientemente del orden en que los tomemos. A priori habría dos posibilidades: que hubiera pares de números para los que esto fuera falso o que no los hubiera. Tenemos un razonamiento que nos convence de que la primera posibilidad es, de hecho, imposible, pero es esencial que antes de tal razonamiento ya sabíamos lo que significaban ambas opciones.

[8] Esto excluye a las dudas que tengan su origen en un escepticismo sistemático.

Un ejemplo más sofisticado: En el capítulo VII definiremos una propiedad de números naturales a la que de momento podemos llamar "ser simpático".[9] Existe un procedimiento para saber si un número dado es simpático o no, exactamente de la misma naturaleza que el que nos permite saber si es primo o no. Pero suceden los siguientes hechos:

a) No es posible probar que todo natural es simpático.

b) Hasta la fecha nadie ha encontrado un natural antipático y es muy dudoso que exista alguno.

Tiene sentido afirmar que todo natural es simpático. Significa que 0 es simpático, 1 es simpático, 2 es simpático ... etc. o sea, que por mucho que uno avance en el examen de números más y más grandes nunca se encuentra una excepción.

La afirmación "Todos los naturales son simpáticos" es metamatemáticamente aceptable porque tiene sentido decir que es verdadera o falsa independientemente de lo que podría hacerse por justificarla (lo que, según lo dicho, es imposible). No sabemos si es verdadera o falsa, pero sabemos lo que es que sea verdadera o falsa.

El concepto de "número simpático" es finitista, pues comprobar si un número es o no simpático se reduce a un número finito de cálculos. No obstante, podemos definir también un número "supersimpático" como un número tal que todos los números mayores que él son simpáticos. Esta noción ya no es finitista. De hecho no tenemos manera de saber si 3 es supersimpático o no, pero lo importante es que tiene sentido: o lo es o no lo es, o hay un número antipático mayor que 3 o no lo hay.

Pensemos ahora en el conjunto A de todos los conjuntos cuyos elementos son números naturales. No podemos asignar un contenido metamatemático a esta definición. Una vez más nos encontramos con el mismo fenómeno: sabemos lo que es el conjunto de los números pares, el de los números primos, el de las potencias de dos, e infinitos más, pero no tenemos ninguna definición precisa de lo que es un conjunto de números naturales en abstracto, ni tenemos, en particular, representación alguna de *la totalidad* de tales conjuntos. Todas las contradicciones de la teoría de conjuntos surgen de la pretensión de hablar de colecciones de objetos en sentido abstracto como si supiéramos de qué estamos hablando.

Quizá el lector crea tener una representación intuitiva del conjunto A, pero deberá reconsiderarlo ante los hechos: los axiomas de la teoría de conjuntos contienen todo lo que los matemáticos saben decir sobre su presunta intuición de los conjuntos abstractos. En particular, de ellos se deducen muchas propiedades de A, tales como que no es numerable. Sin embargo, quedan muchas afirmaciones sobre A que no pueden ser demostradas o refutadas. La más famosa es la hipótesis del continuo: ¿Existe un conjunto infinito $B \subset A$ tal que B no pueda biyectarse con el conjunto de los números naturales y tampoco con A? Si el

[9] Se trata de "no ser el número de Gödel de la demostración de una contradicción en ZFC.

conjunto A tuviera un contenido intuitivo preciso, esta afirmación tendría que ser verdadera o falsa. Ahora bien, veremos que es posible construir modelos de la teoría de conjuntos, es decir, podemos encontrar unos objetos a los que, si los llamamos "conjuntos" satisfacen todos los axiomas que los matemáticos postulan sobre los conjuntos, de modo que la hipótesis del continuo, interpretada como una afirmación sobre estos objetos, resulta ser verdadera, mientras que es posible hacer lo mismo con otra interpretación distinta de la noción de "conjunto" y de tal modo que la hipótesis del continuo resulta ser falsa. Más precisamente, interpretando de formas distintas esa noción de "conjunto" dentro del margen de libertad que nos concede el hecho de que los axiomas de la teoría de conjuntos no la determinan por completo, podemos construir dos objetos A_1 y A_2, ambos con el mismo derecho a ser llamados "la totalidad de los conjuntos de números naturales" (de acuerdo con distintas nociones de "conjunto") y de modo que una cumpla la hipótesis del continuo y la otra no. ¿Cómo se puede digerir esto?

Sólo hay una posibilidad: reconocer que nuestro conocimiento de la noción de "conjunto" es impreciso. Sólo sabemos que los conjuntos han de cumplir unas propiedades básicas, pero existen distintas interpretaciones posibles de la palabra "conjunto" que hacen que esas condiciones básicas sean satisfechas. Cuando decimos que A no tiene un significado metamatemático preciso no queremos decir que A no signifique nada en absoluto, sino más bien que puede significar infinitas cosas distintas y no somos capaces de precisar a cuál de todas queremos referirnos. Por ello nuestra única posibilidad para hablar de A sin caer en vaguedades o contradicciones es postular unos axiomas que recojan lo que estamos suponiendo que cumplen los conjuntos y, a partir de ahí, podremos trabajar con seguridad.

Éste es el origen de todos los temores y recelos del formalista radical. Esta clase de fenómenos son los que —en ciertas situaciones— hacen imposible razonar cabalmente sin el apoyo de una teoría formal. Pero si queremos fundamentar los razonamientos sobre conjuntos abstractos y entenderlos mejor, hemos de empezar por comprender que los problemas están limitados a este terreno: al de los conjuntos abstractos, pues sólo así comprenderemos que es posible una metamatemática basada no en la forma, sino en el contenido de las afirmaciones que involucra.

Este punto de vista nos permite ir un poco más lejos que el finitismo estricto. Así, por ejemplo, ya hemos visto que la afirmación 3 es supersimpático no es finitista pero sí es aceptable. Notemos que involucra un infinito real, en el sentido de que, aunque aparentemente sea una afirmación sobre el número 3, en realidad es una afirmación sobre la totalidad de los números naturales, no sobre una cantidad finita de ellos. Es posible definir una propiedad más débil que la de ser simpático y supersimpático[10] de modo que, en este nuevo sentido, sí pueda probarse que 3 es supersimpático, y sin que esto deje de ser una afirmación sobre la totalidad de los números naturales. La prueba es un argumento que nos convence de que jamás encontraremos un número natural que no sea (débilmente) simpático e involucra esencialmente a los números naturales

[10]Por ejemplo, sin más que sustituir ZFC por la aritmética de primer orden.

como conjunto infinito. De todos modos, los argumentos no finitistas aparecerán en muy contadas ocasiones en la teoría, bien sea porque no aparezcan en sentido estricto, bien porque con pequeñas modificaciones técnicas podrían eliminarse sin dificultad.

Platonismo En contra de lo que podría parecer, nada de lo que acabamos de discutir pretende negar la posibilidad de que sí exista, después de todo, una noción objetiva de "conjunto" en sentido abstracto. Los matemáticos que creen que así es se llaman "realistas" o "platónicos". No intentaré defender una postura que no comparto, pero sí es importante señalar que nada en este libro contradice el platonismo. Lo único que debemos tener presente es que, si existe una interpretación natural de la teoría de conjuntos, la única forma que tenemos de acercarnos a ella con seguridad y rigor es a través de una sucesión de sistemas axiomáticos que vayan incorporando cada vez más axiomas para cubrir los agujeros de los sistemas anteriores, pero nunca metamatemáticamente. El problema, entonces, es decidir cuál de las dos alternativas a que da lugar una afirmación indecidible en un sistema axiomático es la verdadera en esa pretendida interpretación natural de la teoría. Así, si se concluye que la hipótesis del continuo debe ser verdadera tendremos que añadirla como un nuevo axioma y entender que los resultados que se demuestran con la negación de la hipótesis del continuo tratan sobre unos objetos artificiales que no son los conjuntos en el sentido usual. Naturalmente también podría darse el caso contrario y el problema es la falta de criterios para distinguir lo verdadero de lo falso a este nivel.

Capítulo I

Lenguajes formales de primer orden

Nuestro objetivo a medio plazo es hacernos con una definición de demostración matemática precisa y rigurosa, que nos permita manipular con seguridad los conceptos de la matemática abstracta. Si observamos lo que hace un matemático cuando demuestra, vemos que no es sino escribir ordenadamente una afirmación tras otra, por lo que una demostración será una sucesión de afirmaciones. Estas afirmaciones las hace cada matemático en su propia lengua, ya sea en castellano, francés, inglés, alemán o japonés, pero sucede que estos lenguajes son demasiado complejos para ser analizados fructíferamente a nivel teórico. Por ello en primer lugar hemos de construirnos un lenguaje apropiado para nuestro propósito, es decir, un lenguaje que, por una parte esté despojado de relativos, indefinidos, subjuntivos y tantas cosas que tanto enriquecen nuestra lengua, pero que tanto la complican, y que, al mismo tiempo, siga siendo capaz de expresar todo lo que un matemático necesita. Dedicamos este primer tema a presentar y estudiar una familia de lenguajes con estas características.

1.1 Introducción a los lenguajes formales

Ante la posibilidad de que el lector —aun si tiene conocimientos matemáticos— no esté familiarizado con los conceptos básicos que hemos de manejar, vamos a introducirlos aquí de forma poco rigurosa pero más didáctica que en la exposición definitiva que tendrá lugar después. Esta sección no tiene, pues, más objeto que la de familiarizar al lector con las ideas básicas que vamos a manejar. Nada de lo dicho aquí será usado luego. Quien descubra contradicciones entre algo dicho aquí y algo dicho más adelante, que se quede, por supuesto, con lo dicho luego y que piense si no ha sido mejor tener primero una idea equivocada pero clara y después correcta y clara que no tener siempre una idea correcta e ininteligible.

Por razones que sería difícil justificar ahora, resulta conveniente construir

lenguajes para hablar no sólo de lo que ocupa a los matemáticos, sino de cualquier cosa. Construyamos por ejemplo un lenguaje para hablar de todas las personas que habitan la Tierra.

- En primer lugar será conveniente tener nombres para algunas de estas personas. Por ejemplo "p" puede nombrar a Pedro, "j" puede nombrar a Juan, "a" a Ana y "m" a María. A estos signos que usaremos para nombrar los objetos de los que queremos hablar los llamaremos *constantes.* Así, "p", "j", "a" y "m" son constantes de nuestro lenguaje. Valiéndonos del hecho de que sobre la Tierra hay un número finito de personas, podríamos tomar una constante para nombrar a cada una de ellas, pero no es obligatorio hacerlo, podemos, si queremos, quedarnos con estas únicas cuatro constantes.

  Los matemáticos usan constantes como "0", "1", "2", "$\mathbb{N}$", "$\mathbb{R}$", "π" entre otras muchas.

- Podemos ahora tomar signos que expresen hechos, equivalentes a los verbos en las lenguas naturales. Los llamaremos *relatores.* Un relator podría ser "H", que signifique "ser un hombre", de manera que "Hp" significa "Pedro es un hombre". Pongamos que "A" significa "ser amigos", de manera que "Apm" significa "Pedro y María son amigos".

  Diremos que "H" es un relator *monádico* o de rango 1, mientras que "A" es un relator *diádico* o de rango 2. El *rango* de un relator es el número de complementos que necesita para tener sentido. Por supuesto podemos tomar cuantos relatores queramos de cualquier rango no nulo.

  Por conveniencia no vamos a admitir relatores de rango variable. Uno podría pensar que "A" puede usarse con cualquier número de complementos de manera que "Apm" significa "Pedro y María son amigos", "$Apma$" significa "Pedro, María y Ana son amigos", etc. No aceptaremos esto, sino que cada relator tendrá un rango fijo y así, si decidimos que "A" es diádico, convendremos en que "$Apma$" no tiene sentido. La razón es que esto nos evitará complicaciones técnicas y, de todos modos, nuestro lenguaje no pierde capacidad expresiva. En este caso concreto, el intento de afirmación "$Apma$" puede expresarse correctamente usando varias veces el relator "A" como es debido.

  Exigiremos que todo lenguaje tenga un relator diádico que representaremos "$=$" y al que llamaremos *igualador,* cuyo significado será "ser igual" en el sentido de ser una misma cosa. En lugar de escribir "$= pa$" escribiremos "$p = a$", que significa "Pedro es Ana" (afirmación falsa, pero bien escrita).

  Los matemáticos usan muchos relatores, como "$=$", "$\in$", "$\subset$", "$\leq$", etc.

- A partir de unas afirmaciones podemos construir otras más complejas usando para ello los llamados *conectores lógicos.* Son cinco:

  1 El más sencillo es el negador "$\neg$" (léase "no"). Si "Hp" significa "Pedro es un hombre", "$\neg Hp$" significa "Pedro no es un hombre".

En general, si α es una afirmación verdadera, $\neg\alpha$ significa justo lo contrario y, por tanto, es falsa, y viceversa.

2 Otro conector es el *conjuntor* o *conjunción* "$\wedge$" (léase "y"). Así, "$Hp \wedge \neg Ha$" significa "Pedro es un hombre y Ana no es un hombre", es decir, si α y β son dos afirmaciones, $\alpha \wedge \beta$ es la afirmación que afirma lo que afirma α y lo que afirma β. El signo "$\wedge$" se comporta en nuestro lenguaje exactamente igual como se comporta en castellano la conjunción "y".

3 Si el conjuntor es "y", el *disyuntor* o *disyunción* es "o", y lo representaremos por "$\vee$". En castellano hay dos formas de usar la conjunción "o". Cuando a Juanito le dice su papá: "Para tu cumpleaños te puedo regalar la bicicleta o el balón de fútbol", no vale que Juanito responda: "Bien, regálamelos", porque lo que su padre quiere es que elija. Aquí "o" significa "lo uno o lo otro, pero no las dos cosas". Pero cuando a Juanito le dice la abuelita: "Es hora de merendar, come galletas o bizcochos", esta vez Juanito no tiene que elegir, y su abuelita se pondrá muy contenta si come de todo y se hace muy mayor. Aquí "o" significa "lo uno o lo otro, o también las dos cosas". Pues bien, para nosotros "$\vee$" significará siempre esto último. "$Hp \vee Ha$" significa "Pedro es un hombre o Ana es un hombre", lo cual es cierto porque Pedro es un hombre. Pero si digo "$Hp \vee Hj$", sigo estando en lo cierto porque, en sentido no exclusivo —como la abuelita—, Pedro es un hombre o Juan es un hombre: ambos lo son. Todo esto puede resumirse en las tablas siguientes:

α	β	$\alpha \wedge \beta$	$\alpha \vee \beta$
V	V	V	V
V	F	F	V
F	V	F	V
F	F	F	F

Así, si las afirmaciones α y β son ambas verdaderas, entonces $\alpha \wedge \beta$ también lo es, al igual que $\alpha \vee \beta$; si α es verdadera y β falsa $\alpha \wedge \beta$ es falsa, mientras que $\alpha \vee \beta$ es verdadera etc.

Naturalmente también el negador tiene su tabla, más sencilla, puesto que sólo depende de un argumento:

α	$\neg\alpha$
V	F
F	V

4 El siguiente conector es el más polémico. Se llama *implicador* y lo representaremos "$\rightarrow$" (léase "implica"). La idea es que $\alpha \rightarrow \beta$ ha de significar "si α, entonces β", pero esto puede entenderse de varias maneras. El problema se remonta al siglo III a.C., cuando ya los estoicos analizaban este tipo de enunciados y crearon una lógica mucho

mejor que la de Aristóteles pero que, al no estar respaldada por una reputación como la del estagirita, quedó en el olvido.

La controversia sobre cómo interpretar los enunciados $\alpha \to \beta$ fue tal, que Calímaco llegó a decir: "Hasta los cuervos discuten en los tejados este problema". Filón consideraba que $\alpha \to \beta$ era verdadero a no ser que α fuera verdadero pero β falso. Así "Si Pedro es un hombre entonces Ana es una mujer" o "Si Pedro es una mujer entonces Ana es un hombre" son verdaderas aunque el sexo de Pedro poco influya sobre el sexo de Ana. La segunda afirmación es verdadera porque sólo habla de lo que ocurre si Pedro es una mujer, sin decir nada del caso en que Pedro sea un hombre, como de hecho ocurre. En cambio, "Si Pedro es un hombre entonces Ana es un hombre" es falsa, pues Pedro es un hombre y no se cumple lo que, según la frase, debería ocurrir en tal caso, es decir, que Ana sea un hombre.

Por otra parte, Diodoro decía que $\alpha \to \beta$ es verdadero si siempre que α sea verdadero β también lo es. Por ejemplo "Si es verano el cielo está nublado" sería verdadero para Filón en un día de invierno cualquiera, mientras que para Diodoro sería falsa, ya que es posible encontrar días de verano en los que haga sol, es decir, días en los que ocurre α pero no por ello sucede β. La interpretación de "Si ... entonces ..." en las lenguas naturales está más próxima a la de Diodoro que a la de Filón, pues suele depender de relaciones causales y no meramente lógicas. La interpretación de Filón se da ocasionalmente cuando decimos: "Si apruebo este examen lloverán lechugas", que es tanto como decir: "No voy a aprobar este examen". Pero en matemáticas es más práctica la implicación de Filón y así, $\alpha \to \beta$ deberá interpretarse de acuerdo con la tabla siguiente:

α	β	$\alpha \to \beta$
V	V	V
V	F	F
F	V	V
F	F	V

En definitiva, para que $\alpha \to \beta$ sea verdadera ha de ocurrir o bien que α sea verdadera, y en este caso β también ha de serlo, o bien que α sea falsa y entonces da igual lo que le ocurra a β. Notemos por tanto:

- *Una afirmación falsa implica cualquier afirmación:* Si α es falsa, $\alpha \to \beta$ es verdadera, cualquiera que sea β.
- *Una afirmación verdadera es implicada por cualquier afirmación:* Si β es verdadera, $\alpha \to \beta$ es verdadera, cualquiera que sea α.

5 Finalmente tenemos el coimplicador "$\leftrightarrow$" ("si y sólo si") que indica que α y β son ambas verdaderas o ambas falsas, que lo que vale para una, vale para la otra. He aquí su tabla:

α	β	$\alpha \leftrightarrow \beta$
V	V	V
V	F	F
F	V	F
F	F	V

- Los signos que dan mayor fuerza expresiva a los lenguajes formales son los *cuantificadores,* que se usan juntamente con las *variables.*

  Las variables las representaremos con letras cualesquiera x, y, z, u, v, w principalmente. Nombran a objetos indeterminados. Esto es, desde luego, ambiguo, pero lo entenderemos al considerar los cuantificadores. Tomemos por ejemplo el *cuantificador existencial* o *particularizador* "$\bigvee$" (léase "existe"). "$\bigvee xHx$" significa "Existe un x de manera que x es un hombre" o, más brevemente, "Existe un hombre". Aquí, "x" es una variable. En general "$\bigvee x$ algo" significa que "algo" es cierto si la variable "x" se interpreta adecuadamente.

  Otro ejemplo: "$\bigvee xAxa$" significa "Existe un x de manera que x es amigo de Ana", o sea, "Ana tiene un amigo".

  El *cuantificador universal* "$\bigwedge$" (léase "para todo") se usa como sigue: "$\bigwedge x$ algo" significa que "algo" es verdadero se interprete como se interprete la variable "x". Por ejemplo "$\bigwedge xHx$" significa "Para todo x, x es un hombre", o sea, "Sólo hay hombres", lo cual es evidentemente falso. "$\bigwedge x(Axa \rightarrow Axp)$" significa "Para todo x, si x es amigo de Ana entonces x es amigo de Pedro", o sea, "Pedro es amigo de todos los amigos de Ana". Más ejemplos:

Pedro es el único amigo de Ana:	$\bigwedge x(Axa \leftrightarrow x = p)$,
Ana sólo tiene un amigo:	$\bigvee y \bigwedge x(Axa \leftrightarrow x = y)$,
Conmutatividad de la amistad:	$\bigwedge xy(Axy \leftrightarrow Ayx)$,
Ana no es amiga de todos los hombres:	$\bigvee x(Hx \wedge \neg Axa)$,
Ana no es amiga de ningún hombre:	$\bigwedge x(Hx \rightarrow \neg Axa)$,
Ningún hombre es mujer:	$\bigwedge x(Hx \rightarrow \neg Mx)$.

  (El matemático que tenga cierta familiaridad con los cuantificadores debería, no obstante, fijarse en el orden de los signos y en el uso de los paréntesis).

- Otros signos que pueden simplificar la escritura son los funtores. Si queremos decir que los padres de los italianos son italianos podemos usar un relator monádico "I" que signifique "ser italiano" y un relator diádico "P" tal que "Pxy" signifique "x es el padre de y", y construir la afirmación "$\bigwedge xy(Iy \wedge Pxy \rightarrow Ix)$", pero otra alternativa es tomar un funtor monádico "f" de manera que "fx" signifique "el padre de x" y escribir simplemente "$\bigwedge x(Ix \rightarrow Ifx)$".

  Un *funtor* es un signo que, completado con nombres de objetos, nombra a otro objeto, a diferencia de un relator que da lugar a una afirmación.

Un ejemplo de funtor diádico podría ser un "M" que signifique "el mayor de edad de", así "Mpj" es el mayor de edad entre Pedro y Juan. Naturalmente pueden construirse funtores de cualquier rango. Ejemplos de funtores en matemáticas son "$\cup$", "$+$", etc.

- Para acabar nos ocupamos del último signo con que proveeremos a nuestros lenguajes formales: el *descriptor* "$|$" (léase "tal que"). Supongamos que Ana tiene un solo amigo. Entonces "$x \mid Axa$" significa "el x tal que x es amigo de Ana", o sea, "el amigo de Ana". Una expresión de este tipo se llama una *descripción.* En general "x |algo" significa el único x que cumple "algo".

  Es fácil darse cuenta de que esto nos va a crear problemas. Bertrand Russell dijo una vez: "El actual rey de Francia es calvo". Podemos formalizar esta afirmación con un relator monádico "C" que signifique "ser calvo" y otro "F" que signifique "ser rey de Francia ahora". Así nos queda "$C(x \mid Fx)$". Pero sucede que no hay rey en Francia: "$\bigwedge x \neg Fx$". ¿Cómo debemos interpretar esta descripción? ¿El actual rey de Francia es o no calvo? Los matemáticos no se ven libres de este problema: ¿Cuánto vale $\lim\limits_{n}(-1)^n$? Por supuesto $\lim\limits_{n}(-1)^n = x \mid \big((-1)^n \xrightarrow{n} x\big)$.

  El problema es que hemos de encontrar un tratamiento coherente para las descripciones, pues si, por ejemplo, dijéramos que "$C(x \mid Fx)$" es falso, estaríamos obligados a aceptar que "$\neg C(x \mid Fx)$" es verdadero, o sea, que el actual rey de Francia no es calvo, y estaríamos en las mismas.

  Hay dos salidas posibles. Una muy drástica —pero muy natural— consiste en prohibir que se escriban descripciones impropias, o sea, descripciones de la forma "x |algo" donde no hay un único x que cumpla "algo". De este modo "$C(x \mid Fx)$" no es ni verdadera ni falsa porque no es una afirmación.

  Pero hay otra posibilidad más artificial pero más cómoda y que es la que vamos a adoptar: Escójase una "víctima" entre los objetos de los que hablamos, llamémosla descripción impropia o, para abreviar, d. Si estamos hablando de personas, d puede ser una cualquiera, por ejemplo, Julio. Un matemático puede tomar como descripción impropia al conjunto vacío. Al encontrarnos con una descripción "x |algo" pueden pasar dos cosas:

  a) que exista un único x que cumpla "algo". Entonces la descripción se llama propia y convenimos en que "x |algo" representa a ese único x que cumple "algo".

  b) que no exista un único x que cumpla "algo", bien porque no haya ninguno, bien porque haya varios. Entonces la descripción se llama impropia y convenimos en que "x |algo" significa d.

  Según este criterio, el actual rey de Francia es Julio y será calvo o no según si Julio lo es. El hijo de la reina de Inglaterra también es Julio, esta vez porque tiene varios hijos. Este convenio no crea problemas si tenemos siempre en cuenta lo siguiente:

a) Para las descripciones propias: Si existe un único x que cumple "algo" entonces "x | algo" cumple "algo".

b) Para las descripciones impropias: Si no existe un único x que cumpla "algo" entonces "x | algo" es d y, por tanto, no tiene por qué cumplir "algo".

Por ejemplo, el hijo de la reina de Inglaterra no es hijo de la reina de Inglaterra, es Julio. El único cuidado que hemos de tener es el de no ser ingenuos y pensar que todas las descripciones que vemos son propias. Notemos que en cualquier caso "$x \mid x = x$" siempre es una descripción impropia, luego podemos decir que si la descripción "x | algo" es impropia, entonces $(x \mid \text{algo}) = (x \mid x = x)$, y así no hacemos referencia explícita a d, y no es necesario, si no queremos, precisar de quién se trata.

1.2 Definición de lenguaje formal

Pasamos ahora a introducir rigurosamente los conceptos que hemos discutido informalmente en la sección anterior. Para ello hemos de cambiar completamente el planteamiento. Si algo hace útiles a los lenguajes formales es precisamente el hecho de que son formales, es decir, que podemos trabajar con ellos sin necesidad de aludir en ningún momento al significado de sus signos. Así, no podemos permitirnos el lujo de definir las constantes como signos que nombran objetos, pues eso nos obligaría a tener una idea clara de los objetos que nombran las constantes, y eso es precisamente lo que queremos evitar. Queremos un lenguaje que nos permita hablar del conjunto vacío $\varnothing$ sin comprometernos a responder a la pregunta de qué objeto es nombrado por $\varnothing$. Veamos la definición correcta:

Definición 1.1 Un *lenguaje formal de primer orden*[1] $\mathfrak{L}$ es una colección de signos divididos en las categorías siguientes y de modo que cumplan las propiedades que se indican:

Variables Un lenguaje $\mathfrak{L}$ debe tener infinitas variables. Cada variable debe tener asociado un número natural distinto al que llamaremos su *índice,* de tal forma que todo natural es índice de una variable de $\mathfrak{L}$. Llamaremos x_i a la variable de índice i de $\mathfrak{L}$.

Constantes Un lenguaje $\mathfrak{L}$ puede tener cualquier cantidad de constantes, desde ninguna hasta infinitas. En cualquier caso, cada constante debe tener asociado un *índice* natural. Llamaremos c_i a la constante de $\mathfrak{L}$ de índice i (si existe) de modo que si $\mathfrak{L}$ tiene $m+1$ constantes éstas serán $c_0, c_1, \ldots, c_m$, mientras que si $\mathfrak{L}$ tiene infinitas constantes, los índices recorrerán todos los números naturales.

[1] El lector que quiera saber qué significa exactamente "de primer orden" encontrará la respuesta en la sección 10.3.

Relatores Cada relator debe tener asociado un número natural no nulo al que llamaremos su *rango.* Llamaremos relatores n-ádicos a los relatores de rango n. El número de relatores n-ádicos de $\mathfrak{L}$ puede variar entre ninguno e infinitos. Cada relator n-ádico debe llevar asociado un *índice* distinto. Llamaremos R^n_i al relator n-ádico de índice i de $\mathfrak{L}$ (si existe). Así, si $\mathfrak{L}$ tiene $m+1$ relatores n-ádicos, éstos serán $R^n_0, \ldots, R^n_m$.

Todo lenguaje formal debe tener al menos el relator diádico R^2_0, al que llamaremos *igualador* o $=$.

Funtores Cada funtor ha de llevar asociado un *rango* y un *índice* en las mismas condiciones que los relatores. Llamaremos f^n_i al funtor n-ádico de índice i de $\mathfrak{L}$ (si existe).

Negador Llamaremos $\neg$ al negador de $\mathfrak{L}$.

Implicador Llamaremos $\rightarrow$ al implicador de $\mathfrak{L}$.

Cuantificador universal (o generalizador) Lo llamaremos $\bigwedge$.

Descriptor Un lenguaje formal $\mathfrak{L}$ puede tener o no descriptor y, según el caso, diremos que $\mathfrak{L}$ es un lenguaje *con o sin descriptor.* Si existe lo representaremos por $|$.

Cada signo de $\mathfrak{L}$ debe pertenecer a una de estas categorías y sólo a una. Las constantes, los funtores y los relatores distintos del igualador se llaman *signos eventuales* de $\mathfrak{L}$, mientras que los restantes son *signos obligatorios.*

Si $\mathfrak{L}$ es un lenguaje formal con descriptor, llamaremos $\underline{\mathfrak{L}}$ al lenguaje que resulta de eliminarle el descriptor.

Una definición formal Tal y como queríamos, en esta definición no se hace referencia al significado o al uso pretendido de los conceptos mencionados. En ningún momento se dice qué es una constante o un relator. La situación es análoga a la que se daría al describir, por ejemplo, el ajedrez. Si queremos explicarle a alguien cómo se juega al ajedrez, le diremos que se necesita un tablero y 32 piezas divididas en dos grupos de colores de forma que, a una de cada color —no importa cuál— hay que ponerle la etiqueta de rey, a otras la de alfil etc. Cualquier cosa razonable puede hacer el papel de rey en una partida de ajedrez. Lo que convierte en rey a una pieza no es ninguna característica propia, sino tan sólo el convenio que los jugadores adoptan de usarla como rey. Igualmente, para construir un lenguaje formal necesitamos unos cuantos signos, no importa cuáles, y a cada uno le ponemos una etiqueta. Igual que en ajedrez moveremos de forma distinta cada pieza según su nombre, también nosotros usaremos de forma distinta las constantes y los relatores, pero eso vendrá luego.

Conectores y cuantificadores Observemos que sólo hemos exigido dos conectores y un cuantificador. Los restantes los definiremos a partir de éstos.

Ejemplo Para comprobar que la definición no es ambigua, vamos a definir un lenguaje concreto, llamémoslo $\mathcal{L}_0$. Necesitamos signos. No hemos definido "signo", pero tampoco es necesaria una definición de "pieza" para jugar al ajedrez.

- Para las variables necesitamos infinitos signos. Nos sirven, por ejemplo

  $$|, \quad |-, \quad |--, \quad |---, \quad |----, \quad |-----, \quad |------, \quad \cdots$$

  Es decir, la variable x_0 es una raya vertical y, en general, si $i > 0$, la variable x_i es una raya vertical seguida de i rayas horizontales. Con esto quedan perfectamente definidas las variables de $\mathcal{L}_0$. No hay duda de que tenemos infinitas de ellas. Más concretamente, para cada natural i, sabemos perfectamente cuál es la única variable x_i de $\mathcal{L}_0$ a la que le corresponde el índice i.

- El lenguaje $\mathcal{L}_0$ tendrá tres constantes, digamos "$\triangle$", "$\Diamond$", y "$\square$", de índices 0, 1 y 2, respectivamente.

- Tomamos dos relatores, uno monádico "$\blacktriangle$" y el igualador "$\smile$".

- $\mathcal{L}_0$ no tiene funtores.

- Negador: "$\boxtimes$".

- Implicador: "$\triangleleft$".

- Cuantificador universal: "$\bigstar$".

- Descriptor: "$\between$".

Con esto, el lenguaje $\mathcal{L}_0$ queda completamente especificado. Si nos preguntan quién es R_0^1 en $\mathcal{L}_0$ la respuesta es clara: el signo "$\blacktriangle$", y si nos preguntan por R_3^5 hemos de contestar que no existe tal signo en $\mathcal{L}_0$. ∎

Signos y nombres de signos Es importante distinguir entre los signos de $\mathcal{L}_0$ y sus nombres. Debemos tener presente que estas páginas están escritas en castellano, extendido con algunos signos adicionales que, a efectos teóricos, deben considerarse al mismo nivel que cualquier palabra castellana. Así por ejemplo, "c_0" es una palabra castellana que nombra a c_0, que es un signo de $\mathcal{L}_0$, a saber, el signo "$\triangle$". Notar el papel crucial que desempeñan las comillas para que esto sea inteligible. Si decimos "sea x una variable de $\mathcal{L}_0$", aquí la letra "x" no es una variable de $\mathcal{L}_0$, sino que es una palabra castellana, un pronombre indefinido que representa a una cualquiera de las variables de $\mathcal{L}_0$, sea "$|$", o "$|-$", o "$|--$", o cualquier otra.

Infinitos signos La técnica que hemos seguido para definir las variables de $\mathcal{L}_0$ puede refinarse para mostrar que, en realidad, en lugar de trabajar con lenguajes con infinitos signos podríamos trabajar con lenguajes de sólo dos signos, pues con ellos podríamos formar infinitas palabras que desempeñaran el papel que nosotros asignaremos a los signos de nuestros lenguajes. Esto únicamente complicaría técnicamente la exposición y en ningún caso puede considerarse una forma de "evitar" un infinito. Tan sólo de retrasar su aparición. Da igual trabajar con infinitos signos o con infinitas palabras formadas por combinaciones de dos signos. Lo único que cambia es si llamamos signos a trazos "indivisibles" en algún sentido o si llamamos signos a combinaciones finitas de trazos.

1.3 Expresiones, términos y fórmulas

La finalidad primera de los lenguajes formales es, por supuesto, construir afirmaciones con sus signos. Empezaremos con algunas consideraciones generales sobre las sucesiones de signos.

Definición 1.2 Sea $\mathcal{L}$ un lenguaje formal. Una *cadena de signos* de $\mathcal{L}$ es una sucesión finita de signos de $\mathcal{L}$ repetidos o no y en un cierto orden. Si $\zeta_1, \ldots, \zeta_n$ son cadenas de signos de $\mathcal{L}$ llamaremos $\zeta_1 \cdots \zeta_n$ a la cadena que resulta de yuxtaponer las cadenas $\zeta_1, \ldots, \zeta_n$ en este orden. En particular podemos nombrar una cadena nombrando a cada uno de sus signos en el orden en que aparecen.

Dos cadenas de signos ζ_1 y ζ_2 son *idénticas* si constan de los mismos signos en el mismo orden. Lo indicaremos así: $\zeta_1 \equiv \zeta_2$ (y en caso contrario escribiremos $\zeta_1 \not\equiv \zeta_2$), es decir, $\zeta_1 \equiv \zeta_2$ significa que"ζ_1"y "ζ_2"son dos nombres para la misma cadena de signos.

Si ζ es una cadena de signos de $\mathcal{L}$, llamaremos *longitud* de ζ al número de signos que componen ζ, contando cada uno tantas veces como se repita.

Claramente, si $\mathcal{L}$ es un lenguaje formal con descriptor, toda cadena de signos de $\underline{\mathcal{L}}$ lo es de $\mathcal{L}$ y toda cadena sin descriptores de $\mathcal{L}$ lo es de $\underline{\mathcal{L}}$.

Ejemplo Una cadena de signos de $\mathcal{L}_0$ es "$\boxtimes|--||\square\lhd\lhd\lhd$" Su longitud es 8 (¡recordemos que "$|--$" es un solo signo!). De acuerdo con el convenio que hemos adoptado, podemos referirnos a ella con el nombre "$\neg x_2x_0x_0c_2 \to\to\to$". De este modo, podemos decir que "$\neg x_2x_0x_0c_2 \to\to\to$" es una palabra castellana que nombra a una cadena de signos de $\mathcal{L}_0$ y también que $\neg x_2x_0x_0c_2 \to\to\to$ es una cadena de signos de $\mathcal{L}_0$ (¡atención a las comillas!). ∎

Naturalmente, las cadenas de signos del estilo de la que acabamos de considerar no sirven para nada. Ahora hemos de extraer de entre ellas las que tienen "significado", algo así como seleccionar los movimientos "legales" en el ajedrez de entre todos los movimientos posibles. Hay dos casos distintos en los que una cadena de signos puede tener un significado: bien porque nombre a un objeto, bien porque afirme algo. A las cadenas que nombran objetos las llamaremos términos, mientras que a las que afirman algo las llamaremos fórmulas. Ésta

es la idea subyacente, pero no nos sirve como definición porque, además de ser imprecisa, alude a un posible significado de las cadenas de signos, y queremos que la definición sea formal. La definición correcta es la siguiente:

Definición 1.3 Una cadena de signos de un lenguaje formal $\mathcal{L}$ es un *término* o una *fórmula* de $\mathcal{L}$ si puede probarse que lo es a partir de las reglas siguientes (t_i representa a un término, α, β, ... representan fórmulas):

a) x_i es un término.
b) c_i es un término.
c) $R_i^n t_1 \cdots t_n$ es una fórmula.
d) $f_i^n t_1 \cdots t_n$ es un término.
e) $\neg\alpha$ es una fórmula.
f) $\rightarrow\alpha\beta$ es una fórmula.
g) $\bigwedge x_i\, \alpha$ es una fórmula.
h) $\mid x_i\alpha$ es un término (si es que $\mathcal{L}$ tiene descriptor).

En esta definición hemos adoptado un convenio que será útil en lo sucesivo para evitar aclaraciones prolijas. Por ejemplo, la primera condición a) ha de leerse "toda variable es un término", es decir, sobrentendemos que "x_i" hace referencia a una variable arbitraria del lenguaje en cuestión. Similarmente "c_i" hace referencia a una constante arbitraria (supuesto que existan), de modo que la segunda condición es "si $\mathcal{L}$ tiene constantes c_i, éstas son términos", la tercera es "la cadena que resulta de yuxtaponer un relator n-ádico R_i^n con n términos es una fórmula", etc.

Diremos que una cadena de signos θ de un lenguaje $\mathcal{L}$ es una *expresión* si es un término o una fórmula de $\mathcal{L}$.

Ejemplo La cadena $\alpha \equiv \bigwedge x_0 {\rightarrow} {=} x_0 c_1 R_0^1 x_0$ es una fórmula de $\mathcal{L}_0$.

En efecto, para comprobarlo observamos en general que cada una de las reglas anteriores se aplica a cadenas de signos que comienzan por un signo característico. Así, la primera regla sólo se aplica a cadenas que empiecen por una variable, la segunda por una constante, la tercera por un relator, etc. Teniendo esto en cuenta, podemos comprobar que α es una fórmula mediante los pasos siguientes:

a) Como el primer signo es el generalizador, para que α sea una expresión ha de satisfacer la regla g), y en tal caso será una fórmula, como queremos probar. Para que así sea, después del generalizador debe haber una variable, y luego una fórmula. Efectivamente, tras el generalizador tenemos la variable x_0, luego α será una fórmula si y sólo si la cadena de signos $\beta \equiv {\rightarrow} {=} x_0 c_1 R_0^1 x_0$ es una fórmula.

b) El primer signo de β es el implicador, luego para que sea una fórmula ha de cumplir la regla f). Esto sucederá si la cadena $\zeta \equiv {=} x_0 c_1 R_0^1 x_0$ es la yuxtaposición de dos fórmulas.

c) Puesto que su primer signo es el igualador (un relator diádico), para que ζ sea la yuxtaposición de dos fórmulas, la primera ha de cumplir la regla c), lo cual exige que tras el igualador haya dos términos. Por consiguiente $\zeta_1 \equiv x_0 c_1 R_0^1 x_0$ ha de ser la yuxtaposición de dos términos y una fórmula.

d) Puesto que su primer signo es la variable x_0, el primer término tiene que venir dado por la regla a), la cual nos dice, efectivamente, que x_0 es por sí solo un término (y no hay otra posibilidad). Por consiguiente $c_1 R_0^1 x_0$ ha de ser un término seguido de una fórmula.

e) La regla b) nos da que c_1 es un término, luego $\gamma \equiv R_0^1 x_0$ ha de ser una fórmula.

f) La única posibilidad es que γ cumpla la regla c), para lo cual x_0 ha de ser un término. Como así es, acabamos de probar que α es una fórmula. ∎

La definición no es ambigua El ejemplo anterior muestra que la definición de expresión que hemos dado es totalmente precisa y rigurosa. En efecto, el procedimiento que hemos seguido para determinar que α es una fórmula puede aplicarse a cualquier cadena de signos de cualquier lenguaje formal. Barriendo la cadena de izquierda a derecha, cada signo nos remite a una de las reglas, la cual nos da un criterio que, o bien no se cumple, y entonces la cadena no es una expresión, o bien se cumple, y entonces nos remite a analizar una cadena más corta. En otras palabras, es fácil diseñar un algoritmo que decide en un tiempo finito si cualquier cadena dada es un término, una fórmula o no es ni lo uno ni lo otro. ∎

Observaciones Toda fórmula comienza por un relator, el negador, el implicador o el cuantificador universal. Todo término comienza por una variable, una constante, un funtor o el descriptor. Por lo tanto una expresión no puede ser a la vez término y fórmula.

Claramente, si un lenguaje $\mathfrak{L}$ tiene descriptor, toda expresión de $\underline{\mathfrak{L}}$ lo es de $\mathfrak{L}$ y toda expresión sin descriptores de $\mathfrak{L}$ lo es de $\underline{\mathfrak{L}}$.

Convenios de notación El lector se habrá extrañado de las reglas f) y h) de la definición de expresión, donde aparece $\to\alpha\beta$ en lugar de $\alpha \to \beta$ y $|x_i\alpha$ en lugar de $x_i|\alpha$. Quizá el ejemplo anterior le ha hecho ver el por qué de este orden peculiar: simplifica enormemente la gramática, en especial porque evita la necesidad de paréntesis. Efectivamente, las fórmulas que con la notación habitual serían $(\alpha \to \beta) \to \gamma$ y $\alpha \to (\beta \to \gamma)$, para nosotros son $\to\to\alpha\beta\gamma$ y $\to\alpha\to\beta\gamma$, respectivamente. No obstante, esto no quiere decir que en lo sucesivo estemos obligados a nombrar las fórmulas de un modo tan incómodo. Al contrario, solventamos este problema introduciendo los siguientes convenios de notación:

a) $(\alpha \to \beta) \equiv \to\alpha\beta$
b) $(x_i|\alpha) \equiv |x_i\alpha$

c) $(t_1 = t_2) \equiv {=}t_1t_2$
d) $(t_1 \neq t_2) \equiv \neg(t_1 = t_2)$
e) $(\alpha \vee \beta) \equiv (\neg\alpha \to \beta)$
f) $(\alpha \wedge \beta) \equiv \neg(\neg\alpha \vee \neg\beta)$
g) $(\alpha \leftrightarrow \beta) \equiv \big((\alpha \to \beta) \wedge (\beta \to \alpha)\big)$
h) $\bigvee x_i\, \alpha \equiv \neg\bigwedge x_i \neg\alpha$
i) $\overset{1}{\bigvee} x_i\, \alpha \equiv \bigvee x_j \bigwedge x_i \big(\alpha \leftrightarrow (x_i = x_j)\big)$, donde j es el menor natural distinto de i tal que x_j no está en α

Las expresiones de la forma

$$\bigwedge x_{i_1} \cdots \bigwedge x_{i_n} \alpha, \quad \bigvee x_{i_1} \cdots \bigvee x_{i_n} \alpha, \quad \overset{1}{\bigvee} x_{i_1} \cdots \overset{1}{\bigvee} x_{i_n} \alpha$$

las abreviaremos a

$$\bigwedge x_{i_1} \cdots x_{i_n} \alpha, \quad \bigvee x_{i_1} \cdots x_{i_n} \alpha, \quad \overset{1}{\bigvee} x_{i_1} \cdots x_{i_n} \alpha$$

respectivamente. También suprimiremos aquellos paréntesis que no sean imprescindibles para interpretar correctamente una expresión. Para ello tendremos en cuenta que el negador "liga" más fuertemente que la conjunción y la disyunción y que éstas a su vez "ligan" más fuertemente que la implicación y la coimplicación. Por ejemplo "$\neg\alpha \wedge \beta \to \gamma \wedge \delta$" es un nombre para la fórmula $(((\neg\alpha) \wedge \beta) \to (\gamma \wedge \delta))$.

Ejemplos La fórmula $\alpha \equiv \bigwedge x_0 {\to} {=} x_0 c_1 R^1_0 x_0$ del ejemplo anterior puede escribirse alternativamente como $\alpha \equiv \bigwedge x_0 (x_0 = c_1 \to R^1_0 x_0)$. Es importante comprender que no hemos dado una definición de fórmula para abandonarla al párrafo siguiente. Tanto "$\bigwedge x_0 {\to} {=} x_0 c_1 R^1_0 x_0$" como "$\bigwedge x_0 (x_0 = c_1 \to R^1_0 x_0)$" son dos nombres para una cadena de signos de $\mathcal{L}_0$ que en ningún momento resulta alterada. Concretamente, se trata de la cadena "$\bigstar|\lhd \smile |\Diamond\blacktriangle|$". Lo que hemos alterado es el criterio para nombrarla: en el primer caso usábamos el más directo, consistente en yuxtaponer los nombres de sus signos, y en el segundo el más sofisticado, en el que empleamos los convenios de notación que hemos adoptado.

La fórmula $\alpha \equiv (x_0 = c_0) \vee (x_0 = c_1)$ tiene longitud 8 y su primer signo no es un paréntesis ni la variable x_0. Su primer signo es el implicador $\to$ o, más explícitamente "$\lhd$". En efecto, según nuestros convenios:

$$\alpha \equiv (x_0 = c_0) \vee (x_0 = c_1) \equiv \neg(x_0 = c_0) \to (x_0 = c_1) \equiv {\to}\neg{=}x_0c_0{=}x_0c_1.$$

En ningún caso va a tener importancia cuál es el orden concreto en que los signos aparecen en una fórmula, pero debemos ser conscientes de la diferencia entre una fórmula y cualquiera de los nombres que le demos. La línea anterior contiene cuatro nombres igualmente legítimos para una misma fórmula (incluyendo a "α" entre ellos). ∎

Los signos metamatemáticos Observemos que "$\neg$", "$\vee$", "$\wedge$" etc. no son signos de ningún lenguaje formal —sin perjuicio de que podamos definir un lenguaje que los contenga—, sino que son signos metamatemáticos, cuyo status es el mismo que el de las palabras castellanas. El primero nombra a un signo de un lenguaje formal dado (distinto en cada lenguaje, igual que "rey blanco" cambia de significado cuando cambiamos de juego de ajedrez), mientras que los otras dos, "$\wedge$" y "$\vee$", ni siquiera nombran a signo alguno. Son palabras castellanas que intervienen en los convenios que nos permiten abreviar los nombres de fórmulas más complejas. La fórmula $\alpha \vee \neg\alpha$ no contiene ningún signo llamado disyuntor, sino que "$\vee$" indica que hablamos de la fórmula $\neg\alpha \to \neg\alpha$.

Similarmente, $\equiv$ es un signo metamatemático, mera abreviatura de la expresión "es idéntico a".

Por supuesto, en todo esto hay muchas decisiones arbitrarias. Podríamos haber definido lenguajes formales con negador y disyuntor, y definir en términos de ambos el implicador. Incluso podríamos haber estipulado que los lenguajes formales tuvieran los cinco conectores y los dos cuantificadores como signos obligatorios, pero todas estas alternativas terminan siendo equivalentes a la que hemos seguido.

En cualquier caso, no hay ningún riesgo de confusión si en determinados contextos hablamos del conjuntor, el disyuntor, etc. de un lenguaje formal, como si decimos que una fórmula de tipo $\bigvee x\alpha$ "empieza por el cuantificador existencial". ∎

La definición de expresión es formal Insistimos en el carácter formal de la definición de expresión que hemos dado. Un matemático lee

$$\bigwedge xy(\bigwedge u(u \in x \leftrightarrow u \in y) \leftrightarrow x = y)$$

y reconoce que se trata de una fórmula porque le encuentra sentido. Significa que dos conjuntos son iguales si y sólo si tienen los mismos elementos. Sin embargo, para nosotros esto es una fórmula porque tenemos un algoritmo basado exclusivamente en la disposición de los signos en la cadena (en el cual podríamos incorporar, si quisiéramos, los convenios de notación que hemos adoptado) que nos dice que esa combinación particular corresponde a una fórmula. Así, para saber que la cadena anterior es una fórmula no necesitamos saber qué es un conjunto. Éste es el punto de partida para garantizar que los matemáticos puedan hablar de conjuntos sin necesidad de definirlos previamente. ∎

1.4 Variables libres y ligadas

Terminaremos el capítulo introduciendo los mínimos conceptos que vamos a necesitar sobre la sintaxis de los lenguajes formales. En esta sección nos ocuparemos de la noción de variable libre y ligada. La idea es que una variable está libre en una expresión si no está afectada por ningún cuantificador ni por el descriptor. En caso contrario está ligada. Por ejemplo, la fórmula $\bigvee y\ x \in y$ (correspondiente al lenguaje de la teoría de conjuntos) contiene libre la variable x

y ligada (por el particularizador) la variable y. Vamos a dar una definición precisa de estas nociones.

Definición 1.4 Sea $\mathcal{L}$ un lenguaje formal. Diremos que una variable x está *libre* en una expresión de $\mathcal{L}$ si se puede probar que lo está a partir de las reglas siguientes:

a) x está libre en x_i syss $x \equiv x_i$.
b) x nunca está libre en c_i.
c) x está libre en $R_i^n t_1 \cdots t_n$ syss lo está en algún t_j.
d) x está libre en $f_i^n t_1 \cdots t_n$ syss lo está en algún t_j.
e) x está libre en $\neg\alpha$ syss lo está en α.
f) x está libre en $\alpha \to \beta$ syss lo está en α o en β.
g) x está libre en $\bigwedge x_i \alpha$ syss lo está en α y $x \not\equiv x_i$.
h) x está libre en $x_i|\alpha$ syss lo está en α y $x \not\equiv x_i$ (en el caso en que $\mathcal{L}$ tenga descriptor).

Diremos que una variable x está *ligada* en una expresión de $\mathcal{L}$ si se puede probar que lo está a partir de las reglas siguientes:

a) x nunca está ligada en x_i.
b) x nunca está ligada en c_i.
c) x está ligada en $R_i^n t_1 \cdots t_n$ syss lo está en algún t_j.
d) x está ligada en $f_i^n t_1 \cdots t_n$ syss lo está en algún t_j.
e) x está ligada en $\neg\alpha$ syss lo está en α.
f) x está ligada en $\alpha \to \beta$ syss lo está en α o en β.
g) x está ligada en $\bigwedge x_i \alpha$ syss lo está en α o $x \equiv x_i$.
h) x está ligada en $x_i|\alpha$ syss lo está en α o $x \equiv x_i$ (en el caso en que $\mathcal{L}$ tenga descriptor).

Observaciones Estas reglas se traducen en un algoritmo que siempre nos permite determinar sin ambigüedades si una variable dada está o no libre o ligada en una expresión dada.

Una variable x está libre o ligada en $\alpha \vee \beta$, $\alpha \wedge \beta$ o $\alpha \leftrightarrow \beta$ syss lo está en α o en β. Así mismo, x está libre en $\bigvee x_i \alpha$ syss lo está en α y $x \not\equiv x_i$, y x está ligada en $\bigvee x_i \alpha$ syss lo está en α o $x \equiv x_i$. Estas observaciones no forman parte de la definición de variable libre y ligada, sino que se deducen inmediatamente de las definiciones de $\alpha \vee \beta$, etc.

Una variable está en una expresión θ (es decir, es uno de los signos que componen θ) si y sólo si está libre o ligada en θ.

Si θ es una expresión sin descriptores de un lenguaje $\mathcal{L}$ con descriptor, entonces una variable x está libre o ligada en una expresión sin descriptores θ considerada como expresión de $\mathcal{L}$ si y sólo lo está considerada como expresión de $\underline{\mathcal{L}}$. ∎

Ejemplos Observemos que una variable puede estar a la vez libre y ligada en una expresión, así como no estar ni libre ni ligada. Los ejemplos siguientes muestran las cuatro posibilidades para una misma variable x:

$u = v$	x no está ni libre ni ligada.
$u = x$	x está libre y no ligada.
$\bigvee x\ u = x$	x está ligada y no libre.
$x \in y \wedge \bigwedge x\ x = x$	x está libre y ligada.

(Suponemos que las variables x, y, u, v son distintas). ∎

Una expresión es *abierta* si tiene variables libres. En caso contrario es *cerrada.* Un *designador* es un término cerrado. Una *sentencia* es una fórmula cerrada. Por lo tanto las cadenas de signos quedan clasificadas como sigue:

- cadenas de signos
  - expresiones
    - términos
      - designadores
      - términos abiertos
    - fórmulas
      - sentencias
      - fórmulas abiertas
  - no expresivas

1.5 Sustitución de variables

Sea θ una expresión de un lenguaje formal $\mathfrak{L}$ en la que esté libre la variable x y sea t un término de $\mathfrak{L}$. Podemos pensar que θ dice algo de x: si θ es una fórmula, θ significa "a x le pasa tal cosa" y si es un término significa "el tal cosa que depende de x". Nuestra intención ahora es construir una nueva expresión de $\mathfrak{L}$ a la que llamaremos sustitución de x por t en θ o, más brevemente $\mathtt{S}^t_x\theta$, de manera que $\mathtt{S}^t_x\theta$ diga de t lo mismo que θ dice de x. Por ejemplo, si $\theta \equiv Hx$ significa "x es un hombre" y $t \equiv p$ significa "Pedro", entonces $\mathtt{S}^t_x Hx$ debe ser Hp, que significa "Pedro es un hombre". Normalmente, sustituir x por t se reduce en la práctica a escribir t allí donde ponga x, y ésta es la idea que conviene tener in mente, pero, por desgracia, hay un pequeño inconveniente técnico que complica un poco el asunto.

Tomemos $\theta \equiv \bigwedge x(Axy \to Axp)$, que podemos interpretar como "todos los amigos de y son también amigos de Pedro". Tomemos $t \equiv fx$, con el significado de "el padre de x". Entonces $\mathtt{S}^t_y\theta$ debe significar "todos los amigos del padre de x son también amigos de Pedro", pero si nos limitamos a cambiar y por fx en θ nos queda: $\bigwedge x(Axfx \to Axp)$, que significa algo así como "todos los que son amigos de su propio padre son también amigos de Pedro", algo muy diferente de lo que buscábamos.

El problema es que la variable x está libre en $t \equiv fx$ y al poner a éste en lugar de y queda ligada por el cuantificador, mientras que lo que queremos es que siga libre. Lo correcto es $\mathtt{S}^t_x\theta \equiv \bigwedge z(Azfx \to Azp)$, donde z es una variable nueva. Esta fórmula sí cumple lo que queríamos. Teniendo esto en cuenta, la definición de sustitución ha de ser como sigue:

Definición 1.5 Sea $\mathcal{L}$ un lenguaje formal. Definimos la *sustitución* de una variable x por un término t en una expresión θ de $\mathcal{L}$ como la expresión $\mathtt{S}^t_x\theta$ determinada por las reglas siguientes:

a) $\mathtt{S}^t_x x_i \equiv \begin{cases} t & \text{si } x \equiv x_i, \\ x_i & \text{si } x \not\equiv x_i. \end{cases}$

b) $\mathtt{S}^t_x c_i \equiv c_i$.

c) $\mathtt{S}^t_x R^n_i t_1 \cdots t_n \equiv R^n_i S^t_x t_1 \cdots S^t_x t_n$.

d) $\mathtt{S}^t_x f^n_i t_1 \cdots t_n \equiv f^n_i S^t_x t_1 \cdots S^t_x t_n$.

e) $\mathtt{S}^t_x \neg\alpha \equiv \neg \mathtt{S}^t_x \alpha$.

f) $\mathtt{S}^t_x(\alpha \to \beta) \equiv \mathtt{S}^t_x\alpha \to \mathtt{S}^t_x\beta$.

g) $\mathtt{S}^t_x \bigwedge x_i\alpha \equiv \begin{cases} \bigwedge x_i\alpha & \text{si } x \text{ no está libre en } \bigwedge x_i\alpha, \\ \bigwedge x_i \mathtt{S}^t_x \alpha & \text{si } x \text{ está libre en } \bigwedge x_i\alpha \text{ y } x_i \text{ no lo está en } t, \\ \bigwedge x_j \mathtt{S}^t_x \mathtt{S}^{x_j}_{x_i} \alpha & \text{si } x \text{ está libre en } \bigwedge x_i\alpha,\ x_i \text{ está libre en } t \\ & \text{y } j \text{ es el menor índice tal que } x_j \text{ no está} \\ & \text{en } \bigwedge x_i\alpha \text{ ni en } t. \end{cases}$

h) $\mathtt{S}^t_x(x_i|\alpha) \equiv \begin{cases} x_i|\alpha & \text{si } x \text{ no está libre en } x_i|\alpha, \\ x_i|\mathtt{S}^t_x\alpha & \text{si } x \text{ está libre en } x_i|\alpha \text{ y } x_i \text{ no lo está en } t, \\ x_j|\mathtt{S}^t_x \mathtt{S}^{x_j}_{x_i}\alpha & \text{si } x \text{ está libre en } x_i|\alpha,\ x_i \text{ está libre en } t \\ & \text{y } j \text{ es el menor índice tal que } x_j \text{ no está} \\ & \text{en } x_i|\alpha \text{ ni en } t. \end{cases}$

Claramente se cumple

$$\mathtt{S}^t_x(\alpha \vee \beta) \equiv \mathtt{S}^t_x\alpha \vee \mathtt{S}^t_x\beta, \quad \mathtt{S}^t_x(\alpha \wedge \beta) \equiv \mathtt{S}^t_x\alpha \wedge \mathtt{S}^t_x\beta, \quad \mathtt{S}^t_x(\alpha \leftrightarrow \beta) \equiv \mathtt{S}^t_x\alpha \leftrightarrow \mathtt{S}^t_x\beta,$$

$$\mathtt{S}^t_x \bigvee x_i\alpha \equiv \begin{cases} \bigvee x_i\alpha & \text{si } x \text{ no está libre en } \bigvee x_i\alpha, \\ \bigvee x_i \mathtt{S}^t_x \alpha & \text{si } x \text{ está libre en } \bigvee x_i\alpha \text{ y } x_i \text{ no lo está en } t, \\ \bigvee x_j \mathtt{S}^t_x \mathtt{S}^{x_j}_{x_i} \alpha & \text{si } x \text{ está libre en } \bigvee x_i\alpha,\ x_i \text{ está libre en } t \\ & \text{y } j \text{ es el menor índice tal que } x_j \text{ no está} \\ & \text{en } \bigvee x_i\alpha \text{ ni en } t. \end{cases}$$

El lector debe convencerse de que esta definición es natural. Por ejemplo, la regla c) afirma que para sustituir x por t en $t_1 = t_2$ tendremos que sustituir todas las x que haya en t_1 y todas las que haya en t_2, dejando tal cual el igualador. El tercer caso de la regla g) dice que si queremos sustituir la variable x por t en $\bigwedge x_i\alpha$ pero x_i está libre en t (y x está libre en α o, si no no habría nada que sustituir) no podemos limitarnos a cambiar las x por t, pues entonces la x_i libre en t quedaría ligada por el cuantificador. Por ello antes cambiamos todas las x_i de α por una variable nueva x_j, luego sustituimos la x por la t y luego cuantificamos respecto a x_j, con lo que la x_i que estaba libre en t sigue estándolo al sustituir.

El teorema siguiente contiene las propiedades más importantes de la sustitución:

Teorema 1.6 *Se cumple:*

a) $\mathtt{S}^x_x\theta \equiv \theta$.

b) Si y *no está en* θ *entonces* $\mathtt{S}^x_y\mathtt{S}^y_x\theta \equiv \theta$.

c) Si x *no está libre en* θ*, entonces* $\mathtt{S}^t_x\theta \equiv \theta$.

d) Si x *está libre en* θ*, entonces las variables libres de* t *y las variables libres de* θ *distintas de* x *están libres en* $\mathtt{S}^t_x\theta$.

e) Una variable y *está libre en* $\mathtt{S}^t_x\theta$ *syss o bien* y *está libre en* θ *e* $y \not\equiv x$*, o bien* x *está libre en* θ *e* y *está libre en* t.

Todos estos hechos se prueban por inducción sobre la longitud de una expresión. Puesto que no estamos hablando de usar un teorema de inducción de la teoría de conjuntos —que es lo que hacen los matemáticos cuando razonan por inducción— conviene comentar brevemente el argumento en general:

Para probar que todas las expresiones de un lenguaje formal $\mathcal{L}$ cumplen una determinada propiedad P —como es la propiedad b)— basta probar lo siguiente:

a) x_i cumple P (es decir, toda variable cumple P).

b) c_i cumple P (es decir, toda constante cumple P. Naturalmente, si $\mathcal{L}$ no tiene constantes podemos omitir este paso.)

c) Si $t_1, \ldots, t_n$ cumplen P, entonces $R^n_i t_1 \cdots t_n$ cumple P.

d) Si $t_1, \ldots, t_n$ cumplen P, entonces $f^n_i t_1 \cdots t_n$ cumple P.

e) Si α cumple P, entonces $\neg\alpha$ cumple P.

f) Si α y β cumplen P, entonces $\alpha \to \beta$ cumple P.

g) Si α cumple P, entonces $\bigwedge x_i\alpha$ cumple P.

h) Si α cumple P, entonces $x_i|\alpha$ cumple P.

Si hemos probado esto, podemos estar seguros que de toda expresión θ cumple P. Por ejemplo, si $\theta \equiv \bigwedge x_0(x_0 = c_1 \to R^1_0 x_0)$, para que θ cumpla P basta con que $x_0 = c_1 \to R^1_0 x_0$ cumpla P (por g), para lo cual basta con que $x_0 = c_1$ y $R^1_0 x_0$ cumplan P (por f), para lo cual basta con que x_0 y c_1 cumplan P (por c), pero x_0 cumple P (por a) y c_1 cumple P (por b). En general, toda expresión θ contiene un número finito de subexpresiones, la mayor de las cuales es la propia θ. Si θ no cumpliera P, podríamos encontrar una subexpresión de longitud mínima que no cumpliera P, pero ésta tendría que ser de uno de los ocho tipos que permite la definición de expresión, y con ello desmentiríamos la propiedad a)–h) correspondiente a dicho tipo. ∎

Pasemos ya a la prueba del teorema 1.6.

DEMOSTRACIÓN: Todos los apartados se demuestran igual. Veamos sólo la propiedad b). Razonamos por inducción sobre la longitud de θ.

Observamos en primer lugar que $\mathtt{S}^x_y\mathtt{S}^y_x x_i \equiv x_i$. Para ello hay que distinguir dos casos: si $x \equiv x_i$ entonces $\mathtt{S}^x_y\mathtt{S}^y_x x_i \equiv \mathtt{S}^{x_i}_y y \equiv x_i$, mientras que si $x \not\equiv x_i$ entonces $\mathtt{S}^x_y\mathtt{S}^y_x x_i \equiv \mathtt{S}^x_y x_i \equiv x_i$ (porque, por hipótesis $y \not\equiv x_i$).

El caso de c_i es trivial, porque las constantes permanecen invariables. También son inmediatos los casos c), d), e) y f). Veamos g). Supongamos que $\theta \equiv \bigwedge x_i\alpha$.

$$\mathtt{S}^x_y\mathtt{S}^y_x\bigwedge x_i\alpha \equiv \begin{cases} \mathtt{S}^x_y\bigwedge x_i\alpha & \text{si } x \text{ no está libre en } \bigwedge x_i\alpha, \\ \mathtt{S}^x_y\bigwedge x_i\mathtt{S}^y_x\alpha & \text{si } x \text{ está libre en } \bigwedge x_i\alpha. \end{cases}$$

El tercer caso de la definición de sustitución no puede darse, porque exigiría que $x_i \equiv y$, pero estamos suponiendo que y no está en θ. Por este mismo motivo, en el primer caso llegamos a $\bigwedge x_i\alpha$. En el segundo caso es claro que si x está libre en $\bigwedge x_i\alpha$, también lo está en α, luego y está libre en $\mathtt{S}^y_x\alpha$ (esto se prueba fácilmente por inducción), y como $y \not\equiv x_i$, también está libre en $\bigwedge x_i\mathtt{S}^y_x\alpha$. Por consiguiente la sustitución es $\mathtt{S}^x_y\bigwedge x_i\mathtt{S}^y_x\alpha \equiv \bigwedge x_i\mathtt{S}^x_y\mathtt{S}^y_x\alpha \equiv \bigwedge x_i\alpha$ por hipótesis de inducción.

El caso h) es idéntico a g). ∎

La notación matemática Aunque la notación que hemos empleado para la sustitución es la más conveniente para nuestros fines, en matemáticas es habitual escribir $\theta(x)$ para indicar que $\theta(t) \equiv \mathtt{S}^t_x\theta$. No hay que entender, salvo que se diga explícitamente, que la variable x está libre en θ ni que sea la única variable libre de θ. El escribir $\theta(x)$ simplemente nos indica qué variable hay que sustituir por t para interpretar $\theta(t)$. Generalizar esto a varias variables requiere una precaución:

Si escribimos $\theta(y_1,\ldots,y_n)$ en lugar de θ, donde $y_1,\ldots,y_n$ son variables cualesquiera. Entonces se define

$$\theta(t_1,\ldots,t_n) \equiv \mathtt{S}^{t_1}_{z_1}\cdots\mathtt{S}^{t_n}_{z_n}\mathtt{S}^{z_1}_{y_1}\cdots\mathtt{S}^{z_n}_{y_n}\theta,$$

donde $z_1,\ldots,z_n$ son las variables de menor índice que no estén ni en $t_1,\ldots,t_n$ ni en θ.

Notemos que no podemos definir $\theta(t_1,t_2) \equiv \mathtt{S}^{t_1}_{y_1}\mathtt{S}^{t_2}_{y_2}\theta$ porque entonces estaríamos sustituyendo por t_1, no sólo las variables y_1 que hubiera en θ, sino también las que hubiera en t_2. ∎

1.6 Consideraciones finales

Es conveniente insistir en la diferencia entre los signos de un lenguaje formal y los nombres con que nos referimos a ellos. Por ejemplo, cuando en la prueba

del teorema 1.6 distinguimos los casos $x \equiv x_i$ y $x \not\equiv x_i$, esto carecería de sentido si "x" y "x_i" fueran variables en lugar de nombres de variables. La variable x tendrá un cierto índice j, es decir, $x \equiv x_j$, y la distinción es si $i = j$ o si $i \neq j$, es decir, si los nombres metamatemáticos "x_i" y "x" hacen referencia al mismo signo del lenguaje o no. Más en general, el lector debe ser consciente de que jamás vamos a escribir explícitamente los signos de ningún lenguaje formal (salvo a lo sumo en ejemplos aclaratorios sin valor teórico). Cuando un matemático escribe

$$\bigwedge u(u \in x \to u \in y),$$

no debemos pensar que está usando los signos de su lenguaje. La "u" no es una variable, el signo "$\in$" no es un relator diádico, etc., sino que son, respectivamente el nombre de una variable y el nombre del relator de pertenencia, cuyas formas concretas como signos son irrelevantes.

La situación es enteramente análoga a la del ajedrez. Los signos son como piezas de ajedrez. Podemos identificarlos con trozos de madera o de marfil, pero a ningún jugador de ajedrez le preocupan esas piezas. En realidad la teoría del ajedrez no trata sobre la madera o el marfil, sino sobre unos conceptos abstractos, como son "casilla", "alfil del rey blanco", "caballo de la dama negra", etc. Conceptos que se pueden materializar en piezas, en signos sobre un papel, o incluso manipular mentalmente, exactamente igual que ocurre, por ejemplo, con los números naturales.

Cuando escribimos $2 + 2 = 4$ (entendido como una afirmación metamatemática, es decir, como lo entiende cualquier niño en la escuela) sería artificial identificar el número 2 con el signo "2" con el que lo representamos. Más bien "2" es un signo castellano con el que nos referimos a un concepto abstracto, pero perfectamente determinado, como es el número 2. Similarmente, aunque podemos pensar que "$\to$" nombra a un determinado signo, como es "$\triangleleft$" en el lenguaje $\mathcal{L}_0$, es más exacto, dado lo poco que nos importa dicho signo, considerar que "$\to$" nombra en realidad a un objeto abstracto, de la misma naturaleza que el número 2.

Para terminar, hemos de señalar que la mayoría de los conceptos que hemos estudiado aquí (variables, cuantificadores, sustitución, etc.) de hecho, todos excepto el descriptor, que es un concepto más moderno, fueron introducidos y estudiados por Frege, si bien su notación era muy diferente a la moderna, que se debe esencialmente a Peano. Además de usar signos muy diferentes a los actuales, una diferencia notable entre el formalismo de Frege y el de Peano es que el segundo era lineal, es decir, las expresiones son sucesiones de signos que se leen de izquierda a derecha, mientras que la notación de Frege era bidimensional.

También se debe a Frege la distinción entre lenguajes de primer orden y lenguajes de segundo orden. Un lenguaje de segundo orden, además de los signos que nosotros hemos considerado, tiene variables de segundo orden, que pueden ser variables de función o variables de relación, cada una de las cuales tiene asignado un rango. Así, si R es una variable de relación monádica de un lenguaje de segundo orden, podemos escribir la fórmula de segundo orden

$$\bigwedge xy(x = y \leftrightarrow \bigwedge R\big(R(x) \leftrightarrow R(y)\big),$$

que se interpreta como que dos objetos son iguales si y sólo si tienen las mismas propiedades. Observemos que R no es un relator monádico pues, según la sintaxis de primer orden que hemos estudiado, no es correcto escribir $\bigwedge R$ cuando R es un relator. Por otra parte, R no es una variable de primer orden, pues entonces no tendría sentido la expresión $R(x)$. Cuando decimos que R es una variable de relación monádica hemos de entender que R es una variable cuyas interpretaciones posibles no son objetos, sino relaciones monádicas entre objetos. Naturalmente esto exige muchas precisiones en las que no vamos a entrar. La matemática no requiere más que lógica de primer orden para su fundamentación y, de hecho, puede probarse que la lógica de segundo orden no supone ninguna mejora.

Capítulo II

Sistemas deductivos formales

En este capítulo daremos una definición precisa de qué debemos entender por una deducción lógica, es decir, de qué ha de cumplirse para que una fórmula sea una consecuencia lógica de otra u otras fórmulas. De este modo, para determinar qué debemos entender por una demostración matemática sólo nos quedará establecer un sistema de axiomas, de modo que los teoremas matemáticos resultarán ser las consecuencias lógicas de los axiomas.

Debemos tener presente que no podemos dar una definición arbitraria de "deducción lógica", sino que nuestro objetivo es "capturar" la noción de razonamiento que los matemáticos vienen empleando desde hace milenios. En otras palabras, nosotros ya sabemos lo que es un razonamiento lógico, y si alguien no lo sabe de antemano, las páginas que siguen no se lo van a aclarar. Lo que pretendemos hacer con la lógica no es definirla, sino formalizarla, es decir, reducir los razonamientos lógicos a meras manipulaciones mecánicas de fórmulas que no requieran tener en cuenta su posible significado para decidir si son válidas o no. Esto nos permitirá aplicar la lógica sin vacilaciones a la matemática abstracta, donde cualquier intento de confiar en el significado de los conceptos entraña grandes riesgos.

De momento nos limitaremos a presentar una propuesta de cálculo deductivo y constataremos que se ajusta perfectamente a las "costumbres" de razonamiento de los matemáticos. Dejaremos para los capítulos siguientes la justificación de que esta similitud no es aproximada sino exacta, es decir, que nuestro cálculo deductivo es exactamente lo que queremos que sea. Por supuesto, para ello tendremos que precisar qué es lo que queremos que sea, pero eso podemos dejarlo para más adelante.

En una primera aproximación, un razonamiento es una sucesión de premisas tales que cada una es consecuencia de las anteriores en un sentido que hemos de precisar. En realidad esto, tal cual, es absurdo, pues la primera afirmación de un razonamiento no puede ser consecuencia de ninguna otra. Ya hemos comentado que la metamatemática se funda en la posibilidad de garantizar la verdad de

ciertas afirmaciones sin necesidad de razonamiento alguno. Por ejemplo, nuestra conciencia de que podemos añadir una unidad a cualquier número natural nos da la infinitud de los números naturales, pero esa conciencia, ese conocimiento de lo que es contar, no es un razonamiento en el sentido que acabamos de esbozar. Por desgracia, la matemática no puede fundarse en afirmaciones cuya verdad pueda ser establecida metamatemáticamente, pues carecemos de una representación suficientemente clara de un concepto matemático fundamental como es el de "conjunto". Precisamente por ello necesitamos el razonamiento formal, en el cual a las afirmaciones no demostradas les exigimos únicamente que sean declaradas explícitamente de antemano, bajo el nombre de axiomas o premisas.

2.1 El cálculo deductivo de primer orden

Introducimos ahora los conceptos necesarios para precisar las ideas anteriores. A lo largo de todo este capítulo $\mathcal{L}$ será un lenguaje formal prefijado. La definición básica es la siguiente:

Definición 2.1 Un *sistema deductivo formal* F sobre un lenguaje formal $\mathcal{L}$ viene determinado por una colección de fórmulas de $\mathcal{L}$, llamadas *axiomas* de F, y una colección de *reglas primitivas de inferencia* de F, que determinan cuándo una fórmula de $\mathcal{L}$ es *consecuencia inmediata* de otra u otras fórmulas de $\mathcal{L}$.

Esta definición se entenderá mejor tras el ejemplo siguiente:

El sistema deductivo formal $K_{\mathcal{L}}$ Dado un lenguaje formal $\mathcal{L}$, llamaremos $K_{\mathcal{L}}$ al sistema deductivo formal determinado por los siguientes axiomas y reglas de inferencia:

Axiomas de $K_{\mathcal{L}}$: Los axiomas de $K_{\mathcal{L}}$ son todas las fórmulas de los tipos siguientes, donde α, β, γ son fórmulas cualesquiera de $\mathcal{L}$ y t es un término cualquiera de $\mathcal{L}$.

K1	$\alpha \to (\beta \to \alpha)$	
K2	$(\alpha \to (\beta \to \gamma)) \to ((\alpha \to \beta) \to (\alpha \to \gamma))$	
K3	$(\neg\alpha \to \neg\beta) \to (\beta \to \alpha)$	
K4	$\bigwedge x_i \alpha \to \mathtt{S}^t_{x_i}\alpha$	
K5	$\bigwedge x_i(\alpha \to \beta) \to (\alpha \to \bigwedge x_i \beta)$	si x_i no está libre en α,
K6	$\bigwedge x_i(x_i = t \to \alpha) \leftrightarrow \mathtt{S}^t_{x_i}\alpha$	si x_i no está libre en t,
K7	$\overset{1}{\bigvee} x_i \alpha \to \mathtt{S}^{x_i\mid\alpha}_{x_i}\alpha$	si $\mathcal{L}$ tiene descriptor,
K8	$\neg\overset{1}{\bigvee} x_i\alpha \to x_i\mid\alpha = x_j\mid(x_j = x_j)$	si $\mathcal{L}$ tiene descriptor.

Según se indica, una fórmula de tipo $K5$ sólo es un axioma si se cumple la condición indicada sobre la variable x_i. Lo mismo es válido para las fórmulas de tipo $K6$.

Reglas de inferencia de $K_{\mathcal{L}}$:

Modus ponendo ponens (MP): de α y $\alpha \to \beta$ es consecuencia inmediata β.

Introducción del generalizador (IG): de α es consecuencia inmediata $\bigwedge x_i \alpha$.

■

Observemos que "$\alpha \to (\beta \to \alpha)$" no es un axioma de $K_{\mathcal{L}}$, sino lo que se llama un "esquema axiomático", es decir, una expresión metamatemática que determina infinitos axiomas de $K_{\mathcal{L}}$, los que resultan de sustituir las variables metamatemáticas "α" y "β" por fórmulas particulares de $\mathcal{L}$. Así pues, K_L tiene infinitos axiomas, lo cual no supone ninguna imprecisión o ambigüedad: la propiedad "ser un axioma de $K_{\mathcal{L}}$" está perfectamente definida. De hecho, es fácil diseñar un algoritmo que determine en un tiempo finito si cualquier fórmula de $\mathcal{L}$ es o no un axioma de $K_{\mathcal{L}}$. Por ejemplo, $\bigwedge x(x = y) \to y = y$ es un axioma de $K_{\mathcal{L}}$ (de tipo K4), mientras que $y = y$ no lo es.

Lo mismo sucede con las reglas de inferencia: Ahora podemos decir con total precisión que de la fórmula $x = y$ es consecuencia inmediata en $K_{\mathcal{L}}$ la sentencia $\bigwedge x\ x = y$, mientras que no es consecuencia inmediata la fórmula $y = x$.

Los axiomas y las reglas de inferencia son, en principio, arbitrarios. Si tomamos otros axiomas o determinamos de otro modo la noción de "ser consecuencia inmediata" tendremos un sistema deductivo formal distinto de $K_{\mathcal{L}}$. Por supuesto, dentro de toda esta generalidad, vamos a distinguir los sistemas deductivos formales "interesantes" de los carentes de valor. Por ejemplo, podríamos tomar todas las fórmulas de $\mathcal{L}$ como axiomas de un sistema deductivo formal, pero tal sistema resultaría trivial y sin interés para nuestros fines.

Una vez establecido un sistema deductivo formal, todas las nociones básicas de la lógica formal pueden ser definidas con precisión:

Definición 2.2 Una *deducción* en un sistema deductivo formal F a partir de una colección de fórmulas Γ es una sucesión finita $\alpha_1, \ldots, \alpha_n$ de fórmulas de $\mathcal{L}$ tales que cada α_i es un axioma de F, una fórmula de Γ o una consecuencia inmediata de fórmulas anteriores de la sucesión. Las fórmulas de Γ se llaman *premisas* de la deducción. Una *demostración* es una deducción sin premisas.

Una fórmula α es una *consecuencia* en F de una colección de fórmulas Γ si α es la última fórmula de una deducción en F a partir de Γ. Lo representaremos con la notación $\Gamma \underset{F}{\vdash} \alpha$.

Una fórmula α es un *teorema* de F si es la última fórmula de una demostración en F. Lo representaremos mediante $\underset{F}{\vdash} \alpha$.

Si no se indica F explícitamente, se entenderá que se trata de $K_{\mathcal{L}}$, es decir, escribiremos $\Gamma \vdash \alpha$ y $\vdash \alpha$ en lugar de $\Gamma \underset{K_{\mathcal{L}}}{\vdash} \alpha$ y $\underset{K_{\mathcal{L}}}{\vdash} \alpha$.

Los axiomas y teoremas de $K_{\mathcal{L}}$ se llaman *axiomas y teoremas lógicos,* las consecuencias en $K_{\mathcal{L}}$ se llaman *consecuencias lógicas.*

Si la colección de premisas es finita, digamos $\gamma_1, \ldots, \gamma_n$, escribiremos también $\gamma_1, \ldots, \gamma_n \underset{F}{\vdash} \alpha$. La notación $\Gamma, \gamma_1, \ldots, \gamma_n \underset{F}{\vdash} \alpha$ significará, como es obvio, que α es consecuencia en F de las premisas de Γ más las indicadas explícitamente.

Claramente, si $\Gamma \underset{F}{\vdash} \alpha$, existen $\gamma_1, \ldots, \gamma_n$ en Γ tales que $\gamma_1, \ldots, \gamma_n \underset{F}{\vdash} \alpha$.

Axiomas y premisas La diferencia que hemos establecido entre axiomas y premisas (o entre deducciones y demostraciones) es meramente lingüística: no hay ninguna diferencia entre una deducción en un sistema deductivo F a partir de unas premisas Γ y una demostración en el sistema deductivo que se obtiene a partir de F añadiendo las fórmulas de Γ como axiomas. No obstante la distinción resulta útil en la práctica: las premisas son "axiomas temporales", es decir, axiomas que se suponen momentáneamente, como cuando se supone α para probar $\alpha \to \beta$. Esto se verá con más claridad a medida que avancemos. ∎

Veamos un primer ejemplo de demostración formal:

Teorema 2.3 *Si α es una fórmula de $\mathcal{L}$, entonces $\vdash \alpha \to \alpha$.*

Demostración:

(1)	$(\alpha \to ((\alpha \to \alpha) \to \alpha)) \to ((\alpha \to (\alpha \to \alpha)) \to (\alpha \to \alpha))$	(K2)
(2)	$\alpha \to ((\alpha \to \alpha) \to \alpha)$	(K1)
(3)	$(\alpha \to (\alpha \to \alpha)) \to (\alpha \to \alpha)$	(MP 1,2)
(4)	$\alpha \to (\alpha \to \alpha)$	(K1)
(5)	$\alpha \to \alpha$	(MP 3,4)

∎

Observaciones En este punto hemos de señalar varios hechos:

— El teorema anterior no es un teorema en el sentido que acabamos de definir, así como su demostración tampoco es una demostración en el sentido de 2.2. Lo que sigue a la palabra "teorema" no es una fórmula de $\mathcal{L}$, sino una afirmación metamatemática sobre $K_{\mathcal{L}}$, a saber, que todas las fórmulas del tipo $\alpha \to \alpha$ son teoremas de $K_{\mathcal{L}}$. Es lo que también se llama un *metateorema.* La diferencia entre un teorema formal y un metateorema es la misma que hay entre una partida de ajedrez y un resultado como "con un rey y un alfil no se puede dar mate a un rey solo". Esto último no es algo que se hace siguiendo reglas fijas, como una partida de ajedrez, sino que es algo que uno puede concluir analizando adecuadamente dichas reglas. Igualmente, analizando el concepto de demostración en K_L que hemos dado, podemos concluir que tiene la propiedad indicada. Esto no lo concluimos a partir de ciertos axiomas o reglas, sino que nos damos cuenta de ello igual que podemos juzgar sobre qué podemos hacer al jugar a las cartas, o a cualquier otra cosa.

En concreto, lo que hemos hecho es dar un *esquema de demostración,* de manera que si cogemos cualquier fórmula, como $x = y$, por ejemplo, y la ponemos allí donde pone α, obtenemos ciertamente una demostración de $x = y \to x = y$ en el sentido de 2.2.

Notemos que en el capítulo anterior hemos probado ya muchos metateoremas, no sólo el teorema 1.6, que es el único que hemos enunciado como tal, sino que cualquier afirmación metamatemática que requiere una justificación es un metateorema, como el hecho de que una expresión no puede ser a la vez un término y una fórmula.

— Lo siguiente que debemos entender es que esto no es una demostración de que si α entonces α. Sería infantil demostrar que si α entonces α, y mucho más creer que esto lo demuestra, mientras que quien no conoce esta prueba no sabe, en realidad, por qué si α entonces α. Está claro que si α entonces α, todo el mundo lo sabe y si el lector fuera una excepción, haría mejor en no esforzarse más y dejar este libro. Lo que hemos probado —y no es evidente— es que en $K_{\mathcal{L}}$ puede demostrarse que $\alpha \to \alpha$ cualquiera que sea α. ¿Cree el lector que hubiera sido capaz de probarlo él mismo sin conocer nuestra prueba? No es especialmente difícil, pero tampoco es obvio.

El presente capítulo está dedicado en gran parte a conocer $K_{\mathcal{L}}$ y hacernos una idea de lo que puede probarse en él. Este resultado ha sido el primer paso.

— Nótese la numeración a la izquierda de las líneas y las anotaciones a la derecha que indican cómo se obtiene cada línea. Esto no lo exige la definición de demostración pero ayuda a leerla y entenderla, por lo que en adelante lo tomaremos por costumbre.

— Por último destaquemos algo que posiblemente sea lo primero que ha observado el lector: esta demostración es horrible, es lo menos natural que podría imaginarse.

Es cierto, pero no tiene importancia. Lo que hemos hecho ha sido como si un matemático, para calcular $\int_0^1 x^2\,dx$, empezara a calcular particiones de $[0, 1]$, sumas inferiores, y hallara el supremo de estas sumas, es decir, aplicara la definición de integral. Es posible hacerlo, Arquímedes lo hizo, sin embargo un matemático dispone de técnicas que le permiten calcular esa integral en unos segundos. De esta forma el matemático posee una definición teórica de integral, en términos de épsilon y delta, que le es muy útil para tratar problemas teóricos sobre integrales aunque a la hora de calcular resulte incomodísima, pero esto lo suple con resultados prácticos que convierten los cálculos concretos en algo relativamente sencillo.

Igualmente, la definición de demostración en $K_{\mathcal{L}}$ es teóricamente simplicísima y será muy útil a la hora de estudiar las demostraciones. ¿Qué son ocho esquemas de axioma y dos reglas de inferencia frente a las infinitas formas de razonar con las que aparentemente cuenta el matemático? Pero tratar de demostrar cosas en $K_{\mathcal{L}}$ sin más ayuda que la definición de demostración es como querer calcular $\int_0^1 \operatorname{arcsen} x\,dx$ sin más ayuda que la definición de integral.

Como ilustración de esta última observación, vamos a probar un teorema

que vuelve inmediato el teorema anterior. Cuando un matemático quiere probar una fórmula de tipo $\alpha \to \beta$, supone α como premisa y trata de llegar a β. Si lo consigue, da por probado que $\alpha \to \beta$. Vamos a ver que esto es lícito en $K_{\mathcal{L}}$. No figura en la definición de deducción, pero se deduce de ella.

Teorema 2.4 (Teorema de deducción) *Sean α y β fórmulas de $\mathcal{L}$ y sea Γ una colección de fórmulas de $\mathcal{L}$. Si $\Gamma, \alpha \vdash \beta$ y existe una deducción de β en la que no se generalice respecto a variables libres en α, entonces $\Gamma \vdash \alpha \to \beta$.*

DEMOSTRACIÓN: Lo probamos por inducción sobre el número n de líneas de una deducción de β.

Si $n = 1$, entonces β es una deducción de β, luego β es un axioma, o bien β está en Γ o bien $\beta \equiv \alpha$.

a) β es un axioma o β está en Γ:

Veamos una deducción de $\alpha \to \beta$ a partir de Γ:

(1)	β	(axioma o premisa)
(2)	$\beta \to (\alpha \to \beta)$	(K1)
(3)	$\alpha \to \beta$	(MP 1, 2)

b) $\beta \equiv \alpha$: Por el teorema anterior $\vdash \alpha \to \beta$, luego $\Gamma \vdash \alpha \to \beta$.

Supongamos el teorema cierto para fórmulas deducibles en menor de n pasos y supongamos que β se deduce en n pasos.

a) Si β es un axioma o una premisa se razona como antes.

b) Si β se sigue por MP de γ y $\gamma \to \beta$, fórmulas anteriores de la deducción, entonces γ y $\gamma \to \beta$ se deducen de Γ y α en menos de n pasos y sin generalizar respecto a variables libres en α. Por hipótesis de inducción $\Gamma \vdash \alpha \to \gamma$ y $\Gamma \vdash \alpha \to (\gamma \to \beta)$. Veamos una deducción de $\alpha \to \beta$ a partir de Γ:

(1)		
$\vdots$	(deducción de $\alpha \to \gamma$ a partir de Γ)	
(k)	$\alpha \to \gamma$	
(k+1)		
$\vdots$	(deducción de $\alpha \to (\gamma \to \beta)$ a partir de Γ)	
(l)	$\alpha \to (\gamma \to \beta)$	
(l+1)	$(\alpha \to (\gamma \to \beta)) \to ((\alpha \to \gamma) \to (\alpha \to \beta))$	(K2)
(l+2)	$(\alpha \to \gamma) \to (\alpha \to \beta)$	(MP l, l+1)
(l+3)	$\alpha \to \beta$	(MP k, l+2)

c) Si β es consecuencia de γ, fórmula anterior, por IG, entonces $\beta \equiv \bigwedge x_i \gamma$ y, por hipótesis, x_i no está libre en α. Por hipótesis de inducción $\Gamma \vdash \alpha \to \gamma$, ya que γ se deduce de Γ y α en menos de n pasos. Veamos una deducción de $\alpha \to \beta$ a partir de Γ:

(1)		
⋮	(deducción de $\alpha \to \gamma$ a partir de Γ)	
(k)	$\alpha \to \gamma$	
(k+1)	$\bigwedge x_i(\alpha \to \gamma)$	(IG k)
(k+2)	$\bigwedge x_i(\alpha \to \gamma) \to (\alpha \to \bigwedge x_i\gamma)$	(K5)
(k+3)	$\alpha \to \beta$	(MP k+1, k+2)

Conviene observar que en la deducción de $\alpha \to \beta$ se generaliza respecto a las mismas variables que en la de β ∎

Es fácil extraer de la demostración anterior un algoritmo que nos permite construir explícitamente una deducción $\Gamma \vdash \alpha \to \beta$ a partir de una deducción de β en las condiciones del enunciado. Esto es lo que lo convierte en una demostración (metamatemática) rigurosa.

Más adelante podremos probar que la exigencia técnica de que en la deducción de β no se generalice respecto de variables libres en α es realmente necesaria. De momento únicamente podemos mostrar que, de acuerdo con las "costumbres" de los matemáticos, es una restricción razonable. En efecto, ante todo debemos indicar que —en un sentido que será precisado en el capítulo siguiente— el sistema $K_{\mathcal{L}}$ presupone un convenio en cuanto a la interpretación pretendida de las fórmulas de $\mathcal{L}$, y es que una fórmula con variables libres, como $x = y$, significa —por convenio— lo mismo que $\bigwedge xy\ x = y$. Esto es lo que hace "razonable" a la regla de generalización.

Supongamos ahora que queremos probar la sentencia $\bigwedge xy(x = y \to y = x)$. Para ello basta probar la fórmula $x = y \to y = x$, pues si llegamos a ella bastará aplicar dos veces IG para obtener la sentencia buscada. (Este paso equivale a lo que hace el matemático cuando dice: "hemos de probar que todo x y todo y cumplen $x = y \to y = x$, luego fijamos x e y y vamos a ver que lo cumplen".) A continuación, para probar la implicación el matemático supone $x = y$ y se propone demostrar $y = x$, pero a partir de este punto ya no puede generalizar respecto de x e y, ya que ahora x e y son dos objetos concretos que verifican $x = y$, y esto no tienen por qué cumplirlo dos objetos cualesquiera. Si pudiéramos generalizar respecto de x e y podríamos pasar a $\bigwedge xy\ x = y$, pero esto no es "razonable": de suponer que tenemos dos objetos iguales no podemos pasar a que todo par de objetos son iguales. Esta restricción, que el matemático asume casi inconscientemente, aparece explícitamente en el teorema de deducción.

El recíproco del teorema de deducción es inmediato:

Ejercicio: Probar que si $\Gamma \vdash \alpha \to \beta$, entonces $\Gamma, \alpha \vdash \beta$.

2.2 Reglas derivadas de inferencia

Ningún matemático demuestra sus teoremas explícitamente a partir de axiomas. En realidad todo matemático acepta que en una demostración se incluyan sin prueba teoremas que ya han sido demostrados previamente. Es claro

que esto es lícito y nosotros también podemos hacerlo, para evitar así que la demostración del hecho más insignificante ocupe cientos de páginas. Concretamente, si es conocido que $\alpha_1, \ldots, \alpha_n \vdash \alpha$ y entre las líneas de una deducción aparecen $\alpha_1, \ldots, \alpha_n$, admitiremos la escritura de α como línea posterior abreviando la deducción completa, que incluiría las líneas necesarias para deducir α de $\alpha_1, \ldots, \alpha_n$, pero que se sobrentienden por conocidas. En particular todo teorema lógico puede escribirse en una deducción.

Llamaremos *reglas derivadas de inferencia* a ciertos resultados de la forma

$$\alpha_1, \ldots, \alpha_n \vdash \alpha$$

que nos permitirán abreviar considerablemente las deducciones.

Regla de repetición (R): $\alpha \vdash \alpha$.

Es inmediato que de α se deduce α, pues la propia α es una deducción de α con premisa α. Sin embargo aquí queremos algo ligeramente más fuerte, y es que en una deducción podemos repetir una línea anterior si queremos. Esto ya no es evidente, pues α no es consecuencia de α por ninguna de las dos reglas de inferencia. No obstante, si tenemos α en una deducción, podemos escribir $\alpha \to \alpha$ por ser un teorema lógico y a continuación escribir α de nuevo por MP.

Modus Barbara (MB): $\alpha \to \beta,\ \beta \to \gamma \vdash \alpha \to \gamma$.

Demostración:

(1)	$\alpha \to \beta$	(premisa)
(2)	$\beta \to \gamma$	(premisa)
(3)	α	(premisa)
(4)	β	(MP 1,3)
(5)	γ	(MP 2, 4)

Así pues, $\alpha \to \beta,\ \beta \to \gamma,\ \alpha \vdash \gamma$ y, por el teorema de deducción,

$$\alpha \to \beta,\ \beta \to \gamma \vdash \alpha \to \gamma.$$

Reglas de la doble negación (DN): $\neg\neg\alpha \vdash \alpha, \qquad \alpha \vdash \neg\neg\alpha$.

Demostración:

(1)	$\neg\neg\alpha$	(premisa)
(2)	$\neg\neg\alpha \to (\neg\neg\neg\neg\alpha \to \neg\neg\alpha)$	(K1)
(3)	$\neg\neg\neg\neg\alpha \to \neg\neg\alpha$	(MP 1, 2)
(4)	$(\neg\neg\neg\neg\alpha \to \neg\neg\alpha) \to (\neg\alpha \to \neg\neg\neg\alpha)$	(K3)
(5)	$\neg\alpha \to \neg\neg\neg\alpha$	(MP 3, 4)
(6)	$(\neg\alpha \to \neg\neg\neg\alpha) \to (\neg\neg\alpha \to \alpha)$	(K3)
(7)	$\neg\neg\alpha \to \alpha$	(MP 5, 6)
(8)	α	(MP 1, 7)

Así pues, $\neg\neg\alpha \vdash \alpha$. Por el teorema de deducción $\vdash \neg\neg\alpha \to \alpha$. Esto vale para toda fórmula α. Aplicándolo a $\neg\alpha$ obtenemos que $\vdash \neg\neg\neg\alpha \to \neg\alpha$.

(1)	α	(premisa)
(2)	$\neg\neg\neg\alpha \to \neg\alpha$	(teorema lógico)
(3)	$(\neg\neg\neg\alpha \to \neg\alpha) \to (\alpha \to \neg\neg\alpha)$	(K3)
(4)	$\alpha \to \neg\neg\alpha$	(MP 2, 3)
(5)	$\neg\neg\alpha$	(MP 1, 4)

Por el teorema de deducción $\vdash \alpha \to \neg\neg\alpha$. También llamaremos DN a los teoremas $\vdash \alpha \to \neg\neg\alpha$ y $\vdash \neg\neg\alpha \to \alpha$.

Reglas de la negación de la implicación (NI):

$$\alpha \to \beta \vdash \neg\beta \to \neg\alpha \qquad \neg\beta \to \neg\alpha \vdash \alpha \to \beta$$

$$\alpha \to \neg\beta \vdash \beta \to \neg\alpha \qquad \neg\alpha \to \beta \vdash \neg\beta \to \alpha$$

Demostración:

(1)	$\neg\neg\alpha \to \alpha$	(DN)
(2)	$\alpha \to \beta$	(premisa)
(3)	$\neg\neg\alpha \to \beta$	(MB 1, 2)
(4)	$\beta \to \neg\neg\beta$	(DN)
(5)	$\neg\neg\alpha \to \neg\neg\beta$	(MB 3, 4)
(6)	$(\neg\neg\alpha \to \neg\neg\beta) \to (\neg\beta \to \neg\alpha)$	(K3)
(7)	$\neg\beta \to \neg\alpha$	(MP 5, 6)

Las otras variantes se prueban de forma similar.

Modus tollendo tollens (MT): $\alpha \to \beta,\ \neg\beta \vdash \neg\alpha, \qquad \alpha \to \neg\beta,\ \beta \vdash \neg\alpha.$

(Por NI y el recíproco del teorema de deducción.)

Regla de la contradicción (C): $\alpha,\ \neg\alpha \vdash \beta$.

Demostración:

(1)	α	(premisa)
(2)	$\alpha \to (\neg\beta \to \alpha)$	(K1)
(3)	$\neg\beta \to \alpha$	(MP 1, 2)
(4)	$\neg\alpha$	(premisa)
(5)	$\neg\neg\beta$	(MT 3, 4)
(6)	β	(DN 5)

Reglas de equivalencia entre disyunción e implicación (EDI):

$$\alpha \vee \beta \vdash \neg\alpha \to \beta \qquad \neg\alpha \to \beta \vdash \alpha \vee \beta.$$

(Son casos particulares de (R), pues $\alpha \vee \beta \equiv \neg\alpha \to \beta$.)

Reglas de introducción del disyuntor (ID): $\alpha \vdash \alpha \vee \beta \qquad \alpha \vdash \beta \vee \alpha$.

DEMOSTRACIÓN: Por el teorema de deducción aplicado a (C) obtenemos que $\alpha \vdash \neg\alpha \to \beta$, o sea, $\alpha \vdash \alpha \vee \beta$.

(1)	α	(premisa)
(2)	$\alpha \to (\neg\beta \to \alpha)$	(K1)
(3)	$\neg\beta \to \alpha$	(MP 1, 2)
(4)	$\beta \vee \alpha$	(EDI 3)

Modus tollendo ponens (MTP): $\alpha \vee \beta,\ \neg\alpha \vdash \beta, \qquad \alpha \vee \beta,\ \neg\beta \vdash \alpha$.

DEMOSTRACIÓN:

(1)	$\alpha \vee \beta$	(premisa)
(2)	$\neg\alpha \to \beta$	(EDI 1)
(3)	$\neg\alpha$	(premisa)
(4)	β	(MP 2, 3)

La otra es similar.

Leyes de De Morgan (DM):

$$\begin{array}{ll} \alpha \wedge \beta \vdash \neg(\neg\alpha \vee \neg\beta) & \neg(\neg\alpha \vee \neg\beta) \vdash \alpha \wedge \beta \\ \alpha \vee \beta \vdash \neg(\neg\alpha \wedge \neg\beta) & \neg(\neg\alpha \wedge \neg\beta) \vdash \alpha \vee \beta \\ \neg(\alpha \wedge \beta) \vdash \neg\alpha \vee \neg\beta & \neg\alpha \vee \neg\beta \vdash \neg(\alpha \wedge \beta) \\ \neg(\alpha \vee \beta) \vdash \neg\alpha \wedge \neg\beta & \neg\alpha \wedge \neg\beta \vdash \neg(\alpha \vee \beta) \end{array}$$

DEMOSTRACIÓN: Las dos primeras son casos particulares de (R), pues por definición $\alpha \wedge \beta \equiv \neg(\neg\alpha \vee \neg\alpha)$.

(1)	$\alpha \vee \beta$	(premisa)	(1)	$\neg(\neg\alpha \wedge \neg\beta)$	(premisa)
(2)	$\neg\alpha \to \beta$	(EDI 1)	(2)	$\neg\neg(\neg\neg\alpha \vee \neg\neg\beta)$	(R 1)
(3)	$\neg\beta \to \neg\neg\alpha$	(NI 2)	(3)	$\neg\neg\alpha \vee \neg\neg\beta$	(DN 2)
(4)	$\neg\neg\neg\alpha \to \neg\neg\beta$	(NI 3)	(4)	$\neg\neg\neg\alpha \to \neg\neg\beta$	(EDI 3)
(5)	$\neg\neg\alpha \vee \neg\neg\beta$	(EDI 4)	(5)	$\neg\beta \to \neg\neg\alpha$	(NI 5)
(6)	$\neg\neg(\neg\neg\alpha \vee \neg\neg\beta)$	(DN 5)	(6)	$\neg\alpha \to \beta$	(NI 5)
(7)	$\neg(\neg\alpha \wedge \neg\beta)$	(R 6)	(8)	$\alpha \vee \beta$	(EDI 6)

Las restantes se siguen fácilmente de éstas.

Reglas de introducción del conjuntor (IC): $\alpha,\ \beta \vdash \alpha \wedge \beta$.

DEMOSTRACIÓN: Por el teorema de deducción sobre (MTP) se cumple

$$(*): \qquad \neg\neg\beta \vdash \neg\alpha \vee \neg\beta \to \neg\alpha.$$

(1)	β	(premisa)
(2)	$\neg\neg\beta$	(DN 1)
(3)	$\neg\alpha \vee \neg\beta \to \neg\alpha$	$(*)$
(4)	$\neg\neg\alpha \to \neg(\neg\alpha \vee \neg\beta)$	(NI 3)
(5)	α	(premisa)
(6)	$\neg\neg\alpha$	(DN 5)
(7)	$\neg(\neg\alpha \vee \neg\beta)$	(MP 4, 6)
(8)	$\alpha \wedge \beta$	(DM 7)

Reglas de eliminación del conjuntor (EC): $\alpha \wedge \beta \vdash \alpha, \qquad \alpha \wedge \beta \vdash \beta.$

Demostración:

(1)	$\neg\alpha \to \neg\alpha \vee \neg\beta$	(ID)
(2)	$\neg(\neg(\alpha \vee \neg\beta) \to \alpha$	(NI 1)
(3)	$\alpha \wedge \beta$	(premisa)
(4)	$\neg(\neg\alpha \vee \neg\beta)$	(DM 3)
(5)	α	(MP 2, 3)

Análogamente se prueba la otra. También llamaremos (EC) a los teoremas

$$\vdash \alpha \wedge \beta \to \alpha \qquad \vdash \alpha \wedge \beta \to \beta.$$

Regla de eliminación del disyuntor (ED): $\alpha \vee \alpha \vdash \alpha$.

Demostración:

(1)	$\neg\alpha \to \neg\alpha \wedge \neg\alpha$	(IC)
(2)	$\neg(\neg\alpha \wedge \neg\alpha) \to \alpha$	(NI 1)
(3)	$\alpha \vee \alpha$	(premisa)
(4)	$\neg(\neg\alpha \wedge \neg\alpha)$	(DM 3)
(5)	α	(MP 3, 4)

También llamaremos (ED) al teorema $\vdash \alpha \vee \alpha \to \alpha$.

Regla del tertium non datur (TND): $\vdash \alpha \vee \neg\alpha$.

Es un caso particular del teorema $\vdash \alpha \to \alpha$, pues $\alpha \vee \neg\alpha \equiv \neg\alpha \to \neg\alpha$.

Regla de no contradicción (NC): $\vdash \neg(\alpha \wedge \neg\alpha)$.

Demostración:

(1)	$\neg\alpha \vee \neg\neg\alpha$	(TND)
(2)	$\neg(\alpha \wedge \neg\alpha)$	(DM 1)

Regla del dilema (Dil): $\alpha \to \beta,\ \gamma \to \beta \vdash \alpha \vee \gamma \to \beta$.

Demostración:

(1)	$\alpha \to \beta$	(premisa)
(2)	$\gamma \to \beta$	(premisa)
(3)	$\alpha \vee \gamma$	(premisa)
(4)	$\neg\alpha \to \gamma$	(EDI 3)
(5)	$\neg\beta \to \neg\alpha$	(NI 1)
(6)	$\neg\beta \to \gamma$	(MB 4, 5)
(7)	$\neg\beta \to \beta$	(MB 2, 6)
(8)	$\beta \vee \beta$	(EDI 7)
(9)	β	(ED 8)

Por lo tanto $\alpha \to \beta,\ \gamma \to \beta,\ \alpha \vee \gamma \vdash \beta$ y, por el teorema de deducción, $\alpha \to \beta,\ \gamma \to \beta \vdash \alpha \vee \gamma \to \beta$.

Reglas de introducción y eliminación del bicondicionador

$$\begin{array}{ll} \text{(IB)} & \alpha \to \beta,\ \beta \to \alpha \vdash \alpha \leftrightarrow \beta. \\ \text{(EB)} & \alpha \leftrightarrow \beta \vdash \alpha \to \beta, \quad \alpha \leftrightarrow \beta \vdash \beta \to \alpha \end{array}$$

Son casos particulares de (IC), (EC), pues $\alpha \leftrightarrow \beta \equiv (\alpha \to \beta) \wedge (\beta \to \alpha)$.

Regla de eliminación del generalizador (EG): $\bigwedge x\alpha \vdash \mathtt{S}^t_x\alpha$.

Por el axioma (K4) y el recíproco del teorema de deducción.

Reglas de negación del generalizador (NG):

$$\begin{array}{ll} \neg\bigwedge x\neg\alpha \vdash \bigvee x\alpha & \bigvee x\alpha \vdash \neg\bigwedge x\neg\alpha \\ \neg\bigwedge x\alpha \vdash \bigvee x\neg\alpha & \bigvee x\neg\alpha \vdash \neg\bigwedge x\alpha \end{array}$$

Demostración: Las dos primeras son casos particulares de (R), pues $\bigvee x\alpha \equiv \neg\bigwedge x\neg\alpha$.

(1)	$\neg\bigvee x\neg\alpha$	(premisa)
(2)	$\neg\neg\bigwedge x\neg\neg\alpha$	(R 1)
(3)	$\bigwedge x\neg\neg\alpha$	(DN 2)
(4)	$\neg\neg\alpha$	(EG 3) ($\neg\neg\alpha \equiv \mathtt{S}^x_x\neg\neg\alpha$).
(5)	α	(DN 4)
(6)	$\bigwedge x\alpha$	(IG 5)

Por el teorema de deducción se cumple $\vdash \neg\bigvee x\neg\alpha \to \bigwedge x\alpha$, por (NI) tenemos $\vdash \neg\bigwedge x\alpha \to \bigvee x\neg\alpha$, luego $\neg\bigwedge x\alpha \vdash \bigvee x\neg\alpha$.

(1)	$\bigwedge x\alpha$	(premisa)
(2)	α	(EG 1)
(3)	$\neg\neg\alpha$	(DN 2)
(4)	$\bigwedge x\neg\neg\alpha$	(IG 3)

En consecuencia $\vdash \bigwedge x\alpha \to \bigwedge x\neg\neg\alpha$. Por (NI) se cumple

$$\vdash \neg\bigwedge x\neg\neg\alpha \to \neg\bigwedge x\alpha,$$

y por el recíproco del teorema de deducción $\neg\bigwedge x\neg\neg\alpha \vdash \neg\bigwedge x\alpha$ o, lo que es lo mismo, $\bigvee x\neg\alpha \vdash \neg\bigwedge x\alpha$.

Reglas de negación del particularizador (NP):

$$\begin{array}{ll} \neg\bigvee x\alpha \vdash \bigwedge x\neg\alpha & \bigwedge x\neg\alpha \vdash \neg\bigvee x\alpha \\ \neg\bigvee x\neg\alpha \vdash \bigwedge x\alpha & \bigwedge x\alpha \vdash \neg\bigvee x\neg\alpha \end{array}$$

Demostración:

(1)	$\neg\bigvee x\alpha$	(premisa)	(1)	$\bigwedge x\neg\alpha$	(premisa)
(2)	$\neg\neg\bigwedge x\neg\alpha$	(R 1)	(2)	$\neg\neg\bigwedge x\neg\alpha$	(DN 2)
(3)	$\bigwedge x\neg\alpha$	(DN 2)	(3)	$\neg\bigvee x\alpha$	(R 2)

(1)	$\neg\bigvee x\neg\alpha$	(premisa)	(1)	$\bigwedge x\alpha$	(premisa)
(2)	$\neg\neg\bigwedge x\neg\neg\alpha$	(R1)	(2)	α	(EG 1)
(3)	$\bigwedge x\neg\neg\alpha$	(DN 2)	(3)	$\neg\neg\alpha$	(DN 2)
(4)	$\neg\neg\alpha$	(EG 3)	(4)	$\bigwedge x\neg\neg\alpha$	(IG 3)
(5)	α	(DN 4)	(5)	$\neg\neg\bigwedge x\neg\neg\alpha$	(DN 4)
(6)	$\bigwedge x\alpha$	(IG 5)	(6)	$\neg\bigvee x\neg\alpha$	(R5)

Regla de introducción del particularizador (IP): $\mathtt{S}^t_x\alpha \vdash \bigvee x\alpha$.

Demostración:

(1)	$\bigwedge x\neg\alpha \to \mathtt{S}^t_x\neg\alpha$	(K4)
(2)	$\bigwedge x\neg\alpha \to \neg\mathtt{S}^t_x\alpha$	(R1)
(3)	$\mathtt{S}^t_x\alpha$	(premisa)
(4)	$\neg\bigwedge x\neg\alpha$	(MT 2, 3)
(5)	$\bigvee x\alpha$	(NP 4)

Reglas de introducción y eliminación del igualador

(II)	$\mathtt{S}^t_x\alpha \vdash \bigwedge x(x = t \to \alpha)$,	si x no está libre en t.
(EI)	$\bigwedge x(x = t \to \alpha) \vdash \mathtt{S}^t_x\alpha$,	si x no está libre en t.

Por (K6) y (EB).

Regla de la identidad (I): $\vdash t = t$.

Sea x una variable que no esté libre en t.

(1)	$x = t \to x = t$	(teorema lógico)
(2)	$\bigwedge x(x = t \to x = t)$	(IG 1)
(3)	$\mathtt{S}^t_x(x = t)$	(EI 2)
(4)	$t = t$	(R 3)

Regla de la simetría de la identidad (SI): $t_1 = t_2 \vdash t_2 = t_1$.

DEMOSTRACIÓN: Sea x una variable que no esté en t_1 ni en t_2.

(1)	$t_2 = t_2$	(I)
(2)	$\mathtt{S}_x^{t_2}(t_2 = x)$	(R 1)
(3)	$\bigwedge x(x = t_2 \to t_2 = x)$	(II 2)
(4)	$t_1 = t_2 \to t_2 = t_1$	(EG 3)
(5)	$t_1 = t_2$	(premisa)
(6)	$t_2 = t_1$	(MP 4, 5)

Regla de la transitividad de la identidad (TI): $t_1 = t_2,\ t_2 = t_3 \vdash t_1 = t_3$.

DEMOSTRACIÓN: Sea x una variable que no esté libre en t_1, t_2, t_3.

(1)	$t_2 = t_3$	(premisa)
(2)	$\bigwedge x(x = t_2 \to x = t_3)$	(II 1)
(3)	$t_1 = t_2 \to t_1 = t_3$	(EG 2)
(4)	$t_1 = t_2$	(premisa)
(5)	$t_1 = t_3$	(MP 3, 4)

Regla de equivalencia entre términos idénticos (ETI):

$$t_1 = t_2,\ \mathtt{S}_x^{t_2}\alpha \vdash \mathtt{S}_x^{t_1}\alpha.$$

DEMOSTRACIÓN: Sea y una variable que no esté en α, t_1, t_2. Entonces $\mathtt{S}_x^{t_2} \equiv \mathtt{S}_y^{t_2}\mathtt{S}_x^{y}\alpha$ y $\mathtt{S}_x^{t_1} \equiv \mathtt{S}_y^{t_1}\mathtt{S}_x^{y}\alpha$.

(1)	$\mathtt{S}_y^{t_2}\mathtt{S}_x^{y}\alpha$	(premisa)
(2)	$\bigwedge y(y = t_2 \to \mathtt{S}_x^{y}\alpha)$	(II 1)
(3)	$\mathtt{S}_y^{t_1}(y = t_2 \to \mathtt{S}_x^{y}\alpha)$	(EG 2)
(4)	$t_1 = t_2 \to \mathtt{S}_x^{t_1}\alpha$	(R 3)
(5)	$t_1 = t_2$	(premisa)
(6)	$\mathtt{S}_x^{t_1}\alpha$	(MP, 4, 5)

Las dos últimas reglas son para lenguajes formales con descriptor.

Regla de las descripciones propias (DP): $\overset{1}{\bigvee} x\alpha \vdash \mathtt{S}_x^{x|\alpha}\alpha$.

Regla de las descripciones impropias (DI): $\neg\overset{1}{\bigvee} x\alpha \vdash x|\alpha = y|(y = y)$.

Se siguen de (K7), (K8) y el recíproco del teorema de deducción.

Notar que todos los casos de la regla (IG) que aparecen en las pruebas de las reglas de inferencia se aplican a variables que no están libres en las premisas. Esto quiere decir que todas ellas pueden ser usadas incluso en contextos en los que no sea lícito generalizar respecto de ciertas variables, como cuando se aplica el teorema de deducción.

2.3 Técnicas de deducción

Las reglas de inferencia que acabamos de probar son todas "razonables", es decir, consistentes con la interpretación pretendida de los signos lógicos, en la que en ningún momento nos hemos apoyado. Esto refuerza la conjetura de que razonar en $K_{\mathcal{L}}$ equivale a razonar lógicamente en el sentido usual. Ahora probaremos algunos resultados adicionales en esta línea que aproximarán aún más el razonamiento formal en $K_{\mathcal{L}}$ a la forma habitual de razonar de los matemáticos.

Uso práctico del teorema de deducción El teorema de deducción afirma esencialmente que, para demostrar una implicación $\alpha \to \beta$ podemos suponer α y tratar de llegar hasta β. Esto puede ser refinado. Para ello observemos lo siguiente:

Si $\beta_1, \ldots, \beta_m$ es una deducción a partir de la colección de premisas Γ y se cumple que $\Gamma, \beta_1, \ldots, \beta_m, \alpha \vdash \beta$ sin que en la deducción correspondiente se use (IG) respecto a variables libres en α, tenemos por el teorema de deducción que $\Gamma, \beta_1, \ldots, \beta_m \vdash \alpha \to \beta$ y, por tanto, también $\Gamma \vdash \alpha \to \beta$.

Lo nuevo aquí es que si partimos de unas premisas Γ y, en un momento dado, queremos probar $\alpha \to \beta$, para ello bastará deducir β, no sólo de α y de las premisas disponibles Γ, que es lo que afirma el teorema de deducción, sino que también podemos contar con todas las líneas anteriores $\beta_1, \ldots, \beta_m$ ya deducidas. En la práctica, lo que haremos será añadir α en la deducción y seguir razonando. A partir de este momento ya no podremos generalizar respecto de variables libres en α, pero, según lo que acabamos de observar, podremos usar las líneas anteriores a la incorporación de α incluso si en su deducción hubiéramos generalizado respecto de alguna variable libre en α. Cuando lleguemos a β podremos escribir $\alpha \to \beta$ con la garantía de que esta fórmula se sigue exclusivamente de nuestras premisas Γ.

Ahora bien, si $\alpha \to \beta$ no era nuestro objetivo final y queremos seguir razonando, habremos de tener presente que en los pasos siguientes no podremos usar ninguna línea comprendida entre α y β, pues estas líneas las hemos obtenido con α como premisa adicional. Para explicitar esto marcaremos este segmento de la deducción con una línea vertical por la izquierda. De este modo la deducción quedará como sigue:

(1)	β_1
$\vdots$	
(m)	β_m
(m+1)	α
$\vdots$	(deducción $\Gamma,\ \beta_1, \ldots, \beta_m,\ \alpha \vdash \beta$)
(k)	β
(k+1)	$\alpha \to \beta$

Como en $\beta_1, \ldots, \beta_m$ se puede generalizar respecto a cualquier variable, si en la deducción secundaria queremos usar un resultado de la forma $\vdash \gamma$ (por

ejemplo una regla de inferencia) podemos hacerlo aunque en su demostración se generalice respecto a variables libres en α, ya que en teoría podríamos haber escrito γ entre $\beta_1, \ldots, \beta_m$ antes de suponer α. Naturalmente dentro de una deducción secundaria puede repetirse este proceso cuantas veces convenga.

Observemos que esto se ajusta exactamente al modo de proceder del matemático. Por ejemplo, un esbozo de un caso concreto sería:

(1)	—	
⋮		
(7)	$n = m^2$	(m y n son números enteros)
(8)	$3 \mid n$	(Hipótesis)
(9)	$3 \mid m^2$	
(10)	$3 \mid m$	(porque 3 es primo)
(11)	$9 \mid m^2$	
(12)	$9 \mid n$	
(13)	$3 \mid n \to 9 \mid n$	
⋮		

A ningún matemático se le ocurriría usar luego la línea $3 \mid m$, pues no tenemos que 3 divida a m. Eso era cierto bajo la hipótesis (provisional) de que $3 \mid n$.

Demostración por reducción al absurdo Un caso particular de lo visto en el apartado anterior es aquel en que $\alpha \equiv \neg\gamma$ y $\beta \equiv \delta \wedge \neg\delta$. En este caso terminaríamos introduciendo en la deducción $\neg\gamma \to \delta \wedge \neg\delta$ y, como por (C) tenemos que $\delta \wedge \neg\delta \to \gamma$, por (MB) concluiríamos $\neg\gamma \to \gamma$, o sea, $\gamma \vee \gamma$, y por (ED) llegamos a γ.

En definitiva que, si suponiendo $\neg\gamma$ se deduce $\delta \wedge \neg\delta$, podemos concluir γ. La deducción quedará así:

(1)	β_1
⋮	
(m)	β_m
(m+1)	$\neg\gamma$
⋮	(deducción $\Gamma,\ \beta_1, \ldots, \beta_m,\ \neg\gamma \vdash \delta \wedge \neg\delta$)
(k)	$\delta \wedge \neg\delta$
(k+1)	γ
⋮	

Análogamente, suponiendo γ y deduciendo $\delta \wedge \neg\delta$ se concluye $\neg\gamma$. Vemos que esto es exactamente lo que hace el matemático al razonar por reducción al absurdo. Ahora sabemos que este tipo de razonamiento es válido en $K_{\mathcal{L}}$.

Algunas equivalencias Para justificar la eliminación del particularizador en el apartado siguiente vamos a necesitar un hecho sencillo que presentamos aquí,

acompañado de otros similares que también nos harán falta más adelante.

Teorema 2.5 *Se cumplen los siguientes hechos:*

a) Si y no está en α, entonces $\vdash \bigwedge x\alpha \leftrightarrow \bigwedge y\ \mathtt{S}^y_x\alpha$.

b) Si y no está en α, entonces $\vdash \bigvee x\alpha \leftrightarrow \bigvee y\ \mathtt{S}^y_x\alpha$.

c) Si $y \not\equiv x$ no está en α, entonces $\vdash \overset{1}{\bigvee} x\alpha \leftrightarrow \bigvee y \bigwedge x(\alpha \leftrightarrow x = y)$.

DEMOSTRACIÓN: Para probar a), suponemos $\bigwedge\alpha$, pasamos a $\mathtt{S}^y_x\alpha$ por (EG) y pasamos a $\bigwedge y\mathtt{S}^y_x\alpha$ por (IG). Esto nos da una implicación. La otra se prueba igualmente usando que $\mathtt{S}^x_y\mathtt{S}^y_x\alpha \equiv \alpha$ (teorema 1.6).

b) sale de a) aplicado a $\neg\alpha$.

Para probar c) observamos que, por definición,

$$\overset{1}{\bigvee} x\,\alpha \equiv \bigvee z \bigwedge x(\alpha \leftrightarrow x = z),$$

donde z es la variable de menor índice que no está en α y $z \not\equiv x$. Lo que queremos probar es que si cambiamos z por y obtenemos una fórmula equivalente, y esto es un caso particular de a). ∎

Eliminación del particularizador Cuando un matemático dispone en un razonamiento de una fórmula de tipo $\bigvee x\alpha$, considera correcto escribir a continuación "tomemos un y que cumpla α". Este paso se llama "eliminación del particularizador". Vamos a ver que la eliminación del particularizador es legítima en $K_\mathcal{L}$ siempre y cuando impongamos dos restricciones que el matemático se impone instintivamente. Ante todo, planteemos la situación en $K_\mathcal{L}$. Suponemos que estamos deduciendo a partir de unas premisas Γ y tenemos ya escritas las líneas $\beta_1, \ldots, \beta_m$, entre las cuales se encuentra $\bigvee x\alpha$. Nuestra intención es incorporar a la deducción la fórmula $\mathtt{S}^y_x\alpha$, pero de tal modo que las líneas que escribamos a continuación sigan siendo consecuencia de nuestras premisas. Como ya hemos dicho, esto exige dos restricciones.

— La primera es que la variable y ha de ser distinta de cualquier variable que aparezca anteriormente en la deducción. El matemático tiene esto en cuenta porque sabe que no puede suponer a priori que el y que cumple α coincida con ninguno de los otros objetos de los que está hablando.

— La segunda es que a partir del momento que escribamos $\mathtt{S}^y_x\alpha$ ya no podemos generalizar respecto a ninguna variable libre en esta fórmula. Esto es claramente necesario en el caso de la propia variable y: el matemático es consciente de que esta "y" no es una "y" arbitraria sino una "y" particular, por lo que nada de lo que obtenga con ella será válido para todo y, sino sólo para un cierto y. No obstante, hay que insistir en que la restricción afecta a todas las variables que estén libres en $\mathtt{S}^y_x\alpha$.

En efecto, de no ser así, a partir de la fórmula $\bigwedge z \bigvee x\ x = z$ podríamos pasar a $\bigvee x\ x = z$ por (EG), de aquí a $x = z$ por (EP), de aquí a $\bigwedge z\ x = z$ (esto es incorrecto) y de aquí a $\bigvee x \bigwedge z\ x = z$, mientras que todo matemático tiene claro que de la primera sentencia (trivialmente cierta) no se deduce lógicamente la última (que afirma que sólo existe un objeto).

El teorema siguiente demuestra que si cumplimos las dos restricciones indicadas la eliminación del particularizador es legítima.

Teorema 2.6 (Eliminación del particularizador) (EP) *Sea $\beta_1, \ldots, \beta_m$ una deducción con premisas en Γ. Supongamos que $\beta_k \equiv \bigvee x\alpha$ para cierto k y que*

$$\Gamma,\ \beta_1, \ldots, \beta_m,\ \mathtt{S}^y_x\alpha \vdash \beta,$$

donde y no está en α o $y \equiv x$, y no está libre en β y además en la deducción no se generaliza respecto a variables libres en $\mathtt{S}^y_x\alpha$. Entonces $\Gamma \vdash \beta$.

Nota En el teorema exigimos únicamente que y sea distinta de cualquier variable de α, salvo que sea la propia x, mientras que antes hemos dicho que hace falta exigir que y sea distinta de cualquier variable anterior. En realidad las hipótesis del teorema no son suficientes en la práctica por el siguiente motivo: en la deducción de β a partir de Γ se generaliza respecto de la variable y, por lo que y no puede ser ninguna variable respecto a la cual no podamos generalizar. De este modo, si nuestra deducción contiene algunas variables libres $y_1, \ldots, y_r$ que procedan de aplicaciones previas de (EP) o de hipótesis a las que pretendemos aplicar el teorema de deducción (o una reducción al absurdo) la variable y que introduzcamos ha de ser distinta de todas ellas.

Demostración: Por la propiedad b) del teorema 2.5 tenemos $\beta_k \vdash \bigvee y\ \mathtt{S}^y_x\alpha$ y, por lo tanto, $\Gamma \vdash \bigvee y\ \mathtt{S}^y_x\alpha$. Por el teorema de deducción

$$\Gamma,\ \beta_1, \ldots, \beta_m \vdash \mathtt{S}^y_x\alpha \to \beta.$$

Veamos que $\mathtt{S}^y_x\alpha \to \beta,\ \bigvee y\ \mathtt{S}^y_x\alpha \vdash \beta$.

(1)	$\neg\beta$	(hipótesis)
(2)	$\mathtt{S}^y_x\alpha \to \beta$	(premisa)
(3)	$\neg\mathtt{S}^y_x\alpha$	(MT 1, 2)
(4)	$\bigwedge y\ \neg\mathtt{S}^y_x\alpha$	(IG 3)
(5)	$\neg\bigvee y\ \mathtt{S}^y_x\alpha$	(NP 4)
(6)	$\bigvee y\ \mathtt{S}^y_x\alpha$	(premisa)
(7)	$\bigvee y\ \mathtt{S}^y_x\alpha \wedge \neg\bigvee y\ \mathtt{S}^y_x\alpha$	(IC 5, 6) (contradicción)
(8)	β	

De aquí que $\Gamma,\ \beta_1, \ldots, \beta_m \vdash \beta$, luego $\Gamma \vdash \beta$. ∎

En realidad, si γ es una fórmula posterior a una aplicación de (EP) que no tiene libre a la variable y, resulta que γ es deducible a partir de Γ, y también lo serán las que se obtengan al generalizar respecto a cualquier variable, es decir, que en la práctica sí podemos generalizar sobre fórmulas que no tengan libre a y. Notemos que en cualquier momento podemos pasar a una fórmula que no tenga libre a y mediante la regla (IP).

Un ejemplo de demostración Veamos un ejemplo que ilustra las técnicas que hemos introducido hasta ahora. Vamos a probar que hay una persona en el mundo de manera que si ella vive hasta los cien años, toda la humanidad vivirá hasta los cien años, o sea $\bigvee y(Cy \to \bigwedge xCx)$, donde Cx significa "x vivirá hasta los cien años"

El argumento es el siguiente: Si toda la humanidad va a vivir hasta los cien años, cualquier persona sirve; si por el contrario alguien no va a vivir tanto, él cumple lo pedido, pues si él vive cien años (falso) todos viviremos cien años.

(1)	$\bigwedge xCx$	(hipótesis)
(2)	Cy	(hipótesis)
(3)	$\bigwedge xCx$	(R 1)
(4)	$Cy \to \bigwedge xCx$	
(5)	$\bigvee y(Cy \to \bigwedge xCx)$	(IP 4)
(6)	$\bigwedge xCx \to \bigvee y(Cy \to \bigwedge xCx)$	
(7)	$\neg\bigwedge xCx$	(hipótesis)
(8)	$\bigvee x\neg Cx$	(NG 7)
(9)	$\neg Cy$	(EP 8)
(10)	Cy	(hipótesis)
(11)	$Cy \wedge \neg Cy$	(IC 9, 10)
(12)	$\bigwedge xCx$	(C, 11)
(13)	$Cy \to \bigwedge xCx$	
(14)	$\bigvee y(Cy \to \bigwedge xCx)$	(IP 14)
(15)	$\neg\bigwedge xCx \to \bigvee y(Cy \to \bigwedge xCx)$	
(16)	$\bigwedge xCx \vee \neg\bigwedge xCx \to \bigvee y(Cy \to \bigwedge xCx)$	(Dil 6, 15)
(17)	$\bigwedge xCx \vee \neg\bigwedge xCx$	(TND)
(18)	$\bigvee y(Cy \to \bigwedge xCx)$	(MP 16, 17)

Nótese el uso de (EP) en (9). La variable introducida y tiene que cumplir la condición de no estar libre en ninguna línea anterior. En realidad está libre en (2) y (4), por ejemplo, pero esas líneas pertenecen a una deducción secundaria y no pueden ser usadas ya. A partir de (9) no se puede generalizar respecto de y, aunque, como en (14) ligamos y mediante (IP), si la fórmula a la que hemos aplicado (EP) (es decir, $\neg Cy$) hubiera tenido más variables libres y la deducción hubiera seguido, después de (14) ya hubiera sido lícito generalizar respecto a estas otras variables en fórmulas en las que y estuviera ligada con el cuantificador existencial.

Puede parecer sospechosa la línea (11), pero en realidad es un paso en la prueba de que si $\neg Cy$ entonces $Cy \to \bigwedge xCx$. De hecho, si $\neg Cy$ entonces Cy implica cualquier cosa.

Ejercicio: Consideremos las sentencias $\alpha \equiv \bigwedge x(Abx \leftrightarrow Px \wedge \neg Axx)$, $\beta \equiv Pb$, donde A es un relator diádico, P un relator monádico y b una constante. (Estas sentencias pueden interpretarse así: "El barbero afeita a todos los habitantes del pueblo que no se afeitan a sí mismos y sólo a ellos", "El barbero es un habitante del pueblo".) Demostrar que

$$\alpha,\ \beta \vdash b \neq b.$$

(Es decir, deducir de las premisas que "El barbero no es el barbero".)

Más teoremas lógicos Terminamos la sección con algunos teoremas lógicos de interés. En este primer teorema recogemos las propiedades algebraicas del cálculo deductivo (asociatividad y conmutatividad del conjuntor y el disyuntor, etc.) Dejamos las pruebas a cargo del lector.

Teorema 2.7 *Las fórmulas siguientes son teoremas lógicos:*

a) $\alpha \wedge \beta \leftrightarrow \beta \wedge \alpha, \qquad \alpha \vee \beta \leftrightarrow \beta \vee \alpha,$

b) $(\alpha \wedge \beta) \wedge \gamma \leftrightarrow \alpha \wedge (\beta \wedge \gamma), \qquad (\alpha \vee \beta) \vee \gamma \leftrightarrow \alpha \vee (\beta \vee \gamma),$

c) $\alpha \wedge \alpha \leftrightarrow \alpha, \qquad \alpha \vee \alpha \leftrightarrow \alpha,$

d) $\alpha \wedge (\beta \vee \gamma) \leftrightarrow (\alpha \wedge \beta) \vee (\alpha \wedge \gamma), \qquad \alpha \vee (\beta \wedge \gamma) \leftrightarrow (\alpha \vee \beta) \wedge (\alpha \vee \gamma).$

El teorema siguiente contiene los resultados que justifican que si en un contexto dado contamos con una coimplicación $\alpha \leftrightarrow \alpha'$ entonces las fórmulas α y α' son equivalentes, en el sentido de que todo lo que vale para una vale para la otra. Las demostraciones no presentan ninguna dificultad, así que las dejamos una vez más a cargo del lector.

Teorema 2.8 *Las fórmulas siguientes son teoremas lógicos:*

a) $(\alpha \leftrightarrow \alpha') \leftrightarrow (\neg\alpha \leftrightarrow \neg\alpha'),$

b) $\big((\alpha \leftrightarrow \alpha') \wedge (\beta \leftrightarrow \beta')\big) \rightarrow \big((\alpha \rightarrow \beta) \leftrightarrow (\alpha' \rightarrow \beta')\big),$

c) $\big((\alpha \leftrightarrow \alpha') \wedge (\beta \leftrightarrow \beta')\big) \rightarrow \big((\alpha \vee \beta) \leftrightarrow (\alpha' \vee \beta')\big),$

d) $\big((\alpha \leftrightarrow \alpha') \wedge (\beta \leftrightarrow \beta')\big) \rightarrow \big((\alpha \wedge \beta) \leftrightarrow (\alpha' \wedge \beta')\big),$

e) $\big((\alpha \leftrightarrow \alpha') \wedge (\beta \leftrightarrow \beta')\big) \rightarrow \big((\alpha \leftrightarrow \beta) \leftrightarrow (\alpha' \leftrightarrow \beta')\big),$

f) $\bigwedge x(\alpha \leftrightarrow \beta) \rightarrow (\bigwedge x\ \alpha \leftrightarrow \bigwedge x\ \beta),$

g) $\bigwedge x(\alpha \leftrightarrow \beta) \rightarrow (\bigvee x\ \alpha \leftrightarrow \bigvee x\ \beta),$

h) $\bigwedge x(\alpha \leftrightarrow \beta) \rightarrow (\overset{1}{\bigvee} x\ \alpha \leftrightarrow \overset{1}{\bigvee} x\ \beta).$

La definición que hemos dado de unicidad tiene la ventaja de que no involucra sustituciones, pero en muchas ocasiones es más útil la equivalencia siguiente:

Teorema 2.9 *Si la variable y no está en la fórmula $\alpha(x)$ y $x \not\equiv y$, entonces*

$$\vdash \overset{1}{\bigvee} x\alpha(x) \leftrightarrow \bigvee x\, \alpha \wedge \bigwedge xy(\alpha(x) \wedge \alpha(y) \rightarrow x = y).$$

Nota Con la notación que estamos empleando habitualmente para las sustituciones la última parte del teorema se escribe $\bigwedge xy(\alpha \wedge S^y_x\alpha \to x = y)$. Hemos empleado la notación alternativa en el enunciado porque así es como este resultado suele aparecer en la práctica.

DEMOSTRACIÓN: Veamos primero una implicación (por abreviar aplicaremos varias reglas de inferencia simultáneamente).

(1)	$\bigvee x\alpha \wedge \bigwedge xy(\alpha \wedge S^y_x\alpha \to x = y)$	(Hipótesis)
(2)	$\mathtt{S}^z_x\alpha$	(EC, EP 1)
(3)	α	(Hipótesis)
(4)	$\alpha \wedge \mathtt{S}^z_x\alpha \to x = z$	(EC, EG, 1)
(5)	$x = z$	(IC 3, 2; MP 4)
(6)	$\alpha \to x = z$	
(7)	$x = z$	(Hipótesis)
(8)	α	(ETI 2)
(9)	$x = z \to \alpha$	
(10)	$\alpha \leftrightarrow x = z$	(IB 6, 9)
(11)	$\bigwedge x(\alpha \leftrightarrow x = z)$	(IG 10)
(12)	$\bigvee y\bigwedge x(\alpha \leftrightarrow x = y)$	(IP 11)
(13)	$\overset{1}{\bigvee} x\alpha$	(Teorema 2.5)
(14)	$\bigvee x\alpha \wedge \bigwedge xy(\alpha \wedge S^y_x\alpha \to x = y) \to \overset{1}{\bigvee} x\alpha$	

La otra implicación es similar:

(1)	$\overset{1}{\bigvee} x\alpha$	(Hipótesis)
(2)	$\bigvee y\bigwedge x(\alpha \leftrightarrow x = y)$	(Teorema 2.5)
(3)	$\bigwedge x(\alpha \leftrightarrow x = z)$	(EP 2)
(4)	$\mathtt{S}^z_x\alpha \leftrightarrow z = z$	(EG 3)
(5)	$\mathtt{S}^z_x\alpha$	(EB, I, MP, 4)
(6)	$\bigvee x\,\alpha$	(IP 5)
(7)	$\alpha \wedge \mathtt{S}^y_x\alpha$	(Hipótesis)
(8)	$\alpha \to x = z$	(EG 3)
(9)	$\mathtt{S}^y_x\alpha \to y = z$	(EG 3)
(10)	$x = y$	(EC 7, MP 8, MP 9, SI, TI)
(11)	$\alpha \wedge \mathtt{S}^y_x\alpha \to x = y$	
(12)	$\bigwedge xy(\alpha \wedge \mathtt{S}^y_x\alpha \to x = y)$	(IG 11)
(13)	$\overset{1}{\bigvee} x\alpha \to \bigvee x\alpha \wedge \bigwedge xy(\alpha \wedge S^y_x\alpha \to x = y)$	

■

Ejercicio: Comprobar que todas las generalizaciones en la prueba anterior son correctas.

2.4 Teorías axiomáticas

Aunque todavía no lo hemos justificado, los resultados que hemos visto hasta aquí hacen plausible la conjetura de la noción de deducción lógica (en el sentido

de deducción en $K_{\mathfrak{L}}$ coincide con la noción informal de deducción que emplea el matemático. Sin embargo, los teoremas matemáticos no son los teoremas de $K_{\mathfrak{L}}$. En sus demostraciones, los matemáticos aceptan afirmaciones específicas sobre los números, los conjuntos, etc. que no son teoremas lógicos. Más claramente, en $K_{\mathfrak{L}}$ sólo se pueden demostrar trivialidades como que $\alpha \to \alpha$. Si queremos teoremas más profundos hemos de añadir axiomas más profundos. Esto nos lleva a las teorías axiomáticas:

Definición 2.10 Una *teoría axiomática* (de primer orden) sobre un lenguaje formal $\mathfrak{L}$ es un sistema deductivo formal T sobre $\mathfrak{L}$ cuyos axiomas contengan a los de $K_{\mathfrak{L}}$ y cuyas reglas de inferencia sean las de $K_{\mathfrak{L}}$.

En estas condiciones, los axiomas de $K_{\mathfrak{L}}$ se llaman *axiomas lógicos* de T, mientras que los axiomas de T que no sean axiomas de $K_{\mathfrak{L}}$ se llaman *axiomas propios* de T. En la práctica, cuando hablemos de los axiomas de una teoría axiomática se sobrentenderá, salvo que se indique lo contrario, que nos referimos a sus axiomas propios.

Observemos que si Γ es la colección de los axiomas (propios) de una teoría T y α es cualquier fórmula de $\mathfrak{L}$, entonces

$$\underset{T}{\vdash} \alpha \qquad \text{syss} \qquad \Gamma \vdash \alpha.$$

En efecto, una sucesión de fórmulas de $\mathfrak{L}$ es una demostración en T si y sólo si es una deducción en $K_{\mathfrak{L}}$ a partir de Γ. En un caso las fórmulas de Γ se consideran como axiomas y en otro como premisas.

Por ello, todos los resultados que conocemos sobre deducciones en $K_{\mathfrak{L}}$ son válidos inmediatamente para cualquier teoría axiomática.

Ahora podemos decir que nuestro objetivo es encontrar una teoría axiomática cuyos teoremas sean precisamente los teoremas que aceptan como tales los matemáticos. Veremos que no hay una sola. A estas teorías capaces de formalizar toda la matemática se las llama *teorías axiomáticas de conjuntos.*

La aritmética de primer orden Veamos un ejemplo de teoría axiomática. Se trata de la teoría más sencilla no trivial con contenido matemático. Aunque no es, ni mucho menos, una teoría de conjuntos, desde un punto de vista formal no hay ninguna diferencia cualitativa, por lo que nos servirá como ilustración.

Definición 2.11 El *lenguaje de la aritmética de primer orden* es un[1] lenguaje formal (con descriptor) cuyos signos eventuales son una constante 0, un funtor monádico S y dos funtores diádicos $+$ y $\cdot$.

[1]La única razón por la que decimos "un" y no "el" es porque no precisamos la forma concreta de los signos del lenguaje. Si entendemos que dicha forma es irrelevante, entonces el lenguaje que estamos definiendo es único.

En la práctica escribiremos

$$t' \equiv St, \qquad t_1 + t_2 \equiv +t_1t_2, \qquad t_1 \cdot t_2 \equiv \cdot t_1t_2.$$

La *aritmética de primer orden* o *aritmética de Peano* (de primer orden) es la teoría axiomática $\mathcal{P}$ sobre el lenguaje que acabamos de introducir y cuyos axiomas son los siguientes:

P1 $\bigwedge x\ x' \neq 0$,

P2 $\bigwedge xy(x' = y' \to x = y)$,

P3 $\bigwedge x(x + 0 = x)$,

P4 $\bigwedge xy(x + y' = (x + y)')$,

P5 $\bigwedge x(x \cdot 0 = 0)$,

P6 $\bigwedge xy(x \cdot y' = x \cdot y + x)$,

P7 $\alpha(0) \wedge \bigwedge x(\alpha(x) \to \alpha(x')) \to \bigwedge x\alpha(x)$, para toda fórmula $\alpha(x)$ que tenga libre la variable x (no necesariamente la única).

Observemos que P7 no es un axioma, sino un esquema axiomático que determina infinitos axiomas, uno para cada fórmula α posible.

Notas La aritmética de Peano es un prototipo de teoría axiomática formal. Nos permite hablar con rigor sobre los números naturales sin especificar en ningún momento lo que son. Así, podemos pensar que P1 significa que "el 0 no es el siguiente de ningún número natural", pero, "oficialmente", nunca hemos dicho que 0 signifique "cero" ni, mucho menos, que "$\bigwedge x$" signifique "para todo número natural". "Oficialmente" sólo podemos decir que P1 es una fórmula de un lenguaje formal, lo que no quiere decir otra cosa sino que satisface la definición de fórmula. Similarmente, podemos afirmar que la sentencia

$$\bigwedge xyz((x + y) + z = x + (y + z))$$

es un teorema de $\mathcal{P}$, y con ello no queremos decir que la suma de números naturales sea asociativa, sino únicamente que es posible construir una sucesión finita de fórmulas que acaba con la anterior y que satisface nuestra definición de deducción. Vamos a esbozarla: Abreviaremos

$$\alpha(z) \equiv (x + y) + z = x + (y + z).$$

Demostramos $\alpha(0)$ usando P3 y $\alpha(z) \to \alpha(z')$ usando P4, de aquí pasamos a $\bigwedge z(\alpha(z) \to \alpha(z'))$ por (IG), de aquí a $\bigwedge z\,\alpha(z)$ por P7 y de aquí $\bigwedge xyz\,\alpha(z)$ por (IG), que es la sentencia buscada.

Este cuidado por hablar de los números naturales sin reconocer en ningún momento que estamos hablando de números naturales es superfluo en este caso

concreto, pero cuando queramos hablar de conjuntos con la generalidad con la que hablan de ellos los matemáticos será crucial que podamos hacerlo sin necesidad de aclarar qué debemos entender por "conjunto".

Cuando trabajamos con teorías axiomáticas como $\mathcal{P}$, que nos permiten hablar sobre los números naturales, hemos de distinguir entre los números naturales metamatemáticos y los números naturales en los que pensamos cuando "oficialmente" no estamos hablando de nada. Los razonamientos siguientes mostrarán la importancia de este punto.

En el lenguaje de la aritmética de Peano podemos considerar la sucesión de designadores

$$0, \quad 0', \quad 0'', \quad 0''', \quad 0'''', \ldots$$

Conviene representarlos por $0^{(0)}$, $0^{(1)}$, $0^{(2)}$, $0^{(3)}$, $0^{(4)}$, ... de modo que, en general, para cada natural (metamatemático) n, representaremos por $0^{(n)}$ al designador formado por 0 seguido de n aplicaciones del funtor "siguiente". A estos designadores los llamaremos *numerales.*

Es crucial comprender que en "$0^{(n)}$" la "n" es una variable metamatemática (un pronombre indefinido castellano que se refiere a un número natural arbitrario), pero no es una variable del lenguaje formal de la aritmética. Del mismo modo que en 0, $0'$, $0''$, etc. no hay variables libres, ni aparecerá ninguna variable por más comitas que añadamos, en $0^{(n)}$ no hay ninguna variable libre, lo que hay es una constante y n funtores, pero ninguna variable. En particular, es un sinsentido escribir

$$\bigwedge mn\ 0^{(m)} + 0^{(n)} = 0^{(n)} + 0^{(m)}.$$

Si tratamos de interpretar "eso", el cuantificador $\bigwedge$ nos obliga a sobrentender —como hemos hecho hasta ahora muchas veces— que "m" y "n" denotan dos variables de $\mathcal{L}$, como podrían ser $m \equiv x_5$ y $n \equiv x_8$, pero eso nos obligaría a interpretar el "término" $0^{(x_5)}$, y esto no está definido: sabemos lo que es 0, o 0 con una comita, o 0 con dos comitas, o, en general, 0 con n comitas, donde n es un número de comitas, pero nunca hemos definido 0 con x_5 comitas, donde x_5 no es un número, sino una variable.

Lo que sí tiene sentido es el metateorema siguiente:

Teorema Para todo par de números naturales m y n, se cumple

$$\underset{\mathcal{P}}{\vdash} 0^{(m)} + 0^{(n)} = 0^{(n)} + 0^{(m)}. \tag{2.1}$$

Esto es un esquema teoremático, que afirma que las infinitas sentencias que se obtienen sustituyendo m y n por números naturales determinados son, todas ellas, teoremas de $\mathcal{P}$.

Podemos probarlo de dos formas: una es demostrar algo más general, a saber, que

$$\underset{\mathcal{P}}{\vdash} \bigwedge xy\ x + y = y + x \tag{2.2}$$

Una vez establecido esto, cada caso particular de (2.1) se sigue de aquí por la regla de eliminación del generalizador aplicada dos veces, a los designadores $0^{(m)}$ y $0^{(n)}$.

Ejercicio: Demostrar en $\mathcal{P}$ la sentencia $\bigwedge xy(x + y' = x' + y)$. Usar esto para probar la sentencia (2.2).

La otra alternativa es demostrar el metateorema directamente o, mejor aún, deducirlo de este otro metateorema:

Teorema Para cada par de números naturales m y n se cumple

$$\underset{\mathcal{P}}{\vdash} 0^{(m)} + 0^{(n)} = 0^{(m+n)}. \tag{2.3}$$

Aquí hemos de distinguir entre el funtor $+$, que aparece en el miembro izquierdo, de la suma de números naturales $+$, que aparece a la derecha. Este esquema teoremático no es sino la "tabla de sumar". En efecto, afirma que cualquier suma bien calculada, como pueda ser $2+3=5$, es un teorema de $\mathcal{P}$ o, mejor dicho, se corresponde con un teorema de $\mathcal{P}$, en este caso con la sentencia $0'' + 0''' = 0''''$ o, más brevemente, $0^{(2)} + 0^{(3)} = 0^{(5)}$.

Esto se demuestra fácilmente por inducción (metamatemática) sobre n. En efecto, para $n = 0$ lo que hay que probar es

$$\underset{\mathcal{P}}{\vdash} 0^{(m)} + 0 = 0^{(m)},$$

lo cual se sigue de P3, y, supuesto cierto para m, es decir, admitiendo que tenemos una demostración de $0^{(m)} + 0^{(n)} = 0^{(m+n)}$, podemos prolongarla como sigue: eliminando el generalizador en la sentencia $\bigwedge x\ x' = x'$ (que claramente es un teorema lógico) obtenemos $(0^{(m)} + 0^{(n)})' = 0^{(m+n)\prime}$. Está sentencia es idéntica a $(0^{(m)} + 0^{(n)})' = 0^{(m+n+1)}$ y aplicando P4 (y este mismo hecho otra vez) obtenemos $0^{(m)} + 0^{(n+1)} = 0^{(m+n+1)}$, como queríamos probar. Observemos que en esta prueba no hemos usado el axioma P7. Por ello no se trata de una inducción matemática. Sino metamatemática. Lo que acabamos de razonar no es riguroso porque se apoye en unos axiomas prefijados, sino porque es concluyente, en el sentido de que a partir de este argumento cualquiera puede diseñar sin dificultad un algoritmo que nos proporcione una demostración formal de (2.3) para cualquier par de números naturales dados.

Una vez probado (2.3) es fácil deducir (2.1). Basta observar que, dados m y n, sabemos probar

$$\underset{\mathcal{P}}{\vdash} 0^{(m)} + 0^{(n)} = 0^{(m+n)} \wedge 0^{(n)} + 0^{(m)} = 0^{(n+m)},$$

pero como $m + n = n + m$, sucede que $0^{(m+n)} \equiv 0^{(n+m)}$ (ambos designadores son la constante 0 precedida del mismo número de funtores), luego (2.1) se sigue de la simetría y transitividad de la igualdad.

El lector que encuentre inadmisible que usemos la conmutatividad de la suma de números naturales —lo hemos hecho al apoyarnos en que $m + n = n + m$—

para probar la conmutatividad de la suma de números naturales (2.1), debería convencerse de que probar (2.1) no es probar la conmutatividad de los números naturales. Es probar que una cierta sentencia es un teorema de una cierta teoría axiomática. Algo muy distinto. Uno puede saber que la suma de números naturales es conmutativa y no saber que (2.1) es demostrable en $\mathcal{P}$.

En el capítulo VI sistematizaremos estos hechos. Los hemos anticipado aquí para que el lector pueda juzgar si debería meditar más sobre todo lo visto hasta ahora antes de seguir adelante. La consecuencia más importante que debemos extraer de estos hechos es que la formalización de la aritmética no nos exime de trabajar con los números naturales en términos metamatemáticos no formales.

Definiciones formales Cuando definamos el lenguaje y los axiomas de la teoría de conjuntos tendremos definido con todo rigor la noción de demostración matemática, pero nos falta todavía interpretar desde un punto de vista lógico lo que hace el matemático cuando introduce una definición.

Supongamos, por ejemplo que en la teoría $\mathcal{P}$ queremos definir la relación de orden en los números naturales. Una forma de hacerlo de acuerdo con los hábitos del matemático sería mediante:

> Definición: Diremos que un número natural x es menor o igual que otro y, en signos, $x \leq y$, si se cumple $\bigvee z\ x + z = y$.

Hay básicamente dos formas de interpretar esto. Una es considerar que estamos añadiendo un nuevo relator diádico al lenguaje de $\mathcal{P}$, de modo que, visto así, una definición matemática en una teoría T consta de dos partes:

— Añadir un nuevo relator al lenguaje de T, en este caso el relator $\leq$,

— Añadir un nuevo axioma a T, en este caso la fórmula

$$x \leq y \leftrightarrow \bigvee z\ x + z = y.$$

En efecto, la definición se comporta como un axioma: es una afirmación que puede usarse en las demostraciones y a la que no se le exige una demostración.

Este punto de vista tiene varios inconvenientes. Por lo pronto hace que no exista una teoría fija, sino que cada vez que añadimos una definición estamos pasando de una teoría a otra distinta. A su vez esto exige justificar que no importa el orden en que damos las definiciones. Por otra parte, es necesario explicar la diferencia que hay entre los axiomas-axiomas y los axiomas-definiciones. Por ejemplo, en teoría de conjuntos no podemos equiparar la definición de espacio vectorial con el axioma de elección, y la diferencia no es únicamente que el axioma de elección no incorpora un nuevo relator a la teoría. Esencialmente, esto conlleva probar que las definiciones no añaden teoremas esenciales, en el sentido de que cualquier teorema demostrado con el apoyo de una definición es equivalente a un teorema que no requiere en su enunciado el concepto definido y que a su vez puede ser probado sin ayuda de la definición. Afortunadamente todos estos hechos resultan triviales si concebimos las definiciones de otra forma equivalente.

Para nosotros, las definiciones no serán más que abreviaturas de fórmulas. Así, en el ejemplo que discutíamos, la definición de la relación de orden se reduce a establecer que

$$x \leq y \equiv \bigvee z\ x + z = y,$$

es decir, que $x \leq y$ no será sino una forma abreviada de referirnos a la fórmula de la derecha, exactamente igual que hemos convenido que $\alpha \vee \beta$ no es más que una forma abreviada de referirnos a $\neg\alpha \to \beta$. En particular, $x \leq y$ no consta de un relator diádico y dos variables, sino que es una fórmula de longitud 9, cuyo primer signo es el negador, su segundo signo el cuantificador universal, etc.

Concibiendo así las definiciones trabajamos siempre en la misma teoría y es fácil ver que todas las cuestiones planteadas antes se vuelven triviales. Esto sirve para cualquier definición de tipo relacional, es decir, que introduzca una fórmula como $x \leq y$. Más delicada es la cuestión de las definiciones funtoriales, es decir, las que introducen términos. Por ejemplo, cualquier matemático admite esto como definición rigurosa en la aritmética de Peano:

> DEFINICIÓN: Dados dos números naturales x e y tales que $x \leq y$, llamaremos diferencia entre ellos, en signos $y - x$, al único número natural z tal que $y = x + z$.

(Supuesto que previamente se haya demostrado que existe tal z y es único) Desde el punto de vista de ampliación del lenguaje, podemos considerar que añadimos un nuevo funtor diádico $-$ junto con el axioma

$$\bigwedge xy(x \leq y \to y = x + (y - x)). \tag{2.4}$$

De nuevo, esto es una fuente de detalles que hay que precisar. Además te todas las cuestiones que ya hemos comentado antes, hay que especificar las condiciones que han de cumplirse para que sea lícita una definición de este tipo (en este caso el haber probado antes la existencia de la resta, pero ¿y en general?), además deja una laguna a la hora de interpretar expresiones "mal definidas" como $0^{(3)} - 0^{(5)}$.

Nosotros adoptaremos la misma postura que en el caso de las definiciones relacionales, pero hay que destacar que no podríamos hacer tal cosa si no hubiéramos incorporado el descriptor a nuestros lenguajes formales. En el ejemplo que estamos considerando, basta establecer que

$$y - x \equiv z \mid y = x + z.$$

Podemos definir esto en cualquier momento y sin ninguna hipótesis. Ahora bien, para probar en $\mathcal{P}$ el teorema (2.4) necesitamos probar primero el teorema

$$\bigwedge xy(x \leq y \to \overset{1}{\bigvee} z\ y = x + z).$$

Notemos que sólo hay que probar la unicidad, pues la existencia es la definición de $x \leq y$. Una vez probado esto, podemos probar (2.4) usando la regla de

las descripciones propias: suponemos $x \leq y$, con lo que tenemos $\bigvee^1 z\ x + z = y$, luego por (DP) concluimos $S_z^{y-x} y = x + z$, es decir, la tesis de (2.4).

Las descripciones impropias no presentan ningún problema. Una vez probado que $\neg\bigvee z\ 0^{(3)} = 0^{(5)} + z$, la regla de las descripciones impropias nos da que

$$0^{(3)} - 0^{(5)} = x|x = x.$$

Aunque no tiene interés, pues las descripciones impropias no sirven para nada, podríamos precisar esto añadiendo a $\mathcal{P}$ el axioma $0 = x|x = x$, y así podríamos probar, más concretamente, que $0^{(3)} - 0^{(5)} = 0$.

Ejercicio: Definir en $\mathcal{P}$ el máximo de dos números naturales.

2.5 Descriptores

Dedicamos esta sección a estudiar con más detalle el comportamiento de los descriptores en el cálculo deductivo. Dado que el lector puede estar poco familiarizado con ellos aun si está familiarizado con la lógica, empezaremos probando un hecho elemental para familiarizarnos con su uso formal. El primero afirma que no hemos de preocuparnos por la fórmula en concreto que elegimos para definir un concepto mediante un descriptor, pues fórmulas equivalentes dan lugar a descripciones iguales.

Teorema 2.12 *Se cumple* $\vdash \bigwedge x(\alpha \leftrightarrow \beta) \rightarrow x|\alpha = x|\beta$.

Demostración: Daremos un esbozo de la prueba. Suponemos la hipótesis $\bigwedge x(\alpha \leftrightarrow \beta)$ y distinguimos dos casos: $\bigvee^1 x\alpha \vee \neg\bigvee^1 x\alpha$.

En el primer caso, el teorema 2.8 nos da $\bigvee^1 x\alpha \wedge \bigvee^1 x\beta$, y la regla de las descripciones propias nos da $\mathtt{S}_x^{x|\alpha}\alpha \wedge \mathtt{S}_x^{x|\beta}\beta$. Eliminando el generalizador en la hipótesis tenemos $\mathtt{S}_x^{x|\beta}\alpha \leftrightarrow \mathtt{S}_x^{x|\beta}\beta$, luego $\mathtt{S}_x^{x|\alpha}\alpha \wedge \mathtt{S}_x^{x|\beta}\alpha$. Aplicamos 2.9 a la unicidad en α y concluimos que $x|\alpha = x|\beta$.

En el segundo caso, el teorema 2.8 nos da $\neg\bigvee^1 x\alpha \wedge \neg\bigvee^1 x\beta$. La regla de las descripciones impropias nos lleva a que $x|\alpha = x|y = y \wedge x|\beta = y|y = y$, luego también $x|\alpha = x|\beta$. ∎

Nota El lector no debe estudiar esta demostración y otras similares como si estudiara un teorema matemático, en el sentido de que su preocupación principal no ha de ser convencerse de que el resultado es cierto — en algún sentido de la palabra "cierto" del que no hemos hablado—, sino convencerse de que es demostrable en $K_{\mathcal{L}}$. Así, deberá darse por satisfecho cuando se convenza de que el esbozo de prueba que hemos dado puede desarrollarse hasta convertirse en una demostración lógica que no use nada cuyo uso no esté justificado en $K_{\mathcal{L}}$.

En la sección anterior hemos visto cómo el descriptor es necesario para introducir nuevos conceptos a una teoría sin necesidad de extender su lenguaje y

añadir nuevos axiomas. Como hemos comentado allí, esto evita muchos teoremas prolijos sobre las definiciones, pero por otra parte hemos de observar que el descriptor complica notablemente la gramática de los lenguajes formales. En efecto, en un lenguaje sin descriptor podemos definir los términos a partir de las variables, las constantes y los funtores y luego definir las fórmulas a partir de los términos. A su vez esto permite separar los términos de las fórmulas en los razonamientos inductivos, que en un lenguaje con descriptor han de hacerse conjuntamente.

Sin embargo, estos inconvenientes son sólo aparentes, pues, como vamos a ver a continuación, los descriptores pueden eliminarse casi por completo (en algunos casos podemos suprimir el "casi"). La idea es que una descripción puede sustituirse por una perífrasis: en lugar de decir,

"La reina de Inglaterra se llama Isabel"

(o sea, "el x tal que x es reina de inglaterra se llama Isabel") podemos decir

"Existe un x tal que x es reina de Inglaterra y x se llama Isabel",

y aquí ya no hay ninguna descripción. Esta solución no es completamente general, pues presupone que existe una única reina de Inglaterra. Si queremos parafrasear "El rey del país P es calvo", donde no sabemos qué país es P ni, por consiguiente, si tiene o no tiene rey, deberemos decir:

> *"O bien existe un único x tal que x es rey del país P y x es calvo, o bien no existe un único x tal que x es rey del país P y en tal caso $x|x = x$ es calvo.*

Vemos que nos queda una "descripción residual", en la forma $x|x = x$, pero al menos es siempre la misma, no depende de la fórmula que parafraseamos. Por eso hablábamos de eliminar los descriptores casi por completo.

El teorema siguiente contiene la idea central de esta discusión:

Teorema 2.13 *Si $x \not\equiv y$, se cumple*

$$\vdash y = x|\alpha \leftrightarrow \bigwedge x(\alpha \leftrightarrow x = y) \vee (\neg \overset{1}{\bigvee} x\, \alpha \wedge y = z|z = z),$$

Demostración: Esbozamos la prueba. Bajo la hipótesis $y = x|\alpha$, distinguimos dos casos, o bien $\overset{1}{\bigvee} x\ \alpha$ o bien $\neg \overset{1}{\bigvee} x\ \alpha$.

En el primer caso, por la definición de unicidad, tenemos $\bigvee y \bigwedge x(\alpha \leftrightarrow x = y)$. Eliminamos el particularizador (cambiando de variable porque y la tenemos ya libre y no podemos generalizar sobre ella), con lo que $\bigwedge x(\alpha \leftrightarrow x = z)$. Eliminando el generalizador llegamos a $S_x^{x|\alpha}\alpha \leftrightarrow (x|\alpha) = z$, pero la parte izquierda la tenemos por la regla de las descripciones propias, con lo que $z = x|\alpha$. Por hipótesis, $z = y$ y por la equivalencia de términos idénticos $\bigwedge x(\alpha \leftrightarrow x = y)$.

En el segundo caso, la regla de las descripciones impropias nos permite afirmar que $x|\alpha = z|z = z$ y por hipótesis $y = z|z = z$, lo que nos lleva a la conclusión.

Supongamos ahora el término derecho del teorema. Por la regla del dilema basta probar que ambas disyuntivas nos llevan a $y = x|\alpha$.

Si suponemos $(*)$: $\bigwedge x(\alpha \leftrightarrow x = y)$, introduciendo el particularizador $\bigvee y$ obtenemos $\overset{1}{\bigvee} x\,\alpha$, luego la regla de las descripciones propias nos da $\mathtt{S}_x^{x|\alpha}\alpha$. Por otro lado, eliminando el generalizador en $(*)$ obtenemos $\mathtt{S}_x^{x|\alpha}\alpha \leftrightarrow y = x|\alpha$, luego concluimos que $y = x|\alpha$.

Si suponemos $\neg\overset{1}{\bigvee} x\,\alpha \wedge y = z|z = z$, la regla de las descripciones impropias nos da que $x|\alpha = z|z = z$, luego concluimos igualmente que $y = x|\alpha$. ∎

Observemos que el miembro derecho de la fórmula del teorema anterior puede tener descriptores (aparte del que aparece explícitamente) si los tiene α. No obstante, una aplicación reiterada de este resultado más un pequeño truco para eliminar la descripción impropia nos permiten eliminar todos los descriptores de una fórmula. Lo probamos en el teorema siguiente:

Teorema 2.14 (Teorema de cuasieliminación de descriptores) *Sea c una constante de un lenguaje formal $\mathcal{L}$ con descriptor y sea θ una expresión de $\mathcal{L}$. Entonces*

- *Si θ es una fórmula, existe otra fórmula θ' con las mismas variables libres que θ y sin descriptores tal que $c = z|(z = z) \vdash \theta \leftrightarrow \theta'$.*
- *Si θ es un término e y no está libre en θ, existe una fórmula ϕ sin descriptores con las mismas variables libres que $y = \theta$ tal que*

$$c = z|(z = z) \vdash y = \theta \leftrightarrow \phi.$$

Demostración: Por inducción sobre la longitud de θ.

Si $\theta \equiv x_i$ sirve $\phi \equiv y = x_i$.

Si $\theta \equiv c_i$ sirve $\phi \equiv y = c_i$.

Si $\theta \equiv R_i^n t_1 \cdots t_n$, sean $y_1, \ldots, y_n$ variables que no estén en θ. Por hipótesis de inducción existen fórmulas sin descriptores $\phi_1, \ldots, \phi_n$ con las mismas variables libres que $y_i = t_i$ de manera que

$$c = z|(z = z) \vdash y_i = t_i \leftrightarrow \phi_i.$$

Se comprueba que $\theta' \equiv \bigvee y_1 \cdots y_n(\phi_1 \wedge \cdots \wedge \phi_n \wedge R_i^n y_1 \cdots y_n)$ cumple lo pedido.

Si $\theta \equiv f_i^n t_1 \cdots t_n$, sea y una variable que no esté en θ y sean $y_1, \ldots, y_n$ variables que no estén en $y = \theta$. Por hipótesis de inducción hay n fórmulas sin descriptores ϕ_i con las mismas variables libres que $y_i = t_i$ de manera que

$$c = z|(z = z) \vdash y_i = t_i \leftrightarrow \phi_i.$$

Se comprueba que $\phi \equiv \bigvee y_1 \cdots y_n(\phi_1 \wedge \cdots \wedge \phi_n \wedge y = f_i^n y_1 \cdots y_n)$ cumple lo pedido.

Si $\theta \equiv \neg\alpha$, por hipótesis de inducción existe α' sin descriptores y con las mismas variables libres que α de modo que $c = z|(z = z) \vdash \alpha \leftrightarrow \alpha'$.

Claramente $\theta' \equiv \neg\alpha'$ cumple el teorema.

Si $\theta \equiv \alpha \to \beta$, por hipótesis de inducción existen α' y β' con las mismas variables libres que α y β respectivamente y sin descriptores, tales que

$$c = z|(z = z) \vdash \alpha \leftrightarrow \alpha', \qquad c = z|(z = z) \vdash \beta \leftrightarrow \beta'$$

Claramente $\theta' \equiv \alpha' \to \beta'$ cumple el teorema.

Si $\theta \equiv \bigwedge u\alpha$, por hipótesis de inducción existe α' sin descriptores y con las mismas variables libres que α tal que $c = z|(z = z) \vdash \alpha \leftrightarrow \alpha'$.

Se comprueba que $\theta' \equiv \bigwedge u\alpha'$ cumple el teorema.

Si $\theta \equiv x|\alpha$, por hipótesis de inducción existe α' sin descriptores y con las mismas variables libres que α tal que $c = z|(z = z) \vdash \alpha \leftrightarrow \alpha'$.

Según hemos probado antes,

$$\vdash y = x|\alpha \leftrightarrow \bigwedge x(\alpha \leftrightarrow x = y) \vee (\neg \overset{1}{\bigvee} x\alpha \wedge y = z|(z = z)).$$

De aquí se sigue que

$$c = z|(z = z) \vdash y = x|\alpha \leftrightarrow \bigwedge x(\alpha \leftrightarrow x = y) \vee (\neg \overset{1}{\bigvee} x\alpha \wedge y = c),$$

de donde se concluye que $\phi \equiv \bigwedge x(\alpha \leftrightarrow x = y) \vee (\neg \overset{1}{\bigvee} x\alpha \wedge y = c)$ cumple lo pedido. ∎

De este modo, si en la aritmética de Peano añadimos el axioma $0 = x|x = x$, tal y como ya habíamos comentado en la sección anterior, tenemos que toda fórmula es equivalente en $\mathcal{P}$ a una fórmula sin descriptores. Más adelante probaremos que todo teorema sin descriptores puede demostrarse sin descriptores.

2.6 Forma prenexa

Los lenguajes formales permiten definir una noción de "complejidad" de una afirmación que resulta útil en contextos muy variados. Es frecuente que a los estudiantes les cueste asimilar la noción de límite de una función en un punto más de lo que les cuesta comprender otros conceptos del mismo nivel. Uno de los factores que influyen en ello es que empieza más o menos así: *"Para todo $\epsilon > 0$ existe un $\delta > 0$ tal que para todo $x \in \mathbb{R}$, ..."*. La dificultad no está en que haya tres cuantificadores, pues una definición que empiece con *"Para todo ϵ, para todo δ y para todo x se cumple ..."* resulta mucho más sencilla. La complejidad de la definición de límite se debe a que los tres cuantificadores se alternan: *"para todo... existe... para todo ..."*

Vamos a definir la complejidad de una fórmula en términos de la alternancia de sus cuantificadores. Para ello introducimos la noción de forma prenexa:

Definición 2.15 Se dice que una fórmula sin descriptores α de un lenguaje formal $\mathfrak{L}$ está en *forma prenexa* si $\alpha \equiv \pi\alpha_0$, donde α_0 es una fórmula sin cuantificadores y π es una sucesión finita de cuantificadores universales $\bigwedge x$ y/o existenciales $\bigvee x$. A π se le llama *prefijo* de α.

Vamos a demostrar que toda fórmula sin descriptores es lógicamente equivalente a una fórmula en forma prenexa. La prueba se basa en el teorema siguiente, que dejamos como ejercicio:

Teorema 2.16 *Se cumple:*

$$\begin{aligned}
&\vdash (\alpha \to \textstyle\bigwedge x\beta) \leftrightarrow \bigwedge x(\alpha \to \beta) && \textit{si } x \textit{ no está libre en } \alpha,\\
&\vdash (\alpha \to \textstyle\bigvee x\beta) \leftrightarrow \bigvee x(\alpha \to \beta) && \textit{si } x \textit{ no está libre en } \alpha,\\
&\vdash (\textstyle\bigwedge x\alpha \to \beta) \leftrightarrow \bigvee x(\alpha \to \beta) && \textit{si } x \textit{ no está libre en } \beta,\\
&\vdash (\textstyle\bigvee x\alpha \to \beta) \leftrightarrow \bigwedge x(\alpha \to \beta) && \textit{si } x \textit{ no está libre en } \beta,\\
&\vdash (\alpha \vee \textstyle\bigwedge x\beta) \leftrightarrow \bigwedge x(\alpha \vee \beta) && \textit{si } x \textit{ no está libre en } \alpha,\\
&\vdash (\alpha \vee \textstyle\bigvee x\beta) \leftrightarrow \bigvee x(\alpha \vee \beta) && \textit{si } x \textit{ no está libre en } \alpha,\\
&\vdash (\textstyle\bigwedge x\alpha \wedge \beta) \leftrightarrow \bigwedge x(\alpha \wedge \beta) && \textit{si } x \textit{ no está libre en } \beta,\\
&\vdash (\textstyle\bigvee x\alpha \wedge \beta) \leftrightarrow \bigvee x(\alpha \wedge \beta) && \textit{si } x \textit{ no está libre en } \beta.
\end{aligned}$$

Ahora es fácil probar:

Teorema 2.17 *Si α es una fórmula sin descriptores, existe otra fórmula β en forma prenexa con las mismas variables libres que α tal que $\vdash \alpha \leftrightarrow \beta$.*

Demostración: Lo probamos por inducción sobre la longitud de α.

Si $\alpha \equiv R_i^n t_1 \cdots t_n$, entonces ya está en forma prenexa. Tomamos $\beta \equiv \alpha$.

Si $\alpha \equiv \neg\gamma$, por hipótesis de inducción sabemos que $\vdash \gamma \leftrightarrow \pi\delta$, para cierta fórmula $\pi\delta$ en forma prenexa, luego por el teorema 2.8 tenemos $\vdash \neg\gamma \leftrightarrow \neg\pi\delta$. Aplicando (NG) y (NP) podemos "meter" el negador, y así $\vdash \neg\gamma \leftrightarrow \pi'\neg\delta$, donde π' es la sucesión de cuantificadores que resulta de cambiar cada cuantificador universal de π por uno existencial y viceversa.

Si $\alpha \equiv \gamma \to \delta$, por hipótesis de inducción $\vdash \gamma \leftrightarrow \pi\epsilon$ y $\vdash \delta \leftrightarrow \pi'\eta$. Aplicando el teorema 2.5 si es preciso, podemos suponer que las variables que liga π no están en $\pi'\eta$ y viceversa. Por el teorema 2.8 tenemos que $\vdash \alpha \leftrightarrow (\pi\epsilon \to \pi'\eta)$. Por el teorema anterior, $\vdash \alpha \leftrightarrow \pi'(\pi\epsilon \to \eta)$, y así mismo, $\vdash (\pi\epsilon \to \eta) \leftrightarrow \pi''(\epsilon \to \eta)$.

Usando (IG) llegamos a que $\vdash \pi'((\pi\epsilon \to \eta) \leftrightarrow \pi''(\epsilon \to \eta))$. Por 2.8, tenemos que $\vdash \pi'(\pi\epsilon \to \eta) \leftrightarrow \pi'\pi''(\epsilon \to \eta)$ y, por lo tanto, $\vdash \alpha \leftrightarrow \pi'\pi''(\epsilon \to \eta)$.

Si $\alpha \equiv \bigwedge x\gamma$, por hipótesis de inducción $\vdash \gamma \leftrightarrow \pi\delta$. Por (IG) tenemos que $\vdash \bigwedge x(\gamma \leftrightarrow \pi\delta)$ y por 2.8 queda $\vdash \alpha \leftrightarrow \bigwedge x\pi\delta$.

Es fácil comprobar que en cada caso las variables libres de la fórmula construida son las mismas que las de la fórmula dada. ∎

En la práctica es fácil extraer los cuantificadores de cualquier fórmula dada usando el teorema 2.16. Aquí tenemos un primer ejemplo de la necesidad de eliminar los descriptores en ciertos contextos: no es posible extraer los cuantificadores de una descripción por lo que en el teorema anterior teníamos que partir de una fórmula sin descriptores, pero ahora sabemos que, bajo hipótesis mínimas, toda fórmula es equivalente a una fórmula sin descriptores, que a su vez tiene una forma prenexa, luego cualquier fórmula tiene una forma prenexa.

Hemos probado que toda fórmula sin descriptores es lógicamente equivalente a una fórmula en forma prenexa, no necesariamente única, pero dentro de cada teoría axiomática podemos encontrar, en general, formas prenexas más simples (cuya equivalencia con la fórmula dada requiera los axiomas propios de T). La noción de "formas prenexas más simples" queda precisada por la definición siguiente:

Definición 2.18 Una fórmula α es de tipo Σ_n (Π_n) en una teoría T si admite una forma prenexa cuyo prefijo conste de n bloques de cuantificadores alternados empezando por un cuantificador existencial (universal). Las fórmulas sin cuantificadores se llaman fórmulas Δ_0.

Por ejemplo, una fórmula de tipo Σ_3 es

$$\bigvee xy \bigwedge uvw \bigvee z(x + u = y \wedge yvv = z).$$

Naturalmente, una misma fórmula puede ser de varios tipos a la vez y además, como ya hemos comentado, el tipo de una fórmula depende de la teoría en la que trabajemos. De hecho, introduciendo variables en los prefijos es fácil ver que toda fórmula Σ_n o Π_n es también Σ_m y Π_m para todo $m > n$, por lo que con esta clasificación las fórmulas quedan ordenadas en una doble jerarquía que mide su complejidad.

2.7 Consideraciones finales

El cálculo deductivo que hemos presentado aquí, salvo en lo tocante a los descriptores, que son mucho más recientes, se debe esencialmente a Frege. La teoría axiomática $\mathcal{P}$ para describir los números naturales es de Peano, si bien él trabajaba con un sistema de segundo orden, lo que le permitía definir la suma y el producto sin necesidad de tomar sus propiedades como axiomas. Esto reduce los axiomas de siete a tres, a los que hay que añadir los que afirmaban que "el cero es un número natural" y "el siguiente de un número natural es un número natural" (afirmaciones vacías en nuestra teoría), lo que nos da los cinco axiomas clásicos de Peano.

Sin embargo, tanto Frege como Peano concebían sus cálculos deductivos como un estudio explícito de las leyes del razonamiento matemático. El primero en presentar el cálculo deductivo como una teoría formal, es decir, desvinculada de todo posible significado de sus signos, fue Hilbert.

También fue Hilbert el primero en formular la distinción moderna entre matemática, y metamatemática y el primero en discutir los requisitos que debe cumplir el razonamiento metamatemático para ser fiable. Como anunciamos en la introducción, nosotros violaremos en parte los requisitos finitistas que propuso Hilbert, pero de momento aún no lo hemos hecho.

Es fácil encontrar libros que, bajo un título genérico en torno a la palabra "lógica", dedican todas sus páginas a demostrar resultados similares a los de este capítulo, pero llegando a teoremas aparentemente sofisticados y con nombres pomposos como "introducción del generalizador en el consecuente", etc. Un lector ingenuo que hojee uno de estos libros puede sacar fácilmente dos consecuencias igualmente infundadas: que la lógica es muy aburrida o que la lógica es muy interesante. Ambas serán infundadas porque esos libros no tratan de lógica. Más exactamente, la relación entre la lógica y el contenido de estos libros es la misma que entre la matemática y un cuadernillo de multiplicaciones. Esos libros presentan el cálculo deductivo con reverencia, como si nos enseñara algo, cuando la verdad es que todos sus teoremas, por sofisticada que pueda llegar a ser su presentación, son tan triviales como $\alpha \to \alpha$, son hechos que cualquier matemático ya sabe o, al menos, de los cuales se convencería en breves segundos si es que los llegara a necesitar alguna vez. Un lector cabal ha de tener claro que $K_{\mathcal{L}}$ no enseña nada. Nuestro propósito es usar $K_{\mathcal{L}}$ como herramienta para aprender hechos muy profundos sobre el razonamiento matemático, pero los teoremas de $K_{\mathcal{L}}$ en sí mismos son triviales. Esto lo probaremos en el capítulo III, mientras que en el capítulo IV veremos que todas las fórmulas triviales (en un sentido que hemos de precisar) son teoremas de $K_{\mathcal{L}}$.

Con esto podremos justificar que $K_{\mathcal{L}}$ cumple exactamente su misión, que no es la de enseñarnos lógica, sino la de contener toda la lógica que nosotros sabemos, que resulta ser, de hecho, toda la lógica. Con ello habremos reducido toda cuestión sobre la capacidad de razonamiento lógico a un análisis de ocho axiomas y dos reglas de inferencia y, añadiendo unos pocos axiomas más, tendremos bajo el microscopio toda la capacidad de razonamiento matemático. Es a partir de aquí cuando $K_{\mathcal{L}}$ nos resultará útil y cuando de verdad empezaremos a investigar la lógica. En resumen: confundir un libro de lógica con un libro de cálculo deductivo es como confundir un libro de pintura con un libro de pincel.

Capítulo III

Modelos

En los dos capítulos anteriores hemos insistido en que los lenguajes formales posibilitan una definición precisa de lo que son los teoremas de una teoría axiomática sin obligarnos en ningún momento a determinar el significado de sus fórmulas. Esto es crucial para fundamentar la matemática abstracta, pues es fácil especificar unos axiomas para la teoría de conjuntos y sería muy difícil, si no imposible, especificar qué debemos entender por un conjunto. No obstante, el hecho de que no necesitemos precisar el significado de las fórmulas de una teoría axiomática no impide que dichas fórmulas puedan tener un significado muy concreto. Así, podemos afirmar que la sentencia $\bigwedge xy\ x + y = y + x$ es un teorema de la aritmética de Peano y con ello, estrictamente, únicamente estamos diciendo que esta combinación de signos satisface la definición de fórmula y que existe una sucesión de fórmulas que acaba con ella y que satisface la definición de demostración, pero lo cierto es que cualquier matemático ve ahí algo más que el hecho de que se satisfacen unas definiciones combinatorias. Todo matemático entiende que ahí dice que la suma de números naturales es conmutativa. En este capítulo nos vamos a ocupar de la relación entre las teorías axiomáticas y sus interpretaciones posibles. Los conceptos que introduciremos para ello se deben esencialmente a Alfred Tarski.

3.1 Conceptos básicos

Para definir con rigor las noción de interpretación de una fórmula de un lenguaje formal necesitamos hablar de colecciones de objetos y de relaciones y funciones abstractas entre ellos. Esto nos pone al borde de un círculo vicioso. En efecto, las nociones generales de "conjunto", "relación" y "función" están en la base misma de la matemática abstracta, y son la causa de que no podamos fundamentar rigurosamente la matemática sin la ayuda de una teoría axiomática formal. Por otra parte, ahora afirmamos que vamos a usar estas nociones para estudiar la noción de teoría axiomática con la que queremos fundamentar estas nociones.

El círculo se rompe si tenemos presente que, aunque no podemos confiar en nuestra noción intuitiva general de conjunto, relación o función, no es menos cierto que conocemos con precisión algunas colecciones de objetos, algunas relaciones y algunas funciones concretas.

Por ejemplo, sabemos perfectamente lo que decimos cuando hablamos de la colección formada por las piezas de un juego de ajedrez. Podemos decir que consta de 32 elementos, de los cuales 16 serán piezas blancas y 16 negras, etc. Cualquier pregunta sobre las piezas de ajedrez puede ser planteada y respondida con absoluta precisión.

Del mismo modo, podemos hablar con toda precisión de la colección de todos los números naturales. Es cierto que no tenemos por qué saber la respuesta a cualquier pregunta acerca de ellos, debido a que son infinitos, pero sí sabemos qué significa cualquier afirmación sobre los números naturales. También conocemos la colección de los números pares, en el sentido de que no hay duda de qué números naturales son pares y cuáles son impares y, lo que es más importante, de que sabemos qué significa una afirmación sobre la totalidad de los números pares: significa que la cumple el 0, y el 2, y el 4, etc.

Por el contrario, no está claro qué debemos entender por una afirmación sobre la totalidad de las colecciones formadas por números naturales. Conocemos *algunas* colecciones de números naturales, pero no tenemos ninguna representación de *la totalidad* de ellas o, al menos, no una representación suficientemente precisa como para que podamos razonar sobre dicha totalidad sin el auxilio de una teoría axiomática.

Así pues, cuando hablemos de una *colección*[1] *de elementos* U entenderemos que hablamos de una de esas colecciones de objetos que conocemos con precisión, de modo que cualquier afirmación del tipo a es un elemento de U, o b no es un elemento de U y, más aún, cualquier afirmación sobre la totalidad de los elementos de U tenga un significado preciso.

No tenemos ninguna definición precisa de qué es una colección de objetos bien definida, pero sí tenemos definiciones precisas de muchas colecciones de objetos. Siempre está claro si una colección está bien definida o no, pues si hay alguna duda es que no lo está. Las afirmaciones generales que hagamos sobre colecciones de objetos sólo adquirirán un significado concreto cuando se particularicen a colecciones concretas bien definidas. Esto se entenderá mejor en cuanto dispongamos de ejemplos.

Similarmente, una *función n-ádica* en U será un criterio bien definido por el que a cada n elementos de U repetidos o no y en un cierto orden se les asigna otro elemento de U. Si f es una función n-ádica en U, representaremos por $f(a_1, \ldots, a_n)$ al elemento de U asignado por f a los elementos $a_1, \ldots, a_n$ en este orden.

[1] Evitamos la palabra "conjunto" porque la reservamos para el uso técnico que le daremos en teoría de conjuntos. Así, del mismo modo que hemos de distinguir la noción de "número natural" metamatemático de la noción de "número natural" como concepto técnico de la aritmética de Peano, distinguiremos entre colección de elementos metamatemática (como pueda ser la colección de los axiomas de Peano) de la noción de "conjunto" que no la definiremos, sino que la introduciremos axiomáticamente en su momento.

Una *relación* n-ádica en U es un criterio bien definido por el cual seleccionamos ciertos grupos de n elementos de U repetidos o no y en un cierto orden. Si los elementos $a_1, \ldots, a_n$ en este orden constituyen uno de los grupos seleccionados por una relación R, escribiremos $R(a_1, \ldots, a_n)$, y diremos que $a_1, \ldots, a_n$ están relacionados por R.

En general, si U es una colección de elementos, llamaremos $=_U$, o simplemente $=$, a la relación diádica en U dada por $=(a, b)$ syss a y b son el mismo objeto de U. Escribiremos $a = b$ en lugar de $=(a, b)$.

Al igual que con las colecciones, cualquier afirmación que hagamos sobre relaciones o funciones en general deberá entenderse como una afirmación que será verdadera cada vez que se particularice a relaciones o funciones concretas bien definidas.

Con estas nociones, estamos ya en condiciones de definir el concepto de modelo de un lenguaje formal. Observemos que para darle un significado a una sentencia como $\bigwedge x(Hx \to \neg Mx)$ hemos de especificar el universo de objetos de los que pretendemos hablar (con lo que establecemos el significado del cuantificador $\bigwedge x$) y el significado de los relatores H y M. Si, por ejemplo, establecemos que el universo de objetos es la colección de todas las personas vivas ahora, que H ha de interpretarse como la relación monádica "ser un hombre" y que M ha de interpretarse como la relación "ser una mujer", entonces la sentencia pasa a tener un significado preciso: "toda persona que sea hombre no es mujer". La noción de modelo, que introducimos a continuación, recoge todo lo que hay que especificar para que podamos atribuir un significado a una fórmula cualquiera de un lenguaje formal.

Definición 3.1 Un *modelo* M de un lenguaje formal $\mathfrak{L}$ viene determinado por:

a) Una colección de objetos U llamada *universo* de M. La colección U ha de tener al menos un objeto.

b) Un criterio que asocie a cada constante c de $\mathfrak{L}$ un objeto $M(c)$ de U.

c) Un criterio que asocie a cada relator n-ádico R_i^n de $\mathfrak{L}$ una relación n-ádica $M(R_i^n)$ en U. La relación $M(=)$ ha de ser $=$.

d) Un criterio que asocie a cada funtor n-ádico f_i^n de $\mathfrak{L}$ una función n-ádica $M(f_i^n)$ en U.

e) (Si $\mathfrak{L}$ tiene descriptor) un elemento d de U al que llamaremos *descripción impropia* de M.

Claramente, si $\mathfrak{L}$ tiene descriptor, todo modelo de $\mathfrak{L}$ lo es de $\underline{\mathfrak{L}}$ (olvidando la descripción impropia) y todo modelo de $\underline{\mathfrak{L}}$ se convierte en un modelo de $\mathfrak{L}$ fijando una descripción impropia.

Ejemplo Consideremos el lenguaje $\mathfrak{L}$ de la aritmética de Peano. Llamaremos *modelo natural* de $\mathfrak{L}$ al modelo $\mathfrak{M}$ cuyo universo son los números naturales, en el que la constante 0 se interpreta como el número natural 0, el funtor S se

interpreta como la función monádica que a cada natural le asigna su siguiente y los funtores + y · se interpretan como las funciones diádicas suma y producto. La descripción impropia es el cero.

El modelo natural no es el único modelo posible. Otro distinto sería el modelo M que tiene por universo al número 0 únicamente, con las mismas interpretaciones que $\mathcal{M}$ excepto que el funtor S se interpreta como la función constante igual a 0. La diferencia esencial entre ambos modelos consiste en que los axiomas de Peano serán verdaderos respecto al primero mientras que uno de ellos será falso respecto al segundo. Para justificar esto hemos de introducir antes la noción de "verdad" en un modelo. ∎

Observemos que la definición de modelo no atribuye ningún significado a los signos como $\to$ o $\neg$. Ello se debe a que estos signos tendrán siempre el mismo significado, independientemente del modelo. Un hecho más notable es que no asignamos ninguna interpretación a las variables. El motivo es justo el contrario: que deseamos que las variables puedan variar de significado aunque hayamos fijado un modelo. Por ello las interpretamos aparte, mediante el concepto siguiente:

Una *valoración* de un lenguaje formal $\mathfrak{L}$ en un modelo M es un criterio v que asigna a cada variable x de $\mathfrak{L}$ un objeto $v(x)$ del universo de M.

Si v es una valoración de un lenguaje formal $\mathfrak{L}$ en un modelo M, a es un elemento del universo U de M y x es una variable de $\mathfrak{L}$, llamaremos v_x^a a la valoración de M dada por

$$v_x^a(y) = \begin{cases} a & \text{si } y \equiv x, \\ v(y) & \text{si } y \not\equiv x. \end{cases}$$

Llamaremos v_{xy}^{ab} a $(v_x^a)_y^b$, llamaremos v_{xyz}^{abc} a $((v_x^a)_y^b)_z^c$, etc.

Es claro que si $x \equiv y$ entonces v_{xy}^{ab} coincide con v_y^b, mientras que si $x \not\equiv y$, entonces v_{xy}^{ab} coincide con v_{yx}^{ba}.

Ahora ya podemos definir el significado de una expresión arbitraria θ de un lenguaje $\mathfrak{L}$ respecto a un modelo M. Notemos que si θ es un término su significado ha de ser un objeto del universo de M, mientras que si θ es una fórmula su significado ha de ser un valor de verdad: ha de ser verdadera o falsa.

Definición 3.2 Sea v una valoración de un lenguaje formal $\mathfrak{L}$ en un modelo M de universo U. Las condiciones siguientes determinan cuándo M *satisface* la fórmula α respecto a la valoración v (abreviado $M \vDash \alpha[v]$) y cuándo un término t *denota* en M respecto a la valoración v a un objeto que representaremos por $M(t)[v]$:

a) $M(x_i)[v] = v(x_i)$,

b) $M(c_i)[v] = M(c_i)$,

c) $M \vDash R_i^n t_1 \cdots t_n$ syss $M(R_i^n)(M(t_1)[v], \ldots, M(t_n)[v])$,

d) $M(f_i^n t_1 \cdots t_n)[v] = M(f_i^n)(M(t_1)[v], \ldots, M(t_n)[v])$,

e) $M \vDash \neg\alpha[v]$ syss no $M \vDash \alpha[v]$,

f) $M \vDash (\alpha \to \beta)[v]$ syss no $M \vDash \alpha[v]$ o $M \vDash \beta[v]$,

g) $M \vDash \bigwedge x_i \alpha[v]$ syss para todo objeto a de U se cumple que $M \vDash \alpha[v^a_{x_i}]$,

h) Si $\mathfrak{L}$ tiene descriptor

$$M(x_i|\alpha)[v] = \begin{cases} \text{el único } a \text{ de } U \text{ tal que } M \vDash \alpha[v^a_{x_i}] & \text{si existe tal } a, \\ d & \text{en otro caso.} \end{cases}$$

Esta definición requiere varias reflexiones.

— En primer lugar, esta definición no hace sino especificar lo que todos sabemos hacer instintivamente al leer una sentencia formal. Por ejemplo, si consideramos la sentencia

$$\bigwedge xy\ x \cdot x \neq 0'' \cdot (y \cdot y) + 0'''$$

del lenguaje de la aritmética de Peano, sin necesidad de definición alguna todos entendemos que ahí dice que no existen números naturales m y n tales que $m^2 - 2n^2 = 3$. Vamos a ver que si aplicamos la definición anterior con el modelo natural de la aritmética llegamos a la misma conclusión. Fijemos una valoración v arbitraria.

- Por definición b) tenemos que $\mathcal{M}(0)[v]$ es el número natural 0.
- Por la definición d) tenemos que $\mathcal{M}(0')[v]$ es el siguiente de $\mathcal{M}(0)[v]$, o sea, el número 1. Similarmente, $\mathcal{M}(0'')[v] = 2$ y $\mathcal{M}(0''')[v] = 3$.
- Aplicando varias veces a) y d) concluimos que

$$\mathcal{M}(x \cdot x)[v] = v(x)^2, \qquad \mathcal{M}\big(0'' \cdot (y \cdot y) + 0'''\big)[v] = 2 \cdot v(y)^2 + 3.$$

- Por c) y e),

$$\mathcal{M} \vDash \big(x \cdot x \neq 0'' \cdot (y \cdot y) + 0'''\big)[v] \quad \text{syss} \quad v(x)^2 - 2 \cdot v(y)^2 \neq 3.$$

- En particular, si m y n son dos números naturales cualesquiera

$$\mathcal{M} \vDash \big(x \cdot x \neq 0'' \cdot (y \cdot y) + 0'''\big)[v^{mn}_{x\,y}] \quad \text{syss} \quad m^2 - 2 \cdot n^2 \neq 3.$$

- Aplicando dos veces g) concluimos que $\mathcal{M} \vDash \bigwedge xy\ x \cdot x \neq 0'' \cdot (y \cdot y) + 0'''[v]$ syss para todos los números naturales m y n, se cumple $m^2 - 2 \cdot n^2 \neq 3$.

Así pues, en ejemplos concretos nunca necesitaremos aplicar explícitamente la definición de denotación y satisfacción, pues hacerlo nos lleva por un camino más largo al mismo sitio al que llegamos si simplemente "leemos" la expresión dada.

— Otro hecho que hay que destacar es que, al contrario que ocurría con otras definiciones similares, como la de variable libre o la de sustitución, la definición anterior nos dice qué es el objeto denotado por un término y qué significa que una fórmula sea satisfecha, pero no nos proporciona un algoritmo para decidir si se da o no el caso. Por ejemplo, acabamos de ver que $\mathcal{M} \vDash \bigwedge xy\ x\cdot x \neq 0''\cdot(y\cdot y)+0'''[v]$ significa que no existen números naturales m y n tales que $m^2 - 2n^2 = 3$. Eso es lo que obtenemos al aplicar la definición de satisfacción, pero un problema muy distinto es decidir si existen o no tales números naturales. Si en el capítulo anterior insistíamos en que podemos trabajar en un sistema deductivo formal sin apoyarnos en ninguna interpretación de sus fórmulas, ahora hemos de insistir en el polo opuesto: podemos hablar de la satisfacción o no satisfacción de una fórmula con independencia de cualquier razonamiento (formal o informal) que nos convenza de la verdad o falsedad de dicha fórmula. Sabemos lo que significa que no existen números naturales tales que $m^2 - 2n^2 = 3$ con independencia de si sabemos probar que existen o que no existen. Si no fuera así, el modelo $\mathcal{M}$ no estaría bien definido.

— Es en la definición anterior donde por primera vez atribuimos un significado a los signos lógicos. Así, en el apartado e) es donde estipulamos que el signo $\neg$ se ha de interpretar siempre como "no", etc.

La definición de denotación y satisfacción puede completarse con los hechos siguientes:

$M \vDash (\alpha \vee \beta)[v]$ syss $M \vDash \alpha[v]$ o $M \vDash \beta[v]$,

$M \vDash (\alpha \wedge \beta)[v]$ syss $M \vDash \alpha[v]$ y $M \vDash \beta[v]$,

$M \vDash (\alpha \leftrightarrow \beta)[v]$ syss $M \vDash \alpha[v]$ y $M \vDash \beta[v]$ o por el contrario no $M \vDash \alpha[v]$ y no $M \vDash \beta[v]$,

$M \vDash \bigvee x_i\alpha[v]$ syss existe un objeto a de U tal que $M \vDash \alpha[v^a_{x_i}]$,

$M \vDash \overset{1}{\bigvee} x\alpha[v]$ syss existe un único objeto a de U tal que $M \vDash \alpha[v^a_x]$,

$M(x|x = x)[v] = d$.

Estos hechos no forman parte de la definición, sino que son teoremas que se deducen de ella. Por ejemplo, el primero se demuestra aplicando los apartados e) y f), ya que, por definición, $\alpha \vee \beta \equiv \neg\alpha \rightarrow \beta$.

Si $\mathfrak{L}$ tiene descriptor y una fórmula α no tiene descriptores, entonces se cumple $M \vDash \alpha[v]$ considerando a M como modelo de $\mathfrak{L}$ si y sólo si se cumple considerándolo como modelo de $\underline{\mathfrak{L}}$.

Probamos ahora un par de resultados rutinarios sobre denotación y satisfacción. Observemos que para saber si la fórmula $x \cdot x = y$ es satisfecha o no en el modelo natural de la aritmética respecto a una valoración v hemos de conocer $v(x)$ y $v(y)$. Si el segundo es el cuadrado del primero la respuesta será afirmativa, y en caso contrario será negativa. Sin embargo, es irrelevante los valores que tome v sobre cualquier variable distinta de x e y. Esto es un hecho

general: el que una fórmula α sea satisfecha o no en un modelo respecto de una valoración v sólo depende de los valores que toma v sobre las variables libres en α. Lo probamos en el teorema siguiente:

Teorema 3.3 *Si v y w son valoraciones de un lenguaje forma $\mathfrak{L}$ en un modelo M de universo U que coinciden sobre las variables libres de una expresión θ, entonces si θ es un término $M(\theta)[v] = M(\theta)[w]$ y si θ es una fórmula $M \vDash \theta[v]$ syss $M \vDash \theta[w]$.*

DEMOSTRACIÓN: Por inducción sobre la longitud de θ.

Si $\theta \equiv x$ entonces x está libre en θ, luego $M(\theta)[v] = v(x) = w(x) = M(\theta)[w]$.

Si $\theta \equiv c$ entonces $M(\theta)[v] = M(c) = M(\theta)[w]$.

Si $\theta \equiv R_i^n t_1 \cdots t_n$, por hipótesis de inducción $M(t_j)[v] = M(t_j)[w]$. Entonces

$$M \vDash \theta[v] \text{ syss } M(R_i^n)(M(t_1)[v], \ldots, M(t_n)[v])$$
$$\text{syss } M(R_i^n)(M(t_1)[w], \ldots, M(t_n)[w]) \text{ syss } M \vDash \theta[w].$$

Si $\theta \equiv f_i^n t_1 \cdots t_n$, por hipótesis de inducción $M(t_j)[v] = M(t_j)[w]$. Entonces

$$M(\theta)[v] = M(f_i^n)(M(t_1)[v], \ldots, M(t_n)[v])$$
$$= M(f_i^n)(M(t_1)[w], \ldots, M(t_n)[w]) = M(\theta)[w].$$

Si $\theta \equiv \neg\alpha$, por hipótesis de inducción $M \vDash \alpha[v]$ syss $M \vDash \alpha[w]$, luego no $M \vDash \alpha[v]$ syss no $M \vDash \alpha[w]$, o sea $M \vDash \neg\alpha[v]$ syss $M \vDash \neg\alpha[w]$.

Si $\theta \equiv \alpha \to \beta$, por hipótesis de inducción $M \vDash \alpha[v]$ syss $M \vDash \alpha[w]$ y $M \vDash \beta[v]$ syss $M \vDash \beta[w]$. Entonces $M \vDash (\alpha \to \beta)[v]$ syss no $M \vDash \alpha[v]$ o $M \vDash \beta[v]$ syss no $M \vDash \alpha[w]$ o $M \vDash \beta[w]$ syss $M \vDash (\alpha \to \beta)[w]$.

Si $\theta \equiv \bigwedge x\alpha$, sea a un objeto de U. Si y está libre en α entonces y está libre en θ o $y \equiv x$. En cualquier caso $v_x^a(y) = w_x^a(y)$, luego v_x^a y w_x^a coinciden en las variables libres de α. Por hipótesis de inducción $M \vDash \alpha[v_x^a]$ syss $M \vDash \alpha[w_x^a]$ para todo objeto a de U.

En particular $M \vDash \alpha[v_x^a]$ para todo objeto a de U syss $M \vDash \alpha[w_x^a]$ para todo objeto a de U, es decir, $M \vDash \bigwedge x\alpha[v]$ syss $M \vDash \bigwedge x\alpha[w]$.

Si $\theta \equiv x|\alpha$, razonando como antes, para todo objeto a de U tenemos que $M \vDash \alpha[v_x^a]$ syss $M \vDash \alpha[w_x^a]$. Por lo tanto hay un único a en U tal que $M \vDash \alpha[v_x^a]$ syss hay un único a en U tal que $M \vDash \alpha[w_x^a]$, además en tal caso son el mismo objeto a.

Si se da la unicidad $M(\theta)[v] = a = M(\theta)[w]$. Si no se da la unicidad $M(\theta)[v] = d = M(\theta)[w]$. ∎

Cuando definimos la sustitución dijimos que nuestra intención era que $\mathtt{S}_x^t\alpha$ afirmara de t lo que α afirmaba de x. Naturalmente, la definición que hemos dado es formal y en ella no aparece explícitamente esta idea. Ahora que contamos con una noción precisa de significado de una fórmula podemos probar que, desde el punto de vista semántico, la definición de sustitución es exactamente lo que queríamos que fuera.

Teorema 3.4 *Sea v una valoración de un lenguaje formal $\mathfrak{L}$ en un modelo M de universo U. Sea t un término de $\mathfrak{L}$, sea θ una expresión y x una variable. Entonces si θ es un término*

$$M(\mathtt{S}_x^t\theta)[v] = M(\theta)[v_x^{M(t)[v]}]$$

y si θ es una fórmula

$$M \vDash \mathtt{S}_x^t\theta[v] \quad \textit{syss} \quad M \vDash \theta[v_x^{M(t)[v]}].$$

Por ejemplo, la segunda afirmación puede parafrasearse así: "Se verifica $\mathtt{S}_x^t\theta$ si y sólo si se verifica θ cuando x se interpreta como el objeto denotado por t".

DEMOSTRACIÓN: Por inducción sobre la longitud de θ.

Si $\theta \equiv y$ distinguimos dos casos:
a) si $x \not\equiv y$ entonces $M(\mathtt{S}_x^t\theta)[v] = M(y)[v] = M(\theta)[v_x^{M(t)[v]}]$,
b) si $x \equiv y$ entonces $M(\mathtt{S}_x^t\theta)[v] = M(t)[v] = M(\theta)[v_x^{M(t)[v]}]$.

Si $\theta \equiv c$ entonces $M(\mathtt{S}_x^t\theta)[v] = M(c)[v] = M(\theta)[v_x^{M(t)[v]}]$.

Si $\theta \equiv R_i^n t_1 \cdots t_n$ entonces

$$M \vDash \mathtt{S}_x^t\theta[v] \text{ syss } M \vDash R_i^n \mathtt{S}_x^t t_1 \cdots \mathtt{S}_x^t t_n[v] \text{ syss}$$
$$M(R_i^n)(M(\mathtt{S}_x^t t_1)[v], \ldots, M(\mathtt{S}_x^t t_n)[v]) \text{ syss}$$
$$M(R_i^n)(M(t_1)[v_x^{M(t)[v]}], \ldots, M(t_n)[v_x^{M(t)[v]}]) \text{ syss } M \vDash \theta[v_x^{M(t)[v]}].$$

Si $\theta \equiv f_i^n t_1 \cdots t_n$ entonces

$$M(\mathtt{S}_x^t\theta)[v] = M(f_i^n \mathtt{S}_x^t t_1 \cdots \mathtt{S}_x^t t_n)[v] = M(f_i^n)(M(\mathtt{S}_x^t t_1)[v], \ldots, M(\mathtt{S}_x^t t_n)[v])$$
$$= M(f_i^n)(M(t_1)[v_x^{M(t)[v]}], \ldots, M(t_n)[v_x^{M(t)[v]}]) = M(\theta)[v_x^{M(t)[v]}].$$

Si $\theta \equiv \neg\alpha$ entonces $M \vDash \mathtt{S}_x^t\theta[v]$ syss $M \vDash \neg\mathtt{S}_x^t\alpha[v]$ syss no $M \vDash \mathtt{S}_x^t\alpha[v]$ syss no $M \vDash \alpha[v_x^{M(t)[v]}]$ syss $M \vDash \theta[v_x^{M(t)[v]}]$.

Si $\theta \equiv \alpha \to \beta$ entonces $M \vDash \mathtt{S}_x^t\theta[v]$ syss $M \vDash (\mathtt{S}_x^t\alpha \to \mathtt{S}_x^t\beta)[v]$ syss no $M \vDash \mathtt{S}_x^t\alpha[v]$ o $M \vDash \mathtt{S}_x^t\beta[v]$ syss no $M \vDash \alpha[v_x^{M(t)[v]}]$ o $M \vDash \beta[v_x^{M(t)[v]}]$ syss $M \vDash \theta[v_x^{M(t)[v]}]$.

Si $\theta \equiv \bigwedge y\alpha$ distinguimos tres casos:

a) Si x no está libre en $\bigwedge y\alpha$, entonces $M \vDash \mathtt{S}_x^t\theta[v]$ syss $M \vDash \bigwedge y\alpha[v]$ syss $M \vDash \bigwedge y\alpha[v_x^{M(t)[v]}]$ por el teorema anterior.

b) Si x está libre en $\bigwedge y\alpha$ e y no lo está en t, fijemos un objeto a en U. Entonces $M \vDash \mathtt{S}_x^t\alpha[v_y^a]$ syss (hip. de ind.) $M \vDash \alpha[v_{y\,x}^{a\,M(t)[v_y^a]}]$ syss (teor. 1) $M \vDash \alpha[v_{y\,x}^{a\,M(t)[v]}]$ syss $M \vDash \alpha[v_{x\;\;y}^{M(t)[v]\,a}]$ (notar que $x \not\equiv y$ pues x está libre en $\bigwedge y\alpha$ e y no lo está).

Por lo tanto, $M \vDash \mathtt{S}^t_x\theta[v]$ syss $M \vDash \bigwedge y\mathtt{S}^t_x\alpha[v]$ syss para todo a de U se cumple que $M \vDash \mathtt{S}^t_x\alpha[v^a_y]$ syss para todo a de U se cumple que $M \vDash \alpha[v_x^{M(t)[v]a}{}_y]$ syss $M \vDash \bigwedge y\alpha[v_x^{M(t)[v]}]$.

c) Si x está libre en $\bigwedge y\alpha$, y está libre en t y z es la variable de menor índice que no está en $\bigwedge y\alpha$ ni en t, fijemos un objeto a en U.

Entonces $M \vDash \mathtt{S}^t_x\mathtt{S}^z_y\alpha[v^a_z]$ syss (hip. de ind.) $M \vDash \mathtt{S}^z_y\alpha[v_z^{aM(t)[v^a_z]}]$ syss (3.3) $M \vDash \mathtt{S}^z_y\alpha[v_z^{aM(t)[v]}]$ syss $M \vDash \mathtt{S}^z_y\alpha[v_x^{M(t)[v]a}{}_z]$ syss (hip. de ind.) $M \vDash \alpha[v_x^{M(t)[v]aa}{}_{zy}]$ syss (teor. 1) $M \vDash \alpha[v_x^{M(t)[v]a}{}_y]$.

Por lo tanto $M \vDash \mathtt{S}^t_x\theta[v]$ syss $M \vDash \bigwedge z\mathtt{S}^t_x\mathtt{S}^z_y\alpha[v]$ syss para todo a de U se cumple que $M \vDash \mathtt{S}^t_x\mathtt{S}^z_y\alpha[v^a_z]$ syss para todo a de U se cumple que $M \vDash \alpha[v_x^{M(t)[v]a}{}_y]$ syss $M \vDash \bigwedge y\alpha[v_x^{M(t)[v]}]$.

Si $\theta \equiv y|\alpha$ distinguimos tres casos:

a) Si x no está libre en $y \mid \alpha$ entonces, usando el teorema anterior,

$$M(\mathtt{S}^t_x\theta)[v] = M(y|\alpha)[v] = M(\theta)[v_x^{M(t)[v]}].$$

b) Si x está libre en $y|\alpha$ e y no lo está en t, entonces, fijando un objeto a en U, como en el apartado b) del caso anterior concluimos que $M \vDash \mathtt{S}^t_x\alpha[v^a_y]$ syss $M \vDash \alpha[v_x^{M(t)[v]a}{}_y]$, luego existe un único a en U tal que $M \vDash \mathtt{S}^t_x\alpha[v^a_y]$ syss existe un único a en U tal que $M \vDash \alpha[v_x^{M(t)[v]a}{}_y]$, y en tal caso son el mismo.

Si se da la unicidad entonces $M(\mathtt{S}^t_x\theta)[v] = M(x|\mathtt{S}^t_x\theta\alpha) = a = M(\theta)[v_x^{M(t)[v]}]$. En otro caso $M(\mathtt{S}^t_x\theta)[v] = d = M(\theta)[v_c^{M(t)[v]}]$.

c) Si x está libre en $y|\alpha$, y está libre en t y z es la variable de menor índice que no está en $y|\alpha$ ni en t, fijamos un objeto a en U y como en el apartado c) del caso anterior se prueba que $M \vDash \mathtt{S}^t_x\mathtt{S}^z_y\alpha[v^a_z]$ syss $M \vDash \alpha[v_x^{M(t)[v]a}{}_y]$.

Ahora razonamos igual que en el apartado b) de este caso. ∎

3.2 Verdad y validez lógica

Es claro que el hecho de que una fórmula α sea satisfecha en un modelo M es tanto como decir que α es verdadera cuando sus signos se interpretan en M. Sin embargo, conviene restringir la noción de fórmula verdadera para describir una situación ligeramente más fuerte que la mera satisfacción. Esencialmente, nuestro interés es definir la noción de verdad de tal modo que no dependa de la interpretación de sus variables, de modo que una fórmula como $x+y = y+x$ será verdadera en el modelo natural de la aritmética (porque se cumple sean quienes sean x e y), mientras que una fórmula como $x = y + 0'$ no sea ni verdadera ni falsa, porque será satisfecha o no según cómo interpretemos x e y. La definición de verdad es, pues, la siguiente:

Definición 3.5 Una fórmula α de un lenguaje formal $\mathfrak{L}$ es *verdadera* en un modelo M si $M \vDash \alpha[v]$ cualquiera que sea la valoración v de $\mathfrak{L}$ en M. Lo representaremos $M \vDash \alpha$. Diremos que α es *falsa* en M si ninguna valoración v de $\mathfrak{L}$ en M cumple $M \vDash \alpha[v]$. Si Γ es una colección de fórmulas escribiremos $M \vDash \Gamma$ para indicar que todas las fórmulas de Γ son verdaderas en M. Diremos también que M es un *modelo* de Γ.

Nota Esta definición presupone algo que debe ser matizado: que cuando hablamos de la totalidad de las valoraciones de un lenguaje en un modelo sabemos lo que estamos diciendo. Esto no está claro en absoluto: cuando hablamos de que todas las fórmulas de un lenguaje cumplen algo sabemos lo que queremos decir: es fácil enumerarlas explícitamente, y entonces nuestra afirmación significa que la primera cumple lo indicado, y la segunda también, etc. (con independencia de si sabemos probar o no que así es). Por el contrario, no tenemos ninguna representación similar que nos permita atribuir un significado a las afirmaciones que hagamos sobre la totalidad de las valoraciones.

Pese a esto, la definición anterior tiene un sentido preciso gracias al teorema 3.3. Los únicos modelos que vamos a considerar, tanto a nivel teórico como a nivel práctico, (sin entrar en la cuestión de si tendría sentido hablar de modelos más generales) van a ser modelos cuyo universo es una colección numerable, es decir, tal que sabemos establecer una correspondencia biunívoca entre sus objetos y los números naturales (no necesariamente calculable en la práctica). En tal caso, una afirmación sobre la totalidad de los objetos del modelo se entiende como una afirmación válida para el primer objeto, y para el segundo, etc.

En estas circunstancias —que, según lo dicho, son las únicas en las que vamos a trabajar—, también sabemos dar un sentido preciso a cualquier afirmación sobre la totalidad de los grupos $(a_1, \ldots, a_n)$ de n objetos del modelo en un orden dado y con posibles repeticiones. En efecto, podemos enumerar explícitamente todas las n-tuplas de números naturales (ponemos primero todas las formadas por números que sumen 0 (hay una sola), luego todas las formadas por números que sumen 1 (hay n), etc. De este modo, una afirmación sobre la totalidad de los grupos de n objetos es verdadera si se cumple con el primer grupo de n objetos, y con el segundo, etc.

Ahora sólo queda observar que en virtud del teorema 3.3 podríamos haber definido que una fórmula α es verdadera en un modelo M como que α es satisfecha para todas las interpretaciones posibles de sus variables libres, lo cual sí sabemos lo que significa porque son un número finito. Si sucede esto, tendremos que $M \vDash \alpha[v]$ para cualquier valoración v que consideremos, tal y como exige la definición de verdad que hemos dado.

Claramente se cumplen los hechos siguientes:

a) Una fórmula no puede ser verdadera y falsa en un mismo modelo (pues tomando una valoración cualquiera, será satisfecha o no lo será).

b) Toda sentencia es verdadera o falsa en un modelo (por el teorema 3.3).

c) Una fórmula α es verdadera en un modelo M syss $\neg\alpha$ es falsa, α es falsa syss $\neg\alpha$ es verdadera.

d) Si $M \vDash (\alpha \to \beta)$ y $M \vDash \alpha$, entonces $M \vDash \beta$ (pues toda valoración v cumple $M \vDash (\alpha \to \beta)[v]$ y $M \vDash \alpha[v]$, luego $M \vDash \beta[v]$).

e) $M \vDash \alpha$ syss $M \vDash \bigwedge x\alpha$.

En efecto: si $M \vDash \alpha$ y v es una valoración en M, dado a en el universo U de M, se ha de cumplir $M \vDash \alpha[v^a_x]$ y, por lo tanto, $M \vDash \bigwedge x\alpha[v]$, o sea, $M \vDash \bigwedge x\alpha$.

Si $M \vDash \bigwedge x\alpha$ y v es una valoración en M, entonces $M \vDash \bigwedge x\alpha[v]$, luego $M \vDash \alpha[v]$, y así, $M \vDash \alpha$.

f) $M \vDash \alpha$ syss $M \vDash \alpha^c$, donde α^c es la *clausura universal* de α, es decir, la sentencia que resulta de generalizar todas las variables de α (aplicando varias veces el resultado anterior).

g) Si α no tiene descriptores, entonces $M \vDash \alpha$ se cumple o no independientemente del cuál sea la descripción impropia de M. En particular, $M \vDash \alpha$ en $\mathfrak{L}$ syss $M \vDash \alpha$ en $\underline{\mathfrak{L}}$.

Obviamente, la noción de verdad es relativa a un modelo: la sentencia $\bigwedge x\ x' \neq 0$ es verdadera en el modelo natural de la aritmética pero es falsa en el otro modelo que describimos a continuación, en el que el funtor S se interpretaba como la función constante 0. Sin embargo, hay fórmulas que son verdaderas en cualquier modelo, tales como $x = x$. Esto nos lleva a introducir algunos conceptos adicionales:

Definición 3.6 Una fórmula α de un lenguaje formal $\mathfrak{L}$ es *lógicamente válida* si es verdadera en todo modelo de $\mathfrak{L}$. Lo representaremos por $\vDash \alpha$. Diremos que α es *insatisfacible* si es falsa en todo modelo, α es *satisfacible* si es verdadera en algún modelo y es *falseable* si es falsa en algún modelo.

Nota Nos encontramos aquí con una dificultad mayor que la que nos ha planteado la definición de verdad. Estos conceptos presuponen la noción de la "totalidad de los modelos de un lenguaje formal", noción que está muy lejos de ser precisa. Desgraciadamente, ahora no contamos con un análogo del teorema 3.3 que nos permita esquivar la dificultad, por lo que la única posibilidad que nos queda es debilitar el significado pretendido de este tipo de afirmaciones. Así, convenimos que cuando escribamos $\vDash \alpha$ no habremos de entender que α es verdadera en cualquier modelo, lo cual, como estamos explicando, no está claro lo que significa, sino más bien que existe un argumento que nos convence de que α ha de ser verdadera en cualquier modelo.

Por ejemplo, podemos afirmar que $\vDash x = x$, ya que a partir de las definiciones de modelo, satisfacción y verdad se razona fácilmente que la fórmula $x = x$ tiene que ser verdadera en cualquier modelo que tengamos definido o, dicho de otro modo, que es imposible que nos encontremos alguna vez con un modelo en el cual $x = x$ no sea verdadera.

Así pues, si sabemos demostrar que una fórmula α ha de ser verdadera en cualquier modelo, podemos expresar este hecho concreto diciendo que α es lógicamente válida; si por el contrario sabemos construir un modelo donde α es falsa podemos expresar este hecho concreto diciendo que α es falseable; pero si no disponemos de ningún argumento que nos convenza de lo uno o de lo otro, las afirmaciones "α es lógicamente válida" y "α es falseable" carecen de un significado concreto, por lo que, en particular, carece de sentido afirmar que una de las dos ha de ser cierta.

De todos modos, debemos advertir que estas restricciones son provisionales, ya que más adelante justificaremos que es posible atribuir un significado intrínseco a la noción de validez lógica independiente de si sabemos verificarlo o no. ∎

Claramente se cumple:

a) α es lógicamente válida syss $\neg\alpha$ es insatisfacible.

b) α es insatisfacible syss $\neg\alpha$ es lógicamente válida.

c) α no puede ser lógicamente válida e insatisfacible.

d) α es satisfacible syss $\neg\alpha$ es falseable.

e) α es falseable syss $\neg\alpha$ es satisfacible.

f) Si $\vDash \alpha \to \beta$ y $\vDash \alpha$, entonces $\vDash \beta$.

g) $\vDash \alpha$ syss $\vDash \bigwedge x\alpha$.

h) $\vDash \alpha$ syss $\vDash \alpha^c$ (donde α^c es la clausura universal de α).

Todos estos hechos han de entenderse en los términos explicados en la nota anterior. Por ejemplo, la afirmación a) expresa que si existe un argumento que nos convence de que una fórmula α es necesariamente verdadera en cualquier modelo dado, dicho argumento se prolonga inmediatamente hasta otro que nos convence de que $\neg\alpha$ es necesariamente falsa en cualquier modelo dado, etc.

La propiedad c) hace uso de que todo lenguaje formal $\mathcal{L}$ tiene al menos un modelo. En efecto, es fácil definir un modelo cuyo universo tenga, por ejemplo, un único objeto y en el que las constantes, los relatores y los funtores de $\mathcal{L}$ se interpreten de forma obvia.

Las propiedades f) y g) se interpretan como que las dos reglas de inferencia primitivas MP e IG son "lógicamente válidas", en el sentido que que nos llevan siempre de fórmulas lógicamente válidas a fórmulas lógicamente válidas.

Ahora estamos en condiciones de probar un teorema básico sobre el cálculo deductivo:

Teorema 3.7 *Todos los axiomas de $K_{\mathcal{L}}$ son lógicamente válidos.*

DEMOSTRACIÓN: Según lo que hemos explicado, lo que hemos de probar es que disponemos de argumentos que nos convencen de que es imposible tener un modelo de un lenguaje $\mathcal{L}$ en el que alguno de los axiomas de $K_{\mathcal{L}}$ no sea verdadero.

Supongamos, pues, que M es un modelo de un lenguaje $\mathcal{L}$ de universo U y veamos que cualquier axioma ϕ de $K_{\mathcal{L}}$ ha de cumplir $M \models \phi$. A su vez, para ello, fijamos una valoración v en M y trataremos de justificar que $M \models \phi[v]$.

Si ϕ es un axioma de tipo K1, K2 o K3 la comprobación es fácil.

Si ϕ es de tipo K4 entonces $\phi \equiv \bigwedge x\alpha \to \mathtt{S}^t_x\alpha$. Hemos de probar que si $M \models \bigwedge x\alpha[v]$ entonces $M \models \mathtt{S}^t_x\alpha[v]$. Ahora bien, por el teorema 3.4 esto último equivale a que $M \models \alpha[v_x^{M(t)[v]}]$, y esto se cumple bajo nuestra hipótesis.

Si ϕ es de tipo K5 entonces $\phi \equiv \bigwedge x(\alpha \to \beta) \to (\alpha \to \bigwedge x\beta)$, y la variable x no está libre en α. Suponemos que $M \models \bigwedge x(\alpha \to \beta)[v]$ y hemos de probar que $M \models (\alpha \to \bigwedge x\beta)[v]$. A su vez, para ello suponemos que $M \models \alpha[v]$ y hemos de probar que $M \models \bigwedge x\beta[v]$.

Fijemos a en U. Tenemos que $M \models (\alpha \to \beta)[v^a_x]$ y, como x no está libre en α, por el teorema 1 también $M \models \alpha[v^a_x]$, luego $M \models \beta[v^a_x]$. Como esto se cumple para todo a de U, concluimos que $M \models \bigwedge x\beta[v]$, como queríamos probar.

Si ϕ es de tipo K6 entonces $\phi \equiv \bigwedge x(x = t \to \alpha) \leftrightarrow \mathtt{S}^t_x\alpha$, donde x no está libre en t. Hemos de probar que $M \models \bigwedge x(x = t \to \alpha)[v]$ syss $M \models \mathtt{S}^t_x\alpha[v]$.

Si $M \models \bigwedge x(x = t \to \alpha)[v]$, entonces $M \models (x = t \to \alpha)[v_x^{M(t)[v]}]$. Por el teorema 3.4 $M \models \mathtt{S}^t_x(x = t \to \alpha)[v]$, o sea, $M \models (t = t \to \mathtt{S}^t_x\alpha)[v]$. Puesto que obviamente $M \models (t = t)[v]$, también $M \models \mathtt{S}^t_x\alpha[v]$.

Supongamos ahora que $M \models \mathtt{S}^t_x\alpha[v]$ y fijemos un a en U. Hemos de probar que $M \models (x = t \to \alpha)[v^a_x]$. Para ello suponemos que $M \models (x = t)[v^a_x]$, es decir, que $M(x)[v^a_x] = M(t)[v^a_x]$ o, lo que es lo mismo, que $a = M(t)[v]$ (pues x no está libre en t).

Como $M \models \mathtt{S}^t_x\alpha[v]$, por el teorema 3.4 $M \models \alpha[v_x^{M(t)[v]}]$, o sea, $M \models \alpha[v^a_x]$, que es lo que teníamos que probar.

Si ϕ es de tipo K7 entonces $\phi \equiv \overset{1}{\bigvee} x\alpha \to \mathtt{S}^{x|\alpha}_x\alpha$. Suponemos que $M \models \overset{1}{\bigvee} x\alpha[v]$, lo que significa que existe un único a en U tal que $M \models \alpha[v^a_x]$. Por consiguiente $a = M(x|\alpha)[v]$ y así $M \models \alpha[v_x^{M(x|\alpha)[v]}]$. Por el teorema 3.4 $M \models \mathtt{S}^{x|\alpha}_x[v]$, como habíamos de probar.

Si ϕ es de tipo K8 entonces $\phi \equiv \neg\overset{1}{\bigvee} x\alpha \to x|\alpha = z|(z = z)$. Suponemos que $M \models \neg\overset{1}{\bigvee} x\alpha[v]$, con lo que no existe un único a en U tal que $M \models \alpha[v^a_x]$. Por lo tanto $M(x|\alpha)[v] = d = M(z|(z = z))[v]$, luego $M \models (x|\alpha = z|(z = z))[v]$, como teníamos que probar. ∎

Este teorema se generaliza inmediatamente al resultado principal de este capítulo:

Teorema 3.8 (Teorema de corrección) *Todos los teoremas lógicos son lógicamente válidos.*

DEMOSTRACIÓN: Sea α un teorema lógico. Esto significa que existe una demostración lógica que termina con α. Según el teorema anterior, sabemos justificar que cada uno de los axiomas que intervienen en la prueba es lógicamente válido y, por otra parte, sabemos que las reglas de inferencia transforman fórmulas lógicamente válidas en fórmulas lógicamente válidas. De todo esto se desprende que cada demostración lógica determina un razonamiento que nos convence de la validez lógica de cada una de sus fórmulas, en particular de su conclusión. ∎

Conviene enfatizar en este punto que un teorema formal no es, en sentido estricto, un razonamiento, sino una mera sucesión de signos sujeta a ciertos requisitos formales. No obstante, el argumento del teorema anterior muestra que cada teorema lógico codifica un razonamiento lógico: un razonamiento irrefutable de que la conclusión ha de ser verdadera en cualquier modelo que podamos considerar.

Esto nos proporciona una técnica para obtener resultados negativos, es decir, para demostrar que ciertas fórmulas no son teoremas lógicos.

Ejemplo La fórmula $x \neq x$ no es un teorema lógico.

Se entiende que $x \neq x$ es una fórmula de un cierto lenguaje formal $\mathcal{L}$ que no hemos especificado por ser irrelevante. Sea cual sea $\mathcal{L}$, podemos construirle un modelo M cuyo universo contenga un único objeto a y en el que sus constantes, relatores y funtores se interpreten de forma obvia. Es claro que la fórmula $x \neq x$ es falsa en M, luego no puede ser un teorema lógico. ∎

El argumento del teorema de corrección en general —o el del ejemplo anterior en particular— es delicado y, puesto que es un razonamiento metamatemático, no sujeto a unas condiciones de rigor preestablecidas, conviene que el lector reflexione sobre él hasta convencerse de que no deja lugar a dudas: es un argumento irrefutable en virtud del cual podemos estar seguros de que jamás aparecerá alguien con un papel que contenga una sucesión de fórmulas que satisfaga nuestra definición de demostración lógica y que termine con la fórmula $x \neq x$. Veamos un ejemplo más detallado:

Ejemplo Si $x \not\equiv y$, la sentencia $\bigwedge xy(x = y)$ no es un teorema lógico.

En efecto, basta ver que es falsa en un modelo M. Consideremos un modelo M cuyo universo U conste de dos objetos distintos a y b. Interpretemos de cualquier forma los relatores y funtores que pueda tener el lenguaje formal que estemos considerando. Entonces $M \vDash \neg\bigwedge xy(x = y)$, pues si v es una valoración tal que $v(x) = a$ y $v(y) = b$, entonces no $M \vDash (x = y)[v]$, luego tampoco $M \vDash \bigwedge xy(x = y)[v]$, luego no $M \vDash \bigwedge xy(x = y)$. ∎

Ejercicio: Probar que $\neg\bigwedge xy(x = y)$ no es un teorema lógico.

Ejercicio: Probar que la fórmula $x = y$ no es equivalente a su clausura universal $\bigwedge xy(x = y)$, es decir, que $x = y \leftrightarrow \bigwedge xy(x = y)$ no es un teorema lógico.

La demostración del teorema de corrección se generaliza de forma obvia al teorema siguiente:

Teorema 3.9 *Sea M un modelo de un lenguaje formal $\mathcal{L}$ y sea Γ una colección de fórmulas de $\mathcal{L}$. Si $M \models \Gamma$ y $\Gamma \vdash \alpha$, entonces $M \models \alpha$.*

En otras palabras, las consecuencias lógicas de premisas verdaderas son siempre verdaderas. Éste es exactamente el requisito que ha de cumplir un cálculo deductivo para que merezca el calificativo de "lógico" en el sentido tradicional. La distinción entre los auténticos razonamientos lógicos y las falacias no es arbitraria: una falacia es un presunto razonamiento que nos lleva a una conclusión que puede ser falsa aunque sus premisas sean verdaderas. Por el contrario, cualquier razonamiento que nos permita confiar en sus conclusiones exactamente en la misma medida en que confiemos en sus premisas es un razonamiento lógico legítimo. Los dos últimos teoremas justifican que el cálculo deductivo que hemos presentado es ciertamente "correcto" en este sentido. Sólo ahora está justificado el calificativo de teoremas lógicos y deducciones lógicas que hemos dado a estos conceptos introducidos —en principio— de forma arbitraria.

En particular podemos afirmar que el cálculo deductivo es matemáticamente aceptable. Recordemos que no vamos a ser capaces de precisar completamente el significado o la interpretación de las afirmaciones matemáticas. Precisamente ello nos ha obligado a asegurarnos de que es posible emplear rigurosamente el cálculo deductivo sin necesidad de especificar ningún modelo del lenguaje empleado, pero ahora sabemos que si exigimos que los teoremas matemáticos sean consecuencias lógicas de unos axiomas prefijados (en el sentido técnico que hemos dado a "consecuencia lógica"), tendremos la garantía de que todos los teoremas serán verdaderos en la medida en que pueda decirse que los axiomas lo son. Esto es todo lo que podemos pedir como garantía de corrección a un razonamiento matemático.[2]

Por otra parte, debemos señalar que una cosa es que nuestro cálculo deductivo sea correcto, en el sentido de que en él no hay nada objetable, y otra muy distinta que sea adecuado, en el sentido de que sea todo lo potente que debería ser. Más concretamente, queda pendiente la cuestión de si cualquier razonamiento lógico —en el sentido informal de que garantice la verdad de las conclusiones supuesta la verdad de las premisas— puede formalizarse, es decir, convertirse en una deducción lógica en el sentido técnico que hemos dado a esta noción. Por ejemplo, si eliminásemos el esquema axiomático $K6$ en $K_{\mathcal{L}}$ seguiríamos teniendo un cálculo deductivo correcto, aunque entonces $x = x$ ya no sería un teorema lógico (cuando no cabe duda de que debería serlo). La cuestión es, pues, si los axiomas de $K_{\mathcal{L}}$ son suficientes para deducir (formalmente) cualquier consecuencia lógica (en sentido informal) de unas premisas dadas o si, por el contrario, necesitamos añadir más axiomas o reglas de inferencia. En otras

[2]En realidad en algunos contextos es razonable pedir más, como que los razonamientos sean constructivos, pero esto no afecta a lo que estamos tratando ahora.

palabras, si existen o no razonamientos que un matemático daría por correctos pero que no son formalizables en $K_{\mathfrak{L}}$. La respuesta es que $K_{\mathfrak{L}}$ sí es adecuado en este sentido, pero esto lo veremos en el capítulo siguiente.

El teorema anterior nos permite probar que determinadas fórmulas no son consecuencias lógicas de otras dadas:

Ejemplo San Anselmo de Canterbury (1033–1109) propuso la siguiente demostración de la existencia de Dios:

> *Dios es el ser más perfecto que el cual ninguno puede ser pensado. Cualquier ser que exista es más perfecto que un ser que no exista, luego Dios ha de existir.*

Consideremos las sentencias siguientes:

$$\begin{aligned}
\alpha_1 &\equiv d = x|\neg\bigvee y(Py \wedge Myx),\\
\alpha_2 &\equiv Pd,\\
\alpha_3 &\equiv \bigvee x(Px \wedge Ex),\\
\alpha_4 &\equiv \bigwedge xy(Px \wedge Ex \wedge Py \wedge \neg Ey \to Mxy).
\end{aligned}$$

Veamos que de estas sentencias no se puede deducir Ed.

Las sentencias anteriores presuponen un lenguaje formal $\mathfrak{L}$ con descriptor dotado de una constante d, dos relatores monádicos E y P y un relator diádico M. Construyamos un modelo N de este lenguaje de acuerdo con las propiedades siguientes:

- El universo U de N consta de tres objetos a, b y c,
- la descripción impropia es a,
- $N(d) = a$,
- se cumple $N(P)(a)$, $N(P)(b)$ y $N(P)(c)$,
- se cumple $N(E)(b)$ y $N(E)(c)$, pero no $N(E)(a)$,
- se cumple $N(M)(b, a)$ y $N(M)(c, a)$, pero $N(M)$ es falso en cualquier otro caso.

Es fácil ver que $M \vDash \alpha_1, \alpha_2, \alpha_4, \alpha_4$, pero no $M \vDash Ed$. ∎

Ejercicio: En la formalización del argumento de san Anselmo, sustituir la sentencia α_1 por $\alpha_1' \equiv \neg\bigvee y(Py \wedge Myd)$. Probar que $\alpha_1', \alpha_2, \alpha_3, \alpha_4 \vdash Ed$.

Ejercicio: Formalizar las sentencias siguientes: "Cualquiera que sea capaz de superar esta marca ha de ser un atleta", "Juan no es capaz de superar esta marca", "Juan no es un atleta". Probar que la tercera no es consecuencia lógica de las dos primeras.

3.3 Consistencia

Los resultados que hemos obtenido nos permiten probar algunos hechos básicos sobre una noción fundamental en la lógica matemática: la noción de consistencia.

Definición 3.10 Una colección de fórmulas Γ de un lenguaje formal $\mathfrak{L}$ es *contradictoria* si existe una fórmula α de $\mathfrak{L}$ tal que $\Gamma \vdash \alpha$ y $\Gamma \vdash \neg\alpha$. En caso contrario se dice que es *consistente*. Equivalentemente, una teoría axiomática T es *contradictoria* si existe una fórmula α tal que $\underset{T}{\vdash} \alpha$ y $\underset{T}{\vdash} \neg\alpha$. En caso contrario es *consistente*.

La equivalencia consiste en que, obviamente, una teoría axiomática es consistente o contradictoria si y sólo si lo son sus axiomas: Es equivalente decir "la aritmética de Peano es consistente" que decir "los axiomas de la aritmética de Peano son consistentes". En general, todos los resultados sobre consistencia pueden enunciarse equivalentemente en términos de colecciones de fórmulas o de teorías axiomáticas. Nosotros usaremos arbitraria e indistintamente una u otra formulación.

El teorema siguiente muestra la importancia en la lógica matemática de la noción de consistencia.

Teorema 3.11 *Una teoría axiomática T sobre un lenguaje $\mathfrak{L}$ es contradictoria si y sólo sí todas las fórmulas de $\mathfrak{L}$ son teoremas de T.*

Demostración: Si en T puede probarse una contradicción, la regla de inferencia (C) de contradicción nos permite prolongar la prueba hasta una demostración de cualquier fórmula. El recíproco es todavía más evidente. ∎

Así pues, la consistencia es el requisito mínimo que ha de tener una teoría axiomática para que tenga interés trabajar en ella.

Nota A menudo conviene tener presente la versión recíproca del teorema anterior: *Una teoría axiomática es consistente si y sólo si existe una fórmula que no es un teorema.* Por consiguiente, los ejemplos de la sección anterior muestran que $K_{\mathfrak{L}}$ es consistente.

Observemos que la consistencia es una propiedad negativa: una teoría es consistente si ciertas cosas no se pueden demostrar en ella. Esto hace que en general no sea fácil determinar si una teoría dada es consistente o no. De hecho, veremos que en los casos más importantes es imposible. Por ello tienen interés los resultados de consistencia relativa, es decir, pruebas de que si una teoría es consistente sigue siéndolo al añadirle algún axioma más. En esta línea es útil tener presente la siguiente equivalencia:

Teorema 3.12 *Sea Γ una colección de fórmulas y α una sentencia. Entonces $\Gamma \cup \{\alpha\}$ es consistente si y sólo si no $\Gamma \vdash \neg\alpha$.*

DEMOSTRACIÓN: Evidentemente, $\Gamma \cup \{\alpha\}$ representa la colección que resulta de añadir la sentencia α a Γ. La notación conjuntista es mera taquigrafía. El enunciado no corresponde a un teorema de la teoría de conjuntos, sino que es un metateorema no formalizado, como todos los resultados que estamos probando.

Si $\Gamma \vdash \neg\alpha$ entonces $\Gamma \cup \{\alpha\} \vdash \alpha$ y $\Gamma \cup \{\alpha\} \vdash \neg\alpha$, luego $\Gamma \cup \{\alpha\}$ es contradictoria.

Si no $\Gamma \vdash \neg\alpha$, para ver que $\Gamma \cup \{\alpha\}$ es consistente basta probar que no $\Gamma \cup \{\alpha\} \vdash \neg\alpha$, pero en caso contrario, por el teorema de deducción (y aquí usamos que α es una sentencia) tendríamos $\Gamma \vdash \alpha \to \neg\alpha$, es decir, $\Gamma \vdash \neg\alpha \vee \neg\alpha$. Por consiguiente $\Gamma \vdash \neg\alpha$, en contra de lo supuesto. ∎

Por ejemplo, es equivalente decir que el quinto postulado de Euclides es independiente de los demás axiomas de la geometría (es decir, que no se puede demostrar a partir de ellos) que decir que la negación del quinto postulado es consistente con los demás axiomas de la geometría. El lector debería asegurarse de que no concibe el teorema anterior como un mero resultado técnico, sino como un hecho básico que a menudo se usa tácitamente para pasar de un planteamiento en términos de independencia a otro en términos de consistencia y viceversa.

Ejercicio: Mostrar que el teorema anterior es falso si α no es una sentencia. (Considerar, por ejemplo, $\Gamma = \{\bigvee xy\, \neg x = y\}$, $\alpha \equiv x = y$.)

Otro resultado elemental es que la consistencia puede perderse al añadir axiomas a una teoría pero, desde luego, nunca al quitarlos. Esto es lo que afirma en esencia el teorema siguiente:

Teorema 3.13 *Sean Δ y Γ colecciones de fórmulas de un lenguaje formal $\mathcal{L}$ tales que toda fórmula de Δ sea consecuencia de Γ (esto sucede en particular si Δ está contenida en Γ). Entonces, si Γ es consistente Δ también lo es y si Δ es contradictoria Γ también lo es.*

DEMOSTRACIÓN: Una contradicción deducida de las premisas de Δ puede probarse a partir de Γ deduciendo primero todas las premisas empleadas (lo cual es posible por hipótesis) y después continuando con el mismo razonamiento. ∎

Observemos que hemos justificado la consistencia de $K_{\mathcal{L}}$ probando que hay fórmulas que no son teoremas lógicos, lo cual a su vez lo hemos obtenido considerando modelos concretos. Esto puede generalizarse:

Definición 3.14 Un *modelo* de una teoría axiomática es un modelo de sus axiomas, es decir, un modelo de su lenguaje formal en el que son verdaderos todos sus axiomas.

Del teorema 3.9 se sigue que si M es un modelo de una teoría T, entonces todos los teoremas de T son verdaderos en M (aunque puede haber fórmulas verdaderas en M que no sean teoremas de T). Recíprocamente, ninguna fórmula falsa en M puede ser un teorema de T. Puesto que siempre hay fórmulas falsas en un modelo dado, tenemos el teorema siguiente:

Teorema 3.15 *Si una teoría axiomática tiene un modelo, entonces es consistente.*

Advertimos al lector que teorema siguiente no es aceptado como tal por algunos matemáticos: utiliza esencialmente un argumento no finitista:

Teorema 3.16 *La aritmética de Peano es consistente.*

DEMOSTRACIÓN: Basta probar que el modelo natural descrito tras la definición 3.1 es un modelo de la aritmética de Peano. La comprobación es rutinaria (o, si se prefiere, inmediata). Detallemos, por ejemplo, la comprobación de una instancia arbitraria del esquema de inducción, es decir, veamos que

$$\mathcal{M} \vDash \alpha(0) \wedge \bigwedge x(\alpha(x) \to \alpha(x')) \to \bigwedge x\alpha(x).$$

Fijamos una valoración cualquiera v, por ejemplo la constante igual a 0 (aunque esto es irrelevante). Hemos de probar que si

$$\mathcal{M} \vDash \alpha(0)[v] \qquad \text{y} \qquad \mathcal{M} \vDash \big(\bigwedge x(\alpha(x) \to \alpha(x'))\big)[v],$$

entonces $\mathcal{M} \vDash \bigwedge x\alpha(x)$.

La primera parte de la hipótesis es que $\mathcal{M} \vDash \mathtt{S}_x^0\alpha[v]$, que por 3.4 equivale a que $\mathcal{M} \vDash \alpha[v_x^0]$, donde el 0 que aparece junto a v no es la constante 0, sino el número natural 0.

Veamos por inducción que, para todo natural n, se cumple $\mathcal{M} \vDash \alpha[v_x^n]$. Por definición de satisfacción tenemos que

$$\mathcal{M} \vDash (\alpha \to \mathtt{S}_x^{x'}\alpha)[v_x^n],$$

de donde se sigue que $\mathcal{M} \vDash \mathtt{S}_x^{x'}\alpha[v_x^n]$. Aplicando de nuevo 3.4, esto equivale a $\mathcal{M} \vDash \alpha[v_x^{n+1}]$, como queríamos probar. Por definición de satisfacción, tenemos probado que $\mathcal{M} \vDash \bigwedge x\,\alpha$. ∎

Observaciones Podría objetarse que la prueba anterior es vacía, en el sentido de que, por ejemplo, nos apoyamos en el principio de inducción para demostrar que $\mathcal{M}$ satisface el principio de inducción. Esto es cierto. Más concretamente, lo que sucede es que no hemos demostrado —ni hemos pretendido demostrar— el principio de inducción, sino únicamente que el significado de cada instancia del esquema axiomático P7 es un caso particular del principio de inducción (entendido no como una fórmula de un lenguaje formal, sino como un razonamiento lógicamente admisible). El principio de inducción no admite una demostración formal (salvo que partamos de axiomas más fuertes que el propio principio, como son los axiomas de la teoría de conjuntos). Si pudiera demostrarse, Peano no lo habría tomado como axioma; pero esto no significa que no tenga sentido plantearse si es una afirmación verdadera o falsa sobre los números naturales. En la prueba anterior hemos dado por sabido que es verdadero: ciertamente, si comprobamos que el cero cumple una propiedad y argumentamos que si un

número arbitrario la cumple es necesario que la cumpla el siguiente, entonces podemos estar seguros de que todos los números la cumplen o, dicho de otro modo, que es imposible que nos encontremos con una excepción: tal excepción no podría ser el 0, pues hemos argumentado que tiene la propiedad, ni podría ser el 1, pues sabemos argumentar que la tiene (combinando los dos pasos de la inducción), ni podría ser el 2, pues aplicando dos veces el segundo paso sabemos razonar que 2 cumple la propiedad, ni —en general— podría ser el 3 o el 4, etc.

La clave de la prueba de la consistencia de la aritmética no es esta comprobación rutinaria de que los axiomas se corresponden (a través de la definición de verdad) con afirmaciones verdaderas sobre los números naturales, sino el hecho en sí de que las afirmaciones sobre los números naturales tengan un significado objetivo que nos permita dividirlas en verdaderas y falsas. Esta objetividad se basa en que toda afirmación sobre los números naturales puede entenderse como una afirmación sobre unos algoritmos completamente determinados (el de contar, el de sumar y el de multiplicar). Así, por ejemplo, si podemos garantizar que el axioma que afirma que $x + 0 = x$ es verdadero, ello se debe a que cualquiera que conozca el algoritmo de la suma sabe que sumar 0 es no hacer nada, independientemente de cuál sea el número al que le sumamos 0. Cuando afirmamos que la ecuación $x^2 - 2y^2 = 5$ no tiene solución esto puede entenderse como que el algoritmo que a partir de x e y nos da $x^2 - 2y^2$ nunca puede dar 5 como resultado, y esto sabemos lo que quiere decir exactamente aunque no esté claro a priori si es cierto o es falso.

En conclusión, aunque existan afirmaciones sobre números naturales —como la última que hemos considerado— cuya verdad es problemática, las que hemos tomado como axiomas son sin duda verdaderas, lo cual se justifica, no por deducción formal a partir de otras premisas previas, sino por el análisis de unos algoritmos. Una vez sentada la verdad de estos axiomas, la corrección del cálculo deductivo nos garantiza que todos los teoremas aritméticos han de ser verdaderos. Más concretamente, cada demostración proporciona un argumento irrefutable que garantiza la verdad de su conclusión. Por consiguiente, una afirmación falsa como $0^{(2)} + 0^{(2)} = 0^{(5)}$ no puede ser un teorema de la aritmética de Peano. Si lo fuera, tendríamos un argumento irrefutable en virtud del cual al juntar dos canicas con dos canicas nos encontraríamos con cinco canicas.

Destaquemos también que, como ya hemos comentado, la prueba que hemos dado es esencialmente no finitista. Para que se cumplan tan sólo los dos primeros axiomas de la aritmética necesitamos un modelo con un universo infinito, es decir, nos estamos apoyando esencialmente en que podemos hablar consistentemente de la totalidad de los números naturales y no sólo de una cantidad arbitrariamente grande pero finita de números naturales. Éste es un buen momento para que el lector medite sobre si considera aceptable o no una prueba no finitista: la cuestión es si, después de considerar detenidamente el argumento anterior, aún cree posible que exista una demostración formal de la sentencia $0 \neq 0$ a partir de los axiomas de Peano o si, por el contrario, comprende que es imposible que exista tal prueba.

Por otra parte, debemos comentar que Gentzen y después Gödel han ob-

tenido pruebas finitistas de la consistencia de la aritmética, el primero usando la llamada "inducción hasta ϵ_0" y el segundo un formalismo lógico de orden infinito. Estas pruebas tienen la ventaja de que con ellas se obtienen algunos resultados importantes sobre la recursividad de la aritmética de Peano, pero no vamos a entrar en detalles.

Capítulo IV

La completitud semántica

En este capítulo justificaremos que el cálculo deductivo de primer orden, tal y como lo hemos introducido el el capítulo II, captura exactamente la noción de razonamiento matemático riguroso. Para ello demostraremos el teorema de completitud de Gödel, del cual se desprenden, a su vez, resultados interesantísimos sobre las limitaciones del razonamiento formal.

Todos los resultados que vamos a obtener se siguen del siguiente teorema de Gödel que, aunque por su contenido no se ajusta al nombre de "teorema de completitud", lo cierto es que es frecuente citarlo de esta forma, ya que técnicamente es mucho más útil que el teorema de completitud propiamente dicho. Se trata del recíproco del teorema 3.15:

Teorema 4.1 *Si una teoría axiomática es consistente, entonces tiene un modelo numerable.*

Un modelo numerable es un modelo tal que es posible ordenar los objetos de su universo, ya sea en una sucesión finita $a_0, \ldots, a_n$, ya infinita $a_0, a_1, a_2, \ldots$ Ya hemos comentado el alguna ocasión que consideramos dudoso que tenga sentido hablar metamatemáticamente de colecciones de objetos más generales que éstas, pero, precisamente, el teorema de incompletitud nos garantizará que no perdemos generalidad si tratamos únicamente con modelos numerables.

Hilbert consideraba que la consistencia de unos axiomas da derecho a considerar que existen unos objetos que los satisfacen. Esto es esencialmente lo que afirma el teorema anterior. La primera prueba se debe a Gödel, aunque la demostración que veremos aquí es de Henkin.

Debemos advertir que la prueba del teorema de completitud no es finitista, por lo que, dadas las profundas repercusiones, este teorema se convierte en una piedra de toque de la metamatemática no finitista, es decir, el lector se va a ver obligado a decidir si acepta definitivamente el planteamiento de la metamatemática que estamos presentando (si no una versión más fuerte) o renuncia a considerar como hechos probados y fuera de toda duda al teorema de completitud y a sus consecuencias.

4.1 Completitud sintáctica

Una diferencia notable entre la sintaxis (cálculo deductivo) y la semántica (modelos) es que una sentencia ha de ser verdadera o falsa en un modelo dado, mientras que no tiene por qué ser demostrable o refutable en una teoría dada. Para analizar esto con más detalle conviene introducir la noción de completitud (sintáctica) de una teoría axiomática:

Definición 4.2 Una teoría axiomática T es *completa* si toda sentencia α de su lenguaje formal cumple $\underset{T}{\vdash} \alpha$ o $\underset{T}{\vdash} \neg\alpha$.

Hay un caso en el que una teoría dada es indudablemente completa: si es contradictoria. Obviamente este caso carece de interés. Si la consistencia es lo mínimo que ha de cumplir una teoría axiomática para que tenga interés, la completitud es lo máximo que le podemos pedir. Una teoría completa es una teoría capaz de resolvernos cualquier duda sobre su objeto de estudio.

Vamos a probar que toda teoría axiomática puede extenderse hasta una teoría completa. Para ello hemos de definir qué entendemos por extensión de una teoría axiomática.

Definición 4.3 Diremos que una teoría axiomática S es una *extensión* de una teoría T si el lenguaje formal de S contiene a todos los signos del lenguaje de T y todos los axiomas de T son teoremas de S.

Es claro que si S es una extensión de T, entonces todos los teoremas de T son teoremas de S. La forma habitual de extender una teoría T consiste en formar una nueva teoría S que tenga más axiomas, pero la definición que hemos dado permite que se eliminen algunos axiomas de T y en su lugar se añadan otros axiomas más fuertes.

Para probar el teorema de completitud necesitaremos una versión refinada del teorema siguiente, pero incluimos también esta versión porque el resultado tiene interés en sí mismo y muestra más claramente la idea básica. Todas las cuestiones metamatemáticas sobre el carácter no finitista del teorema de completitud pueden tratarse equivalentemente al respecto de este teorema:

Teorema 4.4 *Toda teoría axiomática consistente tiene una extensión consistente y completa.*

Demostración: Sea T una teoría axiomática consistente sobre un lenguaje formal $\mathfrak{L}$. Sea Γ la colección de las clausuras universales de sus axiomas. Sea $\alpha_0, \alpha_1, \alpha_2, \ldots$ una enumeración de las sentencias de $\mathfrak{L}$. Definimos $\Gamma_0 = \Gamma$. Entonces es claro que para toda fórmula α se cumple $\Gamma \vdash \alpha$ si y sólo si $\Gamma_0 \vdash \alpha$. Por el teorema 3.13 tenemos que Γ_0 es una colección consistente de sentencias de $\mathfrak{L}$. Para cada número natural n, definimos

$$\Gamma_{n+1} = \begin{cases} \Gamma_n & \text{si } \Gamma_n \cup \{\alpha_n\} \text{ es contradictorio,} \\ \Gamma_n \cup \{\alpha_n\} & \text{si } \Gamma_n \cup \{\alpha_n\} \text{ es consistente.} \end{cases}$$

Por construcción todas las colecciones Γ_n son consistentes y si $m \leq n$ entonces Γ_m está contenido en Γ_n.

Sea Γ_∞ la unión de todas las colecciones Γ_n. Es claro que Γ_∞ es consistente, pues si de sus sentencias se dedujera una contradicción, ésta se deduciría de hecho de un número finito de ellas, y todas estarían contenidas en un cierto Γ_n, que sería, pues, contradictorio.

Sea S la teoría axiomática cuya colección de axiomas es Γ_∞. Ciertamente es consistente. Como Γ está contenido en Γ_∞, es claro que S es una extensión de T (los teoremas de T son las consecuencias de Γ, luego también son consecuencias de Γ_0 y de Γ_∞). Veamos por último que S es completa.

Sea α una sentencia de $\mathcal{L}$. Entonces $\alpha \equiv \alpha_i$ para un cierto i. Supongamos que no $\underset{S}{\vdash} \neg\alpha_i$, o sea, que no $\Gamma_\infty \vdash \neg\alpha_i$. Entonces tampoco $\Gamma_i \vdash \neg\alpha_i$. Por el teorema 3.12, la colección $\Gamma_i \cup \{\alpha_i\}$ es consistente, y así $\Gamma_{i+1} = \Gamma_i \cup \{\alpha_i\}$, luego α_i está en Γ_∞ y por consiguiente $\underset{S}{\vdash} \alpha_i$. ∎

Observaciones Este resultado es una consecuencia inmediata de 4.1, pues si T tiene un modelo M, basta tomar como S la teoría cuyos axiomas son las sentencias verdaderas en M. Sin embargo, ya hemos comentado que necesitamos una versión más fuerte de este teorema para probar 4.1.

La enumeración $\alpha_0, \alpha_1, \ldots$ de las sentencias de un lenguaje formal puede hacerse explícitamente. Más adelante veremos detalladamente una forma de hacerlo (a través de la numeración de Gödel), por lo que por ahora no insistiremos más en ello. Lo importante es que una tal ordenación es un concepto finitista.

Un punto mucho más delicado es la definición de las colecciones Γ_n, pues, según veremos más adelante, en los casos de interés matemático no es posible calcular explícitamente cada uno de sus términos. Aunque pudiéramos calcular los primeros, digamos hasta Γ_5, nada nos garantiza que seamos capaces de determinar quién es Γ_6 o, más concretamente, si la sentencia α_5 forma parte o no de Γ_6. El problema es que para ello tendríamos que decidir si $\Gamma_5 \cup \{\alpha_5\}$ es o no consistente, y no tenemos ningún algoritmo que nos permita decidir si una colección de fórmulas, aunque sea finita, es consistente o no.

Pese a ello, lo cierto es que $\Gamma_5 \cup \{\alpha_5\}$ será consistente o contradictoria y, según el caso Γ_6 coincidirá con esta colección o se reducirá a Γ_5. Puesto que uno de los dos casos, y sólo uno, ha de ser cierto, podemos afirmar que Γ_6 está bien definida con independencia de si sabemos o no determinar sus elementos.

Así pues, la colección Γ_∞ de los axiomas de la extensión S está completamente determinada por T y por la ordenación de las sentencias de $\mathcal{L}$ que hemos escogido, a pesar de que no sabemos determinar qué sentencias contiene. Estamos ante el tipo de colecciones de objetos más general que vamos a considerar desde un punto de vista metamatemático.

La teoría S es bastante "patológica", pues, aunque conozcamos perfectamente la teoría de partida T, lo cierto es que no sabemos qué sentencias son axiomas de S y, a fortiori, qué sentencias son teoremas de S. Más adelante

veremos que esta patología es inevitable si partimos de una teoría consistente en la que puedan demostrarse los axiomas de Peano.

A diferencia de lo que sucede con el teorema de completitud, este teorema afirma simplemente la existencia de un objeto bien definido que escapa a nuestro control. En sí mismo no tiene repercusiones finitistas. Por ello no es un resultado indicado para valorar si tenemos realmente motivos para aceptar razonamientos no finitistas. Ciertamente, si las técnicas no finitistas nos llevaran únicamente a conclusiones de este tipo, resultarían ser totalmente superfluas. ∎

El papel que desempeña la completitud en la prueba del teorema 4.1 es, a grandes rasgos, el siguiente: un modelo de una teoría axiomática determina si una sentencia dada es verdadera o falsa. Por consiguiente, para construir un modelo debemos contar con toda la información necesaria para aceptar o rechazar cualquier sentencia y, para ello, uno de los primeros pasos que daremos será completar la teoría de partida. Si nos fijamos en la teoría S construida en la prueba del teorema anterior veremos que una sentencia es un teorema de S si y sólo si es un axioma. Para no trabajar con teorías "hinchadas" de axiomas, conviene tratar directamente con la colección de las sentencias demostrables en una teoría axiomática, ahorrándonos así el darles artificialmente rango de axiomas. Ello nos lleva al concepto siguiente:

Definición 4.5 Una colección Γ de sentencias de un lenguaje formal $\mathcal{L}$ es *maximalmente consistente* si Γ es consistente y para toda sentencia α de $\mathcal{L}$ que no esté en Γ se cumple que $\Gamma \cup \{\alpha\}$ es contradictoria.

La relación con la completitud es la siguiente:

Teorema 4.6 *Una teoría axiomática T es consistente y completa si y sólo si la colección Γ de todas las sentencias demostrables en F es maximalmente consistente.*

Demostración: Supongamos que T es consistente y completa. Entonces Γ es consistente, pues las consecuencias lógicas de los teoremas de T son teoremas de T, luego si a partir de Γ pudiera demostrarse α y $\neg\alpha$, estas fórmulas serían teoremas de T.

Sea α una sentencia que no esté en Γ, es decir, tal que no $\underset{T}{\vdash} \alpha$. Como T es completa, $\underset{T}{\vdash} \neg\alpha$, luego $\neg\alpha$ está en Γ. Es claro entonces que $\Gamma \cup \{\alpha\}$ es contradictorio. Por lo tanto Γ es maximalmente consistente.

Recíprocamente, si Γ es maximalmente consistente, entonces T es consistente, pues hay sentencias que no son teoremas de T (las negaciones de las sentencias de Γ). Además, si α es una sentencia, o bien $\underset{F}{\vdash} \alpha$ o, en caso contrario, α no está en Γ, luego $\Gamma \cup \{\alpha\}$ es contradictorio luego, por el teorema 3.12, tenemos que $\Gamma \vdash \neg\alpha$, y a su vez esto implica que $\underset{T}{\vdash} \neg\alpha$. ∎

El teorema siguiente recoge las propiedades básicas de las colecciones maximalmente consistentes. Notemos que en él aparecen conexiones estrictamente sintácticas (es decir, no basadas en ningún modelo) entre los signos lógicos y su significado.

Teorema 4.7 *Sea Γ una colección maximalmente consistente de sentencias de un lenguaje formal $\mathfrak{L}$ y α, β dos sentencias de $\mathfrak{L}$. Entonces*

a) $\Gamma \vdash \alpha$ syss α está en Γ,

b) Si $\vdash \alpha$, entonces α está en Γ,

c) $\neg\alpha$ está en Γ syss α no está en Γ,

d) $\alpha \to \beta$ está en Γ syss α no está en Γ o β está en Γ,

e) $\alpha \vee \beta$ está en Γ syss α está en Γ o β está en Γ,

f) $\alpha \wedge \beta$ está en Γ syss α está en Γ y β está en Γ,

g) $\alpha \leftrightarrow \beta$ está en Γ syss α y β están ambas en Γ o ninguna lo está.

DEMOSTRACIÓN: a) Si $\Gamma \vdash \alpha$ entonces no $\Gamma \vdash \neg\alpha$, porque Γ es consistente, luego $\Gamma \cup \{\alpha\}$ es consistente, por el teorema 3.12, luego α está en Γ. El recíproco es obvio.

b) Es consecuencia de a).

c) Si $\neg\alpha$ está en Γ, entonces α no puede estar en Γ, porque Γ es consistente. Si α no está en Γ entonces $\Gamma \cup \{\alpha\}$ es contradictorio, luego por el teorema 3.12 se cumple que $\Gamma \vdash \neg\alpha$ y por a) concluimos que $\neg\alpha$ está en Γ.

d) Si $\alpha \to \beta$ está en Γ y α está en Γ, entonces $\Gamma \vdash \beta$, luego por 1) concluimos que β está en Γ.

Si α no está en Γ o β está en Γ, por c) $\neg\alpha$ está en Γ o β está en Γ. Por consiguiente $\Gamma \vdash \neg\alpha$ o $\Gamma \vdash \beta$. En cualquier caso $\Gamma \vdash \alpha \to \beta$ y por a) $\alpha \to \beta$ está en Γ.

e), f) y g) se deducen de c) y d) por las definiciones de los conectores. ∎

Necesitamos propiedades análogas a las de este teorema pero en relación a los cuantificadores. Para ello necesitamos una nueva noción.

Definición 4.8 Una colección Γ de sentencias de un lenguaje formal F es *ejemplificada* si cuando $\bigvee x\alpha$ está en Γ existe un designador t de $\mathfrak{L}$ tal que $\mathtt{S}^t_x\alpha$ está en Γ.

Es decir, Γ está ejemplificada si cuando afirma la existencia de un objeto que cumple algo, nos da también un ejemplo concreto t de objeto que cumple lo indicado. En realidad se cumple mucho más:

Teorema 4.9 *Sea Γ una colección de sentencias de un lenguaje formal $\mathfrak{L}$ maximalmente consistente y ejemplificada. Sea α una fórmula de $\mathfrak{L}$ en la que a lo sumo esté libre la variable x. Entonces*

a) $\bigvee x\alpha$ está en Γ syss existe un designador t de $\mathfrak{L}$ tal que $\mathtt{S}^t_x\alpha$ está en Γ.

b) $\bigwedge x\alpha$ está en Γ syss para todo designador t de $\mathfrak{L}$ se cumple que $\mathtt{S}^t_x\alpha$ está en Γ.

DEMOSTRACIÓN: a) Si $\bigvee x\alpha$ está en Γ, hay un designador t de $\mathfrak{L}$ tal que $\mathtt{S}^t_x\alpha$ está en Γ por ser Γ ejemplificada.

Si $\mathtt{S}^t_x\alpha$ está en Γ, por (IP) obtenemos que $\Gamma \vdash \bigvee x\alpha$, con lo que $\bigvee x\alpha$ está en Γ por el teorema anterior.

b) Si $\bigwedge x\alpha$ está en Γ y t es un designador de $\mathfrak{L}$, por (EG) se cumple $\Gamma \vdash \mathtt{S}^t_x\alpha$ y por consiguiente $\mathtt{S}^t_x\alpha$ está en Γ (por el teorema anterior).

Si $\mathtt{S}^t_x\alpha$ está en Γ para todo designador t de $\mathfrak{L}$, entonces el teorema anterior nos da que $\neg\mathtt{S}^t_x\alpha \equiv \mathtt{S}^t_x\neg\alpha$ no está en Γ para todo t. Por a) $\bigvee x\neg\alpha$ no está en Γ y por el teorema 4.7 otra vez concluimos que $\neg\bigvee x\neg\alpha$ sí lo esta. Aplicando (NP) resulta que $\Gamma \vdash \bigvee x\alpha$, luego, por el teorema anterior una vez más, concluimos que $\bigwedge x\alpha$ está en Γ. ∎

Con esto tenemos todos los conceptos necesarios para demostrar el teorema 4.1. Nos dedicamos a ello en la sección siguiente.

4.2 La prueba del teorema de completitud

El teorema 4.4 puede reformularse como que toda colección consistente de sentencias está contenida en una colección maximalmente consistente. Sin embargo, necesitaremos una colección que además sea ejemplificada y ello plantea un problema técnico que hemos de resolver previamente. La clave será el teorema siguiente:

Teorema 4.10 *Sea $\mathfrak{L}$ un lenguaje formal y sea $\mathfrak{L}'$ un lenguaje formal que conste de los mismos signos que $\mathfrak{L}$ más una constante c (aunque admitimos el caso de que c esté en $\mathfrak{L}$ y, por consiguiente, que $\mathfrak{L}$ coincida con $\mathfrak{L}'$. Si c no está en una fórmula α, la variable x no está ligada en α y $\underset{K_{\mathfrak{L}'}}{\vdash} \mathtt{S}^c_x\alpha$, entonces $\underset{K_{\mathfrak{L}}}{\vdash} \alpha$.*

DEMOSTRACIÓN: Como x no está ligada en α, tenemos que x no está (ni libre ni ligada) en $\mathtt{S}^c_x\alpha$. Si en una demostración de $\mathtt{S}^c_x\alpha$ en $K_{\mathfrak{L}'}$ sustituimos todas las intervenciones de la variable x por otra variable que no aparezca en la demostración, obtenemos una demostración de $\mathtt{S}^c_x\alpha$ en la que no aparece la variable x. Veamos que $\underset{K_{\mathfrak{L}}}{\vdash} \alpha$ por inducción sobre el número de líneas de una demostración en estas condiciones.

Si $\mathtt{S}^c_x\alpha$ se demuestra en una línea, entonces es un axioma de $K_{\mathfrak{L}'}$. Veamos que α ha de ser un axioma de $K_{\mathfrak{L}}$. Para probarlo nos basaremos en dos hechos obvios:

A Si θ es una expresión de $\mathfrak{L}'$ que no contiene la variable x, entonces $\theta \equiv \mathtt{S}^c_x\theta_0$, para una expresión θ_0 de $\mathfrak{L}$ que no contiene a c. (θ_0 es la expresión que resulta de cambiar por x cada aparición de c en θ.)

B Si β_0 y β_1 son fórmulas de $\mathfrak{L}$ que no contienen a c y $\mathtt{S}^c_x\beta_0 \equiv \mathtt{S}^c_x\beta_1$ entonces $\beta_0 \equiv \beta_1$.

Así, si por ejemplo $\mathtt{S}^c_x\alpha \equiv \beta \to (\gamma \to \beta)$, entonces por A

$$\mathtt{S}^c_x\alpha \equiv \beta \to (\gamma \to \beta) \equiv \mathtt{S}^c_x\beta_0 \to (\mathtt{S}^c_x\gamma_0 \to \mathtt{S}^c_x\beta_0) \equiv \mathtt{S}^c_x\big(\beta_0 \to (\gamma_0 \to \beta_0)\big),$$

luego por B tenemos que $\alpha \equiv \beta_0 \to (\gamma_0 \to \beta_0)$.

Similarmente, si $\mathtt{S}^c_x\alpha \equiv \bigwedge y\, \beta \to \mathtt{S}^t_y\beta$, entonces

$$\mathtt{S}^c_x\alpha \equiv \bigwedge y\, \mathtt{S}^c_x\beta_0 \to \mathtt{S}^{\mathtt{S}^c_x t_0}_y \mathtt{S}^c_x\beta_0 \equiv \bigwedge y\, \mathtt{S}^c_x\beta_0 \to \mathtt{S}^c_x\mathtt{S}^{t_0}_y\beta_0 \equiv \mathtt{S}^c_x(\bigwedge y\, \beta_0 \to \mathtt{S}^{t_0}_y\beta_0),$$

luego $\alpha \equiv \bigwedge y\, \beta_0 \to \mathtt{S}^{t_0}_y\beta_0$.

La comprobación para los restantes esquemas axiomáticos es análoga.

Si el teorema es cierto para las fórmulas demostrables en menos de n pasos, supongamos que $\mathtt{S}^c_x\alpha$ se demuestra en n pasos.

a) Si $\mathtt{S}^c_x\alpha$ se deduce por (MP) de β y $\beta \to \mathtt{S}^c_x\alpha$, líneas anteriores. Por A) podemos expresar $\beta \equiv \mathtt{S}^c_x\gamma$, donde γ no contiene a c.

Observemos que $\beta \to \mathtt{S}^c_x\alpha \equiv \mathtt{S}^c_x\gamma \to \mathtt{S}^c_x\alpha \equiv \mathtt{S}^c_x(\gamma \to \alpha)$, la constante c no está en γ ni en $\gamma \to \alpha$ y x no está ligada en γ ni en $\gamma \to \alpha$. Por hipótesis de inducción $\underset{K_{\mathcal{L}}}{\vdash} \gamma$ y $\underset{K_{\mathcal{L}}}{\vdash} \gamma \to \alpha$, luego $\underset{K_{\mathcal{L}}}{\vdash} \alpha$.

b) Si $\mathtt{S}^c_x\alpha \equiv \bigwedge y\beta$ se deduce de β por (IG), entonces $y \not\equiv x$, pues x no aparece en la demostración. Sea γ la fórmula resultante de sustituir c por x en β. Como antes $\beta \equiv \mathtt{S}^c_x\gamma$. Además $\mathtt{S}^c_x\alpha \equiv \bigwedge y\mathtt{S}^c_x\gamma \equiv \mathtt{S}^c_x\bigwedge y\gamma$. De aquí se sigue que $\alpha \equiv \bigwedge y\gamma$. Por hipótesis de inducción $\underset{K_{\mathcal{L}}}{\vdash} \gamma$, luego $\underset{K_{\mathcal{L}}}{\vdash} \alpha$. ∎

En definitiva, la prueba del teorema muestra que basta reemplazar todas las apariciones de c por apariciones de x en una demostración de $\mathtt{S}^c_x\alpha$ para tener una demostración de α.

El teorema siguiente es el que nos permitirá ejemplificar una colección consistente de sentencias para volverla a la vez ejemplificada y maximalmente consistente.

Teorema 4.11 *Si $\Gamma \cup \{\bigvee x\, \alpha\}$ es una colección consistente de sentencias de un lenguaje formal $\mathcal{L}$, el lenguaje $\mathcal{L}'$ es como en el teorema anterior y la constante c no está en ninguna sentencia de $\Gamma \cup \{\bigvee x\, \alpha\}$, entonces $\Gamma \cup \{\bigvee x\, \alpha\} \cup \{\mathtt{S}^c_x\alpha\}$ es consistente.*

Demostración: Si $\Gamma \cup \{\bigvee x\, \alpha\} \cup \{\mathtt{S}^c_x\alpha\}$ es contradictorio, por el teorema 3.12 tenemos que $\Gamma \cup \{\bigvee x\, \alpha\} \underset{K_{\mathcal{L}'}}{\vdash} \neg\mathtt{S}^c_x\alpha$.

Existen $\gamma_1, \ldots, \gamma_n$ en Γ tales que $\gamma_1 \wedge \cdots \wedge \gamma_n \wedge \bigvee x\, \alpha \underset{K_{\mathcal{L}'}}{\vdash} \neg\mathtt{S}^c_x\alpha$. Sea y una variable que no esté en $\gamma_1 \wedge \cdots \wedge \gamma_n \wedge \bigvee x\, \alpha$. Entonces

$$\gamma_1 \wedge \cdots \wedge \gamma_n \wedge \bigvee x\, \alpha \underset{K_{\mathcal{L}'}}{\vdash} \mathtt{S}^c_y\mathtt{S}^y_x\neg\alpha,$$

luego por el teorema de deducción

$$\underset{K_{\mathcal{L}'}}{\vdash} \gamma_1 \wedge \cdots \wedge \gamma_n \wedge \bigvee x\, \alpha \rightarrow \mathtt{S}_y^c \mathtt{S}_x^y \neg\alpha,$$

y esto equivale a

$$\underset{K_{\mathcal{L}'}}{\vdash} \mathtt{S}_y^c(\gamma_1 \wedge \cdots \wedge \gamma_n \wedge \bigvee x\, \alpha \rightarrow \mathtt{S}_x^y \neg\alpha),$$

pues y no está libre en $\gamma_1 \wedge \cdots \wedge \gamma_n \wedge \bigvee x\, \alpha$.

Por el teorema anterior $\underset{K_{\mathcal{L}}}{\vdash} \gamma_1 \wedge \cdots \wedge \gamma_n \wedge \bigvee x\, \alpha \rightarrow \mathtt{S}_x^y \neg\alpha$, y de aquí que $\Gamma \cup \{\bigvee x\, \alpha\} \underset{K_{\mathcal{L}}}{\vdash} \neg \mathtt{S}_x^y \alpha$. Aplicando (IG) y (NP) llegamos a que $\Gamma \cup \{\bigvee x\, \alpha\} \underset{K_{\mathcal{L}}}{\vdash} \neg \bigvee y \mathtt{S}_x^y \alpha$, de donde se concluye que $\Gamma \cup \{\bigvee x\, \alpha\} \underset{K_{\mathcal{L}}}{\vdash} \neg \bigvee x\, \alpha$, con lo que $\Gamma \cup \{\bigvee x\, \alpha\}$ resulta ser contradictoria. ∎

Aunque el teorema siguiente no lo necesitaremos hasta un poco más adelante, lo incluimos aquí porque su prueba es completamente análoga a la del teorema anterior.

Teorema 4.12 *Si Γ es una colección consistente de sentencias de un lenguaje formal con descriptor $\mathcal{L}$, el lenguaje $\mathcal{L}'$ es como en el teorema anterior y la constante c no está en $\mathcal{L}$, entonces $\Gamma \cup \{c = x|(x = x)\}$ es consistente.*

Demostración: Si $\Gamma \cup \{c = x|(x = x)\}$ es contradictorio, por el teorema 3.12 tenemos que $\Gamma \underset{K_{\mathcal{L}'}}{\vdash} \neg c = x|(x = x)$. Existen sentencias $\gamma_1, \ldots, \gamma_n$ en Γ tales que $\gamma_1 \wedge \cdots \wedge \gamma_n \underset{K_{\mathcal{L}'}}{\vdash} \neg c = x|(x = x)$ y por el teorema de deducción

$$\underset{K_{\mathcal{L}'}}{\vdash} \gamma_1 \wedge \cdots \wedge \gamma_n \rightarrow \neg c = x|(x = x).$$

Sea y una variable que no esté en la sentencia $\gamma_1 \wedge \cdots \wedge \gamma_n \rightarrow \neg c = x|(x = x)$.

Tenemos que $\underset{K_{\mathcal{L}'}}{\vdash} \mathtt{S}_y^c(\gamma_1 \wedge \cdots \wedge \gamma_n \rightarrow \neg y = x|(x = x))$, luego por el teorema 4.10 también $\underset{K_{\mathcal{L}}}{\vdash} \gamma_1 \wedge \cdots \wedge \gamma_n \rightarrow \neg y = x|(x = x)$. Así pues, $\Gamma \underset{K_{\mathcal{L}}}{\vdash} \neg y = x|(x = x)$. Aplicando (IG) resulta que $\Gamma \underset{K_{\mathcal{L}}}{\vdash} \bigwedge y \neg y = x|(x = x)$, y aplicando (EG) llegamos a $\Gamma \underset{K_{\mathcal{L}}}{\vdash} \neg(x|(x = x)) = (x|(x = x))$, de donde se sigue que Γ es contradictorio. ∎

Ahora ya podemos probar la base de nuestro argumento hacia el teorema de completitud:

Teorema 4.13 *Sea $\mathcal{L}$ un lenguaje formal, sea $\mathcal{L}'$ un lenguaje formal que conste de los mismos signos que $\mathcal{L}$ más una colección de constantes $d_0, d_1, \ldots$ que no estén en $\mathcal{L}$, sea Γ una colección consistente de sentencias de $\mathcal{L}$. Existe una colección Γ_∞ maximalmente consistente y ejemplificada de sentencias de $\mathcal{L}'$ que contiene a Γ.*

DEMOSTRACIÓN: Sea $\alpha_0, \alpha_1, \ldots$ una enumeración de las sentencias de $\mathcal{L}$. Definimos $\Gamma_0 = \Gamma$ y, supuesto definido Γ_n, sea

$$\Gamma_{n+1} = \begin{cases} \Gamma_n & \text{si } \Gamma_n \cup \{\alpha_n\} \text{ es contradictorio,} \\ \Gamma_n \cup \{\alpha_n\} & \text{si } \Gamma_n \cup \{\alpha_n\} \text{ es consistente y } \alpha_n \text{ no es de la forma } \bigvee x\, \beta, \\ \Gamma_n \cup \{\alpha_n\} \cup \{\mathtt{S}_x^{d_k}\beta\} & \text{si } \Gamma_n \cup \{\alpha_n\} \text{ es consistente, } \alpha_n \equiv \bigvee x\, \beta \text{ y } k \text{ es el menor natural tal que } d_k \text{ no está en } \Gamma_n \cup \{\alpha_n\}. \end{cases}$$

Por el teorema 4.11, cada Γ_n es consistente. Sea Γ_∞ la unión de todas las colecciones Γ_n. Como en el teorema 4.4, es claro que Γ_∞ es consistente. De hecho es maximalmente consistente, pues si una sentencia $\alpha \equiv \alpha_i$ de $\mathcal{L}'$ no está en Γ_∞, entonces $\Gamma_i \cup \{\alpha_i\}$ es contradictorio o, de lo contrario, α_i estaría en Γ_{i+1}, luego en Γ_∞. Por consiguiente $\Gamma_\infty \cup \{\alpha_i\}$ también es contradictorio.

Por último veamos que Γ_∞ es ejemplificada. Si $\bigvee x\, \alpha$ está en Γ_∞, entonces $\bigvee x\, \alpha \equiv \alpha_j$ para algún natural j. Como $\Gamma_j \cup \{\bigvee x\, \alpha\}$ está contenido en Γ_∞, ciertamente es consistente. Por construcción $\Gamma_{j+1} = \Gamma_j \cup \{\bigvee x\, \alpha\} \cup \{\mathtt{S}_x^{d_k}\alpha\}$, luego $\mathtt{S}_x^{d_k}\alpha$ está en Γ_∞ y d_k es un designador de $\mathcal{L}'$. ∎

La colección de sentencias Γ_∞ construida en la prueba del teorema anterior tiene las mismas características que la construida en 4.4, es decir, está unívocamente determinada a partir de Γ y de la enumeración fijada de las sentencias de $\mathcal{L}$, pero no tenemos ningún algoritmo que nos permita decidir si una sentencia dada está o no en Γ_∞.

A partir de una colección de sentencias Γ ejemplificada y maximalmente consistente es fácil construir un modelo. Para ello hemos de buscar una colección de objetos que, con las relaciones y funciones adecuadas, verifiquen cuanto se afirma en Γ. Ahora bien, puesto que Γ proporciona designadores concretos que nombran a cada uno de los objetos de los que pretende hablar, podemos tomar como universo del modelo los propios designadores del lenguaje formal de Γ, y arreglar las definiciones de forma que cada designador se denote a sí mismo. Aquí nos aparece un problema técnico, y es que si Γ contiene una sentencia de tipo $t_1 = t_2$, donde t_1 y t_2 son designadores distintos, entonces t_1 y t_2 deben denotar al mismo objeto, lo cual no sucederá si, como pretendemos, cada uno se denota a sí mismo. Para corregir este inconveniente no tomaremos como universo a los designadores exactamente, sino a las clases de equivalencia respecto a la relación que satisfacen dos designadores t_1 y t_2 precisamente cuando $t_1 = t_2$ está en Γ. Veamos los detalles:

Teorema 4.14 *Sea $\mathcal{L}$ un lenguaje formal (con o sin descriptor) y sea Γ una colección consistente de sentencias sin descriptores de $\mathcal{L}$. Entonces Γ tiene un modelo (de universo) numerable.*

DEMOSTRACIÓN: Recordemos que $\underline{\mathcal{L}}$ es el lenguaje que resulta de eliminar el descriptor de $\mathcal{L}$. Sea Γ_∞ una colección de sentencias de $\underline{\mathcal{L}}'$ maximalmente consistente y ejemplificada que contenga a Γ según el teorema anterior. Basta

probar que Γ_∞ tiene un modelo numerable M_∞, pues entonces ciertamente $M_\infty \vDash \Gamma$ y si llamamos M al modelo de $\underline{\mathfrak{L}}$ que se diferencia de M_∞ en que no interpreta las constantes nuevas, es claro que $M \vDash \Gamma$. Finalmente, seleccionando si es necesario una descripción impropia, podemos considerar a M como modelo de $\mathfrak{L}$ (pues la descripción impropia no influye en la interpretación de fórmulas sin descriptores).

Equivalentemente, podemos suponer que $\mathfrak{L}$ no tiene descriptor y que Γ es una colección maximalmente consistente y ejemplificada de sentencias de $\mathfrak{L}$.

Sea T la colección de todos los designadores de $\mathfrak{L}$. Consideramos en T la relación diádica dada por $t_1 \sim t_2$ syss la sentencia $t_1 = t_2$ está en Γ. Veamos que se trata de una relación de equivalencia.[1]

Dado un designador t, ciertamente $\vdash t = t$, luego $t = t$ está en Γ por el teorema 4.7. Por consiguiente la relación es reflexiva.

Dados dos designadores t_1 y t_2 tales que $t_1 \sim t_2$, esto significa que $t_1 = t_2$ está en Γ, luego $\Gamma \vdash t_2 = t_1$ y por consiguiente $t_2 = t_1$ está también en Γ. Esto prueba la simetría y similarmente se prueba la transitividad.

Representaremos por $[t]$ a la clase de equivalencia de t respecto a $\sim$ y llamaremos U a la colección de todas las clases de equivalencia. Es claro que U es una colección numerable, ya que lo es la colección de los designadores de $\mathfrak{L}$.

Definimos como sigue un modelo M de $\mathfrak{L}$:

- El universo de M es U.
- Si c es una constante de $\mathfrak{L}$, entonces $M(c) = [c]$.
- Si R_i^n es un relator n-ádico de $\mathfrak{L}$, entonces $M(R_i^n)$ es la relación n-ádica dada por

$$M(R_i^n)([t_1], \ldots, [t_n]) \quad \text{syss} \quad R_i^n t_1 \cdots t_n \text{ está en } \Gamma.$$

- Si f_i^n es un funtor n-ádico de $\mathfrak{L}$, entonces $M(f_i^n)$ es la función n-ádica en U dada por

$$M(f_i^n)([t_1], \ldots, [t_n]) = [f_i^n t_1 \cdots t_n].$$

Hemos de comprobar que las interpretaciones de los relatores y los funtores están bien definidas. En el caso de los relatores esto significa que si se cumple $[t_1] = [t'_1], \ldots, [t_n] = [t'_n]$ entonces $R_i^n t_1 \cdots t_n$ está en Γ syss $R_i^n t'_1 \cdots t'_n$ está en Γ.

En efecto, tenemos que las sentencias $t_1 = t'_1$, $\ldots$, $t_n = t'_n$ están en Γ, de donde se sigue fácilmente que $\Gamma \vdash R_i^n t_1 \cdots t_n \leftrightarrow R_i^n t'_1 \cdots t'_n$. Por la consistencia maximal, la sentencia $R_i^n t_1 \cdots t_n \leftrightarrow R_i^n t'_1 \cdots t'_n$ está en Γ, y por el teorema 4.7 tenemos que $R_i^n t_1 \cdots t_n$ está en Γ syss $R_i^n t'_1 \cdots t'_n$ está en Γ.

Análogamente se prueba que las funciones $M(f_i^n)$ están bien definidas. También hemos de comprobar que $M(=)$ es la relación de igualdad, pero esto es inmediato a partir de las definiciones.

[1] Si el lector no está familiarizado con las relaciones de equivalencia debe consultar el apéndice A.

Ahora probamos que si t es un designador de $\mathcal{L}$ y v es cualquier valoración, entonces $M(t)[v] = [t]$. Si t es una constante esto es cierto por definición de M. Como $\mathcal{L}$ no tiene descriptores, la única posibilidad adicional es que $t \equiv f_i^n t_1 \cdots t_n$, en cuyo caso, razonando por inducción sobre la longitud de t, podemos suponer que $M(t_i)[v] = [t_i]$, y por definición de $M(f_i^n)$ llegamos a que $M(t)[v] = [t]$.

Finalmente probamos que si α es una sentencia de $\mathcal{L}$, entonces $M \vDash \alpha$ si y sólo si α está en Γ. En particular $M \vDash \Gamma$, que es lo que queríamos probar. Razonamos por inducción sobre el número de signos lógicos ($\neg$, $\to$, $\bigwedge$) que contiene α. Notar que los términos de $\mathcal{L}$ no tienen descriptores, por lo que no pueden contener signos lógicos.

Si α no tiene signos lógicos, entonces $\alpha \equiv R_i^n t_1 \cdots t_n$ y se cumple que

$$M \vDash \alpha \text{ syss } M(R_i^n)([t_1], \ldots, [t_n]) \text{ syss } \alpha \equiv R_i^n t_1 \cdots t_n \text{ está en } \Gamma.$$

Supuesto cierto para sentencias con menos signos lógicos que α, distinguimos tres casos:

a) Si $\alpha \equiv \neg\beta$, entonces $M \vDash \alpha$ syss no $M \vDash \beta$ syss (hip. de ind.) β no está en Γ syss (teorema 4.7) $\neg\beta$ está en Γ syss α está en Γ.

b) Si $\alpha \equiv \beta \to \gamma$, entonces $M \vDash \alpha$ syss no $M \vdash \beta$ o $M \vdash \gamma$ syss (hip. ind.) β no está en Γ o γ está en Γ syss (teorema 4.7) $\beta \to \gamma$ está en Γ syss α está en Γ.

c) Si $\alpha \equiv \bigwedge x\beta$, entonces $M \vDash \alpha$ syss para todo $[t]$ de U (y cualquier valoración v en M) se cumple $M \vDash \beta[v_x^{[t]}]$ syss para todo designador t de $\mathcal{L}$ (y cualquier valoración) $M \vDash \beta[v_x^{M(t)[v]}]$ syss (por 3.4) para todo designador t de $\mathcal{L}$ se cumple $M \vDash \mathtt{S}_x^t\beta$.

Notemos que las sentencias $\mathtt{S}_x^t\beta$ —aunque no tienen necesariamente menor longitud que α— tienen menos signos lógicos, luego podemos aplicarles la hipótesis de inducción: $M \vDash \alpha$ syss $S_x^t\beta$ está en Γ para todo designador t de $\mathcal{L}$ syss (por el teorema 4.9) $\alpha \equiv \bigwedge x\beta$ está en Γ. ∎

Ahora sólo nos queda el problema técnico de ir eliminando hipótesis en el teorema anterior. En primer lugar nos ocupamos de la restricción sobre los descriptores:

Teorema 4.15 *Sea $\mathcal{L}$ un lenguaje formal (con o sin descriptor) y Γ una colección consistente de sentencias de $\mathcal{L}$. Entonces Γ tiene un modelo numerable.*

Demostración: Sea $\mathcal{L}'$ un lenguaje formal que conste de los mismos signos que $\mathcal{L}$ más una nueva constante c. Sea x una variable. Por el teorema 4.12 tenemos que $\Gamma \cup \{c = x|(x = x)\}$ es consistente.

Por el teorema 2.14, para cada sentencia γ de Γ existe una sentencia γ' de $\mathcal{L}'$ sin descriptores y tal que

$$c = x|(x = x) \vdash \gamma \leftrightarrow \gamma'.$$

Sea Δ la colección de todas estas sentencias γ'. Toda sentencia de Δ es consecuencia de $\Gamma \cup \{c = x|(x = x)\}$. Por el teorema 3.13 tenemos que Δ es consistente. Por el teorema anterior Δ tiene un modelo M' de universo numerable. Como las sentencias de Δ no tienen descriptores, si cambiamos la descripción impropia en M' éste no deja de ser un modelo de Δ. Tomamos concretamente $M'(c)$ como descripción impropia, de modo que ahora se cumple $M' \vDash c = x|(x = x)$.

Así pues, $M' \vDash \Delta \cup \{c = x|(x = x)\}$ y, como toda fórmula de Γ es consecuencia de $\Delta \cup \{c = x|(x = x)\}$, por 3.9 tenemos que $M' \vDash \Gamma$. Finalmente, sea M el modelo de $\mathcal{L}$ que se diferencia de M' en que no interpreta la constante c. Claramente M es numerable y $M \vDash \Gamma$. ∎

Finalmente eliminamos la restricción de que las fórmulas sean sentencias, con lo que tenemos 4.1 y, combinándolo con 3.15, tenemos el teorema siguiente:

Teorema 4.16 *Una colección Γ de fórmulas de un lenguaje formal $\mathcal{L}$ es consistente si y sólo si tiene un modelo (que podemos tomar numerable).*

Demostración: Supongamos que Γ es consistente. Sea Γ^c la colección de las clausuras universales de las fórmulas de Γ. Como todas las sentencias de Γ^c se deducen de las de Γ, el teorema 3.13 nos da que Γ^c es consistente. Por el teorema anterior Γ^c tiene un modelo numerable M que, claramente, también es un modelo de Γ. El recíproco es 3.15 ∎

4.3 Consecuencias del teorema de completitud

Del teorema 4.16 se sigue que el cálculo deductivo que hemos introducido en principio de forma arbitraria es exactamente lo que tiene que ser. Esto se sigue de los recíprocos de los teoremas 3.8 y 3.9, que pasamos a demostrar:

Teorema 4.17 (Teorema de adecuación) *Toda fórmula lógicamente válida es un teorema lógico.*

Demostración: Sea α una fórmula y α^c su clausura universal. Si no $\vdash \alpha$, entonces no $\vdash \alpha^c$, luego no $\vdash \neg\neg\alpha^c$. Por el teorema 3.12 obtenemos que $\neg\alpha^c$ es consistente, luego tiene un modelo M, es decir, $M \vDash \neg\alpha^c$, luego no $M \vDash \alpha^c$, luego no $M \vDash \alpha$, es decir, α no es lógicamente válida. ∎

Nota Es crucial comprender que el enunciado de este teorema carecería de un significado metamatemático preciso si no fuera por la propia demostración del teorema. Recordemos que no sabemos dar un sentido a una afirmación del tipo "α es lógicamente válida" salvo en el caso en que dispongamos de un argumento que nos convenza de que α ha de ser verdadera en cualquier modelo. El teorema de corrección nos da que esto sucede siempre que α es un teorema lógico. Lo que ahora hemos probado es que si α no es un teorema lógico (y está claro qué significa esto) entonces sabemos construir un modelo en el que α es falsa, y

esto podemos expresarlo diciendo que α no es lógicamente válida. En resumen, ahora sabemos que toda fórmula α ha de encontrarse en uno de los dos casos siguientes: o bien es un teorema lógico y entonces es verdadera en cualquier modelo, o bien no es un teorema lógico y entonces sabemos describir un modelo concreto en el cual es falsa. Esto nos permite considerar como equivalentes las afirmaciones $\vdash \alpha$ y $\models \alpha$, con lo que, dado que la primera tiene un significado preciso, lo mismo podemos decir, a partir de ahora, de la segunda. ∎

La adecuación del cálculo deductivo queda plasmada más claramente en lo que propiamente se conoce como el *teorema de completitud semántica para la lógica de primer orden:*

Teorema 4.18 (Teorema de completitud semántica (de Gödel)) *Sea T una teoría axiomática consistente y sea α una fórmula de su lenguaje formal. Si α es verdadera en todo modelo (numerable) de T, entonces $\underset{T}{\vdash} \alpha$.*

Demostración: Sea Γ la colección de los axiomas de T. Hemos de probar que $\Gamma \vdash \alpha$. En caso contrario no $\Gamma \vdash \alpha^c$, luego no $\Gamma \vdash \neg\neg\alpha^c$. Por el teorema 3.12 tenemos que $\Gamma \cup \{\neg\alpha^c\}$ es consistente, luego tiene un modelo numerable M. Como $M \models \Gamma$ se cumple que M es un modelo de T, pero no $M \models \alpha$, en contra de lo supuesto. ∎

Así pues, si el teorema 3.9 garantizaba que el cálculo deductivo jamás nos lleva de premisas verdaderas a conclusiones falsas, el teorema de completitud semántica nos garantiza que el cálculo deductivo es completo, no en el sentido sintáctico de que nos responda afirmativa o negativamente a cualquier pregunta, sino en el sentido semántico de que cualquier otro cálculo deductivo "más generoso" que permitiera deducir más consecuencias que el nuestro de unas premisas dadas, necesariamente nos permitiría deducir consecuencias falsas en un modelo a partir de premisas verdaderas en él, por lo que no sería semánticamente aceptable. En resumen, ahora sabemos que nuestro cálculo deductivo se corresponde exactamente con la noción metamatemática de razonamiento lógico, por lo que todas las arbitrariedades de su definición están ahora plenamente justificadas.

En la prueba del teorema de completitud hemos visto un ejemplo representativo de la necesidad que puede surgir en diversos contextos de razonar con fórmulas sin descriptores. Como primera aplicación del teorema de completitud veremos que no sólo podemos eliminar los descriptores de los axiomas y teoremas (encontrando fórmulas equivalentes sin descriptores) sino también de las demostraciones.

Teorema 4.19 *Sea $\mathcal{L}$ un lenguaje formal con descriptor, sea Γ una colección de fórmulas de $\mathcal{L}$ sin descriptores y α una fórmula sin descriptores. Si se cumple $\Gamma \vdash \alpha$, entonces existe una deducción de α a partir de Γ en la que no aparecen descriptores.*

Demostración: Si M es un modelo de Γ (considerando a Γ como colección de fórmulas de $\underline{\mathcal{L}}$), determinando arbitrariamente una descripción impropia obtenemos un modelo M' de $\mathcal{L}$ que obviamente cumple $M' \models \Gamma$. Por consiguiente

$M' \vDash \alpha$ y, como α no tiene descriptores, $M \vDash \alpha$. Así pues, α (como fórmula de $\underline{\mathcal{L}}$) es verdadera en todo modelo de Γ (como colección de fórmulas de $\underline{\mathcal{L}}$). Por el teorema de completitud (para $\underline{\mathcal{L}}$) concluimos que $\Gamma \underset{K_{\underline{\mathcal{L}}}}{\vdash} \alpha$, es decir, α se deduce sin descriptores. ∎

Nota En la prueba del teorema anterior hemos usado el teorema de completitud para un lenguaje sin descriptor. Éste ha sido el motivo por el que hasta ahora hemos trabajado tanto con lenguajes con y sin descriptor. A partir de este momento todos los lenguajes formales que consideremos tendrán descriptor. ∎

Aritmética no estándar Si el teorema de completitud nos ha mostrado que el cálculo deductivo es exactamente lo que tiene que ser, a la vez nos muestra ciertas limitaciones que, por esta misma razón, resultan ser esenciales a toda posible formalización y axiomatización de una teoría matemática.

Observemos que si una colección de fórmulas Γ tiene la propiedad de que todas sus subcolecciones finitas son consistentes, entonces es consistente. En efecto, si a partir de Γ se dedujera una contradicción, en la deducción sólo podría aparecer una cantidad finita de premisas, las cuales formarían una subcolección finita de Γ contradictoria, en contra de lo supuesto. El teorema de completitud traduce este hecho obvio en un hecho nada trivial:

Teorema 4.20 (Teorema de compacidad de Gödel) *Una colección de fórmulas tiene un modelo si y sólo si lo tiene cada una de sus subcolecciones finitas.*

Lo importante en este teorema es que ninguno de los modelos de ninguna de las subcolecciones finitas tiene por qué ser un modelo de la totalidad de las fórmulas y, pese a ello, podemos garantizar que existe un modelo que cumple simultáneamente todas ellas.

De aquí se deduce que la lógica de primer orden no es *categórica,* es decir, que —en la mayoría de casos de interés— es imposible caracterizar unívocamente unos objetos que pretendamos estudiar a través de una colección de axiomas. Concretamente vamos a probarlo con las nociones de "finitud" y de "número natural".

Nosotros hemos presentado los números naturales como los objetos 0, 1, 2, 3, 4, 5, etc., es decir, los objetos generados por un proceso de cómputo perfectamente determinado que nos permite continuar indefinidamente y sin vacilación la sucesión anterior. Así aprenden todos los niños lo que son los números naturales y esta definición les basta para manejarlos en todos los contextos distintos del de la matemática formal. Muchos matemáticos piensan que esta noción "intuitiva", en el más despectivo sentido de la palabra, puede ser suficiente para usos no sofisticados, como contar monedas, o sellos, o piedras, pero no para las matemáticas serias, donde es necesaria una definición más precisa y rigurosa de número natural. Ahora vamos a probar que esto, aunque tiene algo de cierto,

también tiene mucho de falso. Es verdad que la matemática, desde el momento en que pretende estudiar objetos abstractos que involucran la noción general de conjunto, requiere ser axiomatizada en su totalidad, lo cual incluye el tratar axiomáticamente los números naturales. Sin embargo, no es cierto que una presentación axiomática de los números naturales sea más precisa y rigurosa que una presentación no axiomática como la que hemos dado aquí. Al contrario, vamos a demostrar que una presentación axiomática de los números naturales será rigurosa, pero nunca precisa, en el sentido de que será necesariamente ambigua.

En efecto supongamos que el lector cree que puede definir con total precisión los números naturales en el seno de una teoría axiomática. El intento más simple es la aritmética de Peano que ya hemos presentado, pero si el lector considera que es demasiado débil, podemos admitir cualquier otra teoría. Por ejemplo podríamos pensar en una teoría axiomática de conjuntos. No importa cuáles sean sus axiomas en concreto. El argumento que vamos a emplear se aplica a cualquier teoría axiomática T que cumpla los siguientes requisitos mínimos:

- T es consistente (es obvio que una teoría contradictoria no sería una buena forma de presentar los números naturales)
- El lenguaje de T contiene un designador 0, un término x' con x como única variable libre y una fórmula $\mathrm{Nat}\, x$ (léase "x es un número natural") con x como única variable libre de modo que en T puedan demostrarse los teoremas siguientes:

  a) $\mathrm{Nat}\ 0$,

  b) $\bigwedge x(\mathrm{Nat}\, x \to \mathrm{Nat} x')$,

  c) $\bigwedge x(\mathrm{Nat}\, x \to x' \neq 0)$,

  d) $\bigwedge xy(\mathrm{Nat}\, x \wedge \mathrm{Nat}\, y \wedge x' = y' \to x = y)$.

En otras palabras, admitimos que en la teoría T se definan los números naturales, el cero y la operación "siguiente" como se considere oportuno, con tal de que se puedan demostrar las cuatro propiedades elementales que hemos exigido. Si T es la aritmética de Peano, como definición de número natural sirve $\mathrm{Nat}\ x \equiv x = x$, pues en T sólo se puede hablar de números naturales. Si T es una teoría más general entonces $\mathrm{Nat}\ x$ ha de ser una fórmula que especifique qué objetos son números naturales y cuáles no, es decir, lo que cualquier matemático entendería por una "definición de número natural".

Si es posible determinar axiomáticamente los números naturales, la forma de hacerlo será, sin duda, una teoría T que cumpla los requisitos anteriores. Ahora probaremos que existen unos objetos que satisfacen la definición de número natural que ha propuesto el lector —cualquiera que ésta sea— y que, pese a ello, nadie en su juicio los aceptaría como números naturales. Más precisamente, vamos a construir un modelo de T en el que existen objetos que satisfacen la definición de número natural del lector y que son distintos de lo que el lector ha decidido llamar 0, y de lo que el lector ha decidido llamar 1, etc.

Sea $\mathcal{L}$ el lenguaje de T y sea $\mathcal{L}'$ el lenguaje que resulta de añadir a $\mathcal{L}$ una constante c. Consideramos la teoría T' sobre $\mathcal{L}'$ que resulta de añadir a los axiomas de T la siguiente colección de sentencias:

$$\mathrm{Nat}\, c, \quad c \neq 0, \quad c \neq 0', \quad c \neq 0'', \quad c \neq 0''', \quad c \neq 0'''', \quad \ldots$$

Si usamos en T la misma notación que en la aritmética de Peano, hemos añadido el axioma Nat c más los axiomas $c \neq 0^{(n)}$, para todo natural n.

La teoría T' es consistente. En virtud del teorema de compacidad basta encontrar un modelo de cada colección finita de axiomas de T. De hecho, es claro que basta encontrar un modelo de cada teoría T'_n formada por los axiomas de T más los axiomas Nat c, $c \neq 0$, …, $c \neq 0^{(n)}$.

Por hipótesis T es consistente, luego tiene un modelo M. Llamaremos M_n al modelo de $\mathcal{L}'$ que es igual que M salvo por que interpreta la constante c como el objeto denotado por $0^{(n+1)}$. Así, M_n es un modelo de T en el que además es verdadera la sentencia $c = 0^{(n+1)}$. De esta sentencia más las sentencias a), b), c), d) que estamos suponiendo que son teoremas de T se deducen las sentencias Nat c, $c \neq 0$, …, $c \neq 0^{(n)}$, es decir, M_n es un modelo de T'_n.

Por el teorema de compacidad T' tiene un modelo M'. En particular M' es un modelo de T, es decir, sus objetos cumplen todos los axiomas que el lector ha considerado razonables. En particular, los objetos a de M' que cumplen $M' \vDash \mathrm{Nat}\, x[v^a_x]$, para una valoración cualquiera v, satisfacen todos los requisitos que el lector ha tenido bien exigir a unos objetos para que merezcan el calificativo de números naturales.

Llamemos $\xi = M'(c)$. Puesto que $M' \vDash \mathrm{Nat}\, c$, tenemos que ξ es uno de esos objetos que el lector está dispuesto a aceptar como números naturales. Ahora bien, Como $M' \vDash c \neq 0$, tenemos que ξ es distinto del objeto denotado por el designador 0, es decir, es distinto del objeto que satisface todo lo que el lector ha tenido a bien exigir para que merezca el calificativo de "número natural cero". Similarmente, como $M' \vDash c \neq 0'$, tenemos que ξ es distinto de lo que el lector a tenido a bien llamar 1, etc. En resumen, la definición de número natural propuesta por el lector es satisfecha por unos objetos entre los cuales hay uno ξ que no es lo que el lector ha llamado 0, ni 1, ni 2, ni 3, ni, en general, ningún número que pueda obtenerse a partir del 0 por un número finito de aplicaciones de la operación siguiente. Cualquier niño de 10 años al que se le explique esto adecuadamente comprenderá que el lector se equivoca si cree haber definido correctamente los números naturales.

En general, diremos que un modelo M de una teoría que satisface los requisitos que hemos exigido a T es un modelo *no estándar* de la aritmética si en su universo hay un objeto ξ tal que, para una valoración v cualquiera, $M \vDash \mathrm{Nat}\, x[v^\xi_x]$ y para todo número natural n se cumple $M \vDash x \neq 0^{(n)}[v^\xi_x]$. A tales objetos ξ los llamaremos *números no estándar* del modelo M.

Hemos probado que cualquier formalización mínimamente razonable de la aritmética tiene modelos no estándar, modelos en los que hay "números naturales" que no pueden obtenerse a partir del cero en un número finito de pasos.

Vemos así que el razonamiento metamatemático que estamos empleando desde el primer capítulo, aunque inútil para tratar con la matemática abstracta, es mucho más preciso que el razonamiento axiomático formal a la hora de tratar con objetos intuitivamente precisos. Así, aunque la noción de finitud es totalmente precisa y rigurosa, tan simple que hasta un niño de 10 años comprende sin dificultad que hay un número finito de dedos en la mano pero hay infinitos números naturales, resulta que el más sofisticado aparato matemático es incapaz de caracterizarla con precisión.

En efecto, nosotros nunca hemos dado una definición de finitud, pues si el lector no supiera perfectamente lo que es ser finito debería entretenerse leyendo libros más elementales que éste. Ahora bien, el lector no sólo debe ser consciente de que él ya sabe lo que es ser finito, sino que además debe comprender que no estamos "siendo poco rigurosos" al eludir una definición formal de finitud, ya que no se puede pecar de poco riguroso por no hacer algo imposible. Supongamos que el lector se siente capaz de corregirnos y enunciar una definición razonable de "conjunto finito". Sin duda, para ello deberá hacer uso de algunas propiedades elementales de los conjuntos. Todo cuanto utilice podrá enunciarse explícitamente como los axiomas de una teoría de conjuntos T. No importa cuál sea la teoría T. Pongamos que el lector construye el lenguaje formal que considere oportuno y en él enuncia unos axiomas que digan "los conjuntos cumplen esto y lo otro". Sólo exigimos las siguientes condiciones mínimas:

- T es consistente.
- En T pueden definirse los números naturales en las mismas condiciones que antes, y además ha de poder demostrarse un principio de inducción similar al esquema axiomático de la aritmética de Peano. También ha de ser posible definir la relación de orden en los números naturales y demostrar sus propiedades básicas.
- En T puede definirse una fórmula "x es un conjunto finito" con la cual pueda probarse que todo conjunto con un elemento es finito y que si a un conjunto finito le añadimos un elemento obtenemos un conjunto finito. Por lo demás, el lector es libre de exigir cuanto estime oportuno a esta definición para que sea todo lo exacta que considere posible.
- En T tiene que poder demostrarse que para cada número natural x existe el conjunto de los números naturales menores o iguales que x.

Si se cumplen estos requisitos, en la teoría T puede probarse que el conjunto de los números naturales menores o iguales que un número n es finito, es decir, que satisface la definición de finitud que ha decidido adoptar el lector. Ahora bien, la teoría T tiene un modelo M no estándar, en el cual podemos considerar el conjunto Ξ de todos los números naturales menores o iguales que un número no estándar fijo ξ. Este conjunto Ξ satisface, pues, la definición de finitud del lector, pero contiene al objeto que en M satisface la definición de 0 (ya que ξ no es 0 y en T ha de poder probarse que todo número distinto de 0 es mayor que 0), y también contiene a lo que el lector ha llamado 1 (ya que ξ es distinto

de 1 y en T ha de poder probarse que todo números mayor que 0 y distinto de 1 es mayor que 1) y ha de contener a lo que el lector ha llamado 2, y 3, y 4, etc. En definitiva, tenemos un conjunto infinito que satisface la definición de finitud que haya propuesto el lector, cualquiera que ésta sea.

Las nociones de finitud y de número natural están íntimamente relacionadas: si tuviéramos una definición formal precisa de finitud podríamos definir los números naturales definiendo el 0 y la operación siguiente y estipulando que ésta ha de aplicarse un número finito de veces para obtener cada número natural; recíprocamente, si tuviéramos una definición formal precisa de los números naturales podríamos definir a partir de ella la noción de finitud; pero sucede que no existe ni lo uno ni lo otro, lo cual a su vez no es obstáculo para que cualquier niño de 10 años —al igual que el lector— tenga una noción precisa (metamatemática, no axiomática) de lo que es la finitud y de lo que son los números naturales.

Por otra parte, el lector debe tener presente que todos los teoremas de la aritmética de Peano, o de otra teoría similar, son afirmaciones verdaderas sobre los números naturales. Lo que hemos probado es que también son afirmaciones verdaderas sobre otros objetos que no son los números naturales, pero esto no contradice a lo primero, que es lo que realmente importa. Más en general, una teoría axiomática con axiomas razonables nos permite probar cosas razonables, independientemente que que pueda aplicarse también a objetos no razonables.

Aquí el lector se encuentra nuevamente ante un dilema: o concede que el tratamiento metamatemático que estamos dando a los números naturales es legítimo, o concluye que todo lo dicho en este apartado es, no ya falso, sino un completo sinsentido. Por supuesto, los números naturales son simplemente el ejemplo más simple. Lo mismo se puede decir de cualquier concepto de naturaleza "numerable", como puedan ser los números enteros y racionales, las sucesiones finitas de números racionales, los polinomios con coeficientes racionales, los números algebraicos, los grupos finitos, etc. En teoría es posible trabajar metamatemáticamente con todos estos conceptos, aunque en muchos casos puede ser delicado y requerir una extrema atención para no caer en palabras sin significado. Nadie dice que convenga hacerlo, pues la alternativa de trabajar en una teoría axiomática es mucho más ventajosa, pero lo cierto es que es posible. Nosotros sólo trataremos con los estrictamente imprescindibles para estudiar la lógica matemática, donde el uso de una teoría axiomática nos llevaría a un círculo vicioso.

La paradoja de Skolem Veamos ahora una última consecuencia del teorema de completitud del que a su vez se siguen implicaciones muy profundas sobre la naturaleza del razonamiento matemático. En realidad no es nada que no sepamos ya: se trata de enfatizar la numerabilidad de los modelos que sabemos construir. Según el teorema 4.16, una colección de fórmulas tiene un modelo si y sólo si tiene un modelo numerable. Equivalentemente:

Teorema 4.21 (Teorema de Löwenheim-Skolem) *Una teoría axiomática tiene un modelo si y sólo si tiene un modelo numerable.*

En definitiva, este teorema garantiza que no perdemos generalidad si trabajamos con modelos numerables, los únicos que en realidad sabemos entender metamatemáticamente. Lo sorprendente de este resultado estriba en que los matemáticos están convencidos de que en sus teorías tratan con conjuntos no numerables.

Consideremos una teoría axiomática de conjuntos T. No importa cuál en concreto, no importa cuáles sean sus axiomas exactamente. Basta con saber que en ella pueden demostrarse todos los teoremas matemáticos. Las teorías de conjuntos usuales disponen de un relator diádico $\in$ de tal modo que una fórmula como $x \in y$ se interpreta como que x es un elemento del conjunto y. En T puede definirse el conjunto $\mathcal{P}\mathbb{N}$ de todos los subconjuntos del conjunto de los números naturales, y el conocido teorema de Cantor afirma que $\mathcal{P}\mathbb{N}$ no es numerable. Recordemos el argumento: si existiera una aplicación biyectiva $f : \mathbb{N} \longrightarrow \mathcal{P}\mathbb{N}$, podríamos definir el conjunto $A = \{x \in \mathbb{N} \mid x \notin f(x)\}$, el cual nos lleva a una contradicción: debe existir un número $n \in \mathbb{N}$ tal que $A = f(n)$, pero entonces, si $n \in A = f(n)$, por definición de A debería ser $n \notin A$, pero si $n \notin A = f(n)$, por definición de A debería ser $n \in A$.

Supongamos que la teoría T es consistente (en caso contrario deberíamos buscar otra). Entonces T tiene un modelo numerable M. Digamos que los elementos de su universo son

$$c_0, \quad c_1, \quad c_2, \quad c_3, \quad c_4, \quad c_5, \quad \ldots$$

Estos objetos, con las relaciones y funciones adecuadas, satisfacen todos los axiomas y teoremas de la teoría de conjuntos, por lo que podemos llamarlos "conjuntos". A lo largo de este apartado, la palabra "conjunto" se referirá a los objetos c_n y a nada más. Retocando la enumeración si es preciso, podemos suponer que $c_0 = M(\mathbb{N})$, es decir, c_0 es el único objeto que satisface la definición de "conjunto de los números naturales". Así mismo podemos suponer que sus elementos son los conjuntos c_{2^n}, para $n \geq 1$. Concretamente, c_2 es el conjunto que satisface la definición de número natural 0, c_4 el que satisface la definición de número natural 1, c_8 la de 2, etc. Más precisamente,

$$M(0^{(n)}) = c_{2^{n+1}}.$$

También estamos afirmando que $M(\in)(c_n, c_0)$ si y sólo si $n = 2^{k+1}$ para algún k. Con esto estamos suponiendo tácitamente que M es un modelo estándar, es decir, que no tiene números naturales infinitos. No tendría por qué ser así, pero vamos a suponerlo por simplicidad.

No perdemos generalidad si suponemos que $M(\mathcal{P}\mathbb{N}) = c_1$, es decir, que c_1 es el único conjunto que tiene por elementos exactamente a todos los subconjuntos de $\mathbb{N}$, (de c_0). Así mismo podemos suponer que los elementos de c_1 son los conjuntos de la forma c_{3^n}, para $n \geq 1$. Así, c_3 podría ser el conjunto de los

números pares, c_9 el conjunto de los números primos, c_{27} el conjunto vacío, c_{81} el conjunto de los números menores que 1000, etc.[2]

Más concretamente, estamos suponiendo que si un conjunto c_n es un subconjunto de c_0, es decir, si todo c_i que cumpla $M(\in)(c_i, c_n)$ cumple también $M(\in)(c_i, c_0)$, entonces $n = 3^{k+1}$, así como que $M(\in)(c_n, c_1)$ si y sólo si se cumple $n = 3^{k+1}$.

La llamada paradoja de Skolem consiste en que este modelo que estamos describiendo existe realmente, y ello no contradice el hecho de que $\mathcal{P}\mathbb{N}$, es decir, el conjunto cuyos elementos son c_3, c_9, c_{27}, etc. es un conjunto no numerable: no es posible biyectar sus elementos con los números naturales.

Quien crea ver una contradicción en todo esto necesita aclararse algunas ideas confusas. Por ejemplo, una presunta contradicción que probara que en este modelo $\mathcal{P}\mathbb{N}$ sí que es numerable sería considerar la "biyección" $n \mapsto c_{3^{n+1}}$. Pero esto no es correcto. La sentencia de T que afirma que $\mathcal{P}\mathbb{N}$ no es numerable se interpreta en M como que no existe ninguna biyección entre los conjuntos que en M satisfacen la definición de número natural y los conjuntos que en M satisfacen la definición de subconjunto de $\mathbb{N}$. En nuestro caso, lo que tendríamos que encontrar es una biyección entre los elementos de c_0 y los elementos de c_1, es decir, entre los conjuntos c_2, c_4, c_8, etc. y los conjuntos c_3, c_9, c_{27}, etc.

Quizá el lector ingenuo aún crea ver una biyección en estas condiciones, a saber, la dada por $c_{2^k} \mapsto c_{3^k}$. Pero esto tampoco es una biyección. Una biyección entre dos conjuntos es un conjunto que satisface la definición de biyección: un conjunto de pares ordenados cuyas primeras componentes estén en el primer conjunto, sus segundas componentes en el segundo conjunto y de modo que cada elemento del primer conjunto está emparejado con un único elemento del segundo y viceversa. Tratemos de conseguir eso. Ante todo, un teorema de la teoría de conjuntos afirma que dados dos conjuntos x e y existe un único conjunto z tal que $z = (x, y)$, es decir, z es el par ordenado formado por x e y en este orden. Este teorema tiene que cumplirse en nuestro modelo M. Si lo aplicamos a los conjuntos c_{2^k} y c_{3^k}, concluimos que tiene que haber otro conjunto, y reordenando los índices podemos suponer que es c_{5^k}, tal que $M \vDash z = (x, y)[v]$, donde v es cualquier valoración que cumpla $v(x) = c_{2^k}$, $v(y) = c_{3^k}$ y $v(z) = c_{5^k}$.

Así, el "conjunto" formado por los conjuntos c_5, c_{25}, c_{125}, ... *sería* una biyección entre c_0 y c_1, es decir, entre $\mathbb{N}$ y $\mathcal{P}\mathbb{N}$. Lo sería... si fuera un conjunto.

Estamos al borde de la contradicción, pero no vamos a llegar a ella. Tendríamos una contradicción si la *colección* de conjuntos c_5, c_{25}, c_{125}, ... fuera la extensión de un conjunto, es decir, si existiera un conjunto, digamos c_r, cuyos elementos fueran exactamente los conjuntos $c_{5^{k+1}}$, es decir, si para algún r se cumpliera que $M(\in)(c_n, c_r)$ si y sólo si $n = 5^{k+1}$. En tal caso c_r sí que sería una biyección entre $\mathbb{N}$ y $\mathcal{P}\mathbb{N}$ y en M sería falso el teorema que afirma la no

[2]Así suponemos que ningún número natural está contenido en $\mathbb{N}$. De acuerdo con la construcción más habitual del conjunto de los números naturales sucede justo lo contrario: todo número natural es un subconjunto de $\mathbb{N}$, pero es posible construir los números naturales para que esto no suceda y hemos preferido evitar las confusiones que podría producir este tecnicismo.

numerabilidad de $\mathcal{P}\mathbb{N}$. Pero es que no tenemos nada que justifique ha de existir tal conjunto c_r. Justo al contrario, como sabemos que M es un modelo de la teoría de conjuntos T, podemos asegurar que tal c_r no puede existir.

Nos encontramos ante el quid de la cuestión. En general, si M es un modelo de la teoría de conjuntos, cada conjunto c —entendido como un objeto del universo de M— tiene asociada una colección de conjuntos, a saber, la colección de todos los conjuntos d tales que $M(\in)(d, c)$. A esta colección podemos llamarla la *extensión* del conjunto c. La extensión de un conjunto no es sino la colección de sus elementos. De este modo, todo conjunto tiene asociada una colección de conjuntos, pero no podemos afirmar que toda colección de conjuntos sea la extensión de un conjunto. La discusión anterior nos muestra que en todo modelo numerable de la teoría de conjuntos podemos formar colecciones de pares ordenados que biyecten $\mathbb{N}$ con $\mathcal{P}\mathbb{N}$, y ninguna de estas colecciones puede ser la extensión de un conjunto, o de lo contrario el modelo no cumpliría el teorema de Cantor.[3]

Otro ejemplo de esta situación lo proporcionan los modelos no estándar. Si M es un modelo no estándar de la teoría de conjuntos, podemos considerar la colección de todos los conjuntos que satisfacen la definición de número natural pero son números no estándar, es decir, no pueden obtenerse a partir del conjunto que satisface la definición de 0 aplicando un número finito de veces la definición de "siguiente". Tal colección no puede ser la extensión de ningún conjunto, pues un teorema elemental afirma que todo número natural no nulo tiene un inmediato anterior y, como los números no estándar son todos no nulos, resulta que todo número no estándar tiene un anterior. Así, si la colección de los números no estándar fuera la extensión de un conjunto, en M existiría un subconjunto no vacío de $\mathbb{N}$ sin un mínimo elemento, en contradicción con un conocido teorema de la teoría de conjuntos.

Esto no significa que los números no estándar no pertenezcan a ningún conjunto. Al contrario. Un teorema de la teoría de conjuntos afirma que, dado un número natural, existe el conjunto de todos los elementos menores o iguales que él. Si lo aplicamos a un número no estándar de M obtenemos un conjunto, es decir, un objeto del universo de M, que satisface la definición de conjunto finito y cuya extensión es —pese a ello— infinita, pues contiene entre otros a todos los números naturales estándar.

En resumen, al rastrear hasta su base la paradoja de Skolem encontramos que surge de una confusión: la confusión entre una colección (metamatemática) de conjuntos, como es c_5, c_{25}, c_{125}, ... y un conjunto (matemático), es decir, un objeto de un modelo de la teoría de conjuntos.

[3]Tal vez el lector familiarizado con la teoría de conjuntos piense al leer esto en la existencia de clases que no son conjuntos, como la clase de todos los conjuntos o la clase de todos los cardinales. Ciertamente, éste es el ejemplo más elemental del fenómeno del que estamos hablando, pero no el único. Una colección de pares ordenados que biyecte $\mathbb{N}$ con $\mathcal{P}\mathbb{N}$ no es una clase que no sea un conjunto, sino una colección "invisible" (= indefinible) para alguien que sólo vea los conjuntos del modelo (= para un matemático). De hecho, en teoría de conjuntos se prueba que toda subclase de un conjunto como $\mathbb{N} \times \mathcal{P}\mathbb{N}$ ha de ser un conjunto, luego la colección de la que hablamos ni siquiera es una clase.

4.4 Consideraciones finales

El lector no debe considerar anecdóticos o marginales los ejemplos de la sección anterior. Al contrario, contienen una parte importante de los hechos más profundos que vamos a estudiar en este libro. Probablemente, los irá asimilando cada vez mejor a medida que avancemos, pero para que así sea debería volver a meditar sobre ellos cada vez que encuentre nueva información relevante. La dificultad principal con la que se va a encontrar es que, a diferencia de lo que ocurre en contextos similares estrictamente matemáticos, lo necesario para comprenderlos cabalmente no es un mayor o menor grado de inteligencia, conocimientos o destreza, sino asimilar un determinado esquema conceptual mucho más rico que el que requiere la matemática formal.

Esta última sección pretende ser una ayuda para este fin. Afortunadamente, mientras la física moderna requiere pasar del esquema conceptual clásico a otro mucho más extraño, sutil y todavía no comprendido del todo, el esquema conceptual que requiere la lógica moderna —si bien distinto del que tradicionalmente han adoptado los matemáticos— no es extraño y novedoso, sino uno bien familiar y cotidiano.

Supongamos que hemos visto en el cine una película biográfica sobre Napoleón y al salir discutimos sobre ella. No tendremos ninguna dificultad en usar correctamente la palabra "Napoleón" a pesar de que tiene tres significados diferentes según el contexto: Napoleón-histórico, Napoleón-personaje y Napoleón-actor. Las afirmaciones sobre el primero son objetivas y semánticamente completas: o bien Napoleón padecía de gota o no padecía, con independencia de que sepamos cuál era el caso. El segundo es una creación del guionista de la película. Debe tener un cierto parecido con el Napoleón-histórico para que merezca el mismo nombre, pero tampoco tiene por qué coincidir con él. Por ejemplo, podría ser que el Napoleón-histórico padeciera de gota y el Napoleón-personaje no, o viceversa. Más aún, podría ocurrir que en la película no se hiciera ninguna alusión a si el Napoleón-personaje padecía o no gota, y en tal caso carece de sentido preguntar si esta afirmación es verdadera o falsa. Una película es sintácticamente incompleta: lo que no se dice explícita o implícitamente en ella no es verdadero ni falso, es indecidible. Por último, un mismo guión puede ser interpretado de forma diferente por actores diferentes. Los actores pueden precisar aspectos de los personajes que no están determinados por el guión. Nadie tiene dificultad en distinguir una crítica al guionista de una película por la mejor o peor caracterización de un personaje con una crítica a un actor por su mejor o peor interpretación del mismo.

Pues bien, afirmamos que el esquema conceptual necesario para interpretar adecuadamente los ejemplos de la sección anterior es exactamente el mismo que el que espontáneamente empleamos al discutir sobre una película. Estrictamente hablando, una demostración formal no es más que una sucesión de signos en un papel, igual que una película no es más que una sucesión de cuadrículas de celuloide coloreado, pero cuando leemos una demostración formal —al igual que cuando vemos una película— no vemos eso. Vemos una historia sobre unos personajes, los cuales a su vez pueden ser réplicas de objetos reales.

Los números naturales metamatemáticos son como el Napoleón-histórico, son objetos de los que podemos hablar objetivamente, que cumplen o no cumplen ciertas propiedades con independencia de que sepamos o no cuál es el caso. Al trabajar metamatemáticamente con ellos estamos investigándolos igual que un historiador puede investigar a Napoleón: reunimos la información que tenemos a nuestro alcance y a partir de ella tratamos de inferir hechos nuevos. Cuando decidimos formalizar la teoría de los números naturales hacemos como el novelista que prepara una novela histórica, o como el guionista de cine: diseñamos un personaje que pretende ser lo más parecido posible al original. La aritmética de Peano es una película sobre los números naturales. Podemos pensar objetivamente en sus protagonistas, es decir, tratarlos como si fueran objetos reales, al igual que podemos pensar objetivamente en Sherlock Holmes o en el pato Donald, pero debemos pensar que sólo son determinaciones parciales.

Notemos que hay tres clases de personajes de película o de novela: los históricos, que se ciñen a las características de un ser real, los personajes históricos novelados, que se basan en un personaje histórico pero han sido distorsionados por el autor (una caricatura de Napoleón, por ejemplo) y los ficticios, como Sherlock Holmes, sin ninguna relación con la realidad. Sin embargo, esta distinción es externa a la propia película, en el sentido de que un espectador que no sepa más que lo que la película le muestra será incapaz de distinguir a qué tipo pertenece cada personaje. Para hacer la distinción hemos de investigar la realidad y determinar si contiene objetos de características similares a los personajes.

Igualmente, podemos decir que los números naturales-matemáticos (= personajes) que aparecen en la aritmética de Peano son personajes históricos, porque todos los axiomas son afirmaciones verdaderas sobre los números naturales reales. Si extendemos la teoría para formar la aritmética no estándar obtenemos unos personajes históricos-novelados y, por último, una antigua discusión sobre la filosofía de las matemáticas puede enunciarse en estos términos como el dilema de si personajes como el conjunto de los números reales o los cardinales transfinitos son personajes históricos o ficticios. Después volveremos sobre este punto. Lo cierto es que, como meros espectadores, no podemos distinguirlos, pues podemos pensar con la misma objetividad y sentido de la realidad tanto acerca de Don Quijote como de Rodrigo Díaz de Vivar.

El modelo natural de la aritmética de Peano es la película perfecta: la película en la que cada personaje histórico se interpreta a sí mismo. No obstante, hemos visto que el mismo guión puede ser interpretado por actores esperpénticos, que se aprovechan de que el guión no dice explícitamente que no existen números no estándar. Lo peculiar de la situación es que, mientras es fácil exigir en el guión la existencia de naturales no estándar (añadiendo la constante c y los axiomas que dicen que es diferente de cualquier $0^{(n)}$), hemos probado que es imposible escribir un guión que exija la no existencia de números no estándar.

Un modelo numerable de la teoría de conjuntos es una película con efectos especiales. Tanto si queremos hacer una película sobre la llegada del hombre a la luna (hecho histórico) como si queremos hacerla sobre la llegada del hombre a Júpiter (ciencia-ficción), no podemos permitirnos filmar escenas reales y, en

ambos casos, tendremos que recurrir a los efectos especiales. Así pues, sin entrar en la discusión de si existe metamatemáticamente un conjunto no numerable como es $\mathcal{P}\mathbb{N}$, lo cierto es que podemos "simularlo" con efectos especiales.

Un técnico en efectos especiales puede hacer que una pequeña maqueta de plástico parezca una nave espacial, pero si por accidente se viera su mano en la escena, el espectador podría calcular el tamaño real de la "nave", y se daría cuenta de la farsa. Si M es un modelo de la teoría de conjuntos, podemos comparar a las colecciones de elementos de su universo con las personas que realizan la película, y las colecciones que constituyen la extensión de un conjunto con las personas que "se ven" en la pantalla. En el ejemplo de la sección anterior, c_0 es el actor que interpreta el papel de conjunto (= personaje) de los números naturales, mientras que la colección de los conjuntos c_{5^k} es un técnico en efectos especiales. Está ahí, pero, si se viera en escena, el espectador se daría cuenta de que $\mathcal{P}\mathbb{N}$ es en realidad una pequeña colección numerable, y no el conjunto inmenso que pretende parecer.

Si a un matemático le enseñamos únicamente los "actores" de M, es decir, los conjuntos, las colecciones que aparecen en escena, creerá estar viendo el universo del que hablan todos los libros de matemáticas, con sus conjuntos no numerables incluidos, pero si llegara a ver colecciones como la de los conjuntos c_{5^k} o la de los números naturales no estándar, si es que los hay, sería como si el espectador sorprendiera a Napoleón en manos de un maquillador. Estas colecciones "no existen" exactamente en el mismo sentido en que los maquilladores "no existen" a ojos del espectador. Napoleón-actor necesita ser maquillado, Napoleón-personaje no.[4]

En resumen, la mayor dificultad que el lector se encontrará a la hora de interpretar los resultados que hemos visto y vamos a ver, es la de reconocer significados diversos según el contexto en conceptos que para el matemático suelen tener un único significado (p.ej. la terna *colección metamatemática–conjunto matemático como concepto axiomático–conjunto como objeto metamatemático de un modelo concreto).* La única finalidad del juego de analogías que hemos desplegado es la de ayudar al lector a advertir qué distinciones van a ser necesarias y en qué han de consistir. Sin embargo, es importante tener presente que ninguna de estas analogías es un argumento. Todas estas distinciones deben ser entendidas y justificadas directamente sobre los conceptos que estamos tratando: números naturales, conjuntos, signos, etc. Por otra parte, no es menos cierto que —aunque esto no quede justificado sino a posteriori— los esquemas conceptuales son idénticos: cualquier problema conceptual sobre la naturaleza de $\mathcal{P}\mathbb{N}$ puede trasladarse a un problema idéntico sobre Sherlock Holmes y viceversa, y esto puede ser de gran ayuda.

[4]En esta comparación, las clases que no son conjuntos equivalen a personajes de los que se habla en la película e intervienen en la trama, pero que nunca aparecen en escena y, por consiguiente, no son encarnados por ningún actor. Es como Tutank-Amon en una película de arqueólogos. Ciertamente, no es lo mismo Tutank-Amon que un maquillador. El matemático puede hablar de los cardinales aunque no vea ningún conjunto que los contenga a todos, pero no puede hablar de una biyección fantasma entre $\mathbb{N}$ y $\mathcal{P}\mathbb{N}$.

Capítulo V

Teoría de la recursión

En este punto hemos completado la primera parte del programa de fundamentación de la matemática: tenemos una definición precisa de lo que debemos entender por razonamiento matemático, una definición que captura con total fidelidad la noción metamatemática de razonamiento lógico correcto. En el capítulo anterior hemos tenido ocasión de comprobar que al formalizar el razonamiento lógico perdemos capacidad de precisión, pues resulta imposible caracterizar formalmente nociones tan básicas como la de finitud o número natural.[1] Precisamente por ello es fundamental el teorema de completitud, pues nos garantiza que estos inconvenientes son inherentes al razonamiento formal y no son achacables a ninguna arbitrariedad en la que hayamos podido incurrir al definir el cálculo deductivo.

Para completar el proceso de fundamentación de la matemática nos falta determinar los axiomas específicos que vamos a admitir como punto de partida legítimo de las demostraciones matemáticas formales. No obstante, dejaremos esto para más adelante, porque sucede que, al disponer de una caracterización tan simple del razonamiento deductivo, podemos encontrar limitaciones aún más fuertes a cualquier intento de fundamentación de la matemática, y es importante mostrar que tales limitaciones no dependen de ninguna arbitrariedad en la elección de los axiomas. Por ello vamos a establecerlas en un contexto general, para teorías axiomáticas que cumplan unos requisitos mínimos, sin los cuales carecerían de valor. A esto dedicaremos los próximos capítulos. Más concretamente, nos encaminamos a demostrar los llamados *teoremas de incompletitud de Gödel* y sus consecuencias.

Al contrario de lo que sucedía con el teorema de completitud, los teoremas de incompletitud son totalmente finitistas y, más aún, constructivos. Para probarlos con el grado de generalidad que estamos indicando necesitamos establecer primero algunos resultados generales referentes a relaciones y funciones sobre los números naturales. Dedicamos a ello este capítulo, en el que expondremos

[1] Recordemos que esto no significa que no podamos hablar formalmente de conjuntos finitos o de números naturales. Lo que no podemos evitar es que los resultados que demostremos sobre ellos sean aplicables a objetos distintos de los que en principio pretendemos estudiar.

los rudimentos de lo que se conoce como teoría de la recursión.

Es muy importante tener presente que el contenido de este capítulo es enteramente metamatemático, es decir, nada de lo que digamos tiene ninguna relación con ningún sistema deductivo formal ni, en particular, con ningún lenguaje formal. No obstante, por conveniencia usaremos los signos lógicos $\neg$, $\wedge$, $\vee$, etc. como meras abreviaturas de "no", "y", "o", etc. La notación "$\mu x \ldots$" significará "el mínimo natural x tal que ... o bien 0 si no hay ninguno". Mientras no se indique lo contrario, todas las relaciones y funciones que consideremos lo serán sobre el conjunto[2] de los números naturales.

5.1 Funciones recursivas

Las funciones recursivas fueron introducidas por Gödel como un concepto técnico auxiliar que le permitía enunciar en la forma más adecuada algunos resultados previos a sus teoremas de incompletitud. No obstante, pronto se vio que este concepto tenía interés por sí mismo, debido a que las funciones recursivas resultan ser exactamente las funciones calculables mediante un algoritmo. Por "un algoritmo" entendemos cualquier secuencia finita de instrucciones que nos permita obtener tras un número finito de cálculos mecánicos el valor que toma la función sobre unos argumentos dados (el número de pasos necesarios puede depender de la magnitud de los argumentos). Equivalentemente, una función es calculable mediante un algoritmo si se puede diseñar un programa de ordenador que la calcule (supuesto que se implementa en un ordenador con memoria suficiente para realizar los cálculos, y teniendo en cuenta que la cantidad de memoria requerida dependerá de la magnitud de los datos). Antes de entrar en más detalles conviene que veamos la definición:

Funciones recursivas elementales Llamaremos *funciones recursivas elementales* a las siguientes funciones.

- La función monádica c, dada por $c(n) = 0$ para todo n.
- La función monádica s, dada por $s(n) = n + 1$ para todo n.
- Las funciones k-ádicas p_i^k para $1 \leq i \leq k$, dadas por $p_i^k(n_1, \ldots, n_k) = n_i$.

Observar que todas las funciones recursivas elementales se pueden calcular explícitamente en cada caso concreto. Las funciones recursivas son las que pueden obtenerse a partir de éstas mediante la aplicación de un número finito de los procesos de definición que indicamos seguidamente.

[2]Si fuéramos consecuentes deberíamos hablar de la *colección* de los números naturales pues —recordemos— estamos reservando la palabra "conjunto" para su uso técnico matemático y la palabra "colección" para la correspondiente noción metamatemática. No obstante, dado que aquí no puede haber confusión alguna, usaremos la palabra más habitual.

Definición de funciones a) Una función k-ádica f está definida por *composición* a partir de la función r-ádica g y de las funciones k-ádicas $h_1, \ldots, h_r$ syss para todos los naturales $a_1, \ldots, a_k$ se cumple que

$$f(a_1, \ldots, a_k) = g(h_1(a_1, \ldots, a_k), \ldots, h_r(a_1, \ldots, a_k)).$$

Claramente, si tenemos funciones g y $h_1, \ldots, h_r$, la ecuación anterior determina una función f sin ambigüedad alguna. Si disponemos de algoritmos para calcular las funciones g y h_i, es fácil diseñar a partir de ellos un algoritmo que calcule f: basta aplicar los algoritmos de las h_i para calcular las imágenes de los datos y aplicar el algoritmo de g a los resultados que obtengamos.

b) Una función $k+1$-ádica f está definida por *recursión* a partir de la función k-ádica g [o del natural a si $k = 0$] y de la función $k+2$-ádica h syss para todos los naturales $a_1, \ldots, a_k$, n se cumple que

$$\begin{aligned} f(0, a_1, \ldots, a_k) &= g(a_1, \ldots, a_k) \qquad [f(0) = a, \text{ si } k = 0] \\ f(n+1, a_1, \ldots, a_k) &= h(n, f(n, a_1, \ldots, a_k), a_1, \ldots, a_k). \end{aligned}$$

Si tenemos funciones g, h [o un número a y una función h], las ecuaciones anteriores determinan unívocamente una función f. Si disponemos de algoritmos para calcular g y h también tenemos otro para calcular f: calculamos primero $f(0, a_1, \ldots, a_k)$ con el algoritmo de g y después vamos calculando $f(1, a_1, \ldots, a_k)$, $f(2, a_1, \ldots, a_k)$, etc. mediante el algoritmo de h, hasta llegar a $f(n, a_1, \ldots, a_k)$.

c) Una función k-ádica f está definida por *minimización* a partir de una función $k+1$-ádica g si para todos los naturales $a_1, \ldots, a_k$ se cumple

1. $\bigvee n\ g(a_1, \ldots, a_k, n) = 0$,
2. $f(a_1, \ldots, a_k) = \mu n\ g(a_1, \ldots, a_k, n) = 0$.

Dada una función g que cumpla 1), la ecuación b) determina unívocamente una función f que será calculable mediante un algoritmo si lo es g: basta aplicar el algoritmo de g para calcular sucesivamente $g(a_1, \ldots, a_k, 0)$, $g(a_1, \ldots, a_k, 1)$, $g(a_1, \ldots, a_k, 2)$, etc. hasta encontrar el primer n que hace $g(a_1, \ldots, a_k, n) = 0$ (existe por la condición 1.)

Funciones recursivas Una función f es *recursiva primitiva (recursiva)* si existe una sucesión de funciones $f_1, \ldots, f_n$ tales que f_n es f y para todo natural i entre 1 y n, la función f_i es recursiva elemental o bien f_i está definida por composición o recursión (o minimización) a partir de funciones anteriores de la sucesión. Es claro que toda función recursiva primitiva es recursiva.

Así, la única diferencia entre las funciones recursivas y las recursivas primitivas consiste en que en las primeras se admite la minimización como técnica de definición y en las segundas no.

Puesto que las funciones elementales se pueden calcular mediante algoritmos (elementales) y las funciones definidas por composición, recursión o minimización a partir de funciones calculables mediante algoritmos son también calculables mediante algoritmos, es claro que toda función recursiva es calculable mediante un algoritmo. Más concretamente, si f es una función recursiva, una sucesión de funciones $f_1, \ldots, f_n$ según la definición determina un algoritmo para calcular f (en el sentido de que conociendo la sucesión es fácil diseñar el algoritmo correspondiente).

En realidad Gödel llamó funciones recursivas a lo que nosotros hemos llamado funciones recursivas primitivas. Como ya hemos comentado, su definición no tenía más pretensión que la de sistematizar algunos resultados previos a sus teoremas de incompletitud, y el nombre de función recursiva aludía simplemente a que el rasgo más característico de estas funciones era que permiten las definiciones recurrentes. Posteriormente Herbrand introdujo lo que llamó funciones *recursivas generales,* cuya definición permitía procedimientos de construcción más generales, tales como recursiones simultáneas en varias variables, definiciones implícitas por sistemas de ecuaciones que cumplieran ciertos requisitos, etc. Todo ello sin perder la propiedad de que las funciones así obtenidas eran calculables mediante algoritmos. La definición de Herbrand era tan amplia que resultaba natural conjeturar que cualquier función calculable mediante un algoritmo debía de ser recursiva general. Kleene demostró que las funciones recursivas generales de Herbrand coincidían con las que nosotros hemos definido como funciones recursivas, es decir, que toda la generalidad de la definición de Herbrand se obtenía igualmente sin más que añadir la minimización a la definición de Gödel. Desde entonces que los términos antiguos "función recursiva" y "función recursiva general" han sido sustituidos por "función recursiva primitiva" y "función recursiva", tal y como los hemos introducido nosotros. Finalmente, Turing demostró que las funciones recursivas son exactamente las calculables mediante un algoritmo. Esta afirmación se conoce como *Tesis de Church-Turing* y la probaremos al final de este capítulo, pero de momento hemos de empezar por estudiar las funciones recursivas.

Conviene resaltar la similitud formal entre la definición de función recursiva y la definición de teorema en una teoría axiomática: las funciones elementales son el equivalente a los axiomas y los métodos de construcción de funciones son el equivalente a las reglas de inferencia. Como en el caso del cálculo deductivo, la definición que hemos dado de función recursiva es arbitraria, pero cuando probemos la tesis de Church-Turing toda esta arbitrariedad desaparecerá y lo que tendremos será una caracterización extremadamente simple de un concepto aparentemente tan vasto como es el de función computable algorítmicamente. El análogo al teorema de corrección es el hecho que ya hemos constatado de que las funciones recursivas son calculables mediante algoritmos. Nos falta el análogo al teorema de adecuación, que es precisamente el resultado de Turing.

Como en el caso del cálculo deductivo, unas observaciones elementales simplifican notablemente la manipulación de funciones recursivas. En primer lugar, a la hora de mostrar que una función es recursiva primitiva (o recursiva) pode-

mos admitir que en la sucesión se incorporen funciones que ya hayamos probado que son recursivas primitivas (recursivas) como abreviatura de la sucesión completa que tendría, en lugar de dicha función, la sucesión que justifica su carácter recursivo.

Claramente, si f es una función n-ádica recursiva (primitiva) e $i_1, \ldots, i_n$ es una reordenación de $1, \ldots, n$, entonces la función g dada por $g(a_1, \ldots, a_n) = f(a_{i_1}, \ldots, a_{i_n})$ es recursiva (primitiva), pues

$$g(a_1, \ldots, a_n) = f(p^n_{i_1}(a_1, \ldots, a_n), \ldots, p^n_{i_n}(a_1, \ldots, a_n)),$$

es decir, g está definida por composición a partir de f y de las proyecciones.

Esto significa que no hemos de preocuparnos por el orden de los argumentos de las funciones. En particular si g y h son recursivas (primitivas) donde g es k-ádica y h es $k+2$-ádica, también será recursiva (primitiva) la función dada por

$$\begin{aligned} f(a_1, \ldots, a_k, 0) &= g(a_1, \ldots, a_k), \\ f(a_1, \ldots, a_k, n+1) &= h(a_1, \ldots, a_k, n, f(a_1, \ldots, a_k, n)). \end{aligned}$$

De hecho si hubiéramos definido la recursión con estas ecuaciones, la definición de función recursiva (primitiva) correspondiente sería equivalente. Por tanto en adelante usaremos la forma que consideremos más oportuna.

Veamos un primer ejemplo no trivial de función recursiva:

Teorema 5.1 *La suma de números naturales es una función diádica recursiva primitiva.*

DEMOSTRACIÓN:

$h_1(m) = m$		(p^1_1)
$h_2(m, n, p) = p$		(p^3_3)
$h_3(m) = m + 1$		(s)
$h_4(m, n, p) = h_3(h_2(m, n, p))$	$[= p + 1]$	(composición)
$h_5(m, 0) = h_1(m)$	$[= m]$	(recursión)
$h_5(m, n+1) = h_4(m, n, h_5(m, n))$	$[= h_5(m, n) + 1]$	

Claramente $h_5(m, n) = m + n$. ∎

En la práctica abreviaremos estas demostraciones expresando la función en términos de funciones ya probadas recursivas, sobrentendiendo las proyecciones donde proceda. Por ejemplo la prueba del teorema anterior se puede reducir a $m + 0 = m$; $m + (n + 1) = (m + n) + 1$.

Aquí tenemos algunos ejemplos adicionales de funciones recursivas primitivas. Las indicaciones que damos son suficientes para justificar su carácter recursivo.

1) $m \cdot n \qquad m \cdot 0 = 0 \qquad m \cdot (n+1) = m \cdot n + m.$

2) $c_a(n) = a \qquad c_a(0) = a \qquad c_a(n+1) = c_a(n).$

3) $m^n \qquad m^0 = 1 \qquad m^{n+1} = m^n \cdot m.$

4) $n! \qquad 0! = 1 \qquad (n+1)! = n! \cdot (n+1).$

5) $\mathrm{sg}(n) = \begin{cases} 0 & \text{si } n = 0 \\ 1 & \text{si } n \neq 0 \end{cases} \qquad \overline{\mathrm{sg}}(n) = \begin{cases} 1 & \text{si } n = 0 \\ 0 & \text{si } n \neq 0 \end{cases}$

$\mathrm{sg}(0) = 0 \qquad \mathrm{sg}(n+1) = c_1(n)$

$\overline{\mathrm{sg}}(0) = 1 \qquad \overline{\mathrm{sg}}(n+1) = c_0(n)$

6) $\mathrm{pre}(0) = 0 \qquad \mathrm{pre}(n+1) = n$

7) $m \dot{-} 0 = n \qquad m \dot{-} (n+1) = \mathrm{pre}(m \dot{-} n)$

5.2 Relaciones recursivas

Cuando hayamos justificado la tesis de Church-Turing será inmediato que una relación es recursiva —en el sentido que introducimos seguidamente— si y sólo si existe un algoritmo para determinar si se cumple o no sobre unos argumentos dados.

Definición 5.2 Si R es una relación n-ádica, llamaremos *función característica* de R a la función n-ádica dada por

$$\chi_R(a_1, \ldots, a_n) = \begin{cases} 0 & \text{si } R(a_1, \ldots, a_n) \\ 1 & \text{si no } R(a_1, \ldots, a_n) \end{cases}$$

Una relación es *recursiva (primitiva)* si lo es su función característica.

Veamos algunos teoremas elementales:

Teorema 5.3 *Si R y S son relaciones k-ádicas recursivas (primitivas), entonces $\neg R$, $R \vee S$, $R \wedge S$, $R \to S$ y $R \leftrightarrow S$ también lo son (donde $\neg R$ es la relación que se cumple cuando no se cumple R, etc.).*

Demostración: Claramente se cumple

a) $\chi_{\neg R}(a_1, \ldots, a_k) = \overline{\mathrm{sg}}(\chi_R(a_1, \ldots, a_k))$,

b) $\chi_{R \vee S}(a_1, \ldots, a_k) = \chi_R(a_1, \ldots, a_k) \cdot \chi_S(a_1, \ldots, a_k)$,

c) $\chi_{R \wedge S}(a_1, \ldots, a_k) = \mathrm{sg}(\chi_R(a_1, \ldots, a_k) + \chi_S(a_1, \ldots, a_k))$,

d) $R \to S$ es $\neg R \vee S$ y $R \leftrightarrow S$ es $(R \to S) \wedge (S \to R)$,

de donde se sigue fácilmente el teorema. ∎

Teorema 5.4 *Si f y g son funciones k-ádicas recursivas (primitivas), entonces las relaciones R, S, T dadas por*

$R(a_1,\ldots,a_k)$ *syss* $f(a_1,\ldots,a_k) = g(a_1,\ldots,a_k)$,

$S(a_1,\ldots,a_k)$ *syss* $f(a_1,\ldots,a_k) < g(a_1,\ldots,a_k)$,

$T(a_1,\ldots,a_k)$ *syss* $g(a_1,\ldots,a_k) \leq f(a_1,\ldots,a_k)$,

son recursivas (primitivas).

Demostración: Se cumple que

$$\chi_S(a_1,\ldots,a_k) = \overline{\mathrm{sg}}(g(a_1,\ldots,a_k) \dot{-} f(a_1,\ldots,a_k)),$$

luego S es recursiva (primitiva). La relación T es $\neg S$, luego por el teorema anterior es recursiva (primitiva). En particular también es recursiva (primitiva) la relación dada por $H(a_1,\ldots,a_k)$ syss $f(a_1,\ldots,a_k) \leq g(a_1,\ldots,a_k)$ y R es $T \wedge H$, con lo que también es recursiva (primitiva). ∎

Teorema 5.5 *Si R es una relación k-ádica recursiva (primitiva) y $f_1,\ldots,f_k$ son funciones n-ádicas recursivas (primitivas), entonces la relación dada por*

$$S(a_1,\ldots,a_k) \text{ syss } R(f_1(a_1,\ldots,a_k),\ldots,f_n(a_1,\ldots,a_k))$$

es recursiva (primitiva).

Demostración: Basta observar que

$$\chi_S(a_1,\ldots,a_k) = \chi_R(f_1(a_1,\ldots,a_k),\ldots,f_n(a_1,\ldots,a_k)).$$

∎

Teorema 5.6 *Si R es una relación $k+1$-ádica recursiva y para todos los naturales $a_1,\ldots,a_k$ existe un x tal que $R(a_1,\ldots,a_k,x)$, entonces la función dada por $f(a_1,\ldots,a_k) = \mu x\ R(a_1,\ldots,a_k,x)$ es recursiva.*

Demostración: Es claro, ya que $f(a_1,\ldots,a_k) = \mu x\ \chi_R(a_1,\ldots,a_k,x) = 0$.

∎

En general, aunque una relación $R(x,y)$ sea recursiva, la relación dada por $S(y) \leftrightarrow \bigvee x\, R(x,y)$ no tiene por qué ser recursiva, pues no está claro que exista un algoritmo que determine, para un valor de y, si existe o no un x que cumpla $R(x,y)$. Lo máximo que podríamos hacer a priori es ir calculando si se cumple o no $R(0,y)$, $R(1,y)$, $R(2,y)$, etc., pero si no existe un x que cumpla $R(x,y)$ estos cálculos no nos lo mostrarán. La situación es distinta si acotamos la variable x con una función recursiva f, es decir, si consideramos la relación $S(y) \leftrightarrow \bigvee x \leq f(y)\, R(x,y)$. Para comprobar si $S(y)$ es cierto o falso basta calcular $f(y)$ y después comprobar $R(x,y)$ para x entre 0 y $f(y)$, es decir, basta hacer un número finito de comprobaciones. Por consiguiente, siempre acabamos sabiendo si $S(y)$ se cumple o no. Así pues, S debe ser recursiva. La prueba no es trivial.

Teorema 5.7 *Si f es una función k-ádica recursiva (primitiva) y R es una relación $k+1$-ádica recursiva (primitiva), entonces las relaciones S y T y la función g dadas por*

$$S(a_1,\dots,a_k) \text{ syss } \bigvee x(x \le f(a_1,\dots,a_k) \wedge R(x,a_1,\dots,a_k)),$$

$$T(a_1,\dots,a_k) \text{ syss } \bigwedge x(x \le f(a_1,\dots,a_k) \to R(x,a_1,\dots,a_k)),$$

$$g(a_1,\dots,a_k) = \mu x(x \le f(a_1,\dots,a_k) \wedge R(x,a_1,\dots,a_k)),$$

son también recursivas (primitivas).

Demostración: Sea

$$\begin{aligned} p(a_1,\dots,a_k,0) &= \chi_R(0,a_1,\dots,a_k) \\ p(a_1,\dots,a_k,d+1) &= p(a_1,\dots,a_k,d)\cdot\chi_R(d+1,a_1,\dots,a_k), \end{aligned}$$

es decir,

$$p(a_1,\dots,a_k,n) = \prod_{i=0}^{n}\chi_R(i,a_1,\dots,a_k)$$

y por lo tanto $p(a_1,\dots,a_k,n) = 0$ syss $R(i,a_1,\dots,a_k)$ para algún $i \le n$, y en caso contrario se cumple $p(a_1,\dots,a_k,n) = 1$. Además p es recursiva (primitiva). Claramente χ_S es $p(a_1,\dots,a_k,f(a_1,\dots,a_k))$, con lo que S es recursiva (primitiva). Igualmente lo será la relación dada por

$$H(a_1,\dots,a_k) \text{ syss } \bigvee x(x \le f(a_1,\dots,a_k) \wedge \neg R(x,a_1,\dots,a_k))$$

y T es $\neg H$, luego también es recursiva (primitiva). Sea

$$\begin{aligned} m(a_1,\dots,a_k,0) &= 0 \\ m(a_1,\dots,a_k,d+1) &= (d+1)\cdot(p(a_1,\dots,a_k,d) \dot{-} p(a_1,\dots,a_k,d+1)) \\ &\quad + m(a_1,\dots,a_k,d)\cdot\overline{\mathrm{sg}}(p(a_1,\dots,a_k,d) \dot{-} p(a_1,\dots,a_k,d+1)) \end{aligned}$$

Si ningún natural x cumple $R(x,a_1,\dots,a_k)$, entonces $p(a_1,\dots,a_k,d) = 1$ para todo d, luego $p(a_1,\dots,a_k,d) \dot{-} p(a_1,\dots,a_k,d+1) = 0$ y así tenemos que $m(a_1,\dots,a_k,d+1) = m(a_1,\dots,a_k,d)$ para todo d y, como $m(a_1,\dots,a_k,0) = 0$, se cumple $m(a_1,\dots,a_k,d) = 0$ para todo d.

Si x es el menor natural tal que $R(x,a_1,\dots,a_k)$, entonces distinguimos varios casos:

a) Si $x = 0$, entonces se cumple $p(a_1,\dots,a_k,n) = 0$ para todo n, luego $p(a_1,\dots,a_k,d) \dot{-} p(a_1,\dots,a_k,d+1) = 0$ para todo d y, como antes, llegamos a que $m(a_1,\dots,a_k,d) = 0$ para todo d ($\ge x$).

b) Si $x \ne 0$, sea $x = y+1$.

b1) Si $d+1 \le y$, tenemos que $p(a_1,\dots,a_k,d) = p(a_1,\dots,a_k,d+1) = 1$, luego se cumple que $p(a_1,\dots,a_k,d) \dot{-} p(a_1,\dots,a_k,d+1) = 0$ y consecuentemente $m(a_1,\dots,a_k,d+1) = m(a_1,\dots,a_k,d)$, es decir, $m(a_1,\dots,a_k,d) = 0$ para todo $d < y$.

b2) Si $d = y$ tenemos que $p(a_1, \ldots, a_k, d) = 1$ y $p(a_1, \ldots, a_k, d+1) = 0$, luego $p(a_1, \ldots, a_k, d) \dot{-} p(a_1, \ldots, a_k, d+1) = 1$ y resulta $m(a_1, \ldots, a_k, x) = (d+1) \cdot 1 + 0 = x$.

b3) Si $d \geq x$ tenemos que $p(a_1, \ldots, a_k, d) = p(a_1, \ldots, a_k, d+1) = 0$, luego otra vez $p(a_1, \ldots, a_k, d) \dot{-} p(a_1, \ldots, a_k, d+1) = 0$ y queda $m(a_1, \ldots, a_k, d+1) = m(a_1, \ldots, a_k, d)$, es decir, $m(a_1, \ldots, a_k, d) = x$ para todo $d \geq x$.

En resumen, $m(a_1, \ldots, a_k, d) = 0$ hasta el momento en que un x cumple $R(x, a_1, \ldots, a_k)$, y a partir de dicho x vale constantemente x.

Así, $m(a_1, \ldots, a_k, f(a_1, \ldots, a_k))$ es el mínimo natural $x \leq f(a_1, \ldots, a_k)$ tal que $R(x, a_1, \ldots, a_k)$ si existe o bien 0 en otro caso, o sea,

$$m(a_1, \ldots, a_k, f(a_1, \ldots, a_k)) = g(a_1, \ldots, a_k)$$

que será, por tanto, recursiva (primitiva). ∎

Las siguientes relaciones y funciones son recursivas primitivas. La comprobación es una aplicación directa de los resultados anteriores.

1) $y \mid x$ syss $\bigvee z(z \leq x \wedge x = y \cdot z)$ (y divide a x.)

2) Prim x syss $\neg \bigvee z(z \leq x \wedge z \neq 1 \wedge z \neq x \wedge z \mid x) \wedge x \neq 0 \wedge x \neq 1$ (x es un número primo.)

3) $0 \text{ Pr } x = 0$; $\quad (n+1) \text{ Pr } x = \mu y(y \leq x \wedge \text{ Prim } y \wedge y \mid x \wedge y > n \text{ Pr } x)$ (n Pr x es el n-simo primo en orden creciente que divide a x y vale 0 si n es mayor que el número de primos que dividen a x, por ejemplo, $2 \text{ Pr } 60 = 3$.)

4) $\text{Pr } 0 = 0$; $\quad \text{Pr}(n+1) = \mu y(y \leq \text{ Pr}(n)! + 1 \wedge \text{ Prim } y \wedge y > \text{ Pr}(n))$ (Pr(n) es el n-simo primo en orden creciente.)

5) $nNx = \mu y(y \leq x \wedge (\text{Pr } n)^y \mid x \wedge \neg(\text{Pr } n)^{y+1} \mid x)$ (nNx es el exponente del primo $\text{Pr}(n)$ en la descomposición en factores primos de x, por ejemplo, $2N60 = 1$.)

6) Sea f una función monádica recursiva (primitiva)

$$\prod_{0 \leq i \leq 0} f(i) = 1; \qquad \prod_{0 \leq i \leq n+1} f(i) = \Big(\prod_{0 \leq i \leq n} f(i) \Big) \cdot f(n+1)$$

7) $\prod_{0 < i < n} f(i) = \prod_{0 < i \leq \text{pre}\, n} f(i)$.

5.3 Conjuntos recursivos

Definimos ahora los conjuntos recursivos. Cuando hayamos probado la tesis de Church-Turing será inmediato que un conjunto es recursivo si y sólo si existe un algoritmo que determina si un número natural dado pertenece o no al conjunto.

Definición 5.8 Sea A un conjunto de números naturales. Llamaremos *función característica* de A a la función monádica dada por

$$\chi_A(n) = \begin{cases} 0 & \text{si } n \text{ está en } A, \\ 1 & \text{si } n \text{ no está en } A. \end{cases}$$

Diremos que el conjunto A es *recursivo (primitivo)* si la función χ_A es recursiva (primitiva).

Notemos que en realidad un conjunto recursivo es lo mismo que una relación monádica recursiva.

Teorema 5.9 *Todo conjunto finito es recursivo primitivo.*

Demostración: Veamos primero que si k es un número natural, el conjunto $\{k\}$ es recursivo primitivo. Ello se debe a que la relación dada por $R(x)$ syss $x = k$ es recursiva primitiva por el teorema 5.4 (notar que $x = k$ puede escribirse como $p_1^1(x) = c_k(x)$), y claramente $\chi_{\{k\}} = \chi_R$.

Para un conjunto $A = \{k_1, \ldots, k_n\}$ tenemos que $\chi_A = \chi_{\{k_1\}} \cdots \chi_{\{k_n\}}$. ∎

Ejercicio: Demostrar que las uniones, las intersecciones y los complementos de conjuntos recursivos son también conjuntos recursivos.

Ejercicio: Un conjunto de números naturales es *recursivamente numerable* si es la imagen de una función monádica recursiva. Demostrar que todo conjunto recursivo es recursivamente numerable. [ayuda: observar cómo hemos enumerado recursivamente la sucesión de los números primos.] Demostrar que un conjunto es recursivo si y sólo si él y su complementario son ambos recursivamente numerables.

Ejercicio: Diremos que una función $f : \mathbb{N}^n \longrightarrow \mathbb{N}^m$ es *recursiva* si lo son sus funciones coordenadas. Construir una función $f : \mathbb{N} \longrightarrow \mathbb{N}^n$ recursiva y suprayectiva. [ayuda: Considerar los exponentes de los primos que dividen a un número natural dado].

Ejercicio: Probar que un conjunto A es recursivamente numerable si y sólo si existe una relación diádica recursiva R tal que $A = \{x | \bigvee y\, R(x, y)\}$.

5.4 Números de Gödel

Ahora veremos cómo aparecen las funciones y las relaciones recursivas en el estudio de las teorías axiomáticas. El lector que esté interesado en la caracterización de Turing de las funciones recursivas puede saltarse momentáneamente esta sección. Lo que sigue es necesario para demostrar los teoremas de incompletitud, pero no será usado en el resto del capítulo. Alternativamente, el lector que no esté interesado en las funciones recursivas en sí mismas puede continuar con esta sección y después pasar al capítulo siguiente.

Sea $\mathfrak{L}$ un lenguaje formal. Definimos como sigue el *número de Gödel* $g(\zeta)$ de cada signo ζ de $\mathfrak{L}$:

$$\begin{array}{rclrcll} g(\neg) & = & 3, & g(x_i) & = & 17+8i & i=0,1,2,\ldots \\ g(\to) & = & 5, & g(c_i) & = & 11+8i, & i=0,1,2,\ldots \\ g(\bigwedge) & = & 7, & g(R_i^n) & = & 5+8\cdot 2^n\cdot 3^i, & i=0,1,2,\ldots, \quad n=1,2,3,\ldots \\ g(|) & = & 9, & g(f_i^n) & = & 7+8\cdot 2^n\cdot 3^i, & i=0,1,2,\ldots, \quad n=1,2,3,\ldots \end{array}$$

Se entiende que el número de Gödel de una constante o un relator o un funtor está definido sólo si el signo correspondiente está definido en $\mathfrak{L}$.

Notemos que $g(x_i) \equiv 1 \pmod 8$, $g(c_i) \equiv 3 \pmod 8$, $g(R_i^n) \equiv 5 \pmod 8$, $g(f_i^n) \equiv 7 \pmod 8$ y todos estos números son mayores que 9, luego signos distintos tienen números de Gödel distintos. Además $g(\zeta)$ es siempre un número impar.

Si ζ es una cadena de signos de $\mathfrak{L}$ y $\zeta_1,\ldots,\zeta_n$ son los signos que la componen en el orden en que aparecen, definimos el *número de Gödel* de ζ como $g(\zeta) = p_1^{g(\zeta_1)}\cdots p_n^{g(\zeta_n)}$, donde $p_1, p_2, \ldots$ es la sucesión de los números primos $2, 3, 5, \ldots$

Es claro que cadenas distintas tienen números de Gödel distintos, todos ellos números pares con el exponente de 2 impar.

Finalmente, si $\zeta_1,\ldots,\zeta_n$ es una sucesión de cadenas de signos de $\mathfrak{L}$, definimos el *número de Gödel* de la sucesión como $g(\zeta_1,\ldots,\zeta_n) = p_1^{g(\zeta_1)}\cdots p_n^{g(\zeta_n)}$. También es claro que sucesiones distintas tienen números de Gödel distintos, todos ellos números pares con exponente de 2 par.

Observaciones Con la numeración de Gödel, cada signo, cada fórmula y cada demostración de una teoría axiomática tiene asociado un número natural que la identifica completamente. Esto es todo lo que necesitamos, pero a la hora de interpretar los resultados que vamos a obtener conviene tener presente la posibilidad de concebir la numeración de Gödel de otras maneras equivalentes.

Una de ellas, la sugerida por el propio Gödel, consiste en identificar cada concepto lógico con su número de Gödel. Ya comentamos en su momento que en realidad los signos de un lenguaje formal son como las piezas de ajedrez: son conceptos abstractos que podemos identificar, si así lo deseamos, con trozos de madera, plástico o marfil, pero lo cierto es que cuando se discute si en una apertura dada conviene o no mover el caballo del rey, importa poco si el caballo es de madera, de plástico o de marfil. Igualmente, no importa si el implicador de un lenguaje formal tiene esta o aquella forma. Siempre nos referiremos a él como "$\to$" sin preocuparnos de cuál es el signo concreto nombrado por este signo metamatemático. Por ello, nada nos impide considerar que los signos de un lenguaje formal son números naturales. Nada nos impide definir el negador como el número 3, la variable x_0 como el número 17, etc.

Similarmente, en lugar de llamar cadenas de signos a las sucesiones de signos (o de números), podemos definir las cadenas de signos como ciertos números

naturales, concretamente los que son divisibles por primos consecutivos desde el 2 hasta uno dado y cuyos exponentes son signos (= números de Gödel de signos). Así, nos da igual considerar que los signos metamatemáticos $x_0 = x_0$ nombran una sucesión de tres signos en el sentido usual o bien considerar que está nombrando al número natural $2^{37} \cdot 3^{17} \cdot 5^{17}$. Observemos que ni siquiera tenemos que cambiar de notación metamatemática: podemos seguir escribiendo $x_0 = x_0$, sólo que ahora eso es otra forma de escribir un número natural de 32 cifras decimales ¿qué más da que $x_0 = x_0$ sea un número o sea una sucesión de tres signos?

Lo mismo ocurre con las sucesiones de cadenas de signos, en particular con las demostraciones. Cuando escribimos una demostración formal, podemos convenir que lo que estamos haciendo es nombrar (con una notación bidimensional) un cierto número natural.

Desde este punto de vista, sería posible escribir algo del estilo de

Teorema 5.274.552.122.

DEMOSTRACIÓN: : 6.258.554.987.990.000.123.432. ∎

(Salvo por el hecho de que los números deberían ser astronómicos.)

Obviamente, con esto no ganamos nada práctico. No estamos sugiriendo que convenga concebir la matemática como meras manipulaciones de números. Lo que importa es que teóricamente sería posible hacerlo y al comprender que esto es así estamos dando un paso hacia la comprensión de los teoremas de incompletitud.

Así pues, de ahora en adelante suprimiremos a menudo la expresión "número de Gödel de", de tal modo que si decimos que un número natural n es una fórmula, habrá de entenderse que es el número de Gödel de una fórmula, si decimos que el conjunto de las fórmulas de un lenguaje formal es recursivo habremos de entender que el conjunto de los números de Gödel de las fórmulas del lenguaje es recursivo, etc. Podemos entender esto como un abuso de lenguaje, pero —según hemos comentado— también podríamos, si quisiéramos, entenderlo literalmente.

Otra concepción útil de la numeración de Gödel consiste en considerar que los números codifican afirmaciones y razonamientos. Dicho de una forma más sugerente, a través de la numeración de Gödel los números "hablan" de matemáticas. No todos, claro. De entre los números impares hay algunos que no dicen nada (no son números de Gödel) y otros que dicen únicamente "no", o "y", o "para todo", etc. De entre los números pares con exponente de 2 impar, los hay mudos y los hay que dicen "$2+2=4$", o "la suma de funciones continuas es continua", o "toda función continua es derivable", etc.

Finalmente, entre los números pares con exponente de 2 par nos encontramos algunos que demuestran que $2+2=4$, o que la suma de funciones continuas es continua y, probablemente,[3] no habrá ninguno que demuestre que toda función continua es derivable.

[3] ¡Más adelante veremos que no podemos asegurarlo!

A través de la numeración de Gödel tiene sentido afirmar que muchos conceptos lógicos son recursivos, y es fácil probarlo. Por ejemplo, las siguientes funciones y relaciones son recursivas primitivas:

1) $nNx = \mu y(y \leq x \wedge (\Pr\ n)^y \mid x \wedge \neg(\Pr\ n)^{y+1} \mid x)$

Esta relación la definimos ya en la página 127, pero ahora podemos interpretarla de otro modo: si x es (el número de Gödel de) una cadena de signos o de una sucesión de cadenas de signos de un lenguaje forma $\mathfrak{L}$, nNx es el (número de Gödel del) n-simo signo o la n-sima cadena de signos de x.

En lo sucesivo suprimiremos las expresiones "el número de Gödel de", que aquí hemos puesto entre paréntesis, de modo que al decir de un número natural x que es un signo o una cadena de signos hemos de entender que es su número de Gödel.

2) $L(x) = \mu y(y \leq x \wedge \bigwedge u(u \leq x \wedge y + 1 \leq u \to uNx = 0))$

Si x es una cadena o sucesión de cadenas de signos, $L(x)$ es su longitud.

3) $x * y = \mu z(z \leq \Pr(L(x) + L(y))^{x+y} \wedge \bigwedge n(n \leq L(x) \to nNz = nNx) \wedge \bigwedge n(0 < n \leq L(y) \to (n + L(x))Nz = nNy))$

Si x e y son cadenas o sucesiones de cadenas, $x * y$ es su yuxtaposición.

4) $x \equiv y$ (mód n) syss $\bigvee z(z \leq x + y \wedge (x = y + zn \vee y = x + zn))$

5) Var n syss $n \equiv 1$ (mód 8) $\wedge\ n \geq 9$

n es una variable.

6) Si $\mathfrak{L}$ tiene infinitas constantes definimos

$$\text{Const } n \text{ syss } n \equiv 3 \text{ (mód 8)} \wedge n > 3.$$

Si $\mathfrak{L}$ tiene un número finito de constantes y a es el número de Gödel de la mayor ($a = 0$ si no hay constantes), definimos

$$\text{Const } n \text{ syss } n \equiv 3 \text{ (mód 8)} \wedge n > 3 \wedge n \leq a.$$

De este modo Const n significa "n es una constante".

Notemos que la numeración de Gödel nos obliga a distinguir entre la variable x_0 (el número 17) y el término formado únicamente por la variable x_0 (que es el número 2^{17}).

Llamaremos $\text{Var}'x$ a la relación dada por $L(x) = 1 \wedge \text{Var}(1Nx)$, es decir, x es una cadena de signos que consta de una única variable. Similarmente se define $\text{Const}'x$.

Definición Un lenguaje formal $\mathfrak{L}$ es *recursivo* si los conjuntos de los (números de Gödel de los) relatores y funtores de $\mathfrak{L}$ son recursivos. Por ejemplo si, como suele ocurrir, hay un número finito de ellos.

Si $\mathfrak{L}$ es un lenguaje formal recursivo están bien definidas las relaciones Rel x y Fn x que se cumplen para los (números de Gödel de los) relatores y funtores de $\mathfrak{L}$. respectivamente, y son recursivas.

En lo que sigue suponemos que $\mathfrak{L}$ es un formalismo recursivo. Las relaciones y funciones que definimos son recursivas (recursivas primitivas si lo son las relaciones Rel y Fn).

7) n Rel x syss Rel $x \wedge \bigvee k(k \leq x \wedge x = 5 + 8 \cdot 2^n \cdot 3^k)$

x es un relator n-ádico.

8) n Fn x syss Fn $x \wedge \bigvee k(k \leq x \wedge x = 7 + 8 \cdot 2^n \cdot 3^k)$

x es un funtor n-ádico.

9) Neg $x = 2^3 * x$

Si x es α, Neg x es $\neg\alpha$.

10) $\mathrm{Imp}(x, y) = 2^5 * x * y$

Si x es α e y es β, $\mathrm{Imp}(x, y)$ es $\alpha \rightarrow \beta$.

11) $\mathrm{Dis}(x, y) = \mathrm{Imp}(\mathrm{Neg}\ x, y)$

Si x es α e y es β, $\mathrm{Dis}(x, y)$ es $\alpha \vee \beta$.

12) $\mathrm{Conj}(x, y) = \mathrm{Neg}(\mathrm{Dis}(\mathrm{Neg}\ x, \mathrm{Neg}\ y))$

Si x es α e y es β, $\mathrm{Conj}(x, y)$ es $\alpha \wedge \beta$.

13) Coimp $(x, y) = \mathrm{Conj}(\mathrm{Imp}(x, y), \mathrm{Imp}(y, x))$

Si x es α e y es β, $\mathrm{Coimp}(x, y)$ es $\alpha \leftrightarrow \beta$.

14) x Gen $y = 2^7 * 2^x * y$

Si x es x_i e y es α, x Gen y es $\bigwedge x_i \alpha$.

15) x Desc $y = 2^9 * 2^x * y$

Si x es x_i e y es α, x Desc y es $x_i|\alpha$.

Las relaciones siguientes están encaminadas a justificar la recursividad de la relación "ser una expresión", que es complicada porque la definición de expresión involucra una recursión muy compleja.

16) Tx syss $1Nx = 9 \vee \mathrm{Fn}(1Nx) \vee \mathrm{Const}(1Nx) \vee \mathrm{Var}(1Nx)$

Si x es una expresión, Tx syss x es un término.

17) Fx syss $1Nx = 3 \vee 1Nx = 5 \vee 1Nx = 7 \vee \mathrm{Rel}(1Nx)$

Si x es una expresión, Fx syss x es una fórmula.

18) $A(m, n, 0) = (1Nm)Nn$; $A(m, n, k+1) = A(m, n, k) * (((k+2)Nm)Nn)$

Si n es $\zeta_1, \ldots, \zeta_r$ y m es $a_0, \ldots, a_s$ e $i \leq s$ entonces $A(m, n, i)$ es $\zeta_{a_0} \cdots \zeta_{a_i}$.

19) $C(m, n) = A(m, n, L(m) \dot{-} 1)$

Si n es $\zeta_1, \ldots, \zeta_r$ y m es $a_0, \ldots, a_s$, entonces $C(m, n)$ es $\zeta_{a_0} \cdots \zeta_{a_s}$.

20) $\mathrm{Op}(x,y,z)$ syss $x = \mathrm{Neg}y \vee x = \mathrm{Imp}(y,z) \vee \bigvee v(v \leq x \wedge \mathrm{Var}\ v \wedge x = v\mathrm{Gen}y)$

Si x, y, z son fórmulas, x se obtiene de y, z por negación, implicación o generalización.

21) $\mathrm{SE}(n)$ syss $\bigwedge u(u \leq L(n) \wedge u > 0 \to \mathrm{Var}'(uNn) \vee \mathrm{Const}'(uNn) \vee \bigvee mp(m \leq n \wedge p \leq n \wedge (L(m)\mathrm{Rel}\ p \vee L(m)\mathrm{Fn}\ p) \wedge \bigwedge v(0 < v \leq L(m) \to 0 < vNm < u \wedge \mathrm{T}((vNm)Nn)) \wedge uNn = 2^p * C(m,n)) \vee \bigvee yz(y < u \wedge z < u \wedge 0 < y \wedge 0 < z \wedge \mathrm{F}(yNn) \wedge \mathrm{F}(zNn) \wedge \mathrm{Op}(uNn, yNn, zNn)) \vee \bigvee yv(y < n \wedge v < n \wedge 0 < y \wedge \mathrm{Var}\ v \wedge \mathrm{F}(yNn) \wedge uNn = v\mathrm{Desc}(yNn)))$

$\mathrm{SE}(n)$ syss n es una sucesión de expresiones tal que cada una se obtiene de las anteriores en un paso.

22) $\mathrm{Exp}\ x$ syss $\bigvee n(n \leq \mathrm{Pr}(L(x)^2)^{x \cdot L(x)^2} \wedge \mathrm{SE}(n) \wedge x = L(n)Nn)$

x es una expresión.

La acotación de n se obtiene como sigue: Para definir una expresión se necesitan tantos pasos como subexpresiones tenga. Claramente x tiene 1 subexpresión de longitud $L(x)$, a lo sumo 2 de longitud $L(x)-1, \ldots$, a lo sumo $L(x)$ de longitud 1, luego el número de subexpresiones es a lo sumo

$$1 + 2 + \cdots + L(x) = \frac{L(x)(L(x)-1)}{2} \leq L(x)^2.$$

Así pues, el mayor primo que divide a n es a lo sumo $\mathrm{Pr}(L(x)^2)$, cada potencia de primo que divide a n está acotada por $\mathrm{Pr}(L(x)^2)^x$ y el producto de estas potencias de primo está acotado por $\mathrm{Pr}(L(x)^2)^{x \cdot L(x)^2}$, tal y como afirmamos.

23) $\mathrm{Form}\ x$ syss $\mathrm{Exp}\ x \wedge \mathrm{F}x$.

x es una fórmula.

24) $\mathrm{Term}\ x$ syss $\mathrm{Exp}\ x \wedge \mathrm{T}x$.

x es un término.

25) $v\ \mathrm{Lig}\ n, x$ syss $\mathrm{Var}\ v \wedge \mathrm{Exp}\ x \wedge v = nNx \wedge \bigvee abc(a,b,c \leq x \wedge \mathrm{Form}\ b \wedge (x = a*(v\mathrm{Gen}\ b)*c \wedge L(a)+1 \leq n \leq L(a)+L(v\mathrm{Gen}\ b)) \vee (x = a*(v\mathrm{Desc}\ b)*c \wedge L(a)+1 \leq n \leq L(a)+L(v\mathrm{Desc}\ b)))$

La variable v está ligada en el lugar n-simo de la expresión x.

26) $v\ \mathrm{Lib}\ n, x$ syss $\mathrm{Var}\ v \wedge \mathrm{Exp}\ x \wedge v = nNx \wedge \neg v\mathrm{Lig}\ n, x$

La variable v está libre en el lugar n-simo de la expresión x.

27) $v\ \mathrm{Lib}\ x$ syss $\bigvee n(n \leq L(x) \wedge v\mathrm{Lib}\ n, x)$

La variable v está libre en la expresión x.

28) $v\ \mathrm{Lig}\ x$ syss $\bigvee n(n \leq L(x) \wedge v\mathrm{Lig}\ n, x)$

La variable v está ligada en la expresión x.

Ahora nos encaminamos a demostrar la recursividad de la función sustitución, que es más compleja aún que la de la relación "ser una expresión".

29) $\mathrm{VS}(n,t) = \mu y(y \leq n+t+8 \wedge \mathrm{Var}\ y \wedge \neg\bigvee r(r \leq n+t \wedge (y = rNn \vee y = rNt)))$

$\mathrm{VS}(n,t)$ es la menor variable que no está en n ni en t.

30) $B(n,v,0) = 0$; $B(n,v,r+1) = B(n,v,r) + \overline{\mathrm{sg}}\,\chi_R(v,r+1,n)$, donde R es la relación 26.

$B(n,v,r)$ es el número de veces que la variable v está libre en n hasta el lugar r.

31) $D(n,m,t,v,w,u)$ syss $\bigvee x(x \leq L(n) \wedge 0 < x \wedge w\mathrm{Lib}\ x,n \wedge u = B(n,v,x)\cdot(L(t) \dot{-} 1) + x)$

Si la expresión m resulta de sustituir en la expresión n cada variable v libre por el término t, entonces u es una posición de m donde hay una variable w que estaba libre en n.

32) $\mathrm{St}(n,m,t,v,w,k) = \mu p(p \leq m^k \wedge L(p) = L(m) \wedge \bigwedge u(u \leq L(m) \to (\neg D(n,m,t,v,w,u) \to uNp = uNm) \wedge (D(n,m,t,v,w,u) \to uNp = k)))$

Si la expresión m resulta de sustituir en la expresión n cada variable v libre por t, entonces St es la sustitución en m de cada variable w libre en n por la variable k.

33) $\mathrm{Sus}(n,r,v,t)$ syss $SE(n) \wedge SE(r) \wedge L(n) = L(r) \wedge \mathrm{Var}\ v \wedge \mathrm{Term}\ t \wedge \bigwedge u(u \leq L(n) \wedge 0 < u \to (uNn = 2^v \to uNr = t) \wedge (\mathrm{Var}'(uNn) \wedge uNn \neq 2^v \to uNr = uNn) \wedge (\mathrm{Const}'(uNn) \to uNr = uNn) \wedge \bigwedge mp(m \leq n \wedge p \leq n \wedge (L(m)\mathrm{Rel}\ p \vee L(m)\mathrm{Fn}\ p) \wedge \bigwedge w(0 < w \leq L(m) \to 0 < wNm < u \wedge T((wNm)Nn) \wedge uNn = 2^p * C(m,n)) \to uNr = 2^p * C(m,r)) \wedge \bigwedge y(0 < y < u \wedge F(yNn) \wedge uNn = \mathrm{Neg}\ (yNn) \to uNr = \mathrm{Neg}\ (yNr)) \wedge \bigwedge yz((0 < y < u \wedge 0 < z < u \wedge F(yNn) \wedge F(zNn) \wedge uNn = \mathrm{Imp}(yNn, zNn)) \to uNr = \mathrm{Imp}(yNr, zNr)) \wedge \bigwedge yw((0 < y < u \wedge w \leq n \wedge \mathrm{Var}\ w \wedge F(yNn) \wedge uNn = w\mathrm{Gen}(yNn)) \to ((\neg v\mathrm{Lib}(uNn) \to uNr = uNn) \wedge (v\mathrm{Lib}(uNn) \wedge \neg w\mathrm{Lib}\ t \to uNr = w\mathrm{Gen}(yNr)) \wedge (v\mathrm{Lib}(uNn) \wedge w\mathrm{Lib}\ t \to uNr = \mathrm{VS}(uNn,t)\mathrm{Gen}(\mathrm{St}(yNn, yNr, t, v, w, \mathrm{VS}(uNn,t)))))) \wedge \bigwedge yw((0 < y < u \wedge w \leq n \wedge \mathrm{Var}\ w \wedge F(yNn) \wedge uNn = w\mathrm{Desc}(yNn)) \to ((\neg v\mathrm{Lib}(uNn) \to uNr = uNn) \wedge (v\mathrm{Lib}(uNn) \wedge \neg w\mathrm{Lib}\ t \to uNr = w\mathrm{Desc}(yNr)) \wedge (v\mathrm{Lib}(uNn) \wedge w\mathrm{Lib}\ t \to uNr = \mathrm{VS}(uNn,t)\mathrm{Desc}(\mathrm{St}(yNn, yNr, t, v, w, \mathrm{VS}(uNn,t))))))$

n y r son sucesiones de expresiones de la misma longitud y cada expresión de r resulta de sustituir v por t en la correspondiente expresión de n.

34) $\mathrm{Sust}(n,v,t) = \mu r(\neg(\mathrm{SE}(n) \wedge \mathrm{Var}\ v \wedge \mathrm{Term}\ t) \vee \mathrm{Sus}(n,r,v,t))$

Si n es una sucesión de expresiones en el sentido de la relación 21, entonces Sust es la sucesión de expresiones que resulta de sustituir v por t en cada expresión de n. (Notar que en principio hemos probado que Sust es una función recursiva, pero calculando una cota explícita para r podemos probar que es recursiva primitiva.)

35) $S(n) = \mu m(m \leq \Pr(L(n)^2)^{nL(n)^2} \wedge \mathrm{SE}(m) \wedge n = L(m)Nm)$

$S(n)$ es una sucesión de expresiones que termina en n. (Para la cota ver la relación 22.)

36) $\mathtt{S}_v^t n = L(\mathrm{Sust}(S(n), v, t))N\mathrm{Sust}(S(n), v, t)$

Si n es una expresión, v es una variable y t es un término, entonces $\mathtt{S}_v^t n$ es la sustitución definida en el capítulo I.

Ahora es fácil probar la recursividad del cálculo deductivo:

37) $\mathrm{Ax1}(n)$ syss $\bigvee yz(y \leq n \wedge z \leq n \wedge \mathrm{Form}\ y \wedge \mathrm{Form}\ z \wedge n = \mathrm{Imp}(y, \mathrm{Imp}(z, y))$

n es un caso particular de K-1.

38) $\mathrm{Ax2}(n)$ syss $\bigvee xyz(y \leq n \wedge y \leq n \wedge z \leq n \wedge \mathrm{Form}\ x \wedge \mathrm{Form}\ y \wedge \mathrm{Form}\ z \wedge n = \mathrm{Imp}((\mathrm{Imp}(x, \mathrm{Imp}(y, z)), \mathrm{Imp}(\mathrm{Imp}(x, y), \mathrm{Imp}(x, z))))$

n es un caso particular de K-2.

39) $\mathrm{Ax3}(n)$ syss $\bigvee yz(y \leq n \wedge z \leq n \wedge \mathrm{Form}\ y \wedge \mathrm{Form}\ z \wedge n = \mathrm{Imp}(\mathrm{Imp}(\mathrm{Neg}\ y, \mathrm{Neg}\ z), \mathrm{Imp}(z, y)))$

n es un caso particular de K-3.

40) $\mathrm{Ax4}(n)$ syss $\bigvee tvm(t \leq n \wedge v \leq n \wedge m \leq n \wedge \mathrm{Term}\ t \wedge \mathrm{Var}\ v \wedge \mathrm{Form}\ m \wedge n = \mathrm{Imp}(v\mathrm{Gen}\ m, \mathtt{S}_v^t n))$

n es un caso particular de K-4.

41) $\mathrm{Ax5}(n)$ syss $\bigvee yzv(y \leq n \wedge z \leq n \wedge v \leq n \wedge \mathrm{Form}\ y \wedge \mathrm{Form}\ z \wedge \mathrm{Var}\ v \wedge \neg v\mathrm{Lib}\ y \wedge n = \mathrm{Imp}(v\mathrm{Gen}(\mathrm{Imp}(y, z)), \mathrm{Imp}(y, v\mathrm{Gen}\ z)))$

n es un caso particular de K-5.

42) $\mathrm{Ax6}(n)$ syss $\bigvee ytv(y \leq n \wedge t \leq n \wedge v \leq n \wedge \mathrm{Form}\ y \wedge \mathrm{Term}\ t \wedge \mathrm{Var}\ v \wedge \neg v\mathrm{Lib}\ t \wedge n = \mathrm{Coimp}(v\mathrm{Gen}(\mathrm{Imp}(2^{37} * v * t, y), \mathtt{S}_v^t y))$

n es un caso particular de K-6 (notar que $g(=) = 37$)

44) $v\ \mathrm{Part}\ n = \mathrm{Neg}(v\mathrm{Gen}(\mathrm{Neg}\ n))$

Si n es α y v es x, entonces $v\ \mathrm{Part}\ n$ es $\bigvee x\alpha$.

45) $v\ \mathrm{Part1}\ n = \mathrm{VS}(2^v, n)\mathrm{Part}(v\mathrm{Gen}(\mathrm{Coimp}(n, 2^{37} * v * \mathrm{VS}(2^v, n))))$

Si n es α y v es x, entonces $v\ \mathrm{Part1}\ n$ es $\overset{1}{\bigvee} x\alpha$.

46) $\mathrm{Ax7}(n)$ syss $\bigvee vy(v \leq n \wedge y \leq n \wedge \mathrm{Var}\ v \wedge \mathrm{Form}\ y \wedge n = \mathrm{Imp}(v\mathrm{Part1}\ y, \mathtt{S}_v^{v\mathrm{Desc}\ y} y))$.

n es un caso particular de K-7.

47) $\mathrm{Ax8}(n)$ syss $\bigvee uvy(u \leq n \wedge v \leq n \wedge y \leq n \wedge \mathrm{Var}\ u \wedge \mathrm{Var}\ v \wedge \mathrm{Form}\ y \wedge n = \mathrm{Imp}(\mathrm{Neg}(v\mathrm{Part}\ 1\ y), 2^{37} * (v\mathrm{Desc}\ y) * (u\mathrm{Desc}(2^{37} * u * u))))$

n es un caso particular de K-8.

45) Axl(n) syss Ax1(n) $\vee$Ax2(n) $\vee$Ax3(n) $\vee$Ax4(n) $\vee$Ax5(n) $\vee$ Ax6(n) $\vee$Ax7(n) $\vee$Ax8(n)

n es un axioma lógico.

Definición Una teoría axiomática es *recursiva* si su lenguaje formal es recursivo y el conjunto de (los números de Gödel de) sus axiomas es recursivo.

Si T es una teoría axiomática recursiva, la relación Ax n que se cumple si n es un axioma de T es recursiva. En lo que sigue supondremos que T es una teoría axiomática recursiva. Las relaciones siguientes son recursivas (recursivas primitivas si T lo es).

46) CI(x, y, z) syss Form $x \wedge$ Form $y \wedge$ Form $z \wedge (y = \mathrm{Imp}(z, x)$ $\vee \bigvee v(v \leq x \wedge \mathrm{Var}\ v \wedge x = v\mathrm{Gen}y))$

x es consecuencia inmediata de y, z.

47) Dm(n, x) syss $\bigwedge u(0 < u < L(n) \rightarrow (\mathrm{Ax}(uNm) \vee \mathrm{Axl}(uNm)$ $\vee \bigvee yz(0 < y < u \wedge 0 < z < u \wedge \mathrm{CI}(uNn, yNn, zNn)))) \wedge x = L(n)Nn.$

n es una demostración de x.

48) Rf(n, x) syss Dm($n, \mathrm{Neg}\ x$)

n es una refutación de x.

Observaciones En la prueba de los teoremas de incompletitud vamos a necesitar, concretamente, la recursividad de la función $\mathtt{S}_v^t n$ (la número 36) y de las relaciones Dm y Rf.

Conviene recordar, pues, que $\mathtt{S}_v^t n$ es una cierta función que a cada terna (v, t, n) de números naturales le asigna otro número natural que, en general, no tiene ninguna interpretación destacable pero, cuando v es (el número de Gödel de) una variable, t es (el número de Gödel de) un término y n es (el número de Gödel de) una expresión, entonces $\mathtt{S}_v^t n$ es (el número de Gödel de) la sustitución de la variable por el término en la expresión. Equivalentemente:

$$g(\mathtt{S}_x^t \theta) = \mathtt{S}_{g(x)}^{g(t)} g(\theta),$$

donde la sustitución de la izquierda es la definida en el capítulo I y la de la derecha es la definida aquí. De acuerdo con la tesis de Church-Turing, que la sustitución sea recursiva significa simplemente que disponemos de un algoritmo para calcularla. Así mismo, la recursividad de Dm significa que en una teoría axiomática recursiva (es decir, si sabemos distinguir qué es un axioma y qué no lo es) sabemos distinguir mediante un algoritmo si una sucesión de fórmulas dada demuestra o no una fórmula dada.

5.5 Funciones parciales

Dedicamos el resto del capítulo a demostrar la tesis de Church-Turing. Para ello necesitamos introducir una clase más general de funciones, y en primer lugar vamos a ver por qué hacen falta.

Digamos que una función recursiva primitiva es de rango n si puede definirse en n pasos sin usar proyecciones p_i^k con $k > n$. Es claro que sólo hay un número finito de tales funciones, pues sólo hay una cantidad finita de funciones elementales de las que podamos partir y sólo hay un número finito de posibilidades de combinar los dos métodos de construcción de funciones en n pasos. No sería difícil fijar un criterio explícito para ordenar las funciones de rango n. Si enumeramos todas las funciones de rango 1 (sólo hay tres) y a continuación todas las de rango dos, etc., obtenemos una enumeración (con posibles repeticiones) de todas las funciones recursivas primitivas. De ella podemos extraer una enumeración de las funciones monádicas recursivas primitivas: f_0, f_1, f_2, ...

Observemos que disponemos de un algoritmo para calcular $f_m(n)$ para cualquier par de números m y n. Sólo tenemos que enumerar todas las funciones de rango 1 y contar todas las que sean monádicas, luego todas las de rango 2, etc. hasta que la suma total de funciones monádicas iguale o exceda a m. Entonces, si esto ha sucedido al considerar las funciones de rango k, las ordenamos según el criterio explícito prefijado para las definiciones y nos quedamos con la definición de la función monádica que hace el número de orden m. Ésta es f_m y, como tenemos su definición, a partir de ella podemos calcular $f_m(n)$.

De este modo, tenemos definida una función diádica $f(m,n) = f_m(n)$ que sabemos calcular mediante un algoritmo. Aceptando la tesis de Church-Turing, la función f ha de ser recursiva. De hecho, es posible precisar todos los cálculos que hemos descrito para calcular f explícitamente y comprobar que satisface la definición de función recursiva. Se trata de una función universal para las funciones recursivas primitivas, pues un algoritmo que calcula f nos permite calcular cualquier función monádica recursiva primitiva.

Ahora bien, es claro entonces que la función $g(n) = f(n,n) + 1$ también ha de ser recursiva (recursiva primitiva si lo es f), pero por otra parte no puede ser recursiva primitiva, ya que si lo fuera habría de existir un m tal que $g(n) = f_m(n)$ para todo n, de donde $f(m,m) + 1 = g(m) = f_m(m) = f(m,m)$, lo cual es absurdo.

Tenemos así un ejemplo de una función recursiva que no es recursiva primitiva. Ahora nos encontramos con una paradoja: ¿qué ocurre si en lugar de partir de las funciones recursivas primitivas partimos de las funciones recursivas?, es decir, enumeramos las funciones monádicas recursivas g_0, g_1, g_2, ... y consideramos la función universal $g(m,n) = g_m(n)$. Por el mismo argumento anterior g no puede ser recursiva, pero, ¿qué nos impide calcularla igual que f?, ¿estamos ante un contraejemplo a la tesis de Church-Turing?

Ciertamente no. La falacia del argumento anterior está en que no disponemos de ningún algoritmo para enumerar explícitamente las sucesiones que definen funciones recursivas y el problema está, naturalmente, en las definicio-

nes por minimización. Una definición por minimización $h(m) = \mu n\, r(m,n) = 0$ sólo es válida si para todo m existe un n que cumpla $r(m,n) = 0$ y, aunque r sea recursiva, es decir, aunque sepamos calcular $r(m,n)$ para todo par de números, no está claro que sepamos decidir si se cumple o no esta condición. Así pues, aunque podamos enumerar todas las "posibles" definiciones de funciones recursivas de rango k, no está claro que sepamos distinguir cuáles de ellas son válidas. Pese a ello, la función g está bien definida, sólo que no es calculable mediante un algoritmo (o, al menos, no está claro que lo sea) y, desde luego, no es recursiva.

El hecho de no saber si una definición es correcta o no es un inconveniente técnico que necesitamos resolver, para lo cual introducimos las funciones parciales.

Definición 5.10 Una *función parcial* n-ádica f es un criterio bien definido que a ciertos grupos de n naturales $a_1, \ldots, a_n$ repetidos o no y en un cierto orden les asigna otro número natural que representaremos por $f(a_1, \ldots, a_n)$. Diremos en este caso que f *está definida* para $a_1, \ldots, a_n$ o que $f(a_1, \ldots, a_n)$ *está definido.*

Una función parcial k-ádica está definida por *composición parcial* a partir de las funciones parciales g (r-ádica) y $h_1, \ldots, h_r$ (k-ádicas) si f está definida exactamente para aquellos naturales $a_1, \ldots, a_k$ tales que están definidas $h_i(a_1, \ldots, a_k)$ $(i = 1, \ldots, r)$ y $g(h_1(a_1, \ldots, a_k), \ldots, h_r(a_1, \ldots, a_k))$ y se cumple

$$f(a_1, \ldots, a_k) = g(h_1(a_1, \ldots, a_k), \ldots, h_r(a_1, \ldots, a_k)).$$

Una función parcial $k+1$-ádica está definida por *recursión parcial* a partir de la función parcial k-ádica g (o del número natural a si $k = 0$) y la función parcial $k+2$-ádica h si f está definida exactamente para aquellos naturales $a_1, \ldots, a_k$, n tales que

a) $g(a_1, \ldots, a_k)$ está definido [si $k \neq 0$],

b) $f(u, a_1, \ldots, a_k)$ está definido para todo $u < n$,

c) $h(u, f(u, a_1, \ldots, a_k), a_1, \ldots, a_k)$ está definido para todo $u < n$,

y se cumple

$$\begin{aligned} f(0, a_1, \ldots, a_k) &= g(a_1, \ldots, a_k) \\ f(u+1, a_1, \ldots, a_k) &= h(u, f(u, a_1, \ldots, a_k), a_1, \ldots, a_k) \quad \text{si } 0 \leq u < n \end{aligned}$$

Una función parcial n-ádica f está definida por *minimización parcial* a partir de una función parcial $n+1$-ádica g si f está definida exactamente para aquellos naturales $a_1, \ldots, a_n$ tales que existe un natural m que cumple

a) Si $k \leq m$ entonces g está definida para $a_1, \ldots, a_n$, k.

b) $g(a_1, \ldots, a_n, m) = 0$ y se cumple

$$f(a_1, \ldots, a_n) = \mu\, m g(a_1, \ldots, a_n, m) = 0.$$

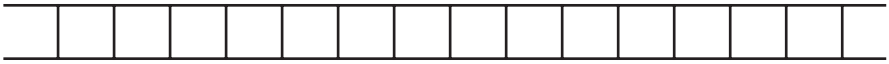
Una función parcial f es *recursiva parcial* si hay una sucesión de funciones $f_1, \ldots, f_n$ tales que f_n es f y cada f_i es recursiva elemental o está definida por composición, recursión o minimización parcial a partir de funciones anteriores de la sucesión.

Obviamente, toda función recursiva es recursiva parcial. Ahora sí podemos enumerar explícitamente todas las funciones recursivas parciales, pero ya no llegamos a ninguna contradicción. Simplemente, la función $g(n,n)+1$ puede ser igual a una función g_n de modo que $g_n(n)$ no esté definido.

5.6 Máquinas de Turing

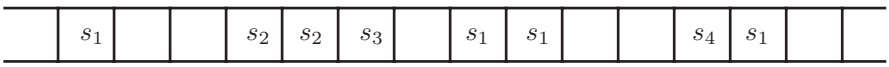
El argumento de Turing para probar que las funciones recursivas coinciden con las funciones calculables mediante un algoritmo se basa en el concepto de máquina de Turing. Una máquina de Turing es un modelo teórico de ordenador con infinita memoria disponible. Aunque la descripción que sigue parezca la descripción de una máquina real, debemos tener presente que una máquina de Turing es un concepto abstracto de la misma naturaleza que una teoría axiomática, es decir, un sistema conceptual que fundamenta una serie de afirmaciones objetivas.

Una *máquina de Turing* consta de una cinta infinita dividida en infinitas casillas contiguas infinitamente prolongable tanto a izquierda como a derecha:

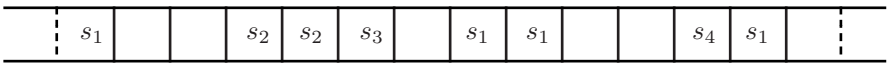


Cada casilla puede estar en blanco o tener impreso un signo de entre los de una lista finita que llamaremos *alfabeto:* $s_1, \ldots, s_j$ $(j \geq 1)$ fija para cada máquina particular. Escribiremos s_0 para nombrar al "blanco" y así la situación posible de una casilla será una de entre $s_0, \ldots, s_j$. En cualquier momento la cinta tendrá un número finito de casillas impresas (con signos distintos de s_0).

Representaremos la cinta con signos así:



Se sobrentiende que el resto de la cinta está en blanco. Si queremos indicar que el resto de la cinta puede no estar en blanco pero que no importa lo que haya impreso, lo representaremos así:



El estado de la cinta en un momento dado lo llamaremos *situación.*

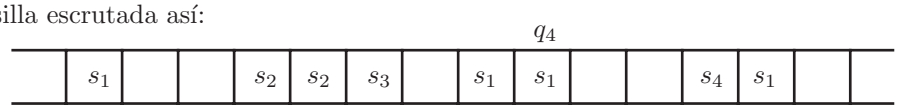
En cada instante la máquina se encontrará en un *estado* de entre un número finito posible de ellos $q_0, \ldots, q_k$ $(k \geq 1)$ fijo para cada máquina particular. Cada estado puede ser *activo* o *pasivo.* El estado q_0 siempre es pasivo, q_1 siempre es

activo. A q_1 le llamaremos *estado inicial,* es el estado en que se encuentra la máquina cuando empieza a funcionar.

En cada instante la máquina lee el signo de una casilla de la cinta. Esta casilla se llama *casilla escrutada* y el signo se llama *signo escrutado.*

El estado de la máquina y el signo escrutado determinan la *configuración* de la máquina en un instante dado. La configuración y la situación determinan la *configuración completa.*

Expresaremos las configuraciones completas indicando el estado sobre la casilla escrutada así:



Si no queremos indicar el estado usaremos un guión "—".

Si en un instante dado una máquina de Turing se encuentra en un estado activo, ésta realizará un *acto.* Un acto consiste en:

a) Leer el signo de la casilla escrutada,

b) Imprimir un signo (quizá s_0) en la casilla escrutada,

c) Mover un lugar la cinta de modo que la nueva casilla escrutada pase a ser la contigua izquierda, la misma casilla o la contigua derecha,

d) Cambiar de estado (pasando quizá al mismo),

de tal modo que el signo que se imprime, el movimiento que se hace y el estado al que se pasa, son función exclusivamente de la configuración de la máquina en ese instante.

Si el estado es pasivo no se produce ningún acto: la máquina está *parada.*

Según esto una máquina de Turing viene determinada por:

a) El alfabeto $s_0, \ldots, s_j$, con $j \geq 1$,

b) El conjunto de estados posibles $q_0, \ldots, q_k$, con $k \geq 1$,

c) Una función que a cada configuración activa (s_a, q_b) le asigna una terna (s_c, M, q_d), donde s_c, M, q_d son, respectivamente el signo impreso, el movimiento realizado I, D o C (izquierda, derecha o centro) y el estado al que se pasa, todo esto cuando la configuración es (s_a, q_b). A esta función se le llama *programa* de la máquina.

En la práctica escribiremos el programa en forma de tabla. Por ejemplo: Sea A la máquina de Turing con alfabeto s_0, s_1, estados q_0, q_1 y programa

A	s_0	s_1
q_1	$s_1 C q_0$	$s_1 D q_1$

La máquina A se mueve sobre la cinta hacia la derecha hasta encontrar una casilla en blanco, donde imprime s_1 y se para.

Según advertíamos al principio, las máquinas de Turing no existen (como objetos físicos). No son ordenadores porque ningún ordenador puede trabajar con una "cinta" de memoria infinita. Son un modelo de ordenador ideal exento de limitaciones de memoria. Lo único importante es que podemos hablar consistentemente de ellas y determinar qué hace una máquina dada a partir de unos datos dados, como acabamos de hacer con la máquina A.

Computabilidad Consideremos una máquina de Turing y sea $s = s_1$. Llamaremos *representación* del número natural n a la situación de la cinta que consta de $n+1$ signos s consecutivos, con el anterior y posterior en blanco.

Llamaremos *representación* de los números $a_1, \ldots, a_n$ a la situación que consta de n secuencias de $a_i + 1$ signos s consecutivos cada una, separadas por un blanco. Por ejemplo, la representación de 2, 0, 3 es

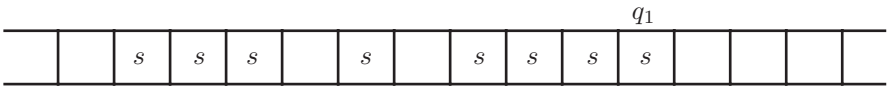
Llamaremos *vacío* en la cinta a dos o más casillas en blanco consecutivas. Llamaremos *representación normal* o *posición normal* de los naturales $a_1, \ldots, a_n$ a la representación de $a_1, \ldots, a_n$ cuando la casilla escrutada es la última casilla impresa de a_n. Por ejemplo 2, 0, 3 en posición normal es

Diremos que una máquina de Turing M *computa* la función parcial n-ádica f si cuando M comienza con los números $a_1, \ldots, a_n$ en posición normal y el resto de la cinta en blanco, termina con la representación normal de

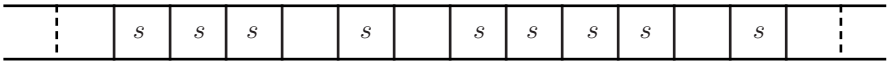
$$a_1, \ldots, a_n, f(a_1, \ldots, a_n)$$

en el caso de que $f(a_1, \ldots, a_n)$ esté definido y no se detiene o no se detiene con $a_1, \ldots, a_n, a$ en posición normal para ningún número a si $f(a_1, \ldots, a_n)$ no está definido.

Por ejemplo si $f(2, 0, 3) = 0$ y M computa f, cuando M comienza con



termina con



No se exige que la posición absoluta de los números en la cinta sea la misma que al comienzo.

Una función parcial es *computable* si hay una máquina de Turing que la computa. Una máquina de Turing M computa $1 \mid 1$ la función parcial n-ádica f si cumple:

a) El alfabeto de M es s_0, s,

b) Si M comienza con $a_1, \ldots, a_n$ en posición normal y el resto de la cinta a la derecha en blanco se cumple:

   1. Las casillas a la izquierda de la representación de $a_1, \ldots, a_n$ (o sea, a la izquierda del blanco anterior a a_1) no son nunca escrutadas.
   2. Si $f(a_1, \ldots, a_n)$ está definido, entonces M acaba con

      $$a_1, \ldots, a_n, f(a_1, \ldots, a_n)$$

      en posición normal de modo que la representación comienza en la misma casilla donde comenzaba la de $a_1, \ldots, a_n$ al principio. Además todas las casillas a la derecha quedan en blanco.
   3. Si $f(a_1, \ldots, a_n)$ no está definido entonces M no se para.

Una función parcial es $1 \mid 1$ *computable* si hay una máquina de Turing que la computa $1 \mid 1$. Vamos a demostrar que una función es computable si y sólo si es $1 \mid 1$ computable si y sólo si es recursiva parcial. El concepto de computabilidad $1 \mid 1$ es un concepto auxiliar técnico para la prueba.

Por el momento trabajaremos con máquinas de un solo signo. Para ellas usaremos la siguiente notación más cómoda:

a) Llamaremos 0 a s_0 y 1 a s_1.

b) Imprimir 1 sobre un 0 lo representaremos E (escribir).

c) Imprimir 0 sobre un 1 lo representaremos B (borrar).

d) Si un signo no se modifica no indicaremos nada.

e) Los estados pasivos serán $0_1, \ldots, 0_n$ (o "0" si sólo hay uno).

f) Los estados activos serán 1, 2, 3, ... (1 es el estado inicial).

Por ejemplo la máquina A de antes se representa ahora así:

A	0	1
1	$E0$	$D1$

Concatenación de máquinas de Turing Si M es una máquina de Turing con estados pasivos $0_1, \ldots, 0_n$ y $N_1, \ldots, N_n$ son otras máquinas de Turing, llamaremos

$$M \left[\begin{array}{c} N_1 \\ \vdots \\ N_n \end{array} \right.$$

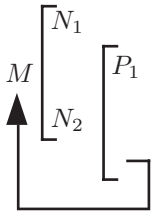
a la máquina de Turing definida como sigue:

Si $q_1, \ldots, q_j$ son los estados activos de M y $q_1^i, \ldots q_{j_i}^i$ son los estados activos de N_i, los estados activos de la nueva máquina son $q_1, \ldots, q_j, q_1^i, \ldots q_{j_i}^i$, para $i = 1, \ldots, j$.

Los estados pasivos son los de las máquinas $N_1, \ldots, N_n$. El estado inicial es q_1, es decir, el estado inicial de M. El programa es como sigue:

Dada una configuración, se realiza el acto marcado por el programa de la máquina a la que pertenece el estado en curso, a excepción del caso en que M deba pasar al estado 0_i, en cuyo caso se pasa al estado q_1^i.

En otras palabras, se trata de la máquina que empieza actuando como M y, cuando se ésta se ha de parar por pasar al estado 0_i, en lugar de ello comienza a actuar la máquina N_i.



La concatenación puede repetirse cuantas veces se quiera, incluso de forma circular. Por ejemplo, la máquina de la figura empieza actuando como M, cuando ésta acaba empieza N_1 o N_2, según el estado pasivo de M al que se llegue; si empieza N_2, cuando ésta acaba empieza P_1 o vuelve a empezar M según el estado pasivo final.

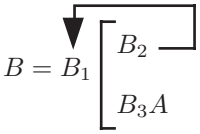
Si M es una máquina de Turing con un único estado pasivo, llamaremos M^n a la máquina que resulta de concatenar M consigo misma n veces.

Construcción de máquinas de Turing Construimos ahora algunas máquinas de Turing concretas e indicamos la actividad que realizan (bajo determinadas condiciones iniciales). Un guión en la tabla del programa indica que no importa la instrucción que pongamos en esa casilla, pues no afecta al comportamiento que se requiere de la máquina.

B_1	0	1
1	$-$	$BI2$
2	$EI3$	$I2$
3	$D0_1$	$D0_2$

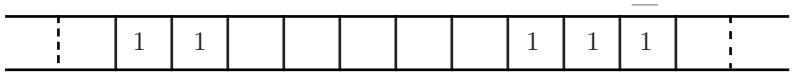
B_2	0	1
1	$I0$	$D1$

B_3	0	1
1	$-$	$BD0$

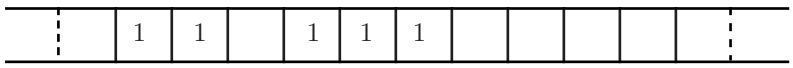


Si B comienza con un número en posición normal y otro a su izquierda, mueve el primero hasta eliminar el vacío que los separa (si hay tal vacío) sin escrutar las casillas a la izquierda del segundo número.

Por ejemplo, partiendo de



B termina así:



sin escrutar ninguna casilla no representada.

C	0	1
1	$-$	$D2$
2	$D3$	$-$
3	$E0$	$-$

D	0	1
1	$I2$	$I1$
2	$I2$	0

E	0	1
1	$-$	$I2$
2	$D0_1$	$D0_2$

F	0	1
1	$-$	$BI0$

G	0	1
1	$-$	$D2$
2	$D2$	$D3$
3	$I0$	$D3$

H	0	1
1	$-$	$D2$
2	$ED2$	$I3$
3	$-$	$BI0$

El comportamiento de estas máquinas es el siguiente:

C Cuando empieza con un número en posición normal va dos lugares a la derecha e imprime.

D Cuando empieza con un número en posición normal que no sea el extremo izquierdo de la cinta, se sitúa en posición normal respecto al número siguiente por la izquierda.

E Cuando empieza con un número en posición normal toma la salida 0_1 o 0_2 según sea 0 o distinto de 0 y termina en posición normal.

F Cuando comienza en una casilla impresa, borra y va una casilla a la izquierda.

G Va un número a la derecha (al revés que D).

H Cuando comienza con un número en posición normal que no sea el extremo derecho de la cinta, lo completa con unos hasta eliminar el vacío (si existe) que lo separa del siguiente número por la derecha y termina con el número completado en posición normal.

$$I_m = CD^m\ E \begin{cases} HG^m \\ FG^mA \end{cases}$$

La máquina I_m, cuando comienza con $a_1, \ldots, a_m$ en posición normal y con las a_1+2 casillas siguientes por la derecha en blanco, termina con $a_1, \ldots, a_m, a_1$, en posición normal.

$$K_m = AI_m^m F D^m F G^m.$$

La máquina K_m, cuando comienza con $a_1, \ldots, a_m$ en posición normal y las $a_1 + \cdots + a_m + 2m + 1$ casillas siguientes por la derecha en blanco, termina

imprimiendo $a_1, \ldots, a_m, , a_1, \ldots, a_m$, donde la doble coma ",," indica un vacío de dos blancos y donde el segundo a_m está en posición normal.

L	0	1
1	$I2$	$I1$
2	$I3$	$BI2$
3	$D4$	$BI2$
4	$D4$	$D5$
5	$I0$	$D5$

Si L comienza con un número en posición normal, borra todos los anteriores a él hasta el primer vacío y vuelve a la posición inicial.

Notemos que ninguna de las máquinas que hemos definido escruta las casillas a la izquierda de los datos.

5.7 La tesis de Church-Turing

Ya sabemos que toda función recursiva es calculable mediante un algoritmo. Una forma más explícita de este hecho es el teorema siguiente:

Teorema 5.11 *Toda función recursiva parcial es* $1 \mid 1$*-computable.*

Demostración: Por inducción sobre el número r de funciones de una sucesión que defina a f. Si $r = 1$ se trata de una función recursiva elemental. La función cero c es computada por la máquina C, la función sucesor s es computada por I_1A y la proyección p_i^k es computada por I_{k-i+1}. (En toda esta prueba, "computable" significará "$1 \mid 1$-computable".)

Supongamos ahora que f se define en r pasos y que todas las funciones definibles en menos de r pasos son computables. Distinguimos tres casos, según que f se defina por composición, recursión o minimización a partir de funciones anteriores (que por hipótesis de inducción serán computables).

Caso a) $f(a_1, \ldots, a_n) = g(h_1(a_1, \ldots, a_n), \ldots, h_m(a_1, \ldots, a_n))$, donde las funciones g y h_i son computables por máquinas M_g y M_{h_i} respectivamente.

Veamos que la función f es computada por la máquina

$$\begin{aligned} M_f \quad = \quad & K_m M_{h_1} I_{n+1}^n M_{h_2} I_{n+1}^n \cdots I_{n+1}^n M_{h_n} \\ & I_{(m-1)(n+1)+1} I_{(m-2)(n+1)+2} \cdots I_{0(n+1)+m} M_g LB. \end{aligned}$$

Supongamos definido $f(a_1, \ldots, a_n)$. Si M_f empieza con $a_1, \ldots, \overline{a}_n$ (el guión sobre a_n indica que está en posición normal), en primer lugar K_n copia $a_1, \ldots, a_n$ con un vacío en medio:

$$a_1, \ldots, a_n, , a_1, \ldots, \overline{a}_n.$$

Luego M_{h_1} calcula $h_1(a_1, \ldots a_n)$:

$$a_1, \ldots, a_n, , a_1, \ldots, a_n, \overline{h_1(a_1, \ldots, a_n)}.$$

Ahora I^n_{n+1} copia $a_1, \ldots, a_n$ y M_{h_2} calcula $h_2(a_1, \ldots, a_n)$:

$$a_1, \ldots, a_n, , a_1, \ldots, a_n, h_1(a_1, \ldots, a_n), a_1, \ldots, a_n, \overline{h_2(a_2, \ldots, a_n)}.$$

Tras haber actuado M_{h_m} tenemos

$$a_1, \ldots, a_n, , a_1, \ldots, a_n, h_1(a_1, \ldots, a_n), \ldots, a_1, \ldots, a_n, \overline{h_m(a_2, \ldots, a_n)}.$$

Seguidamente las máquinas $I_{(m-1)(n+1)+1} I_{(m-2)(n+1)+2} \cdots I_{0(n+1)+m}$ copian $h_1(a_1, \ldots, a_n), \ldots, h_m(a_1, \ldots, a_n)$ y entonces M_g calcula la imagen de estos números por g, o sea, calcula $f(a_1, \ldots, a_n)$. La situación de la cinta es entonces:

$$a_1, \ldots, a_n, , x_1, \ldots, x_r, \overline{f(a_1, \ldots, a_n)}.$$

La máquina L borra $x_1, \ldots, x_r$ y B borra el vacío intermedio, hasta quedar

$$a_1, \ldots, a_n, \overline{f(a_1, \ldots, a_n)}.$$

Las casillas a la izquierda del blanco anterior a a_1 nunca han sido escrutadas durante el cálculo.

Si $f(a_1, \ldots, a_n)$ no está definida, entonces no lo está alguna de las funciones g, $h_1, \ldots, h_n$, por lo que la máquina correspondiente no se para y M_f tampoco.

Caso b) La función f está definida por recurrencia a partir de las funciones g y h, es decir:

$$\begin{aligned} f(0, a_2, \ldots, a_n) &= g(a_2, \ldots, a_n), \\ f(a+1, a_2, \ldots, a_n) &= h(a, f(a, a_2, \ldots, a_n), a_2, \ldots, a_n). \end{aligned}$$

Por hipótesis de inducción existen máquinas M_g y M_h que computan a g y h respectivamente. Razonando de forma similar al caso anterior es fácil ver que la función f es computada por la máquina

$$M_f = K_n M_g I_{n+1} E \left[\begin{array}{l} I_2 L B \\ \\ C I_3 I^{n-1}_{n+3} M_h I_{n+3} F E \end{array} \right. \left[\begin{array}{l} I_2 L B \\ \\ I_{n+3} A \end{array} \right.$$

Una ligera modificación da cuenta del caso $n = 1$.

Caso c) $f(a_1, \ldots, a_n) = \mu x \, g(a_1, \ldots, a_n, x) = 0$. Por hipótesis de inducción existe una máquina M_g que computa a g. Entonces la función f es computada por la máquina

$$M_f = K_n C M_g E \left[\begin{array}{l} I_2 L B \\ \\ I^{n+1}_{n+2} A \end{array} \right.$$

De este modo, para cada función recursiva parcial f sabemos construir explícitamente una máquina de Turing que la computa. ∎

Es claro que cualquier función computable por una máquina de Turing es computable mediante un ordenador (salvo limitaciones de memoria). El recíproco no está claro. Las máquinas de Turing tienen, en principio, muy pocas capacidades de cálculo. No obstante hemos visto que pueden calcular cualquier función recursiva, lo que, a la larga, se traducirá en que la capacidad de cálculo de una máquina de Turing es idéntica a la de cualquier ordenador (superior —de hecho— por carecer de limitaciones de memoria). El punto más delicado de la demostración de la tesis de Church-Turing es probar el recíproco del teorema anterior. La clave del argumento está en que es general, en el sentido de que no sólo es aplicable a máquinas de Turing, sino que meros cambios técnicos permitirían adaptarlo para justificar que cualquier función calculable por un ordenador cualquiera es recursiva. Luego volveremos sobre este hecho.

Numeración de Gödel para máquinas de Turing Sea M una máquina de Turing con alfabeto $s_0, \ldots, s_j$. Sean $s_{u_1}, s_{u_2}, s_{u_3}, \ldots$ los signos impresos de derecha a izquierda a la izquierda de una casilla fija de la cinta. Llamaremos *número de Gödel* de la cinta a la izquierda de dicha casilla al número $u = \prod_{0<i} p_i^{u_i}$, donde $p_1, p_2, p_3, \ldots$ es la sucesión de los números primos 2, 3, 5, ...

El producto es finito, pues a partir de un cierto i todos los u_i serán 0 (la cinta estará en blanco). Los u_i son los exponentes de la descomposición en primos de u. Como esta descomposición es única, dado u, podemos calcular los u_i y saber, en consecuencia, la situación de la cinta a la izquierda de la casilla fijada.

Claramente tenemos que $u = \prod_{0<i<u} p_i^{iNu}$, donde iNu es la función recursiva que da el exponente del primo i-ésimo de u.

Análogamente, se define el *número de Gödel* de la cinta a la derecha de una casilla dada como $v = \prod_{0<i} p_i^{v_i}$, donde $s_{v_1}, s_{v_2}, \ldots$ son los signos impresos a la derecha de la casilla dada leídos de izquierda a derecha. Se cumple que $v = \prod_{0<i<u} p_i^{iNv}$.

Si, en un instante dado, el número de Gödel a la izquierda de la casilla escrutada es u, el signo escrutado es s_a, el estado de M es q_c y el número de Gödel a la derecha de la casilla escrutada es v, el *número de Gödel* de la configuración completa de la máquina en ese instante es, por definición, $w = 2^u \cdot 3^a \cdot 5^c \cdot 7^v$.

Conociendo w podemos descomponerlo en primos y calcular u, a, c y v y así obtenemos la configuración completa de la máquina.

Teorema 5.12 *Toda función parcial computable es recursiva parcial.*

Demostración: Sea ϕ una función parcial computable por una máquina de Turing M. Sean $q_0, \ldots q_k$ sus estados y $s_0, \ldots, s_j$ su alfabeto. Podemos suponer que q_0 es el único estado pasivo.

Definamos, para cada configuración activa (s_a, q_c) una función $\rho_{a,c}$ como sigue:

Si el acto tras (s_a, q_c) es $s_b I q_d$, entonces

$$\rho_{a,c}(u,v) = 2^{\prod_{0<i<u} \Pr(i)^{(i+1)Nu}} \cdot 3^{1Nu} \cdot 5^d \cdot 7^{2^b \cdot \prod_{0<i<u} \Pr(i+1)^{iNv}},$$

donde $\Pr(i)$ es el primo i-ésimo.

Si el acto tras (s_a, q_c) es $s_b C q_d$, entonces

$$\rho_{a,c}(u,v) = 2^u \cdot 3^b \cdot 5^d \cdot 7^v.$$

Si el acto tras (s_a, q_c) es $s_b D q_d$, entonces

$$\rho_{a,c}(u,v) = 2^{2^b \cdot \prod_{0<i<u} \Pr(i+1)^{iNu}} \cdot 3^{1Nv} \cdot 5^d \cdot 7^{\prod_{0<i<u} \Pr(i)^{(i+1)Nv}}.$$

En cualquier caso, si $2^u \cdot 3^a \cdot 5^c \cdot 7^v$ es el número de Gödel de la configuración completa de M en un instante, entonces $\rho_{a,c}(u,v)$ es el número de Gödel de la configuración completa siguiente. Es claro que cada función $\rho_{a,c}$ (como función de u, v únicamente) es recursiva primitiva.

Consideremos la relación dada por $R_{a,c}$ syss $2Nx = a \wedge 3Nx = c$, que claramente es recursiva primitiva, y sea $\chi_{a,c}$ su función característica, que también es recursiva primitiva. Definimos

$$\rho(w) = \sum_{\substack{a=0,\ldots,j \\ c=1,\ldots,k}} \rho_{a,c}(1Nw, 4Nw) \cdot \overline{\mathrm{sg}}\chi_{a,c}(w) + w \cdot \overline{\mathrm{sg}}(3Nw),$$

que es una función recursiva primitiva.

Si w es el número de Gödel de una configuración completa, $\rho(w)$ es el número de Gödel de la siguiente configuración completa (el sumando $w \cdot \overline{\mathrm{sg}}(3Nw)$ recoge el caso de que el estado sea pasivo, o sea, $3Nw = 0$, con lo que $\overline{\mathrm{sg}}(3Nw) = 1$ y así $\rho(w) = w$, es decir, la configuración no cambia).

Definimos ahora la función recursiva primitiva

$$\begin{aligned} \theta(w,0) &= w, \\ \theta(w,z+1) &= \rho(\theta(w,z)). \end{aligned}$$

Si w es el número de Gödel de una configuración completa, $\theta(w,z)$ es el número de Gödel de la configuración completa en que se halla M después de z actos (o la situación final si M se detiene antes).

Para cada número natural n vamos a definir $\tau_n(x_1, \ldots, x_n, c, u, v)$ de modo que si $x_1, \ldots, x_n$ está representado en posición normal, el estado es q_c y los números de Gödel de la cinta a la izquierda y a la derecha de x_1 y x_n son, respectivamente, u y v, entonces τ_n da el número de Gödel de la configuración completa.

Definimos primero

$$\tau_1(x_1, c, u, v) = 2^{\prod_{0<i\leq x_1} \Pr(i)^1 \cdot \prod_{0<i<u} \Pr(x_1+i+1)^{iNu}} \cdot 3^1 \cdot 5^c \cdot 7^{\prod_{0<i<v} \Pr(i+1)^{iNv}}.$$

Supuesta definida τ_n, definimos τ_{n+1} como

$$\tau_{n+1}(x_1,\ldots,x_n,x_{n+1},c,u,v) = \tau_1(x_{n+1},c,1N\tau_n(x_1,\ldots,x_{n-1},x_n{+}1,c,u,v),v).$$

Es fácil ver que las funciones τ_n cumplen lo pedido, así como que son recursivas primitivas.

Digamos que la función ϕ es n-ádica. Si $x_1,\ldots,x_n$ es escrutado en posición normal con estado q_1 y el resto de la cinta en blanco, la configuración completa es $\tau_n(x_1,\ldots,x_n,1,1,1)$. Así mismo, si $x_1,\ldots,x_n,x$ (para un cierto x) es escrutado en posición normal con estado q_0, la configuración completa es $\tau_{n+1}(x_1,\ldots,x_n,x,0,u,v)$, para ciertos u, v, y viceversa.

Así pues, $\phi(x_1,\ldots,x_n)$ está definido si y sólo si existen z, x, u y v tales que

$$\theta(\tau_n(x_1,\ldots,x_n,1,1,1),z) = \tau_{n+1}(x_1,\ldots,x_n,x,0,u,v),$$

y entonces $\phi(x_1,\ldots,x_n) = x$.

Equivalentemente, $\phi(x_1,\ldots,x_n)$ está definido si y sólo si existe un número natural t ($t = 2^z\cdot 3^x\cdot 5^u\cdot 7^v$) tal que

$$\theta(\tau_n(x_1,\ldots,x_n,1,1,1),1Nt) = \tau_{n+1}(x_1,\ldots,x_n,2Nt,0,3Nt,4Nt), \qquad (5.1)$$

y entonces $\phi(x_1,\ldots,x_n) = nNt$.

Si llamamos $S(x_1,\ldots,x_n,t)$ a la relación determinada por (5.1), se trata de una relación recursiva primitiva y tenemos que

$$\phi(x_1,\ldots,x_n) = 2N\big(\mu t\, S(x_1,\ldots,x_n,t)\big),$$

con lo que ϕ es recursiva parcial. ∎

Como consecuencias inmediatas tenemos:

Teorema 5.13 *Una función parcial es recursiva parcial si y sólo si es computable, si y sólo si es* $1 \mid 1$ *computable.*

Teorema 5.14 *Una función es recursiva si y sólo si es computable.*

Nota Según anticipábamos, es fácil convencerse de que la prueba del teorema anterior puede adaptarse para probar que toda función calculable con un programa de ordenador es recursiva. Para ello sólo hay que complicarla teniendo en cuenta la complejidad adicional de un ordenador frente a una máquina de Turing, pero es claro que no es necesario aportar ninguna idea nueva, sino que el esquema general del argumento sería exactamente el mismo. De hecho es teóricamente más simple, pues la cinta infinita se sustituye por una memoria finita, que sólo puede estar en un número finito de configuraciones. El comportamiento del ordenador está determinado por la configuración de su memoria y por el estado de su microprocesador (incluyendo aquí cualquier hecho relevante, aunque no corresponda estrictamente al microprocesador). Podemos introducir una numeración de Gödel para la configuración de la memoria y el estado del

microprocesador y la función que a partir del número de Gödel de la configuración completa calcula el de la siguiente configuración completa (entendiendo que vale 0 si el número de partida no corresponde a ninguna configuración completa posible) es recursiva primitiva (obviamente, pues es una función definida sobre una cantidad finita de números). La función que a partir de la configuración completa inicial calcula la configuración al cabo de n pasos es recursiva, lo cual se prueba exactamente igual que para máquinas de Turing y a su vez nos lleva ya sin ningún cambio a la conclusión del teorema.[4]

Teniendo esto en cuenta, el concepto de función recursiva puede considerarse como una caracterización precisa de la noción de computabilidad, al igual que el concepto de demostración formal es una caracterización precisa del concepto de razonamiento matemático. Algunos autores afirman que la tesis de Church-Turing es indemostrable, porque la noción de computabilidad mediante un algoritmo no admite una definición precisa. Lo que hay de verdad en esta afirmación es que la noción de computabilidad es metamatemática, y cualquier intento de formalización, por ejemplo, a través de la noción de función computable por una máquina de Turing, suscita la duda de si realmente estamos capturando la totalidad de las funciones computables en sentido metamatemático. Ahora bien, acabamos de probar que así es. La demostración es concluyente en el sentido de que nadie que medite sobre ella puede dudar de que cualquier función que pueda calcular un ordenador satisface necesariamente la definición de función recursiva. Ya hemos tenido ocasión de constatar en muchas ocasiones que el hecho de no disponer de una definición formal explícita de algunos conceptos, como la finitud o los números naturales, no nos impide asegurar ciertos hechos sobre ellos.

Ejercicio: Dada una función recursiva f y una máquina de Turing M que la compute $1|1$, construir otra máquina de Turing que cuando empiece con la cinta en blanco vaya escribiendo sucesivamente en la cinta los números $f(0)$, $f(1)$, $f(2)$, etc.

5.8 Consideraciones finales

Terminamos el capítulos con algunas consideraciones adicionales sobre la recursividad y las máquinas de Turing. Ante todo veamos cómo éstas nos proporcionan un ejemplo explícito muy simple de función no recursiva.

Ejemplo Sea n un número natural. Un problema que se ha convertido en entretenimiento de algunos amantes de los acertijos matemáticos es el siguiente: encontrar una máquina de Turing M con dos signos 0 y 1 y a lo sumo n estados con la condición de que cuando empieza con la cinta en blanco se detiene tras haber escrito el máximo número posible de 1's. En otras palabras, se trata de encontrar el "récord" de unos que puede escribir una máquina de Turing con n

[4]En este argumento prescindimos de toda incorporación de nuevos datos durante el cálculo, es decir, suponemos que el ordenador no tiene teclado o conexión con otros ordenadores. Esto no es una restricción, pues únicamente nos interesa el intervalo comprendido desde que el ordenador tiene todos los datos introducidos hasta que termina el cálculo.

estados excluyendo el caso trivial de que no se detenga nunca y escriba infinitos unos.

Más explícitamente, para cada máquina M que acaba deteniéndose cuando empieza con la cinta en blanco, llamamos p_M al número de unos que tiene la cinta cuando esto ocurre. Definimos $\Sigma(n)$ como el máximo de los números p_M, cuando M varía entre las máquinas que acaban deteniéndose al empezar con la cinta en blanco. El problema es, entonces, calcular los valores de Σ.

Teorema 5.15 *La función Σ no es recursiva.*

DEMOSTRACIÓN: Sea f una función recursiva y definamos

$$g(n) = \text{máx}\{f(2n+2), f(2n+3)\}.$$

Es fácil ver que g es recursiva y por lo tanto es computada por una máquina de Turing M. Sea k el número de estados activos de M.

Para cada número natural n sea N_n una máquina que al empezar con la cinta en blanco escriba el número n en la cinta y después actúe como M. Podemos construir N_n con $n+k+2$ estados. Cuando N_n actúa con la cinta en blanco, al acabar está escrito (entre otras cosas) $g(n)$, es decir, hay $g(n)+1$ unos en la cinta como mínimo.

Por lo tanto,

$$\text{máx}\{f(2n+2), f(2n+3)\} + 1 \leq \Sigma(n+k+2) \qquad \text{para todo } n.$$

Si $n \geq k$ tenemos que

$$f(2n+2), f(2n+3) < \Sigma(n+k+2) \leq \Sigma(2n+2) \leq \Sigma(2n+3),$$

ya que Σ es evidentemente creciente. Pero todo número x que cumpla $2k+3 \leq x$ puede expresarse como $x = 2n+2$ o $x = 2n+3$, para un cierto número $n \geq k$, y así $f(x) < \Sigma(x)$.

Hemos probado que Σ supera a cualquier función recursiva a partir de un cierto número natural. En particular Σ no es recursiva. ∎

Estos son algunos datos conocidos sobre Σ:

$$\Sigma(1) = 1, \qquad \Sigma(2) = 4, \qquad \Sigma(3) = 6, \qquad \Sigma(4) = 13,$$

$$\Sigma(5) \geq 17, \qquad \Sigma(6) \geq 35, \qquad \Sigma(7) \geq 22.961, \qquad \Sigma(8) \geq 8 \cdot 10^{44}.$$

El problema de la detención Una pregunta natural que plantea la no recursividad de la función Σ es qué nos impide calcularla. Para calcular $\Sigma(n)$ hay que tomar todas las máquinas de Turing con dos signos y n estados, que son un número finito, seleccionar las que se detienen al empezar con la cinta en blanco y contar el máximo número de unos impreso por cada una de ellas. El único paso que no es evidentemente realizable es determinar cuáles se detienen, por lo que concluimos que no existe un método general para decidir si una máquina de

Turing va a detenerse o no cuando comienza con una situación dada, es decir, el problema de la detención de las máquinas de Turing es insoluble.

En algunos casos podremos concluir algo a partir del análisis del programa, por ejemplo es fácil ver que la máquina D no se detiene cuando empieza con la cinta en blanco. En otros casos, en cambio, lo mejor que podremos hacer será ponerla a funcionar y ver si se detiene. Si lo hace sabremos que se para, pero si no se detiene nos quedaremos con la duda de si se va a parar más adelante o si no se va a parar nunca.

Ejercicio: La conjetura de Goldbach es una afirmación de la teoría de números que hasta ahora no ha sido demostrada ni refutada. Consideremos la función G que es constantemente igual a 1 si la conjetura de Goldbach es cierta y constante igual a 0 si es falsa. ¿Es una función recursiva?

Capítulo VI

Teorías aritméticas

Los teoremas de incompletitud que pretendemos demostrar se aplican a teorías recursivas, (es decir, teorías axiomáticas en las que sabemos reconocer efectivamente si una fórmula es o no un axioma) y aritméticas, es decir, teorías en las que es posible demostrar formalmente las propiedades básicas de los números naturales. Ya conocemos la teoría aritmética más simple: la aritmética de Peano. En este capítulo definiremos el concepto general de teoría aritmética y probaremos los hechos que necesitamos para llegar a los teoremas de incompletitud.

6.1 Definición y propiedades básicas

Definición Una *teoría aritmética* T es una teoría axiomática sobre un lenguaje formal $\mathfrak{L}$ que cumple las condiciones siguientes:

Existen un designador 0 de $\mathfrak{L}$, un término x' de $\mathfrak{L}$ cuya única variable libre es x, dos términos $x + y$, $x \cdot y$, cuyas únicas variables libres son x e y (distintas entre sí) y una fórmula $\mathrm{Nat}\, x$ cuya única variable es x, de tal forma que las fórmulas siguientes son teoremas de T:

N1: $\mathrm{Nat}\, 0$,

N2: $\bigwedge x(\mathrm{Nat}\, x \to \mathrm{Nat}\, x')$,

N3: $\bigwedge x(\mathrm{Nat}\, x \to \neg x' = 0)$,

N4: $\bigwedge xy(\mathrm{Nat}\, x \wedge \mathrm{Nat}\, y \wedge x' = y' \to x = y)$,

N5: $\bigwedge x(\mathrm{Nat}\, x \to x + 0 = x)$,

N6: $\bigwedge xy(\mathrm{Nat}\, x \wedge \mathrm{Nat}\, y \to x + y' = (x + y)')$,

N7: $\bigwedge x(\mathrm{Nat}\, x \wedge x \cdot 0 = 0)$,

N8: $\bigwedge xy(\mathrm{Nat}\, x \wedge \mathrm{Nat}\, y \to x \cdot y' = (x \cdot y) + x)$,

N9: $(\alpha(0) \wedge \bigwedge x(\mathrm{Nat}\, x \wedge \alpha(x) \to \alpha(x'))) \to \bigwedge x(\mathrm{Nat}\, x \to \alpha(x))$, para toda fórmula aritmética $\alpha(x)$ que tenga a x como variable libre (no necesariamente la única), donde una *expresión aritmética* se define como sigue:

a) x es un término aritmético,

b) 0 es un término aritmético,

c) Si t_1 y t_2 son términos aritméticos, también lo son t_1', $t_1 + t_2$, $t_1 \cdot t_2$,

d) Si t_1 t_2 son términos aritméticos, entonces $t_1 = t_2$ es una fórmula aritmética,

e) Si α y β son fórmulas aritméticas, también lo son

$$\neg\alpha, \quad \alpha \to \beta, \quad \bigwedge x(\mathrm{Nat}\, x \to \alpha), \quad \bigvee x(\mathrm{Nat}\, x \wedge \alpha).$$

Notar que, en particular, si α y β son fórmulas aritméticas, también lo son $\alpha \vee \beta$, $\alpha \wedge \beta$ y $\alpha \leftrightarrow \beta$.

En la práctica consideraremos como fórmulas aritméticas a todas las fórmulas que sean lógicamente equivalentes a fórmulas aritméticas. Por ejemplo, la sentencia N4 no es aritmética en sentido estricto, pero es equivalente a

$$\bigwedge x(\mathrm{Nat}\, x \to \bigwedge y(\mathrm{Nat}\, y \to (x' = y' \to x = y))),$$

que sí lo es. Con este convenio, todas las fórmulas N1–N9 son aritméticas. Es fácil ver que las fórmulas de tipo N9, donde α es aritmética en este sentido amplio, son también teoremas de T.

Observemos también que si t es un término aritmético, entonces $\mathrm{Nat}\, t$ es una fórmula aritmética, pues equivale a $\bigvee x(\mathrm{Nat}\, x \wedge x = t)$.

Escribiremos $0^{(0)} \equiv 0$, $0^{(1)} \equiv 0'$, $0^{(2)} \equiv 0''$, $0^{(3)} \equiv 0'''$, etc. A los designadores $0^{(n)}$, donde n es un número natural, los llamaremos *numerales* de T.

Ejemplo La aritmética de Peano es un ejemplo de teoría aritmética. Basta definir $\mathrm{Nat}\, x \equiv x = x$. Es fácil ver que todas las fórmulas de su lenguaje formal son aritméticas. En su modelo natural (el que tiene por universo a los números naturales) cada numeral $0^{(n)}$ denota al número natural n.

Observaciones Conviene insistir en que la definición de teoría aritmética no se exige, por ejemplo, que 0 sea una constante, sino que puede ser un designador arbitrario. Así, mientras en la aritmética de Peano 0 es ciertamente una constante, en teoría de conjuntos es frecuente definir $0 \equiv \varnothing \equiv x \mid \bigwedge y\, y \notin x$, con lo que 0 es un designador de longitud 8. Lo mismo sucede con los demás signos aritméticos, que no tienen por qué ser signos del lenguaje formal considerado, sino que cada signo aritmético puede corresponderse con un término o fórmula más o menos complejo. En otras palabras, admitimos que la aritmética sea definida en T a partir de otros conceptos.

Similarmente, hemos de recordar que las fórmulas N1–N9 no son necesariamente axiomas de una teoría aritmética, sino teoremas. Por ejemplo, ningún axioma de la teoría de conjuntos afirma que el 0 es un número natural, pero esto puede probarse a partir de las definiciones oportunas de "cero" y "número natural". Pese a ello, usualmente nos referiremos a estas fórmulas como los "axiomas de Peano".[1]

Por otra parte, el hecho de que la "definición" de número natural en la aritmética de Peano sea $x = x$ es una muestra extrema de algo que el lector debería asimilar: las definiciones formales nunca *determinan* los objetos definidos, sino que tan sólo los *especifican* entre otros objetos previos. Así, la definición de número natural en teoría de conjuntos no determina qué es un número natural, sino que únicamente especifica qué conjuntos son números naturales, pero a su vez la noción de conjunto carece de definición: en teoría de conjuntos, todo es un conjunto. Una definición de conjunto[2] sería $x = x$. En el caso de la aritmética de Peano sucede que todo es un número natural, por lo que no hay nada que especificar y por ello podemos tomar como definición de número natural la fórmula $x = x$. ∎

Interpretación natural de las fórmulas aritméticas Aunque las teorías aritméticas "pretenden" hablar sobre los números naturales, no es inmediato en qué sentido esto es así. Pensemos en una sentencia como $\alpha \equiv 0'' + 0'' = 0''''$, perteneciente a una teoría aritmética T. Ahí "leemos" $2 + 2 = 4$, pero, ¿en qué sentido podemos decir que ahí pone realmente que $2 + 2 = 4$? Si α fuera una sentencia de la aritmética de Peano, podríamos decir que $2 + 2 = 4$ es su interpretación en el modelo natural, pero si T es una teoría arbitraria, en principio no disponemos de ningún modelo "natural" de T que fundamente una interpretación de α. Tengamos presente que para definir un modelo de T tendríamos que interpretar todos los signos de su lenguaje formal, que pueden ser muchos más de los que se relacionan con la estructura aritmética de la teoría.

Pese a ello, podemos hablar de una interpretación natural de cualquier fórmula aritmética de cualquier teoría aritmética con respecto a la cual la interpretación de α sea la que cabe esperar. Para ello observamos que a cada expresión aritmética θ de una teoría T le podemos asignar una expresión (necesariamente aritmética) $\overline{\theta}$ de la aritmética de Peano $\mathcal{P}$ de forma inductiva:

a) $\overline{0} \equiv 0$ (es decir, asociamos al designador 0 de T la constante 0 de $\mathcal{P}$),

b) $\overline{St} \equiv S\overline{t}, \quad \overline{t_1 + t_2} \equiv \overline{t}_1 + \overline{t}_2, \quad \overline{t_1 \cdot t_2} \equiv \overline{t}_1 \cdot \overline{t}_2,$

[1] Son axiomas en el mismo sentido en que los matemáticos hablan de "los axiomas de espacio vectorial", que no son axiomas de acuerdo con nuestra definición, sino fragmentos de la definición de espacio vectorial. En sentido histórico estricto los axiomas de Peano serían los cinco axiomas que resultan de eliminar los cuatro que definen la suma y el producto.

[2] Hay teorías de conjuntos que establecen una distinción entre "clase" y "conjunto", de modo que todo conjunto es una clase pero el recíproco es falso. En estas teorías el concepto de clase es el que carece de definición o, si queremos, se puede definir como $x = x$, mientras que la definición de conjunto usual es la dada por Gödel: Cto $x \equiv \bigvee y\, x \in y$. También aquí es claro que esta definición no determina la noción de conjunto, sino que únicamente especifica qué ha de cumplir una clase para ser un conjunto.

c) $\overline{t_1 = t_2} \equiv \overline{t}_1 = \overline{t}_2$,

d) $\overline{\neg\alpha} \equiv \neg\overline{\alpha}, \quad \overline{\alpha \to \beta} \equiv \overline{\alpha} \to \overline{\beta}$,

e) $\overline{\bigwedge x(\mathrm{Nat}\, x \to \alpha)} \equiv \bigwedge x\, \overline{\alpha}$.

En otros términos, lo que decimos es que cualquier fórmula aritmética puede leerse "como si fuera" una fórmula de $\mathcal{P}$. Diremos que una sentencia aritmética α es verdadera o falsa en su interpretación natural si $\overline{\alpha}$ es verdadera o falsa en la interpretación natural de $\mathcal{P}$.

Así pues, debemos tener presente que cuando hablamos de la interpretación natural de una sentencia aritmética α de una teoría T, en realidad estamos considerando un modelo de $\mathcal{P}$. Más adelante veremos que puede darse el caso de que una sentencia aritmética α sea verdadera en cualquier modelo de una cierta teoría aritmética T y que, en cambio, sea falsa en su interpretación natural, o viceversa.

6.2 Algunos teoremas en teorías aritméticas

Las sentencias siguientes son teoremas de cualquier teoría aritmética. (Damos algunas indicaciones sobre la demostración.)

1. $\bigwedge xy(\mathrm{Nat}\, x \wedge \mathrm{Nat}\, y \to \mathrm{Nat}(x + y))$

   Por inducción sobre y, es decir, probando la sentencia

   $$\bigwedge x(\mathrm{Nat}\, x \to \bigwedge y(\mathrm{Nat}\, y \to \mathrm{Nat}(x + y))),$$

   para lo cual se prueba previamente (supuesto $\mathrm{Nat}\, x$) la fórmula

   $$\mathrm{Nat}(x + 0) \wedge \bigwedge y(\mathrm{Nat}\, y \wedge \mathrm{Nat}(x + y) \to \mathrm{Nat}(x + y'))$$

   y se aplica N9. Para probar esta fórmula vemos que $x + 0 = x$, luego $\mathrm{Nat}(x + 0)$ y si $\mathrm{Nat}(x + y)$ entonces $x + y' = (x + y)'$, luego también $\mathrm{Nat}(x + y')$, por N2.

2. $\bigwedge xy(\mathrm{Nat}\, x \wedge \mathrm{Nat}\, y \to \mathrm{Nat}(xy))$

   Similar al caso anterior.

3. $\bigwedge x(\mathrm{Nat}\, x \to x' = x + 0^{(1)})$

   $x + 0^{(1)} = x + 0' = (x + 0)' = x'$.

4. $\bigwedge xyz(\mathrm{Nat}\, x \wedge \mathrm{Nat}\, y \wedge \mathrm{Nat}\, z \to (x + y) + z = x + (y + z))$

   Por inducción sobre z: $(x + y) + 0 = x + y = x + (y + 0)$ y supuesto $(x + y) + z = x + (y + z)$ entonces $((x + y) + z)' = (x + (y + z))'$, luego $(x + y) + z' = x + (y + z)' = x + (y + z')$.

5. $\bigwedge xy(\mathrm{Nat}\, x \wedge \mathrm{Nat}\, y \to x + y' = x' + y)$

   Por inducción sobre y: $x + 0' = (x + 0)' = x' = x' + 0$, y si $x + y' = x' + y$ entonces $x + y'' = (x + y')' = (x' + y)' = x' + y'$.

6. $\bigwedge x(\mathrm{Nat}\, x \to 0 + x = x)$

   Por inducción: $0 + 0 = 0$ y si $0 + x = x$, entonces $0 + x' = (0 + x)' = x'$.

7. $\bigwedge xy(\mathrm{Nat}\, x \wedge \mathrm{Nat}\, y \to x + y = y + x)$

   Por inducción sobre y: $x + 0 = x = 0 + x$ y si $x + y = y + x$, entonces $x + y' = (x + y)' = (y + x)' = y + x' = y' + x$.

8. $\bigwedge xyz(\mathrm{Nat}\, x \wedge \mathrm{Nat}\, y \wedge \mathrm{Nat}\, z \to x(y + z) = xy + xz)$

   Por inducción sobre z: $x(y + 0) = xy = xy + x \cdot 0$, y si $x(y + z) = xy + xz$ entonces $x(y+z') = x(y+z)' = x(y+z)+x = (xy+xz)+x = xy+(xz+x) = xy + xz'$.

9. $\bigwedge xyz(\mathrm{Nat}\, x \wedge \mathrm{Nat}\, y \wedge \mathrm{Nat}\, z \to (xy)z = x(yz))$

   Por inducción sobre z: $(xy) \cdot 0 = 0 = x \cdot 0 = x(y \cdot 0)$, y si $(xy)z = x(yz)$, entonces $(xy)z' = (xy)z + xy = x(yz) + xy = x(yz + y) = x(yz')$.

10. $\bigwedge x(\mathrm{Nat}\, x \to 0 \cdot x = 0)$

    Por inducción: $0 \cdot 0 = 0$ y si $0 \cdot x = 0$, entonces $0 \cdot x' = 0 \cdot x + 0 = 0 + 0 = 0$.

11. $\bigwedge x(\mathrm{Nat}\, x \to 0^{(1)} \cdot x = x)$

    Por inducción: $0^{(1)} \cdot 0 = 0$ y si $0^{(1)} \cdot x = x$, entonces $0^{(1)} \cdot x' = 0^{(1)} \cdot x + 0^{(1)} = x + 0^{(1)} = x'$.

12. $\bigwedge xy(\mathrm{Nat}\, x \wedge \mathrm{Nat}\, y \to xy = yx)$

    Por inducción sobre y: $x \cdot 0 = 0 = 0 \cdot x$, y si $xy = yx$ entonces $xy' = xy + x = yx + x = yx + 0^{(1)} \cdot x = (y + 0^{(1)}) \cdot x = y'x$.

13. $\bigwedge x(\mathrm{Nat}\, x \to x = 0 \vee \bigvee y(\mathrm{Nat}\, y \wedge x = y'))$

    Inmediato por inducción sobre x.

14. $\bigwedge xy(\mathrm{Nat}\, x \wedge \mathrm{Nat}\, y \wedge xy = 0 \to x = 0 \vee y = 0)$

    Si $xy = 0$ pero $\neg x = 0$ y $\neg y = 0$ entonces $x = u' = u + 0^{(1)}$, $y = v' = v + 0^{(1)}$, con lo que $0 = xy = uv + u + v + 0^{(1)} = (uv + u + v)'$, contradicción.

15. $\bigwedge xyz(\mathrm{Nat}\, x \wedge \mathrm{Nat}\, y \wedge \mathrm{Nat}\, z \wedge x + z = y + z \to x = y)$

    Por inducción sobre z: si $x + 0 = y + 0$ entonces $x = y$, si $x + z = y + z \to x = y$, entonces si $x + z' = y + z'$ se cumple $(x + z)' = (y + z)'$, luego $x + z = y + z$, luego $x = y$.

16. $\bigwedge xy(\mathrm{Nat}\, x \wedge \mathrm{Nat}\, y \wedge x + y = 0 \to x = 0 \wedge y = 0)$

    Si $\neg y = 0$ entonces $y = v'$, con lo que $0 = x + y = (x + v)'$, contradicción. Así pues $y = 0$. Similarmente se prueba que $x = 0$.

    Usaremos la notación $x \leq y \equiv \bigvee z(\mathrm{Nat}\, z \wedge x + z = y)$

17. $\bigwedge x(\mathrm{Nat}\, x \to x \leq x)$

    $x + 0 = x$.

18. $\bigwedge xy(\mathrm{Nat}\, x \wedge \mathrm{Nat}\, y \wedge x \leq y \wedge y \leq x \to x = y)$

    Si $x \leq y \wedge y \leq x$ entonces $x+z = y \wedge y+w = x$, luego $y+w+z = y = y+0$, luego $w + z = 0$, luego $z = 0$, luego $x = x + 0 = y$.

19. $\bigwedge xyz(\mathrm{Nat}\, x \wedge \mathrm{Nat}\, y \wedge \mathrm{Nat}\, z \wedge x \leq y \wedge y \leq z \to x \leq z)$

    $x + u = y \wedge y + v = z$, luego $x + (u + v) = y + v = z$, luego $x \leq z$.

20. $\bigwedge x(\mathrm{Nat}\, x \to 0 \leq x)$

    $0 + x = x$

21. $\bigwedge x(\mathrm{Nat}\, x \to x \leq x')$

    $x + 0^{(1)} = x'$

22. $\bigwedge xy(\mathrm{Nat}\, x \wedge \mathrm{Nat}\, y \to x \leq y \vee y \leq x)$

    Por inducción sobre y: $0 \leq x$, luego $x \leq 0 \vee 0 \leq x$, si $x \leq y \vee y \leq x$ entonces, en el caso $y \leq x$ tenemos $y + z = x$. Si $z = 0$ es $y = x$, luego $x \leq x' = y'$, luego $x \leq y' \vee y' \leq x$. Si $\neg z = 0$, entonces $z = u'$, $y + u' = x$, $y + (u + 0^{(1)}) = x$, luego $(y + 0^{(1)}) + u = x$, $y' + u = x$, $y' \leq x$, luego también $x \leq y' \vee y' \leq x$. En el caso $x \leq y$ se cumple $x + z = y$, luego $x + z' = y'$, es decir, $x \leq y'$, y también $x \leq y' \vee y' \leq x$.

23. $\bigwedge xy(\mathrm{Nat}\, x \wedge \mathrm{Nat}\, y \wedge x \leq y \wedge y \leq x' \to x = y \vee y = x')$

    $x + u = y \wedge y + v = x'$, luego $x + (u + v) = x' = x + 0^{(1)}$, $u + v = 0^{(1)}$. Si $\neg y = x \wedge \neg y = x'$ entonces $\neg y = 0 \wedge \neg v = 0$, luego $u = k + 0^{(1)} \wedge v = r + 0^{(1)}$, luego $u + v = ((k + r) + 0^{(1)}) + 0^{(1)} = 0^{(1)}$, luego $(k + r) + 0^{(1)} = 0$, es decir, $0 = (k + r)'$, contradicción.

    Usaremos la notación $x < y \equiv x \leq y \wedge \neg x = y$.

24. $\bigwedge x(\mathrm{Nat}\, x \to x < x')$

    Sabemos que $x \leq x'$ y $\neg x = x'$, pues en otro caso $x = x' = x + 0^{(1)}$, luego $0 = 0^{(1)}$, contradicción.

25. $\bigwedge xy(\mathrm{Nat}\, x \wedge \mathrm{Nat}\, y \to (x < y \leftrightarrow x' \leq y))$

    Si $x < y$ entonces $x+u = y$, con $\neg u = 0$, luego $u = v+0^{(1)}$, $x+(v+0^{(1)}) = y$, $(x + 0^{(1)}) + v = y$, $x' + v = y$, $x' \leq y$.

    Si $x' \leq y$, entonces $x' + u = y$, $x + (0^{(1)} + u) = y$, luego $x \leq y$, y si $x = y$ entonces $0^{(1)} + u = 0$, luego $u' = 0$, contradicción. Así pues, $x < y$.

26. $\bigwedge xy(\mathrm{Nat}\, x \wedge \mathrm{Nat}\, y \wedge \neg y = 0 \to \overset{1}{\bigvee} cr(\mathrm{Nat}\, c \wedge \mathrm{Nat}\, r \wedge x = yc + r \wedge r < y))$

    Veamos

    $$\bigwedge xy(\mathrm{Nat}\, x \wedge \mathrm{Nat}\, y \wedge \neg y = 0 \to \bigvee cr(\mathrm{Nat}\, c \wedge \mathrm{Nat}\, r \wedge x = yc + r \wedge r < y))$$

    por inducción sobre x.

    $0 = y \cdot 0 + 0 \wedge 0 < y$. Si $x = yc + r \wedge r < y$, entonces $x' = yc + r'$ con $r' \leq y$. Si $r' < y$ ya lo tenemos. Si $r' = y$ entonces $x' = yc + y = yc' + 0$ con $0 < y$.

    Para probar la unicidad basta ver que si $x = yc + r = y\bar{c} + \bar{r}$ con $r < y$, $\bar{r} < y$, entonces $c = \bar{c} \wedge r = \bar{r}$.

    Podemos suponer que $c \leq \bar{c}$ (el caso $\bar{c} \leq c$ es análogo). Así $\bar{c} = c + u$. Si $\neg c = \bar{c}$ entonces $\neg u = 0$, $u = v + 0^{(1)}$.

    $$x = y\bar{c} + \bar{r} = (yc + yu) + \bar{r} = ((yc + yv) + y \cdot 0^{(1)}) + \bar{r} = (yc + y) + (yv + \bar{r}),$$

    luego $yc + y \leq x$. Como $r < y$, $y = r + k$ con $\neg k = 0$, $x = (yc + y) + t = (yc + (r + k)) + t = (yc + r) + (k + t) = x + (k + t)$, luego $k + t = 0$ y $k = 0$, contradicción.

    Por lo tanto $c = \bar{c}$, luego $yc = y\bar{c}$ y, como $yc + r = y\bar{c} + \bar{r}$, también $r = \bar{r}$.

    Usaremos la notación

    $$x \equiv y \ (\mathrm{mód}\ z) \equiv \bigvee r(\mathrm{Nat}\, r \wedge (x = y + rz \vee y = x + rz))$$

27. $\bigwedge xy(\mathrm{Nat}\, x \wedge \mathrm{Nat}\, y \to x \equiv x \ (\mathrm{mód}\ y))$

    $x = x + 0 \cdot y$.

28. $\bigwedge xyz(\mathrm{Nat}\, x \wedge \mathrm{Nat}\, y \wedge \mathrm{Nat}\, z \wedge x \equiv y \ (\mathrm{mód}\ z) \to y \equiv x \ (\mathrm{mód}\ z))$

    Inmediato.

29. $\bigwedge xyzw(\mathrm{Nat}\, x \wedge \mathrm{Nat}\, y \wedge \mathrm{Nat}\, z \wedge \mathrm{Nat}\, w \wedge x \equiv y \ (\mathrm{mód}\ w) \wedge y \equiv z \ (\mathrm{mód}\ w) \to x \equiv z \ (\mathrm{mód}\ w))$

    Por hipótesis $(x = y + rw \vee y = x + rw) \wedge (y = z + tw \vee z = y + tw)$. Distinguimos cuatro casos:

a) $x = y + rw \wedge y = z + tw$. Entonces $x = x + (tw + rw) = x + (t + r)w$, luego $x \equiv z \pmod{w}$.

b) $x = y + rw \wedge z = y + tw$. Entonces $x + tw = (y + tw) + rw = z + rw$. Si $r \leq t$ entonces $t = r + u$, luego $x + (rw + uw) = z + rw$, $x + uw = z$. Si $t \leq r$ entonces $r = t + u$, luego $x + tw = z + (tw + uw)$, $x = z + uw$. En cualquier caso $x \equiv z \pmod{w}$.

Los casos restantes son similares a éstos.

30. $\bigwedge xz(\operatorname{Nat} x \wedge \operatorname{Nat} z \wedge \neg z = 0 \to \overset{1}{\bigvee} y(\operatorname{Nat} y \wedge y < z \wedge x \equiv y \pmod{z}))$

$\bigvee cy(\operatorname{Nat} c \wedge \operatorname{Nat} y \wedge x = zc + y \wedge y < z)$. Claramente $x \equiv y \pmod{z}$.

Para la unicidad hemos de ver que si $x \equiv y \pmod{z} \wedge x \equiv \bar{y} \pmod{z}$ con $y < z$, $\bar{y} < z$ entonces $y = \bar{y}$. En efecto, en tal caso $y \equiv \bar{y} \pmod{z}$, luego $y = \bar{y} + rz \vee \bar{y} = y + rz$. Supongamos $y = \bar{y} + rz$ (el otro caso es análogo). Entonces, si $\neg r = 0$ tenemos $r = u + 0^{(1)}$, luego $y = \bar{y} + (ux + 0^{(1)} \cdot z) = (y + ux) + z$, luego $z \leq y$, contradicción. Así pues, $r = 0$, con lo que $y = \bar{y}$.

Observaciones El lector se habrá dado cuenta de que hemos usado muchas de las propiedades que acabamos de probar mucho antes de haberlas demostrado aquí. A estas alturas, ya debería darse cuenta por sí mismo de que esto no delata un círculo vicioso, pero de todos modos vamos a discutirlo aquí. Consideremos por ejemplo el teorema 7, que afirma la conmutatividad de la suma de números naturales. Hemos de distinguir dos hechos muy diferentes:

a) La suma de números naturales es conmutativa.

b) En toda teoría aritmética puede probarse que la suma de números naturales es conmutativa.

Hasta ahora, hemos usado siempre que ha sido oportuno el hecho a), que todos sabemos que es cierto. Ahora acabamos de probar el hecho b), que no es evidente ni siquiera para alguien que sepa a). Sucede que b) implica a), pero esto no es inmediato, sino que es un teorema que requiere su justificación: si una fórmula es demostrable en toda teoría aritmética, entonces es demostrable en la aritmética de Peano, luego es verdadera en su interpretación natural. Ahora bien, aquí es esencial lo de "toda" teoría aritmética, pues la única en la que podemos confiar es la aritmética de Peano. Veremos más adelante que existen teorías aritméticas consistentes en las que se pueden demostrar sentencias falsas en su interpretación natural. En resumen:

- No es cierto que a) implique b), es decir, el hecho de que una sentencia sea verdadera en su interpretación natural no garantiza que sea demostrable en toda teoría aritmética.

- No es cierto que si una sentencia es demostrable en una teoría aritmética tenga por ello que ser verdadera en su interpretación natural.

- Sí es cierto que si una sentencia es demostrable en cualquier teoría aritmética entonces es verdadera en su interpretación natural —es decir, que b) implica a)—, pero esto es una afirmación trivial, pues lo único que necesitamos en realidad es que sea demostrable en la aritmética de Peano.

Si el lector se siente desconcertado por estos hechos debería tener en cuenta las consideraciones siguientes. Supongamos que partimos de unos axiomas similares a los de Peano y con ellos probamos que la suma de números naturales no es conmutativa. No sería difícil conseguir unos axiomas adecuados para ello, pero la conclusión que sacaríamos de ahí no sería que, en contra de lo que pensábamos, la suma de números naturales no es conmutativa, sino que los axiomas de partida "no son buenos", es decir, no reflejan las propiedades de los números naturales. Así pues, si una prueba de la no conmutatividad de la suma no nos hace dudar de la conmutatividad de la suma, sino que nos lleva a desechar los axiomas, debemos admitir que una prueba de la conmutatividad de la suma a partir de unos axiomas no debe convencernos de la conmutatividad de la suma, sino de que los axiomas "son buenos". El hecho de que no estemos dispuestos a descartar la conmutatividad a partir de una demostración formal pone de manifiesto que la conmutatividad no puede depender de una demostración formal.

Los matemáticos piensan —y están en lo cierto— que una teoría matemática totalmente rigurosa pasa por demostrar hechos tales como la conmutatividad de la suma. Ahora bien, estamos señalando que esta exigencia no tiene por objeto eliminar todo margen de duda sobre la conmutatividad de la suma —no hay tal duda—, sino garantizar que los axiomas son lo suficientemente correctos y potentes como para demostrar la conmutatividad de la suma. Si de los axiomas se dedujera que la suma no es conmutativa tendríamos que cambiarlos, y si esto no sucediera pero tampoco se dedujera la conmutatividad, tendríamos que añadir nuevos axiomas o sustituir los que tuviéramos por otros más potentes.

Por otra parte, debemos descartar la idea de que si definimos de un modo u otro los números naturales y conseguimos probar que los objetos que hemos definido cumplen los axiomas de Peano ya tenemos garantizado que estamos hablando realmente de los números naturales y que todo lo que probemos sobre los números naturales en nuestra teoría serán afirmaciones verdaderas sobre los números naturales. Los axiomas de Peano únicamente garantizan que nuestros "números naturales" cumplen las propiedades básicas de los números naturales, de modo que, según hemos visto en esta sección, ya no hace falta que nos molestemos en probar la conmutatividad de la suma o la divisibilidad euclídea: todos estos hechos son consecuencias de dichos axiomas y se cumplen automáticamente en cuanto los demostramos. Pero esto no excluye que en nuestra teoría podamos probar afirmaciones falsas sobre los números naturales.

Matemática y metamatemática A partir de la sección siguiente vamos a demostrar muchos resultados en los que argumentos metamatemáticos se combinan sutilmente con argumentos matemáticos (es decir, con demostraciones

formales en una teoría axiomática). El lector no familiarizado con la lógica deberá hacer aquí un gran esfuerzo de comprensión para asimilar esta relación tan delicada si es que quiere entender cabalmente los teoremas de incompletitud. Para ayudarle detallamos aquí la prueba de un teorema elemental:

Teorema 6.1 *Sea T una teoría aritmética. Para todo natural n se cumple*

$$\vdash_T \bigwedge x(\mathrm{Nat}\, x \to (x \leq 0^{(n)} \leftrightarrow x = 0^{(0)} \vee x = 0^{(1)} \vee \cdots \vee x = 0^{(n)})).$$

Demostración: Por inducción (metamatemática) sobre n, es decir, tenemos que probar que infinitas sentencias (una para cada n) son teoremas de T; empezamos por el caso $n = 0$ y después probaremos que si la sentencia n es demostrable la $n + 1$ también lo es. La prueba es constructiva, de modo que proporciona un algoritmo para generar explícitamente una demostración de cualquiera de las sentencias.

Para $n = 0$ hemos de probar

$$\vdash_T \bigwedge x(\mathrm{Nat}\, x \to (x \leq 0 \leftrightarrow x = 0)).$$

Esto se sigue[3] de los teoremas 17, 18 y 20.

Ahora lo suponemos cierto para n, es decir, suponemos que ya hemos escrito una demostración de la sentencia del enunciado (para n) y veremos cómo extenderla para llegar a la misma fórmula para $n+1$. Así pues, nuestro problema ahora es demostrar

$$\vdash_T \bigwedge x(\mathrm{Nat}\, x \to (x \leq 0^{(n+1)} \leftrightarrow x = 0^{(0)} \vee \cdots \vee x = 0^{(n)} \vee x = 0^{(n+1)})).$$

Procedemos, pues, a esbozar[4] una demostración. Suponemos $\mathrm{Nat}\, x$ y, para probar una implicación, suponemos[5] también $x \leq 0^{(n+1)}$. Esto es lo mismo que $x \leq 0^{(n)\prime}$. Por la definición de $<$ y el teorema 25, de aquí se sigue que $x \leq 0^{(n)} \vee x = 0^{(n+1)}$. Por hipótesis de inducción (metamatemática) disponemos de la sentencia del enunciado, que podemos usar en nuestra demostración (matemática). Al aplicarla al caso $x \leq 0^{(n)}$ la disyunción se transforma en

$$x = 0^{(0)} \vee \cdots \vee x = 0^{(n)} \vee x = 0^{(n+1)},$$

que es lo que queríamos probar.

[3]El lector debería completar los detalles tomando como modelo la segunda parte de la prueba.

[4]Los matemáticos nunca dan pruebas detalladas (ahora modus ponens, ahora modus tollens), sino esbozos de prueba, que contienen la información suficiente para cubrir todos los huecos.

[5]Si fuera $n = 5$, un matemático diría "supongamos que x es un número natural menor o igual que 6, pero nosotros precisamos que el "6" es en realidad el numeral $0^{(6)}$ (en general $0^{(n+1)}$) y evitamos decir que x es un número natural porque x no es un número natural, es una variable. Compárese con el novelista que escribe "... Napoleón le habló a un soldado", pero en realidad Napoleón (= 6) no ha hablado con ningún soldado (= número natural), simplemente el novelista ha escrito las palabras "Napoleón" ($= 0^{(6)}$) y "soldado" ($= x$) en una frase, todo esto sin perjuicio de que el novelista esté pensando en Napoleón y en un soldado.

Recíprocamente, si suponemos esta última fórmula, la fórmula de que disponemos por hipótesis de inducción[6] nos da que $x \leq 0^{(n)} \vee x = 0^{(n+1)}$ y, similarmente, se llega a $x \leq 0^{(n+1)}$. ∎

El lector debería reflexionar sobre este argumento y sobre todos los argumentos similares que veremos más adelante. En particular, debería tener claro que no hemos probado nada extraño y abstracto, sino algo tan simple como que la sentencia

$$\bigwedge x(\mathrm{Nat}\, x \to (x \leq 0''' \leftrightarrow x = 0 \vee x = 0' \vee x = 0'' \vee x = 0'''))$$

es demostrable en cualquier teoría aritmética, al igual que lo es cualquiera de las fórmulas análogas que resultan de cambiar el número de comitas y, por consiguiente, el número de disyunciones.

6.3 Expresabilidad y representabilidad

Los teoremas de incompletitud descansan sobre tres pilares básicos. Uno es la posibilidad de representar los conceptos lógicos a través de relaciones y funciones recursivas, gracias a la numeración de Gödel; el segundo lo probaremos en esta sección y el tercero en el capítulo siguiente. Respecto al que nos ocupa, para enunciarlo necesitamos introducir unos conceptos. Antes de dar la definición general probamos el teorema siguiente, que nos proporciona unos primeros ejemplos.

Teorema 6.2 *Sea T una teoría aritmética. Se cumplen los hechos siguientes:*

a) Para todo natural n, se cumple $\vdash_T \mathrm{Nat}\, 0^{(n)}$.

b) Si $m = n$ entonces $\vdash_T 0^{(m)} = 0^{(n)}$.

c) Si $m \neq n$, entonces $\vdash_T \neg 0^{(m)} = 0^{(n)}$.

d) Para todos los números naturales m y n se cumple

$$\vdash_T 0^{(m)} + 0^{(n)} = 0^{(m+n)}, \qquad \vdash_T 0^{(m)} \cdot 0^{(n)} = 0^{(mn)}.$$

Demostración: a) Se cumple $\vdash_T \mathrm{Nat}\, 0^{(0)}$ por N1. Supuesto $\vdash_T \mathrm{Nat}\, 0^{(n)}$ obtenemos $\vdash_T \mathrm{Nat}\, 0^{(n+1)}$ aplicando N2.

b) Si $m = n$ tenemos que $0^{(m)} \equiv 0^{(n)}$, luego $0^{(m)} = 0^{(n)}$ es un teorema lógico.

[6]Notemos que no sería exacto decir "la hipótesis de inducción", porque nuestra demostración matemática no usa ninguna inducción.

c) Si $m \neq n$, digamos que $m = n + r$. Por N3 obtenemos $\underset{T}{\vdash} \neg 0^{(r)} = 0^{(0)}$. De N4 se deduce $\bigwedge xy(\operatorname{Nat} x \wedge \operatorname{Nat} y \wedge \neg x = y \to \neg x' = y')$ y, aplicando esto sucesivamente vamos obteniendo los teoremas

$$\neg 0^{(r+1)} = 0^{(1)}, \qquad \neg 0^{(r+2)} = 0^{(2)}, \qquad \neg 0^{(r+3)} = 0^{(3)}, \quad \ldots$$

hasta llegar a $\neg 0^{(m)} = 0^{(n)}$.

d) Por inducción sobre n: Si $n = 0$ se cumple $\underset{T}{\vdash} 0^{(m)} + 0^{(0)} = 0^{(m+0)}$ por N5. Supuesto cierto para n, es decir, si $\underset{T}{\vdash} 0^{(m)} + 0^{(n)} = 0^{(m+n)}$, por N6 obtenemos $\underset{T}{\vdash} 0^{(m)} + 0^{(n)\prime} = (0^{(m)} + 0^{(n)})'$, y de aquí que $\underset{T}{\vdash} 0^{(m)} + 0^{(n)\prime} = 0^{(m+n)\prime}$. Esta sentencia es idéntica a $\underset{T}{\vdash} 0^{(m)} + 0^{(n+1)} = 0^{(m+n+1)}$.

Con el producto se razona de forma similar. ∎

Observaciones En la página 63 discutimos ya la primera parte del apartado d) para la aritmética de Peano. Lo que afirma el teorema anterior, por ejemplo en el apartado c), es que no sólo es cierto que $2 \neq 3$, como ya sabíamos, sino que en toda teoría aritmética se puede probar (la sentencia cuya interpretación natural es) que $2 \neq 3$. La prueba que damos es constructiva, pues de ella se desprende un algoritmo para general la prueba de cualquier par de números distintos son distintos. En realidad un ejemplo hubiera sido igual de convincente, en el sentido de que cualquiera que comprenda que lo que sigue es un esbozo de prueba de que $4 \neq 2$ comprende que igualmente se puede probar cualquier otra desigualdad similar:

$$0'' \neq 0 \ \text{ (por N3)}, \ \ 0''' \neq 0' \ \text{ (por N4)}, \ \ 0'''' \neq 0'' \ \text{ (por N4)}.$$

Insistimos en que las inducciones que aparecen en la prueba son todas metamatemáticas. Así, por ejemplo, en una demostración de $\operatorname{Nat} 0^{(n)}$ no interviene para nada el axioma N9, sino que se usa el axioma N1 una vez y el axioma N2 n veces. ∎

Tenemos así ejemplos de que (las sentencias que expresan) determinadas afirmaciones sobre números naturales, como las del tipo $m = n$, pueden demostrarse o refutarse en cualquier teoría aritmética según sean verdaderas o falsas. Los conceptos que introducimos a continuación recogen esta idea de forma general.

Definición 6.3 Una relación n-ádica (sobre números naturales) es *expresable* en una teoría aritmética T si existe una fórmula aritmética $\alpha(y_1, \ldots, y_n)$ cuyas variables libres sean a lo sumo $y_1, \ldots, y_n$ tal que para todos los naturales $a_1, \ldots, a_n$ se cumple

a) Si $R(a_1, \ldots, a_n)$ entonces $\underset{T}{\vdash} \alpha(0^{(a_1)}, \ldots, 0^{(a_n)})$,

b) Si no $R(a_1, \ldots, a_n)$ entonces $\underset{T}{\vdash} \neg\alpha(0^{(a_1)}, \ldots, 0^{(a_n)})$.

Una función n-ádica (sobre números naturales) es *representable* en T si la relación $n+1$-ádica dada por $R(a_1,\ldots,a_n,a_{n+1})$ syss $f(a_1,\ldots,a_n)=a_{n+1}$ es expresable en T por una fórmula $\alpha(y_1,\ldots,y_n,y_{n+1})$ de modo que para todos los números naturales $a_1,\ldots,a_n$ se cumpla

$$\underset{T}{\vdash} \overset{1}{\bigvee} y_{n+1}\ \alpha(0^{(a_1)},\ldots,0^{(a_n)},y_{n+1}).$$

Notemos que no exigimos que se cumpla

$$\underset{T}{\vdash} \bigwedge y_1\ldots y_n(\mathrm{Nat}\, y_1 \wedge \ldots \wedge \mathrm{Nat}\, y_n \to \overset{1}{\bigvee} y_{n+1}\ \alpha(y_1,\ldots,y_n,y_{n+1}),$$

que es una condición más fuerte.

En realidad, para probar que una función f es representable en T basta probar que existe una fórmula aritmética $\alpha(y_1,\ldots,y_n,y_{n+1})$ cuyas variables libres sean a lo sumo las indicadas y tal que

a) Si $f(a_1,\ldots,a_n)=a_{n+1}$ entonces $\underset{T}{\vdash} \alpha(0^{(a_1)},\ldots,0^{(a_n)},0^{(a_{n+1})})$,

b) Para todos los naturales $a_1,\ldots,a_n$ se cumple

$$\underset{T}{\vdash} \overset{1}{\bigvee} y_{n+1}\ \alpha(0^{(a_1)},\ldots,0^{(a_n)},y_{n+1}).$$

Según la definición, haría falta probar también que si $f(a_1,\ldots,a_n)\neq a_{n+1}$ entonces $\underset{T}{\vdash} \neg\alpha(0^{(a_1)},\ldots,0^{(a_n)},0^{(a_{n+1})})$, pero no hace falta porque se sigue del teorema 6.2. En efecto, si llamamos $b=f(a_1,\ldots,a_n)$, según el teorema 6.2 en T podemos probar la sentencia $\neg 0^{(a_{n+1})}=0^{(b)}$, y además tenemos

$$\alpha(0^{(a_1)},\ldots,0^{(a_n)},0^{(b)}) \qquad \text{y} \qquad \overset{1}{\bigvee} y_{n+1}\ \alpha(0^{(a_1)},\ldots,0^{(a_n)},y_{n+1}).$$

De estas sentencias es consecuencia lógica $\neg\alpha(0^{(a_1)},\ldots,0^{(a_n)},0^{(a_{n+1})})$.

Del teorema 6.2 se sigue inmediatamente que la relación $=$ es expresada en toda teoría aritmética por la fórmula $x=y$, así como que las funciones suma y producto están representadas por las fórmulas $x+y=z$ y $x\cdot y=z$, respectivamente.

Observaciones Es crucial distinguir entre las propiedades de los números naturales, expresables en términos de relaciones y funciones, y las sentencias de una determinada teoría aritmética. Así, por ejemplo, si nos preguntamos si es posible demostrar en la aritmética de Peano que el 5 es un número primo, antes debemos plantearnos qué sentencia concreta queremos demostrar, porque "5 es primo" no es, una sentencia de ningún lenguaje formal. Una cosa es definir la relación "ser primo" (es fácil hacerlo: un número natural n es primo si es mayor

que 1 y cuando $n = mr$ entonces $m = 1$ o $r = 1$) y otra cosa es definir una sentencia que "signifique" ser primo.Así, podemos definir la fórmula aritmética

$$\operatorname{Prim} x \equiv \operatorname{Nat} x \wedge x > 0^{(1)} \wedge \bigwedge uv(\operatorname{Nat} u \wedge \operatorname{Nat} v \wedge x = uv \to u = 0' \vee v = 0')$$

y decir que "significa" ser primo, entendiendo por esto que la interpretación natural de cualquier sentencia Prim $0^{(n)}$ es "n es primo". Ahora ya podemos concretar nuestro problema, a saber, si la sentencia Prim $0^{(5)}$ es un teorema de la aritmética de Peano. Ciertamente lo es, pero ahora no nos importa esto, sino el hecho de que antes de plantearnos si se puede probar o no que el número 5 satisface la relación "ser primo" hemos tenido que buscar una fórmula que se pueda interpretar como "ser primo".

Puede probarse que la fórmula Prim x expresa la relación "ser primo" en cualquier teoría aritmética T, es decir, que siempre que n es un número primo es posible demostrar en T la sentencia Prim $0^{(n)}$, y siempre que n no es un número es posible demostrar en T la sentencia $\neg$Prim $0^{(n)}$. Notemos que la afirmación "la interpretación natural de Prim es ser primo" es semántica (involucra modelos) mientras que la afirmación "la fórmula Prim expresa la relación ser primo" es sintáctica (tiene que ver únicamente con el cálculo deductivo).

Por otra parte, del mero hecho de que encontremos una fórmula α que "signifique" una relación dada no podemos concluir que la exprese en el sentido de la definición anterior. Más adelante veremos ejemplos. ∎

Ahora podemos enunciar el resultado que perseguimos: hemos de probar que toda función (relación) recursiva es representable (expresable) en toda teoría aritmética. Para ello necesitamos probar primero algunos casos particulares de este hecho general.

Teorema 6.4 *Sea T una teoría aritmética.*

a) La relación de orden de los números naturales está expresada en T por la fórmula $x \leq y \equiv \bigvee z(\operatorname{Nat} z \wedge x + z = y)$.

b) La relación de congruencia en los números naturales está expresada en T por la fórmula $x \equiv y \pmod z \equiv \bigvee r(\operatorname{Nat} r \wedge (x = y + rz \vee y = x + rz))$.

Demostración: a) Si $m \leq n$ sea r la diferencia: $m + r = n$. Entonces $\underset{F}{\vdash} 0^{(m)} + 0^{(r)} = 0^{(n)}$, de donde $\underset{F}{\vdash} 0^{(m)} \leq 0^{(n)}$.

Si no $m \leq n$, se cumple $n \leq m$ y no $n = m$, luego

$$\underset{F}{\vdash} 0^{(n)} \leq 0^{(m)} \wedge \neg 0^{(n)} = 0^{(m)}.$$

Usando el teorema 18 de la sección anterior se concluye que $\underset{T}{\vdash} \neg 0^{(m)} \leq 0^{(n)}$.

Si $m \equiv n \pmod r$, entonces existe un número natural u tal que $m = n + ur$ o $n = m + ur$. Por consiguiente

$$\underset{T}{\vdash} 0^{(m)} = 0^{(n)} + 0^{(ur)} \qquad \text{o} \qquad \underset{T}{\vdash} 0^{(n)} = 0^{(m)} + 0^{(ur)}.$$

Como también $\underset{T}{\vdash} 0^{(u)} \cdot 0^{(r)} = 0^{(ur)}$, queda

$$\underset{T}{\vdash} 0^{(m)} = 0^{(n)} + 0^{(u)} \cdot 0^{(r)} \qquad \text{o} \qquad \underset{T}{\vdash} 0^{(n)} = 0^{(m)} + 0^{(u)} \cdot 0^{(r)}.$$

En ambos casos se concluye $\underset{T}{\vdash} 0^{(m)} \equiv 0^{(n)}$ (mód $0^{(r)}$).

Si no $m \equiv n$ (mód r), sea $\bar{m}$ tal que $m \equiv \bar{m}$ (mód r) y $\bar{m} < r$, e igualmente sea $\bar{n}$ tal que $n \equiv \bar{n}$ (mód r) y $\bar{n} < r$. (Aquí estamos suponiendo que $r \neq 0$. El caso $r = 0$ se comprueba aparte fácilmente.) Entonces

$$\underset{T}{\vdash} 0^{(m)} \equiv 0^{(\bar{m})} \text{ (mód } 0^{(r)}) \wedge 0^{(\bar{m})} < 0^{(r)}, \quad \underset{T}{\vdash} 0^{(n)} \equiv 0^{(\bar{n})} \text{ (mód } 0^{(r)}) \wedge 0^{(\bar{n})} < 0^{(r)}.$$

Como $\bar{m} \neq \bar{n}$, tenemos también $\underset{T}{\vdash} \neg 0^{(\bar{m})} = 0^{(\bar{n})}$. Si a esto unimos que, por el teorema 30 de la sección anterior,

$$\underset{T}{\vdash} \overset{1}{\bigvee} z(\text{Nat}\, z \wedge z < 0^{(r)} \wedge 0^{(\bar{n})} \equiv z \text{ (mód } 0^{(r)})),$$

ahora es fácil probar que $\underset{T}{\vdash} \neg 0^{(\bar{n})} \equiv 0^{(\bar{m})}$ (mód $0^{(r)}$) y, en consecuencia, que $\underset{T}{\vdash} \neg 0^{(n)} \equiv 0^{(m)}$ (mód $0^{(r)}$). ∎

Finalmente, necesitaremos probar la representabilidad de una función adicional de carácter técnico:

Definición 6.5 La *función beta de Gödel* es la función triádica determinada por que $\beta(c, d, i)$ es el mínimo natural z tal que $z \equiv c$ (mód $1 + (i + 1)d$).

El interés de esta ingeniosa función se debe a que es capaz de codificar en términos aritméticos cualquier sucesión finita de números naturales, como muestra el teorema siguiente:

Teorema 6.6 *Sea q una función monádica y n un número natural. Entonces existen naturales c y d tales que, si $0 \leq i \leq n$, entonces $q(i) = \beta(c, d, i)$.*

Demostración: Sea s un número natural mayor que $q(0), \ldots, q(n), n$. Los números $1 + s!, 1 + 2s!, \ldots, 1 + (n + 1)s!$ son primos entre sí dos a dos, pues si dos de ellos, digamos $1 + is!$ y $1 + js!$, con $1 \leq i < j \leq n + 1$ tuvieran un factor primo p en común, entonces p dividiría a la diferencia: $p \mid (j - i)s!$, luego $p|j - i$ o bien $p|s!$, pero como $j - i \leq n \leq s$, se cumple que $(j - i) \mid s!$, luego en cualquier caso $p \mid s!$, de donde $p \mid is!$, pero también $p \mid (1 + is!)$, luego $p \mid 1$, contradicción.

Sea $d = s!$, con lo que los números $1 + (i + 1)d$ son primos entre sí para $0 \leq i \leq n$. Por el teorema chino del resto (ver el apéndice B) existe un número natural c tal que

$$c \equiv q(i) \text{ (mód } 1 + (i + 1)d) \qquad \text{para } 0 \leq i \leq n.$$

Como $q(i) \leq s$, se cumple que $q(i) < 1 + (i+1)s!$, o sea, $q(i) \leq (i+1)d$ y, en resumen, tenemos que $q(i) \equiv c \pmod{1 + (i+1)d}$ y $q(i) \leq (i+1)d$, es decir, $q(i) = \beta(c,d,i)$ para $0 \leq i \leq n$. ∎

En otras palabras, si fijamos c y d adecuadamente, la función beta se convierte en una función monádica cuyos primeros valores coinciden con cualquier sucesión finita prefijada. Esto hace que podremos reducir una afirmación que empiece por "existe una sucesión finita de números naturales tal que ..." a otra que empiece por "existen dos números naturales tales que ..."

Teorema 6.7 *Sea T una teoría aritmética. Entonces la función beta de Gödel está representada en T por la fórmula*

$$b(c,d,i,z) \equiv \big(\mathrm{Nat}\, z \wedge z \equiv c \pmod{0^{(1)} + (i + 0^{(1)})d} \wedge z \leq (i + 0^{(1)})d\big).$$

Demostración: Si $\beta(c,d,i) = z$, entonces $z \equiv c \pmod{1 + (i+1)d}$ y $z \leq (i+1)d$, luego

$$\underset{T}{\vdash} \mathrm{Nat}\, 0^{(z)} \wedge 0^{(z)} \equiv 0^{(c)} \pmod{0^{(1)} + (0^{(i)} + 0^{(1)})0^{(d)}} \wedge 0^{(z)} \leq 0^{(1)} + (0^{(i)} + 0^{(1)})0^{(d)}.$$

Así pues, $\underset{T}{\vdash} b(0^{(c)}, 0^{(d)}, 0^{(i)}, 0^{(z)})$.

Como consecuencia del teorema 30 de la sección anterior tenemos que

$$\underset{T}{\vdash} \bigwedge cdi(\mathrm{Nat}\, c \wedge \mathrm{Nat}\, d \wedge \mathrm{Nat}\, i \to \overset{1}{\bigvee} z\, b(c,d,i,z)). \tag{6.1}$$

De aquí que, dados c, d, i, se cumple $\underset{T}{\vdash} \overset{1}{\bigvee} z\, b(0^{(c)}, 0^{(d)}, 0^{(i)}, z)$. ∎

Notemos que en la prueba anterior las letras c, d, et. representan según el contexto a una variable o a un número natural. Lo hacemos así para simplificar la notación, pero es esencial que el lector sea consciente de cuál es el significado de cada letra en un momento dado. Estando advertido no hay posibilidad de confusión: si escribimos $\mathrm{Nat}\, c$ entonces c ha de ser una variable, mientras que si escribimos $0^{(c)}$ entonces ha de ser un número natural.

Ahora probamos el resultado principal:

Teorema 6.8 *Toda función recursiva es representable en toda teoría aritmética.*

Demostración: Sea T una teoría aritmética. Vamos a probar más de lo que afirma el enunciado. Vamos a construir explícitamente una fórmula $\phi(x_1, \ldots, x_n, x_{n+1})$ que represente a una función recursiva dada f de tal modo que la interpretación natural de la sentencia $\phi(0^{(a_1)}, \ldots, 0^{(a_n)}, 0^{(a_{n+1})})$ sea que $f(a_1, \ldots, a_n) = a_{n+1}$. Esto no es necesario para probar los teoremas de incompletitud, pero sí es conveniente tenerlo en cuenta para interpretarlos adecuadamente. De todos modos, hemos de ser conscientes de que la parte sintáctica del argumento (es decir, la representabilidad) se prueba independientemente de las consideraciones semánticas sobre la interpretación natural de las fórmulas.

Es inmediato comprobar que la función c está representada en T por la fórmula $\phi_c(x,y) \equiv y = 0$, la función s está representada por la función $\phi_s(x,y) \equiv y = x'$ y la función p_i^k está representada por $\phi_i^k(x_1,\ldots,x_k,x_{k+1}) \equiv x_i = x_{k+1}$. También es claro que la interpretación natural de $0^{(n)} = 0$ es $n = 0$, etc.

Para el caso de funciones recursivas arbitrarias razonaremos por inducción sobre el número de funciones necesarias para definirlas. Ya hemos probado que las funciones definibles en un paso son representables. Ahora basta probar que las funciones definidas por composición, recursión o minimización a partir de funciones representables son también representables.

Supongamos que

$$f(x_1,\ldots,x_n) = g(h_1(x_1,\ldots,x_n),\ldots,h_m(x_1,\ldots,x_n))$$

y que existen fórmulas aritméticas $\phi(x_1,\ldots,x_m,x)$, $\psi_i(x_1,\ldots,x_n,x)$ que representan a g y a las funciones h_i respectivamente. Entonces no ofrece ninguna dificultad[7] probar que la función f está representada por la fórmula aritmética

$$\alpha(x_1,\ldots,x_n,x) \equiv \bigvee u_1\cdots u_m(\mathrm{Nat}\,u_1 \wedge \cdots \wedge \mathrm{Nat}\,u_m \wedge \psi_1(x_1,\ldots,x_n,u_1) \wedge \cdots$$
$$\cdots \wedge \psi_m(x_1,\ldots,x_n,u_m) \wedge \phi(u_1,\ldots,u_m,x)).$$

Si suponemos (por hipótesis de inducción) que la interpretación natural de las sentencias $\psi_i(0^{(a_1)},\ldots,0^{(a_n)},0^{(a_{n+1})})$ es $h_i(a_1,\ldots,a_n) = a_{n+1}$ y la de $\phi(0^{(a_1)},\ldots,0^{(a_m)},0^{(a_{m+1})})$ es $g(a_1,\ldots,a_m) = a_{m+1}$ entonces la interpretación natural de $\alpha(0^{(a_1)},\ldots,0^{(a_n)},0^{(a_{n+1})})$ es "*existen números naturales* $b_1,\ldots,b_m$ *tales que* $b_i = h_i(a_1,\ldots,a_n)$ *y* $a_{n+1} = g(b_1,\ldots,b_m)$", y esto equivale a que $a_{n+1} = f(a_1,\ldots,a_n)$.

Supongamos ahora que

$$\begin{aligned} f(0,x_1,\ldots,x_n) &= g(x_1,\ldots,x_n),\\ f(u+1,x_1,\ldots,x_n) &= h(u,f(x_1,\ldots,x_n),x_1,\ldots,x_n),\end{aligned}$$

y que existen fórmulas aritméticas $\phi(x_1,\ldots,x_n,x)$ y $\psi(u,v,x_1,\ldots,x_n,x)$ que representan a g y h respectivamente.

Sea $b(c,d,i,z)$ la fórmula que representa a las función beta de Gödel. Consideremos la fórmula aritmética

$$\alpha(x,x_1,\ldots,x_n,y) \equiv \bigvee cd(\mathrm{Nat}\,c \wedge \mathrm{Nat}\,d \wedge \bigvee t(\mathrm{Nat}\,t \wedge b(c,d,0,t)$$
$$\wedge \phi(x_1,\ldots,x_n,t)) \wedge \bigwedge u(\mathrm{Nat}\,u \wedge u + 0^{(1)} \leq x \to \bigvee vw(\mathrm{Nat}\,v \wedge \mathrm{Nat}\,w$$
$$\wedge b(c,d,u,v) \wedge b(c,d,u+0^{(1)},w) \wedge \psi(u,v,x_1,\ldots,x_n,w))) \wedge b(c,d,x,y)).$$

La interpretación natural de una sentencia $\alpha(0^{(a)},0^{(a_1)},\ldots,0^{(a_n)},0^{(a_{n+1})})$ es la siguiente:

[7] Dado que la prueba completa del teorema es larga y rutinaria, preferimos detallar el caso siguiente, que es más delicado, y dejar éste al lector.

Existen dos números naturales c y d tales que si $t = \beta(c,d,0)$, entonces $t = g(a_1,\ldots,a_n)$ —o sea, $\beta(c,d,0) = g(a_1,\ldots,a_n)$—; para todo número natural $u \leq a$, existen dos números naturales v y w tales que $v = \beta(c,d,u)$ y $w = \beta(c,d,u+1)$ y $w = h(u,v,a_1,\ldots,a_n)$ —equivalentemente, $\beta(c,d,u+1) = h(u,\beta(c,d,u),a_1,\ldots a_n)$— y $a_{n+1} = \beta(c,d,a)$.

Esto a su vez equivale a que existen dos números naturales c y d tales que para todo $u \leq a$ se cumple $\beta(c,d,u) = f(u,a_1,\ldots,a_n)$ y $a_{n+1} = \beta(c,d,a)$. Por la propiedad de la función β esto sucede si y sólo si $a_{n+1} = f(a,a_1,\ldots,a_n)$.

Vamos a esbozar una demostración de que α representa a f. Por abreviar llamaremos

$$\gamma(x_1,\ldots,x_n,u,v,w,c,d) \equiv b(c,d,u,v) \wedge b(c,d,u+0^{(1)},w)$$

$$\wedge\, \psi(u,v,x_1,\ldots,x_n,w),$$

$$\delta(x,x_1,\ldots,x_n,c,d) \equiv \bigwedge u(\mathrm{Nat}\, u \wedge u + 0^{(1)} \leq x \to \bigvee vw(\mathrm{Nat}\, v \wedge \mathrm{Nat}\, w$$

$$\wedge\, \gamma(x_1,\ldots,x_n,u,v,w,c,d))$$

$$\epsilon(x_1,\ldots,x_n,c,d,t) \equiv b(c,d,0,t) \wedge \phi(x_1,\ldots,x_n,t),$$

$$\eta(x,x_1,\ldots,x_n,c,d) \equiv \bigvee t(\mathrm{Nat}\, t \wedge \epsilon(x_1,\ldots,x_n,c,d,t)) \wedge \delta(x,x_1,\ldots,x_n,c,d).$$

Con esta notación

$$\alpha(x,x_1,\ldots,x_n,y) \equiv \bigvee cd(\mathrm{Nat}\, c \wedge \mathrm{Nat}\, d \wedge \eta(x,x_1,\ldots,x_n,c,d) \wedge b(c,d,x,y)).$$

Sean $x,x_1,\ldots,x_n$ números naturales e $y = f(x,x_1,\ldots,x_n)$. Sea $q(i) = f(i,x_1,\ldots,x_n)$. Sean c y d números naturales tales que $\beta(c,d,i) = q(i)$ para $i = 0,\ldots,x$.

Vamos a probar $\eta(0^{(x)},0^{(x_1)},\ldots,0^{(x_n)},0^{(c)},0^{(d)})$.

Como $q(0) = \beta(c,d,0)$, se cumple $\underset{T}{\vdash} b(0^{(c)},0^{(d)},0,0^{(q(0))})$ y como $q(0) = g(x_1,\ldots,x_n)$ y ϕ representa a g, también $\underset{T}{\vdash} \phi(0^{(x_1)},\ldots,0^{(x_n)},0^{(q(0))})$. De aquí se sigue $\epsilon(0^{(x_1)},\ldots,0^{(x_n)},0^{(c)},0^{(d)},0^{(q(0))})$, luego la primera parte de η se cumple con $t = 0^{(q(0))}$. Falta probar $\delta(0^{(x)},0^{(x_1)},\ldots,0^{(x_n)},0^{(c)},0^{(d)})$. Notar que si $x = 0$ es trivial. Supongamos, pues, $x > 0$.

Tomemos un natural $u < x$. Entonces tenemos $q(u+1) = \beta(c,d,u+1)$ y $q(u) = \beta(c,d,u)$. Por consiguiente

$$\underset{T}{\vdash} b(0^{(c)},0^{(d)},0^{(u)}+0^{(1)},0^{(q(u+1))}) \quad \text{y} \quad \underset{T}{\vdash} b(0^{(c)},0^{(d)},0^{(u)},0^{(q(u))}).$$

Por otra parte $q(u+1) = f(u+1,x_1,\ldots,x_n) = h(u,q(u),x_1,\ldots,x_n)$ y, como ψ representa a h tenemos que

$$\underset{T}{\vdash} \psi(0^{(u)},0^{(q(u))},0^{(x_1)},\ldots,0^{(x_n)},0^{(q(u+1))}).$$

Con esto hemos probado $\gamma(0^{(x_1)},\dots,0^{(x_n)},0^{(u)},0^{(q(u))},0^{(q(u+1))},0^{(c)},0^{(d)})$. En particular

$$\underset{T}{\vdash}\bigvee vw(\operatorname{Nat} v\wedge\operatorname{Nat} w\wedge\gamma(0^{(x_1)},\dots,0^{(x_n)},0^{(u)},v,w,0^{(c)},0^{(d)})).$$

Esto vale para todo $u<x$. Por el teorema 6.1 tenemos que

$$\underset{T}{\vdash}\bigwedge u(\operatorname{Nat} u\wedge u+0^{(1)}\leq x\to u=0\vee u=0^{(1)}\vee\cdots\vee u=0^{(x-1)}).$$

Puesto que en cada uno de los casos tenemos probado

$$\bigvee vw(\operatorname{Nat} v\wedge\operatorname{Nat} w\wedge\gamma(0^{(x_1)},\dots,0^{(x_n)},u,v,w,0^{(c)},0^{(d)})),$$

podemos concluir $\delta(0^{(x)},0^{(x_1)},\dots,0^{(x_n)},0^{(c)},0^{(d)})$.

Con esto tenemos probado η. Por último, como $y=q(x)=\beta(c,d,x)$, tenemos $\underset{T}{\vdash}b(0^{(c)},0^{(d)},0^{(x)},0^{(y)})$, luego

$$\eta(0^{(x)},0^{(x_1)},\dots,0^{(x_n)},0^{(c)},0^{(d)})\wedge b(0^{(c)},0^{(d)},0^{(x)},0^{(y)}).$$

De aquí se sigue $\alpha(0^{(x)},0^{(x_1)},\dots,0^{(x_n)},0^{(y)})$ sin más que tomar $c=0^{(c)}$ y $d=0^{(d)}$. (Notar que la c de la izquierda es una variable y la de la derecha un número natural.)

Falta probar que $\underset{T}{\vdash}\overset{1}{\bigvee}y\alpha(0^{(x)},0^{(x_1)},\dots,0^{(x_n)},y)$. Para ello supongamos

$$\alpha(0^{(x)},0^{(x_1)},\dots,0^{(x_n)},y)\wedge\alpha(0^{(x)},0^{(x_1)},\dots,0^{(x_n)},\bar{y}).$$

De aquí

$$\operatorname{Nat} c\wedge\operatorname{Nat} d\wedge b(c,d,0^{(x)},y)\wedge\eta(0^{(x)},0^{(x_1)},\dots,0^{(x_n)},c,d),$$

$$\operatorname{Nat}\bar{c}\wedge\operatorname{Nat}\bar{d}\wedge b(c,d,0^{(x)},\bar{y})\wedge\eta(0^{(x)},0^{(x_1)},\dots,0^{(x_n)},\bar{c},\bar{d}),$$

luego

$$\operatorname{Nat} t\wedge\epsilon(0^{(x_1)},\dots,0^{(x_n)},c,d,t)\wedge\delta(0^{(x)},0^{(x_1)},\dots,0^{(x_n)},c,d),$$

$$\operatorname{Nat}\bar{t}\wedge\epsilon(0^{(x_1)},\dots,0^{(x_n)},\bar{c},\bar{d},\bar{t})\wedge\delta(0^{(x)},0^{(x_1)},\dots,0^{(x_n)},\bar{c},\bar{d}).$$

De ϵ sale $\phi(0^{(x_1)},\dots,0^{(x_n)},t)\wedge\phi(0^{(x_1)},\dots,0^{(x_n)},\bar{t})$ y, como ϕ representa a g, concluimos que $t=\bar{t}=0^{(g(0))}$. Por lo tanto de ϵ se sigue ahora que

$$b(c,d,0^{(0)},0^{(q(0))})\wedge b(\bar{c},\bar{d},0^{(0)},0^{(q(0))}).$$

Veamos por inducción (metamatemática) que si $u\leq x$ entonces

$$\underset{T}{\vdash}b(c,d,0^{(u)},0^{(q(u))})\wedge b(\bar{c},\bar{d},0^{(u)},0^{(q(u))}).$$

Acabamos de probarlo para $u = 0$. Si vale para $u < x$, como $0^{(u)}+0^{(1)} \leq 0^{(x)}$, a partir de δ obtenemos que existen v, w, $\bar{v}$, $\bar{w}$ tales que

$$\mathrm{Nat}\, v \wedge \mathrm{Nat}\, w \wedge \gamma(0^{(x_1)}, \ldots, 0^{(x_n)}, 0^{(u)}, v, w, c, d),$$

$$\mathrm{Nat}\, \bar{v} \wedge \mathrm{Nat}\, \bar{w} \wedge \gamma(0^{(x_1)}, \ldots, 0^{(x_n)}, 0^{(u)}, \bar{v}, \bar{w}, \bar{c}, \bar{d}),$$

y de γ se sigue que

$$b(c, d, 0^{(u)}, v) \wedge b(c, d, 0^{(u+1)}, w) \wedge b(\bar{c}, \bar{d}, 0^{(u)}, \bar{v}) \wedge b(\bar{c}, \bar{d}, 0^{(u+1)}, \bar{w}).$$

Por la hipótesis de inducción y (6.1) tenemos que $v = 0^{(q(u))} = \bar{v}$. De γ se sigue también

$$\psi(0^{(u)}, 0^{(q(u))}, 0^{(x_1)}, \ldots, 0^{(x_n)}, w) \wedge \psi(0^{(u)}, 0^{(q(u))}, 0^{(x_1)}, \ldots, 0^{(x_n)}, \bar{w}).$$

Usando que ψ representa a h llegamos a que $w = 0^{(q(u+1))} = \bar{w}$, luego

$$b(c, d, 0^{(u+1)}, 0^{(q(u+1))}) \wedge b(\bar{c}, \bar{d}, 0^{(u+1)}, 0^{(q(u+1))}).$$

En particular

$$\underset{T}{\vdash} b(c, d, 0^{(x)}, 0^{(q(x))}) \wedge b(\bar{c}, \bar{d}, 0^{(x)}, 0^{(q(x))}),$$

pero también tenemos $b(c, d, 0^{(x)}, y) \wedge b(\bar{c}, \bar{d}, 0^{(x)}, \bar{y})$, luego $y = 0^{(q(x))} = \bar{y}$.

La prueba se adapta fácilmente al caso $n = 0$.

Finalmente, supongamos que $f(x_1, \ldots, x_n) = \mu x\, g(x_1, \ldots, x_n, 0) = 0$, donde la función g cumple que para todos los naturales $x_1, \ldots, x_n$ existe un x tal que $g(x_1, \ldots, x_n, x) = 0$ y está representada en T por una fórmula aritmética $\phi(x_1, \ldots, x_n, x, y)$. Basta tomar

$$\alpha(x_1, \ldots, x_n, x) \equiv \phi(x_1, \ldots, x_n, x, 0) \wedge \bigwedge u(\mathrm{Nat}\, u \wedge u < x \rightarrow$$

$$\neg\phi(x_1, \ldots, x_n, u, 0)).$$

Es fácil comprobar que α representa a f, así como que su interpretación natural es la debida. ∎

Como consecuencia inmediata obtenemos:

Teorema 6.9 *Toda relación recursiva es expresable en toda teoría aritmética.*

Demostración: Sea T una teoría aritmética y sea $R(x_1, \ldots, x_n)$ una relación recursiva. Entonces su función característica χ_R es recursiva, luego existe una fórmula aritmética $\phi(x_1, \ldots, x_n, x)$ que representa a χ_R en T. Entonces la fórmula aritmética $\psi(x_1, \ldots, x_n) \equiv \phi(x_1, \ldots, x_n, 0)$ expresa a R. En efecto, si $R(a_1, \ldots, a_n)$ entonces $\chi_R(a_1, \ldots, a_n) = 0$, luego $\underset{T}{\vdash} \phi(0^{(a_1)}, \ldots, 0^{(a_n)}, 0)$, o sea, $\underset{T}{\vdash} \psi(0^{(a_1)}, \ldots, 0^{(a_n)})$.

Igualmente, si no $R(a_1, \ldots, a_n)$ entonces $\chi_R(a_1, \ldots, a_n) \neq 0$, luego se cumple $\underset{T}{\vdash} \neg\phi(0^{(a_1)}, \ldots, 0^{(a_n)}, 0)$, o sea, $\underset{T}{\vdash} \neg\psi(0^{(a_1)}, \ldots, 0^{(a_n)})$.

Observemos además que si, de acuerdo con la prueba del teorema anterior, la interpretación natural de una sentencia $\phi(0^{(a_1)}, \ldots, 0^{(a_n)}, 0^{(a_{n+1})})$ es $\chi_R(a_1, \ldots, a_n) = a_{n+1}$, entonces la interpretación natural de $\psi(0^{(a_1)}, \ldots, 0^{(a_n)})$ es $R(a_1, \ldots, a_n)$. ∎

En particular, podemos expresar y representar en toda teoría aritmética las relaciones y funciones que definimos en la sección 5.4. Usaremos la misma notación para referirnos tanto a las relaciones y funciones como a sus fórmulas asociadas por los teoremas anteriores. Así, por ejemplo, "Axl" significará igualmente la relación monádica "ser un axioma lógico" y la fórmula aritmética Axl x que lo expresa en una teoría aritmética dada.

Ejercicio: Probar que si una relación (función) es expresable (representable) en una teoría aritmética recursiva, entonces es recursiva.

Terminamos añadiendo una relación recursiva a la lista de la sección 5.4. Si T es una teoría aritmética, s es el número de Gödel del designador 0 y n es el número de Gödel del término x'_0, definimos

$$N(0) = s; \quad N(r+1) = \mathtt{S}_{17}^{N(r)} n.$$

(Recordemos que $17 = g(x_0)$.) Claramente, $N(r) = g(0^{(r)})$, y la función N es recursiva primitiva.

Capítulo VII

Incompletitud

En este capítulo demostraremos y discutiremos los resultados más importantes de la lógica moderna, gracias a los cuales conseguiremos una mejor comprensión del alcance y las limitaciones de nuestra propia capacidad de razonamiento. Más concretamente, nos permitirán perfilar lo que debemos entender por una fundamentación razonable de la matemática abstracta.

7.1 El primer teorema de incompletitud

Ya hemos comentado que los teoremas de incompletitud se basan fundamentalmente en tres ideas: la recursividad de una gran parte de los conceptos de la lógica formal (lo cual tiene sentido gracias a la numeración de Gödel), la posibilidad de expresar cualquier función recursiva en cualquier teoría aritmética y el primer resultado que vamos a probar aquí, un ingenioso teorema en el que se combinan estos dos hechos.

Recordemos que a través de la numeración de Gödel cualquier afirmación sobre una teoría axiomática puede expresarse como una afirmación sobre números naturales, la cual a su vez puede formalizarse en una teoría aritmética, en el sentido de que podemos encontrar una sentencia cuya interpretación natural sea la afirmación que estamos considerando.

Por ejemplo, consideremos una teoría axiomática recursiva T, y consideremos la afirmación "T es consistente". No estamos suponiendo que T es consistente, simplemente vamos a ver cómo puede expresarse esta afirmación (tanto si es verdadera como si es falsa) en términos aritméticos.

Es claro que "T es consistente" equivale a que la fórmula $x_0 \neq x_0$ no sea demostrable en T. En efecto, puesto que $x_0 = x_0$ sí es demostrable, si pudiéramos probar su negación tendríamos una contradicción y, recíprocamente, si una fórmula no es demostrable en T entonces T es consistente.

El número de Gödel de $x_0 \neq x_0$ es $g = 2^3 \cdot 3^{32} \cdot 5^{17} \cdot 7^{17}$ (recordemos que $x_0 \neq x_0 \equiv \neg{=}x_0x_0$). Por consiguiente, la consistencia de T equivale a que g no sea (el número de Gödel de) un teorema de T o, con la notación del capítulo V (ver la página 136), a que no existe ningún número natural n tal que $\mathrm{Dm}(n, g)$.

Definición 7.1 En las condiciones anteriores, sea T' una teoría aritmética y llamemos también Dm a la fórmula que expresa en T' la relación Dm (correspondiente a la teoría T). Definimos la sentencia aritmética

$$\operatorname{Consis} T \equiv \neg\bigvee x(\operatorname{Nat} x \wedge \operatorname{Dm}(x, 0^{(g)})).$$

De este modo, $\operatorname{Consis} T$ es una sentencia aritmética que es verdadera en su interpretación natural si y sólo si la teoría T es consistente. En otras palabras, mediante $\operatorname{Consis} T$ hemos reducido una propiedad lógica, como es la consistencia de una teoría axiomática, a una propiedad aritmética, más concretamente a la no existencia de un número natural con determinadas características aritméticas.

En los capítulos anteriores tenemos todo lo necesario para escribir explícitamente esta sentencia, si bien hay que advertir que su longitud es astronómica (la mera presencia del numeral $0^{(g)}$ —que no es más que una pequeña porción de la sentencia— supone ya más de 10^{42} signos).

Esto es sólo un ejemplo de cómo podemos hablar de una teoría axiomática recursiva T a través de sentencias de una teoría aritmética T'. En realidad esto es un caso particular de algo muy frecuente: los matemáticos (o los físicos) acostumbran a hablar de fluidos y ondas y ruletas a través de sentencias de la teoría de conjuntos.

El caso más interesante se da cuando tomamos una misma teoría aritmética recursiva T para hablar de ella misma (es decir, tomamos como T' la propia T). Entonces, las algunas sentencias de T pueden interpretarse como afirmaciones sobre T. En particular, algunas sentencias de T pueden interpretarse como afirmaciones sobre otras fórmulas de T. Por ejemplo, $\operatorname{Consis} T$ es una sentencia de T que habla sobre la fórmula $x_0 \neq x_0$ de T y, concretamente, dice de ella que no es demostrable en T (sin entrar en si esto es cierto o no).

El teorema que vamos a probar a continuación afirma que es posible encontrar una sentencia que hable concretamente de sí misma y, más aún, que diga de sí misma cualquier cosa prefijada.

Teorema 7.2 *Sea T una teoría aritmética recursiva y $\phi(x)$ una fórmula (aritmética) con x como única variable libre. Entonces existe una sentencia (aritmética) ψ tal que $\vdash_T \psi \leftrightarrow \phi(0^{(n)})$, donde $n = g(\psi)$.*

En otras palabras, dada cualquier propiedad $\phi(x)$, podemos encontrar una sentencia ψ que es equivalente a que su propio número de Gödel cumpla la propiedad ϕ. Si identificamos a las fórmulas con sus números de Gödel, podemos decir más gráficamente que existe una sentencia que afirma "yo cumplo ϕ".

En una primera aproximación, para probar el teorema podríamos tratar de encontrar un número natural n que hiciera que la sentencia $\phi(0^{(n)})$ tuviera número de Gödel n, con lo que bastaría tomar $\psi \equiv \phi(0^{(n)})$. Ahora bien, esto es inviable, pues la sentencia $\phi(0^{(n)})$, al contener el numeral $0^{(n)}$, ha de tener al menos $n+1$ signos, luego su número de Gödel ha de ser divisible al menos entre $n+1$ primos, lo que obliga a que sea mucho mayor que n. Así pues, no sólo

no podemos tomar como ψ a una sentencia $\phi(0^{(n)})$, sino que de hecho sabemos que el numeral $0^{(n)}$ no puede aparecer en ψ. La sentencia que buscamos ha de hablar de $0^{(n)}$ sin nombrarlo explícitamente.

Esto no es difícil de conseguir. Pensemos por ejemplo en la afirmación "100 es un número par". Si en una teoría aritmética definimos

$$\mathrm{Par}\, x \equiv \bigvee z(\mathrm{Nat}\, z \wedge x = 0^{(2)} \cdot z),$$

entonces podemos expresar nuestra afirmación como Par $0^{(100)}$. En el caso de la aritmética de Peano (donde podemos incluso eliminar $\mathrm{Nat}\, z$) obtenemos una sentencia de longitud 111. Ahora bien, otra forma de expresar lo mismo (concretamente en el caso de la aritmética de Peano) es

$$\bigvee xy(y = 0^{(10)} \wedge x = y \cdot y \wedge \mathrm{Par}\, x).$$

El lector puede comprobar que esta sentencia tiene exactamente 47 signos. En general, una forma alternativa de expresar que un número n tiene una propiedad (expresable mediante la fórmula $\phi(x)$) es mediante una sentencia de la forma

$$\bigvee x(\mathrm{Nat}\, x \wedge \theta(0^{(p)}, x) \wedge \phi(x)),$$

donde el número p y la fórmula $\theta(y, x)$ se escogen adecuadamente para que $\theta(0^{(p)}, x)$ equivalga a $x = 0^{(n)}$. En nuestro ejemplo, $p = 10$, $\theta(y, x) \equiv x = y \cdot y$. En otras palabras, el truco es describir n mediante una construcción a partir de p; no decir "n", sino "el número que se obtiene de p haciendo tal operación".

Ésta es la técnica que vamos a emplear para construir la sentencia ψ que pide el teorema. Será de la forma

$$\psi \equiv \bigvee x(\mathrm{Nat}\, x \wedge \theta(0^{(p)}, x) \wedge \phi(x)),$$

para cierto número p y cierta fórmula θ. Más concretamente, $\theta(y, x)$ ha de describir una construcción de un número x a partir de un número y, de modo que cuando apliquemos esta construcción a p obtengamos el número de Gödel de ψ. La demostración comienza definiendo una función recursiva adecuada f que nos da un número a partir de otro, y θ será la fórmula que representa a esta función en la teoría aritmética dada.

Demostración (de 7.2): No perdemos generalidad si suponemos que la variable x es distinta de x_0. Sea $f(u) = \mathtt{S}_{17}^{N(u)} u$, donde $\mathtt{S}$ es la función de sustitución definida en el capítulo V (ver la página 136) y N es la función definida al final del capítulo anterior. Se trata de una función recursiva, luego es representable en T por una fórmula aritmética $\theta(x_0, x)$. Así pues, si $f(u) = v$ se cumple

$$\underset{T}{\vdash} \theta(0^{(u)}, 0^{(v)}) \quad \text{y} \quad \underset{T}{\vdash} \overset{1}{\bigvee} x\, \theta(0^{(u)}, x).$$

Sea $\sigma(x_0) \equiv \bigvee x(\mathrm{Nat}\, x \wedge \theta(x_0, x) \wedge \phi(x))$ y llamemos $p = g(\sigma)$. Definimos

$$\psi \equiv \mathtt{S}_{x_0}^{0^{(p)}} \sigma \equiv \bigvee x(\mathrm{Nat}\, x \wedge \theta(0^{(p)}, x) \wedge \phi(x)).$$

Es claro que la sentencia ψ es aritmética si ϕ lo es. Notemos que

$$n = g(\psi) = g(\mathtt{S}_{x_0}^{0^{(p)}}\sigma) = \mathtt{S}_{17}^{N(p)}p = f(p).$$

Como θ representa a f, tenemos $\underset{T}{\vdash} \theta(0^{(p)}, 0^{(n)})$ y, por la unicidad,

$$\underset{T}{\vdash} \bigwedge x\big(\mathrm{Nat}\, x \to \big(\theta(0^{(p)}, x) \leftrightarrow x = 0^{(n)}\big)\big),$$

de donde obviamente se sigue

$$\underset{T}{\vdash} \psi \leftrightarrow \bigvee x(\mathrm{Nat}\, x \wedge x = 0^{(n)} \wedge \phi(x)),$$

o sea, $\underset{T}{\vdash} \psi \leftrightarrow \phi(0^{(n)})$. ∎

Para enunciar el teorema de incompletitud de Gödel necesitamos un concepto adicional:

Definición 7.3 Una teoría aritmética T es *ω-contradictoria* si existe una fórmula $\alpha(x)$ tal que para todo natural n se cumple $\underset{T}{\vdash} \alpha(0^{(n)})$ y además

$$\underset{T}{\vdash} \bigvee x(\mathrm{Nat}\, x \wedge \neg\alpha(x)).$$

En caso contrario T es *ω-consistente.*

Observaciones Notemos ante todo que alguien que se obstinara en identificar los números naturales con unos objetos definidos formalmente en el seno de una teoría axiomática T (por ejemplo, la teoría de conjuntos), sería incapaz de dar sentido a la afirmación "T es ω-consistente". La noción de ω-consistencia de una teoría aritmética T presupone que los números naturales son algo que existe previamente a T.

Tenemos una ω-contradicción cuando somos capaces de probar que cada número natural cumple una propiedad —donde el "cada" es metamatemático, es decir, sabemos probar que la cumple 0, y $0'$, y $0''$, etc.—, pero también podemos probar que no todo número natural la cumple —donde el "todo" es matemático, es decir, sabemos probar que $\neg\bigwedge x(\mathrm{Nat}\, x \to \alpha(x))$. Hemos de destacar que una ω-contradicción está constituida por infinitos teoremas (uno para cada número natural más la prueba de la existencia de un contraejemplo).

Obviamente, toda teoría aritmética contradictoria es ω-contradictorio (pues en ella podemos probar cualquier cosa, en particular una ω-contradicción), por lo que, recíprocamente, toda teoría aritmética ω-consistente es consistente. Ahora bien, pronto podremos justificar que existen teorías aritméticas a la vez consistentes y ω-contradictorias.

Ejercicio: Probar que si una teoría aritmética admite un modelo estándar entonces es ω-consistente

Teorema 7.4 (Teorema de incompletitud de Gödel) *Toda teoría aritmética recursiva ω-consistente es incompleta.*

DEMOSTRACIÓN: Sea $\phi(x) \equiv \bigwedge y(\mathrm{Nat}\, y \to \neg\, \mathrm{Dm}(y, x))$. Por el teorema 7.2 existe una sentencia aritmética G, de número de Gödel n, tal que

$$\underset{T}{\vdash} G \leftrightarrow \bigwedge x(\mathrm{Nat}\, x \to \neg\, \mathrm{Dm}(x, 0^{(n)})).$$

Así, G significa que ningún número natural codifica una demostración de la fórmula cuyo número de Gödel es n o, lo que es lo mismo, que ningún número natural demuestra a G. En definitiva: que G no es demostrable.

Supongamos que $\underset{T}{\vdash} G$. Sea q el número de Gödel de una demostración de G. Entonces $\mathrm{Dm}(q, n)$, luego $\underset{T}{\vdash} \mathrm{Dm}(0^{(q)}, 0^{(n)})$. Esto implica que

$$\underset{T}{\vdash} \bigvee x(\mathrm{Nat}\, x \wedge \mathrm{Dm}(x, 0^{(n)}))$$

y, por lo tanto, $\underset{T}{\vdash} \neg G$, de donde se sigue que T es contradictoria.

Recíprocamente, si suponemos que T es consistente (pero no necesariamente ω-consistente), podemos asegurar que no $\underset{T}{\vdash} G$.

Así pues, no existen demostraciones de G, luego ningún número natural q cumple $\mathrm{Dm}(q, n)$, luego para todo natural q se cumple $\underset{T}{\vdash} \neg\, \mathrm{Dm}(0^{(q)}, 0^{(n)})$. Si suponemos que T es ω-consistente no puede ocurrir $\underset{T}{\vdash} \bigvee x(\mathrm{Nat}\, x \wedge \mathrm{Dm}(x, 0^{(n)})))$, o sea, no $\underset{T}{\vdash} \neg G$.

En conclusión, la sentencia G no es demostrable ni refutable en T, que es, por consiguiente, incompleta. ∎

Observaciones El hecho más destacable de la demostración del teorema anterior es que es completamente constructiva: dada una teoría aritmética recursiva T, sabemos construir explícitamente una sentencia G con la propiedad de que tenemos un algoritmo que aplicado a una hipotética demostración de G nos produciría una demostración de $\neg G$. Así pues, podemos asegurar que si T es consistente entonces G no es demostrable en T. A su vez, si esto es así, podemos estar seguros de que, tomemos el número natural q que tomemos, no será el número de Gödel de una demostración de G, lo cual a su vez nos garantiza que podremos probar cualquiera de las infinitas sentencias $\underset{T}{\vdash} \mathrm{Dm}(0^{(q)}, 0^{(n)})$. Por otra parte, una demostración de $\neg G$ formaría, junto con estas sentencias, una ω-contradicción en T.

Destaquemos también que para garantizar que G no es demostrable basta con que T sea consistente. Esto tiene interés porque la interpretación natural de G es su propia indemostrabilidad, luego si T es una teoría aritmética consistente, la sentencia de Gödel de T es verdadera (en su interpretación natural) y no demostrable en T.

7.2 El segundo teorema de incompletitud

El teorema de incompletitud de Gödel ha dado pie a muchas falacias, en virtud de las cuales la mente humana no es susceptible de análisis lógico. El argumento general es que, dada cualquier teoría axiomática suficientemente rica (aritmética), el teorema de incompletitud nos permite conocer una sentencia verdadera pero que no es demostrable en la teoría en cuestión, es decir, que nosotros sabemos más de lo que puede contener cualquier teoría axiomática. Variantes de este argumento se han empleado también contra la inteligencia artificial, es decir, para argumentar que un ordenador nunca podrá pensar como un ser humano. Como veremos enseguida, todo esto no tiene ningún fundamento.

Ciertamente, estamos ante una paradoja: no habría problema en admitir la existencia de afirmaciones verdaderas sobre números naturales que no puedan ser demostradas en una teoría dada, pero algo muy distinto es que sepamos encontrarlas explícitamente, es decir, que podamos señalar sentencias concretas de las que sepamos demostrar que son verdaderas pero no demostrables.

Consideremos, por ejemplo, el caso de la aritmética de Peano. Sabemos que $\mathrm{Consis}\,\mathcal{P}$ es una sentencia verdadera sobre números naturales pero que no se deduce de los axiomas de Peano. Ahora bien, para probar el teorema de incompletitud, ¿hemos usado alguna propiedad extraña sobre los números naturales, algo que no se deduzca de los axiomas de Peano? La respuesta es negativa, pero entonces, ¿cómo hemos podido llegar a probar algo que no se deduce de los axiomas de Peano?

Esta paradoja desaparece en cuanto nos damos cuenta de que el teorema de incompletitud no dice que la sentencia de Gödel sea verdadera y no demostrable. Sólo dice que SI la teoría axiomática es recursiva y consistente, entonces G es verdadera y no demostrable. La recursividad es una propiedad muy fácil de comprobar y que satisface cualquier teoría razonable, así que la cuestión se reduce a que SI la teoría T es consistente, ENTONCES G no es demostrable. Notemos que el recíproco es trivialmente cierto.

Esto es lo que realmente hemos demostrado para una teoría aritmética recursiva T. Este hecho puede enunciarse fácilmente mediante una sentencia aritmética: el teorema de incompletitud para una teoría axiomática recursiva T afirma que la sentencia

$$\mathrm{Consis}\,T \leftrightarrow \neg\bigvee x(\mathrm{Nat}\,x \wedge \mathrm{Dm}(x, 0^{(n)})) \tag{7.1}$$

es verdadera en su interpretación natural (donde n es el número de Gödel de la sentencia G).

Tendríamos una auténtica paradoja si esto —que es lo que realmente hemos demostrado— no pudiera probarse a partir de los axiomas de Peano. En tal caso sí tendríamos que preguntarnos qué hemos usado sobre los números naturales que no se deduzca de los axiomas de Peano. Sin embargo, lo cierto es que el teorema de incompletitud, visto así como una afirmación puramente aritmética expresada por la sentencia anterior sí puede ser demostrado exclusivamente a

partir de los axiomas de Peano. En particular se puede demostrar[1] en T, es decir,

$$\underset{T}{\vdash}\big(\operatorname{Consis} T \leftrightarrow \neg\bigvee x(\operatorname{Nat} x \wedge \operatorname{Dm}(x, 0^{(n)}))\big).$$

Ahora bien, teniendo en cuenta que, por construcción de G,

$$\underset{T}{\vdash}\big(G \leftrightarrow \bigwedge x(\operatorname{Nat} x \to \neg\operatorname{Dm}(x, 0^{(n)}))\big),$$

concluimos que

$$\underset{T}{\vdash}\big(\operatorname{Consis} T \leftrightarrow G\big).$$

Consecuentemente, todo lo que sabemos sobre G lo podemos aplicar a la sentencia $\operatorname{Consis} T$. Esto nos lleva al teorema siguiente:

Teorema 7.5 (Segundo teorema de incompletitud de Gödel) *Consideremos una una teoría aritmética recursiva T. Entonces*

$$\underset{T}{\vdash}\operatorname{Consis} T \qquad \textit{syss} \qquad T \textit{ es contradictoria.}$$

Demostración: Una implicación es obvia, y si $\underset{T}{\vdash}\operatorname{Consis} T$, según lo que acabamos de obtener, en T también puede probarse la sentencia de Gödel, luego T ha de ser contradictoria. ∎

Así, mientras la sentencia de Gödel tenía una interpretación autorreferente, su forma equivalente $\operatorname{Consis} T$ tiene una interpretación mucho más simple: la consistencia de la teoría axiomática considerada. Así se ve más claramente que la paradoja que describíamos antes no lo es tal: no hemos probado que la sentencia $\operatorname{Consis} T$ es verdadera y no demostrable, sino que, si es verdadera (es decir, si T es consistente) entonces no es demostrable.

En algunos casos sencillos podemos asegurar que es verdadera. Por ejemplo, si $\mathcal{P}$ es la aritmética de Peano, la sentencia $\operatorname{Consis}\mathcal{P}$ es un ejemplo de afirmación verdadera sobre los números naturales y que no es demostrable a partir de los axiomas de Peano. Esto es posible porque la prueba de la consistencia de $\mathcal{P}$ (es decir, la observación de que el conjunto de los números naturales constituye un modelo de $\mathcal{P}$ junto con las definiciones y teoremas relativos a modelos) involucra esencialmente colecciones infinitas y relaciones y funciones sobre colecciones infinitas, y estos conceptos no pueden definirse en $\mathcal{P}$. Por el contrario, $\operatorname{Consis}\mathcal{P}$ puede demostrarse en cualquier teoría de conjuntos en la que pueda probarse la existencia de conjuntos infinitos.

De este modo, la aritmética de Peano es una teoría axiomática de la que sí está justificado decir que es más limitada que la mente humana. Nosotros

[1]No obstante, no es inmediato que así sea. Hay una dificultad técnica debida a que no es posible hablar de funciones recursivas en cualquier teoría aritmética (a lo sumo podemos hablar de sucesiones finitas a través de la función beta), por lo que no podemos formalizar directamente todos los razonamientos que hemos empleado hasta llegar al teorema de incompletitud. En el capítulo X discutiremos con más detalle el caso en que T es una teoría de conjuntos. El lector interesado encontrará la prueba general en [22]. Esencialmente, la idea es sustituir las funciones por fórmulas.

sabemos más sobre los números naturales de lo que puede probarse a partir de los axiomas de Peano. Concretamente, conocemos algunas afirmaciones cuya prueba requiere hablar de conjuntos infinitos.

Pasemos ahora al extremo opuesto: sea T una teoría axiomática de conjuntos. En el capítulo siguiente describiremos varias de ellas con detalle, pero aquí nos basta saber que una teoría de conjuntos es una teoría axiomática en la que se puede formalizar cualquier razonamiento matemático. Si fuera posible dar un argumento convincente de que T es consistente, no habría ninguna dificultad en convertirlo en una demostración matemática en T de la sentencia $\mathrm{Consis}\,T$ (exactamente igual que cualquier matemático sabe convertir en teoremas de T todos sus razonamientos válidos). El segundo teorema de incompletitud nos daría entonces que T es contradictoria. Más concretamente, nos permitiría construir explícitamente una contradicción en T. Con esto hemos probado algo muy importante:

> *Si la teoría de conjuntos es consistente, no existe ningún argumento que pueda convencernos de que así es.*

Equivalentemente, si T es una teoría de conjuntos, la sentencia $\mathrm{Consis}\,T$ es un ejemplo de una afirmación sobre números naturales tal que, si es verdadera, jamás conseguiremos demostrar que lo es. Ahora estamos ante una teoría axiomática "más potente" que la mente humana, en el sentido de que en ella pueden formalizarse todos los razonamientos que nosotros consideramos convincentes (los razonamientos metamatemáticos) y muchos razonamientos más sobre objetos extraños, como puedan ser conjuntos no numerables, de los que no sabríamos hablar consistentemente sin la guía de la teoría axiomática de conjuntos.

Observemos que no es difícil demostrar $\mathrm{Consis}\,T$ en una teoría adecuada. Por ejemplo, basta llamar T' a la teoría que resulta de añadirle a T el axioma $\mathrm{Consis}\,T$ y, ciertamente, en T' se puede probar la consistencia de la teoría de conjuntos, pero la prueba no nos convence de nada. En general, no hay ningún problema en que la consistencia de una teoría T pueda probarse en otra teoría más fuerte T'. Lo que afirma el segundo teorema de incompletitud es que T' ha de ser *necesariamente* más fuerte que T. Así, puesto que la teoría de conjuntos T es más fuerte que nuestra capacidad de razonamiento metamatemático, sucede que no existen razonamientos metamatemáticos que prueben la consistencia de T. Cualquier demostración de esta consistencia (como el caso trivial que acabamos de considerar) partirá necesariamente de algún principio cuya consistencia es, a su vez, dudosa.

Esto supone una seria limitación a la fundamentación de la matemática. El programa de fundamentación de Hilbert pedía una teoría axiomática de conjuntos cuya consistencia y completitud pudieran ser demostradas mediante técnicas metamatemáticas finitistas. Los teoremas de incompletitud muestran que este programa es irrealizable: la completitud es imposible y la consistencia es indemostrable. Esto no quiere decir que sea imposible fundamentar satisfactoriamente las matemáticas. En el capítulo siguiente veremos varias teorías

axiomáticas de conjuntos que cumplen este objetivo, es decir, proporcionan una noción precisa de lo que debemos entender por una demostración matemática rigurosa. Cualquiera de estas teorías constituye de hecho una fundamentación de la matemática en el sentido de que es la referencia que de hecho toman los matemáticos para precisar en qué consiste su trabajo.[2] Es cierto que no podemos probar que ninguna de estas teorías es aceptable (consistente), pero los matemáticos vienen trabajando en ellas casi un siglo sin que nadie haya encontrado ninguna contradicción. Si unimos a esto la imposibilidad teórica marcada por el segundo teorema de incompletitud, concluimos que no hay motivos para sospechar de que la teoría axiomática de conjuntos no sea todo lo sólida que parece ser. Por otra parte, la completitud que exigía Hilbert no es realmente necesaria para el trabajo del matemático. En ninguna rama del conocimiento se considera necesario tener una garantía de poder responder a cualquier pregunta. Es cierto que las matemáticas parecían ser la única ciencia donde se hubiera podido tener tal garantía, pero el primer teorema de incompletitud no ha hecho sino acercarla a otras ramas del saber, como la física o la biología. Sin duda es imposible saber exactamente cómo, cuándo y dónde apareció el primer organismo vivo sobre la Tierra, pero esto no quita para que podamos determinar con gran precisión el proceso que dio lugar a la aparición de la vida.

Incompletitud y aritmética no estándar Los teoremas de incompletitud nos permiten construir y estudiar más claramente modelos no estándar de la aritmética. En efecto, sea T una teoría aritmética recursiva y consistente. Llamemos $S(x) \equiv \mathrm{Dm}(x, 0^{(g)})$, donde g es el número de Gödel de la fórmula $x_0 \neq x_0$. De este modo,[3]

$$\mathrm{Consis}\, T \equiv \neg \bigvee x(\mathrm{Nat}\, x \wedge S(x)).$$

Ahora bien, puesto que $\mathrm{Consis}\, T$ no es demostrable en T, el teorema 3.12 nos da que la teoría aritmética T' que resulta de añadirle a T el axioma $\neg\, \mathrm{Consis}\, T$ o, equivalentemente,

$$\bigvee x(\mathrm{Nat}\, x \wedge S(x)),$$

es consistente. No es difícil demostrar en T' que

$$\overset{1}{\bigvee} x(\mathrm{Nat}\, x \wedge S(x) \wedge \bigwedge y(\mathrm{Nat}\, y \wedge y < x \to \neg S(y))),$$

[2]No hay que entender aquí que la teoría axiomática de conjuntos explique la naturaleza de las matemáticas. Éste es un problema mucho más amplio. La teoría de conjuntos se limita precisar un patrón de rigor suficiente para que el matemático pueda trabajar sin vacilaciones. No obstante, es posible considerar argumentos informales, por ejemplo de carácter geométrico, que merecen el mismo calificativo de "matemáticas" y que no pueden ser considerados teoremas formales.

[3]La fórmula $\neg S(x)$ representa a la relación recursiva "no ser el número de Gödel de una demostración de $x_0 \neq x_0$". Si T es la teoría de conjuntos, esta propiedad satisface lo que en la introducción (pág. 14) llamábamos "ser simpático": es una propiedad que podemos comprobar explícitamente si la cumple un número dado o no y, aunque no sólo tiene sentido, sino que además es razonable conjeturar que la poseen todos los números naturales, lo cierto es que no existe ningún argumento que pueda justificar este hecho. Es algo que —plausiblemente— cumplen todos los números naturales sin que exista ninguna razón para que lo cumplan.

es decir, que existe un mínimo número natural que cumple $S(x)$. (Se prueba por inducción que, para todo natural x, o bien ningún número $y \leq x$ cumple $S(y)$ o bien hay un mínimo $y \leq x$ que cumple $S(y)$.)

Esto nos permite definir

$$c \equiv x \mid (\mathrm{Nat}\, x \wedge S(x) \wedge \bigwedge y(\mathrm{Nat}\, y \wedge y < x \to \neg S(y))).$$

La interpretación natural de c es que se trata del mínimo número de Gödel de una demostración de que $x_0 \neq x_0$ en T. Como estamos suponiendo que T es consistente, c es una descripción impropia en su interpretación natural. Así, si convenimos que las descripciones impropias son denotadas por el cero, tenemos que la interpretación natural de c es el número 0. Sin embargo, en T' tenemos que c es una descripción propia, por lo que la regla de las descripciones propias nos da que $\underset{T'}{\vdash} S(c)$.

Por otra parte, puesto que T es consistente, ningún número n es el número de Gödel de la demostración de una contradicción en T. En particular no $\mathrm{Dm}(n, g)$, luego $\underset{T}{\vdash} \neg S(0^{(n)})$.

Así pues, en T' podemos probar que c es un número natural que cumple una propiedad que también sabemos probar que no cumple ni 0, ni 1, ni 2, etc. Tenemos así un ejemplo de teoría aritmética consistente ω-contradictoria. Si M es un modelo de T', entonces $M(c)$ es un número natural no estándar. A diferencia de lo que sucedía en el capítulo IV, ahora sabemos definir explícitamente números no estándar en una teoría T: un ejemplo es el mínimo número de Gödel de la demostración de una contradicción en T (o, más concretamente, de $x_0 \neq x_0$). Por supuesto, necesitamos postular su existencia con un axioma, pero no necesitamos introducir una constante no definida para referirnos a él.

Observemos que c no es el mínimo número no estándar (de hecho no existe tal mínimo). En efecto, puesto que podemos probar que $c \neq 0$, de aquí deducimos que existe un número d tal que $c = d'$. Es claro que d también es no estándar, en el sentido de que para todo número natural n sabemos probar que $d \neq 0^{(n)}$.

7.3 El teorema de Rosser

El teorema de incompletitud que hemos probado es esencialmente el mismo que Gödel demostró. Sin embargo, J.B. Rosser demostró más tarde que la hipótesis de ω-consistencia se puede suprimir. Hemos dado, pese a ello, la prueba original porque muestra más claramente las ideas involucradas y éstas bastan para llegar al segundo teorema de incompletitud, pero eliminar la hipótesis de ω-consistencia tiene un gran valor teórico y para ello basta considerar una sentencia levemente más compleja que la de Gödel:

Teorema 7.6 (Teorema de incompletitud (versión de Rosser)) *Toda teoría aritmética recursiva consistente es incompleta.*

Demostración: Sea T una teoría aritmética recursiva y consistente. Consideramos la fórmula

$$\phi(x) \equiv \bigwedge y(\mathrm{Nat}\, y \wedge \mathrm{Dm}(y, x) \to \bigvee z(\mathrm{Nat}\, z \wedge z \leq y \wedge \mathrm{Rf}(z, x))).$$

Por el teorema 7.2 existe una sentencia aritmética R tal que

$$\underset{T}{\vdash} R \leftrightarrow \bigwedge y(\mathrm{Nat}\, y \wedge \mathrm{Dm}(y, 0^{(n)}) \rightarrow \bigvee z(\mathrm{Nat}\, z \wedge z \leq y \wedge \mathrm{Rf}(z, 0^{(n)}))),$$

donde $n = g(R)$. Veamos que R es indecidible en T. Observemos que la interpretación natural de R es "si un número natural me demuestra, hay otro menor que me refuta".

Si $\underset{T}{\vdash} R$, sea q el número de Gödel de una demostración de R en T. Entonces $\mathrm{Dm}(q, n)$, luego $\underset{T}{\vdash} \mathrm{Dm}(0^{(q)}, 0^{(n)})$. Puesto que también tenemos $\underset{T}{\vdash} \phi(0^{(n)})$, en particular

$$\underset{T}{\vdash} \bigvee z(\mathrm{Nat}\, z \wedge z \leq 0^{(q)} \wedge \mathrm{Rf}(z, 0^{(n)})).$$

Por el teorema 6.1,

$$\underset{T}{\vdash} \bigwedge z(\mathrm{Nat}\, z \wedge z \leq 0^{(q)} \rightarrow z = 0 \vee z = 0^{(1)} \vee \cdots \vee z = 0^{(q)}),$$

luego

$$\underset{T}{\vdash} \mathrm{Rf}(0, 0^{(n)}) \vee \mathrm{Rf}(0^{(1)}, 0^{(n)}) \vee \cdots \vee \mathrm{Rf}(0^{(q)}, 0^{(n)}).$$

Ahora bien, si T es consistente, como estamos suponiendo que $\underset{T}{\vdash} R$, no puede ocurrir que $\underset{T}{\vdash} \neg R$, luego todo número natural r cumple $\underset{T}{\vdash} \neg\, \mathrm{Rf}(0^{(r)}, 0^{(n)})$. De aquí se sigue

$$\underset{T}{\vdash} \neg\, \mathrm{Rf}(0, 0^{(n)}) \wedge \neg\, \mathrm{Rf}(0^{(1)}, 0^{(n)}) \wedge \cdots \wedge \neg\, \mathrm{Rf}(0^{(q)}, 0^{(n)}),$$

con lo que tenemos una contradicción en T.

Esto prueba que R no es demostrable en T. Supongamos ahora que $\underset{T}{\vdash} \neg R$. Entonces

$$\underset{T}{\vdash} \bigvee y(\mathrm{Nat}\, y \wedge \mathrm{Dm}(y, 0^{(n)}) \wedge \neg \bigvee z(\mathrm{Nat}\, z \wedge z \leq y \wedge \mathrm{Rf}(z, 0^{(n)}))).$$

Sea q el número de Gödel de una demostración de $\neg R$. Entonces $\mathrm{Rf}(q, n)$, luego $\underset{T}{\vdash} \mathrm{Rf}(0^{(q)}, 0^{(n)})$. De estas dos afirmaciones se deduce

$$\underset{T}{\vdash} \bigvee y(\mathrm{Nat}\, y \wedge y \leq 0^{(q)} \wedge \mathrm{Dm}(y, 0^{(n)})).$$

Usando de nuevo el teorema 6.1 llegamos a que

$$\underset{T}{\vdash} \mathrm{Dm}(0, 0^{(n)}) \vee \mathrm{Dm}(0^{(1)}, 0^{(n)}) \vee \cdots \vee \mathrm{Dm}(0^{(q)}, 0^{(n)}).$$

Por otra parte, como no $\underset{T}{\vdash} R$, ningún número natural r cumple $\mathrm{Dm}(r, n)$, luego

$$\underset{T}{\vdash} \neg Dm(0, 0^{(n)}) \wedge \neg\, \mathrm{Dm}(0^{(1)}, 0^{(n)}) \wedge \cdots \wedge \neg\, \mathrm{Dm}(0^{(q)}, 0^{(n)}),$$

y resulta que T es contradictoria. ∎

Como consecuencia obtenemos que una teoría aritmética consistente T tiene infinitos modelos distintos, en el sentido de que, dados dos de ellos, existe una sentencia aritmética verdadera en uno y falsa en el otro. En efecto, si R es la sentencia de Rosser de T, podemos formal las teorías T_0 y T_1 que resultan de añadir como axioma R y $\neg R$ respectivamente. Ambas son consistentes. A su vez podemos considerar la sentencia de Rosser R_i de T_i y formar las teorías R_{ij}, para $j = 0, 1$, que resultan de añadir como axioma a R_i las sentencias R_i y $\neg R_i$. De este modo podemos ir formado teorías consistentes

$$T,\ \ T_0,\ \ T_1,\ \ T_{00},\ \ T_{01},\ \ T_{10},\ \ T_{11},\ \ T_{000},\ \ T_{001},\ \ \text{etc.}$$

Si consideramos modelos de las 2^n teorías de nivel n, tenemos que T admite al menos 2^n modelos distintos para cada número natural n. Por consiguiente T admite infinitos modelos distintos.

7.4 El teorema de Tarski

Es fácil ver que la hipótesis de recursividad en los teoremas de incompletitud es necesaria. Por ejemplo, podemos considerar la extensión T de la aritmética de Peano que resulta de tomar como axiomas todas las sentencias verdaderas en su interpretación natural. Claramente T es una teoría aritmética consistente (ω-consistente, de hecho) y completa. La única explicación de que no contradiga al teorema de incompletitud es que no sea recursiva. Así pues, podemos concluir que el conjunto de las afirmaciones verdaderas sobre números naturales no es recursivo o, dicho de otro modo, que no existe ningún algoritmo para determinar si una determinada afirmación sobre números naturales es verdadera o falsa. El teorema de Tarski es una versión más elaborada de este razonamiento.

Teorema 7.7 (Teorema de Tarski de indefinibilidad de la verdad) *Sea T una teoría aritmética recursiva y consistente. Sea M un modelo de T.*

a) No existe ninguna fórmula $V(x)$ con x como única variable libre y tal que para toda sentencia ϕ (con número de Gödel n) se cumpla

$$M \vDash \phi \quad \textit{syss} \quad M \vDash V(0^{(n)}).$$

b) En particular la relación monádica dada por $V(n)$ syss V es el número de Gödel de una sentencia verdadera en M no es expresable en T, y por lo tanto no es recursiva.

c) Tampoco puede existir una fórmula $V(x)$ tal que para toda sentencia ϕ (de número de Gödel n) se cumpla $\underset{T}{\vdash} \phi \leftrightarrow V(0^{(n)})$.

Demostración: Supongamos que existe la fórmula $V(x)$ según a). Entonces el teorema 7.2 nos da una sentencia τ tal que

$$\underset{T}{\vdash} \tau \leftrightarrow \neg V(0^{(n)}), \qquad \text{donde } n = g(\tau). \tag{7.2}$$

Notemos que τ significa "yo soy falsa".

Si $M \vDash \tau$, entonces $M \vDash V(0^{(n)})$, pero, por (7.2), tendremos también que $M \vDash \neg\tau$, lo cual es absurdo.

Si $M \vDash \neg\tau$, entonces $M \vDash \neg V(0^{(n)})$, luego por hipótesis $M \vDash \tau$, y tenemos de nuevo un imposible. Así pues, no existe tal V.

Las afirmaciones restantes son consecuencias inmediatas de la primera: Una fórmula que expresara la relación V descrita en b) cumpliría a), al igual que le sucedería a una fórmula que cumpliera c). ∎

Observaciones Vemos, pues, que la situación que comentábamos sobre la aritmética de Peano es mucho más general: el conjunto de (los números de Gödel de) las sentencias verdaderas en un modelo de una teoría aritmética recursiva no es recursivo. En realidad es fácil ver que no es necesario que la teoría aritmética sea recursiva, pues un modelo de una teoría aritmética cualquiera determina un modelo de la aritmética de peano, que sí es recursiva.

Loa apartados a) y c) son dos versiones (semántica y sintáctica respectivamente) de un mismo hecho de gran trascendencia. Su significado se comprende mejor en el contexto de la teoría de conjuntos, donde podemos enunciarlos prescindiendo de la numeración de Gödel. Observemos que si T es una teoría axiomática recursiva, metamatemáticamente, una sentencia como

$$\alpha \equiv \bigwedge xy(\mathrm{Nat}\, x \wedge \mathrm{Nat}\, y \to x + y = y + x)$$

es un "objeto" que incluso podemos identificar con un número natural. No obstante, desde el punto de vista de T no es un objeto, sino una afirmación. Esto quiere decir que no es, en principio, ninguno de los objetos de los que podemos hablar con el lenguaje de T, sino una de las afirmaciones que podemos hacer en T. La numeración de Gödel nos permite salvar en parte esta diferencia, es decir, nos permite identificar α con un objeto de T, a saber con el numeral $0^{(n)}$, donde n es el número de Gödel de α. Así, toda afirmación metamatemática sobre α tiene un correlato en T sobre $0^{(n)}$. Por ejemplo, el hecho de que α es una sentencia se traduce en que en T se puede demostrar $\mathrm{Sent}\, 0^{(n)}$, donde $\mathrm{Sent}\, x$ es la fórmula que expresa en T la relación recursiva "ser una sentencia".[4] Sin embargo no deja de ser cierto que α como sentencia de T y el designador $0^{(n)}$ son dos objetos distintos. Un modo de relacionarlos de forma natural sería definir el "significado" de un número natural, es decir, definir una fórmula $V(x)$ de modo que en T pudiéramos probar que

$$V(0^{(n)}) \leftrightarrow \bigwedge xy(\mathrm{Nat}\, x \wedge \mathrm{Nat}\, y \to x + y = y + x)$$

Esto es justo lo que el apartado c) del teorema de Tarski demuestra que es imposible: No podemos asignar a cada número natural mediante una fórmula

[4] El hecho de que cualquier afirmación sobre cualquier sentencia α tiene un correlato formal en T es cierto en general, pero entendiendo que dicho correlato no tiene por qué ser, como en este caso concreto, un teorema de T. En general tendremos únicamente una sentencia cuya interpretación natural será el hecho dado. Pensemos por ejemplo en "$\bigvee x\, x \neq x$ no se puede demostrar en T", cuyo correlato formal es $\mathrm{Consis}\, T$.

(es decir, mediante una definición en T) la sentencia que codifica, (si es que codifica alguna) o, al menos, una sentencia equivalente. El apartado a) nos dice que el problema no está en la incompletitud de T, sino que es más básico todavía: no es posible definir siquiera V de modo que las equivalencias como la anterior (para toda sentencia α) sean, no ya demostrables, sino verdaderas en un modelo predeterminado. En general, no hay ninguna restricción en cuanto a la formalización completa de la lógica de una teoría aritmética en ella misma, pero sí hay restricciones muy importantes a la hora de relacionar los hechos y conceptos metamatemáticos con sus correspondientes formalizaciones.

Respecto a la prueba del teorema de Tarski, hemos de destacar que se apoya esencialmente en la conocida paradoja de Epiménides, o paradoja del mentiroso. La versión clásica se remonta a Epiménides, que, para rebatir la fama de mentirosos que en la antigüedad tenían los cretenses afirmaba: "Todos los cretenses mienten"; ahora bien, Epiménides era cretense, por lo que cualquiera que le oyera tenía que admitir que, por lo menos, los cretenses dicen la verdad en algunas ocasiones, ya que si suponemos que los cretenses mienten siempre entonces la afirmación de Epiménides sería cierta, pero eso es contradictorio con que la afirme un cretense.

Depurando el argumento, supongamos que un día, a las 12:01, Juan afirma: "Todo lo que hoy ha dicho Juan entre las 12:00 y las 12:02 es falso", y no dice nada más en dicho intervalo. Esta afirmación sería sin duda verdadera o falsa si la hubiera pronunciado cualquiera que no fuera Juan, o incluso si la hubiera pronunciado Juan en otro momento. Pero cuando la pronuncia Juan a las 12:01 se vuelve contradictoria.

Es la misma contradicción a la que llegamos si negamos la conclusión del teorema de Tarski: si en una teoría aritmética T podemos definir la noción de "número de Gödel de una sentencia verdadera en un modelo dado", entonces podemos construir una sentencia que diga "yo soy falsa" y tenemos la paradoja. No obstante, hemos de insistir en que la prueba no es un sofisma, sino que, muy al contrario, es totalmente constructiva: si alguien pudiera definir una fórmula $V(x)$ que cumpla el apartado a) del teorema de Tarski, entonces sabríamos escribir una sentencia τ de la que podríamos probar tanto que es verdadera como que es falsa. Por consiguiente estamos seguros de que la fórmula V no existe.

7.5 Otros resultados afines

En esta sección incluimos un par de resultados adicionales relacionados con los teoremas de incompletitud o con las técnicas con que los hemos probado. En la sección anterior hemos visto que el conjunto de las sentencias de una teoría aritmética que son verdaderas en un modelo dado no es recursivo. El lector puede considerar que no es tan extraño que un concepto tan "delicado" como es el de sentencia verdadera se escape a nuestro control, pero en realidad sucede algo más espectacular:

Teorema 7.8 *El conjunto de los (números de Gödel de los) teoremas de una teoría aritmética recursiva consistente no es recursivo.*

DEMOSTRACIÓN: Sea R la función característica del conjunto indicado, es decir, $R(n)$ syss n es el número de Gödel de un teorema de la teoría en cuestión, llamémosla T. Basta probar que R no es recursiva. En caso contrario, R sería expresable en T por una fórmula $R(x)$. Por el teorema 7.2 existe una sentencia C tal que $\underset{T}{\vdash} C \leftrightarrow \neg R(0^{(n)})$, donde $n = g(C)$.

Veamos que $\underset{T}{\vdash} C$. En otro caso, es decir, si C no fuera un teorema de T, tendríamos no $R(n)$, luego $\underset{T}{\vdash} \neg R(0^{(n)})$ y así $\underset{T}{\vdash} C$, en contra de lo supuesto. Esto es absurdo, luego C es un teorema.

Consecuentemente se cumple $R(n)$, luego $\underset{T}{\vdash} R(0^{(n)})$, luego $\underset{T}{\vdash} \neg C$, y concluimos que T es contradictoria. ∎

Tenemos así un ejemplo muy simple —en el sentido de que es muy fácil de definir— de conjunto no recursivo (del cual se siguen inmediatamente ejemplos de funciones y relaciones no recursivas). No hay, pues, ningún criterio para decidir si una sentencia es o no un teorema de una teoría aritmética consistente (notemos que aquí no hace falta pedir que sea recursiva). Al menos, el conjunto de los teoremas de cualquier teoría axiomática recursiva es recursivamente numerable, es decir, podemos generar sucesivamente todos los teoremas sin más que enumerar sistemáticamente las demostraciones. De este modo, si una sentencia es un teorema siempre podemos saberlo en un tiempo finito: basta esperar a que aparezca en la sucesión de teoremas, pero si no es un teorema nos podemos quedar con la duda de si aparecerá más tarde o si no aparecerá nunca.

El teorema anterior proporciona también un ejemplo muy simple de una relación definida con toda precisión (ser un teorema) y que, pese a ello, no es representable en ninguna teoría aritmética.

Tanto en la prueba del teorema de incompletitud como en la prueba del teorema anterior hemos construido sentencias cuya interpretación natural equivale a su no demostrabilidad. En un caso hemos concluido que era verdadera (de hecho la sentencia ha resultado ser indecidible) y en el otro era falsa (la teoría ha resultado ser contradictoria, por lo que la sentencia sí que era demostrable). Por simple curiosidad, podemos preguntarnos qué sucede si aplicamos el teorema 7.2 para construir una sentencia que afirma su propia demostrabilidad. Específicamente, dada una teoría aritmética recursiva T, sabemos construir una sentencia H tal que

$$\underset{T}{\vdash}(H \leftrightarrow \bigvee x(\mathrm{Nat}\, x \wedge \mathrm{Dm}(x, 0^{(n)})), \qquad \text{donde } n = g(H).$$

Las sentencias con esta propiedad se llaman sentencias de Henkin y no es evidente en principio si son verdaderas o falsas o —lo que en este caso es lo mismo—, si son demostrables o no. La respuesta es que todas son verdaderas, pero la prueba se basa en el segundo teorema de incompletitud.

Teorema 7.9 (Teorema de Löb) *Sea T una teoría aritmética recursiva y H una sentencia con número de Gödel n. Entonces*

$$\underset{T}{\vdash} \bigvee x(\operatorname{Nat} x \wedge \operatorname{Dm}(x, 0^{(n)}) \to H \quad \textit{syss} \quad \underset{T}{\vdash} H.$$

Demostración: Una implicación es obvia. Sea

$$D \equiv \bigvee x(\operatorname{Nat} x \wedge \operatorname{Dm}(x, 0^{(n)})).$$

Sea T^* la extensión de T que resulta de añadirle el axioma $\neg H$. Si no $\underset{T}{\vdash} H$, entonces T^* es consistente. Formalizando este sencillo resultado obtenemos que $\underset{T^*}{\vdash} \neg D \to \operatorname{Consis} T^*$.

Por el segundo teorema de incompletitud no $\underset{T^*}{\vdash} \operatorname{Consis} T^*$, luego no $\underset{T^*}{\vdash} \neg D$, de donde no $\underset{T}{\vdash} \neg H \to \neg D$, es decir, no $\underset{T}{\vdash} D \to H$, como había que probar. ∎

7.6 El teorema de Church

Ya hemos visto que, no sólo el conjunto de las sentencias verdaderas en un modelo de una teoría aritmética recursiva no es recursivo, sino que tampoco lo es el conjunto de las sentencias demostrables en dicha teoría. Podríamos pensar que todo esto se debe a que los números naturales son, contrariamente a lo que parecían, unos objetos demasiado complicados. Sin embargo ahora probaremos que la situación es más espectacular todavía: el conjunto de los teoremas lógicos de casi cualquier lenguaje formal no es recursivo. La prueba que veremos aquí se debe a Gödel. Necesitamos un resultado técnico.

Teorema 7.10 *Sea R una relación monádica recursiva primitiva. Entonces existe una sentencia α de un lenguaje formal recursivo $\mathfrak{L}$ tal que α es satisfacible si y sólo si $\bigwedge x\, Rx$.*

Demostración: Sea f la función característica de R. Se cumple que f es recursiva primitiva y además $\bigwedge x(Rx \leftrightarrow f(x) = 0)$. Sea $f_1, \ldots, f_n$ la sucesión de funciones que definen a f. Podemos suponer que f_1 es la función s y que ninguna otra f_i es s.

Sea $\mathfrak{L}$ un lenguaje formal cuyos signos eventuales sean n funtores $F_1, \ldots, F_n$, de manera que si la función f_i es n-ádica, el funtor F_i sea n-ádico. Claramente $\mathfrak{L}$ es recursivo primitivo.

Para cada i entre 1 y n definimos una fórmula α_i de $\mathfrak{L}$ como sigue:

- $\alpha_1 \equiv \bigwedge x_1 \neg F_1 x_1 = x_0 \wedge \bigwedge x_1 x_2 (F_1 x_1 = F_1 x_2 \to x_1 = x_2)$,
- Si f_i es p_j^k definimos $\alpha_i \equiv \bigwedge x_1 \cdots x_k F_i x_1 \cdots x_k = x_j$,
- Si f_i es c definimos $\alpha_i \equiv \bigwedge x_1 F_i = x_0$.

- Si f_i está definida por composición a partir de f_r, $f_{i_1}, \ldots, f_{i_s}$, definimos

$$\alpha_i \equiv \bigwedge x_1 \cdots x_u F_i x_1 \cdots x_u = F_r F_{i_1} x_1 \cdots x_u \cdots F_{i_s} x_1 \cdots x_u.$$

- Si f_i está definida por recursión a partir de f_r y f_s definimos

$$\alpha_i \equiv \bigwedge x_1 \cdots x_u F_i x_0 x_1 \cdots x_u = F_r x_1 \cdots x_u$$

$$\wedge \bigwedge x_1 \cdots x_{u+1} F_i F_1 x_{u+1} x_1 \cdots x_u = F_s x_{u+1} F_i x_{u+1} x_1 \cdots x_u x_1 \cdots x_u,$$

  (donde f_r es u-ádica, f_s es $u+2$-ádica y f_i es $u+1$-ádica.)

- Si f_i es monádica y está definida por recursión a partir del natural p y de f_s, definimos

$$\alpha_i \equiv F_i x_0 = \overbrace{F_1 \cdots F_1}^{p \text{ veces}} x_0 \wedge \bigwedge x_1 F_i F_1 x_1 = F_s x_1 F_i x_1.$$

Definimos además $\alpha_{n+1} \equiv \bigwedge x_1 F_n x_1 = x_0$. Las fórmulas $\alpha_1, \ldots, \alpha_{n+1}$ tienen libre únicamente la variable x_0. Vamos a probar que la sentencia

$$\alpha \equiv \bigvee x_0 (\alpha_1 \wedge \cdots \wedge \alpha_{n+1})$$

cumple el teorema.

Es claro que si $\bigwedge x R x$ entonces α es satisfacible sin más que considerar el modelo M de $\mathfrak{L}$ cuyo universo es el conjunto de los números naturales y donde $M(F_i) = f_i$. El x_0 cuya existencia afirma α es el 0.

Supongamos ahora que α es verdadera en un modelo M de universo U. Sea $g_i = M(F_i)$. Existe una valoración v de $\mathfrak{L}$ en M tal que

$$M \vDash (\alpha_1 \wedge \cdots \wedge \alpha_{n+1})[v].$$

Sea $a = v(x_0)$. Como $M \vDash \alpha_1[v]$, los objetos $a, g_1(a), g_1(g_1(a)), \ldots$ son todos distintos. Llamémoslos $a^{(0)}, a^{(1)}, a^{(2)}, \ldots$ Sea D la colección de todos los $a^{(n)}$.

Una simple inducción sobre i prueba que si g_i es m-ádica y $a_1, \ldots, a_m$ están en D, entonces $g_i(a_1, \ldots, a_m)$ está en D. Llamemos h_i a la restricción de g_i a D. De nuevo por inducción se prueba que para todos los números naturales $k_1, \ldots k_n$ se cumple

$$h_i(a^{(k_1)}, \ldots, a^{(k_n)}) = a^{(f_i(k_1, \ldots k_n))}$$

(se define q_i como la función dada por $h_i(a^{(k_1)}, \ldots, a^{(k_n)}) = a^{(q_i(k_1, \ldots k_n))}$ y se comprueba por inducción que q_i es f_i.)

En particular tenemos que q_n es f_n, o sea f. Como $M \vDash \alpha_{n+1}[v]$, se cumple que, para todo natural m, $h_n(a^{(m)}) = a^{(0)}$, de donde $f(m) = 0$, lo que equivale a $\bigwedge x\, Rx$ ∎

Definición 7.11 Una teoría axiomática es *decidible* si existe un criterio que nos permite saber si cualquier fórmula dada es o no un teorema, o con más precisión, si el conjunto de (los números de Gödel de) sus teoremas es recursivo.

El teorema 7.8 prueba que ningún sistema aritmético recursivo y consistente es decidible (en realidad, si no es recursivo con mayor razón no es decidible). Ahora probamos que, para muchos lenguajes formales, $K_{\mathcal{L}}$ no es decidible.

Teorema 7.12 (Teorema de Church) *El problema de la decisión es insoluble para los sistemas de primer orden, es decir, no existe ningún criterio para distinguir las fórmulas consistentes de las contradictorias.*

Demostración: Supongamos que contamos con un criterio para distinguir las fórmulas consistentes de las contradictorias en todo lenguaje formal. Sea α una fórmula del lenguaje formal de la aritmética de Peano y sea $n = g(\alpha)$.

Sea Rx syss $\neg\,\mathrm{Dm}(x, n)$. Como $\mathcal{P}$ es recursiva primitiva también R es una relación recursiva primitiva. Por el teorema anterior existe una sentencia ψ que es satisfacible (consistente) syss $\bigwedge x \neg\,\mathrm{Dm}(x, n)$, es decir, syss α no es un teorema.

Si existiera el criterio supuesto también existiría un criterio para decidir cuándo una fórmula α de $\mathcal{P}$ es un teorema o no lo es, en contradicción con el teorema 7.8. ∎

Es de destacar que este teorema usa la tesis de Church-Turing en su demostración, no sólo en la interpretación de su enunciado, como ocurre en los demás que hemos visto. Viendo la demostración del teorema 7.10, el teorema de Church prueba realmente que el problema de la decisión es insoluble para formalismos con suficientes funtores. Es fácil comprobar que vale igualmente para formalismos con suficientes relatores y, afinando un poco más, es posible reducirlo al caso de un formalismo con un único relator diádico distinto del igualador, es decir, no existe un criterio para decidir si una fórmula en la que tan sólo aparezca un relator diádico aparte del igualador es consistente o no. Simplemente se trata de comprobar que los funtores pueden ser "codificados" utilizando un relator diádico y añadiendo premisas.

7.7 Ecuaciones diofánticas

Si T es una teoría aritmética recursiva consistente, el teorema de incompletitud nos dice que la sentencia $\mathrm{Consis}\,T$ no es demostrable en T, a pesar de que es verdadera. Se trata de una afirmación sobre números naturales que hemos descrito explícitamente, si bien su estructura es tan compleja que aunque la escribiéramos con todo detalle no podríamos ver más que una maraña intrincada de signos lógicos. En esta sección demostraremos que $\mathrm{Consis}\,T$ es equivalente a una sentencia con la estructura más simple que puede tener una sentencia aritmética sin caer en la trivialidad: la no existencia de solución de una ecuación diofántica.

En la teoría de números se llaman ecuaciones diofánticas a las ecuaciones polinómicas con coeficientes enteros de las que se buscan soluciones enteras. Por ejemplo, una solución de la ecuación diofántica $x^2 - 2y^2 = 7$ es $x = 3$, $y = 1$. Lo que vamos a probar es que (la interpretación natural de) $\mathrm{Consis}\, T$ equivale a una afirmación de la forma

$$\neg\bigvee x_1 \dots x_n \in \mathbb{Z}\, P(x_1, \dots, x_n) = 0, \tag{7.3}$$

donde $\mathbb{Z}$ representa al conjunto de los números enteros y P es un polinomio con coeficientes enteros. De este modo, si T es la aritmética de Peano tendremos una ecuación diofántica de la que sabemos probar que no tiene solución, pero tal que esto no puede probarse únicamente a partir de los axiomas de Peano. Si T es la teoría de conjuntos, obtenemos una ecuación diofántica que *presumiblemente* no tiene solución pero que, si así es, no existe ningún argumento que lo justifique.

Con esto resolvemos negativamente el llamado *décimo problema de Hilbert,* que pedía un método para decidir si una ecuación diofántica dada tiene o no solución y, en caso afirmativo, calcular explícitamente todas sus soluciones. Si bien hay muchas técnicas para resolver muchos tipos de ecuaciones diofánticas, vemos que algunas no pueden ser decididas por método alguno.

Gödel probó que las sentencias indecidibles que se obtienen de su teorema son equivalentes a sentencias similares a (7.3) pero con una sucesión de cuantificadores universales y existenciales alternados. Llegar a una sentencia tan simple como (7.3) no es trivial en absoluto. La prueba se debe principalmente a las aportaciones de Martin Davis, Julia Robinson y Yuri Matiyacevič.

Conviene observar que la referencia a números enteros puede evitarse. Por una parte, $P(x, y)$ es un polinomio con coeficientes enteros

$$\bigvee xy \in \mathbb{Z}\, P(x, y) = 0 \iff \bigvee xy \in \mathbb{N}\, P(x, y)P(-x, y)P(x, -y)P(-x, -y) = 0.$$

Es claro que lo mismo vale para polinomios con cualquier número de variables, luego una afirmación de tipo (7.3) equivale a otra de la forma

$$\neg\bigvee x_1, \dots, x_n \in \mathbb{N}\, Q(x_1, \dots, x_n) = 0.$$

Por otra parte, separando los monomios de Q con coeficiente positivo de los monomios con coeficiente negativo obtenemos dos polinomios R y S con coeficientes naturales tales que la sentencia anterior equivale a

$$\neg\bigvee x_1, \dots, x_n \in \mathbb{N}\, R(x_1, \dots, x_n) = S(x_1, \dots, x_n).$$

No obstante, resulta mucho más conveniente trabajar con números enteros. Todos los razonamientos que vamos a ver pueden considerarse como informales, es decir, afirmaciones a las que podemos dar un significado y comprobar que son verdaderas sin depender de ninguna teoría axiomática formal. Alternativamente, el lector puede considerarlas como teoremas de la teoría de conjuntos, pero con ello supedita su validez a la consistencia de dicha teoría.

Definición 7.13 Una relación n-ádica R (sobre los números naturales) es *diofántica* si existe un polinomio $P(x_1,\ldots,x_n,y_1,\ldots,y_m)$ con coeficientes enteros tal que para todos los naturales $x_1,\ldots,x_n$ se cumple

$$R(x_1,\ldots,x_n) \leftrightarrow \bigvee y_1\ldots y_m \in \mathbb{N}\, P(x_1,\ldots,x_n,y_1,\ldots,y_m) = 0.$$

Una función n-ádica f es *diofántica* si lo es la relación $f(x_1,\ldots,x_n) = y$.

Conviene probar que la cuantificación sobre números naturales equivale a la cuantificación sobre números enteros:

Teorema 7.14 *Una relación n-ádica R es diofántica si y sólo si existe un polinomio con coeficientes enteros $P(x_1,\ldots,x_n,y_1,\ldots,y_m)$ tal que para todos los naturales $x_1,\ldots,x_n$ se cumple*

$$R(x_1,\ldots,x_n) \leftrightarrow \bigvee y_1\ldots y_m \in \mathbb{Z}\, P(x_1,\ldots,x_n,y_1,\ldots,y_m) = 0.$$

Demostración: Una implicación la hemos esbozado antes: si R cumple esta condición, definimos

$$Q(x_1,\ldots,x_n,y_1,\ldots,y_m) = \prod_{(\epsilon_1,\ldots,\epsilon_m)} P(x_1,\ldots,x_n,\epsilon_1 y_1,\ldots,\epsilon_m y_m),$$

donde $(\epsilon_1,\ldots,\epsilon_m)$ recorre todos los valores posibles con $\epsilon_i = \pm 1$. Entonces

$$R(x_1,\ldots,x_n) \leftrightarrow \bigvee y_1\ldots y_m \in \mathbb{N}\, Q(x_1,\ldots,x_n,y_1,\ldots,y_m) = 0.$$

Para la otra implicación necesitamos un truco diferente. Lagrange demostró que todo número natural es suma de cuatro cuadrados (ver el apéndice B), por lo que si R viene dada por

$$R(x_1,\ldots,x_n) \leftrightarrow \bigvee y_1\ldots y_m \in \mathbb{N}\, P(x_1,\ldots,x_n,y_1,\ldots,y_m) = 0,$$

sólo hemos de definir

$$\begin{aligned} &Q(x_1,\ldots,x_n,p_1,q_1,r_1,s_1,\ldots,p_m,q_m,r_m,s_m)\\ &= P(x_1,\ldots,x_n,p_1^1+q_1^2+r_1^2+s_1^2,\ldots,p_m^1+q_m^2+r_m^2+s_m^2), \end{aligned}$$

con lo que $R(x_1,\ldots,x_n)$ es claramente equivalente a

$$\bigvee p_1q_1r_1s_1\ldots p_mq_mr_ms_m \in \mathbb{Z}\, Q(x_1,\ldots,x_n,p_1,\ldots,s_m) = 0.$$

■

El teorema central que hemos de probar es el siguiente:

Teorema 7.15 *Una función es diofántica si y sólo si es recursiva.*

Admitiendo este teorema podemos probar:

Teorema 7.16 *Sea R una relación n-ádica recursiva. Entonces la afirmación*

$$\bigvee x_1 \dots x_n \in \mathbb{N}\, R(x_1, \dots, x_n)$$

es equivalente a que una cierta ecuación diofántica tenga solución.

DEMOSTRACIÓN: La función característica χ_R es recursiva, luego diofántica. Por consiguiente existe un polinomio $P(x_1, \dots, x_n, x, y_1, \dots, y_m)$ con coeficientes enteros tal que

$$\bigvee x_1 \dots x_n \in \mathbb{N}\, R(x_1, \dots, x_n) \leftrightarrow \bigvee x_1 \dots x_n \in \mathbb{N}\, \chi_R(x_1, \dots, x_n) = 0$$

$$\leftrightarrow \bigvee x_1 \dots x_n \in \mathbb{N} \bigvee y_1 \dots y_n \in \mathbb{Z}\, P(x_1, \dots, x_n, 0, y_1, \dots, y_m) = 0.$$

Usando que todo número natural es suma de cuatro cuadrados igual que en el teorema 7.14 llegamos a que esto equivale a su vez a

$$\bigvee z_1 \dots z_r \in \mathbb{Z}\, Q(z_1, \dots, z_r) = 0,$$

para cierto polinomio Q. ∎

Esto nos lleva a la conclusión que buscábamos, pues la consistencia de una teoría axiomática recursiva T equivale a $\neg\bigvee x\, R(x)$, donde $R(x)$ es la relación "x es el número de Gödel de la demostración de una contradicción en T".

A título de curiosidad notemos que los polinomios de las ecuaciones que obtenemos se pueden tomar de grado 4, pues cada monomio xy puede sustituirse por una nueva variable z añadiendo la ecuación $z = xy$. De este modo obtenemos muchas ecuaciones de grado 2 y, al sumar sus cuadrados (ver el teorema siguiente), queda un polinomio de grado 4. Más complejo es demostrar que, aumentando el grado, el número de variables puede reducirse a un máximo de 14.

Algunos hechos elementales Comenzamos probando un par de hechos sencillos que usaremos en todo momento. El primero es que un sistema de ecuaciones diofánticas equivale en realidad a una única ecuación:

Teorema 7.17 *Sean $P_i(x_1, \dots, x_n, y_1, \dots, y_m)$ polinomios con coeficientes enteros para $i = 1, \dots, r$. Entonces la relación dada por*

$$R(x_1, \dots, x_n) \leftrightarrow \bigvee y_1 \dots y_m \in \mathbb{N}(P_1 = 0 \wedge \dots \wedge P_r = 0)$$

es diofántica.

DEMOSTRACIÓN: Basta observar que

$$R(x_1, \dots, x_n) \leftrightarrow \bigvee y_1 \dots y_m \in \mathbb{N}\, P_1^1 + \dots + P_r^2 = 0.$$

∎

Por otra parte, tenemos unos pocos procedimientos generales para construir unas relaciones diofánticas a partir de otras:

Teorema 7.18 *Si R y S son relaciones diofánticas, también lo son las relaciones $R \wedge S$, $R \vee S$ y $\bigvee x\, R$.*

DEMOSTRACIÓN: Supongamos

$$\begin{aligned} R(x_1,\dots,x_n) &\leftrightarrow \bigvee y_1\cdots y_m \in \mathbb{N}\, P(x_1,\dots,x_n,y_1,\dots,y_m) = 0,\\ S(x_1,\dots,x_n) &\leftrightarrow \bigvee z_1\cdots z_r \in \mathbb{N}\, Q(x_1,\dots,x_n,z_1,\dots,z_r) = 0. \end{aligned}$$

Entonces

$$\begin{aligned} (R\wedge S)(x_1,\dots,x_n) &\leftrightarrow \bigvee y_1\cdots y_m z_1,\dots,z_r \in \mathbb{N}\, P^2+Q^2 = 0,\\ (R\vee S)(x_1,\dots,x_n) &\leftrightarrow \bigvee y_1\cdots y_m z_1,\dots,z_r \in \mathbb{N}\, PQ = 0. \end{aligned}$$

Si $R(x,x_1,\dots,x_n) \leftrightarrow \bigvee y_1\cdots y_m \in \mathbb{N}\, P(x,x_1,\dots,x_n,y_1,\dots,y_m) = 0$, entonces

$$\bigvee x\, R(x,x_1,\dots,x_n) \leftrightarrow \bigvee x y_1\cdots y_m \in \mathbb{N}\, P(x,x_1,\dots,x_n,y_1,\dots,y_m) = 0.$$

■

Las funciones diofánticas son recursivas La parte fácil del teorema 7.15 consiste en probar que las funciones diofánticas son recursivas. Empezamos observando que la función

$$T(n) = 1 + 2 + \cdots + n = \frac{n(n+1)}{2}$$

es diofántica, pues $y = T(x) \leftrightarrow 2y - x(x+1) = 0$.

Con ayuda de esta función vamos a construir una biyección diofántica entre $\mathbb{N}\times\mathbb{N}$ y $\mathbb{N}$. Más precisamente, construiremos una aplicación biyectiva diofántica $P : \mathbb{N}\times\mathbb{N} \longrightarrow \mathbb{N}$ cuya inversa está determinada a su vez por dos funciones diofánticas I, D.

Teorema 7.19 *Existen funciones diofánticas $P(x,y)$, $I(z)$, $D(z)$ tales que*

$$\bigwedge xy \in \mathbb{N}(I(P(x,y)) = x \wedge D(P(x,y)) = y),$$

$$\bigwedge z \in \mathbb{N}(P(I(z),D(z)) = z \wedge I(z) \le z \wedge D(z) \le z).$$

DEMOSTRACIÓN: La función T que acabamos de definir es creciente, luego para todo natural z existe un único n tal que $T(n) \le z < T(n+1)$, luego existe un único $y \le n$ tal que $z = T(n) + y$. Equivalentemente, existen unos únicos x, y tales que $z = T(x+y) + y$. Definimos $I(z) = x$, $D(z) = y$. Sea $P(x,y) = T(x+y) + y$.

Claramente, P, I, D cumplen lo pedido. Sólo falta ver que son diofánticas, pero

$$z = P(x,y) \leftrightarrow z = \frac{(x+y)+(x+y+1)}{2} + y$$

$$\leftrightarrow 2z - (x+y)(x+y+1) - 2y = 0,$$
$$x = I(z) \leftrightarrow \bigvee y \in \mathbb{N}\big(2z - (x+y)(x+y+1) - 2y = 0\big),$$
$$y = D(z) \leftrightarrow \bigvee x \in \mathbb{N}\big(2z - (x+y)(x+y+1) - 2y = 0\big).$$

■

La función siguiente es el equivalente diofántico de la función β de Gödel:

Teorema 7.20 *Existe una función diofántica $S(i,u)$ tal que $S(i,u) \leq u$ y para toda sucesión finita de números naturales $a_0, \ldots, a_n$ existe un $u \in \mathbb{N}$ tal que $S(i,u) = a_i$ para $i = 0, \ldots, n$.*

DEMOSTRACIÓN: Sea $S(i,u)$ el único natural w tal que

$$w \equiv I(u) \text{ (mód } (1 + (i+1)(D(u)+1))) \wedge w \leq (i+1)(D(u)+1).$$

La función S es diofántica porque $w = S(i,u)$ equivale a que las siguientes ecuaciones tengan solución natural en x, y, z, v:

$2u = (x+y)(x+y+1) + 2y$,
$x = w + z(1 + (i+1)(y+1))$,
$(i+1)(y+1) = w + v$.

En efecto, la primera equivale a que $x = I(u) \wedge y = D(u)$, la segunda a que $w \equiv I(u)$ (mód $1 + (i+1)(D(u)+1)$) y la tercera a que $w \leq (i+1)(D(u)+1)$. Notemos que $S(i,u) \leq I(u) \leq u$.

Dados $a_0, \ldots, a_n$, sea r mayor que todos ellos y que n. Sea $y = r!$. Los números $1+y, \ldots, 1+(n+1)y$ son primos entre sí (ver la prueba de 6.6), luego el teorema chino del resto nos da un natural x tal que $x \equiv a_i$ (mód $1+(i+1)y$).

Sea $u = P(x, y-1)$. Así $x = I(u) \wedge y = D(u) + 1$ y para $i = 0, \ldots, n$ tenemos que

$$a_i \equiv I(u) \text{ (mód } 1 + (i+1)(D(u)+1)) \wedge a_i < y = D(u) < (i+1)(D(u)+1),$$

luego se cumple que $a_i = S(i,u)$. ■

Es inmediato comprobar que la función S es recursiva. Consideremos ahora una función diofántica f y vamos a ver que es recursiva. En efecto, existen polinomios P y Q con coeficientes naturales tales que $y = f(x_1, \ldots, x_n)$ equivale a

$$\bigvee y_1 \ldots y_m \in \mathbb{N}\, P(x_1, \ldots, x_n, y, y_1, \ldots, y_m) = Q(x_1, \ldots, x_n, y, y_1, \ldots, y_m).$$

Entonces la función

$$g(x_1, \ldots, x_n) = \mu u \big(P(x_1, \ldots, x_n, S(0,u), S(1,u), \ldots, S(m,u))$$
$$= Q(x_1, \ldots, x_n, S(0,u), S(1,u), \ldots, S(m,u))\big)$$

es recursiva. (Notemos que siempre existe un u que cumple esta condición o de lo contrario $f(x_1, \ldots, x_n)$ no estaría definido.) Claramente,

$$f(x_1, \ldots, x_n) = S(0, g(x_1, \ldots, x_n)),$$

y esta expresión prueba que f es recursiva.

Algunas funciones diofánticas Para probar que toda función recursiva es diofántica necesitamos justificar el carácter diofántico de varias funciones. El punto más difícil es el teorema siguiente, cuya prueba dejamos para el final:

Teorema 7.21 (Matiyasevič) *La función x^y es diofántica.*

Este teorema lo probó Matiyasevič en 1970. Los resultados que veremos a continuación los obtuvo Julia Robinson en 1952 tomando como conjetura el teorema anterior.[5]

En primer lugar demostramos que los números combinatorios son diofánticos, para lo cual nos apoyamos en el teorema siguiente:

Teorema 7.22 *Si $0 \le k \le n$ y $u > 2^n$, entonces*

$$E\left[\frac{(u+1)^n}{u^k}\right] \equiv \binom{n}{k} \text{ (mód } u).$$

Demostración: La E representa, naturalmente, la parte entera. Por la fórmula del binomio tenemos que

$$\frac{(u+1)^n}{u^k} = \sum_{i=0}^{n}\binom{n}{i}u^{i-k} = \sum_{i=k}^{n}\binom{n}{i}u^{i-k} + \sum_{i=0}^{k-1}\binom{n}{i}u^{i-k}.$$

Ahora bien,

$$\sum_{i=0}^{k-1}\binom{n}{i}u^{i-k} < \frac{1}{u}\sum_{i=0}^{k-1}\binom{n}{i} < \frac{1}{u}\sum_{i=0}^{n}\binom{n}{i} = \frac{2^n}{u} < 1.$$

Por consiguiente

$$\sum_{i=k}^{n}\binom{n}{i}u^{i-k} \le \frac{(u+1)^n}{u^k} < \sum_{i=k}^{n}\binom{n}{i}u^{i-k} + 1,$$

es decir,

$$E\left[\frac{(u+1)^n}{u^k}\right] = \sum_{i=k}^{n}\binom{n}{i}u^{i-k} = \binom{n}{k} + u\sum_{i=k+1}^{n}\binom{n}{i}u^{i-k-1} \equiv \binom{n}{u} \text{ (mód } u).$$

■

Teorema 7.23 *La relación $z = \binom{n}{k}$ es diofántica.*

Demostración: Observemos que $\binom{n}{k} \le \sum_{i=0}^{n}\binom{n}{i} = 2^n$. Por el teorema anterior, para cualquier $u > 2^n$ tenemos que $\binom{n}{k}$ es el único número natural congruente con $E\left[\frac{(u+1)^n}{u^k}\right]$ módulo u y menor que u. Por lo tanto

$$z = \binom{n}{k} \leftrightarrow \bigvee uvv \in \mathbb{N}(v = 2^n \wedge u > v$$

[5]En realidad Robinson definió las funciones exponencial-diofánticas como las definibles en términos de ecuaciones diofánticas y de la función exponencial y probó que las funciones que vamos a estudiar ahora eran exponencial-diofánticas.

$$\wedge\, w = E\left[\frac{(u+1)^n}{u^k}\right] \wedge z \equiv w \ (\text{mód } u) \wedge z < u).$$

Por el teorema 7.18, basta probar que las relaciones que aparecen dentro del paréntesis son diofánticas. Ahora bien,

$v = 2^n \leftrightarrow \bigvee x \in \mathbb{N}(x = 2 \wedge v = x^n)$ es diofántica por 7.21.

$u > v \leftrightarrow \bigvee x \in \mathbb{N}\, u = v + x + 1$, diofántica.

$w = E\left[\frac{(u+1)^n}{u^k}\right]$ equivale a

$$\bigvee xyt \in \mathbb{N}(t = u + 1 \wedge x = t^n \wedge y = u^k \wedge w \leq (x/y) < w + 1)$$

y también a

$$\bigvee xyt \in \mathbb{N}(t = u + 1 \wedge x = t^n \wedge y = u^k \wedge wy \leq x < (w + 1)y),$$

claramente diofántica.

$z \equiv w \ (\text{mód } u) \wedge z < u \leftrightarrow \bigvee xy \in \mathbb{N}(w = z + xu \wedge u = z + y + 1)$ diofántica. ∎

Nos ocupamos ahora de la función factorial:

Teorema 7.24 *Si $r > (2x)^{x+1}$, entonces $x! = E[r^x/\binom{r}{x}]$.*

Demostración: Podemos suponer $x > 0$.

$$\frac{r^x}{\binom{r}{x}} = \frac{r^x\, x!}{r(r-1)\cdots(r-x+1)} = \frac{x!}{(1-\frac{1}{r})\cdots(1-\frac{x-1}{r})} < \frac{x!}{(1-\frac{x}{r})^x}.$$

Vamos a usar la desigualdad

$$\frac{1}{1-\frac{x}{r}} < 1 + \frac{2x}{r}.$$

(Al desarrollarla equivale a $r > 2x$.) Se cumple que

$$\left(1 + \frac{2x}{r}\right)^x = \sum_{j=0}^{x} \binom{x}{j} \left(\frac{2x}{r}\right)^j = 1 + \frac{2x}{r} \sum_{j=1}^{x} \binom{x}{j} \left(\frac{2x}{r}\right)^{j-1}$$

$$< 1 + \frac{2x}{r} \sum_{j=1}^{x} \binom{x}{j} < 1 + \frac{2x}{r}\, 2^x.$$

Así pues,

$$\frac{r^x}{\binom{r}{x}} < x! \left(1 + \frac{2x}{r}\, 2^x\right) = x! + \frac{2^{x+1} x^{x+1}}{r} = x! + \frac{(2x)^{x+1}}{r} < x! + 1,$$

con lo que

$$x! \leq \frac{r^x}{\binom{r}{x}} < x! + 1.$$

La primera desigualdad se sigue, por ejemplo, de la primera línea de ecuaciones. ∎

Teorema 7.25 *La función $n!$ es diofántica.*

DEMOSTRACIÓN: Teniendo en cuenta el teorema anterior vemos que

$$m = n! \leftrightarrow \bigvee rstuv \in \mathbb{N}(s = 2x+1 \wedge t = x+1 \wedge r = s^t \wedge u = r^n$$

$$\wedge\, v = \tbinom{r}{n} \wedge mv \leq u < (m+1)v).$$

■

La última función que necesitamos es la siguiente:[6]

Teorema 7.26 *La función $h(a,b,y) = \prod_{k=1}^{y}(a+bk)$ es diofántica.*

DEMOSTRACIÓN: Veamos primero que si $bq \equiv a \pmod M$ entonces

$$\prod_{k=1}^{y}(a+bk) \equiv b^y y!\binom{q+y}{y} \pmod M.$$

Si $y = 0$ es inmediato (entendiendo que el producto vale 1). En otro caso

$$\begin{aligned} b^y y!\binom{q+y}{y} &= b^y(q+y)(q+y-1)\cdots(q+1) \\ &= (bq+yb)(bq+(y-1)b)\cdots(bq+b) \\ &\equiv (a+yb)(a+(y-1)b)\cdots(a+b) \pmod M. \end{aligned}$$

Tomemos ahora $M = b(a+by)^y + 1$. Así $(M,b) = 1$ y $M > \prod_{k=1}^{y}(a+bk)$. Existe un q tal que $bq \equiv a \pmod M$ (ver el apéndice B). Así, $\prod_{k=1}^{y}(a+bk)$ es el único natural congruente con $b^y y!\binom{q+y}{y}$ menor que M.

$$z = \prod_{k=1}^{y}(a+bk) \leftrightarrow \bigvee Mpqrstuvwx \in \mathbb{N}(r = a+by \wedge s = r^y \wedge M = bs+1$$

$$\wedge\, bq = a + Mt \wedge u = b^y \wedge v = y! \wedge z < M$$

$$\wedge\, w = q+y \wedge x = \tbinom{w}{y} \wedge z + Mp = uvx),$$

y el miembro derecho es claramente diofántico. ■

[6]Este teorema fue probado por primera vez por Davis y Putnam en 1958, basándose en ideas de Julia Robinson. Aquí damos una prueba posterior debida a Robinson.

Las funciones recursivas son diofánticas Hemos visto que si R es una relación diofántica entonces $\bigvee x\, R$ también lo es. Esto no es cierto si sustituimos el cuantificador existencial por el cuantificador universal, pero ahora podemos probar que sí es cierto para un cuantificador universal acotado. Con esto ya será fácil probar que las funciones diofánticas coinciden con las recursivas. Los resultados de este apartado son de Martin-Putnam-Robinson (1961).

Teorema 7.27 *Si $R(y, z, x_1, \ldots, x_n)$ es una relación diofántica, también lo es la relación $\bigwedge z \in \mathbb{N}(z \leq y \to R(y, z, x_1, \ldots, x_n))$.*

Demostración: Digamos que

$$R(y, z, x_1, \ldots, x_n) \leftrightarrow \bigvee y_1 \ldots y_m \in \mathbb{N}\, P(y, z, x_1, \ldots, x_n, y_1, \ldots, y_m) = 0$$

y llamemos $S(y, x_1, \ldots, x_n)$ a la relación del enunciado. Es claro que

$$S(y, x_1, \ldots, x_n) \leftrightarrow \bigvee u \in \mathbb{N}(u > 0 \wedge \bigwedge z \in \mathbb{N}(z \leq y \to$$

$$\bigvee y_1 \ldots y_m \in \mathbb{N}(y_1, \ldots, y_m \leq u \wedge P(y, z, x_1, \ldots, x_n, y_1, \ldots, y_m) = 0))).$$

Expresemos el polinomio P como suma de monomios $P = \sum_{r=1}^{N} t_r$, donde

$$t_r = c y^a z^b x_1^{q_1} \cdots x_n^{q_n} y_1^{s_1} \cdots y_m^{s_m}, \qquad c \in \mathbb{Z}.$$

Sea $u_r = |c| y^{a+b} x_1^{q_1} \cdots x_n^{q_n} u^{s_1+\cdots+s_m}$. Sea

$$Q(y, u, x_1, \ldots, x_n) = u + y + \sum_{r=1}^{N} u_r + 1.$$

Se cumple que $Q(y, u, x_1, \ldots, x_n) > u$, $Q(y, u, x_1, \ldots, x_n) > y$ y si $z \leq y$, $y_1, \ldots, y_m \leq u$, entonces

$$|P(y, z, x_1, \ldots, x_n, y_1, \ldots, y_m)| \leq Q(y, u, x_1, \ldots, x_n).$$

Sea $u > 0$. Veamos que

$$\bigwedge z \in \mathbb{N}(z \leq y \to \bigvee y_1 \ldots y_m \in \mathbb{N}(y_1, \ldots, y_m \leq u$$

$$\wedge\, P(y, z, x_1, \ldots, x_n, y_1, \ldots, y_m) = 0)).$$

equivale a

$$\bigvee c t a_1 \cdots a_m \in \mathbb{N}(t > 0 \wedge c > 0 \wedge 1 + ct = \prod_{k=1}^{y} (1 + kt) \wedge t = Q(y, u, x_1, \ldots, x_n)!$$

$$\wedge\, (1 + ct) \mid \prod_{j=0}^{u} (a_1 - j) \wedge \cdots \wedge (1 + ct) \mid \prod_{j=0}^{u} (a_m - j)$$

$$\wedge\, P(y, c, x_1, \ldots, x_n, y_1, \ldots, y_m) \equiv 0 \ (\text{mód } 1 + ct))$$

$$\wedge \bigvee y_1, \ldots y_m \in \mathbb{N}(y_1, \ldots, y_m \leq u \wedge P(y, 0, x_1, \ldots, x_n, y_1, \ldots, y_m) = 0).$$

La condición es suficiente: Tomamos $z \leq y$, y hemos de probar que

$$\bigvee y_1 \dots y_m \in \mathbb{N}(y_1, \dots, y_m \leq u \wedge P(y, z, x_1, \dots, x_n, y_1, \dots, y_m) = 0).$$

Para $z = 0$ es exactamente lo que afirma la última parte de la hipótesis. Supongamos que $0 < z \leq y$. Sea p_z un divisor primo de $1 + zt$, sea y_i^z el resto de dividir a_i entre p_z, para $i = 1, \dots, n$. Se cumple que $y_i^z < p_z$. Veamos que

a) $y_i^z \leq u$.

b) $P(y, z, x_1, \dots, x_n, y_1^z, \dots, y_m^z) = 0$.

En efecto, $p_z \mid (1 + zt) \mid (1 + ct) \mid \prod_{j=0}^{u} (a_i - j)$, luego existe un j tal que $0 \leq j \leq u$ tal que $p_z \mid (a_i - j)$, o sea, $j \equiv a_i \equiv y_i^z \pmod{p_z}$.

Como $t = Q(y, u, x_1, \dots, x_n)!$ y $p_z | (1 + zt)$, ha de ser

$$p_z > Q(y, u, x_1, \dots, x_n) > u.$$

Tenemos que $j \leq u < p_z$, $y_i^z < p_z$ y, como son congruentes, ha de ser $y_i^z = j$, luego se cumple a).

$1 + ct \equiv 0 \equiv 1 + zt \pmod{p_z}$, luego $z + zct \equiv c + zct \pmod{p_z}$ y de aquí que $z \equiv c \pmod{p_z}$. Tenemos también que $y_i^z \equiv a_i \pmod{p_z}$, luego

$$P(y, z, x_1, \dots, x_n, y_1^z, \dots, y_m^z) \equiv P(y, c, x_1, \dots, x_n, a_1, \dots, a_m) \equiv 0 \pmod{p_z}.$$

Pero $|P(y, z, x_1, \dots, x_n, y_1^z, \dots, y_m^z)| \leq Q(y, u, x_1, \dots, x_n) < p_z$, lo que implica que $P(y, z, x_1, \dots, x_n, y_1^z, \dots, y_m^z) = 0$. Esto prueba b) y también la suficiencia.

La condición es necesaria: La última parte es inmediata. Hemos de ver la primera. Para cada $0 < z \leq y$, sean $y_1^z, \dots, y_m^z \leq u$ tales que

$$P(y, u, x_1, \dots, x_n, y_1^z, \dots, y_m^z) = 0.$$

Sea $t = Q(y, u, x_1, \dots, x_n)! > 0$. Como $\prod_{k=1}^{y} (1 + kt) \equiv 1 \pmod{t}$, existe un $c > 0$ tal que $1 + ct = \prod_{k=1}^{y} (1 + kt)$.

Veamos que si $1 \leq k < l \leq y$, entonces $(1 + kt, 1 + lt) = 1$. En efecto, si $p \mid (1 + kt)$ y $p \mid (1 + lt)$ es un divisor primo común, entonces $p \mid (l - k)$, luego $p < y$, pero $Q(y, u, x_1, \dots, x_n) > y$, de donde $p \mid t$ y $p \mid (1 + kt)$, y así $p \mid 1$, contradicción.

Por el teorema chino del resto, para cada $1 \leq i \leq m$ existe un a_i tal que

$$a_i \equiv y_i^z \pmod{1 + zt} \qquad z = 1, \dots, y.$$

Como $1+ct \equiv 0 \pmod{1+zt}$ y $1 \equiv -zt \pmod{1+zt}$, tenemos que $(c-z)t \equiv 0 \pmod{1+zt}$ y, como $(t, 1+zt) = 1$, resulta que $c - z \equiv 0 \pmod{1+zt}$, es decir, $c \equiv z \pmod{1 + zt}$. Ahora,

$$P(y, c, x_1, \dots, x_n, a_1, \dots, a_m) \equiv P(y, z, x_1, \dots, x_n, y_1^z, \dots, y_m^z) = 0 \pmod{1+zt},$$

luego $1+zt \mid P(y, c, x_1, \ldots, x_n, a_1, \ldots, a_m)$ y, como los $1+zt$ son primos entre sí, su producto también divide a P, es decir $1 + ct \mid P(y, c, x_1, \ldots, x_n, a_1, \ldots, a_m)$ o, equivalentemente,

$$P(y, c, x_1, \ldots, x_n, a_1, \ldots, a_m) \equiv 0 \text{ (mód } 1 + ct).$$

Como $a_i \equiv y_i^z \text{ (mód } 1 + zt)$, tenemos que $1 + zt \mid a_i - y_i^z$, de donde se sigue que $1 + zt \mid \prod_{j=0}^{u} (a_i - j)$. Como los $1 + zt$ son primos entre sí, también $1 + ct \mid \prod_{j=0}^{u} (a_i - j)$ y tenemos la condición.

Con esto hemos probado que la relación $S(y, x_1, \ldots, x_n)$ equivale a

$$\bigvee u \in \mathbb{N}(u > 0 \wedge (\bigvee c t a_1 \cdots a_m \in \mathbb{N}(t > 0 \wedge c > 0 \wedge 1 + ct = \prod_{k=1}^{y} (1 + kt)$$

$$\wedge\, t = Q(y, u, x_1, \ldots, x_n)!$$

$$\wedge\, (1 + ct) \mid \prod_{j=0}^{u} (a_1 - j) \wedge \cdots \wedge (1 + ct) \mid \prod_{j=0}^{u} (a_m - j)$$

$$\wedge\, P(y, c, x_1, \ldots, x_n, y_1, \ldots, y_m) \equiv 0 \text{ (mód } 1 + ct))$$

$$\wedge \bigvee y_1, \ldots y_m \in \mathbb{N}(y_1, \ldots, y_m \leq u \wedge P(y, 0, x_1, \ldots, x_n, y_1, \ldots, y_m) = 0)).$$

Claramente esto equivale a su vez a

$$\bigvee u v c t a_1 \ldots a_m e f g_1 \ldots g_m h_1 \ldots h_m i y_1 \ldots y_m \in \mathbb{N}(u > 0 \wedge t > 0 \wedge c > 0$$

$$\wedge\, v = u + 1 \wedge e = 1 + ct \wedge e = \prod_{k=1}^{y} (1 + kt) \wedge f = Q(y, u, x_1, \ldots, x_n) \wedge t = f!$$

$$\wedge\, g_1 = a_1 - v \wedge \cdots \wedge g_m = a_m - v \wedge h_1 = \prod_{k=1}^{v} (g_1 + k) \wedge \cdots \wedge h_m = \prod_{k=1}^{v} (g_m + k)$$

$$\wedge\, e | h_1 \wedge \cdots \wedge e | h_m \wedge i = P(y, c, x_1, \ldots, x_n, a_1, \ldots, a_m) \wedge e \mid i$$

$$\wedge\, y_1, \ldots, y_m \leq u \wedge P(y, 0, x_1, \ldots, x_n, y_1, \ldots, y_m) = 0).$$

Los resultados del apartado anterior muestran que esta última expresión es diofántica. ∎

Ahora ya podemos demostrar el teorema 7.15. Ya hemos visto que las funciones diofánticas son recursivas. Nos falta el recíproco. Para probarlo observamos en primer lugar que las funciones recursivas elementales son claramente diofánticas:

$y = c(x) \leftrightarrow y = 0$,
$y = s(x) \leftrightarrow y = x + 1$,
$y = p_i^k(x_1, \ldots, x_k) \leftrightarrow y = x_i$.

Así mismo, la composición de funciones diofánticas es diofántica, ya que si $h(x_1,\ldots,x_n) = f(g_1(x_1,\ldots,x_n),\ldots,g_m(x_1,\ldots,x_n))$, entonces

$$y = h(x_1,\ldots,x_n) \leftrightarrow \bigvee y_1 \ldots y_m (y_1 = g_1(x_1,\ldots,x_n) \wedge \cdots$$

$$\wedge\, y_m = g_m(x_1,\ldots,x_n) \wedge y = h(y_1,\ldots,y_m)).$$

Veamos ahora que si

$$\begin{aligned} h(x_1,\ldots,x_n,0) &= f(x_1,\ldots,x_n),\\ h(x_1,\ldots,x_n,t+1) &= g(t,h(x_1,\ldots,x_n,t),x_1,\ldots,x_n), \end{aligned}$$

y las funciones f y g son diofánticas, entonces h también lo es. En efecto,

$$y = h(x_1,\ldots,x_n,z) \leftrightarrow \bigvee u \in \mathbb{N}\big(\bigvee v \in \mathbb{N}(v = S(0,u) \wedge v = f(x_1,\ldots,x_n))$$

$$\wedge \bigwedge t \in \mathbb{N}(t \le z \to t = z \vee \bigvee v \in \mathbb{N}(v = S(t+1,u)$$

$$\wedge\, v = g(t,S(t,u),x_1,\ldots,x_n))) \wedge y = S(z,u)\big),$$

con lo que el teorema anterior prueba que h es diofántica. El caso $n = 0$ es similar.

Por último, si $h(x_1,\ldots,x_n) = \mu y\, f(x_1,\ldots,x_n,y) = 0$ y f es diofántica, entonces h también lo es, ya que

$$y = h(x_1,\ldots,x_n) \leftrightarrow \bigvee z \in \mathbb{N}(z = f(x_1,\ldots,x_n,y) \wedge z = 0)$$

$$\wedge \bigwedge t \in \mathbb{N}(t \le y \to t = y \vee \bigvee u \in \mathbb{N}(u = f(x_1,\ldots,x_n,t) \wedge u > 0))),$$

y podemos aplicar de nuevo el teorema anterior.

Es claro que esto prueba que toda función recursiva es diofántica. ∎

Ejercicio: Probar que toda relación recursiva es diofántica. Probar que la relación "ser (el número de Gödel de) un teorema de la aritmética de Peano" es diofántica pero no recursiva.

La ecuación de Pell Nos falta demostrar que la función x^y es diofántica. La prueba se basa en un estudio minucioso de las soluciones de una ecuación diofántica clásica: la ecuación de Pell. Se trata de la ecuación $x^2 - dy^2 = 1$, donde d es un número natural no cuadrado perfecto.

Las soluciones de esta ecuación están relacionadas con el anillo cuadrático $\mathbb{Z}[\sqrt{d}\,] = \{a + b\sqrt{d} \mid a,b \in \mathbb{Z}\}$ (en el apéndice B presentamos las propiedades básicas del cuerpo $\mathbb{Q}(\sqrt{d}\,)$).

La norma $\mathrm{N} : \mathbb{Z}[\sqrt{d}\,] \longrightarrow \mathbb{Z}$ dada por $\mathrm{N}(a+b\sqrt{d}\,) = a^2 - db^2$ es multiplicativa. Llamaremos *unidades* de $\mathbb{Z}[\sqrt{d}\,]$ a los enteros cuadráticos $\alpha \in \mathbb{Z}[\sqrt{d}\,]$ de norma 1. De este modo, un par (a,b) es una solución de la ecuación de Pell si y sólo si $\alpha = a + b\sqrt{d} \in \mathbb{Z}[\sqrt{d}\,]$ es una unidad de $\mathbb{Z}[\sqrt{d}\,]$.

Como la norma es multiplicativa, el producto de dos unidades es una unidad, y el inverso de una unidad es también una unidad. La solución trivial $(1,0)$ se corresponde con la unidad $1 \in \mathbb{Z}[\sqrt{d}]$. En lo que sigue nos restringiremos al caso particular en que $d = a^2 - 1$, con lo que otra unidad es $\epsilon = a + \sqrt{d}$, correspondiente a la solución $(a, 1)$. Vamos a probar que esta unidad genera a todas las demás y, por consiguiente, nos da todas las soluciones de la ecuación de Pell.

Teorema 7.28 *Con la notación anterior, las unidades de $\mathbb{Z}[\sqrt{d}]$ son exactamente las de la forma $\pm\epsilon^n$, donde $n \in \mathbb{Z}$.*

DEMOSTRACIÓN: Observemos que $1 < \epsilon$. La prueba se basa en que no existe ninguna unidad tal que $1 < \alpha < \epsilon$. En efecto, sea $\alpha = x + y\sqrt{d}$. Tenemos que

$$1 = (x + y\sqrt{d}\,)(x - y\sqrt{d}) = (a + \sqrt{d}\,)(a - \sqrt{d}\,),$$

luego

$$\frac{x - y\sqrt{d}}{a - \sqrt{d}} = \frac{a + \sqrt{d}}{x + y\sqrt{d}} = \frac{\epsilon}{\alpha} > 1,$$

de donde $x - y\sqrt{d} > a - \sqrt{d}$ y $-x + y\sqrt{d} < -a + \sqrt{d}$. Por otra parte

$$x - y\sqrt{d} = \frac{1}{x + y\sqrt{d}} < 1,$$

luego $-1 < -x + y\sqrt{d}$. Tenemos, pues, que

$$-1 < -x + y\sqrt{d} < -a + \sqrt{d},$$

$$1 < x + y\sqrt{d} < a + \sqrt{d}.$$

Sumando miembro a miembro queda $0 < 2y\sqrt{d} < 2\sqrt{d}$, luego $0 < y < 1$, pero y es entero.

Sea ahora η cualquier unidad de $\mathbb{Z}[\sqrt{d}]$. Cambiando η por $-\eta$ podemos suponer $\eta > 0$ y cambiando η por η^{-1} podemos suponer que $\eta > 1$. La sucesión ϵ^n es monótona creciente y no está acotada, por lo que existe un n tal que $\epsilon^n \le \eta < \epsilon^{n+1}$. Por consiguiente $1 \le \eta\epsilon^{-n} < \epsilon$. Por lo que acabamos de probar $\eta\epsilon^{-n} = 1$, es decir, $\eta = \epsilon^n$. ∎

Como $\epsilon > 1$, para $n > 0$ se cumple $\epsilon^n > 1$, $0 < \epsilon^{-n} < 1$, $-\epsilon^n < -1$ y $-1 < -\epsilon^{-n} < 0$. Si (x, y) es una solución natural de la ecuación de Pell, entonces $x + y\sqrt{d} \ge 1$, luego ha de ser de la forma ϵ^n.

Definición 7.29 Si n y $a > 1$ son números naturales, definimos $x_n(a)$, $y_n(a)$ como los números naturales que cumplen $\epsilon^n = x_n(a) + y_n(a)\sqrt{d}$, donde $d = a^2 - 1$. Si no hay confusión omitiremos a.

El hecho de que sean números naturales se prueba fácilmente por inducción a partir de las relaciones (7.4) más abajo. Los pares (x_n, y_n) son todas las soluciones naturales de la ecuación de Pell.

La fórmula $\epsilon^{m\pm n} = \epsilon^m \epsilon^{\pm n}$ se traduce inmediatamente en las relaciones

$$x_{m\pm n} = x_m x_n \pm d y_m y_n, \qquad y_{m\pm n} = x_n y_m \pm x_m y_n.$$

En particular

$$x_{m\pm 1} = a x_m \pm d y_m, \qquad y_{m\pm 1} = a y_m \pm x_m. \tag{7.4}$$

La prueba de que m^n es diofántica se basa en los siguientes resultados sobre las soluciones de la ecuación de Pell:

1. $(x_n, y_n) = 1$

   Demostración: Si $p \mid x_n$ y $p \mid y_n$, entonces $p \mid x_n^2 - d y_n^2 = 1$.

2. $y_m \mid y_n$ syss $m \mid n$.

   Demostración: Veamos por inducción sobre k que $y_m \mid y_{mk}$. Para $k = 1$ es obvio.

   $$y_{m(k+1)} = x_m y_{mk} + x_{mk} y_m,$$

   luego si $y_m \mid y_{mk}$, también $y_m \mid y_{m(k+1)}$.

   Supongamos ahora que $y_m \mid y_n$ pero $m \nmid n$. Sea $n = mq + r$, con $0 < r < m$. Entonces $y_n = x_r y_{mq} + x_{mq} y_r$. Por la parte ya probada $y_m \mid y_{mq}$, luego $y_m \mid x_{mq} y_r$. Ahora bien, $(x_{mq}, y_{mq}) = 1$, luego también $(x_{mq}, y_m) = 1$ y entonces ha de ser $y_m \mid y_r$, pero (7.4) implica que la sucesión y_m es estrictamente creciente, con lo que tenemos una contradicción.

3. $y_{nk} \equiv k x_n^{k-1} y_n \pmod{y_n^3}$.

   Demostración:

   $$x_{nk} + y_{nk}\sqrt{d} = \epsilon^{nk} = (x_n + y_n\sqrt{d}\,)^k = \sum_{i=0}^{k} \binom{k}{i} x_n^{k-i} y_n^i d^{i/2},$$

   luego

   $$y_{nk} = \sum_{\substack{i=0 \\ \text{impar}}}^{k} \binom{k}{i} x_n^{k-i} y_n^i d^{(i-1)/2},$$

   pero los sumandos con $i > 1$ son $\equiv 0 \pmod{y_n^3}$, luego tenemos la congruencia pedida.

4. $y_n^2 \mid y_{n y_n}$.

   Demostración: Se sigue inmediatamente de la propiedad anterior para $k = y_n$.

5. Si $y_n^2 \mid y_m$, entonces $y_n \mid m$.

   DEMOSTRACIÓN: Por 2 sabemos que $n \mid m$. Sea $m = nk$. Por 3 tenemos $y_m = kx_n^{k-1}y_n + ry_n^3$, luego $y_n^2 \mid kx_n^{k-1}y_n$. Por 1 ha de ser $y_n \mid k$ y, en consecuencia, $y_n \mid m$.

6. $x_{n+1} = 2ax_n - x_{n-1}, \quad y_{n+1} = 2ay_n - y_{n-1}$.

   Basta sumar las relaciones

$$\begin{aligned} x_{n+1} &= ax_n + dy_n, & y_{n+1} &= ay_n + x_n, \\ x_{n-1} &= ax_n - dy_n, & y_{n-1} &= ay_n - x_m. \end{aligned}$$

7. $y_n \equiv n \pmod{a-1}$.

   DEMOSTRACIÓN: Se cumple para $y_0 = 0$ e $y_1 = 1$. Por inducción y la propiedad anterior:

$$y_{n+1} = 2ay_n - y_{n-1} \equiv 2n - (n-1) \equiv n+1 \pmod{a-1}.$$

8. Si $a \equiv b \pmod{c}$, entonces

$$x_n(a) \equiv x_n(b) \pmod{c}, \qquad y_n(a) \equiv y_n(b) \pmod{c}.$$

   DEMOSTRACIÓN: Para $n = 0$, 1 se da la igualdad. Por inducción:

$$y_{n+1}(a) = 2ay_n(a) - y_{n-1}(a) \equiv 2by_n(b) - y_{n-1}(b) = y_{n+1}(b) \pmod{c}.$$

9. $n \equiv y_n \pmod 2$.

   DEMOSTRACIÓN: $y_{n+1} = 2ay_n - y_{n-1} \equiv y_{n-1} \pmod 2$, luego si n es par $n \equiv y_0 = 0 \pmod 2$ y si n es impar $n \equiv y_1 = 1 \pmod 2$.

10. $x_n - y_n(a-y) \equiv y^n \pmod{2ay - y^2 - 1}$. (Ésta es la propiedad que conecta las soluciones de la ecuación de Pell con la exponencial.)

    DEMOSTRACIÓN: $x_0 - y_0(a-y) = 1$, $x_1 - y_1(a-y)$, luego se cumple para $n = 0$, 1. Si vale para n,

$$\begin{aligned} x_{n+1} - y_{n+1}(a-y) &= 2a(x_n - y_n(a-y)) - (x_{n-1} - y_{n-1}(a-y)) \equiv 2ay^n - y^{n-1} \\ &= y^{n-1}(2ay - 1) \equiv y^{n-1}y^2 = y^{n+1} \pmod{2ay - y^2 - 1}. \end{aligned}$$

11. $n \le y_n < y_{n+1}, \quad a^n \le x_n < x_{n+1}, \quad x_n \le (2a)^n$.

    DEMOSTRACIÓN: Las desigualdades sobre y_n se siguen inmediatamente de (7.4). Para las otras, usando (7.4) y 6:

$$ax_n \le ax_n + dy_n = 2ax_n - (ax_n - dy_n) = 2ax_n - x_{n-1} = x_{n+1} \le 2ax_n.$$

    O sea, $ax_n \le x_{n+1} \le 2ax_n$. En particular $x_n < x_{n+1}$.

    Por inducción sobre n: $a^0 = 1 = x_0 = (2a)^0$. Si $a^n \le x_n \le (2a)^n$, entonces $a^{n+1} \le ax_n \le x_{n+1} \le 2ax_n \le (2a)^{n+1}$.

12. $x_{2n\pm j} \equiv -x_j \pmod{x_n}$.

    DEMOSTRACIÓN:

$$x_{2n\pm j} = x_n x_{n\pm j} + dy_n y_{n\pm j} \equiv dy_n(y_n x_j \pm x_n y_j)$$
$$\equiv dy_n^2 x_j = (x_n^2 - 1)x_j \equiv -x_j \pmod{x_n}.$$

13. $x_{4n\pm j} \equiv x_j \pmod{x_n}$.

    DEMOSTRACIÓN: Por el resultado anterior,

$$x_{4n\pm j} \equiv -x_{2n\pm j} \equiv x_j \pmod{x_n}.$$

14. Si $x_i \equiv x_j \pmod{x_n}$ con $i \le j \le 2n$ y $n > 0$, entonces $i = j$ excepto si $a = 2$, $n = 1$, $i = 0$, $j = 2$.

    DEMOSTRACIÓN: Por 11 tenemos que $1 = x_0 < x_1 < \cdots < x_{n-1}$ y por 12 resulta que $x_{n+1}, x_{n+2}, \ldots, x_{2n-1}, x_{2n}$ son congruentes, respectivamente, con $-x_{n-1}, -x_{n-2}, \ldots - x_1, -x_0 = -1$ módulo x_n, luego vemos que $x_0, \ldots, x_{2n}$ son congruentes módulo x_n con

$$-x_{n-1} < -x_{n-1} < \cdots < -x_1 < -x_0 < x_0 < x_1 < \cdots < x_{n-1} \tag{7.5}$$

    Sea

$$q = \begin{cases} (x_n - 1)/2 & \text{si } x_n \text{ es impar,} \\ x_n/2 & \text{si } x_n \text{ es par.} \end{cases}$$

    En ambos casos, (7.4) nos da que

$$x_{n-1} \le x_n/a \le x_n/2 \le q,$$

    luego los números (7.5) están comprendidos entre $-q$ y q. Si x_n es impar entonces $-q, \ldots, q$ forman un sistema de restos módulo x_n, luego los números (7.5) son no congruentes dos a dos y el resultado está probado. Si x_n es par entonces un sistema de restos lo forman los números $-q+1, \ldots, q$ y la conclusión es la misma salvo si $x_{n-1} = q$, en cuyo caso $i = n - 1$, $j = n + 1$ contradicen lo que queremos probar.

    Por (7.4), si $ax_{n-1} + dy_{n-1} = x_n = 2q = 2x_{n-1}$, entonces $a = 2$, $y_{n-1} = 0$, $n = 1$, $i = 0$ y $j = 2$.

15. Si $x_i \equiv x_j \pmod{x_n}$ con $0 < i \le n$ y $0 \le j < 4n$, entonces $j = i$ o bien $j = 4n - i$.

    DEMOSTRACIÓN: Supongamos que $j \le 2n$. Entonces por el resultado anterior tenemos que $i = j$ salvo si $n = 1$, $i = 2$, $j = 0$, pero entonces $i > n$, lo cual es imposible.

    Supongamos que $j > 2n$. Sea $j' = 4n - j$. Así $0 < j' < 2n$ y por 13 se sigue cumpliendo $x_i \equiv x_j \equiv x_{j'} \pmod{x_n}$, y ahora concluimos que $i = j'$.

16. Si $0 < i \le n$ y $x_i \equiv x_j \pmod{x_n}$, entonces $j \equiv \pm i \pmod{4n}$.

    DEMOSTRACIÓN: Sea $j = 4nq + j'$ con $0 \le j' < 4n$. Por 13 tenemos que $x_j \equiv x_{j'} \pmod{x_n}$. Por el resultado anterior, o bien $i = j'$ o bien $i = 4n - j'$, luego $j \equiv j' \equiv \pm i \pmod{4n}$.

La función exponencial Finalmente estamos en condiciones de probar que la función exponencial $m = n^k$ es diofántica. Éstas son las ecuaciones que la caracterizan:

$$\begin{aligned}
&\text{I} && x^2 - (a^2-1)y^2 = 1,\\
&\text{II} && u^2 - (a^2-1)v^2 = 1,\\
&\text{III} && s^2 - (b^2-1)t^2 = 1,\\
&\text{IV} && v = ry^2,\\
&\text{V} && b = 1 + 4(p+1)y = a + (q+1)u,\\
&\text{VI} && s = x + cu,\\
&\text{VII} && t = k + 4dy,\\
&\text{VIII} && y = k + e\\
&\text{IX} && y = \overline{y} + 1,\ v = \overline{v} + 1,\ t = \overline{t} + 1,\ x = \overline{x} + 1,\\
&\text{X} && (x - y(a-n) - m)^2 = f^2(2an - n^2 - 1)^2,\\
&\text{XI} && m + g + 1 = 2an - n^2 - 1,\\
&\text{XII} && w = n + h + 1 = k + l + 1,\\
&\text{XIII} && a^2 - (w^2-1)(w-1)^2(z+1)^2 = 1.
\end{aligned}$$

Lo probaremos en varios pasos. En primer lugar vemos que las primeras ecuaciones caracterizan a las soluciones de la ecuación de Pell:

Teorema 7.30 *Sean a, x, $k \in \mathbb{N}$, $a > 1$, $k > 0$. El sistema* I – IX *tiene solución natural para las demás variables si y sólo si $x = x_k(a)$.*

Demostración: Supongamos que el sistema tiene solución. Las referencias formadas por un único número arábigo corresponden a los resultados del apartado anterior.

Por V tenemos $b > a > 1$.

Por I, II, III y IX existen i, j, $n > 0$ tales que $x = x_i(a)$, $y = y_i(a)$, $u = x_n(a)$, $v = y_n(a)$, $s = x_j(b)$, $t = y_j(b)$.

Por IV, $y \le v$, luego $i \le n$.

Por V y VI, $b \equiv a \pmod{x_n(a)}$, $x_j(b) \equiv x_i(a) \pmod{x_n(a)}$.

Por 8, $x_j(b) \equiv x_j(a) \pmod{x_n(a)}$, luego $x_i(a) \equiv x_j(a) \pmod{x_n(a)}$.

Por 16,

$$j \equiv \pm i \pmod{4n}. \tag{7.6}$$

Por IV, $y_i(a)^2 \mid y_n(a)$, de donde, por 5, $y_i(a) \mid n$.

Por (7.6) tenemos

$$j \equiv \pm i \pmod{4y_i(a)}. \tag{7.7}$$

Por V, $b \equiv 1$ (mód $4y_i(a)$) y por 7,

$$y_j(b) \equiv j \text{ (mód } 4y_i(a)). \tag{7.8}$$

Por VII

$$y_j(b) \equiv k \text{ (mód } 4y_i(a)). \tag{7.9}$$

Por (7.7), (7.8) y (7.9),

$$k \equiv \pm i \text{ (mód } 4y_i(a)). \tag{7.10}$$

Por VIII, $k \leq y_i(a)$ y por 11, $i \leq y_i(a)$. Como los números $-2y+1, \ldots, 2y$ son un sistema de restos módulo $4y$, necesariamente $k = i$, y así, $x = x_i(a) = x_k(a)$.

Supongamos ahora que $x = x_k(a)$, $y = y_k(a)$. Entonces se cumple I.
Sea $m = 2ky_k(a) > 0$, $u = x_m(a)$, $v = y_m(a)$. Se cumple II.
Por 4 y 2, $y^2 = y_k(a)^2 \mid y_{ky_k(a)}(a) \mid y_m(a) = v$. Se cumple IV.
Por 9, $y_m(a) = v$ es par y por 1, $u = x_m(a)$ cumple $(u, v) = 1$. En particular es impar.
También $(u, 4y) = 1$, pues, como u es impar e $y \mid v$,

$$(u, 4y) = (u, y) \mid (u, v) = 1.$$

Por el teorema chino del resto existe un b_0 tal que

$$b_0 \equiv 1 \text{ (mód } 4y) \qquad \text{y} \qquad b_0 \equiv a \text{ (mód } u),$$

y cualquier $b = b_0 + 4juy$ cumple lo mismo. Se cumple V.
Sean $s = x_k(b)$, $t = y_k(b)$. Se cumplen III y IX.
Como $b > a$, se cumple $s = x_k(b) > x_k(a) = x$.
Por 8, $s \equiv x$ (mód u). Se cumple VI.
Por 11, $k \leq y_k(b) = t$ y por 7, $t \equiv k$ (mód $b-1$).
Por V, $4y \mid b-1$, luego $t \equiv k$ (mód $4y$). Se cumple VII.
Por 11, $k \leq y_k(a) = y$. Se cumple VIII. ∎

Para ocuparnos de las cuatro ecuaciones que faltan, necesitamos una observación sencilla:

Teorema 7.31 *Si $a > y^k$, $y > 0$, $k > 0$, entonces $2ay - y^2 - 1 > y^k$.*

Demostración: Basta observar que

$$2ay - y^2 - 1 = a^2 - 1 - (a-y)^2 \geq a^2 - 1 - (a-1)^2 = 2a - 2 \geq a > y^k.$$

La primera desigualdad es porque $y \leq y^k < a$, y la segunda porque $a \geq 2$). ∎

Teorema 7.32 *Si m, n, k son naturales no nulos, entonces $m = n^k$ si y sólo si el sistema de ecuaciones* I–XIII *tiene solución natural en las demás variables.*

DEMOSTRACIÓN: Supongamos que el sistema tiene solución. Por XII $w > 1$, luego por XIII, $a > 1$. Por 7.30, las primeras ecuaciones implican que $x = x_k(a)$, $y = y_k(a)$.

Por 10, $x_k(a) - y_k(a)(a-n) \equiv n^k \pmod{2an - n^2 - 1}$, luego por X, tenemos que $m \equiv n^k \pmod{2an - n^2 - 1}$.

Por XII, k, $n < w$.

Por XIII, existe un $j > 0$ tal que $a = x_j(w)$, $(w-1)(z+1) = y_j(w)$.

Por 7, $j \equiv (w-1)(z+1) \equiv 0 \pmod{w-1}$. Por consiguiente $j \geq w-1$.

Por 11, $a = x_j(w) \geq w^j \geq w^{w-1} > n^k$.

Por XI, $m < 2an - n^2 - 1$ y, por el teorema anterior, $n^k < 2an - n^2 - 1$.

Como m y n^k son congruentes y menores que el módulo, necesariamente $m = n^k$.

Supongamos ahora que $m = n^k$. Tomemos $w > n$, $w > k$. Se cumple XII.

Sea $a = x_{w-1}(w) > 1$. Por 7, $y_{w-1}(w) \equiv (w-1) \equiv 0 \pmod{w-1}$, luego $y_{w-1}(w) = (w-1)(z+1)$, para un z. Se cumple XIII.

Por 11, $a = x_{w-1}(w) \geq w^{w-1} > n^k$, luego podemos aplicar el teorema anterior y $2an - n^2 - 1 > n^k = m$. Se cumple XI

Sean $x = x_k(a)$, $y = y_k(a)$. Por 10, $x - y(a-n) \equiv n^k \pmod{2an - n^2 - 1}$. Se cumple X. Las ecuaciones I–IX se cumplen por 7.30. ∎

Eliminar la restricción sobre que m, n y k sean no nulos es un simple ejercicio. De todos modos en ningún momento hemos usado más de lo que acabamos de probar.

Segunda parte

La lógica de la teoría de conjuntos

Introducción a la teoría axiomática de conjuntos

Es un hecho aceptado por prácticamente todos los matemáticos de hoy en día que la matemática requiere una fundamentación rigurosa, aunque no muchos sabrían precisar por qué. El desconocimiento de los problemas exactos que hacen necesaria dicha fundamentación produce a menudo unas expectativas desmesuradas, que resultan ser imposibles en virtud de los resultados que hemos expuesto en la primera parte de este libro. En esta primera parte hemos ido explicando los problemas que necesitamos resolver, al tiempo que hemos desarrollado las herramientas necesarias para ello; así mismo hemos discutido las características de un posible proyecto de fundamentación de las matemáticas, tanto las que necesitamos exigir como las que podemos aspirar a conseguir. Ahora ha llegado el momento de concretar ese proyecto.

Sobre la noción de conjunto Hoy sabemos que todos los conceptos de la matemática moderna, desde los números naturales hasta las variedades diferenciables, pueden reducirse a la noción de conjunto o colección de objetos, es decir, todos ellos pueden definirse formalmente a partir de éstos. Así pues, para dar completo rigor a todas las afirmaciones matemáticas basta con dar rigor a las afirmaciones sobre conjuntos. Ahora bien, los matemáticos se encuentran en su trabajo con tres "tipos" de conjuntos:

a) Conjuntos de los que podemos hablar informalmente, porque hacen referencia a colecciones bien definidas sobre las que cualquier afirmación tiene un significado objetivo. Por ejemplo, el conjunto $\mathbb{N}$ de los números naturales. A menudo, para enfatizar que un conjunto pertenece a esta categoría hemos venido usando y seguiremos usando la palabra "colección" en lugar de "conjunto".

b) Conjuntos de los que podemos hablar formalmente sin caer en contradicciones, pero a los que no sabemos asignar un significado objetivo. Por ejemplo, el conjunto $\mathcal{P}\mathbb{N}$ de todos los subconjuntos de $\mathbb{N}$.

c) Conjuntos de los que no podemos hablar sin caer en contradicciones, como el conjunto V de todos los conjuntos.

Ya hemos discutido en varias ocasiones la diferencia entre a) y b). Respecto de c), notemos que el clásico teorema de Cantor afirma que, dado un conjunto X, el conjunto $\mathcal{P}X$ de todos sus subconjuntos tiene mayor número de elementos que X. La prueba es lógicamente irrefutable, pero si aplicamos esto al conjunto V obtenemos una contradicción, pues $\mathcal{P}V$ es un subconjunto de V y otro teorema más elemental aún afirma que entonces debería tener menor o igual número de elementos que V.

Cantor era consciente de este problema, y ello le llevó a distinguir entre lo que llamó *multiplicidades consistentes* o *conjuntos* y *multiplicidades inconsistentes* o *absolutamente infinitas.*[7] Para explicar esta distinción conviene observar antes otra más elemental. La colección de los números naturales es infinita, lo cual significa que no podemos recorrerlos todos en la práctica, pero esto no nos impide hacer afirmaciones sobre todos ellos. Ahora bien, una afirmación sobre todos los números naturales puede ser de dos tipos:

- Afirmaciones que involucran sólo una cantidad finita de números naturales, aunque éstos puedan ser arbitrariamente grandes. Por ejemplo, cuando probamos que existen infinitos números primos, lo que hacemos es probar que, dado cualquier número natural n, podemos construir un número primo $p > n$. Se dice entonces que hablamos de un *"infinito potencial".*

- Afirmaciones que involucran simultáneamente a todos los números naturales. Por ejemplo, cuando afirmamos que $\mathbb{N}$ tiene menos elementos que $\mathcal{P}\mathbb{N}$, estamos estimando el tamaño de la totalidad de los números naturales, comparándolo con el tamaño de otro conjunto. Esta afirmación no es reducible a afirmaciones sobre conjuntos finitos de números naturales, aquí estamos tratando a $\mathbb{N}$ como un todo, como una colección completada. Se dice entonces que hablamos de un *"infinito actual".*

Toda la teoría de conjuntos cantoriana se basa en la posibilidad de pensar las colecciones infinitas como totalidades, es decir, la posibilidad de hablar simultáneamente de todos sus elementos sin tener en cuenta que no disponemos de una representación explícita de dicha totalidad. Las paradojas como la que acabamos de comentar llevaron a Cantor a la conclusión de que este principio de actualización del infinito no es universal, sino que hay colecciones que podemos pensar como un todo (los conjuntos) y las que necesariamente hemos de pensar como no acabadas, y es en este sentido de "esencialmente inacabadas" en el que hay que entender su noción de "multiplicidades absolutamente infinitas". Son colecciones que sólo podemos considerar potencialmente, en el sentido de que podemos hablar de sus elementos, pero que nos llevan a contradicciones si pretendemos tratarlas actualmente como objetos a los que aplicar los teoremas de la teoría de conjuntos.

La distinción de Cantor es muy lúcida. A menudo se le ha objetado que, es definitiva, viene a decir que las multiplicidades inconsistentes no existen, por

[7] Entre éstas últimas habría que incluir muchas más, aparte de la colección de todos los conjuntos: la de todos los cardinales, la de todos los conjuntos finitos o la de todos los espacios vectoriales, por citar sólo unas pocas.

lo que hubiera sido más afortunado distinguir entre propiedades que definen conjuntos (como "ser un subconjunto de $\mathbb{N}$") y propiedades que no definen conjuntos (como "ser un conjunto"). No obstante, esta crítica se desvía del núcleo del problema y hace blanco en una cuestión superficial. El problema está en el paso de las colecciones de tipo a) a las colecciones de tipo b). Si estamos trabajando a nivel metamatemático, podemos equiparar "conjunto" con "colección de objetos", de modo que siempre que tenemos unos objetos bien determinados podemos hablar de la colección que forman. Pero esto sólo es válido mientras dispongamos de criterios para dar un significado a cada afirmación sobre la totalidad de los objetos que pretendemos considerar como un todo. Cuando no es así, cuando no tenemos más alternativa que emplear la palabra "conjunto" formalmente, sin ser capaces de atribuirle un significado preciso, entonces debemos tener presente que "conjunto" y "colección de objetos" no son sinónimos en absoluto. Por el contrario, tenemos conjuntos como $\mathcal{P}\mathbb{N}$ a los que no podemos atribuirles como significado ninguna colección de objetos que conozcamos (conocemos algunas colecciones de objetos que podemos identificar con elementos de $\mathcal{P}\mathbb{N}$, pero no una colección de objetos que podamos identificar con *la totalidad* de sus elementos) y tenemos colecciones de objetos, como la colección de todos los conjuntos, que no podemos identificar con ningún conjunto.

Para entender esto debemos tener presente que ahora "conjunto" es un término técnico. Conjuntos son los objetos que estudiamos en teoría de conjuntos. Admitiendo que los axiomas de la teoría de conjuntos sean consistentes, el teorema de completitud nos dice que el término "conjunto" tiene al menos una interpretación, y puede demostrarse que en tal caso tiene infinitas interpretaciones distintas, es decir, que la teoría de conjuntos tiene infinitos modelos distintos entre sí en el sentido de que para cada par de ellos hay una sentencia verdadera en uno y falsa en el otro. Si fijamos una interpretación de la palabra "conjunto" (un modelo) podemos hablar con toda legitimidad de la colección de todos los conjuntos, es decir, de todos los objetos que hemos acordado en llamar conjuntos, del universo del modelo. Se trata de una colección de objetos bien definida, y la idea cantoriana de que forman una multiplicidad absolutamente infinita se concreta ahora en que dicha totalidad no puede constituir la extensión —la colección de los elementos— de un conjunto y, por consiguiente, no podemos aplicarle ninguno de los teoremas de la teoría (que sólo se refieren a los conjuntos).

De todo esto no se desprende en absoluto que los conjuntos, tal y como los conciben los matemáticos, sean una farsa. Una analogía que puede ser útil es la siguiente: la intuición nos proporciona un significado preciso a las nociones de "círculo" y "esfera", pero no a la noción de "esfera cuatridimensional". Esto hace que sólo podamos conocer formalmente la geometría euclídea de cuatro dimensiones. Podemos considerar evidente que una recta y una circunferencia se cortan a lo sumo en dos puntos o bien demostrarlo a partir de unos axiomas. En un caso trabajamos informalmente y en el otro formalmente. Sin embargo, sólo podemos dar sentido riguroso a una afirmación análoga en dimensiones superiores a través de una geometría axiomática. Pese a ello, cualquier distinción entre la geometría tridimensional euclídea y la geometría cuatridimensional euclídea

es subjetiva: depende de nuestra capacidad de intuición, la cual, a su vez, está condicionada ya por la psicología humana, ya por la física del mundo en que vivimos. A priori, nada nos impide suponer que puedan existir seres conscientes que se representen intuitivamente un espacio de cuatro dimensiones igual que nosotros nos representemos el espacio de tres dimensiones. Para ellos, la geometría que para nosotros es puramente formal tendría un contenido intuitivo. Por ello es más razonable pensar que desde un punto de vista estrictamente matemático la geometría tridimensional es idéntica a la cuatridimensional, y que la distinción "una es intuitiva, la otra no" no tiene contenido geométrico, sino contenido antropológico.

Similarmente, podemos considerar que $\mathcal{P}\mathbb{N}$ es algo que existe objetivamente en el mismo sentido en que existen las esferas de cuatro dimensiones, algo que cae fuera del alcance de nuestra intuición o, por el contrario, pensar que $\mathcal{P}\mathbb{N}$ es como un personaje de ficción, del que podemos hablar coherentemente pero al que sería ingenuo atribuir una realidad objetiva.

Es crucial que nada de lo dicho hasta ahora ni de lo que diremos en un futuro próximo aporta nada a la hora de decidir cuál es el caso. Tanto si podemos atribuir una realidad objetiva a los objetos matemáticos abstractos como si no es así, la teoría axiomática de conjuntos es como una novela, en la que se dice "los conjuntos cumplen esto y lo otro", pero sin que, por mucho que la leamos, podamos distinguir si se trata de una novela histórica o una novela de ficción. Si es posible encontrar una diferencia, ésta habrá de estar "fuera" de la novela misma.

Los conjuntos hereditariamente finitos Para comprender mejor las sutilezas que hemos estado discutiendo, conviene pensar en una teoría de conjuntos simplificada, una teoría de conjuntos de la que sí tenemos un modelo estándar metamatemático.

Definimos $V_0 = \varnothing$, es decir, el conjunto vacío, una colección de objetos perfectamente definida, desde el momento en que sabemos dar sentido a cualquier afirmación sobre la totalidad de sus elementos (cualquiera de tales afirmaciones es verdadera). En general, para cada número natural n podemos definir $V_{n+1} = \mathcal{P}V_n$, el conjunto de todos los subconjuntos de V_n. Por ejemplo,

$$V_1 = \{\varnothing\}, \quad V_2 = \{\varnothing, \{\varnothing\}\}, \quad V_3 = \{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \{\varnothing, \{\varnothing\}\}\}, \quad \ldots$$

y de este modo podemos explicitar los elementos de cualquier V_n. Más aún, podemos definir HF como la unión de todos los conjuntos V_n. Los elementos de HF se llaman conjuntos *hereditariamente finitos* (porque son conjuntos finitos formados por conjuntos finitos, formados a su vez por conjuntos finitos, etc.) Una afirmación sobre la totalidad de los conjuntos hereditariamente finitos será verdadera si la cumplen todos los elementos de V_0 (esto siempre se cumple), y todos los de V_1, y todos los de V_2, etc. En definitiva, podemos hablar de los conjuntos hereditariamente finitos sin necesidad de una teoría axiomática, sabemos lo que significa cualquier afirmación sobre ellos con independencia de si sabemos decidir si es verdadera o falsa.

Llamaremos *rango* de un conjunto hereditariamente finito x al mínimo natural n tal que x está en V_{n+1}. Así, $\varnothing$ es el único conjunto de rango 0, $\{\varnothing\}$ es el único conjunto de rango 1, $\{\{\varnothing\}\}$ y $\{\varnothing, \{\varnothing\}\}$ son los dos únicos conjuntos de rango 2 y, en general, para $n \geq 1$, hay 2^{n-1} conjuntos de rango n. Podemos afirmar, pues, que hay infinitos conjuntos hereditariamente finitos. Esta afirmación, como todas las anteriores, tiene un significado objetivo claro.

Si a los matemáticos les bastara trabajar con conjuntos hereditariamente finitos, la teoría axiomática de conjuntos sería inútil, y aun inconveniente. En efecto, no necesitamos un cálculo deductivo para determinar qué afirmaciones o deducciones son admisibles si por "conjunto" entendemos algo definido con total precisión, como son los conjuntos hereditariamente finitos. Más aún, no necesitaríamos un cálculo deductivo para convencernos de que no existe el conjunto de todos los conjuntos, no porque esta colección de objetos no esté bien definida (sí que lo está, es HF), sino porque es infinita y, por consiguiente, no es un conjunto (hereditariamente finito). Si todos los teoremas matemáticos hablaran de conjuntos (hereditariamente finitos) a nadie le chocaría que muchos de ellos no fueran aplicables al "conjunto" de todos los conjuntos. En definitiva, las multiplicidades absolutamente infinitas de Cantor serían simplemente las colecciones infinitas de conjuntos hereditariamente finitos, de las que sería injusto decir que no existen o que tienen algo de paradójico o contradictorio. Simplemente no formarían parte de los objetos estudiados por la matemática. Equivalentemente, habría propiedades que definen conjuntos y propiedades que no definen conjuntos. Las primeras serían las propiedades satisfechas por una cantidad finita de conjuntos.

Vamos a construir un lenguaje formal para hablar de HF. La única característica relevante de un conjunto de HF es cuáles son sus elementos,[8] por lo que únicamente necesitamos un signo eventual, un relator diádico que nos permita expresar si un conjunto está o no está en otro conjunto. "Está" en griego se dice $\grave{\varepsilon}\sigma\tau\acute{\iota}$, y por ello Peano eligió la letra épsilon ($\in$) para representar este relator, lo cual se ha convertido en tradición.

Si llamamos $\mathfrak{L}$ al lenguaje formal formado por este relator $\in$ más los signos usuales, tenemos que HF es un modelo de $\mathfrak{L}$ que determina lo que podemos llamar la *interpretación natural* de cada sentencia de $\mathfrak{L}$.

Un ejemplo de sentencia verdadera en su interpretación natural es:

$$\bigwedge xy \bigvee z \bigwedge u(u \in z \leftrightarrow u = x \vee u = y).$$

Esta sentencia significa que, dados dos conjuntos (hereditariamente finitos), existe un tercer conjunto (hereditariamente finito) cuyos elementos son exactamente los dos conjuntos dados. Es fácil ver que es así. Otra sentencia verdadera es

$$\neg \bigvee x \bigwedge y\, y \in x.$$

[8] En realidad otro hecho notable es que estamos hablando únicamente de conjuntos cuyos elementos son otros conjuntos. Ésta es una característica común a todas las teorías de conjuntos usuales. En principio podríamos admitir conjuntos cuyos elementos no fueran conjuntos, pero ello sólo complicaría formalmente la teoría sin aportar ningún beneficio.

Significa que no existe ningún conjunto (hereditariamente finito) que contenga a todos los conjuntos (hereditariamente finitos). Como ya hemos explicado antes, esto se debe simplemente a que hay infinitos conjuntos hereditariamente finitos, luego ninguno de ellos puede contenerlos a todos.

La teoría de conjuntos de Zermelo-Fraenkel La primera teoría axiomática diseñada para evitar las paradojas de la teoría de conjuntos la crearon Whitehead y Russell con el nombre de *Principia Mathematica*. (1910-13) Su lógica subyacente es desproporcionadamente complicada (es lógica de orden infinito) y no presenta ventaja alguna frente a otras teorías mucho más simples que surgieron después. Las primeras teorías razonables fueron diseñadas por Skolem y Fraenkel en los años veinte a partir de unos axiomas publicados por Zermelo[9] en 1908. Su lenguaje formal es el mismo que acabamos de describir en el apartado anterior. De hecho, todos los axiomas de la teoría de Zermelo-Fraenkel (ZF) son verdaderos en el modelo HF excepto uno, el que afirma la existencia de conjuntos infinitos. Si llamamos ZF^- a la teoría que resulta de suprimir este axioma, tenemos una teoría consistente que admite como modelo a HF, al que podemos considerar la *interpretación natural* de ZF^-. Los axiomas de ZF^- afirman cosas como que dos conjuntos son iguales si y sólo si tienen los mismos elementos, existe un conjunto sin elementos, para todo conjunto existe otro que contiene a todos sus subconjuntos, y otros hechos similares.

A partir de estos axiomas pueden probarse teoremas "razonables", como que no existe el conjunto de todos los conjuntos. Sucede que ZF^- es una teoría aritmética recursiva y consistente, por lo que en realidad tiene infinitos modelos (incluso si añadimos como axioma la no existencia de conjuntos infinitos). Por ello decíamos antes que si sólo estuviéramos interesados en los conjuntos hereditariamente finitos, una teoría informal sería preferible a una teoría axiomática, ya que las teorías axiomáticas no nos permiten determinar los objetos de los que hablamos, cosa que sí podemos hacer informalmente.

Puesto que, como decimos, HF no cumple el axioma de infinitud, tenemos que HF no es un modelo de ZF. De hecho, dada la potencia de ZF —en cuyo seno es posible demostrar todos los teoremas matemáticos— los teoremas de incompletitud nos llevan a que no es posible construir metamatemáticamente un modelo de ZF. Asumir los axiomas de ZF es asumir que existen unos objetos que comparten un buen número de propiedades con los conjuntos hereditariamente finitos pero que no son los conjuntos hereditariamente finitos, sino que entre ellos hay conjuntos infinitos. Para ello no basta añadir a nuestro modelo finito unos cuantos conjuntos más, sino que los conjuntos que pueden ser generados una vez suponemos que existe al menos un conjunto infinito son tantos y tan complejos que no podemos abarcar intuitivamente un universo tan vasto.

[9] Zermelo diseño su axiomática con el fin de clarificar la teoría Cantoriana, pero consideraba que la matemática no necesitaba de la lógica formal. Por ello, su noción de "propiedad bien definida" (hoy diríamos simplemente "fórmula") resultaba oscura. La idea de añadir los axiomas de Zermelo a un cálculo de predicados de primer orden con igualador (lo que nosotros hemos llamado una teoría axiomática) se debe a Skolem. Fraenkel sustituyó uno de los axiomas de Zermelo por otro más fuerte.

Aquí se pone de manifiesto la utilidad de las teorías axiomáticas. Aunque no tenemos ningún modelo natural de ZF, eso no resta un ápice de rigor a la teoría. Una vez fijados sus axiomas, los teoremas de ZF son exactamente las sentencias de su lenguaje formal deducibles a partir de dichos axiomas, donde las nociones de "sentencia" y "deducible" han sido totalmente precisadas en los primeros capítulos de este libro. Así, la sentencia

$$\bigwedge xy \bigvee z \bigwedge u(u \in z \leftrightarrow u = x \vee u = y).$$

es un axioma de ZF. Ahora no tenemos un modelo natural en el que interpretarla. Podemos leer igualmente: "dados dos conjuntos, existe un tercero que los contiene a ambos y sólo a ambos". Ahora no sabemos qué queremos decir exactamente con "conjunto". Podemos pensar que existen unos objetos llamados conjuntos que cumplen éste y otros hechos similares (muchos de ellos fuera de nuestra intuición, como hemos explicado antes), o podemos pensar que esta sentencia es análoga a aquella de *"Una mañana, tras un sueño agitado, Gregorio Samsa se despertó convertido en un horrible insecto"*, una frase de una novela de ficción, una frase coherente, pero de la que sería ingenuo pensar que se refiere a algo real. La única forma de "realizar" la metamorfosis de un empleado en insecto sería en una película con efectos especiales, e igualmente podría suceder que ZF tuviera únicamente modelos con "efectos especiales" que fingieran que ciertas colecciones son no numerables cuando en realidad son numerables, tal y como vimos en el capítulo IV que puede hacerse.

Pero lo más importante es que, independientemente de cuál sea el caso, la teoría axiomática de conjuntos ZF está ahí, completamente determinada, inmune a esta polémica, proporcionando un camino —presumiblemente— firme para el desarrollo de la matemática. Veremos que su capacidad es tal que podemos identificar los teoremas admisibles para un matemático con los teoremas demostrables en ZF.

La teoría de conjuntos de von Neumann-Bernays-Gödel La teoría de conjuntos de Frege resultó ser contradictoria porque, para cada fórmula $\phi(x)$ había tomado como axioma la fórmula $\bigvee y \bigwedge x(x \in y \leftrightarrow \phi(x))$, es decir, que para toda propiedad existe un conjunto cuyos elementos son exactamente los conjuntos que cumplen $\phi(x)$. De aquí se sigue la existencia del conjunto de todos los conjuntos o del conjunto de todos los conjuntos que no se pertenecen a sí mismos, ambos contradictorios.

La teoría de Zermelo-Fraenkel evita estas paradojas con una estrategia policial: no hay en ella nada parecido a un axioma general de existencia de conjuntos. En su lugar, hay varios axiomas que, bajo ciertas condiciones, permiten justificar que existen determinados conjuntos. Podríamos decir que "nada existe salvo que se demuestre lo contrario". Esto hace que, por una parte, en ocasiones sea necesario evitar de forma más o menos forzada toda referencia a un cierto conjunto hasta el momento en que queda justificada su existencia, lo cual resulta incómodo. Por otra parte, en otras ocasiones nos vemos inducidos a tratar con colecciones "ilegales". Para distinguirlas de los conjuntos (las colecciones "legales"), a las "ilegales" las llamamos *clases.* Así, decir que la clase de los

espacios vectoriales está contenida en la clase de todos los grupos es una forma alternativa de decir que todo espacio vectorial es un grupo. Es una forma ilegal —pues no existe el conjunto de todos los espacios vectoriales ni el conjunto de todos los grupos—, pero a la vez inofensiva, pues la última afirmación tiene pleno sentido en la teoría. La consecuencia que extraemos es que el régimen de censura de Zermelo-Fraenkel es injusto, pues podría relajarse sin perjuicio para nadie. El problema es determinar hasta qué punto puede relajarse.

La primera teoría de conjuntos "permisiva" fue diseñada por von Neumann, aunque en un lenguaje un poco extraño, pues sus términos primitivos eran el de "función" y el de "argumento". Bernays tradujo esta teoría al lenguaje conjuntista usual y así, la teoría de von Neumann-Bernays contenía dos conceptos básicos (si se quiere, dos relatores monádicos) el de "clase" y el de "conjunto". Las clases son colecciones de conjuntos y la teoría cuenta con un axioma similar al de Frege: para cada fórmula $\phi(x)$ que sólo haga referencia a conjuntos, existe una clase cuyos elementos son los conjuntos que cumplen $\phi(x)$. Por otra parte, la teoría cuenta con otros axiomas que, bajo ciertas hipótesis, permiten probar que una clase dada se corresponde con un conjunto que tiene los mismos elementos. De este modo, las "multiplicidades inconsistentes" de Cantor aparecen reflejadas en la teoría a través de las clases, que nos permiten hablar de la clase de todos los grupos o la clase de todos los espacios vectoriales (y —cómo no—, la clase de todos los conjuntos o la clase de todos los conjuntos que no se pertenecen a sí mismos). Por otra parte, resulta mucho más cómodo probar de forma inmediata que existe una clase a la que poder hacer referencia y después probar con más cuidado que tiene asociado un conjunto.

La relación entre clases y conjuntos, que en la teoría de Bernays era un tanto técnica, fue simplificada notablemente por Gödel. En lo que hoy se conoce como teoría de von Neumann-Bernays-Gödel, los conjuntos son un tipo particular de clases, son clases con "derecho de pertenencia", es decir, se define un conjunto como una clase que pertenece al menos a otra clase. Probar que una clase es un conjunto es probar que "tenemos permiso" para tomarla como elemento de otras clases, esto no depende de ellas misma, sino de los axiomas que especifican bajo qué condiciones podemos hacerlo.

Por supuesto, NBG tiene sus propias multiplicidades inconsistentes, como la clase de todas las clases (que puede probarse que no existe), lo que permitiría hablar de clases de segundo nivel si fuera conveniente, pero en general los matemáticos están interesados en estudiar los conjuntos, y para ello les es útil contar con las clases, mientras que las clases de segundo nivel sólo harían falta si pretendiéramos tratar a las clases como objeto de estudio y no como un concepto auxiliar.[10]

En los capítulos siguientes desarrollaremos con detalle estas ideas, describiremos las dos teorías de conjuntos de las que hemos hablado, veremos la relación entre ellas, su relación con la metamatemática y, en fin, veremos cómo se resuelve —o hasta qué punto— el problema de la fundamentación de la matemática.

[10]En realidad este problema surge en la teoría de categorías, pues sus objetos de estudio son normalmente clases propias, pero hay formas de resolver el problema que sería complicado exponer aquí.

Capítulo VIII

Los axiomas de la teoría de conjuntos

En este capítulo presentaremos las teorías de conjuntos más importantes y veremos cómo a partir de sus axiomas pueden probarse todos los teoremas básicos de las matemáticas, es decir, los resultados que los matemáticos dan por obvios en las demostraciones importantes. De este modo habremos reducido un mar difuso de resultados supuestamente "evidentes" a unos pocos axiomas. Esto es imprescindible porque no podemos tener nada por evidente cuando el concepto central sobre el que giran todas las afirmaciones no está bien definido. Empezamos por la teoría de conjuntos de von Neumann-Bernays-Gödel, donde, como ya hemos explicado en la introducción, la noción básica no es la de conjunto, sino la de clase, lo cual nos permite incorporar a la teoría las "multiplicidades inconsistentes" de Cantor, sin más precaución que restringir los axiomas oportunos a conjuntos, para evitar contradicciones.

8.1 La teoría de conjuntos de von Neumann-Bernays-Gödel

En lo sucesivo $\mathcal{L}$ será un lenguaje formal cuyo único signo eventual será un relator diádico R_1^2. Usaremos las abreviaturas siguientes:

$$(t_1 \in t_2) \equiv R_1^2 t_1 t_2 \qquad (t_1 \notin t_2) \equiv \neg R_1^2 t_1 t_2$$
$$\bigwedge X \in t\ \alpha \equiv \bigwedge X(X \in t \to \alpha) \qquad \bigvee X \in t\ \alpha \equiv \bigvee X(X \in t \wedge \alpha)$$
$$\overset{1}{\bigvee} X \in t\ \alpha \equiv \overset{1}{\bigvee} X(X \in t \wedge \alpha) \qquad \mathrm{cto}\, X \equiv \bigvee Y\, X \in Y$$

Usaremos la palabra "clase" para referirnos a un objeto genérico, es decir, convenimos en que una fórmula de tipo $\bigwedge X\, \alpha$ se lee "toda clase X cumple α", etc. La última fórmula que hemos definido, $\mathrm{cto}\, X$, se lee "X es un conjunto". Así pues, hemos definido los conjuntos como las clases que pertenecen al menos

a otra clase. Naturalmente esto es un recurso técnico sin más finalidad que la de simplificar lógicamente la teoría. La distinción entre clases y conjuntos no está contenida en esta definición, sino en los axiomas que veremos después y que determinarán qué propiedades postulamos de las clases en general y cuáles de los conjuntos en particular. Llamaremos *clases propias* a las clases que no son conjuntos.

Nombraremos las variables de $\mathcal{L}$ con letras mayúsculas y usaremos las minúsculas para referirnos a conjuntos, es decir

$$\bigwedge x\ \alpha \equiv \bigwedge X(\mathrm{cto}\, X \to \alpha) \qquad \bigvee x\ \alpha \equiv \bigvee X(\mathrm{cto}\, X \wedge \alpha)$$
$$\bigwedge^1 x\ \alpha \equiv \bigwedge^1 X(\mathrm{cto}\, X \wedge \alpha) \qquad x|\alpha \equiv X|(\mathrm{cto}\, X \wedge \alpha)$$

Usaremos también las abreviaturas siguientes:

$\{Y_1, \ldots, Y_n\} \equiv Z|\bigwedge u(u \in Z \leftrightarrow u = Y_1 \vee \cdots \vee u = Y_n)$

$(Y_1) \equiv Y_1 \quad (Y_1, Y_2) \equiv \{\{Y_1\}, \{Y_1, Y_2\}\} \quad (Y_1, \ldots, Y_n) \equiv ((Y_1, \ldots, Y_{n-1}), Y_n)$

(La última definición la damos para $n \geq 3$, pero hemos de observar que es trivialmente válida si $n = 2$.)

Un $X \equiv \bigwedge uvw((u, v) \in X \wedge (u, w) \in X \to v = w)$.

Nota No debemos caer en el error de valorar estas definiciones en más de lo que son. Por ejemplo, para demostrar que $X \in \{X, Y\}$ no nos basta la definición de $\{X, Y\}$, sino que necesitamos algún axioma que nos garantice que $\{X, Y\}$ es una descripción propia.

Definición 8.1 Llamaremos *teoría de conjuntos básica de von Neumann-Bernays-Gödel* a la teoría NBG* determinada por los axiomas siguientes:

NBG-1	$\bigwedge XY(\bigwedge u(u \in X \leftrightarrow u \in Y) \to X = Y)$	extensionalidad
NBG-2	$\bigwedge XY\bigvee Z\bigwedge u(u \in Z \leftrightarrow u \in X \wedge u \in Y)$	intersección
NBG-3	$\bigwedge X\bigvee Y\bigwedge u(u \in Y \leftrightarrow u \notin X)$	complemento
NBG-4	$\bigwedge uv\bigvee y\bigwedge x(x \in y \leftrightarrow x = u \vee x = v)$	par
NBG-5	$\bigvee A\bigwedge xy((x, y) \in A \leftrightarrow x \in y)$	pertenencia
NBG-6	$\bigwedge A\bigvee B\bigwedge x(x \in B \leftrightarrow \bigvee y(x, y) \in A)$	dominio
NBG-7	$\bigwedge A\bigvee B\bigwedge xy((x, y) \in B \leftrightarrow x \in A)$	prod. cartesiano
NBG-8	$\bigwedge A\bigvee B\bigwedge xy((x, y) \in B \leftrightarrow (y, x) \in A)$	relación inversa
NBG-9	$\bigwedge A\bigvee B\bigwedge xyz((x, y, z) \in B \leftrightarrow (y, z, x) \in A)$	permutación
NBG-10	$\bigwedge A\bigvee B\bigwedge xyz((x, y, z) \in B \leftrightarrow (x, z, y) \in A)$	permutación
NBG-11	$\bigvee x\bigwedge y\, y \notin x$	conjunto vacío
NBG-12	$\bigwedge x\bigvee y\bigwedge uv(u \in v \wedge v \in x \to u \in y)$	unión
NBG-13	$\bigwedge xA(\mathrm{Un}A \to \bigvee y\bigwedge u(u \in y \leftrightarrow \bigvee v \in x\,(v, u) \in A))$	reemplazo

Éstos no son todos los axiomas de la teoría NBG, sino que son sólo los necesarios para que las clases y los conjuntos se comporten como esperamos. Después faltará añadir algunos más para redondear la teoría, por ejemplo el que postula la existencia de conjuntos infinitos. Veamos ahora las primeras consecuencias de estos axiomas.

El axioma NBG-1 es el axioma de extensionalidad, que afirma que si dos clases tienen los mismos elementos entonces son iguales (el recíproco es un teorema lógico). La colección de elementos de una clase se conoce como su *extensión,* por lo que el axioma de extensionalidad afirma que dos clases son iguales si y sólo si tienen la misma extensión o, dicho de otro modo, que podemos identificar cada clase con su extensión y, en definitiva, que las clases pueden ser consideradas como colecciones de objetos. Sin embargo, nada más lejos de la realidad que el recíproco: admitir que toda colección de objetos —o incluso de conjuntos— determina una clase nos lleva directamente a un sin fin de contradicciones. Por ello necesitamos unos axiomas que postulen la existencia de clases que cumplan determinados requisitos: han de ser lo suficientemente permisivos como para que los matemáticos puedan "creer" que toda colección determina una clase (o, dicho de otro modo, para que todas las colecciones que les aparecen a ellos sean clases), y lo suficientemente restrictivos para que no den lugar a contradicciones.

Los dos primeros axiomas de formación de clases son NBG-2 y NBG-3. Uno postula la existencia de la intersección de dos clases, es decir, la de una tercera clase que contiene exactamente a los elementos comunes a ambas, y el otro la existencia de la clase complementaria de una clase dada, es decir, una clase que contiene exactamente a los conjuntos que no están en la clase dada. El axioma de extensionalidad nos da que la intersección y el complemento son únicos, pues dos intersecciones o dos complementos de las mismas clases tendrían los mismos elementos. A partir de estos dos axiomas se puede probar la existencia de muchas otras clases. Esbozamos las pruebas en el apartado siguiente.

1) El álgebra de las clases Las fórmulas siguientes son teoremas de NBG* o bien convenios de notación.

1)	$\bigwedge XY\overset{1}{\bigvee}Z\bigwedge u(u\in Z\leftrightarrow u\in X\wedge u\in Y)$	(Por NBG-1, NBG-2)
2)	$X\cap Y\equiv Z\vert\bigwedge u(u\in Z\leftrightarrow u\in X\wedge u\in Y)$	X intersección Y
3)	$\bigwedge XYu(u\in X\cap Y\leftrightarrow u\in X\wedge u\in Y)$	(Por 1, 2)
4)	$\bigwedge X\overset{1}{\bigvee}Y\bigwedge u(u\in Y\leftrightarrow u\notin X)$	(Por NBG-1, NBG-3)
5)	$\overline{X}\equiv Y\vert\bigwedge u(u\in Y\leftrightarrow u\notin X)$	Complemento de X
6)	$\bigwedge Xu(u\in\overline{X}\leftrightarrow u\notin X)$	(Por 4, 5)
7)	$\bigwedge XY\overset{1}{\bigvee}Z\bigwedge u(u\in Z\leftrightarrow u\in X\vee u\in Y)$	($Z=\overline{\overline{X}\cap\overline{Y}}$+ NBG-1)
8)	$X\cup Y\equiv Z\vert\bigwedge u(u\in Z\leftrightarrow u\in X\vee u\in Y)$	X unión Y
9)	$\bigwedge XYu(u\in X\cup Y\leftrightarrow u\in X\vee u\in Y)$	(Por 7, 8)

10) $\bigwedge XY\overset{1}{\bigvee}Z\bigwedge u(u \in Z \leftrightarrow u \in X \wedge u \notin Y)$ ($Z = X \cap \overline{Y}$+NBG-1)

11) $X \setminus Y \equiv Z|\bigwedge u(u \in Z \leftrightarrow u \in X \wedge u \notin Y)$ X menos Y

12) $\bigwedge XY u(u \in X \setminus Y \leftrightarrow u \in X \wedge u \notin Y)$ (Por 10, 11)

13) $\overset{1}{\bigvee}X\bigwedge u\, u \notin X$ ($X = Y \cap \overline{Y}$+NBG-1)

14) $\varnothing \equiv X|\bigwedge u\, u \notin X$ Clase vacía

15) $\bigwedge X(\bigwedge u\, u \notin X \leftrightarrow X = \varnothing)$ (Por 13, 14)

16) $\overset{1}{\bigvee}X\bigwedge u\, u \in X$ ($X = \overline{\varnothing}$+NBG-1)

17) $\mathrm{V} \equiv X|\bigwedge u\, u \in X$ Clase universal

18) $\bigwedge u\, u \in \mathrm{V}$ (Por 15, 16)

Definimos $X \subset Y \equiv \bigwedge u(u \in X \to u \in Y)$.

Observaciones Conviene incidir en la lógica subyacente a estos resultados. Por ejemplo, 3) se demuestra aplicando a 1) la regla de las descripciones propias, es decir, la intersección se define como $Z \mid \alpha$, 1) prueba que $\overset{1}{\bigvee}Z\,\alpha$, DP nos da entonces $\mathtt{S}_Z^{Z|\alpha}\alpha$, y esto es 3). Lo mismo vale en los demás casos.

En la prueba de 13 se usa que existe al menos una clase. De acuerdo con nuestra formalización de la lógica esto es un teorema lógico (podemos probar $\bigvee X\, X = X$). En otras variantes del cálculo deductivo la existencia de objetos no es un teorema, pero eso no afecta en este contexto pues el axioma NBG-11 afirma de hecho que existe la clase vacía y además es un conjunto.

Todas las propiedades básicas del álgebra de clases se demuestran ahora de forma elemental. Por ejemplo, $X \cap Y \subset X$, $X \cup Y = Y \cup X$, etc.

2) n-tuplas desordenadas y ordenadas El axioma NBG-4 asegura la existencia de un conjunto que tiene por elementos a cualquier par de conjuntos dados (no necesariamente distintos). De él se sigue la existencia de n-tuplas desordenadas y ordenadas. Las fórmulas siguientes son teoremas de NBG* o convenios de notación.

1) $\bigwedge uv\overset{1}{\bigvee}Y\bigwedge x(x \in Y \leftrightarrow x = u \vee x = v)$ (NBG-1, NBG-4)

2) $\bigwedge uvx(x \in \{u, v\} \leftrightarrow x = u \vee x = v)$ (Por 1)

3) $\bigwedge uv\,\mathrm{cto}\{u, v\}$ (Por 1, 2, NBG-4)

4) $\bigwedge X\overset{1}{\bigvee}Y\bigwedge u(u \in Y \leftrightarrow u = X)$ ($Y = \{X, X\}$ si $\mathrm{cto}\, X$, $Y = \varnothing$ si $\neg\,\mathrm{cto}\, X$)

5) $\bigwedge Xu(u \in \{X\} \leftrightarrow u = X)$ (Por 4)

6) $\bigwedge X_1 \cdots X_n\, u(u \in \{X_1\} \cup \cdots \cup \{X_n\} \leftrightarrow u = X_1 \vee \cdots \vee u = X_n)$ (Por inducción)

7) $\bigwedge X_1 \cdots X_n \overset{1}{\bigvee} Y \bigwedge u(u \in Y \leftrightarrow u = X_1 \vee \cdots \vee u = X_n)$
(Por 6 y NBG-1)

8) $\bigwedge X_1 \cdots X_n\, u(u \in \{X_1, \ldots, X_n\} \leftrightarrow u = X_1 \vee \cdots \vee u = X_n)$
(Por 7)

9) $\bigwedge u\, \mathrm{cto}\{u\}$ (Por 3, $\{u\} = \{u, u\}$)

10) $\bigwedge uv\, \mathrm{cto}(u, v)$ (Por 3, 9)

11) $\bigwedge uvwz((u, v) = (w, z) \leftrightarrow u = w \wedge v = z)$ (Rutina)

12) $\bigwedge x_1 \cdots x_n\, \mathrm{cto}(x_1, \ldots, x_n)$ (Por inducción)

13) $\bigwedge x_1 \cdots x_n\, y_1 \cdots y_n((x_1, \ldots, x_n) = (y_1, \ldots, y_n) \leftrightarrow x_1 = y_1 \wedge \cdots \wedge x_n = y_n)$
(Por inducción)

14) $\bigwedge x_1 \cdots x_{n+p}((x_1, \ldots, x_n), x_{n+1}, \ldots, x_{n+p}) = (x_1, \ldots, x_{n+p}))$
(Por inducción)

Observaciones 2) y 3) desarrollan el axioma NBG-4, 4)–8) introducen un convenio técnico: la clase $\{X_1, \ldots, X_n\}$ tiene por elementos a aquellas clases de entre $X_1, \ldots, X_n$ que sean conjuntos, pero se reduce a la clase vacía si ninguna lo es. En cualquier caso, lo cierto es que $\{X_1, \ldots, X_n\}$ tiene sentido. Naturalmente, la prueba de 6) es una inducción metamatemática sobre el número de variables. Hemos de tener presente que 6), 7) y 8) no son teoremas, sino esquemas que recogen infinitos teoremas, uno para cada valor de n. Los últimos teoremas muestran que las n-tuplas ordenadas se comportan como cabe esperar, es decir, $(x_1, \ldots, x_n)$ es un conjunto que determina y está determinado por sus componentes $x_1, \ldots, x_n$ teniendo en cuenta el orden.

Fórmulas primitivas y normales Según explicábamos en la introducción, la ventaja principal de la teoría NBG es que en ella es posible demostrar que toda fórmula $\phi(x)$ que habla únicamente de conjuntos determina una clase cuyos elementos son exactamente los conjuntos que cumplen $\phi(x)$. En realidad la condición que hemos de imponerle a $\phi(x)$ es ligeramente más general: puede hacer referencia a clases propias, pero no debe contener ninguna cuantificación de tipo "para toda clase" o "existe una clase". Veamos en qué consiste exactamente esta condición.

Definición 8.2 Una fórmula de $\mathfrak{L}$ es *primitiva* si está construida según las reglas siguientes:

- $X \in Y$ es una fórmula primitiva.
- $\neg\alpha$, $\alpha \to \beta$, $\bigwedge X(\mathrm{cto}\, X \to \alpha)$ son primitivas si α y β lo son.

Una fórmula de $\mathfrak{L}$ es *normal* si es equivalente a una fórmula primitiva. El concepto de fórmula normal es relativo a la teoría en que trabajemos: puede haber fórmulas que no sean normales en NBG* y sí lo sean en alguna extensión suya. De momento, normalidad significará normalidad en NBG*.

Claramente $\neg\alpha$, $\alpha \to \beta$, $\alpha \vee \beta$, $\alpha \wedge \beta$, $\alpha \leftrightarrow \beta$, $\bigwedge x\alpha$, $\bigvee x\alpha$ son normales si α y β lo son.

Por NBG-1 se cumple que $X = Y \leftrightarrow \bigwedge u(u \in X \leftrightarrow u \in Y)$, luego $X = Y$ es normal. Así mismo, cto $X \leftrightarrow \bigvee y\, X = y$, luego cto X es normal.

En definitiva, una fórmula es normal si no involucra cuantificaciones sobre clases propias. Así, una fórmula que hable exclusivamente de conjuntos —que es el caso más habitual— es normal.

Un término t es *normal* si la fórmula $x \in t$ es normal, donde x es cualquier variable que no esté en t. Claramente, las variables son términos normales. Además se cumple el teorema siguiente:

Teorema 8.3 *Si $\alpha(X)$ y t son normales, entonces $\alpha(t)$ es normal (donde la fórmula α puede tener libres variables cualesquiera, no sólo X).*

Demostración: Por definición $\alpha(X) \leftrightarrow \beta(X)$, donde $\beta(X)$ es una fórmula primitiva. Entonces $\alpha(t) \leftrightarrow \beta(t)$. Basta ver que $\beta(t)$ es normal. Lo haremos por inducción sobre la longitud de β.

- Si $\beta \equiv U \in V$ distinguimos cuatro casos:
  - Ni U ni V son X. Entonces $\beta(t) \equiv \beta$, luego es primitiva.
  - $\beta \equiv X \in V$, donde $V \not\equiv X$. Entonces
    $$\beta(t) \equiv t \in V \leftrightarrow \bigvee y(y = t \wedge y \in V)$$
    $$\leftrightarrow \bigvee y(\bigwedge u(u \in t \leftrightarrow u \in y) \wedge y \in V),$$
    que es normal por serlo t.
  - $\beta \equiv X \in X$. Entonces
    $$\beta(t) \equiv t \in t \leftrightarrow \bigvee y(y = t \wedge y \in t) \leftrightarrow \bigvee y(\bigwedge u(u \in t \leftrightarrow u \in y) \wedge y \in t),$$
    y la última fórmula es normal.
  - $\beta \equiv U \in X$, donde $U \not\equiv X$. Entonces sea $\gamma(Y)$ una fórmula primitiva equivalente a $Y \in t$. Así $\beta(t) \equiv U \in t \leftrightarrow \gamma(U)$ y ésta es primitiva.
- Si $\beta \equiv \neg\gamma$, por hipótesis de inducción $\gamma(t)$ es normal, luego $\neg\gamma(t)$ también.
- Si $\beta \equiv \gamma \to \delta$, por hipótesis de inducción $\gamma(t)$ y $\delta(t)$ son normales, luego $\beta(t) \equiv \gamma(t) \to \delta(t)$ también lo es.
- Si $\beta \equiv \bigwedge u\gamma$ podemos, si es necesario, cambiar la variable u por otra que no esté en t (pues la fórmula resultante es equivalente). Así, $\beta(t) \equiv \bigwedge u\, \gamma(t)$, que es normal por serlo $\gamma(t)$. ∎

De este teorema se sigue que si $t_1(X)$ y t_2 son términos normales, entonces $t_1(t_2)$ es normal, pues $y \in t_1(X)$ es normal, $y \in t_1(t_2)$ también lo es y, por definición, $t_1(t_2)$ también lo es.

Con estos resultados es fácil probar la normalidad de cualquier expresión que ciertamente lo sea.

Ejercicio Probar que $X \subset Y$, $X \cap Y$, $X \cup Y$, $X \setminus Y$, $\{X_1, \ldots, X_n\}$, $(x_1, \ldots, x_n)$ son normales.

3) Clases de n-tuplas El lector debería ahora observar el enunciado del teorema 8.5 y su variante 8.6. Se trata del teorema general que estamos persiguiendo, capaz de convencer a todos los matemáticos de que cualquier colección de conjuntos es una clase (esto es falso, lo que se cumple es que cualquier colección de conjuntos definible mediante una fórmula que involucre únicamente a conjuntos es una clase, pero la diferencia es demasiado sutil para que incomode a un matemático.) Los axiomas NBG-5–NBG-10 son casos particulares de este teorema y —como vamos a ver— junto con los axiomas que ya hemos tratado, bastan para demostrar el caso general. Para ello necesitamos extraer algunas consecuencias de este grupo de axiomas.

Las fórmulas siguientes son teoremas de NBG* o convenios de notación.

1) $\bigwedge A \overset{1}{\bigvee} B \bigwedge x(x \in B \leftrightarrow \bigvee y(x, y) \in A)$ (NBG-1, NBG-6)

2) $\mathcal{D}A \equiv B | \bigwedge x(x \in B \leftrightarrow \bigvee y(x, y) \in A)$ Dominio de A

3) $\bigwedge Ax(x \in \mathcal{D}A \leftrightarrow \bigvee y(x, y) \in A)$ (Por 1, 2)

4) $\bigwedge A \overset{1}{\bigvee} B \bigwedge x(x \in B \leftrightarrow \bigvee y(y, x) \in A)$ (Por 1 y NBG-8)

5) $\mathcal{R}A \equiv B | \bigwedge x(x \in B \leftrightarrow \bigvee y(y, x) \in A)$ Rango de A

6) $\bigwedge Ax(x \in \mathcal{R}A \leftrightarrow \bigvee y(y, x) \in A)$ (Por 4, 5)

7) $\bigwedge A \bigvee B \bigwedge xy((x, y) \in B \leftrightarrow y \in A)$ (NBG-7,NBG-8)

8) a) $\bigwedge A \bigvee B \bigwedge xyz((x, y, z) \in B \leftrightarrow (x, y) \in A)$

   b) $\bigwedge A \bigvee B \bigwedge xyz((x, z, y) \in B \leftrightarrow (x, y) \in A)$

   c) $\bigwedge A \bigvee B \bigwedge xyz((z, x, y) \in B \leftrightarrow (x, y) \in A)$

($\bigwedge A \bigvee B \bigwedge wz((w, z) \in B \leftrightarrow w \in A)$ por NBG-7. Haciendo $w = (x, y)$ tenemos a) y, aplicando NBG-9, NBG-10 se siguen b) y c).)

9) $\bigwedge A \bigvee B \bigwedge x_1 \cdots x_n\, y\, ((x_1, \ldots, x_n, y) \in B \leftrightarrow (x_1, \ldots, x_n) \in A)$

   Por NBG-7, haciendo $x = (x_1, \ldots, x_n)$.

10) $\bigwedge A \bigvee B \bigwedge x_1 \cdots x_n\, y_1 \cdots y_k((x_1, \ldots, x_n, y_1, \ldots, y_k) \in B \leftrightarrow (x_1, \ldots, x_n) \in A)$

    (Por 9 aplicado k veces.)

11) $\bigwedge A \bigvee B \bigwedge x_1 \cdots x_n\, y((x_1, \ldots, x_{n-1}, y, x_n) \in B \leftrightarrow (x_1, \ldots, x_n) \in A)$

    (Por 8 b.)

12) $\bigwedge A \bigvee B \bigwedge x_1 \cdots x_n\, y_1 \cdots, y_k((x_1, \ldots, x_{n-1}, y_1, \ldots y_k, x_n) \in B \leftrightarrow (x_1, \ldots, x_n) \in A)$

    (Por 11 aplicado k veces.)

13) $\bigwedge A \bigvee B \bigwedge y_1 \cdots y_k\, x_1\, x_2((y_1, \ldots, y_k, x_1, x_2) \in B \leftrightarrow (x_1, x_2) \in A)$

    (Por 8 c.)

Notemos que a partir de 7) no tenemos unicidad, pues, por ejemplo, puede haber dos clases que cumplan lo que 7) pide a B pero que difieran en elementos que no sean pares ordenados.

Formación de clases Ahora ya podemos encaminarnos a demostrar el teorema 8.5. El núcleo de la prueba está en el teorema siguiente:

Teorema 8.4 *Consideremos una fórmula primitiva $\phi(x_1,\dots,x_n)$. Entonces*

$$\underset{NBG^*}{\vdash} \bigvee A \bigwedge x_1\cdots x_n((x_1,\dots,x_n)\in A \leftrightarrow \phi(x_1,\dots,x_n))$$

Demostración: Sean $Y_1,\dots,Y_m$ las variables libres distintas de $x_1,\dots,x_n$ que haya en ϕ. Podemos suponer que ϕ tiene libre alguna de las variables x_i, pues de lo contrario sirve $A = \mathrm{V}$ o $A = \varnothing$. Así mismo podemos suponer que las variables Y_i no aparecen nunca como primer término de un relator de pertenencia, pues $Y_i \in t$ puede sustituirse por la fórmula $\bigvee x(x = Y_i \wedge x = t)$ y a su vez esto se sustituye por

$$\neg\bigwedge x\neg(\bigwedge u(u\in x \leftrightarrow u\in Y_i)\wedge x\in t),$$

que sigue siendo primitiva. Similarmente podemos exigir que no haya subfórmulas de tipo $X \in X$. Sea, pues, ϕ una fórmula primitiva en estas condiciones. Demostraremos el teorema por inducción sobre la longitud de ϕ.

a) Si $\phi \equiv x_r \in x_s$ entonces $r < s$ o bien $s < r$. Según el caso obtenemos

$$\bigvee F\bigwedge x_r x_s((x_r,x_s)\in F \leftrightarrow x_r\in x_s) \qquad \text{(por NBG-5)}$$
$$\bigvee F\bigwedge x_r x_s((x_s,x_r)\in F \leftrightarrow x_r\in x_s) \qquad \text{(por NBG-5, NBG-8)}$$

En otros términos, si llamamos p al menor de r, s y q al mayor de ambos, se cumple

$$\bigvee F\bigwedge x_p x_q((x_p,x_q)\in F \leftrightarrow \phi(x_1,\dots,x_n)).$$

Si $p \neq 1$ aplicamos[1] 3.13 para concluir

$$\bigvee F_1\bigwedge x_1\cdots x_p\, x_q((x_1,\dots,x_p,x_q)\in F_1 \leftrightarrow \phi(x_1,\dots,x_n)).$$

Si $q \neq p+1$ aplicamos 3.12 para concluir

$$\bigvee F_2\bigwedge x_1\cdots x_q((x_1,\dots,x_q)\in F_2 \leftrightarrow \phi(x_1,\dots,x_n)).$$

Si $q \neq n$ aplicamos 3.10 y obtenemos

$$\bigvee A\bigwedge x_1\cdots x_n((x_1,\dots,x_n)\in A \leftrightarrow \phi(x_1,\dots,x_n)).$$

[1]Esta referencia y las que siguen remiten a los teoremas del apartado 3 de esta misma sección.

b) Si $\phi \equiv x_r \in Y_k$ distinguimos el caso $n = r = 1$, que es trivial, pues $\bigwedge x_1((x_1) \in Y_k \leftrightarrow x_1 \in Y_k)$, luego $\bigvee A \bigwedge x_1((x_1) \in A \leftrightarrow x_1 \in Y_k)$.

Si $n \neq 1$, o bien $r \neq n$, y entonces por NBG-7

$$\bigvee A \bigwedge x_r x_{r+1}((x_r, x_{r+1}) \in A \leftrightarrow x_r \in Y_k),$$

o bien $r = n$, y entonces aplicamos 3.7:

$$\bigvee A \bigwedge x_{r-1} x_r((x_{r-1}, x_r) \in A \leftrightarrow x_r \in Y_k).$$

En ambos casos, aplicando 3.13 y 3.10 llegamos a

$$\bigvee A \bigwedge x_1 \cdots x_n((x_1, \ldots, x_n) \in A \leftrightarrow \phi(x_1, \ldots, x_n)).$$

c) Si $\phi \equiv \neg\psi$, $\phi \equiv \chi \to \psi$, o bien $\phi \equiv \bigwedge x\, \theta$, en el último caso podemos suponer que $x \not\equiv x_i$ para todo i. Por hipótesis de inducción

$$\bigvee B \bigwedge x_1 \cdots x_n((x_1, \ldots, x_n) \in B \leftrightarrow \psi(x_1, \ldots, x_n)),$$
$$\bigvee C \bigwedge x_1 \cdots x_n((x_1, \ldots, x_n) \in C \leftrightarrow \chi(x_1, \ldots, x_n)),$$
$$\bigvee D \bigwedge x_1 \cdots x_n\, x((x_1, \ldots, x_n, x) \in D \leftrightarrow \theta(x_1, \ldots, x_n, x)).$$

Si $\phi \equiv \neg\psi$ tenemos $\bigvee A \bigwedge x_1 \cdots x_n((x_1, \ldots, x_n) \in A \leftrightarrow \phi(x_1, \ldots, x_n))$ sin más que tomar $A = \overline{B}$. Si $\phi \equiv \chi \to \psi$ llegamos a la misma conclusión tomando $A = \overline{C} \cup D$ y si $\phi \equiv \bigwedge x\, \theta$ damos varios pasos:

$$\bigvee E \bigwedge x_1 \cdots x_n((x_1, \ldots, x_n) \in E \leftrightarrow \bigvee x\, (x_1, \ldots, x_n, x) \in \overline{D}) \quad (E = \mathcal{D}\overline{D})$$
$$\bigvee A \bigwedge x_1 \cdots x_n((x_1, \ldots, x_n) \in A \leftrightarrow \neg \bigvee x\, (x_1, \ldots, x_n, x) \in \overline{D})\ (A = \overline{E}),$$

es decir, $\bigvee A \bigwedge x_1 \cdots x_n((x_1, \ldots, x_n) \in A \leftrightarrow \bigwedge x\, \theta(x_1, \ldots, x_n, x))$, como queríamos probar. ∎

En el teorema anterior no tenemos unicidad en A porque puede haber otras clases con las mismas n-tuplas que cumplan lo mismo. Si nos restringimos a una variable el axioma de extensionalidad nos da la unicidad. Además, pasando a una fórmula primitiva equivalente, es claro que el teorema vale para fórmulas normales cualesquiera.

Teorema 8.5 (Teorema general de formación de clases) *Si $\phi(x)$ es una fórmula normal (con cualesquiera variables libres)*

$$\underset{NBG^*}{\vdash} \overset{1}{\bigvee} Y \bigwedge x(x \in Y \leftrightarrow \phi(x)).$$

Definimos

$$\{x \mid \phi(x)\} \equiv Y|\bigwedge x(x \in Y \leftrightarrow \phi(x)).$$

De este modo, aplicando la regla de las descripciones propias al teorema anterior, concluimos:

Teorema 8.6 *Si $\phi(x)$ es una fórmula normal*

$$\underset{NBG^*}{\vdash} \bigwedge x(x \in \{x \mid \phi(x)\} \leftrightarrow \phi(x)).$$

A su vez, de aquí podemos obtener clases formadas únicamente por n-tuplas, para las que también tenemos unicidad. Concretamente, definimos

$$\{(x_1,\dots,x_n) \mid \phi(x_1,\dots,x_n)\} \equiv \{x \mid \bigvee x_1\cdots x_n\,(x = (x_1,\dots,x_n) \wedge \phi(x_1,\dots,x_n))\}.$$

Claramente, si $\phi(x_1,\dots,x_n)$ es una fórmula normal,

$$\underset{\mathrm{NBG}^*}{\vdash} \bigwedge x(x \in \{(x_1,\dots,x_n) \mid \phi(x_1,\dots,x_n)\}$$

$$\leftrightarrow \bigvee x_1\cdots x_n\,(x = (x_1,\dots,x_n) \wedge \phi(x_1,\dots,x_n)),$$

y también

$$\underset{\mathrm{NBG}^*}{\vdash} \bigwedge x_1\cdots x_n((x_1,\dots,x_n) \in \{(x_1,\dots,x_n) \mid \phi(x_1,\dots,x_n)\} \leftrightarrow \phi(x_1,\dots,x_n)).$$

Nota Es claro que si $\phi(x)$ es una fórmula normal, entonces $\{x \mid \phi(x)\}$ es un término normal. Notemos también que para que $\{x \mid \phi(x)\}$ esté bien definida (es decir, para que sea una descripción propia) no es necesario que $\phi(x)$ sea normal. Basta con que lo sea $\mathrm{cto}\,x \wedge \phi(x)$, pues ambas fórmulas definen la misma clase.

Definimos el *producto cartesiano* de n clases como

$$Y_1 \times \cdots \times Y_n \equiv \{(x_1,\dots,x_n) \mid x_1 \in Y_1 \wedge \cdots \wedge x_n \in Y_n\}.$$

El lector deber ahora ojear el apéndice A, donde hemos recogido los conceptos y resultados básicos de la teoría de conjuntos que sin duda ya conoce, y comprobar que todos ellos son claramente[2] teoremas de NBG*. En lo que sigue los usaremos siempre que convenga.

La teoría de conjuntos de Morse-Kelley Ahora es claro que NBG* admite la siguiente axiomatización alternativa:

1) $\bigwedge XY(\bigwedge u(u \in X \leftrightarrow u \in Y) \to X = Y)$ (extensionalidad)

2) Para toda fórmula primitiva $\phi(x)$ (donde Y no está libre)
$\bigvee Y \bigwedge x(x \in Y \leftrightarrow \phi(x))$ (formación de clases)

3) $\bigwedge uv \bigvee y \bigwedge x(x \in y \leftrightarrow x = u \vee x = v)$ (par)

4) $\bigvee x \bigwedge y\, y \notin x$ (conjunto vacío)

5) $\bigwedge x \bigvee y \bigwedge uv(u \in v \wedge v \in x \to u \in y)$ (unión)

6) $\bigwedge xA(\mathrm{Un}A \to \bigvee y \bigwedge u(u \in y \leftrightarrow \bigvee v \in x\,(v,u) \in A))$ (reemplazo)

[2]Los resultados sobre conjuntos cociente requieren algunos hechos adicionales que discutiremos en el apartado sobre formación de conjuntos, más abajo.

En efecto, partiendo del esquema de formación de clases como axioma, tenemos inmediatamente los teoremas 8.5 y 8.6, los cuales nos permiten a su vez demostrar la existencia de las clases que postulan los axiomas que hemos eliminado. Por ejemplo, la intersección y la unión de dos clases puede definirse ahora como

$$X \cap Y = \{u \mid u \in X \wedge u \in Y\}, \qquad X \cup Y = \{u \mid u \in X \vee u \in Y\}.$$

La clase vacía y la clase universal admiten ahora las definiciones duales:

$$\varnothing = \{x \mid x \neq x\}, \qquad \mathrm{V} = \{x \mid x = x\}.$$

Ésta es la forma más habitual y más simple de presentar NBG*. Al presentarla como lo hemos hecho hemos probado algo nada obvio desde este otro punto de vista: que NBG* es finitamente axiomatizable. Concretamente, lo que hemos probado es que bastan unos pocos casos particulares del esquema de formación de clases para demostrar a partir de ellos todos los demás. Vista así, la restricción a fórmulas primitivas del esquema de formación de clases parece artificial. Más adelante veremos que tiene una interpretación muy importante, pero lo cierto es que nada nos impide eliminarla:

Se llama *teoría de conjuntos de Morse-Kelley* (básica) a la teoría axiomática MK* cuyos axiomas son los anteriores eliminando la restricción a fórmulas primitivas en el esquema axiomático 2).

Obviamente todo teorema de NBG* lo es de MK*. Veremos que si ambas teorías son consistentes, el recíproco no es cierto. Además puede probarse que MK* no es finitamente axiomatizable.

Formación de conjuntos El teorema de formación de clases (o axioma, si se prefiere) nos permite introducir como clases todas las colecciones de conjuntos que un matemático puede necesitar, pero para esto sirva de algo nos falta una serie de teoremas que garanticen que la mayoría de estas clases son de hecho conjuntos. Estos teoremas se siguen de los axiomas de formación de conjuntos. Ya hemos estudiado uno de ellos: el axioma del par, del cual hemos deducido que las n-tuplas desordenadas y ordenadas son conjuntos. Los tres últimos axiomas de NBG* son también axiomas de formación de conjuntos, así como varios de los axiomas que faltan para completar la teoría de NBG.

Por ejemplo, NBG-11 afirma que la clase vacía es un conjunto. De hecho, es el único de los axiomas de NBG* que postula la existencia de un conjunto. Si lo eliminásramos, un modelo de la teoría resultante estaría formado por una única clase sin elementos, que sería a la vez la clase vacía y la clase universal. Por el contrario, a partir del axioma del conjunto vacío podemos probar la existencia de infinitos conjuntos:

$$\varnothing, \quad \{\varnothing\}, \quad \{\{\varnothing\}\}, \quad \ldots$$

En general, podemos probar la existencia de cualquier conjunto hereditariamente finito.

El axioma más importante de formación de conjuntos es el axioma de reemplazo NBG-13. La forma en que lo hemos enunciado se debe a que es conveniente que la estructura lógica de los axiomas sea la más simple posible, pero a la hora de trabajar con él conviene reformularlo en términos más prácticos. Observemos que si $F : A \longrightarrow B$ es una aplicación suprayectiva, entonces F satisface la definición de clase unívoca que aparece en la hipótesis de NBG-13. Si además suponemos que A es un conjunto podemos concluir que

$$\bigvee y \bigwedge u(u \in y \leftrightarrow \bigvee v \in A\, F(v) = u),$$

pero, como F es suprayectiva, esto equivale claramente a $\bigvee y\, y = B$ o, dicho de otro modo, a que B sea un conjunto. En definitiva tenemos:

Teorema 8.7 *Si $F : A \longrightarrow B$ es una aplicación suprayectiva y A es un conjunto, entonces B también es un conjunto.*

Éste es esencialmente el contenido del axioma de reemplazo: si podemos "cubrir" los elementos de una clase B con las imágenes de los elementos de un conjunto A mediante una aplicación, entonces B es un conjunto. En general, cuando citemos el axioma del reemplazo nos referiremos al teorema anterior.

Como primera aplicación tenemos:

Teorema 8.8 *Si A es un conjunto y $B \subset A$, entonces B es un conjunto.*

Demostración: Si $B = \varnothing$, entonces ya sabemos que es un conjunto. En caso contrario tomamos $b \in B$ y definimos la aplicación $F : A \longrightarrow B$ mediante[3]

$$F(x) = \begin{cases} x & \text{si } x \in B, \\ a & \text{si } x \notin B. \end{cases}$$

Es claro que F es suprayectiva, luego por el teorema anterior B es un conjunto. ∎

En particular esto implica que la intersección de conjuntos es un conjunto, pues está contenida en cualquiera de ellos.

El axioma del reemplazo no nos permite probar que la unión de conjuntos es un conjunto, ya que en general no podemos "cubrir" la unión con uno de los conjuntos que la forman. Por ello necesitamos un axioma específico, de hecho algo más general, como es NBG-12.

Si definimos $\bigcup x \equiv \{u \mid \bigvee v \in x\, u \in v\}$, entonces NBG-12 equivale a que $\bigwedge x \,\mathrm{cto} \bigcup x$.

Puesto que $x \cup y = \bigcup\{x, y\}$, el axioma del par junto con el axioma de la unión implican que la unión de conjuntos es un conjunto.

Más en general, si $t(v)$ es un término normal, definimos

$$\bigcup_{v \in X} t(v) \equiv \{u \mid \bigvee v \in X\, u \in t(v)\}$$

[3]Es fácil reducir definiciones como ésta al teorema general de formación de clases. Por ejemplo, en este caso sería $F = \{(u, v) \mid (u \in B \wedge v = u) \vee (u \in A \setminus B \wedge v = b)\}$.

y se cumple que

$$\bigwedge x(\bigwedge v \in x \ \mathrm{cto}\, t(v) \to \mathrm{cto} \bigcup_{v\in x} t(v)). \tag{8.1}$$

En efecto, podemos definir la clase

$$B = \{y \mid \bigvee v \in x \, y = t(v)\},$$

junto con la aplicación $F : x \longrightarrow B$ dada por $F(v) = t(v)$ (notemos que tanto en la definición de B como en la de F usamos que t es normal). Claramente F es suprayectiva, luego el axioma del reemplazo nos da que B es un conjunto, luego el axioma de la unión nos permite concluir que también lo es

$$\bigcup B = \bigcup_{v\in x} t(v).$$

(Notemos que en la última igualdad se usa la hipótesis $\bigwedge v \in x \ \mathrm{cto}\, t(v)$.) ∎

A su vez con todo esto podemos probar:

Teorema 8.9 *Si A y B son conjuntos, entonces también lo es $A \times B$.*

Demostración: Para cada $a \in A$, la aplicación $F : B \longrightarrow \{a\} \times B$ dada por $F(b) = (a, b)$ es claramente biyectiva, luego $\bigwedge a \in A \ \mathrm{cto}\{a\} \times B$. Puesto que $\{a\} \times B$ es ciertamente un término normal, podemos aplicar el resultado anterior y concluir que

$$A \times B = \bigcup_{a\in A} \{a\} \times B$$

es un conjunto. ∎

Ahora es claro que toda relación en un conjunto es un conjunto, o que una función es un conjunto si y sólo si lo es su dominio. También es fácil probar que si A es un conjunto y R es una relación de equivalencia en A, entonces el cociente A/R es un conjunto. En efecto, basta aplicar el axioma del reemplazo a la aplicación $F : A \longrightarrow A/R$ dada por $F(a) = [a]$.

Clases propias La existencia de clases propias es una consecuencia inmediata del teorema (o axioma) de formación de clases. En efecto, basta considerar la clase que da lugar a la paradoja de Russell:

$$R = \{x \mid x \notin x\}.$$

Ciertamente está bien definida, porque $x \notin x$ es una fórmula normal, y no puede ser un conjunto pues, en tal caso, sería $R \in R \leftrightarrow R \notin R$, lo cual es una contradicción. Por consiguiente, R ha de ser una clase propia y en particular $R \notin R$.

Los teoremas del apartado anterior nos permiten construir muchas más clases propias. Por ejemplo, la clase universal V no puede ser un conjunto, ya que si lo fuera $R \subset \mathrm{V}$ también lo sería. A su vez esto implica que si x es un conjunto

entonces $\mathrm{V}\setminus x$ ha de ser una clase propia, ya que en caso contrario $\mathrm{V} = x \cup (\mathrm{V}\setminus x)$ sería un conjunto. En general, cualquier clase biyectable con la clase universal es una clase propia. Es el caso de la clase de todos los conjuntos con un elemento. La biyección es $x \mapsto \{x\}$. Así se pueden poner muchos ejemplos más.

8.2 La teoría de conjuntos de Zermelo-Fraenkel

En un sentido que precisaremos después, la teoría de Zermelo-Fraenkel habla de los mismos conjuntos que la de von Neumann-Bernays-Gödel, pero prescindiendo de las clases propias.

Definición 8.10 Llamaremos teoría de conjuntos (básica) de *Zermelo-Fraenkel* a la teoría axiomática ZF^* cuyo lenguaje formal es el mismo de NBG^* y que consta de los axiomas siguientes:

ZF-1	$\bigwedge UV(\bigwedge X(X \in U \leftrightarrow X \in V) \to U = V)$	(extensionalidad)
ZF-2	$\bigwedge XY\bigvee Z\bigwedge U(U \in Z \leftrightarrow U = X \vee U = Y)$	(par)
ZF-3	$\bigvee Y\bigwedge X\, X \notin Y$	(vacío)
ZF-4	$\bigwedge X\bigvee Y\bigwedge U(U \in Y \leftrightarrow \bigvee V(U \in V \wedge V \in X))$	(unión)
ZF-5	Para toda fórmula $\phi(X, Y)$ (quizá con más variables libres)	
	$\bigwedge XYZ(\phi(X, Y) \wedge \phi(X, Z) \to Y = Z) \to$	
	$\bigwedge A\bigvee B\bigwedge Y(Y \in B \leftrightarrow \bigvee X \in A\, \phi(X, Y))$	(reemplazo)

Observaciones El axioma ZF-1 es el axioma de extensionalidad, y desempeña el mismo papel que su análogo en NBG^*. Observemos que la versión de ZF del axioma del par implica que $\bigwedge X\bigvee Y\, X \in Y$ es decir, que toda clase es un conjunto o, dicho de otra manera, que aquí la distinción entre clase y conjunto carece de valor. Por ello hablaremos únicamente de conjuntos y usaremos indistintamente letras mayúsculas o minúsculas como variables.

Análogamente a lo que hemos visto en NBG^*, del axioma del par se sigue que $\{x\}$, $\{x, y\}$ y (x, y) son descripciones propias[4] que se comportan como deben comportarse. Más concretamente, tenemos los teoremas siguientes:

1)	$\bigwedge xu(u \in \{x\} \leftrightarrow u = x)$	(Por ZF-1, ZF-2)
2)	$\bigwedge xyu(U \in \{x, y\} \leftrightarrow u = x \vee u = y)$	(Por ZF-1, ZF-2)
3)	$\bigwedge xyzw((x, y) = (z, w) \leftrightarrow x = z \wedge y = w)$	(Como en NBG^*)

El axioma del conjunto vacío nos da que $\varnothing$ es una descripción propia, es decir, podemos probar que $\bigwedge x\, x \notin \varnothing$. En esta teoría no disponemos de un análogo al teorema de formación de clases, por lo que los teoremas de existencia de

[4]Podemos tomar las mismas definiciones en ZF^* y en NBG^* de todos los conceptos conjuntistas, aunque de hecho las definiciones pueden simplificarse formalmente en ZF^* eliminando todas las restricciones a conjuntos, que ahora son superfluas.

conjuntos dependen principalmente del axioma del reemplazo. Si lo observamos con detenimiento, vemos que el papel que en NBG* desempeña una clase unívoca A, aquí lo desempeña una fórmula $\phi(x,y)$ a la que exigimos una condición de unicidad. Observemos que en NBG*, si $\phi(x,y)$ es una fórmula normal, se cumple que

$$\bigwedge xyz(\phi(x,y) \wedge \phi(x,z) \to y = z) \to \mathrm{Un}\{(x,y) \mid \phi(x,y)\},$$

y el caso particular del axioma de reemplazo en ZF* asociado a la fórmula $\phi(x,y)$ nos lleva a la misma conclusión que el axioma de reemplazo en NBG* particularizado a la clase $A = \{(x,y) \mid \phi(x,y)\}$. Del axioma del reemplazo se deduce una versión limitada del teorema de formación de clases:

Teorema 8.11 (Esquema de especificación) *Para toda fórmula $\phi(x)$ (tal vez con más variables libres) la fórmula siguiente es un teorema de* ZF**:*

$$\bigwedge A \overset{1}{\bigvee} B \bigwedge x(x \in B \leftrightarrow x \in A \wedge \phi(x)).$$

Demostración: Distingamos dos casos:

$$\neg\bigvee x \in A\, \phi(x) \vee \bigvee x \in A\, \phi(x).$$

En el primer caso $B = \varnothing$ cumple el teorema, en el segundo sea $b \in A \wedge \phi(b)$. Consideramos la fórmula $\psi(x,y) \equiv (\phi(x) \wedge y = x) \vee (\neg\phi(x) \wedge y = b)$. Por ZF-5 $\bigvee B \bigwedge y(y \in B \leftrightarrow \bigvee x(x \in A \wedge \psi(x,y)))$, de donde

$$\bigvee B \bigwedge y(y \in B \leftrightarrow y \in A \wedge \phi(y)).$$

La unicidad se sigue del axioma de extensionalidad. ∎

Observaciones Zermelo tomó como axioma el esquema de especificación y no consideró el axioma del reemplazo. Éste fue incorporado por Fraenkel, con lo que el esquema de especificación dejaba de ser un axioma y se convertía en un teorema.

Conviene comparar la prueba del teorema anterior con la del teorema 8.8. Más concretamente, consideremos el caso en que tenemos un conjunto A en NBG* y queremos probar que la subclase $B = \{x \in B \mid \phi(x)\}$ es un conjunto. Es la situación equivalente a la del teorema anterior, donde tenemos A y queremos probar que existe B. El caso $\neg\bigvee x \in A\, \phi(x)$ equivale a $B = \varnothing$, y lo tratamos como en 8.8, mediante el axioma del conjunto vacío. En el otro caso, al tomar un $b \in A$ tal que $\phi(b)$, estamos tomando un $b \in B$ (sólo que aquí no podemos nombrar a B, porque aún no hemos demostrado que existe). Aplicar el axioma del reemplazo a la fórmula $\psi(x,y)$ es el equivalente a aplicar en NBG* el axioma del reemplazo a la clase

$$F = \{(x,y) \mid \psi(x,y)\} = \{(x,y) \mid (x \in B \wedge y = x) \vee (x \notin B \wedge y = b)\},$$

pero ésta es exactamente[5] la función F a la que aplicamos el axioma de reemplazo en 8.8.

En definitiva, en ZF* no podemos garantizar que una fórmula determine un conjunto (en general esto es falso), pero sí es cierto que determina —especifica— un subconjunto de cualquier conjunto dado. Para enfatizar que necesitamos tener un conjunto mayor dado de antemano a la hora de aplicar el esquema de especificación conviene introducir la notación

$$\{x \in A \mid \phi(x)\} \equiv \{x \mid x \in A \wedge \phi(x)\}.$$

El esquema de especificación nos asegura que estos términos siempre son descripciones propias, es decir, para toda fórmula $\phi(X)$ (quizá con más variables libres) tenemos que

$$\bigwedge A x(x \in \{x \in A \mid \phi(x)\} \leftrightarrow x \in A \wedge \phi(x)).$$

■

El esquema de especificación asegura la existencia de la intersección de dos conjuntos, pues

$$x \cap y = \{u \in x \mid u \in y\}.$$

Al igual que ocurría en NBG*, la unión se escapa del alcance del axioma del reemplazo y requiere un axioma propio. El axioma ZF-4 nos asegura que $\bigcup z$ es una descripción propia para todo conjunto z y, aplicado a un par $z = \{x, y\}$, nos da que existe la unión de dos conjuntos cualesquiera, es decir,

$$\bigwedge xyu(u \in x \cup y \leftrightarrow u \in x \vee u \in y).$$

Estamos siguiendo punto por punto el camino que hemos seguido en el apartado de formación de conjuntos de la sección anterior. Según esto, el siguiente paso es probar que, para todo término $t(x)$ (quizá con más variables libres) $\bigcup_{v \in x} t(v)$ es una descripción propia. Más concretamente, hemos de probar que

$$\bigvee y \bigwedge u(u \in y \leftrightarrow \bigvee v \in x\, u \in t(v)).$$

Para ello, en NBG* considerábamos la clase $B = \{y \mid \bigvee v \in x\, y = t(v)\}$ y probábamos que es un conjunto mediante el axioma del reemplazo. Ahora hemos de probar que existe este conjunto mediante el axioma del reemplazo. Concretamente lo aplicamos a la fórmula $\phi(x, y) \equiv y = t(x)$. Claramente cumple la hipótesis de unicidad, por lo que obtenemos

$$\bigvee B \bigwedge y(y \in B \leftrightarrow \bigvee v \in x\, y = t(v)).$$

Ahora aplicamos el axioma de la unión a B y obtenemos el conjunto y que buscamos. ■

[5] En realidad esta F corresponde a la fórmula $x \in B \wedge \psi(x, y)$. Podríamos haber usado esta fórmula en lugar de ψ en la prueba del teorema de especificación y así el paralelismo habría sido completo, pero no lo hemos hecho porque hubiera sido una complicación innecesaria.

El lector debería comprobar que ha captado en la práctica la analogía entre NBG* y ZF* traduciendo la prueba en NBG* de que el producto cartesiano de conjuntos es un conjunto a una prueba en ZF* de que existe el producto cartesiano de dos conjuntos cualesquiera, es decir, hay que probar que

$$\bigwedge xy\overset{1}{\bigvee}z\bigwedge w(w \in z \leftrightarrow \bigvee uv(u \in x \wedge v \in y \wedge w = (u,v))).$$

De aquí se sigue que $x \times y$ es una descripción propia, es decir, que

$$\bigwedge xyw(w \in x \times y \leftrightarrow \bigvee uv(u \in x \wedge v \in y \wedge w = (u,v))).$$

Otro concepto cuya existencia no puede justificarse por especificación es el de dominio de una función. Más en general, se cumple que, para todo conjunto f (tanto si es una función como si no), existe otro conjunto formado por las primeras componentes de los pares ordenados que contenga f. En NBG* esto se demuestra considerando el subconjunto $P = \{x \in f \mid \bigvee uvx = (u,v)\}$ (cuya existencia en ZF* se tiene por especificación) y después la aplicación $g : P \longrightarrow \mathcal{D}f$ dada por $g(u,v) = u$. El análogo en ZF* de esta aplicación g es la fórmula que la define, es decir, $\phi(x,y) \equiv \bigvee z\, x = (y,z)$. Aplicando el axioma del reemplazo asociado a ϕ obtenemos

$$\bigvee B\bigwedge y(y \in B \leftrightarrow \bigvee x \in P\bigvee z\, x = (y,z)),$$

lo que claramente equivale a $\bigvee B\bigwedge y(y \in B \leftrightarrow \bigvee z\,(y,z) \in f)$. El axioma de extensionalidad nos da la unicidad y, por consiguiente, $\mathcal{D}f$ es una descripción propia, es decir,

$$\bigwedge fy(y \in \mathcal{D}f \leftrightarrow \bigvee z\,(y,z) \in f).$$

En realidad podríamos habernos ahorrado la definición de P y haber trabajado con f en su lugar. Similarmente se prueba la existencia del rango de un conjunto.

Ejercicio: Demostrar la existencia del inverso x^{-1} de un conjunto x, es decir,

$$\bigwedge x\overset{1}{\bigvee}y\bigwedge z(z \in y \leftrightarrow \bigvee uv(z = (u,v) \wedge (v,u) \in x)).$$

Con esto ya es inmediato que todos los resultados del apéndice A son definiciones lícitas y teoremas de ZF*. El único punto donde hemos de usar el axioma del reemplazo (de forma obvia) es a la hora de justificar la existencia del conjunto cociente de una relación de equivalencia.

Clases propias en ZF En principio, las clases que en NBG* resultan ser conjuntos no tienen equivalente en ZF*. Así, por ejemplo, el análogo en ZF* al teorema que afirma que la clase $R = \{x \mid x \notin x\}$ no es un conjunto es el teorema

$$\neg\bigvee R\bigwedge x(x \in R \leftrightarrow x \notin x).$$

El argumento es claro: si existiera tal R tendríamos $R \in R \leftrightarrow R \notin R$. Similarmente, podemos probar que

$$\neg\bigvee V\bigwedge x\, x \in V,$$

pues si existiera tal conjunto V por especificación tendríamos la existencia del conjunto R que acabamos de probar que no existe.

De este modo, en ZF* se puede probar que no existe el conjunto de todos los grupos o el conjunto de todos los espacios vectoriales y, en general, cualquiera de las "multiplicidades inconsistentes" de Cantor. Ésta es la forma que tiene la teoría de Zermelo-Fraenkel de evitar las paradojas. Sin embargo, esto no impide que los matemáticos puedan hablar de la clase de todos los grupos o la clase de todos los espacios vectoriales, aunque sea violando los principios de la teoría. Vamos a ver cómo puede hacerse esto sin poner en peligro el edificio que hemos construido.

Definimos una *clase* como una fórmula $\phi(x, x_1, \ldots, x_n)$ del lenguaje de la teoría de conjuntos en la que hemos destacado una variable libre x. Cuando pensemos en una fórmula como clase, usaremos la notación alternativa

$$\{x \mid \phi(x, x_1, \ldots, x_n)\},$$

lo cual no es más que una forma conveniente de referirnos a la fórmula $\phi(x)$, indicando al mismo tiempo que la variable destacada es x. Recordemos que una clase no es una fórmula, sino una fórmula con una variable destacada. Si $\phi(x, y)$ tiene dos variables libres, con ella podemos determinar dos clases, a saber,

$$\{x \mid \phi(x, y)\} \qquad \text{y} \qquad \{y \mid \phi(x, y)\}.$$

Si $A(x_1, \ldots, x_n) \equiv \{x \mid \phi(x, x_1, \ldots, x_n)\}$ es una clase, convenimos en que $x \in A(x_1, \ldots, x_n) \equiv \phi(x, x_1, \ldots, x_n)$. Por ejemplo, si

$$A = \{x \mid \bigvee yz(y \neq z \wedge \{y, z\} \subset x)\}, \quad B = \{x \mid \bigvee yz\, x \subset \{y, z\}\},$$

(podemos leer: A es "la clase de todos los conjuntos con al menos dos elementos" y que B es "la clase de todos los conjuntos con a lo sumo dos elementos"), las sentencias $\varnothing \notin A$, $\varnothing \in B$ son teoremas de ZF*. En efecto, $\varnothing \notin A$ no es más que una forma de nombrar la sentencia $\neg\bigvee yz(y \neq z \wedge \{y, z\} \subset \varnothing)$ y $\varnothing \in B$ es $\bigvee yz\, \varnothing \subset \{y, z\}$.

Ejercicio: Probar en ZF* que no existen conjuntos que contengan a todos los conjuntos con al menos dos elementos ni a todos los conjuntos con a lo sumo dos elementos.

De este modo, cualquier fórmula que contenga a A o a B o a cualquier otra clase a la derecha de uno o varios relatores de pertenencia puede ser interpretada de forma única como una fórmula del lenguaje de la teoría de conjuntos (sin más que cambiar $x \in A$ por la fórmula que define a A). Por ejemplo, se cumple que

$$\{x \mid \bigvee yz\, x = \{y, z\}\} \subset B,$$

pues esta sentencia es, más explícitamente,

$$\bigwedge x(x \in \{x \mid \bigvee yz\, x = \{y, z\}\} \to x \in B),$$

y al sustituir las clases por las fórmulas que las definen obtenemos ciertamente un teorema.

Podemos admitir clases como complementos del igualador si convenimos que, entre clases, $A = B$ es, por definición, $\bigwedge u(u \in A \leftrightarrow u \in B)$, y análogamente si A es una clase y B un conjunto o al revés. La práctica totalidad de los conceptos conjuntistas se puede aplicar a clases propias de forma natural. Por ejemplo, la sentencia

$$A \cap B = \{x \mid \bigvee yz\, x = \{y, z\}\}$$

se interpreta como

$$\bigwedge x(x \in A \wedge x \in B \leftrightarrow x \in \{x \mid \bigvee yz\, x = \{y, z\}\}),$$

y esto es claramente un teorema. No vamos a teorizar sobre qué condiciones ha de cumplir una fórmula que involucre clases para que pueda ser interpretada como una fórmula con sentido en ZF*. En último extremo, sistematizar el uso de clases propias en ZF* es trabajar en NBG*. No obstante, no necesitamos una teoría general para interpretar adecuadamente fórmulas concretas como las anteriores. En cualquier caso, no podemos dejar de ser conscientes de que las clases —pese a que las tratemos como tales— no son términos del lenguaje de la teoría de conjuntos, sino fórmulas. Por ello, la sentencia

$$\bigwedge w(w \neq A)$$

es un teorema de ZF* y no conduce a ninguna contradicción. Sólo expresa el hecho de que ningún conjunto contiene a todos los conjuntos con al menos dos elementos. Para deducir de aquí una contradicción tendríamos que sustituir la variable w por A para llegar a $A \neq A$, pero esto no es lícito, porque A no es un término que podamos sustituir por una variable. Esto se ve más claro si desarrollamos la sentencia:

$$\bigwedge w\neg\bigwedge x(x \in w \leftrightarrow \bigvee yz(y \neq z \wedge \{y, z\} \subset x)).$$

¿Qué podríamos sustituir aquí para tener una contradicción? Lo que hemos de tener claro es que en la expresión $\bigwedge w(w \neq A)$ *parece* que podamos sustituir w por A pero, al verla más de cerca, comprendemos que es sólo una apariencia. En realidad no hay nada que sustituir por nada.

El lector debería comprobar que todos los conceptos que hemos definido para clases en NBG* (incluidos los del apéndice A) tienen sentido también en ZF*.

Ejercicio: En ZF*. sea $\mathrm{V} = \{x \mid x = x\}$ la clase universal y sea $F : \mathrm{V} \longrightarrow A$ la aplicación dada por $F(x) = \{(x, \varnothing), (x, \{\varnothing\})\}$. Escribir explícitamente la fórmula que define a F. Probar que F es una aplicación inyectiva.

En términos de clases, el axioma del reemplazo puede enunciarse de forma mucho más parecida al de NBG*:

Teorema 8.12 *Sean F y B clases. Entonces*

$$\bigwedge A(F : A \longrightarrow B \text{ suprayectiva} \to \bigvee x\, x = B).$$

DEMOSTRACIÓN: Sea $F = \{x \mid \phi(x)\}$. El hecho de que F sea una función implica que $(u, v) \in F \wedge (u, w) \in F \to v = w$, pero esto es lo mismo que $phi(u, v) \wedge \phi(u, w) \to v = w$. Así pues, la fórmula $\phi(u, v) \equiv \mathtt{S}_x^{(u,v)}\phi$ cumple la hipótesis del axioma de reemplazo, luego

$$\bigvee x \bigwedge u(u \in x \leftrightarrow \bigvee v \in A\, \phi(u, v)),$$

es decir, $\bigvee x \bigwedge u(u \in x \leftrightarrow \bigvee v \in A\, u = F(v))$. El hecho de que F sea suprayectiva nos da que $\bigvee x \bigwedge u(u \in x \leftrightarrow u \in B)$, lo que equivale a $\bigvee x\, x = B$. ∎

Usando esta versión del axioma del reemplazo las pruebas que lo utilizan se vuelven prácticamente idénticas a sus análogas en NBG*. Notemos que hemos separado F y B en las hipótesis del teorema. Más detalladamente, su estructura es la siguiente: fijadas dos clases F y B (es decir, dos fórmulas) la fórmula $\bigwedge A(F : A \longrightarrow B$ suprayectiva $\to \bigvee x\, x = B)$ es un teorema de ZF*. Tenemos un teorema distinto para cada elección de las clases F y A, infinitos teoremas en total, no un único teorema. Sería absurdo escribir $\bigwedge FBA(\cdots$, tan absurdo como si escribiéramos $\bigwedge x = xBA(\cdots$, porque F no es una variable que podamos cuantificar, es una fórmula. En general, no tiene sentido cuantificar sobre clases propias. Un "para toda clase" sólo tiene sentido a nivel metamatemático como en el teorema anterior: "para toda clase, la fórmula siguiente es un teorema".

En realidad, la razón última de que este uso de clases en ZF* sea consistente es que existe NBG*, en el mismo sentido en que los antiguos algebristas podían manejar con tanta seguridad los números imaginarios gracias a que existe la teoría de los números complejos (aunque ellos no la conocieran como tal teoría). En el capítulo siguiente profundizaremos en la conexión entre ZF* y NBG* y volveremos a encontrarnos con estas ideas. De momento, lo que el lector debería asimilar es que si se encuentra con teoremas sobre clases en un libro que trabaja únicamente en Zermelo-Fraenkel, deberá ser capaz de interpretarlos como sentencias en las que no intervengan clases (en el caso más sofisticado, como esquemas teoremáticos en los que intervengan una o más fórmulas arbitrarias).

8.3 Los axiomas restantes de NBG y ZF

Hasta aquí hemos estudiado las teorías de conjuntos NBG*, MK* y ZF*, que contienen los axiomas básicos para que sea razonable afirmar que los objetos que describen son colecciones de objetos. Sin embargo, las teorías NBG, MK y ZF contienen algunos axiomas adicionales que garantizan que los conjuntos verifican una serie de hechos convenientes, y a menudo imprescindibles, para que se satisfagan los resultados básicos que los matemáticos admiten. En esta sección presentamos y describimos brevemente estos axiomas que luego estudiaremos con más atención, de modo que el lector pueda tener ya una visión completa de las teorías axiomáticas que sirven de fundamento a la matemática moderna.

El axioma de infinitud Como su nombre indica, el axioma de infinitud postula la existencia de un conjunto infinito. La forma más simple de enunciarlo es a través de la definición de infinitud de Dedekind, según la cual un conjunto x es infinito si y sólo si existe una aplicación $f : x \longrightarrow x$ inyectiva y no suprayectiva. Ésta no es la definición de infinitud de nosotros adoptaremos, pero, ciertamente, un conjunto con esta propiedad ha de ser infinito. Por consiguiente podemos tomar como axioma de infinitud la sentencia

$$AI \equiv \bigvee f x(f : x \longrightarrow x \text{ inyectiva y no suprayectiva}).$$

El axioma de partes En NBG* podemos definir la *clase de las partes* de una clase X como

$$\mathcal{P}X \equiv \{u \mid u \subset X\}.$$

El axioma de partes afirma que la clase de partes de un conjunto es de nuevo un conjunto. Podríamos, pues, enunciarlo como $\bigwedge x \operatorname{cto} \mathcal{P}x$, pero conviene definirlo como una sentencia que pueda entenderse indistintamente como axioma de NBG* o ZF*. Así pues, definimos

$$AP \equiv \bigwedge x \bigvee y \bigwedge u(u \subset y \to u \in y).$$

Así, AP afirma que para todo conjunto x existe un conjunto y que contiene a todos los subconjuntos de x. Esto es suficiente para deducir en NBG* que $\mathcal{P}x$ es un conjunto y en ZF* que existe $\mathcal{P}x$.

Con el axioma de partes es posible simplificar algunas pruebas de existencia. Por ejemplo, teniendo en cuenta la definición de par ordenado es claro que

$$\bigwedge XY\ X \times Y \subset \mathcal{P}\mathcal{P}(X \cup Y),$$

lo que nos da que el producto cartesiano de conjuntos es un conjunto por un argumento mucho más directo que el basado en el axioma del reemplazo.

Así como sin el axioma de infinitud no es posible demostrar la existencia de conjuntos infinitos, sin el axioma de partes no es posible demostrar la existencia de conjuntos no numerables. *Grosso modo*, el axioma de partes sólo es necesario para formalizar los argumentos que involucran de un modo u otro conjuntos no numerables. Por ello no ha de extrañarnos que pueda evitarse en la prueba de que el producto cartesiano de conjuntos es un conjunto, ya que el producto cartesiano de conjuntos no produce conjuntos no numerables a no ser que partamos ya de factores no numerables. En cambio, el axioma de partes es imprescindible en otros casos. Por ejemplo, podemos definir

$$Y^X = \{f \mid f : X \longrightarrow Y\}.$$

Observemos que si X no es un conjunto entonces $Y^X = \varnothing$. En cambio, si X es un conjunto Y^X es una clase no vacía (que puede ser no numerable

aunque lo sean X e Y). Por ello, para demostrar que $\bigwedge xy \operatorname{cto} y^x$, es necesario el axioma de partes. Para ello observamos que $y^x \subset \mathcal{PP}(x \cup y)$.

Así mismo es necesario el axioma de partes para justificar que el producto cartesiano de un conjunto de conjuntos es un conjunto. Más concretamente, dado un conjunto X, se define

$$\prod_{u\in X} u \equiv \{f \mid f : X \longrightarrow \bigcup_{u\in X} u \wedge \bigwedge u \in X\, f(u) \in u\}.$$

La prueba de que si X es un conjunto entonces $\prod_{u\in X} u$ es un conjunto se basa en la inclusión $\prod_{u\in X} u \subset \Big(\bigcup_{u\in X} u\Big)^X$. Análogamente a (8.1) se prueba que si $t(x)$ es un término normal entonces

$$\bigwedge x(\bigwedge u \in x \operatorname{cto} t(u) \to \operatorname{cto} \prod_{u\in x} t(u)),$$

donde el producto se define de forma obvia.

Ejercicio: Probar que los conceptos que acabamos de definir ($\mathcal{P}X$, Y^X, etc.) son normales.

El axioma de regularidad Este axioma, conocido usualmente como V = R, admite diversas formulaciones equivalentes. La más simple formalmente es la siguiente:

$$\bigwedge x(x \neq \varnothing \to \bigvee y \in x\, y \cap x = \varnothing).$$

En el capítulo XII veremos que, admitiendo AP, este axioma afirma en definitiva que todos los conjuntos pueden obtenerse a partir del conjunto vacío mediante sucesivas aplicaciones de la operación $\mathcal{P}$, en el mismo sentido en que definimos los conjuntos hereditariamente finitos en la introducción de esta segunda parte. Ahora bien, si admitimos AI, entonces el número de pasos que es necesario dar para obtener un conjunto dado a partir de $\varnothing$ no es necesariamente finito, sino que puede ser infinito e incluso no numerable. Para formalizar esta idea hemos de sustituir los números naturales (que usábamos en la definición de HF) por los números ordinales que introduciremos en el capítulo XI.

De momento, digamos tan sólo que el axioma de regularidad excluye la existencia de conjuntos "patológicos", como sería un conjunto x tal que $x = \{x\}$, o un par de conjuntos x, y tales que $x \in y \wedge y \in x$.

El axioma de elección Si, en el transcurso de una demostración, llegamos a una fórmula de tipo $x \neq \varnothing$ y, por consiguiente, $\bigvee u\, u \in x$, la regla de eliminación del particularizador nos permite tomar un conjunto $u_0 \in x$. En términos conjuntistas podemos decir que con esto hemos "elegido" un elemento de x. Vemos, pues, que elegir un elemento de un conjunto no vacío es una operación legítima desde un punto de vista lógico. Sin

embargo, nos encontramos con un problema cuando tenemos un conjunto x formado por conjuntos no vacíos y queremos elegir simultáneamente un elemento de cada $u \in x$. Para ello no podemos basarnos en ninguna regla de inferencia lógica, pues para enunciar con rigor lo que queremos hacer necesitamos recurrir al lenguaje de la teoría de conjuntos. Concretamente, el axioma de elección es la sentencia

$$AE \equiv \bigwedge x \bigvee f(f \text{ es una función} \wedge \mathcal{D}f = x \wedge \bigwedge u \in x(u \neq \varnothing \to f(u) \in u)).$$

En otros términos, para todo conjunto x existe una función f que elige un elemento de cada elemento no vacío de x. Se dice entonces que f es una *función de elección* sobre x. Vemos, pues, que el axioma de elección involucra la noción de "función". En realidad puede evitarse, pues es fácil probar que AE equivale a que para todo conjunto x formado por conjuntos no vacíos disjuntos dos a dos existe un conjunto y que contiene exactamente un elemento de cada elemento de x. En cualquier caso, necesitamos de forma esencial la noción general de conjunto, y ningún teorema lógico nos permite concluir ninguna propiedad específica sobre conjuntos.

No obstante, podría ocurrir que AE fuera demostrable (¡o refutable!) a partir de los axiomas restantes de la teoría de conjuntos. Hoy sabemos que no es así, sino que AE no puede demostrarse o refutarse a partir de los otros axiomas (salvo, naturalmente, si éstos fueran contradictorios). El axioma de elección ha suscitado polémicas entre los matemáticos debido a su carácter esencialmente no constructivo: las construcciones conjuntistas que se apoyan en funciones de elección nos dan objetos de los que no tenemos ninguna descripción explícita. No vamos a entrar aquí en detalles sobre esta polémica porque en este punto no disponemos de suficientes elementos de juicio para aportar nada de valor a la cuestión.

Así pues, la teoría de conjuntos de von Neumann-Bernays-Gödel (NBG) consta de un total de 17 axiomas: los 13 de NBG* más los cuatro que acabamos de definir. La teoría de conjuntos de Morse-Kelley (MK) consta del esquema de formación de clases más otros nueve axiomas (los cinco de MK* y los cuatro de esta sección). Restringiendo el esquema de formación de clases tenemos una axiomatización alternativa de NBG.

Es costumbre llamar teoría de conjuntos de Zermelo-Fraenkel (ZF) a la teoría que consta de los axiomas de ZF* más los axiomas de regularidad, infinitud y partes, pero no el axioma de elección (en total, siete axiomas más un esquema axiomático). La teoría que resulta al añadir el axioma de elección se suele representar por ZFC, donde —al parecer— la "C" hace referencia al inglés "choice" (elección).

En lo sucesivo seguiremos trabajando en NBG* o ZF* salvo que explícitamente se indique lo contrario. Esto no obedece a ninguna clase de recelo hacia los otros axiomas, sino que, por razones técnicas, es conveniente saber cuáles de los axiomas AI, AP o AE intervienen en cada teorema de la teoría de conjuntos. Respecto al axioma de regularidad, simplemente no hace falta para nada.

8.4 Los números naturales

Observemos que, aunque venimos hablando de números naturales desde mucho antes de introducir la teoría de conjuntos, no podemos ahora decir "Sea $\mathbb{N}$ el conjunto de los números naturales, es decir,

$$\mathbb{N} = \{0, \quad 1, \quad 2, \quad 3, \quad 4, \quad \dots \quad \}$$

en el que tenemos definidas dos operaciones: la suma $+ : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N}$ y el producto $\cdot : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N}$".

Si queremos demostrar teoremas de NBG o ZFC en los que aparezca $\mathbb{N}$ hemos de *definir* $\mathbb{N}$, y los puntos suspensivos de arriba no son una definición. Definir $\mathbb{N}$ es llamar $\mathbb{N}$ a un cierto designador del lenguaje de la teoría de conjuntos que hemos de especificar. Lo mismo vale para la suma y el producto. En esta sección construiremos los números naturales siguiendo las ideas de Dedekind, para lo cual hemos de postular el axioma de infinitud. No obstante, en el capítulo XI daremos otra construcción que no exige este axioma (*AI* sólo hará falta para probar que la clase de los números naturales es un conjunto o, equivalentemente, para probar que existe un conjunto que contiene a todos los números naturales, pero la definición de número natural puede darse sin necesidad de este axioma.)

Partimos, pues, del axioma de infinitud AI, que nos da un conjunto X y una aplicación $s : X \longrightarrow X$ inyectiva y no suprayectiva. Tomamos un elemento[6] $0 \in X \setminus s[X]$. Siguiendo a Dedekind, definimos un *conjunto inductivo* como un conjunto Y que cumpla

$$0 \in Y \wedge \bigwedge u \in Y\, s(u) \in Y.$$

Ciertamente existen conjuntos inductivos, por ejemplo X. Ahora definimos

$$\mathbb{N} \equiv \{x \in X \mid \bigwedge Y(Y \text{ inductivo} \to x \in Y)\},$$

es decir, $\mathbb{N}$ es la intersección de todos los conjuntos inductivos. El teorema siguiente es inmediato

Teorema 8.13 (Axiomas de Peano) *Se cumple:*

1. $0 \in \mathbb{N}$,
2. $\bigwedge x \in \mathbb{N}\, s(x) \in \mathbb{N}$,
3. $\bigwedge x \in \mathbb{N}\, s(x) \neq 0$,
4. $\bigwedge xy \in \mathbb{N}(s(x) = s(y) \to x = y)$,
5. $\bigwedge Y(Y \subset \mathbb{N} \wedge 0 \in Y \wedge \bigwedge u \in Y\, s(u) \in Y \to Y = \mathbb{N})$.

[6] Conviene observar que 0, en este contexto, es una variable que resulta de eliminar un particularizador, y no un designador. Después corregiremos esto.

Todo lo que sigue es válido para cualquier conjunto $\mathbb{N}$ que cumpla estos axiomas.[7] Del principio de inducción (el quinto axioma de Peano) se sigue inmediatamente que

$$\bigwedge u \in \mathbb{N}(u = 0 \vee \bigvee v \in \mathbb{N}\, u = s(v)).$$

En efecto, el conjunto $Y = \{u \in \mathbb{N} \mid u = 0 \vee \bigvee v \in \mathbb{N}\, u = s(v)\}$ es trivialmente inductivo, luego $Y = \mathbb{N}$.

Ahora vamos a definir la relación de orden en $\mathbb{N}$, para lo cual definiremos las secciones iniciales $I_n = \{0, \ldots, n\}$:

Diremos que I es una *sección inicial* de $n \in \mathbb{N}$ si cumple

$$I \subset \mathbb{N} \wedge 0 \in I \wedge n \in I \wedge s(n) \notin I \wedge \bigwedge u \in I(u \neq n \to s(u) \in I)$$

$$\wedge \bigwedge u \in \mathbb{N}(s(u) \in I \to u \in I).$$

Veamos ahora que cada $n \in \mathbb{N}$ tiene una única sección inicial.

En primer lugar probamos la unicidad: si I e I' son dos secciones de un mismo número natural n, entonces consideramos el conjunto

$$Y = \{u \in \mathbb{N} \mid u \notin I \vee u \in I'\}.$$

Claramente es inductivo, pues $0 \in Y$ (ya que $0 \in I'$) y si $u \in Y$, entonces, o bien $u = n$, en cuyo caso $s(u) \notin I'$, luego $s(u) \in Y$, o bien $u \neq n$, en cuyo caso distinguimos dos posibilidades: o bien $s(u) \notin I$, con lo que $s(u) \in Y$, o bien $s(u) \in I$, con lo que $u \in I$, luego (por hipótesis de inducción) $u \in I'$, luego $s(u) \in I'$, luego $s(u) \in Y$.

Con esto hemos probado que $Y = \mathbb{N}$, lo que implica que $I \subset I'$. Similarmente se prueba la otra inclusión, luego $I = I'$.

Veamos ahora la existencia por inducción sobre n. Ciertamente, $\{0\}$ es una sección inicial de 0 y si n tiene una sección inicial I, es fácil ver que $I' = I \cup \{s(n)\}$ es una sección inicial de $s(n)$.

Así pues, podemos definir $I_n \equiv I \mid (I$ es una sección inicial de $n)$. De la prueba anterior se sigue que $I_0 = \{0\}$ y que $\bigwedge n \in \mathbb{N}\, I_{s(n)} = I_n \cup \{s(n)\}$.

Veamos ahora que si $m \in I_n$ entonces $I_m \subset I_n$. Consideramos el conjunto $Y = \{m \in \mathbb{N} \mid m \notin I_n \vee I_m \subset I_n\}$. Ciertamente $0 \in Y$, pues $I_0 = \{0\} \subset I_n$. Si $m \in \mathbb{N}$, entonces, o bien $s(m) \notin I_n$, en cuyo caso $s(m) \in Y$, o bien $s(m) \in I_n$, con lo que $m \in I_n$ y por hipótesis de inducción $I_m \subset I_n$. Por consiguiente, $I_{s(m)} = I_m \cup \{s(m)\} \subset I_n$, luego $s(m) \in I$. Esto prueba que $Y = \mathbb{N}$, de donde se sigue lo pedido.

[7] Notemos que no hemos definido un conjunto $\mathbb{N}$, en el sentido de que $\mathbb{N}$ no es la abreviatura de ningún designador, sino que a partir del axioma de infinitud hemos demostrado que existe un conjunto que cumple estos axiomas. Lo que hemos definido como $\mathbb{N}$ es un término con dos variables libres, 0 y s (ambas aparecen libres en la definición de "conjunto inductivo", la X puede sustituirse por $\mathcal{D}s$). Por consiguiente tenemos un conjunto $\mathbb{N}$ distinto para cada elección de 0 y s.

Seguidamente probamos que si $m \in \mathbb{N} \setminus I_n$ entonces $I_n \subset I_m$, para lo cual consideramos el conjunto $Y = \{m \in \mathbb{N} \mid m \in I_n \vee I_n \subset I_m\}$. Tenemos que $0 \in Y$ porque $0 \in I_n$. Si $m \in Y$, o bien $m \notin I_n$, en cuyo caso $I_n \subset I_m \subset I_{s(m)}$, luego $s(m) \in Y$, o bien $m \in I_n$, en cuyo caso distinguimos dos posibilidades: $s(m) \in I_n$, en cuyo caso $s(m) \in Y$, o bien $s(m) \notin I_n$, lo que obliga a que $m = n$, con lo que $I_{s(m)} = I_{s(n)}$ y también $s(m) \in Y$. Por consiguiente $Y = \mathbb{N}$ y tenemos lo pedido.

Con esto podemos definir la relación de orden en $\mathbb{N}$ dada por

$$m \leq n \quad \leftrightarrow \quad I_m \subset I_n.$$

Trivialmente es una relación de orden y los dos resultados anteriores prueban que es una relación de orden total. Además para cada $n \in \mathbb{N}$ tenemos que $I_n = \{m \in \mathbb{N} \mid m \leq n\}$. Es claro que 0 es el mínimo número natural. Veamos ahora que todo subconjunto no vacío de $\mathbb{N}$ tiene un mínimo elemento. En efecto, sea $A \subset \mathbb{N}$, $A \neq \varnothing$. Basta considerar el conjunto

$$Y = \{u \in \mathbb{N} \mid \bigwedge v \in \mathbb{N}(v \leq u \to v \notin A) \vee \bigvee v \in \mathbb{N}(v \leq u \wedge v \text{ es mínimo de } A)\}.$$

Es fácil ver que Y es inductivo, con lo que $Y = \mathbb{N}$. Tomamos $a \in A$, de modo que el hecho de que $a \in Y$ implica que existe un mínimo elemento de A menor o igual que a.

Teniendo en cuenta que $I_{s(n)} = I_n \cup \{n\}$, es fácil ver que $s(n)$ es el siguiente de n en el sentido del orden, es decir, es el menor número natural mayor que n.

Ahora probamos el siguiente teorema de recursión:

Teorema 8.14 *Sea $g : X \longrightarrow X$ una aplicación cualquiera y sea $x \in X$. Entonces existe una única aplicación $f : \mathbb{N} \longrightarrow X$ de manera que $f(0) = x$ y $\bigwedge n \in \mathbb{N}\, f(s(x)) = g(f(x))$.*

Demostración: La unicidad es inmediata: si f y f' cumplen lo mismo, entonces una simple inducción nos da que $\bigwedge n \in \mathbb{N}\, f(n) = f'(n)$. Para probar la existencia veamos en primer lugar que

$$\bigwedge n \in \mathbb{N} \overset{1}{\bigvee} f(f : I_n \longrightarrow X \wedge f(0) = x \wedge \bigwedge u \in I_n(u \neq n \to f(s(u)) = g(f(u)))).$$

En efecto, la unicidad se prueba igual que antes. Respecto a la existencia, para $n = 0$ vale $f = \{(0, x)\}$. Si existe f para I_n, entonces $f \cup \{(s(n), g(f(n))\}$ sirve para $s(n)$.

Podemos definir

$$f_n = f \mid f : I_n \longrightarrow X \wedge f(0) = x \wedge \bigwedge u \in I_n(u \neq n \to f(s(u)) = g(f(u))),$$

y es claro entonces que si $m \leq n$ entonces $f_n|_{I_m}$ cumple la propiedad que define a f_m, luego $f_n|_{I_m} = f_m$. De aquí se sigue que $f = \bigcup_{n \in \mathbb{N}} f_n$ es la aplicación buscada. ∎

Ahora podemos corregir el "defecto" lógico que comentábamos en las notas al pie. Notemos que la prueba del teorema anterior se adapta trivialmente para probar lo siguiente:

$$\bigvee f(f \text{ es una función} \wedge \mathcal{D}f = \mathbb{N} \wedge f(0) = \varnothing \wedge \bigwedge n \in \mathbb{N}\, f(n) = n \cup \{n\}).$$

Simplemente hemos de cambiar cada aparición de un término $g(t)$ por $t \cup \{t\}$. La diferencia es que ahora la "función" $u \mapsto u \cup \{u\}$ no está definida sobre ningún conjunto X en particular. Equivalentemente, basta observar que el teorema anterior vale igualmente cuando $G : \mathrm{V} \longrightarrow \mathrm{V}$ es la aplicación dada por $G(u) = u \cup \{u\}$, sólo que ahora G y V son clases propias y no conjuntos.

Llamemos $\tilde{\mathbb{N}} = f[\mathbb{N}]$, de modo que $f : \mathbb{N} \longrightarrow \tilde{\mathbb{N}}$ suprayectiva. Vamos a probar que, de hecho, f es biyectiva. Para ello definimos $I'_n = \{m \in \mathbb{N} \mid m < n\}$, y una simple inducción nos da que $\bigwedge n \in \mathbb{N}\, f(n) = f[I'_n]$.

Así, si suponemos que f no es inyectiva, podemos tomar un mínimo número natural n tal que $f(n)$ tenga otra antiimagen, digamos $f(n) = f(m)$, con $n < m$. Entonces ha de ser $m = s(u)$, con $n \leq u$ y, puesto que $u \in I'_m$, tenemos que $f(u) \in f[I'_m] = f(m) = f(n) = f[I'_n]$, luego $f(u) = f(v)$, con $v < n$ (y en particular $v \neq u$), lo que contradice la minimalidad de n.

Ahora definimos $\tilde{s} : \tilde{\mathbb{N}} \longrightarrow \tilde{\mathbb{N}}$ mediante $\tilde{s}(x) = x \cup \{x\}$. Está bien definida pues si $x \in \tilde{\mathbb{N}}$ entonces $x = f(n)$ para cierto $n \in \mathbb{N}$, luego $x \cup \{x\} = f(s(n)) \in \tilde{\mathbb{N}}$. Además es inyectiva, pues si $\tilde{s}(x) = \tilde{s}(y)$, entonces $x = f(m)$, $y = f(n)$ y tenemos $f(s(m)) = f(s(n))$, luego $s(m) = s(n)$, luego $m = n$, luego $x = y$. Trivialmente $\tilde{s}$ no es suprayectiva, pues $\varnothing = f(0) \in \tilde{\mathbb{N}}$ no es de la forma $\tilde{s}(x)$.

Es inmediato comprobar que $\tilde{\mathbb{N}}$, $\tilde{s}$ y $0 \equiv \varnothing$ cumplen también los axiomas de Peano, sólo que ahora $\tilde{s}$ y 0 están explícitamente definidos. Más concretamente, observemos que, para estas elecciones, un conjunto Y es inductivo si cumple

$$\varnothing \in Y \wedge \bigwedge u \in Y\, u \cup \{u\} \in Y,$$

lo cual es una fórmula con Y como única variable libre, y que $\mathbb{N}$ puede definirse ahora mediante el designador

$$\mathbb{N} \equiv X \mid (X \text{ es inductivo} \wedge \bigwedge Y(Y \text{ inductivo} \to X \subset Y)). \tag{8.2}$$

En efecto, hemos probado que existe un conjunto con estas propiedades (a saber, $\tilde{\mathbb{N}}$), y es inmediato que éste es único, luego la descripción anterior es propia.[8] En definitiva, lo que hemos hecho ha sido reemplazar un número natural 0 indeterminado por $0 \equiv \varnothing$. Definimos

$1 \equiv s(0) = 0 \cup \{0\} = \{0\}$,

$2 \equiv s(1) = 1 \cup \{1\} = \{0, 1\}$,

$3 \equiv s(2) = 2 \cup \{2\} = \{0, 1, 2\}$, etc.

[8]Con más detalle: a partir del axioma de infinitud hemos construido un conjunto $\mathbb{N}$ al que vamos a llamar ahora $\mathbb{N}_0$. A partir de $\mathbb{N}_0$ hemos construido $\tilde{\mathbb{N}}$, $\tilde{s}$ y $0 \equiv \varnothing$ y es inmediato probar que $\tilde{\mathbb{N}} = \mathbb{N}$, donde $\mathbb{N}$ es el designador dado por (8.2).

Ahora ya es fácil probar que $\mathrm{NBG}^* + AI$ o $\mathrm{ZF}^* + AI$ son teorías aritméticas: sólo necesitamos definir la suma y el producto de números naturales, pero para ello basta usar el teorema 8.14, que, para cada $n \in \mathbb{N}$, nos garantiza la existencia de una única aplicación $n+ : \mathbb{N} \longrightarrow \mathbb{N}$ tal que

$$n + 0 = n \wedge \bigwedge m \in \mathbb{N}\, n + s(m) = s(n + m).$$

Definimos $+ : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N}$ mediante $n + m = (n+)(m)$. Similarmente se trata el producto. Ahora todos los teoremas del capítulo VI son válidos ahora para $\mathrm{NBG}^* + AI$ y $\mathrm{ZF}^* + AI$. En particular tenemos la conmutatividad y la asociatividad de la suma y el producto, etc.

A partir de $\mathbb{N}$ es fácil construir el conjunto $\mathbb{Z}$ de los números enteros y el conjunto $\mathbb{Q}$ de los números racionales. En el apéndice B están esbozadas estas construcciones. Concretamente, lo usual es definir $\mathbb{Z}$ como el conjunto cociente de $\mathbb{N} \times \mathbb{N}$ respecto a la relación de equivalencia dada por $(m, n)\ R\ (m', n')$ syss $m + n' = n + m'$, mientras que $\mathbb{Q}$ se define como el cociente de $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ respecto a la relación de equivalencia dada por $(m, n)\ R\ (m', n')$ syss $mn' = nm'$.

La construcción del conjunto $\mathbb{R}$ de los números reales a partir de $\mathbb{Q}$ requiere el axioma de partes AP. Las dos construcciones más frecuentes son la de Cantor, que requiere considerar el conjunto de sucesiones de números racionales, es decir, $\mathbb{Q}^{\mathbb{N}}$ (y para probar que esto es un conjunto hace falta AP) o la de Dedekind, que requiere el conjunto $\mathcal{P}\mathbb{Q}$ de todos los subconjuntos de $\mathbb{Q}$. A partir de aquí ningún matemático encontrará dificultades en formalizar en NBG o ZFC cualquier teorema del que conozca una demostración rigurosa.

Observaciones Terminamos la sección con algunas observaciones sobre la construcción que hemos dado de los números naturales. A estas alturas el lector ya debería tener claro que dicha construcción no puede considerarse como "la definición" de los números naturales, en el sentido de que alguien que crea que saber lo que son los números naturales es conocer dicha construcción u otra similar, será incapaz de entender nada de lo dicho en la primera parte de este libro.

Dedekind sí creía haber dado con "la definición" de los números naturales. Según él mismo explicó, su mayor problema fue encontrar una definición que no presupusiera el concepto de finitud, es decir, que evitara decir explícitamente "los objetos que se obtienen del cero aplicando un número finito de veces la función s". Según hemos visto, lo logró a través de la noción de conjunto inductivo, de modo que $\mathbb{N}$ se define esencialmente como el menor conjunto inductivo. Con ello, Dedekind logró reducir limpiamente la noción de "número natural" a la de "conjunto", y esta reducción es justo lo que necesitamos para probar que la teoría de conjuntos es una teoría aritmética. Ahora bien, en contra de lo que Dedekind pensaba, con ello no aclaró o precisó en modo alguno la noción de "número natural", pues redujo algo perfectamente conocido, como son los números naturales, a algo a lo que no sabemos atribuirle ningún significado preciso, como es "la totalidad de los conjuntos inductivos", algo que no podríamos

tratar sin vacilaciones sin recurrir a una teoría axiomática que nos exima de precisar de qué estamos hablando.

Lo único que hemos probado es que cualquier colección de objetos que cumpla los axiomas de la teoría de conjuntos (cualquier modelo de la teoría de conjuntos) contiene necesariamente unos objetos que *se parecen* a los números naturales en la medida en que satisfacen los axiomas de teoría aritmética. No obstante, nada nos asegura que esos objetos se comporten realmente como los números naturales. Sabemos que (si la teoría es consistente) existen modelos en los que entre los números naturales así definidos hay algunos que no pueden obtenerse a partir del cero en un número finito de pasos. Al menos, no según la noción que todos tenemos de finitud, si bien es cierto que todos ellos se obtendrán del cero en un número finito de pasos en el sentido de "finitud" que se deriva de la propia noción de número natural que hemos definido. En otras palabras, si esto ocurre, el modelo tendrá conjuntos con infinitos elementos pero que satisfarán la definición usual de finitud (un conjunto es finito si se puede biyectar con una sección I_n).

En resumen, la construcción de los números naturales en una teoría axiomática de conjuntos debe entenderse como un proceso técnico enmarcado en el programa general de reducir todos los conceptos básicos a unos pocos axiomas. Pensemos, por ejemplo, que en NBG* hemos demostrado la existencia de la unión de clases a partir de un axioma que postula la existencia de la intersección de clases. Esta posibilidad de reducir el álgebra de clases a dos únicas operaciones básicas (la intersección y el complemento en nuestro caso) tiene valor lógico, pero no matemático, pues desde un punto de vista matemático la unión es tan elemental como pueda serlo la intersección, e igualmente podríamos haber postulado la existencia de la unión y deducir de ahí la existencia de la intersección. Similarmente, la construcción de los números naturales es un proceso similar de reducción lógica de unos resultados básicos a unos pocos axiomas. Si de los axiomas de la teoría de conjuntos no se dedujera la existencia de un conjunto que cumple los axiomas de Peano, no deberíamos por ello deducir que no existen los números naturales, sino más bien que nuestros axiomas serían demasiado débiles, por lo que tendríamos que sustituirlos por otros más fuertes o, simplemente, postular directamente la existencia de $\mathbb{N}$. Así pues, si de la imposibilidad de demostrar la existencia de $\mathbb{N}$ (a partir de unos axiomas arbitrarios) no deduciríamos que $\mathbb{N}$ no existe, tampoco es justo afirmar que la razón definitiva por la que $\mathbb{N}$ existe es que lo hemos definido a partir de unos axiomas arbitrarios.

8.5 Eliminación de descriptores

Se comprueba fácilmente que si $M \models T$, donde T es cualquiera de las teorías de conjuntos que hemos considerado, entonces el modelo M' que difiere de M tan sólo en que $M'(x|x = x) = M(\varnothing)$ cumple también $M' \models T$ (debido a que los axiomas no dicen nada de la descripción impropia).

Por consiguiente, la consistencia de T no se pierde si añadimos como axioma

$\varnothing = x|(x = x)$. Esto puede considerarse más como un convenio que como un axioma propiamente dicho: estamos conviniendo que cualquier cosa mal definida es el conjunto vacío. Así podemos mejorar el teorema 2.14:

Teorema 8.15 (Teorema de eliminación de descriptores) *Para toda fórmula α, existe una fórmula α' sin descriptores de modo que a partir del axioma de extensionalidad y el axioma del conjunto vacío (y el convenio $\varnothing = x|(x = x)$) se demuestra que es equivalente a α.*

DEMOSTRACIÓN: La prueba es la misma que la del teorema 2.14, salvo que en el último paso se usa

$$Y = X|\alpha \leftrightarrow \bigwedge X(\alpha \leftrightarrow X = Y) \vee (\neg \overset{1}{\bigvee} X\alpha \wedge \bigwedge u\, u \notin Y).$$

A su vez esta fórmula se sigue inmediatamente de 2.13 y de los axiomas supuestos. ∎

En particular todos los axiomas de cualquiera de las teorías de conjuntos que hemos estudiado tienen formas equivalentes sin descriptores. De hecho es fácil construir explícitamente formas equivalentes más sencillas que las que proporciona la prueba del teorema anterior. Como consecuencia del teorema de completitud hemos visto que una consecuencia sin descriptores de unas premisas sin descriptores puede deducirse sin descriptores, lo que nos permite, si es necesario, suponer que el lenguaje de la teoría de conjuntos no tiene descriptor.

Capítulo IX

Modelos de la teoría de conjuntos

De los teoremas de incompletitud de Gödel se sigue la imposibilidad de demostrar la consistencia de las teoría de conjuntos y, en particular, de construir explícitamente un modelo de cualquiera de ellas. Esto lo veremos con más detalle en el capítulo siguiente, pero ahora nos ocuparemos de estudiar la relación entre los modelos de la teoría de Zermelo-Fraenkel y los modelos de la teoría de von Neumann-Bernays-Gödel. Recordemos que, por el teorema de completitud, si cualquiera de estas teorías es consistente, entonces tiene un modelo, luego es razonable estudiar tales modelos aunque no sepamos construir ninguno: si, como es plausible, las teorías de conjuntos son consistentes, dichos modelos existen y cuanto deduzcamos en este capítulo serán resultados verdaderos sobre las mismas.

9.1 La consistencia de ZFC−AI

Pese a lo que acabamos de comentar, es muy interesante observar que sí podemos construir un modelo (y, por consiguiente, probar la consistencia) de ZFC sin el axioma de infinitud (brevemente, ZFC$-AI$). En efecto, un modelo de esta teoría lo constituye la colección HF de los conjuntos hereditariamente finitos, descrita en la introducción a la segunda parte de este libro. Esto ya lo comentábamos allí. Ahora que conocemos explícitamente los axiomas de ZFC lo único que falta es comprobar que, efectivamente, todos ellos son verdaderos en HF. La comprobación no ofrece ninguna dificultad. Veamos como ejemplo el axioma del reemplazo. Tomamos una fórmula $\phi(x,y)$, quizá con más variables libres, y hemos de ver que, para cualquier valoración v, se cumple

$$\mathrm{HF} \vDash \big(\bigwedge xyz(\phi(x,y) \wedge \phi(x,z) \to y = z) \to$$

$$\bigwedge A \bigvee B \bigwedge y(y \in B \leftrightarrow \bigvee x \in A\, \phi(x,y))\big)[v].$$

Para ello suponemos la hipótesis, es decir, que para toda terna u, v, w de conjuntos hereditariamente finitos, si se cumple $\mathrm{HF} \vDash \phi[v_{xy}^{uv}]$ y $\mathrm{HF} \vDash \phi[v_{xy}^{uw}]$ entonces $v = w$. Ahora tomamos un conjunto hereditariamente finito a y hemos de encontrar otro b que cumpla la tesis.

El conjunto a tiene una cantidad finita de elementos (quizá ninguno) y para cada uno de ellos u existe a lo sumo un conjunto v tal que $\mathrm{HF} \vDash \phi[v_{xy}^{uv}]$. Así pues, existe a lo sumo una cantidad finita de conjuntos hereditariamente finitos v para los que existe un u en HF tal que $\mathrm{HF} \vDash \phi[v_{xy}^{uv}]$. Por consiguiente, existe un número natural n tal que todos estos conjuntos están en un cierto V_n (ver la introducción). El conjunto b formado por todos ellos es un elemento de V_{n+1} y, en particular, un elemento de HF. Claramente b cumple lo pedido, es decir,

$$\mathrm{HF} \vDash \bigwedge y(y \in B \leftrightarrow \bigvee x \in A\, \phi(x,y))[v_{AB}^{a\,b}].$$

Naturalmente, esta prueba está sujeta a observaciones similares a las que hicimos en la página 91 sobre la consistencia de la aritmética de Peano. No es finitista y no todo el mundo la acepta como concluyente. No obstante, lo cierto es que tenemos criterios para dar un significado objetivo a cualquier afirmación sobre los conjuntos hereditariamente finitos (con independencia de si sabemos determinar si es verdadera o falsa), por lo que la demostración de una contradicción en ZFC$-AI$ nos daría un razonamiento concluyente de que, por ejemplo, el conjunto vacío tiene elementos y, como esto es falso (¡sabemos lo que esto significa y sabemos que es falso!) tal razonamiento y tal demostración no pueden existir. En realidad, la consistencia de ZFC$-AI$ es equivalente a la consistencia de la aritmética de Peano (y esto sí es universalmente aceptado). La prueba se basa en el ejercicio siguiente:

Ejercicio: Consideremos la relación R en el conjunto de los números naturales dada por $m\,R\,n$ si y sólo si el m-simo dígito de la expresión binaria de n es igual a 1. Entonces el conjunto de los números naturales es un modelo de ZFC$-AI$ interpretando la relación de pertenencia como la relación R.

De hecho, el modelo que describe este ejercicio es "isomorfo" a HF, en el sentido obvio de la palabra (existe una biyección entre ambos que conserva la relación de pertenencia).

Recíprocamente, a partir de un modelo de ZFC$-AI$ podemos construir un modelo de la aritmética de Peano. Esto todavía no lo hemos probado, pues para construir los números naturales hemos usado el axioma de infinitud. No obstante, como ya hemos anunciado, en el capítulo XI daremos una construcción alternativa que no requiere este axioma, de modo que los objetos de un modelo de ZFC$-AI$ que verifiquen la definición de número natural, junto con las funciones determinadas por la definición conjuntista de suma y producto constituyen claramente un modelo de la aritmética de Peano.

9.2 Consis NBG implica Consis ZFC

En esta sección veremos que a partir de un modelo de NBG podemos obtener un modelo de ZFC, y la forma de hacerlo será la natural: quitar las clases

propias. Más precisamente, entre esta sección y la siguiente probaremos que los conjuntos de los que habla NBG son, en cierto sentido, los mismos conjuntos de los que habla ZFC. Para precisar estas ideas hemos de introducir la noción de relativización de una expresión:

Definición 9.1 Definimos la *relativización* θ^{V} de una expresión θ del lenguaje de la teoría de conjuntos como la expresión dada por las reglas siguientes:

$$\begin{aligned}
&X^{\mathrm{V}} \equiv X\\
&(t_1 \in t_2)^{\mathrm{V}} \equiv t_1^{\mathrm{V}} \in t_2^{\mathrm{V}} && (t_1 = t_2)^{\mathrm{V}} \equiv t_1^{\mathrm{V}} = t_2^{\mathrm{V}}\\
&(\neg\alpha)^{\mathrm{V}} \equiv \neg\alpha^{\mathrm{V}} && (\alpha \to \beta)^{\mathrm{V}} \equiv \alpha^{\mathrm{V}} \equiv \beta^{\mathrm{V}}\\
&\textstyle(\bigwedge X\alpha)^{\mathrm{V}} \equiv \bigwedge X(\mathrm{cto}\, X \to \alpha^{\mathrm{V}}) && (X|\alpha)^{\mathrm{V}} \equiv X|(\mathrm{cto}\, X \wedge \alpha^{\mathrm{V}})
\end{aligned}$$

Claramente se cumple:

$$(\alpha \vee \beta)^{\mathrm{V}} \leftrightarrow \alpha^{\mathrm{V}} \vee \beta^{\mathrm{V}}, \qquad (\alpha \wedge \beta)^{\mathrm{V}} \leftrightarrow \alpha^{\mathrm{V}} \wedge \beta^{\mathrm{V}}, \qquad (\alpha \leftrightarrow \beta)^{\mathrm{V}} \leftrightarrow (\alpha^{\mathrm{V}} \leftrightarrow \beta^{\mathrm{V}}),$$

$$(\textstyle\bigvee X\alpha)^{\mathrm{V}} \leftrightarrow \bigvee X(\mathrm{cto}\, X \wedge \alpha^{\mathrm{V}}), \qquad (\overset{1}{\bigvee} X\alpha)^{\mathrm{V}} \leftrightarrow \overset{1}{\bigvee} X(\mathrm{cto}\, X \wedge \alpha^{\mathrm{V}}).$$

Es claro que θ y θ^{V} tienen las mismas variables libres. Una simple inducción prueba que θ^{V} es una expresión normal.

En definitiva θ^{V} es la expresión que resulta de exigir que todas las variables ligadas que aparecen en θ hagan referencia a conjuntos. En la práctica, si usamos el convenio de que las letras minúsculas en NBG representan conjuntos, relativizar una expresión es hacer minúsculas todas sus variables ligadas.

Definición 9.2 Sea M un modelo de NBG* de universo U. Llamaremos $\underline{M}$ al modelo del lenguaje de la teoría de conjuntos determinado por:

- El universo de $\underline{M}$ es la colección $\underline{U}$ de los objetos a de U que satisfacen la fórmula $\mathrm{cto}\, X$, es decir, tales que, para cualquier valoración v, cumplen $M \vDash \mathrm{cto}\, X[v^a_X]$.
- Si a y b están en $\underline{U}$, se cumple $\underline{M}(\in)(a,b)$ syss $M(\in)(a,b)$.
- La descripción impropia de $\underline{M}$ es la misma que la de M y, según convinimos en el capítulo anterior, es $M(\varnothing)$. Notemos que $M(\varnothing)$ está en $\underline{U}$, pues $M \vDash \mathrm{cto}\, \varnothing$.

La relación entre los modelos M y $\underline{M}$ viene dada por el teorema siguiente:

Teorema 9.3 *Sea M un modelo de NBG*, sea v una valoración en $\underline{M}$ (también lo será en M) y sea θ una expresión. Entonces*

Si θ es una fórmula $M \vDash \theta^{\mathrm{V}}[v]$ syss $\underline{M} \vDash \theta[v]$.

Si θ es un término $M(\theta^{\mathrm{V}})[v] = \underline{M}(\theta)[v]$.

DEMOSTRACIÓN: Por inducción sobre la longitud de θ.

Si $\theta \equiv X$ entonces $\theta^{\mathrm{V}} \equiv X$ y claramente $M(\theta^{\mathrm{V}})[v] = v(X) = \underline{M}(\theta)[v]$.

Si $\theta \equiv t_1 \in t_2$ entonces $\theta^{\mathrm{V}} \equiv t_1^{\mathrm{V}} \in t_2^{\mathrm{V}}$ y

$$M \vDash \theta^{\mathrm{V}}[v] \text{ syss } M(\in)(M(t_1^{\mathrm{V}})[v], M(t_2^{\mathrm{V}})[v])$$
$$\text{syss } \underline{M}(\in)(\underline{M}(t_1)[v], \underline{M}(t_2)[v]) \text{ syss } \underline{M} \vDash \theta[v].$$

El caso $\theta \equiv t_1 = t_2$ es análogo.

Si $\theta \equiv \neg\alpha$ entonces $\theta^{\mathrm{V}} \equiv \neg\alpha^{\mathrm{V}}$ y

$$M \vDash \theta^{\mathrm{V}}[v] \text{ syss no } M \vDash \alpha^{\mathrm{V}}[v] \text{ syss no } \underline{M} \vDash \alpha[v] \text{ syss } \underline{M} \vDash \theta[v].$$

El caso $\theta \equiv \alpha \to \beta$ es análogo.

Si $\theta \equiv \bigwedge X\alpha$, entonces $\theta^{\mathrm{V}} \equiv \bigwedge X(\mathrm{cto}\, X \to \alpha^{\mathrm{V}})$ y tenemos que $M \vDash \theta^{\mathrm{V}}[v]$ syss para todo a de U se cumple $M \vDash (\mathrm{cto}\, X \to \alpha^{\mathrm{V}})[v^a_X]$ syss para todo a de U se cumple no $M \vDash \mathrm{cto}\, X[v^a_X]$ o $M \vDash \alpha^{\mathrm{V}}[v^a_X]$ syss para todo a de $\underline{U}$ se cumple $\underline{M} \vDash \alpha[v^a_X]$ syss $\underline{M} \vDash \theta[v]$.

Si $\theta \equiv X|\alpha$, entonces $\theta^{\mathrm{V}} \equiv X|(\mathrm{cto}\, X \wedge \alpha^{\mathrm{V}})$. Si hay un único a en U tal que $M \vDash (\mathrm{cto}\, X \wedge \alpha^{\mathrm{V}})[v^a_X]$, entonces, como $M \vDash \mathrm{cto}\, X[v^a_X]$ tenemos que a está en $\underline{U}$, y además a es el único elemento de $\underline{U}$ que cumple $\underline{M} \vDash \alpha[v^a_X]$, pues si otro b cumple esto mismo, entonces $M \vDash (\mathrm{cto}\, X \wedge \alpha^{\mathrm{V}})[v^b_X]$, luego $b = a$. Por consiguiente en este caso $M(\theta^{\mathrm{V}})[v] = a = \underline{M}(\theta)[v]$.

Si no hay un único a en U que cumpla $M \vDash (\mathrm{cto}\, X \wedge \alpha^{\mathrm{V}})[v^a_X]$, se comprueba igualmente que no hay un único a en $\underline{U}$ que cumpla $\underline{M} \vDash \alpha[v^a_X]$, y así concluimos que $M(\theta^{\mathrm{V}})[v] = \underline{M}(\theta)[v] = M(\varnothing)$. ∎

En particular tenemos:

Teorema 9.4 *Sea α una sentencia del lenguaje de la teoría de conjuntos y M un modelo de NBG*. Entonces $M \vDash \alpha^{\mathrm{V}}$ si y sólo si $\underline{M} \vDash \alpha$.*

Teorema 9.5 *Si M es un modelo de NBG* entonces $\underline{M}$ es un modelo de ZF*.*

DEMOSTRACIÓN: Basta probar que si α es un axioma de ZF* entonces α^{V} es un teorema de NBG*, con lo que $M \vDash \alpha^{\mathrm{V}}$ y por el teorema anterior tenemos que $\underline{M} \vDash \alpha$. Notemos que los axiomas de ZF* son sentencias salvo los correspondientes al esquema de reemplazo, pero éstos podemos sustituirlos por sus clausuras universales para así poder aplicar el teorema anterior.

La prueba no ofrece ninguna dificultad. ∎

Más en general, observemos que los axiomas adicionales, AI, AP, $\mathrm{V} = \mathrm{R}$, AE de NBG son exactamente las relativizaciones de los axiomas correspondientes de ZFC, por lo que también tenemos que si M es un modelo de NBG entonces $\underline{M}$ es un modelo de ZFC.

Está claro que la interpretación de estos hechos es la que ya hemos comentado: los conjuntos de los que hablamos mediante NBG satisfacen los axiomas de ZFC. En la sección siguiente precisaremos más aún la relación entre ambas teorías.

9.3 Consis ZFC implica Consis NBG

Ahora recorreremos el camino inverso al de la sección anterior: partiendo de un modelo M de ZF* construiremos un modelo de NBG* con los mismos conjuntos, es decir, que únicamente añadiremos a M las clases propias necesarias para que se satisfagan los axiomas de NBG*, pero no añadiremos ningún conjunto. Ésta es la parte más delicada, pues en la sección anterior sólo teníamos que eliminar las clases propias, mientras que ahora tenemos que crearlas. Para ello nos basaremos en las ideas que esbozamos en el capítulo anterior, a la vez que las precisaremos, según las cuales las clases propias de NBG* son simplemente las colecciones de conjuntos definibles mediante una fórmula con parámetros.

Sea, pues, M un modelo de ZF* de universo U. Sea $\alpha_0, \alpha_1, \alpha_2, \ldots$ una enumeración de las fórmulas con alguna variable libre. Sea $m_i + 1$ el número de variables libres de α_i y sean éstas $y_0^i, \ldots, y_{m_i}^i$ en orden creciente de índices.

Si $a_1, \ldots, a_{m_i}$ son elementos de U, llamaremos $R_i[a_1, \ldots, a_{m_i}]$ a la relación monádica en U dada por

$$R_i[a_1, \ldots, a_{m_i}](b) \quad \text{syss} \quad M \vDash \alpha_i[v],$$

donde v es cualquier valoración en M que cumpla $v(y_0^i) = b$, $v(y_j^i) = a_j$, para $j = 1, \ldots, m_i$.

Observemos que una relación monádica en U como $R_i[a_1, \ldots, a_{m_i}]$ es lo mismo que una colección de objetos de U, la colección de todos los objetos (conjuntos) que la satisfacen. Concretamente, $R_i[a_1, \ldots, a_{m_i}]$ puede verse como la colección de todos los conjuntos que satisfacen la fórmula α_i donde las variables distintas de la primera se interpretan como los parámetros $a_1, \ldots, a_{m_i}$.

Nuestra intención es tomar como clases estas colecciones de conjuntos, pero no podemos hacerlo tan directamente, ya que dos de estas relaciones podrían tener la misma extensión (una misma clase puede ser definida por varias fórmulas distintas).

Sea $[\alpha_i; a_1, \ldots, a_{m_i}]$ la colección de todas las m_j-tuplas $(\alpha_j, b_1, \ldots, b_{m_j})$ tales que la relación $R_j[b_1, \ldots, b_{m_j}]$ coincide con $R_i[a_1, \ldots, a_{m_i}]$.

Esto es tanto como considerar entre las relaciones $R_i[a_1, \ldots, a_{m_i}]$ la relación de equivalencia "tener la misma extensión" y quedarnos con las clases de equivalencia. Estas clases de equivalencia $[\alpha_i; a_1, \ldots, a_{m_i}]$ son mejores candidatos a clases, pero tampoco podemos tomarlas todas, ya que si la extensión de una de estas clases coincide con la extensión de un conjunto de M, entonces dicha "clase" tiene que ser identificada con el conjunto que ya tenemos o, en otras palabras, no es una clase propia que debamos añadirle a M.

Sea C la colección de todas las clases de equivalencia $[\alpha_i; a_1, \ldots, a_{m_i}]$ tales que no existe ningún b en U para el que la relación monádica en U dada por $R_b(x)$ syss $M(\in)(x, b)$ coincida con $R_i[a_1, \ldots, a_{m_i}]$.

Es decir, C es (o pretende ser) la colección de todas las clases de M que no se corresponden con ningún conjunto, es decir, las colecciones de conjuntos que

hemos de añadir a M para formar un modelo de NBG*. Vamos a formar tal modelo:

Llamemos $\overline{U}$ a la colección formada por todos los objetos de U y los de C. Sea $\overline{M}$ el modelo de $\mathfrak{L}$ dado por

- El universo de $\overline{M}$ es $\overline{U}$.
- La descripción impropia de $\overline{M}$ es la misma que la de M, es decir, $M(\varnothing)$.
- $\overline{M}(\in)(a,b)$ syss $\begin{cases} a \text{ y } b \text{ están en } U \text{ y } M(\in)(a,b) \text{ o} \\ a \text{ está en } U,\ b = [\alpha_i; a_1, \ldots, a_{m_i}] \text{ está en } C \\ \text{y } R_i[a_1, \ldots, a_{m_i}](a). \end{cases}$

Así pues, los elementos de $\overline{M}$ son los conjuntos de M y las clases propias de C, y hemos establecido que un conjunto a está en una clase propia b si y solo si a satisface cualquiera de las fórmulas que define a b. El teorema siguiente prueba que todo esto nos lleva al objetivo deseado.

Teorema 9.6 *Si M es un modelo de ZF* entonces $\overline{M}$ es un modelo de NBG*.*

Demostración: Empezamos probando algunos hechos generales sobre $\overline{M}$.

Sea v una valoración en $\overline{M}$. Entonces $\overline{M} \vDash \mathrm{cto}\, X[v]$ syss $\overline{M} \vDash (\bigvee Y\, X \in Y)[v]$ syss existe un b en $\overline{U}$ tal que $\overline{M}(\in)(v(X), b)$ syss $v(X)$ está en U (notemos que todos los objetos de U pertenecen a otro objeto de U).

De aquí se sigue que $\overline{M} \vDash \bigwedge X(\mathrm{cto}\, X \to \alpha)[v]$ syss para todo a en U se cumple $\overline{M} \vDash \alpha[v^a_X]$ e igualmente $\overline{M} \vDash \bigvee X(\mathrm{cto}\, X \wedge \alpha)[v]$ syss existe un a en U tal que $\overline{M} \vDash \alpha[v^a_X]$.

$(*)$ Dados $a_1, \ldots, a_{m_i}$ en U, existe un a en $\overline{U}$ tal que para todo u en $\overline{U}$ se cumple

$$\overline{M}(\in)(u,a) \quad \text{syss} \quad R_i[a_1, \ldots, a_{m_i}](u),$$

pues, o bien existe tal a en U o, en caso contrario, $[\alpha_i; a_1, \ldots, a_{m_i}]$ está en C y cumple lo pedido.

Veamos ahora que todos los axiomas de NBG* son verdaderos en $\overline{M}$.

NBG-1: $\overline{M} \vDash \bigwedge XY(\bigwedge u(u \in X \leftrightarrow u \in Y) \to X = Y)$.

Sea v una valoración en M, sean a y b en $\overline{U}$. Hemos de comprobar que

$$\overline{M} \vDash (\bigwedge u(u \in X \leftrightarrow u \in Y) \to X = Y)[v^{a\,b}_{XY}].$$

Para ello suponemos que $\overline{M} \vDash \bigwedge u(u \in X \leftrightarrow u \in Y)[v^{a\,b}_{XY}]$, es decir, que para todo c de U se cumple $\overline{M}(\in)(c,a)$ syss $\overline{M}(\in)(c,b)$.

Notemos que no puede ocurrir que a esté en U y $b = [\alpha_j; b_1, \ldots, b_{m_j}]$ esté en C, pues entonces $R_a(c)$ equivaldría a $R_j[b_1, \ldots, b_{m_j}](c)$ y esto contradice la definición de C.

Igualmente es imposible que a esté en C y b esté en U. Así pues, ambos están en U o ambos están en C.

Si a y b están ambos en U tenemos que $M \vDash \bigwedge U(U \in X \leftrightarrow U \in Y)[v_{XY}^{a\,b}]$, y como $M \vDash$ ZF-1 resulta que $M \vDash (X = Y)[v_{XY}^{a\,b}]$, o sea, $a = b$.

Si $a = [\alpha_i; a_1, \ldots, a_{m_i}]$ y $b = [\alpha_j; b_1, \ldots, b_{m_j}]$ entonces $R_i[a_1, \ldots, a_{m_i}]$ y $R_j[b_1, \ldots, b_{m_j}]$ coinciden sobre todos los elementos de U, luego $a = b$.

En cualquier caso tenemos que $\overline{M} \vDash (X = Y)[v_{XY}^{a\,b}]$, como queríamos probar.

NBG-2: $\overline{M} \vDash \bigwedge XY \bigvee Z \bigwedge u(u \in Z \leftrightarrow u \in X \wedge u \in Y)$.

Sean a y b en $\overline{U}$. Hay que comprobar que existe un c en $\overline{U}$ tal que para todo d en U se cumple $\overline{M}(\in)(d, c)$ syss $\overline{M}(\in)(d, a)$ y $\overline{M}(\in)(d, b)$.

Si a y b están en U, entonces usamos que

$$M \vDash \bigwedge XY \bigvee Z \bigwedge U(U \in Z \leftrightarrow U \in X \wedge U \in Y),$$

pues es un teorema de ZF*, e interpretando X e Y como ay b obtenemos el c que buscamos.

Si $a = [\alpha_i; a_1, \ldots, a_{m_i}]$ y $b = [\alpha_j; b_1, \ldots, b_{m_j}]$ están en C, podemos suponer que las variables libres de α_i son $x_0, \ldots, x_{m_i}$ y que las variables libres de α_j son $x_0, x_{m_i+1}, \ldots, x_{m_i+m_j}$. Sea $\alpha_k \equiv \alpha_i \wedge \alpha_j$. Así, para todo d en U se cumple

$$R_k[a_1, \ldots, a_{m_i-1}, b_1, \ldots, b_{m_j-1}](d) \quad \text{syss} \quad M \vDash \alpha_i \wedge \alpha_j[w],$$

donde w es una valoración que cumple

$$w(x_i) = \begin{cases} d & \text{si } i = 0, \\ a_i & \text{si } 1 \le i \le m_i, \\ b_{m_i+i} & \text{si } m_i < i \le m_i + m_j. \end{cases}$$

A su vez, esto equivale a que $M \vDash \alpha_i[w]$ y $M \vDash \alpha_j[w]$, o también a $R_i[a_1, \ldots, a_{m_i-1}](d)$ y $R_j[b_1, \ldots, b_{m_j-1}](d)$, es decir, equivale a $\overline{M}(\in)(d, a)$ y $\overline{M}(\in)(d, b)$. Por $(*)$ existe un c en $\overline{U}$ tal que para todo d de U,

$$\overline{M}(\in)(d, c) \qquad \text{syss} \qquad \overline{M}(\in)(d, a) \text{ y } \overline{M}(\in)(d, b),$$

como habíamos de probar.

Si a está en U y $b = [\alpha_j; b_1, \ldots, b_{m_j}]$ está en C, sean $y_0, \ldots, y_{m_j}$ las variables libres de α_j y sea y_{m_j+1} otra variable de índice posterior. Consideremos la fórmula $\alpha_k \equiv \alpha_j \wedge y_0 \in y_{m_j+1}$. Para todo d en U se cumple $R_k[b_1, \ldots, b_{m_j}, a](d)$ syss $M \vDash (\alpha_j \wedge y_0 \in y_{m_j+1})[w]$, donde la valoración w cumple

$$w(y_i) = \begin{cases} d & \text{si } i = 0, \\ b_i & \text{si } 1 \le i \le m_j, \\ a & \text{si } i = m_j + 1. \end{cases}$$

Esto equivale a que $M \vDash \alpha_j[w]$ y $M(\in)(d, a)$, o sea, a que $\overline{M}(\in)(d, a)$ y $\overline{M}(\in)(d, b)$. Como en el caso anterior, ahora basta aplicar $(*)$. Si a está en C y b está en U se razona análogamente.

NBG-3: $\overline{M} \vDash \bigwedge X \bigvee Y \bigwedge u(u \in Y \leftrightarrow u \notin X)$.

Sea v una valoración en M. Sea a en $\overline{U}$. Hemos de probar que existe un b en $\overline{U}$ tal que para todo c en U se cumple $\overline{M}(\in)(c,b)$ syss no $\overline{M}(\in)(c,a)$.

Si a está en U tomamos $\alpha_i \equiv x_0 \notin x_1$. Para todo c en U se cumple

$$R_i[a](c) \quad \text{syss} \quad M \vDash x_0 \notin x_1[v^{c\ a}_{x_0 x_1}] \quad \text{syss} \quad \text{no } M(\in)(c,a).$$

Por $(*)$ existe un b en $\overline{U}$ que cumple lo pedido.

Si $a = [\alpha_i; a_1, \ldots, a_{m_i}]$ está en C tomamos $\alpha_j \equiv \neg\alpha_i$ y se comprueba fácilmente que para todo c en U

$$R_j[a_1, \ldots, a_{m_i}] \quad \text{syss} \quad \text{no } \overline{M}(\in)(d,a),$$

y llegamos a la misma conclusión.

NBG-4: $\overline{M} \vDash \bigwedge uv \bigvee y \bigwedge x(x \in y \leftrightarrow x = u \vee x = v)$.

Es consecuencia inmediata de que $M \vDash$ ZF-2. Teniendo en cuenta que tanto M como $\overline{M}$ cumplen los axiomas de extensionalidad y del par, es claro que si v es una valoración en M entonces $M(\{x,y\})[v] = \overline{M}(\{x,y\})[v]$, pues ambos términos denotan al único elemento de U al cual pertenecen exactamente $v(x)$ y $v(y)$. De aquí se sigue a su vez que $M((x,y))[v] = \overline{M}((x,y))[v]$.

NBG-5: $\overline{M} \vDash \bigvee A \bigwedge xy((x,y) \in A \leftrightarrow x \in y)$.

Sea $\alpha_i \equiv \bigvee x_1 x_2(x_0 = (x_1, x_2) \wedge x_1 \in x_2)$. Para todo b de U se cumple $R_i(b)$ syss $M \vDash \bigvee x_1 x_2(x_0 = (x_1, x_2) \wedge x_1 \in x_2)[v^b_{x_0}]$, lo que a su vez equivale a $\overline{M} \vDash \bigvee x_1 x_2(x_0 = (x_1, x_2) \wedge x_1 \in x_2)[v^b_{x_0}]$.

Por $(*)$ hay un c en $\overline{U}$ de modo que para todo b de U se cumple $\overline{M}(\in)(b,c)$ syss $\overline{M} \vDash \bigvee x_1 x_2(x_0 = (x_1, x_2) \wedge x_1 \in x_2)[v^b_{x_0}]$. De aquí se sigue que

$$\overline{M} \vDash \bigvee A \bigwedge x_0(x_0 \in A \leftrightarrow \bigvee xy(x_0 = (x,y) \wedge x \in y).$$

Usando únicamente los axiomas que ya hemos probado que se cumplen en $\overline{M}$, es fácil ver que esta sentencia implica NBG-5, luego $\overline{M} \vDash$ NBG-5.

NBG-6: $\overline{M} \vDash \bigwedge A \bigvee B \bigwedge x(x \in B \leftrightarrow \bigvee y\,(x,y) \in A)$.

Sea v una valoración en U. Hemos de probar que, fijado a en $\overline{U}$, existe un b en $\overline{U}$ tal que para todo c de U se cumple $\overline{M}(\in)(c,b)$ syss $\overline{M} \vDash \bigvee y\,(x,y) \in A[v^{a\ c}_{Ax}]$.

Si a está en U, entonces usamos que

$$M \vDash \bigwedge A \bigvee B \bigwedge X(X \in B \leftrightarrow \bigvee Y\,(X,Y) \in A),$$

pues es un teorema de ZF*, e interpretando A como a obtenemos el b que buscamos.

Supongamos ahora que $a = [\alpha_i; a_1, \ldots, a_{m_i}]$ está en C. Sean $y_0, \ldots, y_{m_i}$ las variables libres de α_i. Sea $\alpha_j \equiv \bigvee y\, \mathsf{S}^{(y_0,y)}_{y_0} \alpha_i$. Para todo c de U se cumple $R_j[a_1, \ldots, a_{m_i}](c)$ syss $M \vDash \bigvee y\, \mathsf{S}^{(y_0,y)}_{y_0} \alpha_i[w]$, donde

$$w(y_i) = \begin{cases} c & \text{si } i = 0, \\ a_i & \text{si } 1 \le i \le m_i. \end{cases}$$

Esto equivale a que existe un d en U tal que $M \vDash \mathtt{S}_{y_0}^{(y_0,y)}\alpha_i[w_y^d]$ o, lo que es lo mismo, $M \vDash \alpha_i[w_{yy_0}^{dM((y_0,y))[w_y^d]}]$. A su vez, esto es $\overline{M}(\in)(M((y_0,y))[w_y^d], a)$.

Según hemos visto tras la prueba de NBG-4, se cumple $M((y_0,y))[w_y^d] = \overline{M}((y_0,y))[w_y^d]$. Por lo tanto, para todo c de U, se cumple

$$R_j[a_1,\dots,a_{m_i}](c) \quad \text{syss} \quad \overline{M} \vDash \bigvee y\,(y_0,y) \in A[v_{y_0A}^{c\,a}].$$

Aplicando $(*)$ obtenemos el b que buscamos.

NBG-7: $\overline{M} \vDash \bigwedge A \bigvee B \bigwedge xy((x,y) \in B \leftrightarrow x \in A)$.

Sea v una valoración en M. Sea a en $\overline{U}$. Si a está en U consideramos $\alpha_i \equiv \bigvee UV(x_0 = (U,V) \wedge U \in x_1)$. Así, para todo c de U se cumple $R_i[a](c)$ syss $M \vDash \bigvee UV(x_0 = (U,V) \wedge U \in x_1)[v_{x_0x_1}^{c\ a}]$ syss existen d y e en U tales que $c = M((U,V))[v_{UV}^{d\,e}]$ y $M(\in)(d,a)$.

Por $(*)$ existe un b en $\overline{U}$ tal que para todo c en U se cumple $\overline{M}(\in)(c,b)$ syss existen d y e en U tales que $c = \overline{M}((U,V))[v_{UV}^{d\,e}]$ y $\overline{M}(\in)(d,a)$. Por consiguiente

$$\overline{M} \vDash \bigwedge A \bigvee B \bigwedge z(z \in B \leftrightarrow \bigvee xy(z = (x,y) \wedge x \in A)), \tag{9.1}$$

y esta sentencia implica NBG-7.

Supongamos ahora que $a = [\alpha_i; a_1, \dots, a_{m_i}]$ está en C y consideremos la fórmula $\alpha_j \equiv \bigvee UV(y_0 = (U,V) \wedge \mathtt{S}_{y_0}^U\alpha_i)$. Para todo c de U se cumple $R_j[a_1,\dots,a_{m_i}](c)$ syss $M \vDash \bigvee UV(y_0 = (U,V) \wedge \mathtt{S}_{y_0}^U\alpha_i)[w]$, donde

$$w(y_i) = \begin{cases} c & \text{si } i = 0, \\ a_i & \text{si } 1 \le i \le m_i, \end{cases}$$

syss existen d y e en U tales que $c = M((U,V))[w_{UV}^{d\,e}]$ y $\overline{M}(\in)(d,a)$. Por $(*)$ existe un b en $\overline{U}$ tal que para todo c de U se cumple $\overline{M}(\in)(c,b)$ syss $c = \overline{M}((U,V))[w_{UV}^{d\,e}]$ y $\overline{M}(\in)(d,a)$. Por consiguiente $\overline{M}$ cumple (9.1) y también NBG-7.

La comprobación de NBG-8, NBG-9 y NMG-10 es similar a la de NBG-7. Los axiomas NBG-11 y NBG-12 (axioma del conjunto vacío y de la unión) se cumplen en $\overline{M}$ como consecuencia directa de que M cumple los axiomas análogos de ZF* (ZF-3 y ZF-4).

NBG-13: $\overline{M} \vDash \bigwedge xA(\mathrm{Un}A \to \bigvee y \bigwedge u(u \in y \leftrightarrow \bigvee v \in x\,(v,u) \in A))$.

Sea v una valoración en M. Tomamos a en $\overline{U}$ y b en U y suponemos que $\overline{M} \vDash \mathrm{Un}\,A[v_A^a]$. Si a está en U obtenemos la tesis como consecuencia de que

$$M \vDash \bigwedge AX \bigvee Y \bigwedge U(U \in Y \leftrightarrow \bigvee V \in X\,(V,U) \in A)).$$

Es efecto, esto es un teorema de ZF* (basta tomar $Y = \mathcal{R}(B)$, donde $B = \{Z \in A \mid \bigvee UV(Z = (U,V) \wedge U \in X)\}$).

Supongamos ahora que $a = [\alpha_i; a_1, \ldots, a_{m_i}]$ está en C. Sean $y_0, \ldots, y_{m_i}$ las variables libres de α_i. Como $\overline{M} \vDash \mathrm{Un}\, A[v^a_A]$, tenemos que para todos los c, d, e de U, si se cumple $\overline{M}(\in)(\overline{M}((U,V))[v^{c\,d}_{UV}], a)$ y $\overline{M}(\in)(\overline{M}((U,V))[v^{c\,e}_{UV}], a)$, entonces $d = e$. En otros términos, si

$$M \vDash \alpha_i[w_{y_0}^{M((U,V))[w^{c\,d}_{UV}]}] \quad \text{y} \quad M \vDash \alpha_i[w_{y_0}^{M((U,V))[w^{c\,e}_{UV}]}], \text{ entonces } d = e,$$

donde $w(y_i) = a_i$ para $1 \leq i \leq m_i$.

Si llamamos $\phi(U,V) \equiv \mathsf{S}^{(U,V)}_{y_0}\alpha_i$ esto se traduce en que si $M \vDash \phi[w^{c\,d}_{UV}]$ y $M \vDash \phi[w^{c\,e}_{UV}]$ entonces $d = e$, es decir,

$$M \vDash \bigwedge UVW(\phi(U,V) \wedge \phi(U,W) \to V = W)[w].$$

Puesto que M satisface ZF-5, de aquí se sigue que

$$M \vDash \bigwedge A \bigvee B \bigwedge V(V \in B \leftrightarrow \bigvee U(U \in A \wedge \phi(U,V)))[w].$$

En particular, existe un c en U tal que para todo d de U se cumple $\overline{M}(\in)(d,c)$ syss existe un e en U tal que $\overline{M}(\in)(e,b)$ y $M \vDash \phi(U,V)[w^{e\,d}_{UV}]$. Pero

$$M \vDash \phi(U,V)[w^{e\,d}_{UV}] \text{ syss } M \vDash \alpha_i[w_{y_0}^{M((U,V))[w^{e\,d}_{UV}]}] \text{ syss}$$

$$\overline{M}(\in)(\overline{M}((U,V))[w^{e\,d}_{UV}], a) \text{ syss } \overline{M} \vDash (U,V) \in A[w^{e\,d\,a}_{UVA}].$$

Por consiguiente tenemos que $\overline{M}(\in)(d,c)$ syss existe un e en U de manera que $\overline{M}(\in)(e,b)$ y $\overline{M} \vDash (U,V) \in A[w^{e\,d\,a}_{UVA}]$. Equivalentemente:

$$\overline{M} \vDash \bigvee y \bigwedge u(u \in y \leftrightarrow \bigvee v \in x\,(v,u) \in A))[w^{b\,a}_{xA}].$$

Esto es lo que teníamos que probar. ∎

De aquí se deduce un resultado notable:

Teorema 9.7 *Una sentencia α es un teorema de* ZF* *si y sólo si α^{V} lo es de* NBG*.

Demostración: Si $\underset{\mathrm{ZF}^*}{\vdash} \alpha$ y M es un modelo (numerable) de NBG*, entonces $\underline{M}$ es un modelo de ZF*, luego $\underline{M} \vDash \alpha$ y según el teorema 9.3 se cumple que $M \vDash \alpha^{\mathrm{V}}$. Así pues, α^{V} es verdadera en todos los modelos (numerables) de NBG*. Por el teorema de completitud $\underset{\mathrm{NBG}^*}{\vdash} \alpha^{\mathrm{V}}$.

Recíprocamente, si $\underset{\mathrm{NBG}^*}{\vdash} \alpha^{\mathrm{V}}$ y M es un modelo (numerable) de ZF*, entonces $\overline{M}$ es un modelo de NBG*, luego $\overline{M} \vDash \alpha^{\mathrm{V}}$, luego $\underline{\overline{M}} \vDash \alpha$, pero por construcción es claro que $\underline{\overline{M}}$ no es sino el modelo M de partida. Así pues, $M \vDash \alpha$ y, como α es verdadera en todo modelo (numerable) de ZF*, por el teorema de completitud se cumple $\underset{\mathrm{ZF}^*}{\vdash} \alpha$. ∎

Más en general:

Teorema 9.8 *Sea* Γ *una colección de sentencias y sea* Γ^{V} *la colección de sus relativizaciones. Sea* α *otra sentencia. Entonces*

$$\Gamma \underset{\mathrm{ZF}^*}{\vdash} \alpha \quad \textit{si y sólo si} \quad \Gamma^{\mathrm{V}} \underset{\mathrm{NBG}^*}{\vdash} \alpha^{\mathrm{V}}.$$

Demostración: Si $\Gamma \underset{\mathrm{ZF}^*}{\vdash} \alpha$, existen sentencias $\gamma_1, \ldots, \gamma_n$ en Γ tales que $\gamma_1 \wedge \cdots \wedge \gamma_n \underset{\mathrm{ZF}^*}{\vdash} \alpha$. Por el teorema de deducción $\underset{\mathrm{ZF}^*}{\vdash} \gamma_1 \wedge \cdots \wedge \gamma_n \to \alpha$. Por el teorema anterior $\underset{\mathrm{NBG}^*}{\vdash} \gamma_1^{\mathrm{V}} \wedge \cdots \wedge \gamma_n^{\mathrm{V}} \to \alpha^{\mathrm{V}}$, luego $\Gamma^{\mathrm{V}} \underset{\mathrm{NBG}^*}{\vdash} \alpha^{\mathrm{V}}$.

La implicación contraria es análoga. ∎

De este modo, a cada extensión T de ZF^* mediante unos axiomas adicionales Γ le corresponde una extensión T' de NBG^* (la que tiene por axiomas Γ^{V}) equivalente en el sentido del teorema anterior: una sentencia α es un teorema de T si y sólo si su relativización lo es de T'. En particular, la extensión de NBG^* correspondiente a ZFC es NBG, por lo que podemos afirmar que NBG es consistente si y sólo si ZFC es consistente ($\varnothing \neq \varnothing$ es un teorema de una si y sólo si lo es de la otra). Lo mismo vale para $\mathrm{ZFC}{-}AI$ y $\mathrm{NBG}{-}AI$, pero en la sección primera hemos probado que $\mathrm{ZFC}{-}AI$ es consistente, luego ahora podemos afirmar que $\mathrm{NBG}{-}AI$ también lo es. Más aún, hemos visto cómo construir un modelo explícito de $\mathrm{NBG}{-}AI$ a partir del modelo HF.

Por otro lado, no debemos pensar que las extensiones de ZF^* se corresponden con las extensiones de NBG^*: si añadimos a NBG^* un axioma que hable de clases propias y, por consiguiente, que no sea la relativización de ninguna sentencia, entonces obtenemos una extensión de NBG^* que no se corresponde necesariamente con ninguna extensión de ZF^*.

Capítulo X

La formalización de la lógica en teoría de conjuntos

La teoría de conjuntos (cualquiera de ellas) es una teoría axiomática lo suficientemente potente como para formalizar cualquier razonamiento matemático. En su seno se demuestran resultados sobre números, sobre geometría, sobre juegos de azar, sobre el movimiento de los planetas, sobre fluidos, sobre electrones y rayos de luz, etc. Nada nos impide, pues, usar la teoría de conjuntos para estudiar la lógica matemática, ahora ya libres de las precauciones que nos exigía el trabajar metamatemáticamente. Con ello no sólo convertiremos a la lógica en una rama más de la matemática, al lado del álgebra, la geometría o el análisis matemático, sino que la formalización de la lógica nos permitirá examinar más de cerca los razonamientos que conducen a los teoremas de incompletitud en el caso de más interés. A todo ello nos dedicaremos en este capítulo. Mientras no se indique lo contrario, todas las demostraciones de este capítulo se hacen en NBG^* más el axioma de infinitud. Llamaremos NBG^- a esta teoría axiomática. Equivalentemente, podemos trabajar en ZF^-, es decir, $\mathrm{ZF}^* + AI$.

10.1 Lenguajes formales

Empezamos formalizando los conceptos que introdujimos en los capítulos I y II. Necesitamos algunos conceptos elementales relacionados con las sucesiones finitas:

Sucesiones finitas Dada una clase A, llamaremos

$$A^{<\omega} = \bigcup_{n\in\mathbb{N}} A^n,$$

es decir, $A^{<\omega}$ es la clase de todas las funciones cuyo dominio es un número natural y cuya imagen está en A o, equivalentemente, la clase de todas las sucesiones finitas de elementos de A (recordemos que, por construcción, un

número natural es el conjunto de los números menores que él mismo). Es claro que si A es un conjunto entonces $A^{<\omega}$ también lo es (se prueba por inducción que A^n es un conjunto y luego se aplica el axioma de la unión).

Si $s \in A^{<\omega}$ llamaremos *longitud* de s a $\operatorname{long} s = \mathcal{D}s \in \mathbb{N}$. A menudo representaremos las sucesiones de longitud n de la forma $\{a_i\}_{i<n}$.

Lenguajes formales Llamaremos *lenguaje formal* (de primer orden) a toda óctupla ordenada $\mathfrak{L} = (\neg, \to, \bigwedge, |, x, c, R, f)$ que cumpla las condiciones siguientes:

a) $x : \mathbb{N} \longrightarrow \mathrm{V}$ inyectiva. Escribiremos x_i en lugar de $x(i)$.

b) $c : \mathbb{N} \longrightarrow \mathrm{V}$ inyectiva o bien existe un $n \in \mathbb{N}$ tal que $c : n \longrightarrow \mathrm{V}$ inyectiva. Si $i \in \mathcal{D}c$, escribiremos c_i en lugar de $c(i)$.

c) $R : A \subset (\mathbb{N} \setminus \{0\}) \times \mathbb{N} \longrightarrow \mathrm{V}$ inyectiva. Si $(n, i) \in \mathcal{D}R$ escribiremos R_i^n en lugar de $R(n, i)$. Se cumple que $(2, 0) \in \mathcal{D}R$ y a R_0^2 lo llamaremos $=$.

d) $f : B \subset (\mathbb{N} \setminus \{0\}) \times \mathbb{N} \longrightarrow \mathrm{V}$ inyectiva. Si $(n, i) \in \mathcal{D}f$ escribiremos f_i^n en lugar de $f(n, i)$.

e) Los conjuntos $\operatorname{Var} \mathfrak{L} = \mathcal{R}x$, $\operatorname{Const} \mathfrak{L} = \mathcal{R}c$, $\operatorname{Rel} \mathfrak{L} = \mathcal{R}R$ y $\operatorname{Fun} \mathfrak{L} = \mathcal{R}f$ son disjuntos dos a dos y no contienen a ninguno de los conjuntos $\neg$, $\to$, $\bigwedge$, $|$, los cuales son también distintos entre sí.

Llamaremos *variables, constantes, relatores y funtores* de $\mathfrak{L}$ a los elementos de los conjuntos $\operatorname{Var} \mathfrak{L}$, $\operatorname{Const} \mathfrak{L}$, $\operatorname{Rel} \mathfrak{L}$ y $\operatorname{Fun} \mathfrak{L}$ respectivamente. A los conjuntos $\neg$, $\to$, $\bigwedge$ y $|$ los llamaremos *negador, implicador, generalizador y descriptor* de $\mathfrak{L}$ respectivamente.

Llamaremos *relatores n-ádicos* y *funtores n-ádicos* de $\mathfrak{L}$ a los elementos de los conjuntos $\operatorname{Rel}_n \mathfrak{L} = \{R_i^n \mid (n, i) \in \mathcal{D}R\}$ y $\operatorname{Fun}_n \mathfrak{L} = \{f_i^n \mid (n, i) \in \mathcal{D}f\}$ respectivamente.

Llamaremos *signos* de $\mathfrak{L}$ a los elementos del conjunto

$$\operatorname{Sig} \mathfrak{L} = \{\neg, \to, \bigwedge, |\} \cup \operatorname{Var} \mathfrak{L} \cup \operatorname{Const} \mathfrak{L} \cup \operatorname{Rel} \mathfrak{L} \cup \operatorname{Fun} \mathfrak{L}.$$

Nota Si tratamos con un lenguaje formal $\mathfrak{L}$ en el sentido que acabamos de definir, hemos de prestar atención para no confundir los signos de $\mathfrak{L}$ con los signos (metamatemáticos) del lenguaje de la teoría de conjuntos que representamos con la misma notación. Por ejemplo, ahora $===$ es una fórmula correcta, siempre y cuando entendamos que los signos de los extremos representan ambos al conjunto $=\in \operatorname{Sig} \mathfrak{L}$, mientras que el signo central es el igualador metamatemático.

Cadenas de signos Llamaremos *cadenas de signos* de $\mathfrak{L}$ a los elementos del conjunto

$$\operatorname{Cad} \mathfrak{L} = (\operatorname{Sig} \mathfrak{L})^{<\omega} \setminus \{\varnothing\}.$$

Si ζ_1 y $\zeta_2 \in$Cad $\mathfrak{L}$ definimos $\zeta_1\zeta_2 : \operatorname{long}\zeta_1 + \operatorname{long}\zeta_2 \longrightarrow$ Sig $\mathfrak{L}$ mediante

$$\zeta_1\zeta_2(i) = \begin{cases} \zeta_1(i) & \text{si } i < \operatorname{long}\zeta_1, \\ \zeta_2(i - \operatorname{long}\zeta_1) & \text{si } \operatorname{long}\zeta_1 \leq i. \end{cases}$$

Es fácil ver que esta operación en Cad $\mathfrak{L}$ es asociativa. Si $\{\zeta_i\}_{i=0}^n$ es una sucesión de cadenas de signos, definimos inductivamente

$$\prod_{i=0}^{0}\zeta_i = \zeta_0, \qquad \prod_{i=0}^{j+1}\zeta_i = \left(\prod_{i=0}^{j}\zeta_i\right)\zeta_{j+1}, \qquad \text{para } j < n.$$

En la práctica escribiremos $\zeta_0 \cdots \zeta_n$ en lugar de $\prod_{i=0}^{n}\zeta_i$. Así mismo, no distinguiremos entre signos y cadenas con un solo signo. Por ejemplo, si $\zeta \in$ Cad $\mathfrak{L}$, cuando escribamos $\zeta \in$ Var $\mathfrak{L}$ querremos decir que long $\zeta = 1$ y $\zeta(0) \in$ Var $\mathfrak{L}$.

Términos y fórmulas Definimos por recurrencia la sucesión que a cada $k \in \mathbb{N}$ le asigna el par $(\operatorname{Term}_k\mathfrak{L}, \operatorname{Form}_k\mathfrak{L})$ del modo siguiente:

- $\operatorname{Term}_0\mathfrak{L} = \text{Var } \mathfrak{L} \cup \text{Const } \mathfrak{L}, \qquad \operatorname{Form}_0\mathfrak{L} = \varnothing.$
- $\operatorname{Term}_{k+1}\mathfrak{L} = \operatorname{Term}_k\mathfrak{L} \cup \{ft_0 \cdots t_{n-1} \mid n \in \mathbb{N} \wedge f \in \operatorname{Fun}_n\mathfrak{L} \wedge \{t_i\}_{i<n} \in (\operatorname{Term}_k\mathfrak{L})^n\} \cup \{|x\alpha \mid x \in \text{Var } \mathfrak{L} \wedge \alpha \in \operatorname{Form}_k\mathfrak{L}\}.$
- $\operatorname{Form}_{k+1}\mathfrak{L} = \operatorname{Form}_k\mathfrak{L} \cup \{Rt_0 \cdots t_{n-1} \mid n \in \mathbb{N} \wedge f \in \operatorname{Rel}_n\mathfrak{L} \wedge \{t_i\}_{i<n} \in (\operatorname{Term}_k\mathfrak{L})^n\} \cup \{\neg\alpha \mid \alpha \in \operatorname{Form}_k\mathfrak{L}\} \cup \{\rightarrow\alpha\beta \mid \alpha, \beta \in \operatorname{Form}_k\mathfrak{L}\} \cup \{\bigwedge x\alpha \mid x \in \text{Var } \mathfrak{L} \wedge \alpha \in \operatorname{Form}_k\mathfrak{L}\}.$

Definimos los *términos* y las *fórmulas* de $\mathfrak{L}$ como los elementos de los conjuntos

$$\text{Term } \mathfrak{L} = \bigcup_{k\in\mathbb{N}} \operatorname{Term}_k\mathfrak{L}, \qquad \text{Form } \mathfrak{L} = \bigcup_{k\in\mathbb{N}} \operatorname{Form}_k\mathfrak{L},$$

respectivamente. Llamaremos *expresiones* de $\mathfrak{L}$ a los elementos del conjunto Exp $\mathfrak{L}$ =Term $\mathfrak{L} \cup$ Form $\mathfrak{L}$.

Usaremos los convenios habituales de notación, es decir, escribiremos $\alpha \rightarrow \beta$ en lugar de $\rightarrow\alpha\beta$, etc.

Es fácil probar el siguiente principio de inducción:

Teorema 10.1 *Sea A un conjunto de expresiones de un lenguaje formal $\mathfrak{L}$ tal que*

a) Var $\mathfrak{L} \subset A$,

b) Const $\mathfrak{L} \subset A$,

c) si $t_1, \ldots, t_n \in A$ y $R \in \operatorname{Rel}_n\mathfrak{L}$, entonces $Rt_1 \cdots t_n \in A$,

d) si $t_1, \ldots, t_n \in A$ *y* $f \in \mathrm{Fun}_n\mathfrak{L}$*, entonces* $ft_1 \cdots t_n \in A$*,*

e) si $\alpha \in A$ *entonces* $\neg\alpha \in A$*,*

f) si $\alpha, \beta \in A$ *entonces* $\alpha \to \beta \in A$*,*

g) si $\alpha \in A$ *y* $x \in Var\ \mathfrak{L}$ *entonces* $\bigwedge x\alpha \in A$*,*

h) si $\alpha \in A$ *y* $x \in Var\ \mathfrak{L}$ *entonces* $x|\alpha \in A$*,*

entonces $A = \mathrm{Exp}\ \mathfrak{L}$*.*

La prueba consiste en demostrar que $\mathrm{Term}_k\mathfrak{L} \cup \mathrm{Form}_k\mathfrak{L} \subset A$ por inducción sobre k. Así mismo se demuestra un principio de recursión en virtud del cual, para definir una función sobre $\mathrm{Exp}\ \mathfrak{L}$, basta definirla sobre las variables y constantes, definirla sobre las fórmulas $Rt_1 \cdots t_n$ supuesta definida sobre $t_1, \ldots, t_n$, definirla sobre los términos $ft_1 \cdots t_n$ supuesta definida sobre $t_1, \ldots, t_n$, definirla sobre $\neg\alpha$ supuesta definida sobre α, etc.

Ahora es fácil definir formalmente los conceptos de *variable libre, variable ligada* y la función *sustitución* $\mathtt{S} : \mathrm{Exp}\ \mathfrak{L} \times \mathrm{Var}\ \mathfrak{L} \times \mathrm{Term}\ \mathfrak{L} \longrightarrow \mathrm{Exp}\ \mathfrak{L}$. Lo dejamos como ejercicio.

Tampoco ofrece ninguna dificultad la definición de los axiomas lógicos. Por ejemplo,

$$K_1(\mathfrak{L}) = \{\gamma \in \mathrm{Form}\ \mathfrak{L} \mid \bigvee \alpha\beta \in \mathrm{Form}\ \mathfrak{L}\ \gamma = \alpha \to (\beta \to \alpha)\},$$

e igualmente se definen $K_2(\mathfrak{L}), \ldots, K_8(\mathfrak{L})$. Definimos los *axiomas lógicos* de $\mathfrak{L}$ como los elementos del conjunto $\mathrm{Axl}(\mathfrak{L}) = K_1(\mathfrak{L}) \cup \cdots \cup K_8(\mathfrak{L})$.

Definimos $\Gamma \underset{\mathfrak{L}}{\overset{D}{\vdash}} \alpha \equiv \alpha \in \mathrm{Form}\ \mathfrak{L} \wedge \Gamma \subset \mathrm{Form}\ \mathfrak{L} \wedge \bigvee n \in \mathbb{N}(D \in (\mathrm{Form}\ \mathfrak{L})^{n+1}$ $\wedge\ D_n = \alpha \wedge \bigwedge i \in n+1(D_i \in \mathrm{Axl}\ (\mathfrak{L}) \vee D_i \in \Gamma \vee \bigvee uv \in i(D_u = D_v \to D_i \vee \bigvee x \in \mathrm{Var}\ \mathfrak{L}\ D_i = \bigwedge xD_u)))$.

Así mismo, definimos $\Gamma \underset{\mathfrak{L}}{\vdash} \alpha \equiv \bigvee D \in (\mathrm{Form}\ \mathfrak{L})^{<\omega}\ \Gamma \underset{\mathfrak{L}}{\overset{D}{\vdash}} \alpha$.

A partir de estas definiciones, todos los resultados de los capítulos I y II pueden considerarse como teoremas de la teoría de conjuntos. Notemos que no decimos que "pueden demostrarse", sino que "pueden considerarse teoremas", en el sentido de que las demostraciones vistas en su momento satisfacen ahora todos los requisitos de rigor que los matemáticos exigen a sus pruebas. Desde un punto de vista estricto deben considerarse como esbozos de demostraciones, pues contienen muchos saltos lógicos que el lector puede llenar, pero debemos tener presente que absolutamente todos los libros de matemáticas publicados en el mundo demuestran sus teoremas a este mismo nivel de semiformalización, ya que las pruebas completamente formalizadas (indicado que esto es por modus ponens y aquello por modus tollens) serían insoportables de leer.

10.2 Modelos

Nos ocupamos ahora de formalizar los conceptos básicos que hemos visto de teoría de modelos. Empezamos por la definición:

Definición 10.2 Un *modelo* de un lenguaje formal $\mathfrak{L}$ es una terna $M = (U, I, a)$, donde U es un conjunto, $a \in U$ e I es una función cuyo dominio es el conjunto $\text{Const}\,\mathfrak{L} \cup \text{Rel}\,\mathfrak{L} \cup \text{Fun}\,\mathfrak{L}$ tal que

- Si $c \in \text{Const}\,\mathfrak{L}$, entonces $I(c) \in U$.
- Si $R \in \text{Rel}_n\mathfrak{L}$, entonces $I(R) \subset U^n$. Además

$$I(=) = \{s \in U^2 \mid s(0) = s(1)\}.$$

- Si $f \in \text{Fun}_n\mathfrak{L}$, entonces $I(f) : U^n \longrightarrow U$.

En la práctica escribiremos M en lugar de U o I.

Una *valoración* en M es una aplicación $v : \text{Var}\,\mathfrak{L} \longrightarrow M$. Llamaremos $\text{Val}(M)$ al conjunto[1] de todas las valoraciones en M.

Si $v \in \text{Val}(M)$, $x \in \text{Var}\,\mathfrak{L}$ y $u \in M$, llamaremos

$$v_x^u = \big(v \setminus \{(x, v(x))\}\big) \cup \{(x, u)\}.$$

La definición de los conceptos de denotación y satisfacción presenta algunos detalles técnicos delicados. Veámosla en primer lugar y después los comentamos:

Si M es un modelo de un lenguaje formal $\mathfrak{L}$, sean V y F dos conjuntos cualesquiera que no estén en M. Llamamos s a la única aplicación que a cada $\theta \in \text{Exp}\,\mathfrak{L}$ le asigna una función $s(\theta) : \text{Val}(M) \longrightarrow M \cup \{V, F\}$ de acuerdo con las condiciones siguientes:

a) $s(x)(v) = v(x)$,

b) $s(c)(v) = M(c)$,

c) $s(Rt_1 \cdots t_n)(v) = V$ si $M(R)(s(t_1)(v), \ldots, s(t_n)(v))$ y $s(Rt_1 \cdots t_n)(v) = F$ en caso contrario.

d) $s(ft_1 \cdots t_n)(v) = M(f)(s(t_1)(v), \ldots, s(t_n)(v))$.

e) $s(\neg\alpha)(v) = V$ si $s(\alpha)(v) = F$ y $s(\neg\alpha)(v) = F$ en caso contrario.

f) $s(\alpha \to \beta)(v) = V$ si $s(\alpha)(v) = F$ o $s(\beta)(v) = V$ y en caso contrario $s(\alpha \to \beta)(v) = F$.

[1] Para probar $\text{Val}(M)$ que es un conjunto hace falta el axioma de partes. No obstante, conviene tener presente que podríamos trabajar únicamente con valoraciones definidas sobre conjuntos finitos de variables, y entonces no necesitaríamos este axioma.

g) $s(\bigwedge x\alpha)(v) = V$ si $\bigwedge u \in M\, s(\alpha)(v^u_x) = V$ y en otro caso $s(\bigwedge x\alpha)(v) = F$.

h) $s(x|\alpha)(v) =$ el único $u \in M$ tal que $s(\alpha)(v^u_x) = V$ si existe tal único elemento o $s(x|\alpha)(v) = d$ (la descripción impropia de M) en caso contrario.

Escribiremos $M(t)[v]$ en lugar de $s(t)(v)$ y $M \vDash \alpha[v]$ en lugar de $s(\alpha)(v) = V$. Con esta notación, las condiciones que definen a s se convierten en las de la definición 3.2.

Así pues, para definir la fórmula $M \vDash \alpha[v]$ hemos definido recurrentemente una función s que asigna a α y a v el valor V o F según si $M \vDash \alpha[v]$ se cumple o no. El teorema de recursión que justifica la existencia y unicidad de s se remite en último extremo al teorema 8.14, de modo que lo que hacemos es asignar a cada número natural k una función $s_k : \mathrm{Term}_k\mathcal{L} \cup \mathrm{Form}_k\mathcal{L} \longrightarrow (M \cup \{V, F\})^{\mathrm{Val}\ M}$ y después definir s como la unión de todas las funciones s_k. El punto a destacar es que la definición de s pasa por la construcción de una función $k \mapsto s_k$, lo cual exige que cada s_k sea un conjunto (pues (k, s_k) ha de ser un elemento de la función referida). A su vez, esto exige que cada $s_k(\theta) : \mathrm{Val}\ M \longrightarrow M \cup \{V, F\}$ sea un conjunto, lo cual sucede si y sólo si $\mathrm{Val}\ M$ lo es, y esto a su vez sucede si y sólo si (el universo de) M es un conjunto.

En resumen, que nuestra definición de denotación y satisfacción no sería válida si hubiéramos permitido que los modelos tuvieran como universo una clase propia. Desde luego esto no justifica que no sea posible definir la denotación y satisfacción en modelos que sean clases propias mediante otra construcción diferente de la que acabamos de dar, pero veremos que no es así, que es esencialmente imposible trabajar con modelos que sean clases propias.

A partir de aquí podemos considerar como teoremas de la teoría de conjuntos todos los resultados vistos hasta el capítulo IV.

10.3 Lógica de segundo orden

Al estudiar la lógica desde la teoría de conjuntos en lugar de hacerlo metamatemáticamente podemos permitirnos muchas variaciones que metamatemáticamente serían más que cuestionables: podemos considerar lenguajes con fórmulas infinitas, o con una cantidad no numerable de signos o lenguajes de órdenes superiores. En esta sección esbozaremos un caso sencillo de lógica de segundo orden no porque tenga gran interés en general, sino porque es conveniente que el lector tenga claro a qué nos referimos cuando hablamos de lógica de primer orden.

Definición 10.3 Un *lenguaje formal con variables monádicas de segundo orden* se define como $\mathcal{L} = (\neg, \to, \bigwedge, |, x, X, c, R, f)$, donde $\mathcal{L}' = (\neg, \to, \bigwedge, |, x, c, R, f)$ es un lenguaje de primer orden y $X : \mathbb{N} \longrightarrow \mathrm{V}$ es una aplicación inyectiva en cuya imagen no hay ningún signo de $\mathcal{L}'$.

En definitiva, $\mathcal{L}$ se obtiene añadiendo a un lenguaje de primer orden un conjunto infinito de signos $X_i = X(i)$ a los que llamaremos *variables de segundo*

orden. A las variables x_i de $\mathcal{L}'$ las llamaremos ahora *variables de primer orden* de $\mathcal{L}$, y usaremos la notación $\mathrm{Var}_1\mathcal{L}$ y $\mathrm{Var}_2\mathcal{L}$ para referirnos a los conjuntos de variables de primer y segundo orden de $\mathcal{L}$.

Los conjuntos Sig $\mathcal{L}$ y Cad $\mathcal{L}$ se definen igual que para lenguajes de primer orden, salvo que ahora incluimos las variables de segundo orden entre los signos de $\mathcal{L}$. Las definiciones de términos y fórmulas tienen que ser retocadas. La intención es que, en un modelo, las variables de primer orden recorran los objetos del universo y las variables de segundo orden recorran las relaciones monádicas del universo o, lo que es lo mismo, los subconjuntos del universo. Por ello, la definición de los conjuntos Term $\mathcal{L}$ y Form $\mathcal{L}$ es idéntica salvo que ahora añadimos la condición:

Si t es un término de $\mathcal{L}$ y X es una variable de segundo orden, entonces $X(t)$ es una fórmula de $\mathcal{L}$.

Notemos que los paréntesis son superfluos. Podríamos escribir simplemente Xt sin que ello introdujera ambigüedades sintácticas (recordemos que, oficialmente, nuestros lenguajes formales no tienen paréntesis). En la parte de la definición que establece que $\bigwedge x\alpha$ es una fórmula ahora admitimos que x sea una variable de primer o de segundo orden. No obstante, por simplicidad exigiremos que, en la parte que establece que $x|\alpha$ es un término, la variable x sea de primer orden.[2]

Ejemplo Llamamos *lenguaje de la aritmética de Peano de segundo orden* al lenguaje de la aritmética de Peano (de primer orden) aumentado con un juego de variables de segundo orden. Llamaremos *axiomas de Peano de segundo orden* al siguiente conjunto se sentencias de este lenguaje:

a) $\bigwedge x\, \neg x' = 0$,

b) $\bigwedge xy(x' = y' \to x = y)$,

Ind $\bigwedge X\big((X(0) \wedge \bigwedge x(X(x) \to X(x')) \to \bigwedge x\, X(x)\big)$.

De los cinco axiomas de Peano 8.13, dos de ellos son triviales en este contexto (en un lenguaje que sólo puede hablar de números naturales, el cero tiene que ser —obviamente— un número natural y el siguiente de un número natural será otro número natural). Lo relevante es que el principio de inducción, que en la lógica de primer orden tiene que ser expresado (en realidad, parcialmente) por un esquema axiomático (es decir, mediante infinitos axiomas, uno para cada fórmula), con la lógica de segundo orden cabe en una única sentencia Ind.

Informalmente, la interpretación de Ind es la siguiente: "Para toda relación monádica X sobre $\mathbb{N}$ —o, equivalentemente, para todo subconjunto de $\mathbb{N}$—, si

[2]Podemos admitir variables de segundo orden en las descripciones, pero entonces hemos de distinguir entre términos de primer orden y términos de segundo orden y así, la estipulación de que $X(t)$ es una fórmula se habría de cambiar por "si T es un término de segundo orden y t es un término de primer orden, entonces $T(t)$ es una fórmula".

0 cumple X y cuando un número natural x cumple X lo mismo le ocurre a x', entonces todo número natural cumple X".

Para justificar esta interpretación debemos definir las nociones de modelo, denotación y satisfacción para la lógica de segundo orden.

Definición 10.4 Un *modelo* M de un lenguaje de segundo orden $\mathcal{L}$ se define como un modelo del lenguaje de primer orden que resulta de eliminar las variables de segundo orden. Una *valoración* de $\mathcal{L}$ en M es una aplicación v que a cada variable de primer orden x de $\mathcal{L}$ le asigne un objeto $v(x) \in M$ y a cada variable de segundo orden X de $\mathcal{L}$ le asigne un subconjunto $v(X) \subset M$.

La definición del objeto denotado por un término t respecto de una valoración v (representado por $M(t)[v]$) y de la satisfacción de una fórmula α respecto a una valoración (representada $M \vDash \alpha[v]$) se modifica únicamente en los puntos siguientes:

Se cumple $M \vDash X(t)[v]$ syss $M(t)[v] \in v(X)$.

Se cumple $M \vDash \bigwedge X\, \alpha[v]$ syss para todo $A \subset M$ se cumple $M \vDash \alpha[v^A_X]$, donde v^A_X es la valoración que difiere de v tan sólo en que $v^A_X(X) = A$.

Por ejemplo, ahora es inmediato que el modelo natural de la aritmética de Peano (es decir, el que tiene por universo $\mathbb{N}$ con la interpretación obvia de cada signo) es un modelo de los axiomas de Peano de segundo orden. Conviene destacar que 8.13 (5) implica que todos los casos particulares del esquema axiomático de inducción de la aritmética de primer orden, son verdaderos en $\mathbb{N}$, pero de aquí no se deduce 8.13 (5), pues sólo permite probarlo para subconjuntos $X \subset \mathbb{N}$ de la forma

$$X = \{n \in \mathbb{N} \mid \mathbb{N} \vDash \alpha[v^n_x]\},$$

para cierta $\alpha \in \mathrm{Form}\,\mathcal{L}$ y cierta valoración v (es decir, para subconjuntos de $\mathbb{N}$ definidos por una propiedad aritmética). En cambio, 8.13 (5) es equivalente a $\mathbb{N} \vDash \mathrm{Ind}$.

Vemos, pues, que la noción de satisfacción de una fórmula de segundo orden en un modelo involucra la noción de "la totalidad de los subconjuntos del modelo", noción que está perfectamente precisada en la teoría de conjuntos, pero de la que sería dudoso que pudiéramos hablar metamatemáticamente. Ahora es buen momento para advertir nuestra definición de lenguajes de segundo orden no es todo lo general que podría ser.

Observaciones La diferencia esencial entre la lógica de primer orden y la de segundo orden es que en las fórmulas de primer orden sólo podemos cuantificar sobre objetos, es decir, sólo podemos decir "para todo objeto" y "existe un objeto", mientras que mediante una fórmula de segundo orden podemos decir además "para toda relación monádica sobre los objetos" o "existe una relación monádica sobre los objetos".

Más en general, es posible definir lenguajes de segundo orden con un juego de variables relacionales n-ádicas para cada $n \geq 1$, y establecer que en un modelo

M han de recorrer la totalidad de las relaciones n-ádicas $R \subset M^n$. Más aún, podemos añadir variables funtoriales n-ádicas, que recorran la totalidad de las funciones $f : M^n \longrightarrow M$. Esto sólo supone algunos retoques obvios en las definiciones.

Conviene pensar que las variables (relacionales) de segundo orden son a los relatores como las variables de primer orden son a las constantes, es decir, una constante nombra un objeto fijo, mientras que una variable de primer orden varía entre los objetos. Similarmente, un relator n-ádico es lo que podríamos llamar una constante de relación n-ádica, pues denota una relación n-ádica fija, mientras que una variable relacional n-ádica varía entre las relaciones n-ádicas (todo esto respecto a un modelo fijo). Similarmente, los funtores son constantes de función.

Una forma desafortunada de distinguir la lógica de primer orden de la de segundo orden es afirmar que la lógica de primer orden cuantifica sobre objetos y la de segundo orden cuantifica, además, sobre propiedades de objetos. Aunque se parece a lo que hemos dicho en el primer párrafo de estas observaciones, no es lo mismo. No es lo mismo cuantificar sobre relaciones que cuantificar sobre propiedades, ya que mediante la lógica de primer orden podemos cuantificar sobre propiedades (= fórmulas) a través de los esquemas axiomáticos, como ilustran el esquema de inducción de primer orden (que podríamos parafrasear como "para toda propiedad, si 0 la cumple, etc.") o el esquema de formación de clases de MK ("para toda propiedad, existe la clase de todos los conjuntos que la satisfacen").

Debemos tener claro que tanto la aritmética de Peano (de primer orden) como las teorías de conjuntos ZFC o MK son teorías axiomáticas de primer orden pese a contar con esquemas axiomáticos que "cuantifican" sobre propiedades. La diferencia está en que esa cuantificación es metamatemática y no formal a través de una variable de segundo orden. Sería absurdo escribir el esquema de formación de clases como

$$\bigwedge\phi\bigvee X\bigwedge y(y \in X \leftrightarrow \phi(y)). \tag{10.1}$$

El principio "$\bigwedge\phi$" es asintáctico: ϕ es una fórmula, y detrás de un cuantificador sólo puede ir una variable. En todo caso, esto tendría sentido si decidiéramos emplear un lenguaje de segundo orden para la teoría de conjuntos e interpretáramos ϕ, no como una fórmula, sino como una variable de segundo orden. Ahora bien, esto no nos llevaría donde el lector podría estar pensando. Imaginemos que a partir de (10.1), entendido como un axioma de segundo orden, queremos demostrar la existencia de la intersección de clases $Y \cap Z$. Si lo entendiéramos como un esquema de primer orden (sin el $\bigwedge\phi$ inicial) sería muy fácil: bastaría considerar la fórmula $\phi(y) \equiv y \in Y \wedge y \in Z$, ahora bien, para deducir algo de (10.1) tal cual, hemos de apoyarnos necesariamente en la fórmula

$$\bigvee\phi\bigwedge y(\phi(y) \leftrightarrow y \in Y \wedge y \in Z),$$

pero eso es casi lo mismo que lo que queremos probar. En general, toda teoría axiomática de segundo orden requiere entre sus axiomas, si queremos que de

ellos se deduzca algo, un esquema axiomático de *comprensión:*

$$\text{(Comp)}\quad \text{Para toda fórmula } \phi(x), \qquad \bigvee X \bigwedge x(X(x) \leftrightarrow \phi(x)),$$

es decir, un axioma que afirme que toda fórmula determina una relación. Así pues, con la lógica de segundo orden complicamos el formalismo y no nos libramos de los esquemas axiomáticos.

Por otra parte, el lector no debe pensar que la lógica de segundo orden es equivalente a la de primer orden, en el sentido de que son dos formas distintas de expresar las mismas cosas. Al contrario, ambas difieren en hechos esenciales, por lo que insistir en que las teorías de conjuntos que hemos estudiado son teorías de primer orden no es una mera cuestión de lenguaje. Mostraremos esto en el apartado siguiente.

Diferencias entre la lógica de primer y segundo orden Como primera diferencia citaremos que la lógica de segundo orden es categórica en más casos que la lógica de primer orden. Se dice que un conjunto de axiomas es *categórico* si tiene un único modelo salvo (la noción obvia de) isomorfismo. Por ejemplo, en un lenguaje de primer orden sin ningún signo eventual, la sentencia $\alpha \equiv \bigwedge xy(x = y)$ es categórica, pues sólo puede tener un modelo con un elemento. Si el lenguaje tiene, digamos un relator monádico R, entonces α ya no es categórica, pues admite dos modelos esencialmente distintos, uno en el que $M(R)$ es verdadera sobre el único objeto y otro en el que es falsa. No obstante esto se arregla tomando

$$\bigwedge xy(x = y) \wedge \bigwedge x\, Rx.$$

Ahora bien, a raíz del teorema de compacidad, la aritmética de Peano de primer orden no es categórica. Más en general, cualquier colección de fórmulas verdaderas en el modelo natural de la aritmética admite también un modelo no estándar, luego no es categórica.

No ocurre lo mismo con la lógica de segundo orden. De hecho, si M es un modelo de los tres axiomas de Peano a), b) e Ind, entonces la aplicación $i : \mathbb{N} \longrightarrow M$ dada por $i(n) = M(0^{(n)})$ es biyectiva (y entonces es claramente un isomorfismo de modelos). El hecho de que M cumpla los axiomas a) y b) implica que i es inyectiva y el axioma Ind aplicado al conjunto $i[\mathbb{N}]$ prueba que es suprayectiva.

El argumento falla para la lógica de primer orden porque si M es un modelo no estándar entonces $i[\mathbb{N}]$ es el conjunto de los números naturales estándar de M, y no está definido por ninguna fórmula del lenguaje de la aritmética, por lo que no podemos aplicar ningún caso particular del esquema de inducción.

Esto puede parecer una ventaja de la lógica de segundo orden frente a la de primer orden, pero es más aparente que real. Como consecuencia, el teorema de completitud es falso para la lógica de segundo orden, es decir, existen conjuntos consistentes de axiomas de segundo orden que no tienen modelos. En efecto, una forma de verlo es basándonos en que los teoremas de incompletitud son igualmente válidos para la lógica de segundo orden, por lo que toda teoría

aritmética recursiva consistente (de segundo orden) T tiene una sentencia G que no puede demostrarse ni refutarse. Puesto que o bien G o bien $\neg G$ ha de ser falsa en $\mathbb{N}$ (el único modelo de T, si es que tiene alguno) al añadirla como axioma tenemos una teoría consistente sin modelos.

Otra forma de probarlo es mediante el argumento que nos daba la existencia de modelos no estándar a partir del teorema de compacidad (lo que, de paso, nos prueba que el teorema de compacidad es falso para la lógica de segundo orden): añadiendo una constante c al lenguaje de la aritmética y añadiendo a los tres axiomas de Peano de segundo orden los axiomas $c \neq 0$ $c \neq 0'$, $c \neq 0''$, etc. obtenemos otra teoría consistente sin modelos (es consistente porque todo subconjunto finito tiene un modelo, luego es consistente, y no tiene modelos porque si el único posible sería $\mathbb{N}$, pero entonces la constante c no podría ser interpretada).

El hecho de que el teorema de completitud falle implica a su vez que no existe un cálculo deductivo satisfactorio para la lógica de segundo orden, es decir, que puede darse el caso de que una fórmula sea consecuencia lógica de unas premisas (en el sentido de que es necesariamente verdadera en todo modelo en que lo sean las premisas) y, pese a ello, no sea deducible formalmente de las premisas, y ello independientemente de lo que nos esforcemos por afinar el cálculo deductivo (sirva como ejemplo la sentencia de Gödel de la aritmética de Peano de segundo orden).

La lógica de segundo orden tampoco cumple el teorema de Löwenheim-Skolem, es decir, que una teoría que tenga un modelo no tiene necesariamente un modelo numerable. Por ejemplo, consideremos un lenguaje de primer orden que disponga de dos funtores diádicos $+$ y $\cdot$, dos constantes 0 y 1 y un relator diádico $\leq$. Los axiomas de cuerpo ordenado pueden expresarse fácilmente mediante fórmulas de primer orden de este lenguaje. Por ejemplo,

$$\bigwedge x(x \neq 0 \to \bigvee y\, xy = 1).$$

Si ahora añadimos variables de segundo orden, podemos expresar la completitud mediante la sentencia

$$\bigwedge X(\bigvee y\, X(y) \wedge \bigvee y \bigwedge x(X(x) \to x \leq y) \to \bigvee s(\bigwedge x(X(x) \to x \leq s) \\ \wedge \bigwedge y(\bigwedge x(X(x) \to x \leq y) \to s \leq y))).$$

Tenemos así un conjunto Γ de sentencias de segundo orden (una de ellas) que tiene por modelo a $\mathbb{R}$ con las interpretaciones usuales de la suma, el producto, etc. y tal que todo modelo M de Γ tiene estructura de cuerpo ordenado completo. Es conocido que todo cuerpo ordenado completo es no numerable, por lo que Γ tiene modelos pero no tiene modelos numerables. Aguzando el ingenio podemos expresar la propiedad arquimediana mediante una fórmula de segundo orden:

$$\bigwedge x X(X(0) \wedge \bigwedge y(X(y) \to X(y+1)) \to \bigvee y(X(y) \wedge x \leq y)),$$

y sucede que todo cuerpo ordenado arquimediano y completo es isomorfo a $\mathbb{R}$, luego al añadir esta fórmula a Γ obtenemos un conjunto categórico de sentencias de segundo orden con $\mathbb{R}$ como único modelo.

Consideraciones finales No hemos entrado en la descripción de un cálculo deductivo de segundo orden porque, como hemos visto, no existe ninguno que sea semánticamente completo. De todos modos, es fácil construir uno aceptable sin más que retocar levemente el cálculo de primer orden y añadir el esquema axiomático de comprensión. La lógica de segundo orden tiene más interés desde el punto de vista de la teoría de modelos por su mayor capacidad de expresión. De todos modos hemos de insistir en que la semántica de segundo orden involucra la totalidad de los subconjuntos de un conjunto dado, por lo que es cuestionable que los resultados que acabamos de comentar tengan una interpretación metamatemática.

A efectos prácticos, la lógica de primer orden supera con creces a la de segundo orden. No sólo porque es técnicamente más simple y basta para fundamentar la teoría de conjuntos, sino porque los métodos más importantes para obtener pruebas de consistencia e independencia en teoría de conjuntos dependen esencialmente del hecho de que la lógica subyacente es de primer orden.

Una lógica de tercer orden sería una lógica con variables para representar relaciones y funciones entre relaciones y funciones de objetos, y así sucesivamente. También existe la lógica de órdenes, que no es sino la lógica con variables de todos los órdenes posibles y, más en general, está la lógica de tipos, en la que los tipos de variables no están asociados a un orden en $\mathbb{N}$, sino que pueden tener casi cualquier estructura.

10.4 El lenguaje de la teoría de conjuntos

Ahora vamos a estudiar con más detalle la formalización en teoría de conjuntos del lenguaje de la propia teoría de conjuntos. Esto puede llevar a confusiones de notación más peligrosas que el ejemplo $===$ que comentábamos antes, por lo que conviene introducir los llamados *ángulos de Quine*. Para empezar, en lugar de llamar 0, 1, 2, etc. a los números naturales, usaremos la notación $\ulcorner 0\urcorner$, $\ulcorner 1\urcorner$, $\ulcorner 2\urcorner$, ..., es decir,

$$\ulcorner 0\urcorner \equiv \varnothing, \quad \ulcorner 1\urcorner \equiv \{\ulcorner 0\urcorner\}, \quad \ulcorner 2\urcorner \equiv \{\ulcorner 0\urcorner, \ulcorner 1\urcorner\}, \quad \dots$$

En otras palabras, para cada natural n, el designador $\ulcorner n\urcorner$ es lo que en una teoría aritmética arbitraria representábamos por $0^{(n)}$. Preferimos ahora la notación de Quine porque se generaliza fácilmente a otros conceptos:

Definición 10.5 Definimos el lenguaje formal $\ulcorner\mathcal{L}\urcorner = (\ulcorner 0\urcorner, \ulcorner 1\urcorner, \ulcorner 2\urcorner, \ulcorner 3\urcorner, x, c, R, f)$, en el que $x : \mathbb{N} \longrightarrow \mathbb{N}$ es la función dada por $x_i = i + \ulcorner 6\urcorner$, $c = \varnothing$, R es la función de dominio $\{(\ulcorner 2\urcorner, \ulcorner 0\urcorner), (\ulcorner 2\urcorner, \ulcorner 1\urcorner)\}$ dada por $R^{\ulcorner 2\urcorner}_{\ulcorner 0\urcorner} = \ulcorner 4\urcorner$, $R^{\ulcorner 2\urcorner}_{\ulcorner 1\urcorner} = \ulcorner 5\urcorner$ y $f = \varnothing$.

Escribiremos $\ulcorner\neg\urcorner \equiv \ulcorner 0\urcorner$, $\ulcorner\to\urcorner = \ulcorner 1\urcorner$, $\ulcorner\bigwedge\urcorner = \ulcorner 2\urcorner$, $\ulcorner|\urcorner = \ulcorner 3\urcorner$, $\ulcorner=\urcorner = \ulcorner 4\urcorner$, $\ulcorner\in\urcorner = \ulcorner 5\urcorner$ y $\ulcorner x_i\urcorner = x_{\ulcorner i\urcorner}$.

De este modo, para cada signo ζ de $\mathcal{L}$ está definido el designador $\ulcorner\zeta\urcorner$ de $\mathcal{L}$ de modo que $\underset{\text{NBG}^-}{\vdash} \ulcorner\zeta\urcorner \in \text{Sig}\ulcorner\mathcal{L}\urcorner$.

Observaciones Con esta notación, fórmulas "desconcertantes" como $===$ resultan más claras (en este caso $\ulcorner=\urcorner = \ulcorner=\urcorner$). Conviene prestar atención a algunos detalles. Si, trabajando en teoría de conjuntos, decimos, "sea $=$ el igualador de un lenguaje $\mathfrak{L}$", hemos de entender que $=$ y $\mathfrak{L}$ son, metamatemáticamente, dos variables del lenguaje de la teoría de conjuntos. Podríamos haber dicho igualmente "sea x el igualador de un lenguaje y". Si usamos concretamente los nombres $=$ y $\mathfrak{L}$ para las variables es meramente por motivos mnemotécnicos.

Ahora, en cambio, $\ulcorner\mathfrak{L}\urcorner$ no es una variable del lenguaje $\mathfrak{L}$ de la teoría de conjuntos, sino un designador, como pueda serlo $\mathbb{N}$ o $\varnothing$ (informalmente, $\ulcorner\mathfrak{L}\urcorner$ representa a un conjunto concreto, al igual que $\mathbb{N}$ y $\varnothing$). Esto se ve más claramente en el caso de $\ulcorner=\urcorner$, que es exactamente el designador $\ulcorner 4\urcorner$ (lo que los matemáticos representan habitualmente como 4). ∎

Si ζ es una cadena de signos de $\mathfrak{L}$ formada por los signos $\zeta_0, \dots, \zeta_n$, definimos

$$\ulcorner\zeta\urcorner = \{(\ulcorner 0\urcorner, \ulcorner\zeta_0\urcorner), \dots, (\ulcorner n\urcorner, \ulcorner\zeta_n\urcorner)\}. \tag{10.2}$$

Es claro que $\underset{\mathrm{NBG}^-}{\vdash} \ulcorner\zeta\urcorner \in \mathrm{Cad}\ \ulcorner\mathfrak{L}\urcorner$.

Si S es una sucesión de cadenas de signos de $\mathfrak{L}$, formada por las cadenas $S_0, \dots, S_n$, definimos

$$\ulcorner S\urcorner = \{(\ulcorner 0\urcorner, \ulcorner S_0\urcorner), \dots, (\ulcorner n-1\urcorner, \ulcorner S_n\urcorner)\},$$

y es claro que $\underset{\mathrm{NBG}^-}{\vdash} \ulcorner S\urcorner \in (\mathrm{Cad}\ \ulcorner\mathfrak{L}\urcorner)^{<\omega}$.

Más aún, es evidente que si ζ es una cadena de signos de $\mathfrak{L}$,

- si ζ es un término entonces $\underset{\mathrm{NBG}^-}{\vdash} \ulcorner\zeta\urcorner \in \mathrm{Term}\ \ulcorner\mathfrak{L}\urcorner$,
- si ζ es una fórmula entonces $\underset{\mathrm{NBG}^-}{\vdash} \ulcorner\zeta\urcorner \in \mathrm{Form}\ \ulcorner\mathfrak{L}\urcorner$.

Entendamos bien esto: Consideremos por ejemplo la fórmula $x \in y$. Aunque a menudo hablamos de "la fórmula $x \in y$", lo cierto es que esto es ambiguo, pues no hemos especificado cuáles son concretamente las variables x e y. Normalmente esto carece de importancia, pero ahora sí la tiene. Concretemos al caso $x_0 \in x_1$. De acuerdo con nuestros convenios de notación, esto representa a la fórmula formada (en este orden) por los signos $\in$, x_0, x_1. Por consiguiente,

$$\ulcorner x_0 \in x_1\urcorner \equiv \{(\ulcorner 0\urcorner, \ulcorner\in\urcorner), (\ulcorner 1\urcorner, \ulcorner x_0\urcorner), (\ulcorner 2\urcorner, \ulcorner x_1\urcorner)\} \equiv \{(\ulcorner 0\urcorner, \ulcorner 5\urcorner), (\ulcorner 1\urcorner, \ulcorner 6\urcorner), (\ulcorner 2\urcorner, \ulcorner 7\urcorner)\}.$$

En definitiva, $\ulcorner x_0 \in x_1\urcorner$ es lo que un matemático escribiría habitualmente como $\{(0,5),(1,6),(2,7)\}$ o incluso como $(5,6,7)$. Con la notación que se prefiera, se trata de la sucesión de longitud 3 cuyos términos son 5, 6, 7. Es claro que $\ulcorner x_0 \in x_1\urcorner$ satisface la definición de fórmula. De hecho, el mismo argumento (metamatemático) que nos convence de que $x_0 \in x_1$ es una fórmula se convierte inmediatamente en una demostración en NBG^- de que $\ulcorner x_0 \in x_1\urcorner \in \mathrm{Form}\ \ulcorner\mathfrak{L}\urcorner$. Esto es totalmente general: sería fácil programar a un ordenador para que, al darle una fórmula α, nos proporcionara una demostración en NBG^- de la sentencia $\ulcorner\alpha\urcorner \in \mathrm{Form}\ \ulcorner\mathfrak{L}\urcorner$.

Observaciones Ahora el lector debe esforzarse por comprender lo siguiente: aunque podamos probar $\ulcorner x_0 \in x_1 \urcorner \in \mathrm{Form}\,\ulcorner \mathfrak{L} \urcorner$, es decir, dicho más informalmente, que $\ulcorner x_0 \in x_1 \urcorner$ es una fórmula de $\ulcorner \mathfrak{L} \urcorner$, al mismo tiempo es cierto que $\ulcorner x_0 \in x_1 \urcorner$ es un término de $\mathfrak{L}$. De hecho es un designador: podemos demostrar (es trivial) que $\ulcorner x_0 \urcorner$ es una variable libre en $\ulcorner x_0 \in x_1 \urcorner$ pero, al mismo tiempo, ni $\ulcorner x_0 \urcorner$ es una variable (es el número 6) ni está libre en $\ulcorner x_0 \in x_1 \urcorner$. Esto es claro[3] si recordamos que $\ulcorner x_0 \in x_1 \urcorner$ no es otra cosa sino $\{(0,5),(1,6),(2,7)\}$. ¿Cuántas variables libres hay aquí? ∎

Definimos

$$\ulcorner \mathrm{NBG}^- \urcorner \equiv \big\{ \ulcorner \text{NBG-1} \urcorner, \ulcorner \text{NBG-2} \urcorner, \ulcorner \text{NBG-3} \urcorner, \ulcorner \text{NBG-4} \urcorner,$$
$$\ulcorner \text{NBG-5} \urcorner, \ulcorner \text{NBG-6} \urcorner, \ulcorner \text{NBG-7} \urcorner, \ulcorner \text{NBG-8} \urcorner, \ulcorner \text{NBG-9} \urcorner,$$
$$\ulcorner \text{NBG-10} \urcorner, \ulcorner \text{NBG-11} \urcorner, \ulcorner \text{NBG-12} \urcorner, \ulcorner \text{NBG-13} \urcorner, \ulcorner \text{AI} \urcorner \big\}.$$

Más en general, si T es cualquier extensión (recursiva[4]) de NBG^- (por ejemplo todo NBG), podemos definir de igual modo el correspondiente conjunto $\ulcorner T \urcorner$.

Es claro que si S es una demostración en T de una fórmula α de $\mathfrak{L}$, entonces $\vdash_{\mathrm{NBG}^-} \ulcorner T \urcorner \vdash_{\ulcorner \mathfrak{L} \urcorner} \ulcorner S \urcorner \ulcorner \alpha \urcorner$. Simplemente, el mismo razonamiento que nos convence de que S es ciertamente una demostración de α, vale como prueba en NBG^- de que $\ulcorner S \urcorner$ es una demostración de $\ulcorner \alpha \urcorner$.

Observaciones En una teoría aritmética arbitraria necesitábamos introducir la numeración de Gödel para hablar a través de ella de la propia teoría, de sus términos, fórmulas y teoremas. Sin embargo, en el caso de la teoría de conjuntos no necesitamos la numeración de Gödel. Para hablar de una fórmula α, aunque podríamos, por supuesto, definir su número de Gödel, resulta mucho más natural e inmediato utilizar su "imagen" $\ulcorner \alpha \urcorner$. Quizá para el lector sea más difícil distinguir claramente entre α y $\ulcorner \alpha \urcorner$ que entre α y su número de Gödel, pero ello se debe precisamente a que $\ulcorner \alpha \urcorner$ es una imagen mucho más fiel de α, lo cual en teoría no es un inconveniente, sino una muestra más de la potencia de

[3] El lector que se pierda debe pensar en un ejemplo más sencillo: la pregunta ¿qué es $\varnothing$? tiene dos respuestas paralelas: por una parte es un designador formado por ocho signos, por otra es el conjunto vacío, es decir, el único conjunto que no tiene elementos. Cualquier intento de mezclar ambos puntos de vista lleva a sinsentidos: es absurdo preguntarse cuántos signos tiene el conjunto vacío o cuántos elementos tiene un designador. Más informalmente aún: ante la pregunta ¿quién es el que aparece en esta foto? dos respuestas igualmente válidas pueden ser: Harrison Ford o Indiana Jones. Sin embargo, afirmaciones verdaderas sobre Harrison Ford pueden ser falsas sobre Indiana Jones y viceversa: uno es actor, el otro arqueólogo, etc.

[4] En realidad, desde este punto de vista, no tenemos un argumento general que nos pruebe la existencia de $\ulcorner T \urcorner$ para cualquier extensión recursiva T. Lo único que afirmamos es que, dada cualquier extensión recursiva concreta T, tendremos una descripción explícita de sus axiomas, a partir de la cual será inmediata la definición de $\ulcorner T \urcorner$. Ahora no estamos interesados en la generalidad de estos conceptos, sino en su aplicabilidad a casos concretos, como NBG^-, NBG, y unas pocas variantes más que se diferencian de NBG en un número finito de axiomas, y en este caso finito sí es clara en general la existencia de $\ulcorner T \urcorner$

la teoría de conjuntos. Si el lector consigue finalmente asimilar la relación y la diferencia entre cada concepto metamatemático y su equivalente formalizado, en la sección siguiente encontrará el argumento de los teoremas de incompletitud en su forma más transparente.

10.5 Los teoremas de incompletitud

Vamos a probar los teoremas de incompletitud para una extensión recursiva T de NBG^- (o, equivalentemente, de ZF^-) basándonos en la formalización de la lógica en NBG^- como sustituto de la numeración de Gödel. Así, en lugar de la fórmula $\phi(x)$ que representa a la relación recursiva "ser una fórmula", tenemos la fórmula $x \in \mathrm{Form}\,\ulcorner\mathfrak{L}\urcorner$ y, como sustituto de la fórmula $\mathrm{Dm}(m,n)$, tenemos la fórmula $\underset{\mathrm{NBG}^-}{\vdash}\ \ulcorner T\urcorner\ \overset{S}{\underset{\ulcorner\mathfrak{L}\urcorner}{\vdash}}\ \alpha$, con variables S (en lugar de m) y α (en lugar de n).

Solamente nos falta el análogo de la fórmula que representa a la función recursiva N que cumple $N(n) = g(0^{(n)})$. Ante todo, en lugar de $0^{(n)}$ ahora escribimos $\ulcorner n\urcorner$. Esto es un mero cambio de notación. En segundo lugar, la función g podemos simplemente eliminarla, pues ahora ya no necesitamos pasar por números de Gödel para poder hablar de NBG^- en NBG^-. Por consiguiente, nos basta la función $N(n) = \ulcorner n\urcorner$. En realidad, no estamos interesados en la función N, sino en la fórmula que la representa en NBG^-. En nuestro caso, tenemos una función metamatemática N y necesitamos su análogo formal, que podemos llamar $\ulcorner N\urcorner$. Puesto que N asigna a cada número natural n el término $\ulcorner n\urcorner$ de $\mathfrak{L}$, hemos de definir $\ulcorner N\urcorner$ de modo que se cumpla $\ulcorner N\urcorner : \mathbb{N} \longrightarrow \mathrm{Term}\,\ulcorner\mathfrak{L}\urcorner$.

Para cada natural n, se cumple $N(n) = \ulcorner n\urcorner$. El análogo formal de N es $\ulcorner N\urcorner$ (la función que hemos de definir), el análogo formal de n es $\ulcorner n\urcorner$ (el numeral que representa al número metamatemático n) y —éste es el punto crucial— el análogo formal del numeral $\ulcorner n\urcorner$ es el designador $\ulcorner\ulcorner n\urcorner\urcorner$. Por consiguiente, queremos definir $\ulcorner N\urcorner$ de modo que, para cada natural n, se pueda probar que $\ulcorner N\urcorner(\ulcorner n\urcorner) = \ulcorner\ulcorner n\urcorner\urcorner$. Veamos un ejemplo concreto de los dobles ángulos:

Tenemos que $\ulcorner 0\urcorner \equiv \varnothing \equiv x \mid \bigwedge y\, y \notin x$. Por consiguiente

$$\ulcorner\ulcorner 0\urcorner\urcorner \equiv \ulcorner\varnothing\urcorner \equiv \ulcorner x \mid \textstyle\bigwedge y\, y \notin x\urcorner \equiv \{(\ulcorner 0\urcorner, \ulcorner |\urcorner), (\ulcorner 1\urcorner, \ulcorner x\urcorner), (\ulcorner 2\urcorner, \ulcorner \textstyle\bigwedge\urcorner), (\ulcorner 3\urcorner, \ulcorner y\urcorner),$$

$$(\ulcorner 4\urcorner, \ulcorner\neg\urcorner), (\ulcorner 5\urcorner, \ulcorner\in\urcorner), (\ulcorner 6\urcorner, \ulcorner y\urcorner), (\ulcorner 7\urcorner, \ulcorner x\urcorner)\}.$$

La diferencia es clara: $\ulcorner 0\urcorner$ es el conjunto vacío, no tiene elementos, mientras que $\ulcorner\ulcorner 0\urcorner\urcorner$ es la sucesión formada por los ocho signos que definen a $\ulcorner 0\urcorner$. Similarmente $\ulcorner\ulcorner 1\urcorner\urcorner$ es una sucesión de signos mucho más larga, la que define a $\ulcorner 1\urcorner \equiv \ulcorner 0\urcorner \cup \{\ulcorner 0\urcorner\}$, etc.

Para definir $\ulcorner N\urcorner$ sólo tenemos que considerar cuál es la definición metamatemática de N, expresada en términos del teorema de recursión:

$$N(0) = \ulcorner 0\urcorner, \qquad N(n+1) = \ulcorner n+1\urcorner \equiv \ulcorner n\urcorner \cup \{\ulcorner n\urcorner\} \equiv \mathbb{S}_{x_0}^{N(n)}(x_0 \cup \{x_0\}).$$

Aunque hasta aquí ha sido conveniente mantener los ángulos de Quine, dado que en ningún momento vamos a necesitar la función metamatemática N, en la definición que sigue escribimos simplemente N para representar lo que en la discusión previa hemos llamado $\ulcorner N\urcorner$.

Definición 10.6 Sea $N : \mathbb{N} \longrightarrow \text{Term}\ \ulcorner\mathcal{L}\urcorner$ la aplicación dada por

$$N(\ulcorner 0\urcorner) = \ulcorner\ulcorner 0\urcorner\urcorner \wedge \bigwedge x \in \mathbb{N}\, N(x + \ulcorner 1\urcorner) = \mathtt{S}^{N(x)\ulcorner x_0\urcorner}_{\ulcorner x_0\urcorner}\, x_0 \cup \{x_0\}\urcorner. \tag{10.3}$$

Veamos que para todo natural (metamatemático) n se cumple

$$\underset{\text{NBG}^-}{\vdash} N(\ulcorner n\urcorner) = \ulcorner\ulcorner n\urcorner\urcorner.$$

Para ello nos basaremos en el hecho siguiente, que se prueba fácilmente por inducción (metamatemática): Si θ es una expresión, t es un término y x es una variable, entonces

$$\underset{\text{NBG}^-}{\vdash} \ulcorner \mathtt{S}^t_x\theta\urcorner = \mathtt{S}^{\ulcorner t\urcorner}_{\ulcorner x\urcorner}\ulcorner\theta\urcorner.$$

Observemos que la $\mathtt{S}$ del miembro izquierdo es la sustitución metamatemática, mientras que la $\mathtt{S}$ de la derecha es su análogo formal (quizá deberíamos escribir $\ulcorner\mathtt{S}\urcorner$). La demostración se basa en que la definición formal de $\mathtt{S}$ es completamente paralela a la definición metamatemática.

Teniendo esto en cuenta, (10.3) se prueba por inducción (metamatemática) sobre n. En efecto, para $n = 0$ tenemos $\underset{\text{NBG}^-}{\vdash} N(\ulcorner 0\urcorner) = \ulcorner\ulcorner 0\urcorner\urcorner$ por definición de N. Supuesto cierto para n, es decir, si tenemos probado

$$N(\ulcorner n\urcorner) = \ulcorner\ulcorner n\urcorner\urcorner,$$

razonamos como sigue:

$$\begin{aligned} N(\ulcorner n+1\urcorner) &= N(\ulcorner n\urcorner + \ulcorner 1\urcorner) = \mathtt{S}^{\ulcorner\ulcorner n\urcorner\urcorner\ulcorner x_0\urcorner}_{\ulcorner x_0\urcorner}\, x_0 \cup \{x_0\}\urcorner = \ulcorner\mathtt{S}^{\ulcorner n\urcorner}_{x_0}(x_0 \cup \{x_0\})\urcorner \\ &= \ulcorner\ulcorner n\urcorner \cup \{\ulcorner n\urcorner\}\urcorner = \ulcorner\ulcorner n+1\urcorner\urcorner. \end{aligned}$$

En particular, si ζ es un signo de $\mathcal{L}$ tenemos

$$\underset{\text{NBG}^-}{\vdash} N\big(\ulcorner\zeta\urcorner\big) = \ulcorner\ulcorner\zeta\urcorner\urcorner.$$

Por desgracia, la función N no es la que necesitamos para probar los teoremas de incompletitud, sino la función que cumple la propiedad análoga a la que acabamos de enunciar pero donde ζ es una cadena de signos de $\mathcal{L}$, en lugar de un signo. Más concretamente, queremos formalizar la función metamatemática que a cada cadena de signos $\zeta \equiv \zeta_0 \dots \zeta_n$ le asigna el término

$$\ulcorner\zeta\urcorner \equiv \{(\ulcorner 0\urcorner, \ulcorner\zeta_0\urcorner), \dots, (\ulcorner n\urcorner, \ulcorner\zeta_n\urcorner)\}$$

$$\equiv x_1 \mid \bigwedge x_0(x_0 \in x_1 \leftrightarrow x_0 = (\ulcorner 0\urcorner, \ulcorner \zeta_0\urcorner) \vee \cdots \vee x_0 = (\ulcorner n\urcorner, \ulcorner \zeta_n\urcorner)).$$

Lo único que hemos de hacer es formalizar sistemáticamente la definición del último miembro. Nos ocupamos primero de su parte derecha, formada por una concatenación de disyunciones. Empezamos definiendo explícitamente la función disyunción

$$\mathrm{Dis} : \mathrm{Form}\,\ulcorner\mathfrak{L}\urcorner \times \mathrm{Form}\,\ulcorner\mathfrak{L}\urcorner \longrightarrow \mathrm{Form}\,\ulcorner\mathfrak{L}\urcorner$$

dada por $\mathrm{Dis}(\alpha,\beta) = \ulcorner{\to}\urcorner\ulcorner\neg\urcorner\alpha\beta$, o sea, $\mathrm{Dis}(\alpha,\beta) = \alpha \vee \beta$. Extendámosla de modo que $\mathrm{Dis}(\varnothing,\alpha) = \alpha$.

Concretamente, queremos concatenar una disyunción de fórmulas de tipo

$$x_0 = (\ulcorner n\urcorner, \ulcorner \zeta_n\urcorner) \equiv \mathfrak{S}^{\ulcorner n\urcorner}_{x_1}\mathfrak{S}^{\ulcorner \zeta_n\urcorner}_{x_2}(x_0 = (x_1, x_2)).$$

Teniendo en cuenta que la formalización de $\ulcorner n\urcorner$ y $\ulcorner \zeta_n\urcorner$ es $N(n)$ y $N(\zeta_n)$, la concatenación que queremos es la función $F : \big(\mathrm{Sig}\,\ulcorner\mathfrak{L}\urcorner\big)^{<\omega} \longrightarrow \mathrm{Form}\big(\ulcorner\mathfrak{L}\urcorner\big) \cup \{\varnothing\}$ definida como sigue por recurrencia sobre la longitud de una cadena. Para la cadena de longitud 0 definimos $F(\varnothing) = \varnothing$ y, supuesto que F está definida para cadenas de longitud n, si ζ tiene longitud $n+1$ hacemos

$$F(\zeta) = \mathrm{Dis}\big(F(\zeta|_n), \mathfrak{S}^{N(n)}_{\ulcorner x_1\urcorner}\mathfrak{S}^{N(\zeta_n)\ulcorner}_{x_2\urcorner}x_0 = (x_1, x_2)^{\urcorner}\big)$$

De este modo, si $\zeta \equiv \zeta_0 \cdots \zeta_n$ es una cadena de signos de $\mathfrak{L}$, entonces

$$\underset{\mathrm{NBG}^-}{\vdash} F\big(\ulcorner\zeta\urcorner\big) = {}^{\ulcorner}x_0 = \big(\ulcorner 0\urcorner, \ulcorner \zeta_0\urcorner\big) \vee \cdots \vee x_0 = \big(\ulcorner n\urcorner, \ulcorner \zeta_n\urcorner\big)^{\urcorner}.$$

En efecto, lo probamos por inducción sobre n. Si $n = 0$ tenemos

$$\begin{aligned} F\big(\ulcorner\zeta_0\urcorner\big) &= \mathrm{Dis}(F(\varnothing), \mathfrak{S}^{\ulcorner\ulcorner 0\urcorner\urcorner}_{\ulcorner x_1\urcorner}\mathfrak{S}^{\ulcorner\ulcorner\zeta_0\urcorner\urcorner}_{\ulcorner x_2\urcorner}{}^{\ulcorner}x_0 = (x_1, x_2)^{\urcorner}\big) \\ &= {}^{\ulcorner}\mathfrak{S}^{\ulcorner 0\urcorner}_{x_1}\mathfrak{S}^{\ulcorner\zeta_0\urcorner}_{x_2}x_0 = (x_1, x_2)^{\urcorner} = {}^{\ulcorner}x_0 = \big(\ulcorner 0\urcorner, \ulcorner\zeta_0\urcorner\big)^{\urcorner}. \end{aligned}$$

Si es cierto para cadenas de longitud n y ζ tiene longitud $n+1$ entonces

$$\begin{aligned} F(\ulcorner\zeta\urcorner) &= \mathrm{Dis}\big(F(\ulcorner\zeta\urcorner|_{\ulcorner n\urcorner}), \mathfrak{S}^{\ulcorner\ulcorner n\urcorner\urcorner}_{\ulcorner x_1\urcorner}\mathfrak{S}^{\ulcorner\ulcorner\zeta_n\urcorner\urcorner}_{\ulcorner x_2\urcorner}{}^{\ulcorner}x_0 = (x_1, x_2)^{\urcorner}\big) \\ &= F(\ulcorner\zeta\urcorner|_{\ulcorner n\urcorner}) \vee {}^{\ulcorner}x_0 = \big(\ulcorner n\urcorner, \ulcorner\zeta_n\urcorner\big)^{\urcorner} \\ &= {}^{\ulcorner}x_0 = \big(\ulcorner 0\urcorner, \ulcorner\zeta_0\urcorner\big) \vee \cdots \vee x_0 = \big(\ulcorner n-1\urcorner, \ulcorner\zeta_{n-1}\urcorner\big)^{\urcorner} \vee {}^{\ulcorner}x_0 = \big(\ulcorner n\urcorner, \ulcorner\zeta_n\urcorner\big)^{\urcorner} \\ &= {}^{\ulcorner}x_0 = \big(\ulcorner 0\urcorner, \ulcorner\zeta_0\urcorner\big) \vee \cdots \vee x_0 = \big(\ulcorner n\urcorner, \ulcorner\zeta_n\urcorner\big)^{\urcorner}. \end{aligned}$$

Finalmente definimos la aplicación $[\ ] : \big(\mathrm{Sig}\,\ulcorner\mathfrak{L}\urcorner\big)^{<\omega} \longrightarrow \mathrm{Term}\,\ulcorner\mathfrak{L}\urcorner$ dada por $[\zeta] = x_1|\bigwedge x_0(x_0 \in x_1 \leftrightarrow F(\zeta))$.

De este modo, si ζ es una cadena de signos de $\mathfrak{L}$ de longitud $n+1$ se cumple

$$\underset{\mathrm{NBG}^-}{\vdash} \big[\ulcorner\zeta\urcorner\big] = {}^{\ulcorner}x_1|\bigwedge x_0(x_0 \in x_1 \leftrightarrow x_0 = \big(\ulcorner 0\urcorner, \ulcorner\zeta_0\urcorner\big) \vee \cdots \vee x_0 = \big(\ulcorner n\urcorner, \ulcorner\zeta_n\urcorner\big))^{\urcorner},$$

o, lo que es lo mismo,

$$\underset{\mathrm{NBG}^-}{\vdash} \left[\ulcorner\zeta\urcorner\right] = \ulcorner\{\left(\ulcorner 0\urcorner, \ulcorner\zeta_0\urcorner\right), \ldots, \left(\ulcorner n\urcorner, \ulcorner\zeta_n\urcorner\right)\}\urcorner,$$

pero según (10.2) esto es

$$\underset{\mathrm{NBG}^-}{\vdash} \left[\ulcorner\zeta\urcorner\right] = \ulcorner\ulcorner\zeta\urcorner\urcorner. \tag{10.4}$$

Observaciones Si el lector considera que esto es farragoso, debería pararse a pensar que lo es mucho más la prueba de que toda función recursiva es representable en toda teoría aritmética (donde hemos tenido que valernos de trucos como la función beta de Gödel y de comprobaciones mucho más laboriosas que las que hemos visto aquí). Si el lector asimila adecuadamente lo que hemos hecho, se dará cuenta de que, una vez fijado el objetivo de encontrar una función que cumpla (10.4), todos los pasos a seguir son mecánicos y no requieren ninguna idea. Basta escribir explícitamente las definiciones metamatemáticas involucradas para "copiarlas" formalmente. ∎

Ahora podemos probar el análogo al teorema 7.2:

Teorema 10.7 *Sea $\phi(x)$ una fórmula de $\mathfrak{L}$ cuya única variable libre sea x. Entonces existe una sentencia ψ de $\mathfrak{L}$ tal que*

$$\underset{\mathrm{NBG}^-}{\vdash} \psi \leftrightarrow \phi\left(\ulcorner\psi\urcorner\right).$$

Demostración: Consideramos la fórmula $\sigma(x_0) = \phi(\mathsf{S}^{[x_0]}_{\ulcorner x_0\urcorner} x_0)$. Sea

$$\psi \equiv \sigma(\ulcorner\sigma\urcorner) \equiv \mathsf{S}^{\ulcorner\sigma\urcorner}_{x_0}\sigma \equiv \phi(\mathsf{S}^{[\ulcorner\sigma\urcorner]}_{\ulcorner x_0\urcorner}\ulcorner\sigma\urcorner).$$

Usando (10.4) obtenemos $\underset{\mathrm{NBG}^-}{\vdash} \psi \leftrightarrow \phi(\mathsf{S}^{\ulcorner\ulcorner\sigma\urcorner\urcorner}_{\ulcorner x_0\urcorner}\ulcorner\sigma\urcorner)$, pero

$$\phi(\mathsf{S}^{\ulcorner\ulcorner\sigma\urcorner\urcorner}_{\ulcorner x_0\urcorner}\ulcorner\sigma\urcorner) \equiv \phi\left(\ulcorner \mathsf{S}^{\ulcorner\sigma\urcorner}_{x_0}\sigma\urcorner\right) \equiv \phi\left(\ulcorner\psi\urcorner\right).$$

∎

Sea T una extensión recursiva de NBG^-. Para probar el teorema de incompletitud en T podemos aplicar el teorema anterior para obtener una sentencia G de $\mathfrak{L}$ tal que

$$\underset{T}{\vdash} G \leftrightarrow \neg\ulcorner T\urcorner \underset{\ulcorner\mathfrak{L}\urcorner}{\vdash} \ulcorner G\urcorner.$$

Supuesto que T sea consistente podemos asegurar que no $\underset{T}{\vdash} G$, pues si S fuera una demostración de G en T se tendría

$$\underset{\mathrm{NBG}^-}{\vdash} \ulcorner T\urcorner \underset{\ulcorner\mathfrak{L}\urcorner}{\overset{\ulcorner S\urcorner}{\vdash}} \ulcorner G\urcorner,$$

luego lo mismo se demostraría en T (por ser una extensión de NBG^-) y concluiríamos $\underset{T}{\vdash} \neg G$.

Así pues, hemos probado que si T es consistente entonces no $\underset{T}{\vdash} G$. El recíproco es trivial, luego tenemos la doble implicación:

$$T \text{ es consistente syss no } \underset{T}{\vdash} G.$$

Hemos llegado a esta coimplicación mediante un argumento metamatemático. Ahora bien, nada nos impide considerar a todo el argumento que nos ha llevado hasta aquí como un argumento en NBG^-. Para ello no hay que añadir nada a la demostración: basta pensar que cuando hablábamos de la fórmula (metamatemática) G nos referíamos en realidad a la fórmula $\ulcorner G\urcorner$, mientras que cuando hablábamos de $\ulcorner G\urcorner$ nos referíamos a $\ulcorner\ulcorner G\urcorner\urcorner$, etc. Todos los conceptos metamatemáticos que hemos usado los tenemos debidamente formalizados. De este modo, podemos interpretar la equivalencia anterior como un teorema de NBG^-. Concretamente,

$$\underset{\mathrm{NBG}^-}{\vdash} \mathrm{Consis}\,\ulcorner T\urcorner \leftrightarrow \neg\ulcorner T\urcorner \underset{\ulcorner\mathfrak{L}\urcorner}{\vdash} \ulcorner G\urcorner,$$

donde, en general, $\mathrm{Consis}\,\Gamma$ tiene la definición obvia:

$$\mathrm{Consis}\,\Gamma \equiv \Gamma \subset \mathrm{Form}\,\ulcorner\mathfrak{L}\urcorner \wedge \neg\bigvee\alpha \in \mathrm{Form}\,\ulcorner\mathfrak{L}\urcorner\,\big(\Gamma \underset{\ulcorner\mathfrak{L}\urcorner}{\vdash} \alpha \wedge \Gamma \underset{\ulcorner\mathfrak{L}\urcorner}{\vdash} \neg\alpha\big).$$

Ahora bien, por construcción de G, lo que tenemos es

$$\underset{\mathrm{NBG}^-}{\vdash} \mathrm{Consis}\,\ulcorner T\urcorner \leftrightarrow G$$

y, como sabemos que G no es demostrable en T si éste es consistente, tenemos el segundo teorema de incompletitud:

$$\underset{T}{\vdash} \mathrm{Consis}\,\ulcorner T\urcorner \qquad \text{syss} \qquad T \text{ es contradictorio.}$$

Observemos que un modelo de $\ulcorner\mathfrak{L}\urcorner$ puede identificarse con un par (M, R), donde M es un conjunto no vacío y $R \subset M \times M$ es la relación que interpreta al relator $\ulcorner\in\urcorner$. En la práctica suele escribirse M en lugar de (M, R). Por el teorema de completitud (considerado como teorema de NBG^-)

$$\underset{\mathrm{NBG}^-}{\vdash} \bigwedge\Gamma(\Gamma \subset \mathrm{Form}\,\mathfrak{L} \to (\mathrm{Consis}\,\Gamma \leftrightarrow \bigvee M\, M \vDash \Gamma)).$$

En particular, si T es una extensión recursiva consistente de NBG^-, no es posible demostrar en T la existencia de un modelo M tal que $M \vDash \ulcorner T\urcorner$.

Por otra parte, todos los teoremas del capítulo anterior pueden considerarse como teoremas de NBG^-. En particular,

$$\underset{\mathrm{NBG}^-}{\vdash} \mathrm{Consis}\,\ulcorner \mathrm{ZFC} - \mathrm{AI}\urcorner, \qquad \underset{\mathrm{NBG}^-}{\vdash} \mathrm{Consis}\,\ulcorner \mathrm{NBG} - \mathrm{AI}\urcorner,$$

$$\underset{\mathrm{NBG}^-}{\vdash} \mathrm{Consis}\,\ulcorner\mathrm{ZF}^-\urcorner \leftrightarrow \mathrm{Consis}\,\ulcorner\mathrm{NBG}^-\urcorner$$

Más en general, la última equivalencia es cierta para cualquier extensión recursiva de ZF^- y la extensión correspondiente de NBG^- (la que resulta de añadir como axiomas las relativizaciones de los axiomas añadidos a ZF^-).

Por consiguiente, la existencia de un modelo de (una extensión consistente y recursiva de) ZF^- no puede probarse en (la extensión correspondiente de) NBG^-. Ésta es la consecuencia más importante de los teoremas de incompletitud en teoría de conjuntos.

Finalmente, notemos que ahora ya debería ser obvia la imposibilidad de demostrar la consistencia de la teoría de conjuntos. Cualquier argumento que pudiera convencernos de que ZFC es consistente podría ser considerado, tal cual, sin cambio alguno, como una demostración en ZFC de Consis ZFC, lo cual nos llevaría a que ZFC es contradictorio.

10.6 Modelos que son clases propias

En esta sección estudiaremos la relación entre las teorías NBG y MK. Concretamente vamos a probar que

$$\underset{\mathrm{MK}}{\vdash} \mathrm{Consis}\, NBG,$$

con lo que, si NBG es consistente, estamos ante un teorema de MK que no es un teorema de NBG. En realidad lo que probaremos en MK será Consis ZFC que, según acabamos de comentar, es equivalente a Consis NBG. Para ello, lo que haremos será probar que la clase universal V es un modelo de ZFC. Ciertamente, esto ya lo vimos en el capítulo anterior, el problema es que esto no puede probarse en NBG debido a que no podemos definir el concepto de satisfacción de sentencias sobre un modelo que sea una clase propia. Así pues, la dificultad de probar que V es un modelo no es probar que satisface todos los axiomas de ZFC, sino definir la noción de satisfacción.

Podemos trabajar más en general en la teoría MK^-, es decir, MK^* más el axioma de infinitud. Por abreviar llamaremos $E_k = \mathrm{Term}_k\ulcorner\mathcal{L}\urcorner \cup \mathrm{Form}_k\ulcorner\mathcal{L}\urcorner$ (recordemos las definiciones de la página 267). Llamaremos $W = \{v \mid v : \mathrm{Var}\,\ulcorner\mathcal{L}\urcorner \longrightarrow \mathrm{V}\}$ (donde V es la clase universal).

Definición 10.8 Una *interpretación* $I : E_k \times W \longrightarrow \mathrm{V}$ de nivel k es una aplicación que cumpla las propiedades siguientes:

a) Si $x \in \mathrm{Var}\,\ulcorner\mathcal{L}\urcorner$ y $v \in W$, entonces $I(x, v) = v(x)$.

b) Si $(t_1 = t_2) \in E_k$ y $v \in W$ entonces

$$I(t_1 = t_2, v) = \begin{cases} 1 & \text{si } I(t_1, v) = I(t_2, v), \\ 0 & \text{si } I(t_1, v) \neq I(t_2, v). \end{cases}$$

c) Si $(t_1 \in t_2) \in E_k$ y $v \in W$ entonces

$$I(t_1 \in t_2, v) = \begin{cases} 1 & \text{si } I(t_1, v) \in I(t_2, v), \\ 0 & \text{si } I(t_1, v) \notin I(t_2, v). \end{cases}$$

d) Si $\neg\alpha \in E_k$ y $v \in W$, entonces

$$I(\neg\alpha, v) = \begin{cases} 1 & \text{si } I(\alpha, v) = 0, \\ 0 & \text{en caso contrario.} \end{cases}$$

e) Si $\alpha \to \beta \in E_k$ y $v \in W$, entonces

$$I(\alpha \to \beta, v) = \begin{cases} 1 & \text{si } I(\alpha, v) = 0 \text{ o } I(\beta, v) = 1, \\ 0 & \text{en caso contrario.} \end{cases}$$

f) Si $(\bigwedge x\alpha) \in E_k$ y $v \in W$, entonces

$$I(\bigwedge x\alpha, v) = \begin{cases} 1 & \text{si } \bigwedge a\, I(\alpha, v_x^a) = 1, \\ 0 & \text{en caso contrario.} \end{cases}$$

g) Si $(x|\alpha) \in E_k$ y $v \in W$, entonces

$$I(x|\alpha, v) = \begin{cases} \text{el único } a \text{ tal que } I(\alpha, v_x^a) = 1 & \text{si existe tal } a, \\ \varnothing & \text{en otro caso.} \end{cases}$$

Una *interpretación* es una aplicación $I : \mathrm{Exp}^{\ulcorner}\mathfrak{L}^{\urcorner} \times W \longrightarrow \mathrm{V}$ que cumpla estas mismas condiciones.

Teorema 10.9 *Si $k \in \mathbb{N}$ e I, J son dos interpretaciones (de nivel k), entonces $I = J$.*

Demostración: Hay que demostrar que para toda expresión θ en E_k (o en $\mathrm{Exp}^{\ulcorner}\mathfrak{L}^{\urcorner}$) se cumple que $\bigwedge v \in W\, I(\theta, v) = J(\theta, v)$, lo cual se prueba fácilmente por inducción sobre θ. ∎

Teorema 10.10 *Para cada $k \in \mathbb{N}$ existe una interpretación de nivel k.*

Demostración: Es una simple inducción sobre k, pero hemos de observar que dicha inducción no puede realizarse en NBG^-. En efecto, con detalle, la inducción consiste en considerar el conjunto

$$Z = \{k \in \mathbb{N} \mid \bigvee I(I \text{ es una interpretación de nivel } k)\}, \tag{10.5}$$

y probar que $Z = \mathbb{N}$ mediante el principio de inducción, pero la fórmula que define Z no es normal, ya que las interpretaciones son clases propias y, por consiguiente, el cuantificador existencial no puede restringirse a conjuntos. Así pues, el conjunto Z no es definible (o al menos no está justificado que lo sea) en NBG^-.

Admitiendo la existencia de Z (que es inmediata en MK^-) no ofrece ninguna dificultad probar que $0 \in Z$ y, supuesto que $k \in Z$, existe una interpretación I de nivel k, que es fácil extender a una interpretación de nivel $k+1$ siguiendo la propia definición de interpretación. ∎

Los dos teoremas anteriores nos permiten definir

$$I_k \equiv I \mid I \text{ es una interpretación de nivel } k$$

y se cumple que

$$\bigwedge k \in \mathbb{N}\,(I_k : E_k \times W \longrightarrow V \text{ es una interpretación de nivel } k).$$

La unicidad justifica también que si $k < n$ entonces $I_k = I_n|_{E_k \times W}$, con lo que, si definimos

$$I = \bigcup_{k\in\mathbb{N}} I_k,$$

es claro que $I : \mathrm{Exp}\,{}^{\ulcorner}\mathfrak{L}^{\urcorner} \times W \longrightarrow \mathrm{V}$ es una interpretación (la única, de hecho). Notar que I es un designador de $\mathfrak{L}$. Definimos

$$\mathrm{V} \vDash \alpha[v] \equiv I(\alpha, v) = 1, \qquad \mathrm{V} \vDash \alpha \equiv \bigwedge v \in W\, \mathrm{V} \vDash \alpha[v].$$

Más en general, definimos

$$\mathrm{V} \vDash \Gamma \equiv \Gamma \subset \mathrm{Form}\,{}^{\ulcorner}\mathfrak{L}^{\urcorner} \wedge \bigwedge \alpha \in \Gamma\, \mathrm{V} \vDash \alpha.$$

La prueba del teorema 3.7, según el cual todos los axiomas lógicos son verdaderos en cualquier modelo, es válida en este contexto, con lo que tenemos que $\underset{\mathrm{MK}^-}{\vdash} \mathrm{V} \vDash \mathrm{Axl}$.

Por otra parte, en el capítulo anterior vimos que si en un modelo de NBG^- (en particular en un modelo de MK^-) nos quedamos únicamente con las clases que son conjuntos, obtenemos un modelo de ZF^-. El mismo argumento nos permite probar ahora

$$\underset{\mathrm{MK}^-}{\vdash} \mathrm{V} \vDash {}^{\ulcorner}\mathrm{ZF}^{-\urcorner}$$

También disponemos de los argumentos del capítulo III, en virtud de los cuales las consecuencias de las fórmulas verdaderas en un modelo son verdaderas en el modelo. Esto se traduce en que

$$\underset{\mathrm{MK}^-}{\vdash} \bigwedge \alpha \in \mathrm{Form}\,{}^{\ulcorner}\mathfrak{L}^{\urcorner} \left({}^{\ulcorner}\mathrm{ZF}^{-\urcorner} \underset{{}^{\ulcorner}\mathfrak{L}^{\urcorner}}{\vdash} \alpha \to \mathrm{V} \vDash \alpha \right).$$

Claramente entonces,

$$\underset{\mathrm{MK}^-}{\vdash} \mathrm{Consis}\,{}^{\ulcorner}\mathrm{ZF}^{-\urcorner},$$

pues, por reducción al absurdo, si suponemos $\neg\,\mathrm{Consis}\,{}^{\ulcorner}\mathrm{ZF}^{-\urcorner}$, llegamos a que existe $\alpha \in \mathrm{Form}\,{}^{\ulcorner}\mathfrak{L}^{\urcorner}$ tal que $\mathrm{V} \vDash \alpha$ y $\mathrm{V} \vDash \neg\alpha$, es decir, $\mathrm{V} \vDash \alpha$ y $\neg\mathrm{V} \vDash \alpha$, lo cual es absurdo.

Según comentábamos al principio de la sección, de aquí llegamos a

$$\underset{\mathrm{MK}^-}{\vdash} \mathrm{Consis}\,\ulcorner\mathrm{NBG}^-\urcorner.$$

Observemos que el único paso de la demostración de este hecho que no está justificado en NBG^- es la construcción del conjunto (10.5), luego ahora podemos asegurar que si NBG^- es consistente entonces no puede probarse la existencia de (10.5), y tenemos así un ejemplo concreto de un axioma de MK^- que no es un teorema de NBG^-.

Claramente todo esto vale igual si sustituimos ZF^- por una extensión recursiva T, NBG^- por la extensión correspondiente T' y MK^- por la extensión de T' que resulta de añadir el esquema axiomático de formación de clases de MK.

Observaciones Notemos que $\mathrm{V} \vDash \alpha$ es una fórmula que tiene a α como única variable libre. Una simple inducción metamatemática prueba que, para toda fórmula $\alpha(u_1, \ldots, u_n)$, se cumple

$$\underset{\mathrm{MK}^-}{\vdash} \big(\mathrm{V} \vDash \ulcorner\alpha\urcorner[v] \leftrightarrow \alpha^{\mathrm{V}}(v(\ulcorner u_1\urcorner), \ldots, v(\ulcorner u_n\urcorner))\big),$$

donde α^{V} es la relativización de α que definimos en el capítulo anterior. (Hay que probar simultáneamente el hecho análogo para términos.) De aquí se sigue en particular que, para toda sentencia α,

$$\underset{\mathrm{MK}^-}{\vdash} \big(\mathrm{V} \vDash \ulcorner\alpha\urcorner \leftrightarrow \alpha^{\mathrm{V}}\big).$$

Esto es "casi" lo que el teorema de Tarski afirma que no puede ocurrir (en una teoría consistente). Si fuera

$$\underset{\mathrm{ZF}^-}{\vdash} \big(\mathrm{V} \vDash \ulcorner\alpha\urcorner \leftrightarrow \alpha\big).$$

tendríamos que ZF^- sería contradictorio, pues habríamos definido la verdad de una sentencia arbitraria, es decir, tendríamos una fórmula $\mathrm{V} \vDash \alpha$ con α como única variable libre, que nos permitiría recuperar (el significado de) α a partir de su formalización $\ulcorner\alpha\urcorner$. Concretamente, el teorema 10.7 nos permitiría construir una sentencia T tal que

$$\underset{\mathrm{ZF}^-}{\vdash} \big(T \leftrightarrow \neg\mathrm{V} \vDash \ulcorner T\urcorner\big),$$

de donde se sigue la contradicción $T \leftrightarrow \neg T$.

En definitiva, lo que afirma el teorema de Tarski para ZF^- y sus extensiones es que, aunque podemos definir, para toda sentencia, la noción de "ser verdadera en un modelo" (que sea un conjunto), no podemos hacer lo mismo cuando el modelo es la clase universal, pues definir "ser verdadera en la clase universal" sería tanto como definir "ser verdadera", y ello nos permitiría construir una sentencia que dijera "yo soy falsa". En la práctica, lo que nos impide definir la verdad en la clase universal es que "oficialmente" la clase universal no existe en ZFC. Puesto que NBG es, en cierto sentido, equivalente a ZFC, tampoco podemos

definir en esta teoría la verdad la clase universal, y ahora el impedimento es que necesitamos definir un conjunto de números naturales mediante una fórmula en la que aparece una cuantificación sobre clases propias. Nada nos impide tomar esto como axioma (y entonces estamos en MK), pero con ello hemos definido en MK la verdad en ZFC, lo cual no nos permite llegar a ninguna contradicción (que sepamos).

Los resultados del capítulo anterior sobre la equivalencia entre NBG y ZFC y lo que acabamos de obtener explican completamente el sentido de la restricción en NBG sobre que las cuantificaciones sobre clases arbitrarias no definan clases. Esta restricción resulta ser la condición necesaria y suficiente para que las clases propias en NBG sean "eliminables", es decir, para que cualquier afirmación sobre conjuntos demostrada con el apoyo de clases propias (en NBG) sea también demostrable en ZFC y, por consiguiente, sin necesidad de clases propias. En el capítulo anterior vimos que esto es cierto y ahora hemos visto que deja de serlo si admitimos que cualquier fórmula defina una clase. En tal caso tenemos a nuestra disposición argumentos que nos permiten probar cosas sobre conjuntos (por ejemplo, Consis ZFC) en los que el uso de clases propias resulta esencial. Las clases propias de MK ya no son meros "auxiliares" introducidos por comodidad, sino nuevos conceptos con nuevas implicaciones.

Volvamos ahora sobre el conjunto (10.5) cuya existencia no puede probarse en NBG. Es fácil probar (por inducción metamatemática), que para cada número natural (metamatemático) k existe una interpretación de nivel k (de nivel $\ulcorner k\urcorner$, para ser más claros). Esto significa que, en cualquier modelo de NBG, todos los números naturales estándar, es decir, todos los que dejan bajo sí un número finito de números naturales, cumplen la propiedad que define a (10.5). Si el modelo no contiene números no estándar, entonces (10.5) existe en él y es el conjunto de todos los números naturales, pero también puede ocurrir que el modelo contenga números no estándar y que para algunos de ellos no existan interpretaciones del nivel correspondiente. Por ejemplo podría suceder que los únicos números para los que existieran interpretaciones fueran los estándar y, efectivamente, en un modelo no estándar no puede existir ningún conjunto que contenga exactamente a los números estándar. Por consiguiente (10.5) no existiría en tal modelo.

Como última observación, señalamos que si ZFC es consistente, entonces también lo es ZFC$+\neg$ Consis ZFC, y un modelo de esta teoría es un modelo de ZFC que, según sabemos, puede extenderse a un modelo de NBG, pero no a un modelo de MK, pues en todo modelo de MK ha de ser verdadero Consis ZFC.

Tercera parte

La teoría de conjuntos

Introducción a la teoría de conjuntos

En la segunda parte hemos sentado las bases lógicas de la teoría de conjuntos, es decir, hemos precisado cómo pueden entenderse las afirmaciones que hacen los matemáticos: como teoremas de una determinada teoría axiomática. Todos estos términos han sido cuidadosamente definidos. Sin embargo, no hemos demostrado nada que un matemático no tenga por evidente en su trabajo cotidiano: existe la unión, la intersección, los números naturales, etc. Si hasta ahora hemos explorado lógicamente la teoría de conjuntos, esta tercera parte está dedicada a explorarla matemáticamente.

La mayoría de los resultados que presentamos constituyen una exposición sistemática de los descubrimientos de Cantor a finales del siglo XIX. En efecto, lo que hoy se conoce propiamente como "teoría de conjuntos" es una vasta rama de la matemática cuyos fundamentos fueron establecidos por Cantor prácticamente en solitario. Podemos decir que "todo empezó" cuando, a instancias de Heine, Cantor abordó el problema de la unicidad de los desarrollos en series trigonométricas de funciones arbitrarias. Pronto obtuvo un resultado válido para series convergentes sobre todo el intervalo $[0, 2\pi]$, si bien Cantor observó que si la convergencia fallaba en algunos puntos excepcionales la unicidad de la serie seguía siendo válida. Para precisar qué excepciones eran admisibles, introdujo la noción de *conjunto derivado* de un conjunto $P \subset [0, 2\pi]$, que no es sino el conjunto P' de todos los puntos de acumulación de P. Más en general, es posible calcular derivados sucesivos P, P', P'', P''', ... o, mejor, $P^{(1)}$, $P^{(2)}$, $P^{(3)}$, ... Cantor definió un conjunto de *primera especie* como un conjunto P tal que, para algún n, se cumple $P^{(n)} = \varnothing$. A los conjuntos que no eran de primera especie los llamó de *segunda especie.* En estos términos, cantor probó el teorema siguiente:

> *Si se tiene la igualdad* $0 = d_0 + \sum_{n=1}^{\infty} c_n \operatorname{sen} nx + d_n \cos nx$, *para todo* $x \in [0, 2\pi]$ *salvo a lo sumo en un conjunto de puntos de primera especie, entonces todos los coeficientes* c_n, d_n *son nulos.*

Esto le llevó a tratar de comprender cuál era la diferencia entre los conjuntos de primera y segunda especie. Un hecho relevante es que todo conjunto de primera especie es numerable, pero esta noción era completamente desconocida a

la sazón. No obstante, Cantor la intuyó, y se preguntó si sería posible establecer una correspondencia biunívoca entre los números naturales y los números reales.

El 23 de noviembre de 1873 formuló la pregunta en una carta a su amigo Richard Dedekind, el cual le contestó que era incapaz de encontrar una razón por la que no pudiera existir tal correspondencia, pero antes de que acabara el año Cantor ya había probado que no podía existir. A principios de 1874, en una nueva carta, Cantor preguntaba si sería posible biyectar los puntos de una superficie, por ejemplo un cuadrado, con los de un segmento de recta. La respuesta parecía ser obviamente negativa, y muchos de los matemáticos a los que les planteó la cuestión la tomaron por ridícula. Sin embargo, tres años después, en 1877, Cantor anunciaba a Dedekind que, en contra de la opinión general y por asombroso que resultara, tal correspondencia sí era posible. De hecho cualquier espacio de n dimensiones podía biyectarse con la recta real.

A partir de estos resultados, Cantor llegó a la convicción de que tenía pleno sentido hablar del número de elementos de un conjunto infinito, lo que él llamó su "potencia", de modo que dos conjuntos tienen la misma potencia si y sólo si sus elementos pueden ponerse en correspondencia biunívoca. Ya había probado que existen al menos dos potencias distintas: la potencia común a todos los "continuos" (es decir, $\mathbb{R}^n$) y la potencia de los conjuntos "discontinuos", como $\mathbb{N}$ o $\mathbb{Q}$. Se planteó, no obstante, la posibilidad de que existieran potencias mayores que la del continuo, problema que sólo respondió en toda su generalidad mucho después, con el célebre teorema que lleva su nombre.

De momento, Cantor se centró en el estudio de los subconjuntos de $\mathbb{R}$. Su conjetura era que cualquier subconjunto infinito de $\mathbb{R}$ tenía que ser, o bien de la potencia del continuo, es decir, comparable con la totalidad de los números reales, o bien numerable. Para estudiar si esto era correcto continuó su investigación sobre los conjuntos derivados de puntos. Probó que, ciertamente, los conjuntos de primera especie son numerables. Ahora bien, si un conjunto P es de segunda especie, es decir, si todos sus derivados sucesivos son no vacíos, éstos forman una sucesión decreciente:

$$P^{(1)} \supset P^{(2)} \supset P^{(3)} \supset \cdots$$

por lo que podía considerar lo que llamó[5] $P^{(\omega)} = \bigcap_{n=1}^{\infty} P^{(n)}$. A partir de este conjunto derivado infinito podemos formar nuevos derivados $P^{(\omega+1)}$, $P^{(\omega+2)}$, $P^{(\omega+3)}$, ... Si todos ellos son no vacíos, todavía podemos continuar la sucesión formado el conjunto $P^{(\omega+\omega)} = \bigcap_{n=1}^{\infty} P^{(\omega+n)}$, a partir del cual, a su vez, podemos formar los derivados $P^{(\omega+\omega+1)}$, $P^{(\omega+\omega+2)}$, etc.

Uno de los mayores logros de Cantor fue darse cuenta de que los superíndices que le aparecían en su análisis de los conjuntos de puntos, a los que se refería en los teoremas como "símbolos infinitos", tenían entidad matemática propia. En su trabajo de 1893 "Fundamentos de una teoría general de conjuntos" los

[5]Su primera notación fue $P^{(\infty)}$, pero adoptaremos en todo momento la notación que, tiempo después, tomó por definitiva y que es la habitual hoy en día.

presentó con el nombre de "números transfinitos". Según explicaba, los números transfinitos se obtienen mediante dos principios. El "primer principio de generación" consiste en añadir una unidad. Es el principio que, por sí sólo, genera los números naturales: 0, 1, 2, 3, ... A éstos los llamó "números transfinitos de primera especie". Ahora bien, Cantor afirmaba que, cuando tenemos una sucesión inacabada de números transfinitos, siempre podemos postular la existencia de un nuevo número transfinito como inmediato posterior a todos ellos, y a esto lo llamó el "segundo principio de generación". Así, tras la sucesión de todos los números de primera especie, el segundo principio nos da la existencia de un nuevo número transfinito, el primero de los números de segunda especie, al que cantor llamó ω. A éste podemos aplicarle de nuevo el primer principio, para obtener $\omega+1$, $\omega+2$, etc. En definitiva, combinando la aplicación de ambos principios vamos obteniendo la sucesión transfinita:

$$0,\ 1,\ 2,\ \ldots,\ \omega,\ \omega+1,\ \omega+2,\ \ldots,\ \omega\cdot 2,\ \omega\cdot 2+1,\ \omega\cdot 2+2,\ \ldots,$$

$$\omega\cdot 3,\ \omega\cdot 3+1,\ \omega\cdot 3+2,\ \ldots,\ \omega\cdot 4,\ \omega\cdot 4+1,\ \omega\cdot 4+2,\ \ldots,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\omega^2,\ \omega^2+1,\ \omega^2+2,\ \ldots,\ \omega^2+\omega,\ \omega^2+\omega+1,\ \omega^2+\omega+2,\ \ldots\ \omega^2+\omega\cdot 2,\ \ldots$$

Aquí hemos llamado $\omega\cdot 2=\omega+\omega$ al menor número transfinito que continúa la sucesión de números $\omega+n$ y $\omega^2=\omega\cdot\omega$ al menor número transfinito que continúa la sucesión de números $\omega\cdot m+n$. Estos convenios de notación sugieren la existencia de una aritmética transfinita que Cantor desarrollaría después.

Cantor definió los números transfinitos de segunda especie como los números transfinitos que dejan tras de sí una cantidad numerable de números transfinitos. Por ejemplo, todos los números de la lista anterior a partir de ω son de segunda especie, pero —desde luego— no son todos. Cantor probó que el conjunto de todos los números de segunda especie es no numerable, más aún, probó que cualquier conjunto infinito no numerable debe contener al menos tantos elementos como números de segunda especie, es decir, la potencia de los números de segunda especie es la menor potencia que sigue a la de los números de primera especie.

Podría objetarse que los principios de generación no dejan claro cuál es esa totalidad de los números de segunda especie de cuya potencia estamos hablando, pero Cantor precisó mucho más estas ideas con la ayuda de los conceptos de ordinal y cardinal. Aunque ya introdujo estas nociones en sus "Fundamentos", éstas aparecen mucho más claras en su último trabajo importante, publicado en dos partes en 1895 y 1897. Se trata de las "Contribuciones a la fundamentación de la teoría de conjuntos transfinita".

Aquí define Cantor el "ordinal" de un conjunto ordenado como el concepto al que llegamos cuando hacemos abstracción de la naturaleza de sus elementos y conservamos únicamente su ordenación, de modo que dos conjuntos tienen el mismo ordinal si y sólo si sus elementos pueden ponerse en correspondencia biunívoca conservando el orden. Por otra parte, el "cardinal" de un conjunto

es el concepto al que llegamos cuando hacemos abstracción de la naturaleza de sus elementos, así como de toda posible ordenación.

Para el caso de los conjuntos finitos, ordinal y cardinal son equivalentes. Así, un conjunto ordenado como $a < b < c < d$ tiene tanto ordinal 4 como cardinal 4, en el sentido de que sus elementos están ordenados como los números transfinitos anteriores a 4 o, por otra parte, que tiene tantos elementos como números anteriores a 4. Sin embargo, dos conjuntos infinitos del mismo cardinal pueden tener ordenaciones distintas correspondientes a distintos ordinales. Por ejemplo $\mathbb{N}$ y $\mathbb{Q}$ tienen ambos el mismo cardinal, pero sus ordenaciones son diferentes. Para el cardinal de $\mathbb{N}$, Cantor introdujo la notación $\aleph_0$ (alef cero[6]). Así, podemos decir que el cardinal de $\mathbb{Q}$, o incluso el de los números algebraicos, es también $\aleph_0$, lo cual significa que todos ellos son biyectables con el conjunto de los números naturales. Además Cantor probó que éste era el menor cardinal que podía tener un conjunto infinito, en el sentido de que todo conjunto infinito ha de tener un subconjunto con $\aleph_0$ elementos.

Los números transfinitos se corresponden con los ordinales de los conjuntos bien ordenados. Por ejemplo, podemos decir que ω es el ordinal del conjunto de los números naturales, en el sentido de que $\mathbb{N}$ está ordenado igual que los números menores que ω. No obstante, si ordenamos $\mathbb{N}$ en la forma

$$3, 4, 5, 6, \ldots, 0, 1, 2,$$

seguimos teniendo el mismo conjunto (luego el mismo cardinal) pero ahora su ordinal es $\omega+3$, pues su ordenación es la misma que la de los números transfinitos menores que $\omega + 3$. Similarmente, la ordenación

$$0, 2, 4, 6, \cdots, 1, 3, 5, 7, \cdots$$

tiene ordinal $\omega + \omega$.

Esta correspondencia entre los números transfinitos y los ordinales de conjuntos bien ordenados hace que hoy se llama simplemente "ordinales" a los números transfinitos.[7]

Así, los ordinales de segunda especie están en correspondencia biunívoca con las formas esencialmente distintas de ordenar bien (o sea, de modo que todo subconjunto no vacío tenga mínimo) un conjunto numerable. Lo que cantor había probado en los "Fundamentos" es que la cantidad de buenos órdenes posibles en un conjunto numerable es no numerable y es, de hecho, la menor potencia no numerable. A esta potencia la llamó $\aleph_1$. De este modo, Cantor había justificado que existe un menor cardinal posterior a $\aleph_0$. Su hipótesis inicial sobre el cardinal del continuo (la que llamó hipótesis del continuo) podía enunciarse ahora como que el cardinal del continuo es exactamente $\aleph_1$.

Cantor desarrolló una aritmética ordinal que sistematizaba expresiones como $\omega + \omega = \omega \cdot 2$ y una aritmética cardinal, que le permitía relacionar los cardinales

[6] Alef es la primera letra del alfabeto hebreo.

[7] No obstante, hemos de tener en cuenta que Cantor hablaba de ordinales de conjuntos arbitrarios. Así, para Cantor, $\mathbb{R}$ y el intervalo abierto $]0, 1[$ tenían el mismo ordinal (y por consiguiente el mismo cardinal), mientras que $\mathbb{R}$ y $[0, 1]$ tenían el mismo cardinal pero distinto ordinal.

de distintos conjuntos. En términos de esta aritmética, pudo calcular que la potencia del continuo era $2^{\aleph_0}$, lo cual expresa esencialmente que todo número real viene determinado por una sucesión decimal binaria de $\aleph_0$ ceros o unos. En estos términos, la hipótesis del continuo admitía una expresión muy elegante, a saber, la ecuación $2^{\aleph_0} = \aleph_1$.

Las ideas de Cantor se encontraron con la oposición de una poderosa corriente de matemáticos de la época, encabezada por Kronecker, que propugnaba una fundamentación finitista de la matemática. Naturalmente, el trabajo de Cantor constituía la antítesis de este programa. No obstante, a finales de siglo el trabajo de Cantor contaba ya con gran aceptación. Sus "Contribuciones" fueron rápidamente traducidas y difundidas entre toda la comunidad matemática y, en 1900, Hilbert puso la hipótesis del continuo a la cabeza de su lista de los 23 problemas más importantes que tenía planteada la matemática del siglo XX.

La actividad de Cantor no había cesado por aquel entonces. Llegó a definir la sucesión completa de los alefs

$$0,\ 1,\ 2,\ \ldots\ \aleph_0,\ \aleph_1,\ \aleph_2,\ \ldots\ \aleph_\omega,\ \aleph_{\omega+1},\ \aleph_{\omega+2},\ldots\ \aleph_{\omega\cdot 2},\ \aleph_{\omega\cdot 2+1},\ \ldots$$

En general, probó que podemos hablar de $\aleph_\alpha$, donde α es cualquier ordinal, y sabía que los cardinales de esta forma se corresponden con los cardinales de los conjuntos que pueden ser bien ordenados. En trabajos anteriores a las "Contribuciones" había dado por evidente que todo conjunto puede ser bien ordenado, pero al tratar de exponer su teoría con el máximo rigor no consideró justificada esta hipótesis. La importancia de esto residía en que Cantor sabía que los cardinales $\aleph_\alpha$ estaban bien ordenados (por la relación en virtud de la cual un cardinal $\mathfrak{m}$ es menor que un cardinal $\mathfrak{n}$ si todo conjunto de cardinal $\mathfrak{n}$ contiene un subconjunto de cardinal $\mathfrak{m}$), pero para cardinales cualesquiera no era capaz de probar siquiera que estuvieran totalmente ordenados, es decir, que dados dos cardinales cualesquiera, uno tuviera que ser menor que el otro. Cantor creía esencial que se cumpliera esta propiedad como justificación para considerar a los cardinales como números en el pleno sentido del término.

En una carta a Dedekind en 1899 daba una demostración de que todo cardinal es un alef, lo cual zanjaba el problema. En la prueba hacía uso del concepto de "multiplicidad inconsistente", del que ya hemos hablado en la introducción a la segunda parte y que se corresponde con lo que ahora llamamos "clase propia". Recordemos que Cantor había introducido este concepto como respuesta a las paradojas que plagaban su teoría. Por ejemplo, al postular que toda sucesión de números transfinitos puede prolongarse con un nuevo número transfinito, la sucesión de todos los números transfinitos resulta ser contradictoria. Esto es esencialmente lo que se conoce como la "antinomia de Burali-Forti". Más concretamente, en su teoría sobre las multiplicidades inconsistentes Cantor afirmó que la colección Ω de todos los ordinales era una multiplicidad inconsistente, de modo que, aunque estaba bien ordenada, no le correspondía un número transfinito porque no podía ser considerada propiamente como una totalidad. En su prueba de que todo cardinal es un alef usaba que una multiplicidad es contradictoria si y sólo si contiene una parte biyectable con Ω, lo cual no era evidente

en absoluto.

La primera demostración enteramente satisfactoria de que todo conjunto puede ser bien ordenado (lo cual equivale a que todo cardinal es un alef) la dio Ernst Zermelo en 1904. En ella partía del siguiente principio:

> *Dado cualquier conjunto M, existe una aplicación que a cada elemento no vacío de M le asigna uno de sus elementos.*

Zermelo llamó a esta hipótesis (al igual que hemos hecho nosotros) el *axioma de elección* y consideraba que sin él era imposible probar el teorema de buena ordenación que había "demostrado" Cantor y que hoy se conoce como teorema de Zermelo. En realidad, en ese mismo año, Bertrand Russell y Alfred Whitehead descubrieron casualmente dicho axioma infiltrado en la demostración de un resultado de los Principia Mathematica, tal y como ya habíamos comentado.

El axioma de elección dio lugar a grandes controversias sobre su legitimidad. Por ejemplo, Peano lo consideraba inaceptable debido a que no se deducía de los axiomas de su sistema de lógica, que al parecer para él reflejaban todas las posibilidades válidas de razonamiento. Según él, una demostración consiste en concluir la verdad de una afirmación reduciéndola lógicamente a otras más elementales y, en último extremo, a principios elementales indiscutibles (los suyos). Cuando Russell conjeturó que probablemente el axioma no podría demostrarse a partir de los axiomas de sus Principia Mathematica también mostró su rechazo. Russell consideraba que una noción lógicamente correcta de "colección de objetos" estaba asociada a la noción de "propiedad", de manera que las colecciones, multiplicidades, conjuntos, o como se las quisiera llamar, debían en cierto sentido venir definidas por propiedades. Sólo con esta noción determinada de colección de objetos se evitaba de forma natural las paradojas de la teoría de conjuntos. El axioma de elección afirmaba en último extremo la existencia de un conjunto cuyos elementos no estaban relacionados necesariamente por ninguna propiedad común. Si tengo infinitas cajas de zapatos puedo formar la colección de todos los zapatos derechos que hay en las cajas, pero si tengo cajas de calcetines iguales, el axioma de elección nos permitiría construir un conjunto formado por un calcetín de cada caja, pero ¿cómo podemos admitir la existencia de una colección indefinible?

Zermelo, en cambio, difería de Peano y de Russell, que veían a la lógica matemática como algo cerrado y completamente contenida en sus respectivos sistemas formales, y consideraba que el razonamiento matemático estaba abierto a cualquier nuevo principio que pudiera calificarse de intuitivamente evidente y ése era el caso del axioma de elección. Para Zermelo, que por lo visto tenía una idea más amplia de lo que era un conjunto, no había diferencia entre este axioma y los demás principios que Peano o Russell habían admitido como axiomas válidos. De hecho, muchas demostraciones conocidas desde antiguo y de gran importancia en todas las ramas de la matemática lo usaban tácitamente sin que nadie hasta entonces hubiera objetado nada. El hecho de que muchos matemáticos lo hubieran supuesto sin vacilación era una prueba de su carácter intuitivo.

Cantor murió en 1918 a los setenta y tres años de edad. Murió sin conocer

la respuesta al problema al cual se había dedicado casi exclusivamente desde el momento en que él mismo lo planteó y que le había obsesionado hasta la locura. Una respuesta completa a la hipótesis del continuo no se encontró hasta mucho después, hasta el año 1963, y la naturaleza de la respuesta era bien diferente a lo que cualquiera en sus tiempos pudiera haber imaginado.

En los años que siguieron a la muerte de Cantor el status lógico de la matemática quedó completamente clarificado. Todas las demostraciones matemáticas conocidas podían obtenerse mediante unas reglas lógicas conocidas a partir de unos principios elementales conocidos, los que hemos estudiado en la segunda parte de este libro, donde el único axioma que suscitaba controversias era a lo sumo el axioma de elección.

Esto permitía un enfoque lógico de algunos problemas que hasta el momento caían en el campo exclusivo de la filosofía matemática. En efecto, en 1938 Gödel anunció que si NBG$-AE$ es consistente —cosa que, según sabemos, no puede probarse—, también lo es NBG (con el axioma de elección) más la hipótesis del continuo. En realidad Gödel probó la consistencia de dos afirmaciones más fuertes que éstas.

A partir de aquí la solución definitiva del viejo problema cantoriano era de suponer: Gödel había demostrado que en la teoría de conjuntos había sentencias sobre numeros naturales que no podían ser decididas y, si algo tan sencillo como la aritmética resultaba incompleto, con mayor razón debía de haber sentencias indecidibles en la teoría abstracta de conjuntos. De hecho se sabía que ciertas afirmaciones sobre cardinales infinitos implican la existencia de modelos de ZFC, por lo que resultan indecidibles en virtud de los teoremas de incompletitud. Aunque la hipótesis del continuo no era una de ellas, la prueba de consistencia de Gödel dejaba entrever que los axiomas de la teoría de conjuntos eran demasiado imprecisos en cierto sentido para implicar la hipótesis del continuo.

Ciertamente Cantor había perseguido un imposible. En 1963 Paul Cohen desarrolló una nueva técnica muy potente que le permitió probar la independencia (es decir, la indemostrabilidad) del axioma de elección y la hipótesis del continuo, siempre suponiendo la consistencia de la teoría de conjuntos, como prescribe el teorema de incompletitud. Quedaba así cerrada la cuestión que Cantor planteara ochenta años atrás.

La hipótesis del continuo es sólo una de las innumerables afirmaciones que no pueden decidirse a partir de los axiomas de ZFC o NBG. Existen muchas más no sólo en la teoría de conjuntos propiamente dicha, sino también en la topología, el análisis matemático o el álgebra. Aún hoy resultan incómodas a muchos matemáticos no especializados en teoría de conjuntos. La situación es muy similar a la conmoción que produjo el descubrimiento de las geometrías no euclídeas. En un principio dichas geometrías eran inconcebibles, después eran posibles pero absurdas y, finalmente, los matemáticos han comprendido que todas las geometrías tienen el mismo valor matemático. Así mismo, durante una época los matemáticos creyeron necesario decantarse por alguna de las alternativas que permite la indecidibilidad: o se acepta, o no se acepta el axioma de elección, o se considera razonable, o no se considera razonable la hipótesis del continuo, etc. Sin embargo, no es descabellado afirmar que estos dilemas han

sido superados, al menos por los matemáticos que realmente se enfrentan a ellos en su trabajo. Hay axiomas que contradicen el axioma de elección, pero que tienen consecuencias tan interesantes como las de este axioma, las matemáticas con la hipótesis del continuo son unas, y las matemáticas sin la hipótesis del continuo son otras, ni más ni menos interesantes.

Sin embargo, muchos matemáticos a los que estos hechos les pillan más de lejos, siguen desconcertados ante ellos. Su error es suponer que existe una noción natural y bien definida de conjunto, igual que los antiguos creían que existe una noción natural y bien definida de recta. En realidad la situación es peor, pues si acordamos entender por recta lo que intuitivamente entendemos por recta, entonces podemos asegurar que la única geometría que habla realmente de rectas es la geometría tridimensional euclídea. Las otras geometrías son posibles a costa de cambiar la noción de recta por otras nociones diferentes. Sin embargo, en lo que respecta a los conjuntos carecemos de una noción intuitiva de conjunto que se corresponda con los axiomas de ZFC, en el sentido de que si nos presentan dos modelos distintos de ZFC no podemos decir cuál de ellos se corresponde con una presunta intuición y cuál no.

Lo único que nos dice nuestra intuición (sería más exacto decir nuestro entendimiento) es que si tenemos dada una colección de objetos, tenemos derecho a considerarla como un todo al que tratar como un objeto más, y esto es un conjunto. El problema es que en ZFC hablamos de conjuntos que no proceden de considerar una colección dada como un todo. El ejemplo más simple es $\mathcal{P}\mathbb{N}$. No tenemos nada a lo que llamar "la totalidad de los subconjuntos de $\mathbb{N}$", y pese a ello postulamos un objeto que nos permita considerar como un todo a no sabemos qué. En realidad $\mathcal{P}\mathbb{N}$ es un "saco" en el que podemos ir metiendo todos los subconjuntos de $\mathbb{N}$ que nos vayamos encontrando, pero no es así como lo tratamos en ZFC, pues ahí presuponemos que $\mathcal{P}\mathbb{N}$ tiene todos sus elementos dados de antemano. Para Arquímedes, la circunferencia era "la curva plana que describe el extremo de un segmento cuando gira sobre su otro extremo", mientras que para Euclides era "el conjunto de los puntos del plano que equidistan de otro llamado centro". Arquímedes construía, Euclides describía algo que estaba ahí. El problema es que la diferencia no es meramente psicológica, sino que tiene consecuencias prácticas: si suponemos que $\mathcal{P}\mathbb{N}$ es algo que está ahí, a través de él podemos construir nuevos subconjuntos de $\mathbb{N}$ y, como suponemos que $\mathcal{P}\mathbb{N}$ es inmutable, debemos concluir que los nuevos subconjuntos ya estaban ahí. Si no hubiéramos atribuido realidad a $\mathcal{P}\mathbb{N}$ nunca habríamos obtenido esos conjuntos. Este tipo de razonamiento, de validez dudosa, lleva a contradicciones si en lugar de aplicarlo a $\mathbb{N}$ lo aplicamos a un hipotético conjunto de todos los conjuntos. Así pues, nos vemos obligados a admitir que no tenemos ningún derecho a hablar de una hipotética totalidad de los conjuntos que en modo alguno conocemos. En cambio, cuando lo aplicamos a $\mathbb{N}$ no obtenemos ninguna contradicción, sino sólo una mentira: que $\mathcal{P}\mathbb{N}$ es un conjunto no numerable.

En contra de lo que podría parecer, esto no descalifica a la matemática abstracta como un fraude. Para entender esto pensemos mejor en los números reales, de los que se podría decir lo mismo. Imaginemos que quiero estudiar la función $y = x^3 - 6x$ sobre los números racionales. Si me restrinjo a los

números racionales todo cuanto diga tendrá un significado intuitivo (metamatemático) claro. Si hago que un ordenador me dibuje (punto a punto) la gráfica de esta función obtendré una figura con unas peculiaridades (unos máximos, unos mínimos, unos puntos de inflexión, etc.) y resulta que la mejor forma que tengo de estudiar esas peculiaridades es a través del cálculo diferencial, sobre $\mathbb{R}$, naturalmente. Si deduzco que tiene extremos en los puntos (irracionales) $\pm\sqrt{2}$, esta información se aplica certeramente a lo que me muestra la gráfica. Posiblemente podría haberla obtenido sin recurrir a la quimérica totalidad de los números reales, pero, sin duda alguna, el formalismo necesario para ello sería muchísimo más complejo que el que he necesitado con los medios convencionales y hubiera oscurecido las ideas fundamentales del análisis.

En resumen, para hacer matemáticas necesitamos un universo de conjuntos. En ningún momento es esencial que se trate de una hipotética "totalidad" de los conjuntos (afortunadamente, porque no existe) y, sin embargo, la forma más sencilla de trabajar con un universo de conjuntos suficiente para hacer matemáticas es pensar que se trata de la totalidad de los conjuntos. Al fin y al cabo, es la totalidad de los conjuntos ... que estamos considerando. Visto así, ni siquiera necesitamos considerar un fraude la no numerabilidad de $\mathbb{R}$. Si la entendemos bien, lo que afirma no es una propiedad de $\mathbb{R}$, sino una propiedad de los conjuntos que estamos considerando: ninguno de ellos es una biyección entre $\mathbb{N}$ y $\mathbb{R}$.

Si entendemos la matemática no estudia unos objetos dados a priori llamados conjuntos sino unas estructuras llamadas modelos de ZFC, comprenderemos que es igual de absurdo sorprenderse de que haya "conjuntos con hipótesis del continuo" y "conjuntos sin hipótesis del continuo" como lo sería sorprenderse de que haya anillos conmutativos y anillos no conmutativos. Si un matemático se propone estudiar los anillos dirá "sea A un anillo" y, a partir de ahí, demostrará cosas. Si a partir de un momento necesita que sea conmutativo, dirá "supongamos que A es conmutativo", pero no se planteará si realmente su anillo (arbitrario) es o no conmutativo. Igualmente, un teorema que requiere la hipótesis del continuo es un teorema que cumplirán los modelos de ZFC que cumplan la hipótesis del continuo y no lo cumplirán necesariamente los otros modelos.

Si, pese a estas reflexiones, todavía hay alguien que se sorprenda de que existan modelos de las dos alternativas, lo más probable es que su sorpresa se disipe en cuanto vea las construcciones explícitas de modelos correspondientes. No obstante, no entraremos en ello en esta tercera parte del libro, donde —como ya hemos anunciado— únicamente pretendemos completar la teoría que hemos introducido en la parte anterior, ahora en su parte propiamente matemática.

Capítulo XI

Números ordinales

Empezamos nuestro estudio de la teoría de conjuntos con la construcción de los números ordinales. Los ordinales resultan ser el sustituto indispensable de los números naturales a la hora de trabajar con conjuntos arbitrarios, que en general serán mucho mayores que $\mathbb{N}$. Mientras no indiquemos lo contrario trabajamos en NBG* aunque, como ya sabemos, "mordiéndonos la lengua" en los momentos oportunos y dando rodeos, podríamos eliminar toda referencia a clases propias, por lo que también podemos suponer que trabajamos en ZF*. En particular no usamos el axioma de infinitud ni, por consiguiente, suponemos definidos los números naturales. Así obtendremos una construcción alternativa de los números naturales que no requiere AI.

11.1 La construcción de los ordinales

Desde el punto de vista cantoriano, los ordinales se correspondían con todas las formas posibles de ordenar bien un conjunto (es decir, de modo que todo subconjunto no vacío tenga mínimo). Por ello, el primer intento de dar rigor a esta noción de ordinal fue considerar que el ordinal de un conjunto bien ordenado $(X, \leq)$ estaba formado por todos los conjuntos ordenados igual, es decir, todos los conjuntos bien ordenados $(Y, \leq)$ tales que exista una biyección $f : X \longrightarrow Y$ que conserva el orden. Ahora bien, es fácil ver que incluso el ordinal 1, formado por todos los conjuntos bien ordenados con un elemento, sería así una clase propia, pues podríamos definir una aplicación inyectiva $f : \mathrm{V} \longrightarrow 1$ que a cada conjunto x le asignara el par $(\{x\}, \leq)$, donde $\leq$ es el orden obvio.

Fue von Neumann el que tuvo la idea de salvar esta dificultad definiendo un ordinal, no como la clase de todos los conjuntos con una ordenación dada, sino como un conjunto canónico representante de dicha ordenación. Así, identificó el ordinal 0 con el conjunto vacío $\varnothing$, el ordinal 1 tenía que ser un conjunto ordenado con un elemento, y lo más simple era tomar $1 = \{0\}$, con el orden obvio. Similarmente, $2 = \{0, 1\}$, $3 = \{0, 1, 2\}$, etc. Obtenemos así los números naturales, tal y como los definimos en el capítulo VIII. Allí seguimos la construcción de Dedekind, que partía de un conjunto infinito arbitrario y producía

un conjunto $\mathbb{N}$ dependiente del conjunto de partida, pero al final pasamos a los números naturales en el sentido de von Neumann, para tener así los números que se usan habitualmente en la actualidad. Sin embargo, la técnica de von Neumann no se detiene ahí, sino que ahora podemos definir el ordinal ω que representa el orden usual de $\mathbb{N}$ como el propio $\mathbb{N}$, es decir,

$$\omega = \{0,\ 1,\ 2,\ 3,\ \ldots\ \}.$$

Similarmente,

$$\begin{aligned}
\omega+1 &= \{0,\ 1,\ 2,\ 3,\ \ldots\ \omega,\},\\
\omega+2 &= \{0,\ 1,\ 2,\ 3,\ \ldots\ \omega,\ \omega+1\},\\
\omega+\omega &= \{0,\ 1,\ 2,\ 3,\ \ldots\ \omega,\ \omega+1,\ \omega+2,\ \ldots\}.
\end{aligned}$$

En general, cada ordinal coincidirá con el conjunto de todos los ordinales menores que él. Es fácil ver entonces que la relación de orden es simplemente la inclusión.

Definición 11.1 Una clase Y es *transitiva* si $\bigwedge x(x \in Y \to x \subset Y)$ o, equivalentemente, si $\bigwedge uv(u \in v \wedge v \in Y \to u \in Y)$.

Diremos que Y es *conexa* si $\bigwedge uv \in Y(u \in v \vee v \in u \vee u = v)$.

La clase Y está *bien fundada* si $\bigwedge X(X \subset Y \wedge X \neq \varnothing \to \bigvee u \in X\, u \cap X = \varnothing)$.

Diremos que Y es un *ordinal* si es una clase transitiva, conexa y bien fundada.

Es claro que los ejemplos de ordinales que acabamos de esbozar cumplen esta definición de ordinal. Observemos que una clase Y está bien fundada si toda subclase[1] no vacía X tiene lo que llamaremos un *elemento minimal,* es decir un elemento u que está en X pero ninguno de cuyos elementos está en X. El axioma de regularidad afirma que todo conjunto no vacío tiene un elemento minimal luego, si supusiéramos este axioma, podríamos decir que un conjunto es un ordinal si y sólo si es transitivo y conexo.

Es inmediato comprobar que toda subclase de una clase conexa o bien fundada es también conexa o bien fundada. Esto no es cierto para clases transitivas (basta pensar en $\{2,3\} \subset \{0,1,2,3\}$). Vamos a necesitar el siguiente hecho técnico:

Teorema 11.2 *Si X es una clase bien fundada entonces $X \notin X$.*

Demostración: Si $X \in X$ entonces $\{X\} \subset X \wedge \{X\} \neq \varnothing$ (X es un conjunto precisamente porque $X \in X$). Sea u un elemento minimal de $\{X\}$. Necesariamente, $u = X$, pero $X \in X \cap \{X\}$, contradicción. ∎

Hemos de probar que un ordinal está compuesto por los ordinales menores que él. Por lo pronto tenemos:

[1] Observemos que "X está bien fundada" no es una fórmula normal, pues en ella aparece una cuantificación $\bigwedge X$ sobre una clase que no tiene por qué ser un conjunto. Sin embargo, "cto $Y \wedge Y$ está bien fundado" sí que es normal, pues ahora podemos cambiar $\bigwedge X$ por $\bigwedge x$. Lo mismo sucede con "Y es un ordinal".

Teorema 11.3 *Los elementos de los ordinales son ordinales.*

DEMOSTRACIÓN: Sea Y un ordinal y sea $X \in Y$. Por transitividad $X \subset Y$ y por consiguiente X es conexa y bien fundada. Falta probar que es transitiva, es decir, que $\bigwedge uv(u \in v \wedge v \in X \to u \in X)$.

Si $u \in v \wedge v \in X$, tenemos $v \in X \wedge X \in Y$ y, como Y es transitiva, $v \in Y$, e igualmente $u \in Y$. Así pues, $\{u, v, X\} \subset Y$. Como Y está bien fundada se cumplirá

$$u \cap \{u, v, X\} = \varnothing \quad \vee \quad v \cap \{u, v, X\} = \varnothing \quad \vee \quad X \cap \{u, v, X\} = \varnothing,$$

pero $u \in v \cap \{u, v, X\}$ y $v \in X \cap \{u, v, X\}$, luego ha de ser $u \cap \{u, v, X\} = \varnothing$. Como Y es conexa ha de ser $u \in X \vee X \in u \vee u = X$, pero si $X \in u$ entonces $X \in u \cap \{u, v, X\} = \varnothing$, y si $X = u$ entonces $v \in u \cap \{u, v, X\} = \varnothing$. Así pues, se ha de cumplir $u \in X$, como buscábamos. ∎

Ahora podemos probar:

Teorema 11.4 *Todo ordinal está bien ordenado por la inclusión.*

DEMOSTRACIÓN: Sea Y un ordinal. Hay que probar que la relación R dada por $u\,R\,v$ syss $u \subset v$ es un buen orden en Y. Claramente es un orden parcial. Sea $X \subset Y$, $X \neq \varnothing$ y veamos que tiene mínimo.

Sea $u \in X$ un elemento minimal, es decir, $u \cap X = \varnothing$. Vamos a comprobar que u es el mínimo de X, o sea, que $\bigwedge z \in X\, u \subset z$.

Dado $z \in X$, se cumple que $u, z \in Y$, luego $u \in z \vee z \in u \vee u = z$. No puede ser $z \in u$, pues entonces $z \in u \cap X = \varnothing$. Por lo tanto nos queda $u \in z \vee u = z$. Por el teorema anterior, z es un ordinal, luego si $u \in z$, tendremos $u \subset z$, y trivialmente también si $u = z$, por lo que en cualquiera de los dos casos $u \subset z$. Así pues Y está bien ordenado.[2] ∎

Con esto tenemos probado lo esencial sobre (los elementos de) un ordinal, pero ahora veremos que podemos ordenar la totalidad de los ordinales. Para ello necesitamos algunos resultados previos:

Teorema 11.5 *Si X e Y son ordinales y $X \subset Y$, entonces $X \in Y \vee X = Y$.*

DEMOSTRACIÓN: Si $X \neq Y$ entonces $Y \setminus X \neq \varnothing$. Como Y es un ordinal, $Y \setminus X$ tiene un elemento minimal u, es decir, $u \cap (Y \setminus X) = \varnothing$.

Si $z \in u$, entonces $z \notin Y \setminus X$ y $z \in Y$ (pues $z \in u \in Y$), luego $z \in X$. Por lo tanto $u \subset X$.

Si $z \in X$, entonces tenemos $z, u \in Y$, luego $z \in u \vee u \in z \vee z = u$. Si $u \in z$, entonces $u \in z \in X$, luego $u \in X$, contradicción ($u \in Y \setminus X$). Si $z = u$ entonces de nuevo $u \in X$, contradicción. Por lo tanto $z \in u$, y así $X \subset u$.

En definitiva $X = u \in Y$. ∎

[2]Recordemos que un buen orden es siempre un orden total, pues si u, v son dos elementos, se cumplirá $u \leq v$ o $v \leq u$ según quién sea el mínimo del conjunto $\{u, v\}$.

Teorema 11.6 *La intersección de dos ordinales es un ordinal.*

DEMOSTRACIÓN: Sean X, Y ordinales. Como $X \cap Y \subset X$, trivialmente $X \cap Y$ es conexa y bien fundada. Falta ver que es transitiva:

Si $u \in X \cap Y$, entonces $u \in X \wedge u \in Y$, $u \subset X \wedge u \subset Y$, luego $u \subset X \cap Y$. ∎

Con esto tenemos el resultado que buscábamos:

Teorema 11.7 *Sean X, Y ordinales. Entonces $X \in Y \vee Y \in X \vee X = Y$.*

DEMOSTRACIÓN: $X \cap Y$ es un ordinal, $X \cap Y \subset X$ y $X \cap Y \subset Y$. Por el teorema 11.5 tenemos $(X \cap Y \in X \vee X \cap Y = X) \wedge (X \cap Y \in Y \vee X \cap Y = Y)$. De aquí se sigue

$$(X \cap Y \in X \wedge X \cap Y \in Y) \vee (X \cap Y \in X \wedge X \cap Y = Y)$$

$$\vee (X \cap Y = X \wedge X \cap Y \in Y) \vee (X \cap Y = X \wedge X \cap Y = Y),$$

o sea $X \cap Y \in X \cap Y \quad \vee \quad Y \in X \quad \vee \quad X \in Y \quad \vee \quad X = Y$. El primer caso se descarta por el teorema 11.2. ∎

Estos últimos resultados se resumen en uno si introducimos la clase de todos los ordinales:

Definición 11.8 Sea[3] $\Omega \equiv \{x \mid x \text{ es un ordinal}\}$, es decir, Ω es la clase de todos los *conjuntos* que son ordinales.

¿Existen ordinales que sean clases propias? Sólo uno:

Teorema 11.9 *Ω es un ordinal.*

DEMOSTRACIÓN: Ω es transitiva por el teorema 11.3 y es conexa por el teorema 11.7. Veamos que está bien fundada.

Sea $X \subset \Omega$, $X \neq \varnothing$. Sea $y \in X$. Si $y \cap X = \varnothing$ entonces y es un elemento minimal de X. Supongamos ahora que $y \cap X \neq \varnothing$. Como y es un ordinal e $y \cap X \subset y$, $y \cap X$ tiene un elemento minimal u, es decir $u \in y \cap X \wedge u \cap y \cap X = \varnothing$.

Como $u \in y$, también $u \subset y$, de donde $u \in X \wedge u \cap X = \varnothing$, es decir, u es un elemento minimal de X. Por lo tanto Ω es un ordinal. ∎

Es claro que Ω no puede ser un conjunto, o de lo contrario sería $\Omega \in \Omega$, en contra del teorema 11.2. Del teorema 11.7 se sigue que Ω es el único ordinal que no es un conjunto, pues si Y es cualquier ordinal, ha de cumplirse

$$\Omega \in Y \quad \vee \quad Y \in \Omega \quad \vee \quad Y = \Omega,$$

y la primera opción se descarta precisamente porque Ω no es un conjunto.

[3]Esta definición es correcta porque ya hemos observado que la fórmula "cto $x \wedge x$ es un ordinal" es normal.

Llamaremos *números ordinales* a los elementos de Ω, es decir, a los conjuntos ordinales. Los representaremos con letras griegas, de modo que $\bigwedge\alpha$ deberá entenderse como $\bigwedge\alpha(\alpha \in \Omega \to \cdots$, y análogamente con el particularizador.

Si α y β son ordinales, escribiremos $\alpha \leq \beta \equiv \alpha \subset \beta$ y, en virtud de 11.5, $\alpha < \beta \equiv \alpha \in \beta$. Hemos probado que la inclusión $\leq$ es un buen orden en Ω.

Según decíamos, el hecho de que Ω es un ordinal, junto con que está bien ordenado por la relación de pertenencia (como orden estricto), resume todos los resultados que hemos probado hasta ahora.[4]

Introducimos la notación $0 \equiv \varnothing$, $x' = x \cup \{x\}$. Los principios con los que Cantor introdujo los ordinales se convierten ahora en el teorema siguiente:

Teorema 11.10 *Se cumple:*

a) 0 *es el mínimo ordinal.*

b) Si α *es un ordinal, entonces* α' *también lo es, y es el mínimo ordinal mayor que* α*.*

c) Todo conjunto de ordinales $x \subset \Omega$ *tiene supremo, el cual es concretamente*

$$\sigma = \bigcup_{\alpha\in x} \alpha.$$

Demostración: a) Es inmediato que $\varnothing$ es un ordinal. Además es el mínimo porque el orden es la inclusión.

b) Si $\alpha \in \Omega$, es claro que $\alpha' \subset \Omega$, luego α' es un conjunto conexo y bien fundado. Falta probar que es transitivo, pero si $u \in \alpha'$, entonces $u \in \alpha \vee u = \alpha$, y en ambos casos $u \subset \alpha \subset \alpha'$.

Obviamente $\alpha < \alpha'$ y si $\alpha < \beta$ entonces $\alpha \subset \beta$ por transitividad, luego $\alpha' = \alpha \cup \{\alpha\} \subset \beta$, es decir, $\alpha' \leq \beta$. Por consiguiente α' es el menor ordinal mayor que α.

c) Como todo $\alpha \in x$ está contenido en Ω, es claro que $\sigma \subset \Omega$, luego es un conjunto conexo y bien fundado. Hemos de probar que es transitivo, pero si $\beta \in \sigma$, entonces existe un $\alpha \in x$ tal que $\beta \in \alpha$, luego por la transitividad de α es $\beta \subset \alpha \subset \sigma$. Por consiguiente $\sigma \in \Omega$.

Teniendo en cuenta que el orden es la inclusión, es inmediato que σ es el supremo de x. ∎

Notemos que el recíproco del apartado 3 se cumple trivialmente, pues si una subclase X de Ω tiene una cota superior α entonces $X \subset \alpha'$, luego X es un conjunto. Ahora podemos definir los números naturales:

Definición 11.11 $\mathrm{Nat}\, x \equiv x \in \Omega \wedge \bigwedge\alpha(\alpha \in x' \to \alpha = 0 \vee \bigvee\beta \in \alpha\, \alpha = \beta')$, o sea, un número natural es un ordinal tal que tanto él como sus antecesores no nulos tienen un inmediato anterior. Llamaremos $\omega \equiv \mathbb{N} \equiv \{x \mid \mathrm{Nat}\, x\}$, es decir, a la clase de todos los números naturales.

[4] En ZF* la afirmación de que Ω es un ordinal debe entenderse desarrollando la definición de ordinal, es decir, como que los elementos de los ordinales son ordinales, etc.

No podemos demostrar que ω es un conjunto sin usar el axioma de infinitud, pero esto no es necesario para probar los resultados básicos.

Teorema 11.12 *ω es un ordinal.*

DEMOSTRACIÓN: Como $\omega \subset \Omega$, basta ver que es transitiva. Si $u \in v \wedge v \in \omega$, entonces v es un número natural. Por definición tenemos que

$$\bigwedge\alpha(\alpha \in v' \to \alpha = 0 \vee \bigvee\beta \in \alpha\, \alpha = \beta'),$$

como $u < v$, también $u' \subset v'$ en particular

$$\bigwedge\alpha(\alpha \in u' \to \alpha = 0 \vee \bigvee\beta \in \alpha\, \alpha = \beta'),$$

luego $u \in \omega$. ∎

Teorema 11.13 (Axiomas de Peano) *Se cumple:*

1. $0 \in \omega$,
2. $\bigwedge n \in \omega\, n' \in \omega$,
3. $\bigwedge n \in \omega\, n' \neq 0$,
4. $\bigwedge mn \in \omega(m' = n' \to m = n)$,
5. $\bigwedge Y(Y \subset \omega \wedge 0 \in Y \wedge \bigwedge n \in Y\, n' \in Y \to Y = \omega)$.

DEMOSTRACIÓN: 1) es trivial, si $n \in \omega$ y $\alpha \in n''$, entonces, o bien $\alpha \in n'$ o bien $\alpha = n'$. En el primer caso $\alpha = 0 \vee \bigvee\beta \in \alpha\, \alpha = \beta'$, porque $n \in \omega$. Esto también se cumple en el segundo caso, tomando $\beta = n$. Por consiguiente $n' \in \omega$.

Las propiedades 3) y 4) son trivialmente válidas para ordinales cualesquiera (teniendo en cuenta que n' es el menor ordinal mayor que n). Veamos 5).

Si $Y \subset \omega \wedge 0 \in Y \wedge \bigwedge n \in Y\, n' \in Y$ pero $Y \neq \omega$, entonces $\omega \setminus Y$ tendrá un mínimo elemento n, que será no nulo por hipótesis. Por la definición de número natural $n = m'$ para cierto $m < n$, luego $m \in Y$ y, por hipótesis $n = m' \in Y$, lo cual es una contradicción. ∎

Es claro que la clase ω es la misma clase $\mathbb{N}$ construida en el capítulo VIII, si bien aquí la hemos obtenido sin el axioma de infinitud y, por consiguiente, sin poder garantizar que sea un conjunto. En efecto:

Teorema 11.14 *Son equivalentes:*

a) El axioma de infinitud: $\bigvee f x(f : x \longrightarrow x$ *inyectiva y no suprayectiva*$)$,

b) ω *es un conjunto,*

c) $\omega \neq \Omega$,

d) $\omega \in \Omega$.

DEMOSTRACIÓN: En el capítulo VIII, partiendo de a) hemos construido un conjunto $\mathbb{N}$ que cumple los axiomas de Peano con el mismo 0 y la misma función siguiente, de donde se sigue inmediatamente que $\mathbb{N} = \omega$, luego ω es un conjunto.

Puesto que ω es un ordinal, es claro que $2 \to 3$. Si tenemos 3, entonces $\omega \neq \Omega$ porque Ω no es un conjunto. $4 \to 5$ por el teorema 11.7. Finalmente, si $\omega \in \Omega$, entonces $x = \omega$ y $f : \omega \longrightarrow \omega$ dada por $f(n) = n'$ cumplen el axioma de infinitud. ∎

Definición 11.15 Un ordinal $\alpha \in \Omega$ es un *ordinal sucesor* si existe un $\beta \in \Omega$ tal que $\alpha = \beta'$. Un *ordinal límite* es un ordinal $\lambda \in \Omega$ que no es nulo ni un ordinal sucesor.

Obviamente, si $\omega = \Omega$ entonces los únicos ordinales son los números naturales y no existen ordinales límite, pero si $\omega \in \Omega$ entonces ω es un ordinal límite. Además es el menor de todos.

Aunque todavía no estemos en condiciones de probarlo con rigor, el axioma de infinitud implica, de hecho, que existen ordinales límite arbitrariamente grandes. En efecto, dado un ordinal α, podemos considerar el supremo de los ordinales α, α', α'', α''', ...

Tenemos, pues, que todo ordinal α se encuentra en uno y sólo uno de los tres casos siguientes:

$$\alpha = 0, \qquad \bigvee\beta\, \alpha = \beta', \qquad \alpha \text{ es un ordinal límite.}$$

Usaremos la letra λ para referirnos a ordinales límite, es decir, $\bigwedge\lambda$ deberá entenderse como $\bigwedge\lambda(\lambda \text{ es un ordinal límite} \to \cdots$

11.2 Inducción y recursión transfinita

Los ordinales presentan propiedades muy parecidas a las de los números naturales, y entre ellas se encuentran los teoremas de inducción y recursión que, debidamente generalizados, son válidos también para ordinales.

El principio de inducción es válido en general para toda clase bien ordenada en la forma siguiente: si Y es una clase bien ordenada y $X \subset Y$ tiene la propiedad de que si $\bigwedge u \in Y(\bigwedge v(v \in Y \wedge v < u \to v \in X) \to u \in X)$, entonces $X = Y$. Es decir, si podemos probar que un $u \in Y$ está en X bajo la hipótesis de inducción de que todos los elementos anteriores a u están en X, entonces podemos asegurar que X contiene a todos los elementos de Y. La razón es que si no fuera así entonces $Y \setminus X$ sería no vacío y tendría un mínimo elemento u. Ahora bien, entonces todos los elementos anteriores a u estarán en X y, de acuerdo con la hipótesis, tendría que ser $u \in X$, lo que nos da una contradicción.

Más concretamente, si tomamos una fórmula normal $\phi(x)$ y aplicamos esto a la clase $\{\alpha \in \Omega \mid \phi(\alpha)\}$ obtenemos el teorema siguiente:

Teorema 11.16 (Inducción transfinita) *Para toda fórmula normal $\phi(x)$ (quizá con más variables libres) la fórmula siguiente es un teorema de NBG*:*

$$\bigwedge\alpha(\bigwedge\beta(\beta<\alpha\to\phi(\beta))\to\phi(\alpha))\to\bigwedge\alpha\,\phi(\alpha).$$

En otras palabras: para probar que todo ordinal tiene una propiedad (normal) basta suponer que la cumplen todos los menores que uno dado α y probar que α también la cumple. A veces es más cómoda esta otra forma del principio de inducción:

Teorema 11.17 (Inducción transfinita) *Para toda fórmula normal $\phi(x)$ (quizá con más variables libres) la fórmula siguiente es un teorema de NBG*:*

$$\phi(0)\wedge\bigwedge\alpha(\phi(\alpha)\to\phi(\alpha'))\wedge\bigwedge\lambda(\bigwedge\delta(\delta<\lambda\to\phi(\delta))\to\phi(\lambda))\to\bigwedge\alpha\,\phi(\alpha).$$

Es decir, para probar que todo ordinal tiene una propiedad basta ver que la cumple el 0, que si la cumple un ordinal α la cumple el siguiente y que si la cumplen todos los ordinales menores que un límite λ entonces la cumple λ.

La demostración es esencialmente la misma: si existiera un ordinal α que no cumpliera $\phi(\alpha)$, podríamos tomar el mínimo de todos ellos. Por hipótesis no puede ser $\alpha=0$. Si fuera $\alpha=\beta'$, la minimalidad de α nos daría $\phi(\beta)$, luego la hipótesis nos daría también $\phi(\alpha)$, lo cual es absurdo. Finalmente, si α es un límite, la minimalidad nos da que todos los ordinales menores cumplen ϕ, luego de nuevo por hipótesis α cumpliría ϕ y, en cualquier caso, tendríamos una contradicción.

Ocupémonos ahora de la recursión transfinita. Lo que vamos a probar es que para definir una función $F:\Omega\longrightarrow \mathrm{V}$ es suficiente definir $F(\alpha)$ supuesto que F ya está definida sobre los ordinales menores que α, es decir, supuesta definida $F|_\alpha$. Más precisamente, las definiciones de la forma $F(\alpha)=G(F|_\alpha)$, a pesar de su aparente circularidad, son lícitas. Veámoslo.

Teorema 11.18 (Recursión transfinita) *Sea $G:\mathrm{V}\longrightarrow\mathrm{V}$ una función arbitraria. Entonces existe una única función $F:\Omega\longrightarrow\mathrm{V}$ caracterizada por que $\bigwedge\alpha\,F(\alpha)=G(F|_\alpha)$.*

Demostración: Diremos que $f:\beta\longrightarrow\mathrm{V}$ es una *β-aproximación* si para todo $\alpha<\beta$ se cumple $f(\alpha)=G(f|_\alpha)$. Es claro que si existe una β-aproximación entonces es única. En efecto, supongamos que f y g son dos β-aproximaciones. Entonces sea $\alpha<\beta$ el mínimo ordinal en el que difieran (si es que existe). Esto significa que $f|_\alpha=g|_\alpha$, pero que $f(\alpha)\neq g(\alpha)$. Ahora bien, esto es absurdo, pues $f(\alpha)=G(f|_\alpha)=G(g|_\alpha)=g(\alpha)$.

Ahora veamos por inducción que existen β-aproximaciones para todo β.

Es claro que $\varnothing$ es trivialmente una 0-aproximación. Si $f:\alpha\longrightarrow\mathrm{V}$ es una α-aproximación, entonces $g=f\cup\{(\alpha,G(f))\}$ es una α'-aproximación. En efecto, tenemos que $g:\alpha'\longrightarrow\mathrm{V}$ y si $\beta<\alpha'$, o bien $\beta<\alpha$, en cuyo caso $g(\beta)=f(\beta)=G(f|_\beta)=G(g|_\beta)$, o bien $\beta=\alpha$, en cuyo caso $g(\beta)=G(f)=G(g|_\beta)$.

Finalmente, supongamos que existen δ-aproximaciones para todo $\delta < \lambda$ y veamos que existe una λ-aproximación.

Podemos definir $f_\delta \equiv f|f$ es una δ-aproximación, y así[5] $\bigwedge\delta(\delta < \lambda \to f_\delta$ es una δ-aproximación).

De la definición se sigue inmediatamente que si $\delta < \epsilon < \lambda$ entonces $f_\epsilon|_\delta$ es una δ-aproximación, luego la unicidad implica que $f_\epsilon|_\delta = f_\delta$. De aquí se sigue que

$$f = \bigcup_{\delta<\lambda} f_\delta : \lambda \longrightarrow \mathrm{V},$$

y f es una λ-aproximación, pues si $\delta < \lambda$ entonces $f(\delta) = f_{\delta'}(\delta) = G(f_{\delta'}|_\delta) = F(f|_\delta)$.

Con esto hemos probado que existen α-aproximaciones para todo ordinal α. Por el mismo argumento que en el caso límite de la inducción podemos definir $f_\alpha : \alpha \longrightarrow \mathrm{V}$ como la única α aproximación y, de nuevo, la unicidad nos da que si $\alpha < \beta$ entonces $f_\beta|_\alpha = f_\alpha$, lo cual nos permite definir

$$F = \bigcup_{\alpha\in\Omega} f_\alpha : \Omega \longrightarrow \mathrm{V}.$$

Claramente F cumple lo pedido, y el mismo argumento que probaba la unicidad de las aproximaciones prueba que F es única. ∎

En la práctica no es necesario describir explícitamente la función G que determina la recurrencia. Es suficiente con determinar $F(\alpha)$ a partir de los valores que toma F sobre los ordinales menores que α. Veamos un caso particular de especial interés:

Teorema 11.19 *Sea $\beta \in \Omega$ y $H : \Omega \longrightarrow \Omega$. Entonces existe una única aplicación $F : \Omega \longrightarrow \Omega$ caracterizada por:*

$$F(0) = \beta \quad \wedge \quad \bigwedge\alpha\, F(\alpha') = H(F(\alpha)) \quad \wedge \quad \bigwedge\lambda\, F(\lambda) = \bigcup_{\delta<\lambda} F(\delta).$$

Demostración: La existencia de una función $F : \Omega \longrightarrow \mathrm{V}$ que cumpla estas propiedades (aunque $H(F(\alpha))$ pudiera ser una descripción impropia) es consecuencia inmediata del teorema anterior. Por una vez vamos a explicitar G aunque, como ya hemos dicho, está de más:

$$G(f) = \begin{cases} \beta & \text{si } f = \varnothing, \\ H(f(\alpha)) & \text{si } f \text{ es una función de dominio } \alpha', \\ \bigcup\limits_{\delta<\lambda} f(\delta) & \text{si } f \text{ es una función de dominio } \lambda, \\ \varnothing & \text{en cualquier otro caso.} \end{cases}$$

Una simple inducción prueba que $\bigwedge\alpha\, F(\alpha) \in \Omega$ así como que si F_1 y F_2 cumplen el teorema, entonces $\bigwedge\alpha\, F_1(\alpha) = F_2(\alpha)$. ∎

[5]Los suspicaces que vean en todas partes el axioma de elección deberían esforzarse en entender que aquí no lo estamos usando gracias a la unicidad: no hay ninguna elección $\delta \mapsto f_\delta$, sino que f_δ *se define* como la única δ-aproximación. Técnicamente usamos la regla de las descripciones propias.

Como aplicación de los teoremas de inducción y recursión vamos a probar que los ordinales representan todas las formas posibles de ordenar bien un conjunto. Para ello necesitamos la noción de semejanza:

Definición 11.20 Diremos que $F : (X, \leq_1) \longrightarrow (Y, \leq_2)$ es una *semejanza* entre entre dos clases ordenadas si $F : X \longrightarrow Y$ biyectiva y

$$\bigwedge uv \in X(u \leq_1 v \leftrightarrow F(u) \leq_2 F(v)).$$

Es fácil ver que si X está totalmente ordenada podemos sustituir la doble implicación por la implicación $\rightarrow$. La inversa de una semejanza es una semejanza y la composición de semejanzas es una semejanza. Dos clases ordenadas son *semejantes* si existe una semejanza entre ellas. Lo representaremos con la notación[6] $(X, \leq_1) \cong (Y, \leq_2)$.

Teorema 11.21 *Todo conjunto bien ordenado es semejante a un ordinal.*

Demostración: Sea $(x, \leq)$ un conjunto bien ordenado. Podemos suponer que no es vacío. Sea m su mínimo. Definimos $F : \Omega \longrightarrow x$ como la única aplicación que cumple

$$F(\alpha) = \begin{cases} \text{mín}(x \setminus F[\alpha]) & \text{si } F[\alpha] \neq x, \\ m & \text{en caso contrario.} \end{cases}$$

La aplicación F no puede ser inyectiva, pues en tal caso x sería una clase propia. Por consiguiente, existen ordinales $\beta < \alpha$ tales que $F(\beta) = F(\alpha)$. Podemos tomar el mínimo ordinal α para el cual existe un $\beta < \alpha$ con la misma imagen. De este modo, $f = F|_\alpha : \alpha \longrightarrow x$ inyectiva.

Además f es suprayectiva, ya que si $F[\alpha] \neq x$ sería $F(\alpha) \in x \setminus F[\alpha]$, cuando estamos suponiendo que $F(\alpha) = F(\beta) \in F[\alpha]$. Así pues, f es biyectiva.

Para probar que es una semejanza basta ver que para todo $\gamma < \alpha$ se cumple que $f[\gamma] = \{u \in x \mid u < f(\gamma)\}$. Lo probamos por inducción. Supongamos que se cumple para todo $\delta < \gamma$. Entonces, si $u < f(\gamma)$, por definición de f ha de ser $u \in f[\gamma]$. Recíprocamente, si $u \in f[\gamma]$, entonces $u = f(\delta)$, para un $\delta < \gamma$. Todo $v < u$ cumple $v < f(\delta)$ luego, por hipótesis de inducción, $v \in f[\delta] \subset f[\gamma]$. Vemos, pues, que todo $v \leq u$ cumple $v \in f[\gamma]$ y, como $f(\gamma) \notin f[\gamma]$, ha de ser $u < f(\gamma)$.

Ciertamente entonces, si $\delta < \gamma < \alpha$ tenemos que $f(\delta) \in f[\gamma]$, luego se cumple $f(\delta) < f(\gamma)$ y así f es una semejanza. ∎

En realidad todo conjunto bien ordenado es semejante a un único ordinal, pero la unicidad se basa en un hecho general que conviene enunciar aisladamente:

Teorema 11.22 *Toda aplicación estrictamente creciente $f : \alpha \longrightarrow \beta$ entre dos ordinales cumple $\bigwedge \delta \in \alpha\ \delta \leq f(\delta)$. En particular si existe tal f ha de ser $\alpha \leq \beta$.*

[6] Si X e Y son clases propias, no debemos (no podemos) entender que $(X, \leq_1)$ representa a un par ordenado con primera componente X, pues X no puede formar parte de ningún par ordenado. Simplemente hemos de entender $(X, \leq_1) \cong (Y, \leq_2)$ como una abreviatura de la fórmula "$\leq_1$ es un orden en X, $\leq_2$ es un orden en Y y existe una semejanza entre X e Y con dichos órdenes".

DEMOSTRACIÓN: En caso contrario, sea $\delta < \alpha$ el mínimo ordinal tal que $f(\delta) < \delta$. Entonces $f(\delta) < \alpha$, luego podemos calcular $f(f(\delta)) < f(\delta) < \delta$, con lo que $f(\delta)$ contradice la minimalidad de δ.

Si fuera $\beta < \alpha$ tendríamos necesariamente que $f(\beta) < \beta$. ∎

Teorema 11.23 *Todo conjunto bien ordenado es semejante a un único ordinal.*

DEMOSTRACIÓN: Sólo falta la unicidad. Si un mismo conjunto fuera semejante a dos ordinales $\alpha > \beta$, éstos serían semejantes entre sí, con lo que tendríamos una semejanza $f : \alpha \longrightarrow \beta$ en contradicción con el teorema anterior. ∎

Definición 11.24 Llamaremos *ordinal* de un conjunto bien ordenado $(x, \leq)$ al único ordinal semejante a x. Lo representaremos por $\operatorname{ord}(x, \leq)$.

Así pues, los ordinales son, como pretendíamos, representantes distintos de todas las formas posibles de ordenar bien un conjunto. Observemos que también tenemos unicidad en la semejanza:

Teorema 11.25 *Si dos conjuntos bien ordenados son semejantes, entonces existe una única semejanza entre ellos.*

DEMOSTRACIÓN: Basta probar que la única semejanza de un conjunto bien ordenado en sí mismo es la identidad (ya que entonces, si f y g son dos semejanzas entre dos conjuntos bien ordenados, $f \circ g^{-1}$ ha de ser la identidad, luego $f = g$). El argumento es el mismo que el que hemos usado en 11.22:

Sea $f : x \longrightarrow x$ una semejanza. Para todo $u \in x$, ha de ser $u \leq f(u)$, pues si fuera $f(u) < u$, tomando el mínimo u con esta propiedad tendríamos una contradicción, ya que $f(f(u)) < f(u) < u$ nos da que $f(u)$ cumple lo mismo y es menor.

Aplicando esto a f^{-1} obtenemos la desigualdad contraria, luego $f(u) = u$ para todo $u \in x$. ∎

Una consecuencia útil de 11.22 es la siguiente:

Teorema 11.26 *Si y es un conjunto bien ordenado y $x \subset y$, entonces se cumple que* $\operatorname{ord} x \leq \operatorname{ord} y$.

DEMOSTRACIÓN: Sean $\alpha = \operatorname{ord} x$, $\beta = \operatorname{ord} y$ y consideremos las semejanzas $f : x \longrightarrow \alpha$ y $g : y \longrightarrow \beta$. Si fuera $\beta < \alpha$ tendríamos que $f^{-1} \circ g : \alpha \longrightarrow \beta$ sería estrictamente creciente, en contradicción con 11.22. ∎

Si una clase propia bien ordenada es semejante a un ordinal, ha de ser semejante a Ω, pero esto no tiene por qué ser cierto. El teorema siguiente nos da una condición necesaria y suficiente para que así sea.

Teorema 11.27 *Una clase propia X bien ordenada es semejante a Ω si y sólo si, para todo $u \in X$, la clase $X_u = \{v \in X \mid v \leq u\}$ es un conjunto.*

DEMOSTRACIÓN: La condición es claramente necesaria: si existe una semejanza $F : X \longrightarrow \Omega$ y $F(u) = \alpha$, entonces X_u se corresponde con α' a través de F, luego ha de ser un conjunto.

Si se cumple la condición, para cada $u \in X$ tenemos que X_u es un conjunto bien ordenado, luego podemos considerar su ordinal α_u. Sea $f_u : X_u \longrightarrow \alpha_u$ la (única) semejanza entre ellos.

Si $u < v$ es claro que f_v transforma X_u en el conjunto de todos los ordinales menores o iguales que que $f_v(u)$, es decir, en el ordinal $f_v(u)'$. Así pues, tenemos que $f_v|_{X_u} : X_u \longrightarrow f_v(u)'$ es una semejanza y, por la unicidad, $\alpha_u = f_v(u)'$ y $f_v|_{X_u} = f_u$. Esto nos permite definir

$$F = \bigcup_{u\in X} f_u : X \longrightarrow \Omega.$$

Es claro que F es una semejanza. ∎

Así pues, una clase propia bien ordenada es semejante a Ω si y sólo si sus secciones iniciales son conjuntos. Por ejemplo, toda subclase de Ω que no sea un conjunto es semejante a Ω.

Ejemplo Consideremos en $\Omega \times \Omega$ el orden lexicográfico, es decir, el dado por

$$(\alpha, \beta) \leq (\gamma, \delta) \leftrightarrow \beta < \delta \vee (\beta = \delta \wedge \alpha \leq \gamma).$$

Es fácil ver que $\Omega \times \Omega$ es una clase bien ordenada, pero no es semejante a Ω, ya que todos los pares $(\alpha, 0)$ son menores que el par $(0, 1)$, es decir, tenemos una aplicación inyectiva $\Omega \longrightarrow (\Omega \times \Omega)_{(0,1)}$, luego esta sección no es un conjunto. ∎

Definición 11.28 Definimos el *orden canónico* en $\Omega \times \Omega$ como el orden dado por

$$(\alpha, \beta) \leq (\gamma, \delta) \leftrightarrow \text{máx}\{\alpha, \beta\} < \text{máx}\{\gamma, \delta\} \vee$$

$$(\text{máx}\{\alpha, \beta\} = \text{máx}\{\gamma, \delta\} \wedge (\beta < \delta \vee (\beta = \delta \wedge \alpha \leq \gamma))).$$

Es decir, para comparar dos pares, primero comparamos sus máximas componentes, en caso de empate comparamos las de la derecha y en caso de empate comparamos las de la izquierda. Es fácil comprobar que es un buen orden, y sus secciones iniciales son segmentos, ya que la clase de pares menores que (γ, δ) está contenida en el conjunto $\text{máx}\{\gamma', \delta'\} \times \text{máx}\{\gamma', \delta'\}$. Por el teorema anterior existe una (única) semejanza $F : \Omega \times \Omega \longrightarrow \Omega$.

Ejercicio: (Suponiendo AI) probar que $F[\omega \times \omega] = \omega$.

11.3 Funciones normales

Los teoremas de la sección anterior nos permiten introducir y estudiar la aritmética ordinal, pero antes vamos a estudiar una familia de funciones de ordinales que aparecen en éste y otros contextos.

Definición 11.29 Sea Λ un ordinal límite o bien $\Lambda = \Omega$. Diremos que una función $F : \Lambda \longrightarrow \Omega$ es *normal* si

$$\bigwedge\alpha \in \Lambda\, F(\alpha) < F(\alpha') \wedge \bigwedge\lambda \in \Lambda\, F(\lambda) = \bigcup_{\delta<\lambda} F(\delta).$$

Por ejemplo, si aplicamos el teorema 11.19 a una función H que cumpla la propiedad $\bigwedge\alpha\, \alpha < H(\alpha)$, entonces la función F que obtenemos es normal.

La normalidad es fácil de comprobar y tiene varias consecuencias útiles:

Teorema 11.30 *Toda función normal F es estrictamente monótona, es decir, si $\alpha < \beta$ entonces $F(\alpha) < F(\beta)$. En particular F es inyectiva.*

Sea Λ el dominio de F. Fijado $\alpha \in \Lambda$, veamos que

$$\bigwedge\beta \in \Lambda(\alpha < \beta \to F(\alpha) < F(\beta))$$

por inducción sobre β. Para $\beta = 0$ es trivialmente cierto. Si vale para β y tenemos $\alpha < \beta'$, entonces $\alpha < \beta$ o $\alpha = \beta$. Por hipótesis de inducción en el primer caso y trivialmente en el segundo, $F(\alpha) \leq F(\beta)$ y como F es normal $F(\alpha) < F(\beta')$.

Si es cierto para todo $\delta < \lambda$ y $\alpha < \lambda \in \Lambda$, entonces $\alpha < \alpha' < \lambda$ y, por hipótesis de inducción $F(\alpha) < F(\alpha')$. De nuevo por la normalidad de F es $F(\alpha) < F(\lambda)$. ∎

En particular, las funciones normales cumplen el teorema 11.22, es decir, si $F : \Lambda \longrightarrow \Lambda$ es normal, entonces $\bigwedge\alpha \in \Lambda\, \alpha \leq F(\alpha)$.

Teorema 11.31 *Si $F : \Lambda \longrightarrow \Omega$ es una función normal y $\lambda \in \Lambda$, entonces $F(\lambda)$ es un ordinal límite.*

Demostración: Como $0 < \lambda$, es $0 \leq F(0) < F(\lambda)$, luego $F(\lambda) \neq 0$. Si $\alpha < F(\lambda)$, por la normalidad $\alpha < F(\delta)$, para un cierto $\delta < \lambda$. Entonces $\delta < \delta' < \lambda$, luego $\alpha' \leq F(\delta) < F(\delta') \leq F(\lambda)$. Así pues, $F(\lambda) \neq \alpha'$ para todo α. ∎

Teorema 11.32 *Si F, $G : \Lambda \longrightarrow \Lambda$ son funciones normales, entonces $F \circ G$ también lo es.*

Demostración: Claramente, si $\alpha \in \Lambda$ tenemos que $F(\alpha) < F(\alpha')$, luego $G(F(\alpha)) < G(F(\alpha'))$. Tomemos ahora un ordinal límite $\lambda \in \Lambda$. Hemos de probar que

$$G(F(\lambda)) = \bigcup_{\delta<\lambda} G(F(\delta)).$$

Si $\alpha \in G(F(\lambda))$, como $F(\lambda)$ es un ordinal límite tenemos que $\alpha < G(\eta)$, para un $\eta \in F(\lambda)$. A su vez, $\eta \in F(\delta)$ con $\delta < \lambda$. En total $\alpha < G(\eta) < G(F(\delta))$, luego α está en el miembro derecho de la igualdad.

Recíprocamente, si $\alpha \in G(F(\delta))$, con $\delta < \lambda$, entonces $F(\delta) < F(\lambda)$, luego $\alpha < G(F(\delta)) < G(F(\lambda))$. ∎

11.4 La aritmética ordinal

Vamos a definir una suma, un producto y una exponenciación entre ordinales que generalizan a las operaciones análogas sobre los números naturales. Estas operaciones resultan útiles para definir ordinales y aplicaciones entre ordinales.

Suma de ordinales Si A y B son dos conjuntos ordenados, podemos definir su suma como el conjunto $A \oplus B = A \times \{0\} \cup B \times \{1\}$ con el orden dado por $(u, v) < (w, x) \leftrightarrow v < x \vee (v = x \wedge u \leq w)$.

En definitiva, $A \oplus B$ consta de un primer tramo semejante a A seguido de un segundo tramo semejante a B. Es fácil ver que la suma de conjuntos bien ordenados está bien ordenada. Podríamos definir la suma $\alpha + \beta$ de dos ordinales como el ordinal de la suma $\alpha \oplus \beta$, es decir, el ordinal que representa el orden que empieza como α y termina como β. Por ejemplo, $\omega + 1$ es el ordinal del conjunto

$$(0,0) < (1,0) < (2,0) < \cdots (0,1).$$

Es claro que este conjunto es semejante a $\omega' = \{0, 1, 2, \ldots, \omega\}$. Así pues, $\omega + 1 = \omega'$. En cambio, $1 + \omega$ es el ordinal del conjunto

$$(0,0) < (0,1) < (1,1) < (2,1) < \cdots$$

y es claro entonces que $1 + \omega = \omega$. Así pues, $\omega + 1 \neq 1 + \omega$, luego vemos que la suma de ordinales no es conmutativa. En otras palabras, lo que sucede es que si añadimos un elemento a la sucesión de los números naturales por la izquierda "no se nota", pero si lo añadimos por la derecha sí.

Por comodidad vamos a introducir la suma con una definición recurrente más manejable. De todos modos, cuando contemos con las propiedades básicas será fácil ver que se trata de la misma operación que acabamos de considerar.

Definición 11.33 Para cada ordinal $\alpha \in \Omega$ definimos $\alpha+ : \Omega \longrightarrow \Omega$ como la única aplicación que cumple

$$\alpha + 0 = \alpha \quad \wedge \quad \bigwedge\beta\, \alpha + \beta' = (\alpha + \beta)' \quad \wedge \quad \bigwedge\lambda\, \alpha + \lambda = \bigcup_{\delta<\lambda} (\alpha + \delta).$$

Naturalmente, esta definición es correcta[7] por el teorema 11.19, y es claro que $\alpha+$ es una aplicación normal. Esto nos da ya algunas propiedades de la

[7] Observemos además que $\alpha + \beta$ es un término normal. En efecto, $u \in \alpha + \beta$ es equivalente a $\alpha \in \Omega \wedge \beta \in \Omega \wedge \bigvee\gamma \in \Omega \bigvee f(f : \gamma \longrightarrow \Omega \wedge \beta < \gamma \wedge f(0) = \alpha \wedge \ldots \wedge u \in f(\beta)))$, donde los puntos suspensivos son las propiedades que definen la $(\alpha+)$. El mismo argumento justifica la normalidad del producto y la exponenciación que introduciremos después.

suma, como la monotonía:

$$\bigwedge\alpha\beta\gamma(\beta<\gamma\to\alpha+\beta<\alpha+\gamma),\qquad \bigwedge\alpha\beta\,\beta\le\alpha+\beta.$$

o el hecho de que los ordinales $\alpha+\lambda$ son ordinales límite. Si definimos $1=0'$, la definición de suma nos da que $\alpha+1=\alpha+0'=(\alpha+0)'=\alpha'$. Por ello, a partir de ahora ya no volveremos a usar la notación α', sino que escribiremos siempre $\alpha+1$. Notemos que la segunda propiedad de la definición de la suma se escribe ahora $\alpha+(\beta+1)=(\alpha+\beta)+1$.

Todas las propiedades de la suma se demuestran por inducción. Por ejemplo, es inmediato comprobar que $\bigwedge\alpha\ 0+\alpha=\alpha$. Veamos un ejemplo detallado:

Teorema 11.34 $\bigwedge\alpha\beta\gamma(\alpha\le\beta\to\alpha+\gamma\le\beta+\gamma)$.

Demostración: Lo probamos por inducción sobre γ. Para $\gamma=0$ es obvio. Si vale para γ, entonces

$$\alpha+(\gamma+1)=(\alpha+\gamma)+1\le(\beta+\gamma)+1=\beta+(\gamma+1).$$

Si es cierto para todo $\delta<\lambda$, entonces $\alpha+\delta\le\beta+\delta\le\beta+\lambda$ y, tomando el supremo en δ, $\alpha+\lambda\le\beta+\lambda$. ∎

De las desigualdades que hemos probado se sigue sin dificultad (sin necesidad de más inducciones) el siguiente resultado general de monotonía:

$$\bigwedge\alpha\beta\gamma\delta(\alpha\le\beta\wedge\gamma<\delta\to\alpha+\gamma<\beta+\delta),$$

del cual se sigue, obviamente, el caso en que todas las desigualdades son no estrictas.

Una simple inducción demuestra que la suma de números naturales es un número natural. Suponiendo el axioma de infinitud (para que tenga sentido operar con ω) vemos que si $n\in\omega$ entonces

$$\omega\le n+\omega=\bigcup_{m\in\omega}n+m\le\omega,$$

luego $\bigwedge n\in\omega\,n+\omega=\omega$, como ya habíamos anticipado.

Pasemos ahora a las propiedades algebraicas de la suma. Como las funciones $\alpha+$ son normales —luego inyectivas— los sumandos son simplificables por la izquierda:

$$\bigwedge\alpha\beta\gamma(\alpha+\beta=\alpha+\gamma\to\beta=\gamma).$$

En cambio (suponiendo AI), tenemos que $5+\omega=8+\omega$ y no podemos simplificar. El teorema siguiente ilustra el uso de la normalidad en el caso límite de una inducción:

Teorema 11.35 $\bigwedge\alpha\beta\gamma((\alpha+\beta)+\gamma=\alpha+(\beta+\gamma))$.

DEMOSTRACIÓN: Por inducción sobre γ. Para $\gamma = 0$ es trivial. Si vale para γ, entonces

$$(\alpha+\beta)+(\gamma+1) = ((\alpha+\beta)+\gamma)+1 = (\alpha+(\beta+\gamma))+1$$
$$= \alpha+((\beta+\gamma)+1) = \alpha+(\beta+(\gamma+1)).$$

Si vale para todo $\delta < \lambda$, entonces

$$(\alpha+\beta)+\lambda = \bigcup_{\delta<\lambda}(\alpha+\beta)+\delta = \bigcup_{\delta<\lambda}\alpha+(\beta+\delta)$$
$$= \bigcup_{\delta<\lambda}((\beta+)\circ(\alpha+))(\delta) = ((\beta+)\circ(\alpha+))(\lambda) = \alpha+(\beta+\lambda),$$

donde en el penúltimo paso hemos usado la normalidad de la composición de las dos sumas. ∎

Como último resultado general sobre la suma probamos lo siguiente:

Teorema 11.36 $\bigwedge\alpha\beta(\alpha \le \beta \to \overset{1}{\bigvee}\gamma\, \alpha+\gamma = \beta)$.

DEMOSTRACIÓN: Sabemos que $\beta \le \alpha+\beta < \alpha+\beta+1$, luego podemos tomar el mínimo ordinal η tal que $\beta < \alpha+\eta$. Obviamente no puede ser $\eta = 0$ y si η fuera un límite existiría $\delta < \eta$ tal que $\beta < \alpha+\delta$, en contra de la minimalidad de η. Así pues, $\eta = \gamma+1$ para cierto γ tal que $\alpha+\gamma \le \beta < \alpha+\gamma+1$. Claramente $\beta = \alpha+\gamma$. La unicidad se sigue de que $\alpha+$ es normal. ∎

Ejercicio: Probar que, tal y como explicábamos al principio de este apartado, se cumple $\alpha+\beta = \mathrm{ord}(\alpha \oplus \beta)$.

Producto de ordinales Aunque vamos a definir el producto mediante una relación recurrente análoga a la de la suma, también en este caso podríamos dar una definición en términos de buenos órdenes. Concretamente, si A y B son dos conjuntos ordenados, podemos considerar $A \times B$ con el *orden lexicográfico,* es decir, el orden dado por

$$(u,v) \le (w,x) \leftrightarrow v < x \vee (v = x \wedge u \le w).$$

Es fácil ver que el producto de dos conjuntos bien ordenados está bien ordenado, lo que nos permitiría definir $\alpha \cdot \beta = \mathrm{ord}(\alpha \times \beta)$. Por ejemplo, $\omega \cdot 2$ sería el ordinal de

$$(0,0) < (1,0) < (2,0) < \cdots (0,1) < (1,1) < (2,1) < \cdots$$

y es claro que este conjunto es semejante a

$$\omega+\omega = \{0,1,2,\ldots,\omega,\omega+1,\omega+2,\ldots\}.$$

En cambio, $2 \cdot \omega$ es el ordinal de

$$(0,0) < (1,0) < (0,1) < (1,1) < (0,2) < (1,2) < \cdots$$

por lo que $2 \cdot \omega = \omega$.

Definición 11.37 Para cada ordinal $\alpha \in \Omega$ definimos $\alpha\cdot : \Omega \longrightarrow \Omega$ como la única aplicación que cumple

$$\alpha \cdot 0 = 0 \quad \wedge \quad \bigwedge\beta\ \alpha \cdot (\beta + 1) = (\alpha \cdot \beta) + \alpha \quad \wedge \quad \bigwedge\lambda\ \alpha \cdot \lambda = \bigcup_{\delta<\lambda} (\alpha \cdot \delta).$$

Es claro que si $\alpha \neq 0$ entonces $\alpha\cdot$ es una función normal, mientras que una simple inducción prueba que $\bigwedge\alpha\ 0 \cdot \alpha = 0$. Tampoco ofrece dificultad alguna demostrar que $\bigwedge\alpha(\alpha \cdot 1 = 1 \cdot \alpha = \alpha)$.

Como consecuencia inmediata de la normalidad tenemos la monotonía:

$$\bigwedge\alpha\beta\gamma(\alpha < \beta \wedge \gamma \neq 0 \to \gamma \cdot \alpha < \gamma \cdot \beta)$$

Si multiplicamos por la derecha la desigualdad ha de ser no estricta:

$$\bigwedge\alpha\beta\gamma(\alpha \leq \beta \to \alpha \cdot \gamma \leq \beta \cdot \gamma).$$

Esto se prueba exactamente igual que el resultado análogo para la suma. Combinando estas desigualdades tenemos:

$$\bigwedge\alpha\beta\gamma\delta(\alpha \leq \beta \wedge \gamma < \delta \wedge \beta \neq 0 \to \alpha \cdot \gamma < \beta \cdot \delta).$$

De aquí se sigue, en particular, que $\bigwedge\alpha\beta(\alpha \cdot \beta = 0 \leftrightarrow \alpha = 0 \vee \beta = 0)$. También es claro que los factores no nulos se simplifican por la izquierda en las igualdades (por normalidad).

Ejercicio: (Usando el axioma de infinitud) probar que $\bigwedge n \in \omega(n \neq 0 \to n\omega = \omega)$.

Veamos ahora dos propiedades algebraicas:

Teorema 11.38 $\bigwedge\alpha\beta\gamma\ \alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$.

Demostración: Podemos suponer $\alpha \neq 0$. Lo probamos por inducción sobre γ. Para $\gamma = 0$ es obvio. Si vale para γ, entonces

$$\alpha(\beta + \gamma + 1) = \alpha(\beta + \gamma) + \alpha = \alpha\beta + \alpha\gamma + \alpha = \alpha\beta + \alpha(\gamma + 1).$$

Si vale para todo $\delta < \lambda$, entonces, usando que la composición de funciones normales es normal,

$$\alpha(\beta + \lambda) = ((\beta+) \circ (\alpha\cdot))(\lambda) = \bigcup_{\delta<\lambda} ((\beta+) \circ (\alpha\cdot))(\delta) = \bigcup_{\delta<\lambda} (\alpha(\beta + \delta))$$

$$= \bigcup_{\delta<\lambda} (\alpha\beta + \alpha\delta) = \bigcup_{\delta<\lambda} ((\alpha\cdot) \circ (\alpha\beta+))(\delta) = ((\alpha\cdot) \circ (\alpha\beta+))(\lambda) = \alpha\beta + \alpha\lambda.$$

■

Exactamente igual se demuestra:

Teorema 11.39 $\bigwedge\alpha\beta\gamma\ (\alpha\beta)\gamma = \alpha(\beta\gamma)$.

Por último resulta que la división euclídea es válida para ordinales cualesquiera:

Teorema 11.40 $\bigwedge\alpha\beta(\beta \neq 0 \to \bigvee^1\gamma\delta(\alpha = \beta\gamma + \delta \wedge \delta < \beta))$.

DEMOSTRACIÓN: Como $1 \leq \beta$, tenemos que $\alpha \leq \beta\alpha < \beta\alpha + \beta = \beta(\alpha + 1)$. Sea η el mínimo ordinal tal que $\alpha < \beta\eta$. Obviamente no puede ser $\eta = 0$ y tampoco puede ser un ordinal límite, ya que entonces sería $\alpha < \beta\epsilon$, para $\epsilon < \eta$, en contra de la minimalidad de η. Así pues, $\eta = \gamma + 1$, para cierto γ. Tenemos que

$$\beta\gamma \leq \alpha < \beta(\gamma + 1) = \beta\gamma + \beta.$$

Por el teorema 11.36 existe un δ tal que $\alpha = \beta\gamma + \delta$. Como $\beta\gamma + \delta < \beta\gamma + \beta$, por la normalidad de $\beta\gamma+$ se sigue que $\delta < \beta$.

Veamos la unicidad. Si tenemos dos soluciones γ_1, γ_2, δ_1, δ_2 y $\gamma_1 < \gamma_2$, entonces

$$\alpha = \beta\gamma_1 + \delta_1 < \beta\gamma_1 + \beta = \beta(\gamma_1 + 1) \leq \beta\gamma_2 \leq \beta\gamma_2 + \delta_2 = \alpha,$$

lo cual es contradictorio. Similarmente es imposible que $\gamma_2 < \gamma_1$, luego $\gamma_1 = \gamma_2$. Por consiguiente, $\beta\gamma_1 + \delta_1 = \beta\gamma_1 + \delta_2$, de donde $\delta_1 = \delta_2$. ∎

Ejercicio: Probar que la aplicación $\beta \times \eta \longrightarrow \beta \cdot \eta$ dada por $(\delta, \gamma) \mapsto \beta\gamma + \delta$ es biyectiva, así como que es una semejanza respecto al orden lexicográfico en $\beta \times \eta$. Por lo tanto $\beta\alpha = \mathrm{ord}(\beta \times \eta)$.

Exponenciación de ordinales Describir la exponenciación en términos de conjuntos ordenados no es tan sencillo como en el caso de la suma y el producto, así que nos limitaremos a dar la definición recurrente:

Definición 11.41 Para cada ordinal $\alpha \neq 0$ definimos $\alpha^{(\ )} : \Omega \longrightarrow \Omega$ como la única función que cumple

$$\alpha^0 = 1 \wedge \bigwedge\beta\ \alpha^{\beta+1} = \alpha^\beta \cdot \alpha \wedge \bigwedge\lambda\ \alpha^\lambda = \bigcup_{\delta<\lambda} \alpha^\delta.$$

Convenimos en que $0^\alpha = \begin{cases} 1 & \text{si } \alpha = 0 \\ 0 & \text{en otro caso.} \end{cases}$

Una simple inducción nos da que $\bigwedge\alpha\beta(\alpha \neq 0 \to \alpha^\beta \neq 0)$, de donde se sigue que si $\alpha > 1$ entonces $\alpha^{(\ )}$ es una función normal.

Omitimos las demostraciones de las propiedades siguientes, pues todas ellas son similares a los resultados análogos para la suma y el producto (a menudo hay que tratar aparte los casos en los que la base es 0 o 1).

a) $\bigwedge\alpha\ 1^\alpha = 1$,

b) $\bigwedge\alpha\ \alpha^1 = \alpha$,

c) $\bigwedge\alpha\beta\gamma(\alpha < \beta \wedge 1 < \gamma \to \gamma^\alpha < \gamma^\beta)$,

d) $\bigwedge\alpha\beta\gamma(\alpha \leq \beta \to \alpha^\gamma \leq \beta^\gamma)$,

e) $\bigwedge\alpha\beta\gamma(\alpha \leq \beta \wedge 1 < \gamma \wedge \gamma^\alpha = \gamma^\beta \to \alpha = \beta)$,

f) $\bigwedge\alpha\beta\gamma\ \alpha^{\beta+\gamma} = \alpha^\beta \cdot \alpha^\gamma$,

g) $\bigwedge\alpha\beta\gamma\ (\alpha^\beta)^\gamma = \alpha^{\beta\cdot\gamma}$.

Ejercicio: (Usando el axioma de infinitud) probar que $\bigwedge n \in \omega(1 < n \to n^\omega = \omega)$.

La no conmutatividad del producto hace que, en general, $(\alpha\beta)^\gamma \neq \alpha^\gamma\beta^\gamma$. Por ejemplo,

$$(2 \cdot 2)^\omega = \omega \neq \omega^2 = 2^\omega \cdot 2^\omega.$$

Aritmética natural En particular, tenemos definida una suma y un producto sobre los números naturales que cumplen (por definición) los axiomas de teoría aritmética (sin necesidad del axioma de infinitud). Por consiguiente hemos probado que NBG* (luego también ZF*) es una teoría aritmética. La aritmética natural cumple algunas propiedades que no son ciertas para ordinales arbitrarios, como la conmutatividad de la suma y el producto. No vamos a probarlo aquí porque ya lo hemos hecho en el capítulo VI para teorías aritméticas arbitrarias. En todo caso, faltaría probar algunos hechos elementales sobre la exponenciación natural, tales como que $\bigwedge mnr \in \omega\ (mn)^r = m^r n^r$, etc., pero son comprobaciones sencillas que dejamos al lector.

11.5 La forma normal de Cantor

Para familiarizarnos con la aritmética ordinal que acabamos de introducir probaremos un resultado de Cantor que requiere hacer un uso sistemático de las propiedades que hemos establecido. Trabajamos en NBG$^-$, es decir, NBG*+AI. Demostramos varios resultados previos.

Teorema 11.42 *Si $\alpha\omega \leq \beta$ entonces $\alpha + \beta = \beta$.*

Demostración: Intuitivamente, la hipótesis afirma que β empieza por infinitas copias de α, por lo que si añadimos una más "no se nota".

Sabemos que existe un γ tal que $\beta = \alpha\omega + \gamma$, por lo que

$$\alpha + \beta = \alpha + \alpha\omega + \gamma = \alpha(1 + \omega) + \gamma = \alpha\omega + \gamma = \beta.$$

■

Ejercicio: Probar el recíproco del teorema anterior.

Teorema 11.43 *Si $\alpha < \beta$ entonces $\omega^\alpha + \omega^\beta = \omega^\beta$.*

Demostración: Es un caso particular del teorema anterior, puesto que se cumple $\omega^\alpha\omega = \omega^{\alpha+1} \leq \omega^\beta$. ■

Teorema 11.44 *Si $\alpha \neq 0$ existen unos únicos η y β tales que $\alpha = \omega^\eta + \beta$, con $\beta < \alpha$. Además η es concretamente el único ordinal que cumple $\omega^\eta \leq \alpha < \omega^{\eta+1}$.*

DEMOSTRACIÓN: Como la función $\omega^{(\ )}$ es normal, $\alpha \leq \omega^\alpha < \omega^{\alpha+1}$, luego podemos tomar el mínimo γ tal que $\alpha < \omega^\gamma$. No puede ser $\gamma = 0$ ni tampoco que sea un límite, luego $\gamma = \eta + 1$ y tenemos $\omega^\eta \leq \alpha < \omega^{\eta+1}$.

Es claro que η es único. Existe un $\beta \leq \alpha$ tal que $\alpha = \omega^\eta + \beta$, pero ha de ser $\beta < \alpha$, pues si se da la igualdad

$$\alpha = \omega^\eta + \alpha = \omega^\eta + \omega^\eta + \alpha = \omega^\eta + \omega^\eta + \omega^\eta + \alpha = \cdots$$

y, en general, $\omega^\eta \cdot n \leq \alpha$, para todo $n \in \omega$. Por consiguiente, $\omega^\eta\omega = \omega^{\eta+1} \leq \alpha$, contradicción.

Recíprocamente, si $\alpha = \omega^\eta + \beta$ con $\beta < \alpha$, ha de ser $\omega^\eta \leq \alpha < \omega^{\eta+1}$ o, de lo contrario, por 11.42 tendríamos que $\alpha = \omega^\eta + \alpha = \omega^\eta + \beta$ y sería $\beta = \alpha$. De aquí se sigue la unicidad de η, que a su vez implica la de β. ∎

Teorema 11.45 *Si $\alpha \neq 0$ existe una única sucesión finita decreciente de ordinales $\eta_0 \geq \eta_1 \geq \cdots \geq \eta_n$ tal que $\alpha = \omega^{\eta_0} + \cdots + \omega^{\eta_n}$.*

DEMOSTRACIÓN: Aplicamos el teorema anterior repetidamente, con lo que expresamos $\alpha = \omega^{\eta_0} + \alpha_1$, con $\alpha_1 < \alpha$, luego $\alpha_1 = \omega^{\eta_1} + \alpha_2$, con $\alpha_2 < \alpha_1$, etc. Como no podemos tener una sucesión decreciente de ordinales (no tendría mínimo), algún $\alpha_n = 0$, lo que nos da la expresión buscada.

Si fuera $\eta_i < \eta_{i+1}$ para algún i, entonces

$$\alpha_i = \omega^{\eta_i} + \alpha_{i+1} = \omega^{\eta_i} + \omega^{\eta_{i+1}} + \alpha_{i+2} = \omega^{\eta_{i+1}} + \alpha_{i+2} = \alpha_{i+1},$$

contradicción.

Para probar la unicidad observamos que si $\alpha = \omega^{\eta_0} + \cdots + \omega^{\eta_n}$ y los exponentes son decrecientes, entonces

$$\alpha = \omega^{\eta_0} + \cdots + \omega^{\eta_n} \leq \omega^{\eta_0} + \cdots + \omega^{\eta_0} = \omega^{\eta_0} \cdot n < \omega^{\eta_0}\omega = \omega^{\eta_0+1},$$

es decir, $\omega^{\eta_0} \leq \alpha < \omega^{\eta_0+1}$, luego η_0 está unívocamente determinado por α. Si tuviéramos dos expresiones distintas, ambas tendrían el mismo primer término, luego podríamos cancelarlo y de aquí deduciríamos que tendrían el mismo segundo término, y así sucesivamente. En definitiva, ambas serían la misma. ∎

El teorema de cantor se sigue del que acabamos de probar sin más que agrupar los términos con el mismo exponente:

Teorema 11.46 (Forma normal de Cantor) *Si $\alpha \neq 0$ existe una única sucesión finita estrictamente decreciente de ordinales $\eta_0 > \eta_1 > \cdots > \eta_n$ y una única sucesión finita $k_0, \ldots, k_n$ de números naturales no nulos tal que $\alpha = \omega^{\eta_0}k_0 + \cdots + \omega^{\eta_n}k_n$.*

La forma normal de Cantor es especialmente descriptiva para ordinales pequeños. Por ejemplo, si $\alpha < \omega^\omega$ entonces es fácil ver que η_0 ha de ser un número natural, luego tenemos que los ordinales menores que ω^ω se expresan de forma única como polinomios en ω con coeficientes naturales.

Podemos ir algo más lejos, para lo cual conviene definir

$$\omega^{(0)} = \omega, \qquad \omega^{(n+1)} = \omega^{\omega^{(n)}}, \qquad \epsilon_0 = \bigcup_{n\in\omega} \omega^{(n)}.$$

Así, $\omega^{(1)} = \omega$, $\omega^{(2)} = \omega^\omega$, $\omega^{(3)} = \omega^{\omega^\omega}$, etc. y ϵ_0 es el supremo de esta sucesión.

Si $\delta < \epsilon_0$, entonces se cumple $\delta < \omega^{(n)}$ para cierto $n \in \omega$, luego tenemos que $\omega^\delta \le \omega^{\omega^{(n)}} = \omega^{(n+1)} \le \epsilon_0$. Tomando el supremo en δ concluimos que $\omega^{\epsilon_0} \le \epsilon_0$. El recíproco es obvio, luego $\omega^{\epsilon_0} = \epsilon_0$. Cantor llamó *números épsilon* a los números con esta propiedad.

Se cumple que ϵ_0 es el menor número épsilon. Más precisamente, para cada $\alpha < \epsilon_0$ no nulo, llamamos *rango* de α al único número natural n tal que $\omega^{(n)} \le \alpha < \omega^{(n+1)}$.

Es claro que entonces $\omega^{(n+1)} \le \omega^\alpha < \omega^{(n+2)}$, es decir, tenemos que

$$\text{rango } \omega^\alpha = 1 + \text{rango } \alpha.$$

En particular $\omega^\alpha \ne \alpha$, luego α no es un número ϵ.

Los números naturales (no nulos) son los ordinales de rango 0, los números entre ω y ω^ω son los ordinales de rango 1 (y son, como hemos visto, los polinomios en ω con coeficientes naturales).

Más en general, es inmediato comprobar que el rango de un $\alpha < \epsilon_0$ es una unidad mayor que el de su exponente director η_0, es decir, que todos los exponentes de α tienen rango inferior al de α. De aquí se sigue (por inducción) que todo ordinal $0 < \alpha < \epsilon_0$ se obtiene a partir de los números naturales mediante un número finito de sumas, productos por naturales y aplicaciones de la exponencial de base ω.

Esto permite considerar a estos ordinales como objetos finitistas, similares a las funciones recursivas. Los teoremas que hemos probado permiten calcular la forma normal de la suma o el producto de dos números ordinales cuya forma normal es conocida. En particular, el rango de la suma o el producto es el máximo de los rangos de los sumandos o factores.

Ahora bien, todo esto deja de ser válido para ordinales mayores. La forma normal de ϵ_0 es ω^{ϵ_0}, lo cual no dice mucho.

Capítulo XII

Relaciones bien fundadas

Recordemos el axioma de regularidad (V=R):

$$\bigwedge x(x \neq \varnothing \to \bigvee y \in x\, y \cap x = \varnothing).$$

Sin duda, es el más técnico de todos los axiomas de la teoría de conjuntos, aunque también es uno de los menos relevantes. No obstante, tiene una interpretación muy natural. Sin este axioma, nada impide que exista toda suerte de conjuntos "patológicos". Pensemos, por ejemplo, en un conjunto x que cumpliera $x = \{x\}$, o en un par de conjuntos x e y tales que $x = \{y\}$, $y = \{x\}$. Imaginemos un conjunto $x = \{x_1\}$, donde $x_1 = \{x_2\}$, y a su vez $x_2 = \{x_3\}$, y $x_3 = \{x_4\}$, y así hasta el infinito.

En la introducción comentábamos que el axioma de partes contradice sutilmente la concepción de un conjunto como una agrupación de elementos existentes de antemano, pero la existencia de un conjunto $x = \{x\}$, al igual que cualquiera de los otros ejemplos que acabamos de considerar, la contradice descaradamente. Sin embargo, mientras que admitir la existencia de $\mathcal{P}x$ es extremadamente útil, la existencia de los "monstruos" anteriores es prácticamente inútil, por lo que es conveniente un axioma que los descarte, y ése es el axioma de regularidad. De todos modos, en la mayoría de las ocasiones los matemáticos no necesitan descartar tales patologías, sino que les basta con no hablar de ellas, y ésta es la razón por la que decíamos que el axioma de regularidad no es muy relevante.

Según veremos enseguida, el axioma de regularidad está relacionado con un tipo de relaciones especialmente interesante —las relaciones bien fundadas— porque son la clase más general de relaciones sobre las que podemos razonar por inducción o definir objetos recurrentemente. Los resultados que obtendremos sobre este tipo de relaciones, además de generalizar a cuanto hemos visto sobre inducción y recursión sobre ordinales, se aplicarán al estudio de la relación de pertenencia (que resulta estar bien fundada sobre todos los conjuntos de interés para los matemáticos) y, por otro lado, también tienen gran interés en teoría de modelos pues, en muchos casos, las relaciones que interpretan a la pertenencia en los modelos de la teoría de conjuntos están bien fundadas.

En este capítulo trabajaremos en NBG^- o, equivalentemente en ZF^-, es decir, añadimos a la teoría básica el axioma de infinitud.

12.1 Conceptos básicos

En primer lugar introducimos las relaciones que nos proponemos estudiar:

Definición 12.1 Una relación R está *bien fundada* en una clase A si cumple:

$$\bigwedge X(X \subset A \wedge X \neq \varnothing \to \bigvee y \in X \bigwedge z \in X \neg z\, R\, y).$$

En estas condiciones decimos que y es un *R-minimal* de X.

Equivalentemente, una relación R está bien fundada en una clase A si y sólo si no podemos encontrar en A sucesiones infinitas decrecientes de la forma

$$x_2\, R\, x_1, \quad x_3\, R\, x_2, \quad x_4\, R\, x_3, \quad x_5\, R\, x_4, \quad \ldots$$

En efecto:

Teorema 12.2 *Si R es una relación en una clase A y existe una sucesión $\{x_n\}_{n\in\omega}$ de elementos de A (no necesariamente distintos) tal que*

$$\bigwedge n \in \omega\ x_{n+1}\, R\, x_n,$$

entonces R no está bien fundada sobre A.

Demostración: Basta observar que el conjunto $\{x_n \mid n \in \omega\}$ no tiene R-minimal. ∎

Ejercicio: Usar el axioma de elección para probar el recíproco del teorema anterior.

Por ejemplo, si un $x \in A$ cumple $x\, R\, x$, entonces R ya no está bien fundada en A, pues tomando $x_n = x$ para todo n tenemos una sucesión decreciente (alternativamente, porque el conjunto $\{x\}$ no tiene R-minimal).

Notemos que si R es la relación de pertenencia, un minimal de una clase X es un $y \in X$ tal que $y \cap X = \varnothing$. De este modo, $\in$ está bien fundada en una clase X si y sólo si X está bien fundada en el sentido de la definición 11.1. El axioma de regularidad afirma que todo conjunto tiene un minimal para $\in$, por lo que claramente equivale a que $\in$ está bien fundada en todo conjunto. Todos los ejemplos patológicos que hemos puesto al principio del capítulo dan lugar a conjuntos donde $\in$ no está bien fundada.

Necesitamos imponer una restricción técnica a las relaciones que consideremos:

Definición 12.3 Una relación R es *conjuntista* en una clase A si para todo $x \in A$ la clase de los anteriores de x

$$A_x^R = \{y \in A \mid y\, R\, x\}$$

es un conjunto.

Obviamente toda relación es conjuntista en todo conjunto. Esto es —como decíamos— una mera restricción técnica necesaria por las restricciones que la teoría impone al trabajo con clases propias. La relación $\in$ es conjuntista en cualquier clase, pues $A_x^\in = x \cap A$.

Observemos que ya nos hemos encontrado con esta restricción en una ocasión: en el capítulo anterior hemos demostrado que una clase propia bien ordenada es semejante a Ω si y sólo si su relación de orden es conjuntista.

Definición 12.4 Sea R una relación definida sobre una clase A. Diremos que una subclase $B \subset A$ es *R-A-transitiva* si

$$\bigwedge xy \in A(x\, R\, y \wedge y \in B \to x \in B).$$

Es decir, B es R-A-transitiva si cuando partimos de elementos de B y vamos tomando anteriores nunca salimos de B. Si R es la relación de pertenencia y $A = \mathrm{V}$, una clase B es transitiva en este sentido si y sólo si lo es en el sentido usual definido en 11.1.

Si R es una relación definida sobre una clase a A y x es un subconjunto de A, es claro que al considerar los anteriores de x y los anteriores de los anteriores, etc. obtenemos un conjunto R-A-transitivo. En realidad, para que la definición recurrente de este proceso sea correcta hemos de exigir que R sea conjuntista. Veámoslo con detalle:

Definición 12.5 Sea R una relación conjuntista en una clase A y $x \in A$. Definimos

$$(A_x^R)_0 = A_x^R, \qquad \bigwedge n \in \omega (A_x^R)_{n+1} = \bigcup_{y \in (A_x^R)_n} A_y^R.$$

Un simple razonamiento inductivo prueba que cada $(A_x^R)_n$ es un conjunto (si vale para n, entonces $(A_x^R)_{n+1}$ es unión de un conjunto de conjuntos). Esto nos permite definir[1] la *clausura* de x respecto a R en A como el conjunto

$$\mathrm{cl}_A^R(x) = \bigcup_{n\in\omega} (A_x^R)_n \subset A.$$

Cuando R es la relación de pertenencia y $A = \mathrm{V}$, la clausura $\mathrm{cl}_A^R(x)$ se conoce como la *clausura transitiva* de x y se representa por $\mathrm{ct}\ x$. Es claro que admite una definición más sencilla (notemos que ahora $A_x^R = x$):

$$\mathrm{ct}_0 x = x, \qquad \bigwedge n \in \omega\ \mathrm{ct}_{n+1} x = \bigcup_{y \in \mathrm{ct}_n x} y, \qquad \mathrm{ct}\ x = \bigcup_{n\in\omega} \mathrm{ct}_n x.$$

Así, $\mathrm{ct}\ x$ está formada por los elementos de x, los elementos de los elementos de x, etc.

Nuestra intención al definir la clausura de un elemento era formar un conjunto R-A-transitivo. Vamos a ver que, efectivamente, así es. Más concretamente, $\mathrm{cl}_A^R(x)$ es el menor conjunto R-A-transitivo que contiene a A_x^R:

[1] El lector debería esforzarse en entender que si R no fuera conjuntista no estaría justificado el uso del teorema de recursión en la definición de $(A_x^R)_n$.

Teorema 12.6 *Sea R una relación conjuntista en una clase A y sea $x \in A$. Se cumple*

a) $A_x^R \subset \mathrm{cl}_A^R(x)$.

b) $\mathrm{cl}_A^R(x)$ *es un conjunto R-A-transitivo.*

c) Si $A_x^R \subset T$ *y* $T \subset A$ *es una clase R-A-transitiva, entonces* $\mathrm{cl}_A^R(x) \subset T$.

d) $\mathrm{cl}_A^R(x) = A_x^R \cup \bigcup_{y \in A_x^R} \mathrm{cl}_A^R(y)$.

Demostración: a) $A_x^R = (A_x^R)_0 \subset \mathrm{cl}_A^R(x)$.

b) Supongamos que u, $y \in A$ cumplen $u\,R\,y \wedge y \in \mathrm{cl}_A^R(x)$. Entonces existe un $n \in \omega$ tal que $y \in (A_x^R)_n$, con lo que $u \in A_y^R \subset (A_x^R)_{n+1} \subset \mathrm{cl}_A^R(x)$.

c) Una simple inducción prueba que $(A_x^R)_n \subset T$. En efecto, para 0 lo tenemos por hipótesis y si vale para n, entonces todo $u \in (A_x^R)_{n+1}$ cumple $u \in A_y^R$, para cierto $y \in (A_x^R)_n$, con lo que $u\,R\,y \wedge y \in T$. Por transitividad $u \in T$.

Por definición de clausura concluimos que $\mathrm{cl}_A^R(x) \subset T$.

d) Si $y \in A_x^R$, entonces $A_y^R \subset (A_x^R)_1 \subset \mathrm{cl}_A^R(x)$, luego por b) y c) obtenemos que $\mathrm{cl}_A^R(y) \subset \mathrm{cl}_A^R(x)$. Por consiguiente el conjunto de $T = A_x^R \cup \bigcup_{y \in A_x^R} \mathrm{cl}_A^R(y)$ está contenido en $\mathrm{cl}_A^R(x)$.

Para demostrar la otra inclusión basta probar T es transitivo y aplicar c). Sean, pues, u, $v \in A$ tales que $u\,R\,v \wedge v \in T$. Si $v \in \mathrm{cl}_A^R(y)$ para un $y \in A_x^R$, entonces, por la transitividad de la clausura $u \in \mathrm{cl}_A^R(y)$, luego $u \in T$.

Si $v \in A_x^R$, entonces $u \in \mathrm{cl}_A^R(v) \subset T$. ∎

Conviene observar la particularización de este teorema al caso de la relación de pertenencia sobre la clase universal:

Teorema 12.7 *Sea x un conjunto arbitrario. Entonces*

a) $x \subset \mathrm{ct}\ x$.

b) $\mathrm{ct}\ x$ *es un conjunto transitivo.*

c) Si $x \subset T$ *y* T *es una clase transitiva, entonces* $\mathrm{ct}\ x \subset T$.

d) $\mathrm{ct}\ x = x \cup \bigcup_{y \in x} \mathrm{ct}\ y$.

e) x es transitivo si y sólo si $x = \mathrm{ct}\ x$.

La última propiedad es consecuencia inmediata de las anteriores. Como primera aplicación del concepto de clausura demostramos un resultado técnico:

Teorema 12.8 *Sea R una relación conjuntista en una clase A. Entonces R está bien fundada en A si y sólo si todo subconjunto no vacío de A tiene un R-minimal.*

DEMOSTRACIÓN: Una implicación es obvia. Para la otra, suponemos que todo subconjunto no vacío tiene un R-minimal y hemos de probar que lo mismo vale para toda subclase no vacía B. Tomemos un $x \in B$. Si x no es ya un R-minimal de B, entonces existe un $y \in B$ tal que $y\,R\,x$, luego $y \in B \cap \mathrm{cl}_A^R(x)$, que es un subconjunto no vacío de A. Por hipótesis tiene un R-minimal, digamos z.

Vamos a ver que z es un R-minimal de B. En efecto, si existiera un $v \in B$ tal que $v\,R\,z$, entonces, por la transitividad de la clausura, $v \in B \cap \mathrm{cl}_A^R(x)$, pero esto contradice la minimalidad de z. ∎

Esto implica que el concepto de relación bien fundada es, pese a lo que en principio podría parecer, una fórmula normal (pues el cuantificador "para toda subclase no vacía" puede sustituirse por "para todo subconjunto no vacío").

12.2 Inducción y recursión transfinita

Como habíamos anticipado, el interés principal de las relaciones bien fundadas se debe a que sobre ellas es posible razonar por inducción y dar definiciones por recurrencia. El teorema de inducción es muy simple. Afirma que si R es una relación conjuntista y bien fundada sobre una clase A y queremos probar una inclusión $A \subset B$ (es decir, queremos probar que todos los elementos de A cumplen algo) podemos suponer que todos los elementos anteriores a un cierto x están en B y probar que $x \in B$:

Teorema 12.9 (Teorema general de inducción transfinita) *Sea R una relación conjuntista y bien fundada en una clase A. Entonces*

$$\bigwedge x \in A(A_x^R \subset B \to x \in B) \to A \subset B.$$

DEMOSTRACIÓN: En caso contrario $A \setminus B$ es una subclase no vacía de A, luego tiene un R-minimal x. Por definición de minimal, $A_x^R \subset B$, luego por hipótesis $x \in B$, con lo que tenemos una contradicción. ∎

El teorema general de recursión transfinita afirma que para definir una función F sobre una clase donde hay definida una relación conjuntista y bien fundada, podemos definir $F(x)$ suponiendo que F está ya definida sobre todos los elementos anteriores a x.

Teorema 12.10 (Teorema general de recursión transfinita) *Sea R una relación conjuntista y bien fundada en una clase A y sea $G : \mathrm{V} \longrightarrow \mathrm{V}$ una aplicación arbitraria. Entonces existe una única función $F : A \longrightarrow \mathrm{V}$ tal que*

$$\bigwedge x \in A\, F(x) = G(x, F|_{A_x^R}).$$

DEMOSTRACIÓN: La prueba se basa en las mismas ideas que el teorema análogo para ordinales, si bien es técnicamente más delicada. Por abreviar, a lo largo de esta prueba, "transitivo" significará R-A-transitivo.

Si $d \subset A$ es un conjunto transitivo, diremos que $h : d \longrightarrow \mathrm{V}$ es una *d-aproximación* si

$$\bigwedge x \in d\; h(x) = G(x, h|_{A_x^R}).$$

Para cada $x \in A$, definimos

$$\hat{x} = \{x\} \cup \mathrm{cl}_A^R(x)$$

Es claro que $\hat{x}$ es transitivo y $x \in \hat{x}$ (de hecho, es el menor conjunto transitivo que contiene a x). Dividimos la prueba en varios pasos:

1) Si h es una d-aproximación y h' es una d'-aproximación, entonces se cumple $h|_{d\cap d'} = h'|_{d\cap d'}$. En particular, para cada conjunto transitivo $d \subset A$ existe a lo sumo una d-aproximación.

Lo probamos por inducción en $d \cap d'$, es decir, vamos a probar que todo elemento de $d \cap d'$ está en $\{u \in d \cap d' \mid h(u) = h'(u)\}$. Para ello tomamos $x \in d \cap d'$ y suponemos que $h(u) = h(u')$ siempre que $u \in (d \cap d')_x^R$. Ahora bien, es inmediato que $d \cap d'$ es transitivo, de donde se sigue que $(d \cap d')_x^R = A_x^R$. Por consiguiente tenemos que $h|_{A_x^R} = h'|_{A_x^R}$, luego

$$h(x) = G(x, h|_{A_x^R}) = G(x, h'|_{A_x^R}) = h'(x).$$

2) Para todo $x \in A$ existe una $\hat{x}$-aproximación.

Lo probamos por inducción sobre x, es decir, suponemos que para todo $u \in A_x^R$ existe una $\hat{u}$-aproximación. Por 1) es única, luego podemos definir $h_u \equiv h|h$ es una $\hat{u}$-aproximación. Definimos $h = \bigcup_{u\in A_x^R} h_u$. De nuevo por 1) tenemos que h es una función y su dominio es

$$\bigcup_{u\in A_x^R} \hat{u} = \bigcup_{u\in A_x^R} (\{u\} \cup \mathrm{cl}_A^R(u)) = A_x^R \cup \bigcup_{u\in A_x^R} \mathrm{cl}_A^R(u) = \mathrm{cl}_A^R(x),$$

donde hemos aplicado el teorema 12.6.

Si $v \in \mathrm{cl}_A^R(x)$, entonces $h(v) = h_u(v)$, para cierto $u \in A_x^R$ tal que $v \in \hat{u}$. Puesto que $h_u \subset h$ y $A_v^R \subset \hat{u}$ (por ser $\hat{u}$ transitivo) tenemos que $h_u|_{A_v^R} = h|_{A_v^R}$. Como h_u es una $\hat{u}$-aproximación,

$$h(v) = h_u(v) = G(v, h_u|_{A_v^R}) = G(v, h|_{A_v^R}),$$

con lo que h resulta ser una $\mathrm{cl}_A^R(x)$-aproximación.

Puede probarse que $x \notin \mathrm{cl}_A^R(x)$, pero no es necesario, en cualquier caso podemos definir

$$h' = h \cup \{(x, G(x, h|_{A_x^R}))\},$$

de modo que $h : \hat{x} \longrightarrow \mathrm{V}$ y es inmediato que para todo $v \in \hat{x}$ se cumple $h'|_{A_x^R} = h|_{A_x^R}$, de donde se sigue claramente que h' es una $\hat{x}$-aproximación.

3) Definimos $F = \bigcup_{x\in A} h_x$, donde $h_x \equiv h|h$ es una $\hat{x}$-aproximación.

La unicidad de 1) hace que $F : A \longrightarrow \mathrm{V}$, y los mismos razonamientos que hemos aplicado a h en el paso anterior prueban que para todo $x \in A$ se cumple $F(x) = G(x, F|_{A_x^R})$.

4) La unicidad de F se prueba igual que 1) ∎

Como primera aplicación de este teorema, dada una clase con una relación conjuntista y bien fundada, vamos a asociar a cada uno de sus elementos un ordinal que exprese su "altura" en la relación, entendiendo que un elemento es más alto cuantos más elementos tiene por debajo.

Definición 12.11 Sea R una relación conjuntista y bien fundada en una clase A. Definimos $\operatorname{rang} : A \longrightarrow \Omega$ como la única aplicación que cumple

$$\bigwedge x \in A \ \operatorname{rang}_A^R x = \bigcup_{y \in A_x^R} (\operatorname{rang}_A^R y + 1).$$

Observemos que hemos definido el rango de un elemento supuesto definido el rango de los elementos anteriores a él. Recordemos que la unión de un conjunto de ordinales no es más que su supremo. Hemos de entender que el supremo del conjunto vacío es 0 (lo cual es cierto, pues 0 es la menor cota superior de $\varnothing$).

De este modo, los minimales de A tienen todos rango 0 y, en general, el rango de un elemento es el mínimo ordinal estrictamente mayor que los rangos de todos sus anteriores.

Teorema 12.12 *Sea R una relación conjuntista y bien fundada en una clase A. Sean x, $y \in A$. Si $x \in \operatorname{cl}_A^R(y)$, entonces* $\operatorname{rang}_A^R x < \operatorname{rang}_A^R y$.

Demostración: Por la definición de clausura, basta probar que esto es cierto para todo $x \in (A_y^R)_n$, para todo $n \in \omega$. Si $n = 0$, tenemos que $x \in A_y^R$, luego, por definición de rango, $\operatorname{rang}_A^R x < \operatorname{rang}_A^R x + 1 \leq \operatorname{rang}_A^R y$.

Si es cierto para n, supongamos $x \in (A_y^R)_{n+1}$, con lo que $x \in A_v^R$, para cierto $v \in (A_y^R)_n$. Por el caso $n = 0$ tenemos que $\operatorname{rang}_A^R x < \operatorname{rang}_A^R v$ y por hipótesis de inducción $\operatorname{rang}_A^R v < \operatorname{rang}_A^R y$. ∎

Con la ayuda del rango podemos demostrar teoremas de inducción y recursión aún más potentes. En el caso de la inducción, vamos a ver que podemos tomar como hipótesis de inducción, no ya que todos los elementos anteriores a uno dado cumplen lo que queremos probar, sino que todos los elementos de su clausura lo cumplen (o sea, los anteriores, y los anteriores de los anteriores, etc.).

Teorema 12.13 (Teorema general de inducción transfinita) *Sea R una relación conjuntista y bien fundada sobre una clase A. Entonces*

$$\bigwedge x \in A(\operatorname{cl}_A^R(x) \subset B \to x \in B) \to A \subset B.$$

Demostración: Si no se da la inclusión podemos tomar un $x \in A \setminus B$ de rango mínimo. Si $u \in \operatorname{cl}_A^R(x)$, entonces $\operatorname{rang}_A^R u < \operatorname{rang}_A^R x$, luego por minimalidad $u \in B$. Pero entonces la hipótesis nos da que $x \in B$, lo cual es absurdo. ∎

Similarmente, para definir una función podemos suponer que está ya definida sobre la clausura de un elemento dado:

Teorema 12.14 (Teorema general de recursión transfinita) *Sea R una relación conjuntista y bien fundada en una clase A y sea $G : \mathrm{V} \longrightarrow \mathrm{V}$ una aplicación arbitraria. Entonces existe una única función $F : A \longrightarrow \mathrm{V}$ tal que*

$$\bigwedge x \in A\, F(x) = G(x, F|_{\mathrm{cl}_A^R(x)}).$$

La prueba de este teorema es idéntica a la de 12.9, salvo que el paso 1) y la unicidad de F se demuestran usando la versión fuerte del teorema general de recursión transfinita en lugar de la débil.

Como aplicación de los teoremas de inducción y recursión vamos a generalizar el teorema que afirma que todo conjunto bien ordenado es semejante a un único ordinal.

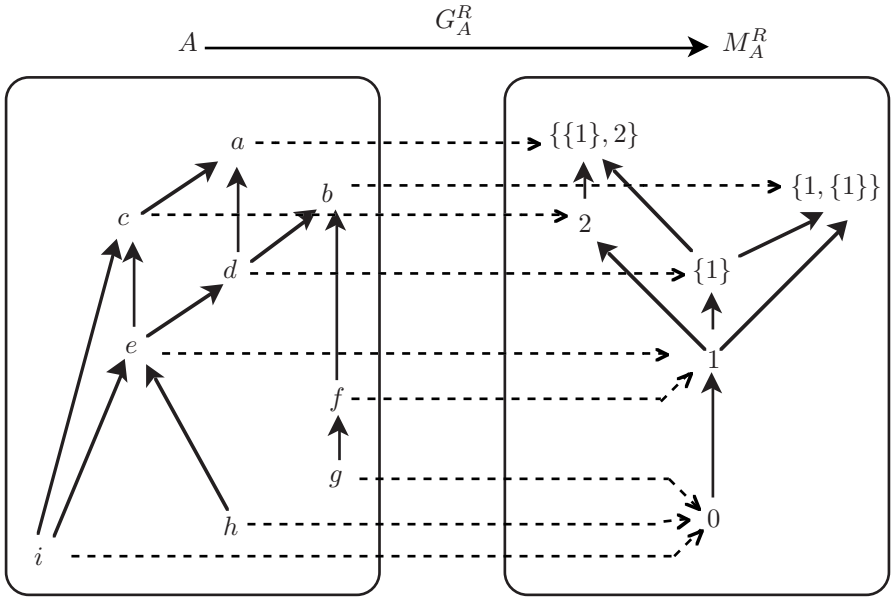
Definición 12.15 Sea R una relación conjuntista y bien fundada en una clase A. Llamaremos *función colapsante de Mostowski* a la función $G_A^R : A \longrightarrow \mathrm{V}$ dada por

$$G_A^R(x) = \{G_A^R(y) \mid y \in A \wedge y\, R\, x\}.$$

Así pues, $G_A^R(x)$ es el conjunto de las imágenes por G de los anteriores de x.

Definimos el *colapso transitivo* de A como el rango de G_A^R, y lo representaremos por M_A^R, de modo que $G_A^R : A \longrightarrow M_A^R$ suprayectiva.

Ejemplo La figura muestra un conjunto A con nueve elementos y las flechas representan una relación bien fundada R. A la derecha tenemos su colapso transitivo y las líneas discontinuas representan la función colapsante.



Ejercicio: Calcular el rango de cada elemento del conjunto A.

Vamos a probar que, bajo condiciones adecuadas, M_A^R es una generalización del ordinal de un conjunto bien ordenado. Como prueba el teorema siguiente, M_A^R es siempre una clase transitiva y bien fundada, de modo que sólo le falta ser conexa para ser un ordinal.

Teorema 12.16 *Sea R una relación conjuntista y bien fundada sobre una clase A. Entonces M_A^R es una clase transitiva y bien fundada y*

$$\bigwedge xy \in A(x\,R\,y \to G_A^R(x) \in G_A^R(y)).$$

La última afirmación es consecuencia inmediata de la definición de G. Respecto a la transitividad, tomemos $u \in v \in M_A^R$. Entonces $v = G_a^R(x)$, para cierto $x \in A$, luego $u = G_A^R(y)$, para cierto $y \in A$ tal que $y\,R\,x$. Por consiguiente $u \in M_A^R$.

Veamos ahora que M_A^R está bien fundada. Tomamos $X \subset M_A^R$, $X \neq \varnothing$. Entonces $(G_A^R)^{-1}[X]$ es una subclase no vacía de A, luego tiene un R-minimal y. Veamos que $G_A^R(y)$ es un minimal de X, es decir, que $X \cap G_A^R(y) = \varnothing$. En efecto, si $u \in X \cap G_A^R(y)$, entonces $u = G_A^R(x)$, para un cierto $x \in A$ tal que $x\,R\,y$. Entonces $x \in (G_A^R)^{-1}[X]$, luego contradice la minimalidad de y. ∎

El ejemplo que hemos visto muestra que la función colapsante no es necesariamente inyectiva. Lo será cuando la relación R sea un buen orden (estricto, para que sea una relación bien fundada) y entonces el colapso será simplemente su ordinal. Sin embargo, para que la función colapsante sea biyectiva no es necesario exigir tanto de R, sino que basta con que cumpla una propiedad obviamente necesaria:

Definición 12.17 Una relación R es *extensional* sobre una clase A si

$$\bigwedge xy \in A(\bigwedge u \in A(u\,R\,x \leftrightarrow u\,R\,y) \to x = y).$$

En otras palabras, R es extensional si cumple lo que el axioma de extensionalidad postula de la relación de pertenencia: que dos objetos distintos tienen elementos anteriores distintos. Es claro que la relación de pertenencia es extensional en cualquier clase transitiva. Toda relación de orden total es extensional.

Teorema 12.18 *Sea R una relación conjuntista, extensional y bien fundada en una clase A. Entonces $G_A^R : A \longrightarrow M_A^R$ biyectiva y*

$$\bigwedge xy \in A(x\,R\,y \leftrightarrow G_A^R(x) \in G_A^R(y)).$$

Demostración: Si G_A^R no fuera inyectiva podríamos tomar un R-minimal x de la clase

$$B = \{x \in A \mid \bigvee y \in A(x \neq y \wedge G_A^R(x) = G_A^R(y)\}.$$

Entonces existe un $y \in A$ tal que $y \neq x$ y $G^R_A(x) = G^R_A(y)$. Como R es extensional, existe un $z \in A$ tal que

$$(z\,R\,x \wedge \neg z\,R\,y) \vee (z\,R\,y \wedge \neg z\,R\,x).$$

Si $z\,R\,x \wedge \neg z\,R\,y$, entonces $G^R_A(z) \in G^R_A(x) = G^R_A(y)$, luego, por definición de $G^R_A(y)$, ha de ser $G^R_A(z) = G^R_A(w)$, para cierto $w \in A$, con $w\,R\,y$. Además, como $\neg z\,R\,y$, ha de ser $w \neq z$, o sea, $z \in B \wedge z\,R\,x$, contradicción.

Si $z\,R\,y \wedge \neg z\,R\,x$, entonces $G^R_A(z) \in G^R_A(y) = G^R_A(x)$, luego $G^R_A(z) = G^R_A(w)$, con $w \in A$, $w\,R\,x$, luego en particular $w \neq z$. Por consiguiente $w \in B$ contradice la minimalidad de x.

Con esto tenemos la biyectividad. Finalmente, si $G^R_A(x) \in G^R_A(y)$, ha de ser $G^R_A(x) = G^R_A(u)$, para cierto $u \in A$, $u\,R\,y$. Como G^R_A es biyectiva, ha de ser $x = u$, luego $u\,R\,y$. ∎

Así, del mismo modo que el ordinal de un conjunto bien ordenado es un representante de su ordenación, concretamente, el único representante transitivo en el que la relación de orden es la pertenencia, ahora vemos que cualquier clase dotada de una relación conjuntista, extensional y bien fundada tiene asociada una clase transitiva en la cual la relación de pertenencia se comporta exactamente igual que la relación dada. Para terminar de perfilar la situación probamos la unicidad tanto del colapso transitivo como de la función colapsante:

Teorema 12.19 (Teorema del colapso de Mostowski) *Si R es una relación conjuntista, bien fundada y extensional en una clase A, existe una única clase transitiva M y una única aplicación $G : A \longrightarrow M$ biyectiva tal que*

$$\bigwedge xy \in A(x\,R\,y \leftrightarrow G(x) \in G(y)).$$

Demostración: Sólo falta probar la unicidad. Ahora bien, si G si G y M cumplen estas propiedades, se ha de cumplir que

$$\bigwedge x \in A\ G(x) = \{G(y) \mid y \in A \wedge y\,R\,x\}.$$

En efecto, si $u \in G(x) \in M$, por transitividad $u \in M$, luego $u = G(y)$, para cierto $y \in A$. Entonces $G(y) \in G(x)$, luego $x\,R\,y$. La otra inclusión es inmediata.

Por consiguiente G es la función colapsante (por la unicidad de su definición recurrente) y M ha de ser el colapso transitivo. ∎

Nota En el capítulo IV demostramos el teorema de Löwenheim-Skolem usando una técnica de Henkin, que de hecho nos lleva al teorema de completitud. No obstante, la técnica original de Skolem puede aplicarse para obtener modelos numerables de porciones finitas de ZFC con la particularidad de que la relación de pertenencia se interpreta como ella misma. Así, si M es uno de estos modelos y satisface el axioma de extensionalidad, ello implica que $\in$ es extensional en M (de hecho es equivalente). Si además suponemos el axioma de regularidad,

resulta que todo conjunto no vacío tiene un minimal para $\in$, por lo que $\in$ está bien fundada en todo conjunto, en particular en M. Por consiguiente, podemos aplicar el teorema anterior, que nos da un modelo con las mismas propiedades que M (isomorfo) pero que tiene la propiedad adicional de ser transitivo. Por razones que ahora sería largo detallar, la transitividad simplifica enormemente el trabajo con un modelo, por lo que ésta es una de las principales aplicaciones del teorema de Mostowski: transitivizar modelos.

El nombre de "colapso" hace referencia a que en este caso en el que la relación de partida es la relación de pertenencia el colapso supone una simplificación, una compresión de la clase dada. Por ejemplo, el colapso transitivo de $\{\{\{\varnothing\}\}\}$ es $\{\varnothing\}$. ∎

Ejercicio: Probar que si R es un buen orden estricto (y conjuntista) en una clase A entonces M_A^R es un ordinal y G_A^R es una semejanza. Deducir de aquí pruebas alternativas de los teoremas 11.21 y 11.27.

Ejercicio: Deducir del teorema anterior que si $F : A \longrightarrow B$ es una biyección entre clases transitivas tal que $\bigwedge xy \in A(x \in y \leftrightarrow F(x) \in F(y))$ entonces $A = B$ y F es la identidad.

12.3 Conjuntos regulares

En esta sección estudiaremos más detenidamente el significado y las repercusiones del axioma de regularidad. Según explicábamos, la finalidad de este axioma es negar la existencia de conjuntos "perversos" que no puedan interpretarse como construidos a partir de elementos previos, como ocurriría con una pareja $x = \{y\}$, $y = \{x\}$.

Para describir con precisión este tipo de patologías introducimos la noción de conjunto regular:

Definición 12.20 Un conjunto x es *regular* si la relación $\in$ está bien fundada en su clausura transitiva. Llamaremos R a la clase de todos los conjuntos regulares.

Así, si a partir de un conjunto x podemos construir una sucesión de conjuntos tal que

$$x \quad \ni \quad x_0 \quad \ni \quad x_1 \quad \ni \quad x_2 \quad \ni \quad x_3 \quad \ni \quad \cdots$$

entonces todos ellos están en ct x, por lo que el teorema 12.2 nos da que $\in$ no está bien fundada en ct x, es decir, que x no es regular. Usando el axioma de elección también tenemos el recíproco.

Notemos que no hubiera servido definir un conjunto regular como un conjunto en el que $\in$ está bien fundada, pues si $x = \{y\}$, $y = \{x\}$, entonces $\in$ está bien fundada tanto en x como en y, pero no lo está en ct $x = \{x, y\}$.

El axioma de regularidad equivale a que todo conjunto es regular, es decir, a la igualdad V = R, y es por esto que nos hemos referido a él como V = R incluso antes de haber definido la clase R.

En efecto: lo que afirma propiamente el axioma de regularidad es que todo conjunto no vacío tiene un minimal (para $\in$). Si esto es así, entonces $\in$ está bien fundada en todo conjunto, en particular en la clausura transitiva de todo conjunto. Recíprocamente, si $\mathrm{V} = \mathrm{R}$ y x es cualquier conjunto no vacío, entonces $x \subset \mathrm{ct}\ x$ y $\in$ está bien fundada en ct x, luego x tiene un minimal.

Así pues, en lo sucesivo $\mathrm{V} = \mathrm{R}$ no será ya un convenio de notación para referirnos al axioma de regularidad, sino una forma equivalente del mismo.

Veamos ahora las propiedades de los conjuntos regulares. Notemos que para ello no necesitamos suponer el axioma de regularidad. Para empezar probamos que los elementos de los conjuntos regulares son de nuevo conjuntos regulares, o en otros términos:

Teorema 12.21 *R es una clase transitiva.*

Demostración: Si $u \in v \in \mathrm{R}$, entonces $u \in \mathrm{ct}\ v$, luego $u \subset \mathrm{ct}\ v$, luego $\mathrm{ct}\ u \subset \mathrm{ct}\ v$. Por hipótesis $\in$ está bien fundada en ct v, luego también lo está en ct u, es decir, $u \in \mathrm{R}$. ∎

Por otra parte:

Teorema 12.22 *La relación de pertenencia está bien fundada en R.*

Demostración: Sea $X \subset R$ una subclase no vacía. Sea $u \in X$. Si u no es ya un minimal de X, entonces $u \cap X \neq \varnothing$, luego $\mathrm{ct}\ u \cap X \neq \varnothing$. Sea v un minimal de $\mathrm{ct}\ u \cap X$ (que existe porque u es regular) y vamos a probar que v es, de hecho un minimal de X. En efecto, si existiera $w \in v \cap X$, entonces, por transitividad, $w \in \mathrm{ct}\ u \cap X$ y $w \in v$, en contra de la minimalidad de v. ∎

De aquí extraemos ya algunas consecuencias notables:

Teorema 12.23 *R no es un conjunto.*

Demostración: Si lo fuera, como es transitiva sería $\mathrm{ct}\ \mathrm{R} = \mathrm{R}$ y por el teorema anterior R sería un conjunto regular, es decir, $R \in R$, pero entonces $\{R\}$ sería un subconjunto de R sin minimal, en contradicción con el teorema anterior. ∎

Teorema 12.24 $\mathcal{P}\mathrm{R} = \mathrm{R}$.

Demostración: Como R es transitiva, $\mathrm{R} \subset \mathcal{P}\mathrm{R}$. Si $x \in \mathcal{P}\mathrm{R}$, entonces $x \subset \mathrm{R}$. Como R es transitiva, $\mathrm{ct}\ x \subset \mathrm{R}$ y como $\in$ está bien fundada en R también lo está en ct x. Así pues, x es regular, luego $x \in \mathrm{R}$. ∎

Si particularizamos a R el teorema general de inducción transfinita obtenemos esencialmente que R es la menor clase con esta propiedad. En efecto, este principio dice que para probar $\mathrm{R} \subset A$ podemos suponer que un $x \in \mathrm{R}$ cumple $x \subset A$ y demostrar que $x \in A$, es decir:

Teorema 12.25 *Para toda clase A:*

$$\mathrm{R} \cap \mathcal{P}A \subset A \to \mathrm{R} \subset A.$$

En particular: $\mathcal{P}A \subset A \to \mathrm{R} \subset A$.

Así pues, para probar que todo conjunto regular cumple algo podemos suponer que todos los elementos de un conjunto dado x lo cumplen y probar que x también lo cumple. Por supuesto, si lo deseamos, podemos tomar como hipótesis de inducción que todos los elementos de ct x cumplen lo que queremos probar.

Así mismo disponemos de los teoremas de recursión transfinita:

Teorema 12.26 *Dada $G : \mathrm{R} \longrightarrow \mathrm{V}$, existe una única función $F : \mathrm{R} \longrightarrow \mathrm{V}$ tal que*

$$\bigwedge x \in \mathrm{R}\ F(x) = G(F|_x),$$

o, alternativamente:

$$\bigwedge x \in \mathrm{R}\ F(x) = G(x, F|_{\mathrm{ct}\ x}).$$

No ponemos $G(x, F|_x)$ porque $x = \mathcal{D}F|_x$, es decir, si conocemos $F|_x$ conocemos x y, dada la generalidad de G, sería redundante hacerla depender de x.

Para terminar con estas propiedades básicas de R observamos que los colapsos transitivos están formados siempre por conjuntos regulares. Más en general:

Teorema 12.27 *Si A es una clase transitiva en la que la relación de pertenencia está bien fundada, entonces $A \subset \mathrm{R}$.*

Demostración: Se prueba inmediatamente por inducción transfinita en A. Observemos que si $x \in A$ entonces $x \subset A$, luego $A^{\in}_x = x$. Suponemos que $x \in A$ y, como hipótesis de inducción, que $x \subset \mathrm{R}$, pero entonces $x \in \mathcal{P}\mathrm{R} = \mathrm{R}$. ∎

En la sección anterior definimos el rango de un elemento respecto a una relación bien fundada. En el caso concreto de los conjuntos regulares la definición del rango se particulariza a

$$\mathrm{rang}\ x = \bigcup_{y \in x} (\mathrm{rang}\ y + 1).$$

Para cada $\alpha \in \Omega$, definimos

$$\mathrm{R}_\alpha = \{x \in \mathrm{R} \mid \mathrm{rang}\ x < \alpha\}.$$

Así obtenemos una descripción clara de la clase R:

Teorema 12.28 *Se cumple:*

$$\mathrm{R}_0 = \varnothing \quad \wedge \quad \bigwedge \alpha\ \mathrm{R}_{\alpha+1} = \mathcal{P}\mathrm{R}_\alpha \quad \wedge \quad \bigwedge \lambda\ \mathrm{R}_\lambda = \bigcup_{\delta < \lambda} \mathrm{R}_\delta,$$

$$\mathrm{R} = \bigcup_{\alpha \in \Omega} \mathrm{R}_\alpha.$$

DEMOSTRACIÓN: Probamos únicamente el caso de $\mathrm{R}_{\alpha+1}$, pues los otros son inmediatos. Si $x \in \mathcal{P}\mathrm{R}_\alpha$, entonces $x \subset \mathrm{R}_\alpha$, luego

$$\operatorname{rang} x = \bigcup_{y\in x} (\operatorname{rang} y + 1) \leq \alpha < \alpha + 1,$$

luego $x \in \mathrm{R}_{\alpha+1}$. Recíprocamente, si $x \in \mathrm{R}_{\alpha+1}$, entonces todo $y \in x$ cumple $\operatorname{rang} y + 1 \leq \operatorname{rang} x < \alpha + 1$, luego $\operatorname{rang} y < \alpha$, luego $y \in \mathrm{R}_\alpha$. Así pues, $x \subset \mathrm{R}_\alpha$. ∎

Así pues, los conjuntos regulares son los conjuntos que pueden obtenerse a partir del conjunto vacío en una cantidad transfinita de pasos mediante el operador $\mathcal{P}$. Notemos que, en particular, R_ω no es sino el conjunto de todos los conjuntos hereditariamente finitos, de los que ya hemos hablado (metamatemáticamente) en otras ocasiones.

Ejercicio: Probar que las clases R_α son transitivas.

El teorema anterior nos proporciona una definición alternativa de la clase de los conjuntos regulares supuesto que aceptemos el axioma de partes, pues sin este axioma no podemos justificar que $\mathrm{R}_{\omega+1}$ y las clases posteriores sean conjuntos y no podemos aplicar el teorema de recursión transfinita. Por el contrario, suponiendo AP, una inducción transfinita nos da inmediatamente que todas las clases R_α son conjuntos.

Ejercicio: Supuesto definido R mediante el teorema anterior (aceptando AP), probar las propiedades básicas de los conjuntos R_α (son transitivos, forman una sucesión creciente, definen el rango de un conjunto regular, etc.).

Del teorema 12.27 se sigue ya que $\Omega \subset \mathrm{R}$. Veamos con más detalle la relación entre ambas clases. En particular, el teorema siguiente prueba que hay conjuntos regulares de todos los rangos posibles:

Teorema 12.29 $\Omega \subset \mathrm{R}$, $\bigwedge\alpha \operatorname{rang} \alpha = \alpha$, $\bigwedge\alpha\ \mathrm{R}_\alpha \cap \Omega = \alpha$.

DEMOSTRACIÓN: Veamos que $\operatorname{rang} \alpha = \alpha$ por inducción. Supongamos que es cierto para todo $\beta < \alpha$. Entonces

$$\operatorname{rang} \alpha = \bigcup_{\beta<\alpha} (\operatorname{rang} \beta + 1) = \bigcup_{\beta<\alpha} (\beta + 1) = \alpha.$$

Por último, $u \in \mathrm{R}_\alpha \cap \Omega$ si y sólo si $u \in \Omega \wedge \operatorname{rang} u < \alpha$ si y sólo si $u \in \Omega \wedge u < \alpha$ si y sólo si $u \in \alpha$. Así pues, $\mathrm{R}_\alpha \cap \Omega = \alpha$. ∎

Ahora es claro que todos los conjuntos de interés para los matemáticos son regulares. Por ejemplo, un par ordenado (m, n) de números naturales está en R_ω, un número entero z se define como un conjunto de pares de números naturales, luego $z \subset \mathrm{R}_\omega$, con lo que $z \in \mathrm{R}_{\omega+1}$ y, por consiguiente, $\mathbb{Z} \in \mathrm{R}_{\omega+2}$. Similarmente se comprueba que $\mathbb{Q} \in \mathrm{R}_{\omega+5}$. Si construimos $\mathbb{R}$ por el método de Dedekind, según el cual un número real es un subconjunto de $\mathbb{Q}$, tenemos $\mathbb{R} \in \mathrm{R}_{\omega+6}$, etc.

Ejercicio: Calcular el rango de $\mathbb{R}$ si lo suponemos construido por el método de Cantor, es decir, si un número real es una clase de equivalencia de sucesiones de Cauchy de números racionales.

Esto hace que, en la práctica, el axioma de regularidad sea superfluo: no es necesario postular que no existen conjuntos regulares, basta con limitarse a trabajar con conjuntos regulares. Ahora bien, dado que puede probarse que los axiomas de ZFC o NBG sin V = R (supuesto que sean consistentes) no implican la existencia de conjuntos no regulares, podemos simplificar la teoría postulando que no existen.

Cuando se supone el axioma de regularidad, a los conjuntos R_α se les suele llamar V_α, de modo que la clase universal queda estructurada en la jerarquía transfinita creciente de clases transitivas (conjuntos si suponemos AP) dada por:

$$\mathrm{V}_0 = \varnothing \quad \wedge \quad \bigwedge\alpha\ \mathrm{V}_{\alpha+1} = \mathcal{P}\mathrm{V}_\alpha \quad \wedge \quad \bigwedge\lambda\ \mathrm{V}_\lambda = \bigcup_{\delta<\lambda} \mathrm{V}_\delta \quad \wedge \quad \mathrm{V} = \bigcup_{\alpha\in\Omega} \mathrm{V}_\alpha,$$

de modo que todo conjunto puede pensarse como construido a partir de $\varnothing$ en una cantidad transfinita de pasos, en el sentido de que si rastreamos sus elementos y los elementos de sus elementos, etc. siempre terminamos en $\varnothing$.

Todo conjunto x tiene definido un rango, que es el menor ordinal α tal que $x \subset \mathrm{V}_\alpha$ o, equivalentemente, $x \in \mathrm{V}_{\alpha+1}$. El rango de un conjunto es una medida de su complejidad o, equivalentemente, de la cantidad de pasos que hemos de dar para construirlo a partir de $\varnothing$.

Si suponemos AP, tenemos una distinción clara entre los conjuntos y las clases propias:

Teorema 12.30 *Una clase es propia si y sólo si contiene conjuntos de rango arbitrariamente grande.*

Demostración: Si todos los elementos de una clase X tienen rango menor que un ordinal α entonces $X \subset \mathrm{V}_\alpha$, luego X es un conjunto. Recíprocamente, si X es un conjunto, la imagen de X por la aplicación rango es un subconjunto de Ω (por el axioma del reemplazo), luego está acotado. ∎

12.4 Átomos

Aunque, como acabamos de comentar, de los axiomas de la teoría de conjuntos no se deduce la existencia de conjuntos no regulares, lo cierto es que tampoco se deduce su no existencia (siempre suponiendo que sean consistentes), lo que nos deja abierta la posibilidad de postular la existencia de conjuntos no regulares. Vamos a exponer ahora una alternativa al axioma de regularidad que, por ciertas razones, no carece por completo de interés.

Observemos que, de acuerdo con las definiciones que hemos dado, $3 = \{0, 1, 2\}$. Ahora bien, esto es un mero convenio arbitrario. El número 3, en

su sentido metamatemático, no es un conjunto, y carece de sentido preguntarse cuáles son sus elementos. Un niño de diez años conoce perfectamente los números naturales y se quedaría perplejo si alguien le preguntara por los elementos de 3. Lo que sucede es que la teoría de conjuntos presupone desde su misma base que todos los objetos sobre los que trata son conjuntos (o clases, que para el caso es lo mismo). Ello fuerza a identificar artificialmente con determinados conjuntos conceptos que, en principio, no corresponden a conjunto alguno.[2]

Sería conceptualmente más natural trabajar en una teoría que admitiera la existencia de objetos básicos que no puedan considerarse formados por objetos más elementales, objetos que podríamos llamar "átomos". Ahora bien, lo cierto es que también seria técnicamente más complicado, pues, por lo pronto, exigiría modificar la teoría de conjuntos desde su raíz, y la "naturalidad" del resultado no compensaría el esfuerzo, ya que, con toda su "artificialidad" la teoría de "sólo-conjuntos" cumple perfectamente su finalidad.

No obstante, hay una forma de "burlar" a la teoría de conjuntos respetando todos sus principios e incorporando átomos al mismo tiempo. Consiste en identificar un átomo x con un conjunto de la forma $x = \{x\}$. En las secciones anteriores afirmábamos que un objeto así es patológico sobre la base de que pretendíamos hablar de conjuntos, pero, por otra parte, se trata de un convenio muy natural si queremos hablar de átomos. En último extremo, decir que el único elemento de x es x equivale a decir que es inútil preguntar por los elementos de x. Notemos que no serviría postular literalmente que los átomos no tienen elementos, porque entonces el axioma de extensionalidad forzaría a que todos los átomos son el conjunto vacío, es decir, a que sólo hay un átomo. En cambio, con este convenio el axioma de extensionalidad sólo dice que dos átomos son iguales si y sólo si son iguales, lo cual es perfecto.

Por simplicidad, vamos a trabajar en $\mathrm{NBG}^* + AI + AP$. Definimos la clase de los *átomos* como

$$\mathrm{A} = \{x \mid x = \{x\}\}.$$

Por ejemplo, el axioma de regularidad implica $\mathrm{A} = \varnothing$. De momento tomamos como axioma únicamente que A es un conjunto y vamos a definir los conjuntos regulares respecto de A. En realidad esta definición es válida para un conjunto arbitrario X:

Definición 12.31 Llamaremos clase $\mathrm{R}(X)$ de los conjuntos *regulares* respecto a un conjunto X como la dada por

$$\mathrm{R}_0(X) = \mathrm{ct}\ X \quad \wedge \quad \bigwedge\alpha\ \mathrm{R}_{\alpha+1}(X) = \mathcal{P}\mathrm{R}_\alpha(X) \quad \wedge \quad \bigwedge\lambda\ \mathrm{R}_\lambda(X) = \bigcup_{\delta<\lambda} \mathrm{R}_\delta(X),$$

$$\mathrm{R}(X) = \bigcup_{\alpha\in\Omega} \mathrm{R}_\alpha(X).$$

[2]Andrej Mostowski expresó su insatisfacción por este fenómeno con su protesta: ¡yo no soy un conjunto!

Una simple inducción prueba que cada $\mathrm{R}_\alpha(X)$ es un conjunto transitivo. En efecto, es claro para $\alpha = 0$ y trivial para λ. Supuesto cierto para α, entonces $\mathrm{R}_{\alpha+1}(X)$ es un conjunto por AP y si suponemos que $u \in v \in \mathrm{R}_{\alpha+1}(X)$, tenemos que $u \in v \subset \mathrm{R}_\alpha(X)$, luego $u \in \mathrm{R}_\alpha(X)$, luego (por hipótesis de inducción) $u \subset \mathrm{R}_\alpha(X)$, luego $u \in \mathrm{R}_{\alpha+1}(X)$.

Consecuentemente $\mathrm{R}(X)$ es una clase transitiva, luego $\mathrm{R}(X) \subset \mathcal{P}\mathrm{R}(X)$. De hecho se da la igualdad $\mathrm{R}(X) = \mathcal{P}\mathrm{R}(X)$. En efecto, si $x \subset \mathrm{R}(X)$, a cada $y \in x$ le podemos asignar el mínimo ordinal α_y tal que $y \in \mathrm{R}_\alpha(X)$. El axioma del reemplazo nos da que $\{\alpha_y \mid y \in x\}$ es un conjunto de ordinales, luego tiene supremo α y así $x \subset \mathrm{R}_\alpha(X)$, luego $x \in \mathrm{R}_{\alpha+1}(X) \subset \mathrm{R}(X)$.

La igualdad $\mathrm{R}(X) = \mathcal{P}\mathrm{R}(X)$ implica, por el teorema 12.25, que $\mathrm{R} \subset \mathrm{R}(X)$. No obstante, una simple inducción nos da la relación más fina $\bigwedge\alpha\ \mathrm{R}_\alpha \subset \mathrm{R}_\alpha(X)$.

Ejercicio: Probar que $\mathrm{R} = \mathrm{R}(X)$ si y sólo si $X \in \mathrm{R}$.

En general, el comportamiento de $\mathrm{R}(X)$ es algo caótico porque si X (o su clausura transitiva) contiene conjuntos regulares éstos aparecen "antes de tiempo" en la jerarquía. Ahora bien, esto no sucede si partimos del conjunto de átomos.

Definición 12.32 Llamaremos NBGA (resp. ZFCA) a la teoría de conjuntos NBG (resp. ZFC) sin el axioma de regularidad más el axioma[3]

$$\mathrm{cto}\,\mathrm{A} \wedge \mathrm{V} = \mathrm{R}(\mathrm{A}).$$

Notemos que este axioma no contradice a $\mathrm{V} = \mathrm{R}$, sino que equivale a éste si añadimos que $\mathrm{A} = \varnothing$. Por otra parte, la teoría de conjuntos con átomos puede completarse con varios axiomas, como por ejemplo $\bigvee f\ f : \omega \longrightarrow \mathrm{A}$ biyectiva, que nos da un conjunto numerable de átomos. También conviene observar que A es un conjunto transitivo, por lo que $\mathrm{R}_0(\mathrm{A}) = \mathrm{A}$.

Una simple inducción prueba que $\mathrm{R} \cap \mathrm{R}_\alpha(\mathrm{A}) = \mathrm{R}_\alpha$, por lo que si definimos el rango de un conjunto x como el mínimo ordinal α tal que $x \subset \mathrm{R}_\alpha(\mathrm{A})$, resulta que este rango coincide sobre R con el que ya teníamos definido. Los conjuntos de rango 0 son los formados por átomos (incluyendo a los propios átomos y, trivialmente, a $\varnothing$). La relación $x \in y \to \mathrm{rang}\ x < \mathrm{rang}\ y$ se cumple salvo si y tiene rango 0.

Los teoremas de inducción y recursión se adaptan fácilmente a este contexto (la relación de pertenencia ya no está bien fundada, pero pueden probarse mediante inducción y recursión ordinal sobre el rango). Por ejemplo, es fácil ver que

$$\bigwedge X(\mathrm{A} \subset X \wedge \mathcal{P}X \subset X \to X = \mathrm{V}).$$

[3]Para que esta sentencia tenga sentido como axioma de ZFCA hemos de observar que puede reformularse omitiendo toda referencia a clases propias. Concretamente:

$$\bigvee x(\bigwedge y(y \in x \leftrightarrow y = \{y\}) \wedge \bigwedge u \bigvee \alpha\ u \in \mathrm{R}_\alpha(x)).$$

No vamos a desarrollar más allá la teoría de conjuntos con átomos. Comentemos únicamente que von Neumann empleó esta teoría para probar la independencia del axioma de elección, es decir, demostró que en ZFA (es decir, la teoría con átomos pero sin el axioma de elección) no es posible demostrar el axioma de elección. Para ello probó la consistencia de que existieran contraejemplos al axioma de elección, pero estos contraejemplos eran necesariamente átomos. No obstante, hoy se sabe que todos los resultados de consistencia que pueden obtenerse con la técnica de von Neumann para conjuntos con átomos, pueden obtenerse también para conjuntos sin átomos, por lo que la hipótesis de que existan átomos resulta ser una simplificación no esencial. De aquí que esta teoría no esté totalmente desprovista de interés.

Capítulo XIII

Números cardinales

Entramos ahora en la parte más profunda de la teoría cantoriana de conjuntos: el estudio de los cardinales infinitos. En esta parte intervienen en mayor o menor medida todos los axiomas de la teoría de conjuntos: el axioma de infinitud para asegurar la existencia de conjuntos infinitos, el axioma de partes para asegurar la existencia de conjuntos no numerables y el axioma de elección para asegurar que los cardinales se comportan de forma razonable. Conviene ser consciente de los axiomas en los que se basa cada teorema. Así, los resultados más elementales no requieren más que la teoría básica ZF^* o NBG^*. No obstante, para no ser prolijos no llevaremos un control sistemático, sino que nos limitaremos a hacer algunas observaciones generales de tanto en tanto. La única excepción la haremos con el axioma de elección, pues sí tiene gran interés saber qué papel desempeña exactamente en la teoría. Por ello establecemos que vamos a trabajar en ZF o $\mathrm{NBG}-AE$, de modo que no supondremos el axioma de elección salvo que lo indiquemos explícitamente. Veremos que, en ausencia del axioma de elección, incluso el modesto axioma de regularidad contribuye en una pequeña parte a la fundamentación de la teoría de cardinales.

Puesto que hasta ahora no hemos usado en ningún momento el axioma de elección, dedicamos la primera sección a estudiar sus consecuencias principales a nivel general, para que después sea más fácil relacionarlo con la teoría de cardinales en concreto.

13.1 El axioma de elección

Recordemos que el axioma de elección de Zermelo es la sentencia

$$AE \equiv \bigwedge x \bigvee f(f \text{ es una función} \wedge \mathcal{D}f = x \wedge \bigwedge u \in x(u \neq \varnothing \to f(u) \in u)).$$

En definitiva, AE afirma que, dado cualquier conjunto x de conjuntos no vacíos, existe una función que asigna a cada elemento de x uno de sus elementos. A una función de estas características se la llama una *función de elección* sobre x.

Naturalmente no siempre es necesario el axioma de elección para asegurar la existencia de una función de elección. Éste sólo hace falta cuando no tenemos

medios para justificar su existencia mediante los otros axiomas de la teoría de conjuntos.

Por ejemplo, sin necesidad de AE, el conjunto $\mathcal{P}\omega$ tiene una función de elección, la que a cada conjunto no vacío de números naturales le asigna su mínimo. En general $\mathcal{P}x$ tiene una función de elección siempre que existe un buen orden en x.

Otro caso en que podemos asegurar (sin AE) que un conjunto x tiene una función de elección es cuando sus elementos tienen un único elemento cada uno, es decir, si $\bigwedge y \in x \bigvee z \in y\ y = \{z\}$. En tal caso, la (única) función de elección puede definirse mediante

$$f = \{(u, v) \mid u \in x \wedge v \in u\}.$$

Esto puede parecer trivial, pero lo cierto es que algunos matemáticos no excesivamente familiarizados con la teoría de conjuntos tienen dificultades para determinar si una demostración usa o no el axioma de elección, y no es raro que las confusiones provengan de casos análogos a éste aunque no tan explícitos.

Ejercicio: Probar que el axioma de elección equivale a que si x es una familia de conjuntos disjuntos dos a dos, entonces existe un conjunto y que contiene exactamente a un elemento de cada elemento de x.

Una variante de AE consiste en elegir elementos, no de los elementos de un conjunto dado, como afirma el axioma, sino de una familia de conjuntos no vacíos $X = \{X_i\}_{i\in I}$, donde I es un conjunto de índices arbitrario. Es decir, más explícitamente tenemos que $X : I \longrightarrow \mathrm{V}$ es una aplicación cualquiera y llamamos $X_i = X(i)$. La diferencia esencial es que ahora podemos tener repeticiones, es decir, que puede ocurrir que $X_i = X_j$ aunque $i \neq j$, mientras que no tiene sentido decir que un conjunto tiene "elementos repetidos".

Una *función de elección* para una familia de conjuntos X no es sino un elemento de su *producto cartesiano:*

$$\prod_{i\in I} X_i \equiv \{x \mid x : I \longrightarrow \mathrm{V} \wedge \bigwedge i \in I\ x_i \in X_i\}.$$

Así pues, la equivalencia de AE de la que hablamos es la siguiente:

Teorema 13.1 *El axioma de elección equivale a que el producto cartesiano de toda familia no vacía de conjuntos no vacíos es no vacío.*

Demostración: Sea $\{X_i\}_{\in I}$ una familia de conjuntos no vacíos. Aplicamos el axioma de elección al conjunto $\{X_i \mid i \in I\}$, con lo que obtenemos una función de elección f. Basta definir $x_i = f(X_i)$ y tenemos un x en el producto cartesiano.

Recíprocamente, si X es un conjunto cualquiera, es fácil ver que no perdemos generalidad si suponemos que no contiene a $\varnothing$. Una función de elección en X es simplemente un elemento de $\prod_{y\in X} y$. ∎

De nuevo conviene advertir que hay muchas familias de conjuntos de las que podemos probar que su producto cartesiano es no vacío sin necesidad del axioma de elección.

Hay un resultado de uso frecuente que requiere el axioma de elección y es fácil que pase desapercibido en parte por su simplicidad y en parte porque se parece mucho a otros dos resultados que no lo requieren en absoluto.

En efecto, *si* $f : X \longrightarrow Y$ *y* $g : Y \longrightarrow X$ *son dos aplicaciones que cumplen* $f \circ g = I_X$ *(la identidad en* X*), entonces* f *es inyectiva y* g *es suprayectiva.*

Esto es elemental: si $a, b \in X$ cumplen $f(a) = f(b)$, entonces $a = g(f(a)) = g(f(b)) = b$ y si $a \in X$, entonces $f(a)$ es una antiimagen de a por g.

Similarmente, *si* $X \neq \varnothing$ *y* $f : X \longrightarrow Y$ *inyectiva, existe* $g : Y \longrightarrow X$ *(suprayectiva) tal que* $f \circ g = I_X$.

En efecto, basta tomar un $a \in X$ y definir $g(y) = \begin{cases} f^{-1}(y) & \text{si } y \in f[X], \\ a & \text{en otro caso.} \end{cases}$

El resultado que sí requiere el axioma de elección es el siguiente: *Si* Y *es un conjunto y* $g : Y \longrightarrow X$ *suprayectiva, entonces existe* $f : X \longrightarrow Y$ *(inyectiva) tal que* $f \circ g = I_X$.

En efecto, tomamos una función de elección h sobre $\{g^{-1}[x] \mid x \in X\}$ y definimos $f(x) = h(g^{-1}[x])$.

De hecho el axioma de elección es necesario para probar la existencia de una aplicación inyectiva cualquiera $f : X \longrightarrow Y$ a partir de la existencia de una aplicación suprayectiva $g : Y \longrightarrow X$.

Terminamos con las equivalencias más importantes del axioma de elección. Si X es un conjunto parcialmente ordenado, una *cadena* en X es un subconjunto totalmente ordenado por la relación de orden de X. Un conjunto X está *inductivamente ordenado* si $X \neq \varnothing$ y toda cadena en X tiene una cota superior.

Teorema 13.2 *Las sentencias siguientes son equivalentes:*

a) El axioma de elección de Zermelo.

b) El teorema de numerabilidad: $\bigwedge x \bigvee f \alpha\ f : \alpha \longrightarrow x$ *biyectiva.*

c) El teorema de buena ordenación de Zermelo: Todo conjunto puede ser bien ordenado.

d) El lema de Zorn: Todo conjunto inductivamente ordenado tiene un elemento maximal.

e) El lema de Zorn (variante): Todo conjunto de un conjunto inductivamente ordenado está por debajo de un elemento maximal.

Demostración: Probaremos las implicaciones a) → b) → c) → a) por una parte y a) → d) → e) → a) por otra.

a) → b). El argumento es el mismo que empleamos para probar que todo conjunto bien ordenado es semejante a un ordinal. Allí definíamos recurrentemente una función que a cada ordinal le asignaba el mínimo elemento de X

no asignado ya a ordinales menores. Aquí sustituimos el buen orden que nos permite elegir dicho mínimo por una función de elección arbitraria.

Sea $s : \mathcal{P}X \longrightarrow X$ una función de elección (podemos suponer que $s(\varnothing) \in X$). El teorema de recursión transfinita nos da que existe una única función $F : \Omega \longrightarrow X$ tal que $\bigwedge\alpha\ F(\alpha) = s(X \setminus F[\alpha])$.

Si F fuera inyectiva concluiríamos que Ω es un conjunto y tendríamos una contradicción. Podemos tomar, pues, el mínimo ordinal α tal que existe un $\beta < \alpha$ con $F(\beta) = F(\alpha)$. Necesariamente entonces, $f = F|_\alpha : \alpha \longrightarrow X$ es inyectiva.

Por otra parte, f ha de ser suprayectiva, ya que si $f[\alpha] \neq X$, entonces $F(\alpha) = s(X \setminus F[\alpha]) \in X \setminus F[\alpha]$, cuando por otra parte $F(\alpha) = F(\beta) \in F[\alpha]$.

b) → c) Si $f : \alpha \longrightarrow X$ es biyectiva, es claro que

$$\{(f(\beta), f(\gamma)) \mid \beta \leq \gamma < \alpha\}$$

es un buen orden en X.

c) → a) Dado un conjunto X, consideramos un buen orden en $\bigcup_{u\in X} u$. Entonces, una función de elección en X es la que a cada $u \in X$ le asigna su mínimo respecto al buen orden considerado.

a) → d) Sea A un conjunto inductivamente ordenado. Supongamos que para toda cadena B de A se cumple $\bigvee z \in A \bigwedge u \in B\ u < z$. Llegaremos a una contradicción, con lo que concluiremos que existe una cadena B en A tal que $\bigvee z \in A \bigwedge u \in B\ u \leq z$ (toda cadena tiene una cota superior), pero para la cual esto es falso con la desigualdad estricta. De aquí se concluye que la cota z ha de ser un maximal de A, pues si existiera $v \in A$ tal que $z < v$, dicho v sería una cota estricta de B, cuando hemos dicho que no existen.

Para cada $B \subset A$, sea $C_B = \{z \in A \mid \bigwedge u \in B\ u < z\}$. Estamos suponiendo que si B es una cadena en A entonces $C_B \neq \varnothing$. Sea $f : \mathcal{P}A \longrightarrow A$ una función de elección. Por el teorema de recursión existe una única función $F : \Omega \longrightarrow A$ tal que $\bigwedge\alpha\ F(\alpha) = f(C_{F[\alpha]})$.

Veamos por inducción que $\bigwedge\alpha\ F|_\alpha : \alpha \longrightarrow F[\alpha]$ semejanza. Lo suponemos cierto para todo $\beta < \alpha$. Así, si $\beta < \gamma < \alpha$, tenemos que $F[\gamma]$ es una cadena en A, luego $C_{F[\gamma]} \neq \varnothing$, luego $F(\gamma) \in C_{F[\gamma]}$, luego $F(\beta) < F(\gamma)$.

De aquí se sigue que $F : \Omega \longrightarrow A$ es inyectiva, lo cual es absurdo.

d) → e) Sea A un conjunto inductivamente ordenado y sea $a \in A$. Llamemos $B = \{x \in A \mid a \leq x\}$. Es fácil probar que B está inductivamente ordenado, luego tiene un maximal m, que también es maximal en A y está por encima de a.

e) → a) Sea X un conjunto no vacío y sea A el conjunto de las funciones de elección sobre los subconjuntos de X. Claramente A es un conjunto no vacío, pues los subconjuntos de la forma $\{x\}$ tienen obviamente una función de elección, y está inductivamente ordenado por la inclusión, pues la unión de una cadena de funciones de elección es una función de elección. Sea f un maximal

de A y sea $Y \subset A$ su dominio. Basta probar que $Y = A$, pero en caso contrario podemos tomar cualquier $v \in A \setminus Y$ y cualquier $u \in v$, de modo que $f \cup \{(v, u)\}$ es una función de elección por encima de f, en contra de su maximalidad. ∎

13.2 Cardinalidad

Cantor desarrolló su teoría de conjuntos a partir de su descubrimiento de que es posible hablar de la cantidad de elementos de un conjunto infinito de forma completamente análoga a como hablamos habitualmente del número de elementos de un conjunto finito. En esta sección demostramos los teoremas que justifican esta afirmación.

Definición 13.3 Diremos que dos conjuntos X e Y son *equipotentes,* y lo representaremos por $\overline{\overline{X}} = \overline{\overline{Y}}$, si existe una aplicación $f : X \longrightarrow Y$ biyectiva. Diremos que X es *minuspotente* a Y, y lo representaremos por $\overline{\overline{X}} \leq \overline{\overline{Y}}$, si existe $f : X \longrightarrow Y$ inyectiva. Diremos que X es *estrictamente minuspotente* a Y, en signos $\overline{\overline{X}} < \overline{\overline{Y}}$, si $\overline{\overline{X}} \leq \overline{\overline{Y}}$ y no $\overline{\overline{X}} = \overline{\overline{Y}}$.

Observaciones Es importante advertir que no hemos definido un término $\overline{\overline{X}}$, sino únicamente las fórmulas $\overline{\overline{X}} = \overline{\overline{Y}}$, $\overline{\overline{X}} \leq \overline{\overline{Y}}$, $\overline{\overline{X}} < \overline{\overline{Y}}$, es decir, que $\overline{\overline{X}} = \overline{\overline{Y}}$ no es una igualdad de dos términos, sino únicamente una abreviatura de la fórmula $\bigvee f\ f : X \longrightarrow Y$ biyectiva.

Si X es un conjunto ordenado, Cantor representaba por $\overline{X}$ su ordinal, es decir, el concepto resultante de abstraer la naturaleza de los elementos de X pero conservando su ordenación, y por $\overline{\overline{X}}$ su cardinal, es decir, el resultado de abstraer tanto la naturaleza de sus elementos como su ordenación. Así, la doble barra indicaba "doble abstracción". Naturalmente, para nosotros la doble barra se interpreta de acuerdo con la definición precedente y la observación anterior.

Observemos que suponiendo AE tenemos que, para conjuntos no vacíos, $\overline{\overline{X}} \leq \overline{\overline{Y}} \leftrightarrow \bigvee g\ g : Y \longrightarrow X$ suprayectiva. En esta sección no vamos a necesitar este hecho.

El teorema siguiente justifica que las definiciones que hemos dado contienen realmente una noción razonable de "número de elementos" de un conjunto.

Teorema 13.4 *Sean X, Y, Z, W conjuntos cualesquiera. Se cumple:*

a) $\overline{\overline{X}} = \overline{\overline{X}}$,

b) $\overline{\overline{X}} = \overline{\overline{Y}}$ *si y sólo si* $\overline{\overline{Y}} = \overline{\overline{X}}$,

c) Si $\overline{\overline{X}} = \overline{\overline{Y}}$ *y* $\overline{\overline{Y}} = \overline{\overline{Z}}$, *entonces* $\overline{\overline{X}} = \overline{\overline{Z}}$,

d) $\overline{\overline{X}} \leq \overline{\overline{X}}$,

e) Si $\overline{\overline{X}} \leq \overline{\overline{Y}}$ *y* $\overline{\overline{Y}} \leq \overline{\overline{X}}$, *entonces* $\overline{\overline{X}} = \overline{\overline{Y}}$,

f) Si $\overline{\overline{X}} \leq \overline{\overline{Y}}$ *y* $\overline{\overline{Y}} \leq \overline{\overline{Z}}$, *entonces* $\overline{\overline{X}} \leq \overline{\overline{Z}}$,

g) Si $\overline{\overline{X}} = \overline{\overline{Y}}$ *y* $\overline{\overline{Z}} = \overline{\overline{W}}$, *entonces* $\overline{\overline{X}} \leq \overline{\overline{Z}}$ *si y sólo si* $\overline{\overline{Y}} \leq \overline{\overline{W}}$.

Todas estas propiedades excepto e) son consecuencias inmediatas de los hechos básicos sobre aplicaciones entre conjuntos. Debemos insistir en que no deben confundirse, pese a la notación, con teoremas lógicos. Por ejemplo, b) no se cumple por la simetría de la igualdad, ya que no tiene nada que ver con igualdades. Se cumple porque si existe una biyección $f : X \longrightarrow Y$ entonces $f^{-1} : Y \longrightarrow X$ es también una biyección.

Como decimos, la propiedad e) no es evidente en absoluto. Explícitamente, afirma que si existen aplicaciones inyectivas $f : X \longrightarrow Y$ y $g : Y \longrightarrow X$ entonces existe una aplicación biyectiva $h : X \longrightarrow Y$. La forma de construir h a partir de f y g no es inmediata. Cantor demostró este hecho para conjuntos bien ordenados, luego su prueba sólo vale en general si aceptamos el axioma de elección. Al parecer, el primero que probó este hecho sin hacer uso del axioma de elección fue Dedekind, si bien su demostración permaneció inédita hasta 1932. Schröder publicó en 1897 una prueba, pero resultó ser incorrecta, aunque ese mismo año F. Bernstein publicó la primera demostración válida de lo que hoy se conoce como teorema de Cantor-Bernstein. Para probarlo nos apoyaremos en un resultado previo (notemos que, pese a las apariencias, no usa AP).

Teorema 13.5 *Sea X un conjunto y $F : \mathcal{P}X \longrightarrow \mathcal{P}X$ una aplicación tal que si $u \subset v \subset X$ entonces $F(u) \subset F(v)$. Entonces existe un $z \in \mathcal{P}X$ tal que $F(z) = z$.*

Demostración: Sea $A = \{u \in \mathcal{P}X \mid F(u) \subset u\}$. Se cumple que A es una clase no vacía (pues contiene a X). Llamemos $z = \bigcap_{u \in A} u$. Claramente $z \in \mathcal{P}X$ (porque X es un conjunto).

Si $u \in A$, entonces $z \subset u$, luego $F(z) \subset F(u) \subset u$, con lo que $F(z) \subset z$.

Por la hipótesis, $F(F(z)) \subset F(z)$, luego $F(z) \in A$, luego $z \subset F(z)$, lo que nos da la igualdad $F(z) = z$. ∎

Teorema 13.6 (Teorema de Cantor-Bernstein) *Sean X e Y conjuntos tales que existen aplicaciones inyectivas $f : X \longrightarrow Y$ y $g : Y \longrightarrow X$. Entonces existe $h : X \longrightarrow Y$ biyectiva.*

Demostración: Sea $F : \mathcal{P}X \longrightarrow \mathcal{P}X$ la aplicación dada por $F(u) = X \setminus g[Y \setminus f[u]]$. Estamos en las hipótesis del teorema anterior, pues si $u \subset v \subset X$, entonces

$$f[u] \subset f[v], \quad Y \setminus f[v] \subset Y \setminus f[u], \quad g[Y \setminus f[v]] \subset g[Y \setminus f[u]],$$

$$X \setminus g[Y \setminus f[u]] \subset X \setminus g[Y \setminus f[v]],$$

luego $F(u) \subset F(v)$.

En consecuencia existe un subconjunto $z \subset X$ tal que $F(z) = z$, es decir, $X \setminus g[Y \setminus f[z]] = z$ o, equivalentemente, $X \setminus z = g[Y \setminus f[z]]$. Por consiguiente, $f|_z : z \longrightarrow f[z]$ y $g|_{Y\setminus f[z]} : Y \setminus f[z] \longrightarrow X \setminus z$ son ambas biyectivas, luego la unión de la primera con la inversa de la segunda nos da la aplicación h buscada. ∎

Veamos ahora el famoso teorema con que Cantor demostró que existen infinitas potencias infinitas:

Teorema 13.7 (Teorema de Cantor) *Si X es un conjunto, $\overline{\overline{X}} < \overline{\overline{\mathcal{P}X}}$.*

DEMOSTRACIÓN: La aplicación $f : X \longrightarrow \mathcal{P}X$ dada por $f(x) = \{x\}$ es claramente inyectiva, luego $\overline{\overline{X}} \leq \overline{\overline{\mathcal{P}X}}$. Si se diera la igualdad, existiría una aplicación $g : X \longrightarrow \mathcal{P}(X)$ biyectiva, pero entonces podríamos considerar el conjunto $R = \{x \in X \mid x \notin g(x)\} \in \mathcal{P}X$. Sea $r \in X$ tal que $g(r) = R$. Por definición de R tenemos que $r \in R$ si y sólo si $r \notin g(r) = R$, lo cual es una contradicción. ∎

Estos resultados justifican que podemos hablar coherentemente del número de elementos de los conjuntos infinitos y que ello no nos lleva a resultados triviales. Ahora bien, hasta ahora hemos definido lo que significa que un conjunto tenga mayor, menor o igual número elementos que otro, pero no hemos definido lo que debemos entender concretamente por el número de elementos de un conjunto. Más explícitamente, nos gustaría reinterpretar las fórmulas $\overline{\overline{X}} = \overline{\overline{Y}}$ como auténticas igualdades, de modo que podamos leerlas realmente como "el cardinal de X es igual al cardinal de Y", para lo cual hemos de asociar a todo conjunto X un cardinal $\overline{\overline{X}}$ de modo que dos conjuntos tengan el mismo cardinal si y sólo si son equipotentes.

Como en el caso de los ordinales, una forma natural de hacerlo sería definir el cardinal de un conjunto x como la clase de todos los conjuntos equipotentes a x, es decir,

$$\overline{\overline{x}} = \{y \mid \overline{\overline{y}} = \overline{\overline{x}}\}.$$

Esto es teóricamente posible, pero no es conveniente, ya que entonces los cardinales serían clases propias y no podríamos hablar de la clase de todos los cardinales. Una forma de evitar este problema requiere los axiomas de partes y regularidad (sirve la versión con átomos). Consiste en definir el cardinal de un conjunto como el conjunto de todos los conjuntos equipotentes a x *de rango mínimo*. Es decir, tomamos el mínimo ordinal α tal que existe $y \in \mathrm{V}_\alpha$ con $\overline{\overline{x}} = \overline{\overline{y}}$, y definimos el cardinal de x, que representaremos por $\overline{\overline{x}}$, como el conjunto de todos los $y \in \mathrm{V}_\alpha$ equipotentes a x (que es un conjunto por serlo V_α, aquí usamos AP). Explícitamente:

$$\overline{\overline{x}} = \{y \mid \overline{\overline{x}} = \overline{\overline{y}} \wedge \neg\bigvee z(\mathrm{rang}\ z < \mathrm{rang}\ y \wedge \overline{\overline{x}} = \overline{\overline{z}}\ )\}. \tag{13.1}$$

Para que podamos decir que $\overline{\overline{x}}$ es un cardinal, hemos de definir "cardinal" adecuadamente.

Definición 13.8 Un *cardinal* es un conjunto $\mathfrak{p}$ tal que

a) $\bigwedge xy \in \mathfrak{p}(\overline{\overline{x}} = \overline{\overline{y}})$,

b) $\bigvee\alpha\bigwedge x \in \mathfrak{p}\ \text{rang}\ x = \alpha$,

c) $\bigwedge xy(x \in \mathfrak{p} \wedge \overline{\overline{x}} = \overline{\overline{y}} \wedge \text{rang}\ x = \text{rang}\ y \to y \in \mathfrak{p})$,

d) $\neg\bigvee xy(y \in \mathfrak{p} \wedge \overline{\overline{x}} = \overline{\overline{y}} \wedge \text{rang}\ x < \text{rang}\ y)$.

Llamaremos $\mathfrak{C}$ a la clase de todos los cardinales.

Aunque la definición sea técnicamente compleja, esto carece de importancia, pues podemos olvidarnos de ella en cuanto nos convenzamos de lo siguiente:

Teorema 13.9 *Se cumple:*

a) Para cada conjunto x tenemos definido $\overline{\overline{x}} \in \mathfrak{C}$ y para todo $\mathfrak{p} \in \mathfrak{C}$ existe un conjunto x tal que $\overline{\overline{x}} = \mathfrak{p}$.

b) Dados dos conjuntos x e y, se cumple $\overline{\overline{x}} = \overline{\overline{y}}$ si y sólo si x e y son equipotentes.

Así pues, hemos conseguido nuestro propósito: las fórmulas $\overline{\overline{x}} = \overline{\overline{y}}$ tienen el mismo significado que les hemos dado en 13.3, pero ahora son auténticas igualdades.

Veamos que podemos hacer lo mismo con la relación de minuspotencia. Para ello definimos

$$\mathfrak{p} \leq \mathfrak{q} \quad \text{si} \quad \mathfrak{p} = \overline{\overline{x}} \wedge \mathfrak{q} = \overline{\overline{y}} \wedge x \text{ es minuspontente a } y.$$

(Convenimos en que las letras góticas $\mathfrak{p}$, $\mathfrak{q}$, ... denotarán siempre cardinales, aunque no se indique explícitamente). Esta definición no depende de la elección de x e y en virtud del último apartado del teorema 13.4. Además es claro que, para cualquier par de conjuntos x e y, se cumple

$$\overline{\overline{x}} \leq \overline{\overline{y}} \quad \text{si y sólo si} \quad x \text{ es minuspotente a } y,$$

es decir, las fórmulas $\overline{\overline{x}} \leq \overline{\overline{y}}$ tienen el mismo significado que tenían en 13.3, pero ahora son auténticas desigualdades entre cardinales.

El teorema 13.4 implica inmediatamente que la relación que acabamos de definir es ciertamente una relación de orden (no necesariamente de orden total) sobre la clase $\mathfrak{C}$.

Notemos que podríamos haber definido el ordinal de un conjunto ordenado como una clase de equivalencia restringida de conjuntos ordenados semejantes exactamente igual que hemos hecho con los cardinales. Sin embargo, no lo hicimos así porque los ordinales de von Neumann eran una solución mucho más elegante. Von Neumann también definió el concepto de cardinal con esta misma técnica, pero ello requiere el axioma de elección.

En efecto, en la sección anterior hemos visto que AE implica que todo conjunto es equipotente a un ordinal, lo cual nos permite definir el cardinal de un conjunto como el mínimo ordinal equipotente a él. Con más detalle:

Definición 13.10 La clase de los *cardinales de von Neumann* es la clase[1]

$$K = \{\alpha \in \Omega \mid \neg\bigvee\beta < \alpha\ \overline{\beta} = \overline{\alpha}\}.$$

Usaremos las letras griegas κ, μ, ν, ... para referirnos a cardinales de von Neumann, aunque no lo indiquemos explícitamente.

Diremos que un conjunto es *bien ordenable* si admite una buena ordenación o, equivalentemente, si es equipotente a un ordinal. Para cada conjunto bien ordenable x, el menor ordinal equipotente a x no puede ser equipotente a ningún ordinal anterior, luego es un cardinal de von Neumann. En definitiva, si definimos

$$|x| = \kappa \mid (\kappa \in K \wedge \overline{\kappa} = \overline{\overline{x}}\,),$$

Tenemos que para todo conjunto bien ordenable x se cumple que $|x| \in K$ y si x e y son bien ordenables entonces $|x| = |y|$ si y sólo si x es equipotente a y. También es obvio que dos cardinales de von Neumann distintos no pueden ser equipotentes (pues en tal caso uno de ellos sería equipotente a un ordinal anterior, luego no sería un cardinal).

Así pues, si aceptamos AE, todo conjunto tiene asociado un cardinal de von Neumann, luego podemos olvidarnos de la definición de $\mathfrak{C}$ y trabajar únicamente con K (además así no necesitamos AP ni V = R).

Por otra parte, si no suponemos AE, la relación entre ambas definiciones es que tenemos una inmersión $K \longrightarrow \mathfrak{C}$ dada por $\kappa \mapsto \overline{\kappa}$, es decir, a cada cardinal de von Neumann le asociamos su cardinal en el sentido de (13.1). La aplicación es inyectiva, pues si $\overline{\kappa} = \overline{\mu}$, entonces κ y μ son cardinales equipotentes, luego son iguales.

Si esta inmersión es suprayectiva, entonces para todo conjunto x tenemos que $\overline{\overline{x}} = \overline{\kappa}$, para cierto $\kappa \in K$, luego x es equipotente a κ y, por consiguiente, bien ordenable. En suma, la suprayectividad de la inmersión de K en $\mathfrak{C}$ equivale al axioma de elección.

Por otra parte, si consideramos en K el orden de Ω, tenemos que la inmersión es una semejanza en la imagen, es decir, $\kappa \leq \mu$ si y sólo si $\overline{\kappa} \leq \overline{\mu}$.

En efecto, si $\kappa \leq \mu$, entonces $\kappa \subset \mu$, luego es obvio que $\overline{\kappa} \leq \overline{\mu}$. Recíprocamente, si $\overline{\kappa} \leq \overline{\mu}$, no puede ser $\mu < \kappa$, pues entonces $\overline{\mu} \leq \overline{\kappa}$, luego $\overline{\kappa} = \overline{\mu}$, luego $\kappa = \mu$, contradicción. Así pues, $\kappa \leq \mu$.

En particular, si x e y son conjuntos bien ordenables, tenemos que $|x| \leq |y|$ si y sólo si x es minuspotente a y.

Puede probarse que sin el axioma de elección es imposible demostrar que la relación de orden en $\mathfrak{C}$ sea un orden total, mientras que con el axioma de elección

[1] Por seguir la tradición cantoriana, escribiremos $\overline{\alpha}$ en lugar de $\overline{\overline{\alpha}}$ cuando α sea un ordinal. Recordemos que para Cantor una barra significaba "ordinal" y una barra sobre el ordinal (o sea, dos barras sobre un conjunto) significaba "cardinal".

$\mathfrak{C}$ es semejante a K y, por consiguiente, $\mathfrak{C}$ resulta estar, no sólo totalmente ordenado, sino incluso bien ordenado.

En la práctica identificaremos K con su imagen en $\mathfrak{C}$, en el sentido de que si $\mathfrak{p} \in \mathfrak{C}$ y afirmamos que $\mathfrak{p} \in K$ deberemos entender que $\mathfrak{p} = \overline{\kappa}$ para un cierto $\kappa \in K$. Por ejemplo, es obvio que si $\overline{\overline{X}} \leq \overline{\overline{Y}}$ e Y es bien ordenable, entonces X también lo es. Alternativamente, podemos expresar esto diciendo que si $\mathfrak{p} \leq \kappa$, entonces $\mathfrak{p} \in K$.

Veamos ahora algunos resultados básicos sobre los cardinales de von Neumann (recordemos que no usamos el axioma de elección salvo que lo indiquemos explícitamente, y éste no es el caso).

Teorema 13.11 $\omega \subset K$.

Demostración: Probamos por inducción que todo número natural es un cardinal. Obviamente 0 no es equipotente a ningún ordinal anterior, luego $0 \in K$. Supongamos que $n \in K$ pero que $n+1 \notin K$. Entonces existe un ordinal anterior $m < n+1$ y una biyección $f : n+1 \longrightarrow m$. Es claro que m no puede ser 0, luego $m = r+1$. Veamos que podemos suponer que $f(n) = r$. En caso contrario, sea $n' = f^{-1}(r)$. Definimos

$$f' = (f \setminus \{(n, f(n)), (n', r)\}) \cup \{(n, r), (n', f(n))\},$$

y es claro que f' es una biyección como f pero tal que $f(n) = r$.

Ahora bien, $r < n$ y $f'|_n : n \longrightarrow r$ biyectiva, lo cual contradice que n sea un cardinal. ∎

Teorema 13.12 $\omega \in K$.

Demostración: En caso contrario existiría un $n \in \omega$ tal que $\overline{n} = \overline{\omega}$, pero como $n \subset n+1 \subset \omega$ es claro que $\overline{n} \leq \overline{n+1} \leq \overline{\omega} = \overline{n}$, luego sería $\overline{n} = \overline{n+1}$ y $n+1$ no sería un cardinal, en contra del teorema anterior. ∎

El siguiente cardinal ya no es tan fácil de encontrar. Ciertamente no puede ser $\omega + 1$, como muestra el teorema siguiente:

Teorema 13.13 $\bigwedge\kappa(\omega \leq \kappa \to \kappa$ *es un ordinal límite*$)$.

Demostración: Vamos a ver que no puede existir un ordinal α tal que $\kappa = \alpha + 1$. En efecto, en tal caso podríamos definir una aplicación $f : \kappa \longrightarrow \alpha$ biyectiva mediante

$$f(\beta) = \begin{cases} \beta & \text{si } \beta \in \alpha \setminus \omega, \\ \beta + 1 & \text{si } \beta \in \omega, \\ 0 & \text{si } \beta = \alpha. \end{cases}$$

Por consiguiente κ no sería un cardinal. ∎

La forma más natural de encontrar un cardinal mayor que ω es tomar un ordinal equipotente a $\mathcal{P}\omega$ y aplicar el teorema de Cantor. No obstante, no podemos encontrar dicho ordinal sin el axioma de elección, pues sin él no puede probarse que $\mathcal{P}\omega$ pueda ser bien ordenado. Sin embargo, es posible probar la existencia de cardinales arbitrariamente grandes sin necesidad del axioma de elección. El método es esencialmente el mismo que usó Cantor para construir la sucesión de los alefs. En cualquier caso, sigamos el método que sigamos, es importante destacar que necesitaremos sin duda el axioma de partes, pues sin él es imposible demostrar la existencia de conjuntos no numerables. Así, el teorema siguiente es el primero en el que usamos esencialmente *AP*.

Teorema 13.14 $\bigwedge\alpha\bigvee\kappa\ \alpha < \kappa$.

DEMOSTRACIÓN: Sea $A = \{R \in \mathcal{P}(\alpha \times \alpha) \mid R \text{ es un buen orden en } \alpha\}$, es decir, A es el conjunto de todos los buenos órdenes posibles en α. Se cumple que es un conjunto por el axioma de partes.

Sea $f : A \longrightarrow \Omega$ la aplicación dada por $f(R) = \mathrm{ord}(\alpha, R)$. Por el axioma del reemplazo $f[A]$ es un subconjunto de Ω, luego está acotado. Sea $\beta \in \Omega$ tal que $\bigwedge\delta \in f[A]\, \delta < \beta$.

Si R es la relación de orden usual en α, tenemos que $R \in A$ y $f(R) = \alpha$, luego $\alpha < \beta$. Si fuera $\overline{\alpha} = \overline{\beta}$, entonces tendríamos una biyección $g : \alpha \longrightarrow \beta$, la cual nos permitiría definir la relación en α dada por $\delta\, R\, \epsilon$ si y sólo si $g(\delta) < g(\epsilon)$. Claramente R es un buen orden en α y $g : (\alpha, R) \longrightarrow \beta$ es una semejanza. Por consiguiente $f(R) = \beta \in f[A]$, en contradicción con la elección de β.

Así pues, como obviamente $\overline{\alpha} \leq \overline{\beta}$, ha de ser $\overline{\alpha} < \overline{\beta}$.

Llamemos κ al mínimo ordinal tal que $\overline{\alpha} < \overline{\kappa}$. Claramente $\kappa \in K$, pues si existiera un $\gamma < \kappa$ tal que $\overline{\gamma} = \overline{\kappa}$, también tendríamos que $\overline{\alpha} < \overline{\gamma}$, en contra de la definición de κ.

Además $\alpha < \kappa$, pues de lo contrario sería $\overline{\kappa} \leq \overline{\alpha}$, y esto contradice a $\overline{\alpha} < \overline{\kappa}$, por el teorema de Cantor-Bernstein. ∎

Definición 13.15 Dado un ordinal α llamaremos *cardinal siguiente* de α al mínimo cardinal mayor que α y lo representaremos por α^+.

Según hemos visto, $\bigwedge n \in \omega\ n^+ = n+1$, mientras que si α es infinito esto ya no es cierto, pues entonces α^+ es un ordinal límite.

Ahora ya tenemos demostrada la existencia de infinitos cardinales infinitos. Más aún, hemos probado que K no está acotado en Ω, lo que implica que la clase de todos los cardinales no es un conjunto.

Para construir la sucesión de los alefs necesitamos un último resultado:

Teorema 13.16 *El supremo de un conjunto de cardinales es un cardinal.*

DEMOSTRACIÓN: Sea $A \subset K$ un conjunto y sea $\kappa = \bigcup_{\mu \in A} \mu$. Ciertamente $\kappa \in \Omega$ y hemos de probar que es un cardinal. Si existiera un $\alpha < \kappa$ tal que

$\overline{\alpha} = \overline{\kappa}$, entonces existe un $\mu \in A$ tal que $\alpha < \mu$. Entonces

$$\overline{\alpha} \leq \overline{\mu} \leq \overline{\kappa} = \overline{\alpha}.$$

Por consiguiente $\overline{\alpha} = \overline{\mu}$, en contra de que μ sea un cardinal. ∎

Definición 13.17 Llamaremos $\aleph : \Omega \longrightarrow \Omega$ (función alef) a la única aplicación que cumple

$$\aleph_0 = \omega \quad \wedge \quad \bigwedge\alpha\, \aleph_{\alpha+1} = \aleph_\alpha^+ \quad \wedge \quad \bigwedge\lambda\, \aleph_\lambda = \bigcup_{\delta<\lambda} \aleph_\delta.$$

Es claro que se trata de una función normal. Vamos a probar que recorre todos los cardinales infinitos.

Teorema 13.18 *$\aleph : \Omega \longrightarrow K \setminus \omega$ biyectiva.*

Demostración: Como $\aleph$ es normal sabemos que es inyectiva, luego basta probar que es suprayectiva. Una simple inducción demuestra que $\bigwedge\alpha\, \aleph_\alpha \in K$ (el caso límite es el teorema 13.16). Esto significa que $\aleph[\Omega] \subset K$. Como $\aleph$ es creciente y $\aleph_0 = \omega$, ciertamente $\aleph[\Omega] \subset K \setminus \omega$.

Sólo nos falta probar que si $\kappa \in K \setminus \omega$ existe un α tal que $\kappa = \aleph_\alpha$.

Por el teorema 11.22, tenemos que $\kappa \leq \aleph_\kappa < \aleph_{\kappa+1}$. Sea β el mínimo ordinal tal que $\kappa < \aleph_\beta$. No puede ser $\beta = 0$, pues entonces $\kappa \in \aleph_0 = \omega$. Por la definición de $\aleph$ tampoco puede ocurrir que β sea un ordinal límite. Consecuentemente, $\beta = \alpha + 1$ y tenemos que $\aleph_\alpha \leq \kappa < \aleph_{\alpha+1} = \aleph_\alpha^+$. Necesariamente entonces $\kappa = \aleph_\alpha$. ∎

Según esto, tenemos que $\aleph_0 = \omega$, aunque es costumbre no usar las dos notaciones indiscriminadamente, sino que se usa $\aleph_0$ cuando lo consideramos como un cardinal y ω cuando lo consideramos como un ordinal. Similarmente, es costumbre representar $\aleph_\alpha$ como ω_α cuando lo consideramos como un ordinal. De este modo, la sucesión de los cardinales infinitos empieza así:

$$\aleph_0,\ \aleph_1,\ \aleph_2,\ \ldots\ \aleph_\omega,\ \aleph_{\omega+1},\ \aleph_{\omega+2},\ \ldots\ \aleph_{\omega_1},\ \aleph_{\omega_1+1},\ \ldots\ \aleph_{\omega_1+\omega},\ \ldots$$

Es claro que el axioma de elección equivale a que todo cardinal infinito es un aleph. Naturalmente, por cardinal infinito entendemos un cardinal que no sea un número natural. De todos modos, lo cierto es que no hemos definido todavía la finitud:

Definición 13.19 Un conjunto es *finito* si es equipotente a un número natural. Una clase es *infinita* si no es finita.

Un conjunto X es *numerable* si es minuspotente[2] a ω, es decir, si es bien ordenable y $|X| \leq \aleph_0$.

[2]No es infrecuente definir un conjunto numerable como un conjunto *equipotente* a ω, es decir, de cardinal exactamente $\aleph_0$.

El teorema siguiente recoge algunas propiedades básicas de los conjuntos finitos. El primer apartado se demuestra fácilmente por inducción sobre el cardinal. Los otros dos son consecuencias inmediatas del primero:

Teorema 13.20 *a) Si X es un conjunto finito parcialmente ordenado y $a \in X$, entonces existen u, $v \in X$ tales que $u \leq a \leq v$, u es un minimal y v es un maximal.*

b) Si X es un conjunto finito no vacío totalmente ordenado, entonces X tiene máximo y mínimo.

c) Todo conjunto finito totalmente ordenado está bien ordenado.

13.3 La aritmética cardinal

La aritmética cardinal permite reducir el cálculo del cardinal de un conjunto al de otros conocidos. Empezamos estudiando la suma y el producto de cardinales, que, a diferencia de lo que ocurre con la exponenciación, son muy fáciles de calcular. La definición se apoya en el siguiente hecho elemental:

Si X, Y, X', Y' son conjuntos cualesquiera y $\overline{\overline{X}} = \overline{\overline{X'}}$, $\overline{\overline{Y}} = \overline{\overline{Y'}}$, entonces

$$\overline{\overline{X \times \{0\} \cup Y \times \{1\}}} = \overline{\overline{X' \times \{0\} \cup Y' \times \{1\}}}, \qquad \overline{\overline{X \times Y}} = \overline{\overline{X' \times Y'}}.$$

La comprobación no ofrece ninguna dificultad.

Definición 13.21 Dados dos cardinales $\mathfrak{p} = \overline{\overline{X}}$, $\mathfrak{q} = \overline{\overline{Y}}$, definimos

$$\mathfrak{p} + \mathfrak{q} = \overline{\overline{X \times \{0\} \cup Y \times \{1\}}}, \qquad \mathfrak{p}\mathfrak{q} = \overline{\overline{X \times Y}}.$$

La observación anterior justifica que esta definición no depende de la elección de los conjuntos X e Y. Más aún, es fácil probar:

Teorema 13.22 *a) Si X e Y son conjuntos disjuntos, $\overline{\overline{X \cup Y}} = \overline{\overline{X}} + \overline{\overline{Y}}$.*

b) Si X e Y son conjuntos cualesquiera, $\overline{\overline{X \times Y}} = \overline{\overline{X}} \cdot \overline{\overline{Y}}$.

Observemos que si X e Y son conjuntos bien ordenables, también lo son $X \cup Y$ y $X \times Y$. En efecto, si $X \cap Y = \varnothing$, entonces un buen orden en $X \cup Y$ se obtiene lexicográficamente a partir de buenos órdenes en X y en Y, es decir, consideramos el orden para el cual $x \leq y$ si

$$(x \in X \wedge y \in Y) \vee (x \in X \wedge y \in X \wedge x \leq y) \vee (x \in Y \wedge y \in Y \wedge x \leq y).$$

Si la intersección no es vacía observamos que $Y \setminus X$ también es bien ordenable y $X \cup Y = X \cup (Y \setminus X)$, donde ahora la unión es disjunta. Respecto a $X \times Y$, basta considerar el producto lexicográfico de los órdenes de X e Y.

De aquí se sigue que si κ y μ son cardinales de von Neumann podemos definir $\kappa + \mu = |\kappa \times \{0\} \cup \mu \times \{1\}|$ y $\kappa\mu = |\kappa \times \mu|$, con lo que tenemos una suma y

un producto en K que se conservan por la inclusión de K en $\mathfrak{C}$, es decir, que se cumple $\overline{\kappa+\mu}=\overline{\kappa}+\overline{\mu}$, $\overline{\kappa\cdot\mu}=\overline{\kappa}\cdot\overline{\mu}$.

El teorema siguiente se demuestra sin dificultad sin más que manipular de forma obvia aplicaciones entre conjuntos:

Teorema 13.23 *Para todos los cardinales* $\mathfrak{p}$, $\mathfrak{q}$, $\mathfrak{r}$, $\mathfrak{s}$ *se cumple:*

a) $(\mathfrak{p}+\mathfrak{q})+\mathfrak{r}=\mathfrak{p}+(\mathfrak{q}+\mathfrak{r})$,

b) $\mathfrak{p}+\mathfrak{q}=\mathfrak{q}+\mathfrak{p}$,

c) $\mathfrak{p}+0=\mathfrak{p}$,

d) $\mathfrak{p}\le\mathfrak{q}\wedge\mathfrak{r}\le\mathfrak{s}\to\mathfrak{p}+\mathfrak{r}\le\mathfrak{q}+\mathfrak{s}$,

e) $(\mathfrak{p}\mathfrak{q})\mathfrak{r}=\mathfrak{p}(\mathfrak{q}\mathfrak{r})$,

f) $\mathfrak{p}\mathfrak{q}=\mathfrak{q}\mathfrak{p}$,

g) $\mathfrak{p}\cdot 0=0\wedge\mathfrak{p}\cdot 1=\mathfrak{p}$,

h) $\mathfrak{p}(\mathfrak{q}+\mathfrak{r})=\mathfrak{p}\mathfrak{q}+\mathfrak{p}\mathfrak{r}$,

i) $\mathfrak{p}\le\mathfrak{q}\wedge\mathfrak{r}\le\mathfrak{s}\to\mathfrak{p}+\mathfrak{r}\le\mathfrak{q}+\mathfrak{s}\wedge\mathfrak{p}\mathfrak{r}\le\mathfrak{q}\mathfrak{s}$.

Estas propiedades permiten operar fácilmente con cardinales. Veamos un ejemplo:

Teorema 13.24 *Para todo par de conjuntos* X, Y *se cumple*

$$\overline{\overline{X}}+\overline{\overline{Y}}=\overline{\overline{X\cup Y}}+\overline{\overline{X\cap Y}}.$$

En particular $\overline{\overline{X\cup Y}}\le\overline{\overline{X}}+\overline{\overline{Y}}$.

Demostración: Claramente X se descompone en la unión disjunta $X=(X\setminus(X\cap Y))\cup(X\cap Y)$, luego $\overline{\overline{X}}=\overline{\overline{X\setminus(X\cap Y)}}+\overline{\overline{X\cap Y}}$. Por lo tanto

$$\overline{\overline{X}}+\overline{\overline{Y}}=\overline{\overline{X\setminus(X\cap Y)}}+\overline{\overline{Y}}+\overline{\overline{X\cap Y}}=\overline{\overline{(X\setminus(X\cap Y))\cup Y}}+\overline{\overline{X\cap Y}},$$

donde hemos usado que los dos primeros sumandos del término central son disjuntos. Es claro que el último miembro coincide con $\overline{\overline{X\cup Y}}+\overline{\overline{X\cap Y}}$.

La desigualdad se sigue del último apartado del teorema anterior. ∎

No hay que confundir la suma cardinal en K con la suma ordinal. Por ejemplo, la primera es conmutativa y la segunda no. Más explícitamente, $\omega+1$ (suma ordinal) no es un cardinal, luego ha de ser distinto de $\aleph_0+1$ (suma cardinal). La relación entre ambas operaciones es la siguiente:

Teorema 13.25 $\bigwedge\alpha\beta\ |\alpha+\beta|=|\alpha|+|\beta|$, $\quad\bigwedge\alpha\beta\ |\alpha\beta|=|\alpha|\,|\beta|$.

DEMOSTRACIÓN: Hemos de entender que en ambas igualdades la suma de la izquierda es ordinal y la de la derecha cardinal. Al estudiar la suma de ordinales vimos que $\alpha + \beta$ es semejante (en particular equipotente) a $\alpha \times \{0\} \cup \beta \times \{1\}$ con el orden lexicográfico, luego $|\alpha + \beta| = |\alpha \times \{0\} \cup \beta \times \{1\}| = |\alpha| + |\beta|$. Similarmente ocurre con el producto. ∎

Ninguno de los resultados vistos hasta ahora nos permite calcular explícitamente ninguna operación no trivial entre cardinales. Por ejemplo, no sabemos cuánto vale $\aleph_0 + \aleph_1$. Resolvemos este problema con los teoremas siguientes.

Teorema 13.26 *Sobre los números naturales, la suma cardinal coincide con la suma ordinal.*

DEMOSTRACIÓN: Es consecuencia inmediata del teorema anterior, teniendo en cuenta que todo $n \in \omega$ cumple $|n| = n$. ∎

La suma y el producto en K quedan completamente determinados por el teorema siguiente:

Teorema 13.27 *Para todo alef κ se cumple $\kappa\kappa = \kappa$.*

DEMOSTRACIÓN: Lo probamos por inducción, es decir, suponemos que para todo aleph $\mu < \kappa$ se cumple $\mu\mu = \mu$. Entonces, $\bigwedge \mu < \kappa\ \mu\mu < \kappa$, pues μ ha de ser un alef o bien un número natural.

Consideramos en $\kappa \times \kappa$ la restricción del orden canónico de $\Omega \times \Omega$ definido en 11.28.

Sea $\alpha = \text{ord}\,(\kappa \times \kappa)$. Entonces $\alpha \geq |\alpha| = \kappa\kappa \geq \kappa$. Supongamos que fuera $\kappa < \alpha$ y sea $f : \alpha \longrightarrow \kappa \times \kappa$ la semejanza. Sea $f(\kappa) = (\beta, \gamma)$. Como κ es un ordinal límite, podemos tomar $\delta < \kappa$ tal que β, $\gamma < \delta$.

Como κ está formado por los ordinales menores que κ, tenemos que $f[\kappa]$ está formado por los pares menores que (β, γ). Ahora bien, por la definición del orden canónico, si $(\delta, \epsilon) < (\beta, \gamma)$, entonces δ, $\epsilon < \delta$, es decir, $f[\kappa] \subset \delta \times \delta$.

Por consiguiente, $\kappa = |f[\kappa]| \leq |\delta \times \delta| = |\delta|\,|\delta| < \kappa$, contradicción. Por consiguiente $\alpha = \kappa$, lo cual prueba que $\kappa \times \kappa$ es equipotente a κ. ∎

Como consecuencia:

Teorema 13.28 *Se cumple:*

$$\bigwedge \kappa\mu(\kappa \leq \mu \wedge \aleph_0 \leq \mu \to \kappa + \mu = \mu),$$

$$\bigwedge \kappa\mu(\kappa \leq \mu \wedge \aleph_0 \leq \mu \wedge \kappa \neq 0 \to \kappa\mu = \mu).$$

DEMOSTRACIÓN: $\mu \leq \kappa + \mu \leq \mu + \mu = 2\mu \leq \mu\mu = \mu$, luego $\kappa + \mu = \mu$.
$\mu \leq \kappa\mu \leq \mu\mu = \mu$, luego $\kappa\mu = \mu$. ∎

Así pues, la aritmética de K es muy sencilla:

$$\aleph_0 + \aleph_1 = \aleph_1, \quad \aleph_{\omega_{15}} + \aleph_3 = \aleph_{\omega_{15}}, \quad 3\aleph_7 = \aleph_7, \quad \aleph_{23}\aleph_7 = \aleph_{23}, \quad \text{etc.}$$

Ejercicio: Probar que si Y es un conjunto infinito bien ordenable y $|X| < |Y|$, entonces $|Y \setminus X| = |Y|$.

Respecto a la aritmética de los cardinales no bien ordenables, poco podemos decir. Un concepto útil en su estudio es el siguiente:

Definición 13.29 Sea X un conjunto infinito. Sea

$$B = \{R \mid R \text{ es un buen orden en un subconjunto de } X\}.$$

Se cumple que B es un conjunto porque $B \subset \mathcal{P}(X \times X)$. Llamaremos *número de Hartogs* de X a $\aleph(X) = \{\mathrm{ord}(\mathcal{D}R, R) \mid R \in B\}$.

Como $\aleph(X)$ es imagen de B, por el axioma del reemplazo es un conjunto de ordinales. Es claro que $\alpha \in \aleph(X)$ si y sólo si existe $f : \alpha \longrightarrow X$ inyectiva, de donde se sigue claramente que $\aleph(X)$ es un conjunto transitivo y, por consiguiente, un ordinal.

Más aún, si $|\alpha| = |\beta|$ y $\beta < \aleph(X)$, entonces $\alpha < \aleph(X)$, de donde se sigue que $\aleph(X)$ es, de hecho, un cardinal.

Una simple inducción prueba que si X es infinito existe $f : n \longrightarrow X$ inyectiva para todo n, luego $\aleph(X)$ es un cardinal infinito, es decir, un alef.

También es inmediato que si $\overline{\overline{X}} = \overline{\overline{Y}}$ entonces $\aleph(X) = \aleph(Y)$, luego, para cada cardinal $\mathfrak{p}$, podemos definir $\aleph(\mathfrak{p}) = \aleph(X)$, donde X es cualquier conjunto tal que $\overline{\overline{X}} = \mathfrak{p}$.

Es claro que $\aleph(\mathfrak{p})$ es el menor alef κ que no cumple $\kappa \leq \mathfrak{p}$ (notemos que si fuera $\aleph(\mathfrak{p}) \leq \mathfrak{p}$ entonces tendríamos $\aleph(\mathfrak{p}) < \aleph(\mathfrak{p})$). En particular, si κ es un alef, se cumple $\aleph(\kappa) = \kappa^+$.

Teorema 13.30 *Sean $\mathfrak{p}$ y κ cardinales infinitos tales que $\mathfrak{p} + \kappa = \mathfrak{p}\kappa$. Entonces $\mathfrak{p} \leq \kappa$ o $\kappa \leq \mathfrak{p}$. En particular, si $\mathfrak{p} + \aleph(\mathfrak{p}) = \mathfrak{p}\aleph(\mathfrak{p})$, entonces $\mathfrak{p}$ es un alef.*

Demostración: Sea X un conjunto de cardinal $\mathfrak{p}$. Por hipótesis existen conjuntos disjuntos A y B tales que $X \times \kappa = A \cup B$, $\overline{\overline{A}} = \mathfrak{p}$, $\overline{\overline{B}} = \kappa$.

Si existe un $x \in X$ tal que $\{(x, \alpha) \mid \alpha < \kappa\} \subset A$, entonces claramente $\kappa \leq \mathfrak{p}$.

En caso contrario, para cada $x \in X$ existe un mínimo $\alpha_x \in \kappa$ tal que $(x, \alpha_x) \notin A$, de donde $\{(x, \alpha_x) \mid x \in X\} \subset B$, por lo que $\mathfrak{p} \leq \kappa$.

En el caso particular en que $\kappa = \aleph(\mathfrak{p})$ no puede ocurrir $\aleph(\mathfrak{p}) \leq \mathfrak{p}$, luego ha de ser $\mathfrak{p} \leq \aleph(\mathfrak{p})$ y, por consiguiente, $\mathfrak{p}$ es un alef. ∎

Como aplicación tenemos un resultado interesante:

Teorema 13.31 *El axioma de elección equivale a que $\mathfrak{p}\mathfrak{p} = \mathfrak{p}$ para todo cardinal infinito $\mathfrak{p}$.*

Demostración: Si suponemos el axioma de elección entonces todo cardinal infinito es un alef y basta aplicar el teorema 13.27. Para el recíproco basta probar que todo cardinal infinito $\mathfrak{p}$ es un alef y, a su vez, para ello basta probar que

$\mathfrak{p} + \aleph(\mathfrak{p}) = \mathfrak{p}\aleph(\mathfrak{p})$. De hecho basta ver que $\mathfrak{p}\aleph(\mathfrak{p}) \leq \mathfrak{p} + \aleph(\mathfrak{p})$, ya que la otra desigualdad se da siempre trivialmente. Ahora bien:

$$\mathfrak{p} + \aleph(\mathfrak{p}) = (\mathfrak{p} + \aleph(\mathfrak{p}))(\mathfrak{p} + \aleph(\mathfrak{p})) = \mathfrak{p}\mathfrak{p} + 2\mathfrak{p}\aleph(\mathfrak{p}) + \aleph(\mathfrak{p})\aleph(\mathfrak{p}) \geq \mathfrak{p}\aleph(\mathfrak{p}).$$

■

Finalmente introducimos la exponenciación de cardinales, una operación tan natural como la suma y el producto pero cuyo comportamiento es muy diferente.

Recordemos que $A^B = \{f \mid f : B \longrightarrow A\}$. La definición de la exponenciación de cardinales se apoya en el siguiente hecho obvio:

Si A, A', B y B' son conjuntos tales que $\overline{\overline{A}} = \overline{\overline{A'}}$ y $\overline{\overline{B}} = \overline{\overline{B'}}$ entonces se cumple $\overline{\overline{A^B}} = \overline{\overline{A'^{B'}}}$.

Definición 13.32 Dados dos cardinales $\mathfrak{p}$ y $\mathfrak{q}$, definimos $\mathfrak{p}^\mathfrak{q} = \overline{\overline{A^B}}$, donde $\overline{\overline{A}} = \mathfrak{p}$ y $\overline{\overline{B}} = \mathfrak{q}$.

La observación precedente demuestra que $\mathfrak{p}^\mathfrak{q}$ no depende de la elección de los conjuntos A y B. Las propiedades siguientes se demuestran sin dificultad:

Teorema 13.33 *Para todos los cardinales $\mathfrak{p}$, $\mathfrak{q}$, $\mathfrak{r}$ se cumple:*

a) $\mathfrak{p} \neq 0 \to 0^\mathfrak{p} = 0$,

b) $\mathfrak{p}^0 = 1, \quad 1^\mathfrak{p} = 1, \quad \mathfrak{p}^1 = \mathfrak{p}$,

c) $\mathfrak{q} \leq \mathfrak{r} \to \mathfrak{p}^\mathfrak{q} \leq \mathfrak{p}^\mathfrak{r}$,

d) $\mathfrak{p} \neq 0 \wedge \mathfrak{q} \neq 0 \wedge \mathfrak{q} \leq \mathfrak{r} \to \mathfrak{p}^\mathfrak{q} \leq \mathfrak{p}^\mathfrak{r}$,

e) $\mathfrak{p}^{\mathfrak{q}+\mathfrak{r}} = \mathfrak{p}^\mathfrak{q}\mathfrak{p}^\mathfrak{r}$,

f) $(\mathfrak{p}\mathfrak{q})^\mathfrak{r} = \mathfrak{p}^\mathfrak{r}\mathfrak{q}^\mathfrak{r}$.

Así mismo, una simple inducción basada en estas propiedades demuestra que la exponenciación cardinal sobre los números naturales coincide con la exponenciación ordinal. Otro resultado notable es el siguiente:

Teorema 13.34 *Para todo conjunto X, se cumple $\overline{\overline{\mathcal{P}X}} = 2^{\overline{\overline{X}}}$.*

Demostración: Basta observar que la aplicación $f : 2^X \longrightarrow \mathcal{P}X$ dada por $h \mapsto h^{-1}[\{1\}]$ es biyectiva. ■

Como consecuencia, el teorema de Cantor admite una formulación aritmética:

Teorema 13.35 (Teorema de Cantor) *Para todo cardinal $\mathfrak{p}$ se cumple*

$$\mathfrak{p} < 2^\mathfrak{p}.$$

Observaciones El hecho de que la exponenciación cardinal esté tan vinculada al operador $\mathcal{P}$ hace que los axiomas de la teoría de conjuntos dejen sin decidir los hechos más importantes sobre la misma. En efecto, dichos axiomas se limitan a imponer la existencia de los conjuntos que un matemático necesita, es decir, garantizan la existencia de uniones, intersecciones, de los conjuntos de números, etc., pero no dicen nada sobre qué clase de cosa es un conjunto y, en particular, no dicen nada sobre qué contiene $\mathcal{P}X$, ni de lo grande o pequeño que pueda ser este conjunto. Tal y como explicábamos en la introducción, el teorema de Cantor, al igual que una ligera generalización que veremos más adelante, no es más que una contradicción atenuada, es decir, al postular la existencia de un conjunto que no existe, como es $\mathcal{P}X$, podríamos haber llegado a una contradicción, como sucede al postular la existencia del conjunto de todos los conjuntos pero, afortunadamente, sólo llegamos a una mentira: su diferencia de tamaño.

Como muestra de lo "reacia" que es la teoría de conjuntos a pronunciarse sobre la exponenciación cardinal, observamos que sin el axioma de elección no podemos asegurar que A^B o incluso $\mathcal{P}A$ sean bien ordenables aunque A y B lo sean. Esto hace que no podamos desarrollar una aritmética de la exponenciación de K que no requiera el axioma de elección, tal y como hemos hecho con la suma y el producto. El hecho de que cualquier resultado no trivial requiera el axioma de elección es una muestra de que, por muy bien que conozcamos los conjuntos A y B, en realidad no sabemos casi nada de A^B (y lo mismo vale para $\mathcal{P}X$). De todos modos, aunque el axioma de elección hace que la exponenciación cardinal tenga un comportamiento bastante razonable, lo cierto es que no resuelve en absoluto las cuestiones centrales en torno a ella. Por ejemplo, recordemos que la hipótesis del continuo es la conjetura de Cantor según la cual $2^{\aleph_0} = \aleph_1$. Sin el axioma de elección hemos probado que $2^{\aleph_0} > \aleph_0$ y, con el axioma de elección (¡pero no sin él!), lo único que podemos añadir a esto es que $2^{\aleph_0} \geq \aleph_1$. Naturalmente, el problema es idéntico para cardinales mayores.

Definición 13.36 Se llama *hipótesis del continuo generalizada* a la siguiente sentencia:

(HCG) Si $\mathfrak{p}$ es un cardinal infinito, no existe ningún cardinal $\mathfrak{q}$ tal que

$$\mathfrak{p} < \mathfrak{q} < 2^{\mathfrak{p}}.$$

Con el axioma de elección esto equivale a que $2^\kappa = \kappa^+$ para todo cardinal infinito κ o, equivalentemente, a que

$$\bigwedge \alpha \; 2^{\aleph_\alpha} = \aleph_{\alpha+1}.$$

Veremos que la HCG determina completamente la exponenciación cardinal, si bien dista mucho de ser un teorema de ZFC. Dejamos esto —junto a estudio en profundidad de la exponenciación cardinal— para el capítulo siguiente. De momento terminaremos esta sección demostrando un hecho sorprendente: La hipótesis del continuo generalizada implica el axioma de elección. Esto fue anunciado por Hausdorff, si bien la primera prueba publicada fue de Sierpinski.

La demostración que veremos aquí es posterior. Necesitamos algunos resultados previos.

En primer lugar, sin el axioma de elección hemos probado que, para todo ordinal infinito α, se cumple $|\alpha \times \alpha| = |\alpha|$. Ahora necesitamos construir, sin el axioma de elección, una aplicación que a cada ordinal infinito α le asigne una biyección $f_\alpha : \alpha \times \alpha \longrightarrow \alpha$. Por ejemplo,la prueba de 13.27 muestra que si α es un cardinal entonces $\alpha \times \alpha$ con el orden canónico es semejante a α, luego si nos bastara trabajar con cardinales podríamos definir f_α como la única semejanza entre $\alpha \times \alpha$ y α. El problema es que necesitamos esto para cualquier ordinal $\alpha \geq \omega$. Resolveremos esto en varios pasos.

a) *Para cada par de ordinales α y β, podemos definir explícitamente una biyección entre $\alpha + \beta$ y $\beta + \alpha$.*

   En efecto, sabemos definir biyecciones explícitas entre $\alpha + \beta$ y $\alpha \times \{0\} \cup \beta \times \{1\}$ y $\beta + \alpha$ y $\beta \times \{0\} \cup \alpha \times \{1\}$, a saber, las semejanzas cuando consideramos el orden lexicográfico en los conjuntos de pares. Ahora bien, es inmediato cómo definir explícitamente una biyección entre estos dos conjuntos de pares, luego podemos construir la biyección buscada.

b) *Para todos los ordinales $\eta_0, \ldots, \eta_n$ sabemos definir una biyección explícita entre $\omega^{\eta_0} + \cdots + \omega^{\eta_n}$ y $\omega^{\eta_n} + \cdots + \omega^{\eta_0}$*

   En efecto, basta aplicar repetidas veces el apartado anterior.

c) *Si $\alpha = \omega^{\eta_0} k_0 + \cdots + \omega^{\eta_n} k_n$ es la forma normal de Cantor del ordinal α, sabemos definir una biyección entre α y $\omega^{\eta_0} k_0$.*

   En efecto, por el apartado anterior sabemos invertir el orden de los sumandos, y por el teorema 11.43 la suma en orden inverso es $\omega^{\eta_0} k_0$.

d) *En las condiciones del apartado anterior, sabemos definir una biyección entre α y ω^{η_0}.*

   En efecto, razonando como en el apartado a) pero para el producto en lugar de la suma sabemos definir una biyección entre $\omega^{\eta_0} k_0$ y $k_0 \omega^{\eta_0} = \omega^{\eta_0}$.

   Así pues, si llamamos η_α al exponente director de la forma normal de α, sabemos definir explícitamente una biyección entre α y ω^{η_α}.

e) *Si α es un ordinal infinito, entonces $\eta_\alpha = \eta_{\alpha+\alpha}$.*

   En efecto, $\alpha + \alpha = \alpha \cdot 2$, por lo que la forma normal de $\alpha + \alpha$ se diferencia de la de α en que sus coeficientes están multiplicados por 2 (pero los exponentes son idénticos).

f) *Sabemos definir una biyección entre cada ordinal infinito α y $\alpha + \alpha$*

   Basta biyectar α con $\omega^{\eta_\alpha} = \omega^{\eta_{\alpha+\alpha}}$ y éste con $\alpha + \alpha$.

g) *Cada elemento no nulo de ω^η se expresa de forma única como*

$$\delta = \omega^{\eta_0}k_0 + \cdots + \omega^{\eta_n}k_n,$$

donde $\eta > \eta_0 > \cdots > \eta_n$ y los k_i son naturales no nulos.

En efecto, si δ es de esta forma tenemos que

$$\delta \le \omega^{\eta_0}k_0 + \cdots + \omega^{\eta_0}k_n = \omega^{\eta_0}(k_0 + \cdots + k_n) < \omega^{\eta_0+1} \le \omega^\eta.$$

Recíprocamente, si $\delta < \omega^\eta$, el exponente η_0 de su forma normal ha de ser necesariamente menor que η. La unicidad es la de la forma normal.

h) *Sabemos definir una biyección entre ω^η y $\omega^{\eta+\eta}$.*

Si η es infinito, por el apartado anterior, cada elemento no nulo de ω^η está determinado por una aplicación de un subconjunto finito arbitrario de η en ω. La aplicación de f) nos permite biyectar los subconjuntos finitos de η con los de $\eta + \eta$, luego sabemos biyectar los elementos no nulos de ω^η con los elementos no nulos de $\omega^{\eta+\eta}$.

Si η es finito (no nulo), es fácil definir explícitamente la biyección partiendo de una biyección entre ω y $\omega \times \omega$.

i) *Sabemos biyectar α con $\alpha \times \alpha$.*

Basta biyectar α con ω^{η_α}, éste con $\omega^{\eta_\alpha+\eta_\alpha} = \omega^{\eta_\alpha}\cdot\omega^{\eta_\alpha}$, éste con $\omega^{\eta_\alpha}\times\omega^{\eta_\alpha}$ y éste con $\alpha \times \alpha$.

El siguiente paso es probar lo que sin el axioma de elección es una leve generalización del teorema de Cantor:

Teorema 13.37 *Si $\mathfrak{p} \ge 5$ entonces no $2^\mathfrak{p} \le \mathfrak{p}^2$.*

Demostración: Para cardinales finitos se demuestra fácilmente por inducción que $\mathfrak{p}^2 < 2^\mathfrak{p}$. Supongamos, pues, que $\mathfrak{p}$ es un cardinal infinito y sea $\overline{\overline{X}} = \mathfrak{p}$. Entonces tenemos una aplicación $f : \mathcal{P}X \longrightarrow X \times X$ inyectiva.

Vamos a construir una aplicación $G : \Omega \longrightarrow X$ inyectiva, con lo que tendremos una contradicción. En primer lugar veremos que podemos construir $g_\omega : \omega \longrightarrow X$ inyectiva.

Es fácil ver que el conjunto de todos los subconjuntos finitos de ω es biyectable con ω, por lo que podemos fijar un buen orden en él. Tomemos elementos distintos $x_0, \ldots, x_4 \in X$ y definamos $g_\omega(i) = x_i$, para $i < 5$.

Supuesta definida $g_\omega|n : n \longrightarrow X$ inyectiva, para $n \ge 5$, sea $C_n = g_\omega[n]$. Como $|\mathcal{P}C_n| = 2^n > n^2 = |C_n \times C_n|$, existe un subconjunto U de C_n tal que $f(U) \notin C_n \times C_n$. Elegimos el que cumple que $g_\omega^{-1}[U]$ es mínimo respecto al buen orden que hemos fijado en los subconjuntos finitos de ω. Si $f(U) = (x, y)$, definimos $g_\omega(n+1) = x$ si $x \notin C_n$ y $g_\omega(n+1) = y$ en caso contrario. Con esto

tenemos que $g_\omega|_{n+1} : n+1 \longrightarrow X$ es inyectiva. El teorema de recursión nos garantiza la existencia de g_ω.

Pasemos ahora a la construcción de $G : \Omega \longrightarrow X$. Para ello nos apoyaremos en las biyecciones $f_\alpha : \alpha \longrightarrow \alpha \times \alpha$ que hemos definido para todo ordinal infinito α sin el axioma de elección. Suponemos definida $G|_\alpha : \alpha \longrightarrow X$ inyectiva. Sea $C_\alpha = G[\alpha]$.

Definimos $g : \alpha \longrightarrow \mathcal{P}X$ como sigue: dado $\beta < \alpha$ calculamos $f_\alpha(\beta) = (\gamma, \delta)$ y tomamos $g(\beta) = f^{-1}(G(\gamma), G(\delta))$ si el par $(G(\gamma), G(\delta))$ tiene antiimagen por f y $g(\beta) = \varnothing$ en caso contrario.

Sea $U = \{G(\beta) \mid \beta < \alpha \wedge G(\beta) \notin g(\beta)\}$. Sea $f(U) = (x, y)$.

Si $(x, y) \in C_\alpha \times C_\alpha$ entonces $(x, y) = (G(\gamma), G(\delta))$ para ciertos γ, $\delta < \alpha$. Sea $\beta = f_\alpha^{-1}(\gamma, \delta)$, de modo que $g(\beta) = U$ y tenemos una contradicción tanto si $G(\beta) \in U$ como en caso contrario. Por consiguiente $(x, y) \notin C_\alpha \times C_\alpha$, luego podemos definir $G(\alpha) = x$ si $x \notin C_\alpha$ o $G(\alpha) = y$ en caso contrario.

El teorema de recursión transfinita nos da entonces la existencia de G. ∎

Nota Sin el axioma de elección no puede probarse en general que $\mathfrak{p}^2 \leq 2^\mathfrak{p}$.

Teorema 13.38 *La hipótesis generalizada del continuo implica el axioma de elección.*

Demostración: Por el teorema 13.31, basta probar que $\mathfrak{p}^2 = \mathfrak{p}$ para todo cardinal infinito $\mathfrak{p}$.

En primer lugar probamos que $\mathfrak{p} = \mathfrak{p} + 1$.

Es fácil ver que $\mathfrak{p} \leq \mathfrak{p} + 1 \leq 2^\mathfrak{p}$, pero si fuera $\mathfrak{p} + 1 = 2^\mathfrak{p}$, tendríamos que $2^\mathfrak{p} \leq \mathfrak{p} + 1 \leq \mathfrak{p} + \mathfrak{p} \leq \mathfrak{p}\mathfrak{p}$, en contradicción con el teorema anterior. Así pues, la HCG implica que $\mathfrak{p} = \mathfrak{p} + 1$

Ahora veamos que $\mathfrak{p} = 2\mathfrak{p}$.

En efecto, $\mathfrak{p} \leq 2\mathfrak{p} \leq 2 \cdot 2^\mathfrak{p} = 2^{\mathfrak{p}+1} = 2^\mathfrak{p}$, pero no puede ser $2\mathfrak{p} = 2^\mathfrak{p}$ ya que entonces $2^\mathfrak{p} = 2\mathfrak{p} \leq \mathfrak{p}\mathfrak{p}$, de nuevo en contra del teorema anterior. La HCG nos da, pues, la igualdad $\mathfrak{p} = 2\mathfrak{p}$.

Así, $\mathfrak{p} \leq \mathfrak{p}^2 \leq (2^\mathfrak{p})^2 = 2^{2\mathfrak{p}} = 2^\mathfrak{p}$. El teorema anterior y la HCG nos dan la igualdad $\mathfrak{p}^2 = \mathfrak{p}$. ∎

13.4 Sumas y productos infinitos

El cálculo explícito del cardinal de determinados conjuntos requiere considerar sumas y productos infinitos de otros cardinales conocidos. Prácticamente todos los resultados sobre estas sumas y productos dependen del axioma de elección, pues cuando tenemos infinitos conjuntos a menudo es imprescindible escoger una biyección entre cada uno de ellos y su cardinal. Así pues, en esta sección usaremos libremente dicho axioma sin mención explícita.

Definición 13.39 La *suma* de una familia de cardinales $\{\kappa_i\}_{i\in I}$ se define como

$$\sum_{i\in I}\kappa_i = \Big|\bigcup_{i\in I}\kappa_i\times\{i\}\Big|.$$

El resultado fundamental sobre sumas infinitas es el siguiente:

Teorema 13.40 *Para cualquier familia de conjuntos $\{X_i\}_{i\in I}$ se cumple que*

$$\Big|\bigcup_{i\in I}X_i\Big| \le \sum_{i\in I}|X_i|,$$

y si los conjuntos son disjuntos dos a dos entonces se da la igualdad.

Demostración: Por el axioma de elección existe una familia de aplicaciones biyectivas $f_i : |X_i|\times\{i\} \longrightarrow X_i$. Claramente

$$\bigcup_{i\in I} f_i : \bigcup_{i\in I}|X_i|\times\{i\} \longrightarrow \bigcup_{i\in I}X_i$$

es suprayectiva y si los conjuntos X_i son disjuntos dos a dos es biyectiva. Consecuentemente

$$\Big|\bigcup_{i\in I}X_i\Big| \le \Big|\bigcup_{i\in I}|X_i|\times\{i\}\Big| = \sum_{i\in I}|X_i|,$$

y se da la igualdad si los conjuntos son disjuntos. ∎

El teorema siguiente se demuestra sin dificultad:

Teorema 13.41 *Se cumple*

a) *Si $\bigwedge i\in I\ \kappa_i \le \mu_i$, entonces $\sum_{i\in I}\kappa_i \le \sum_{i\in I}\mu_i$,*

b) $\sum_{i\in I}\kappa = |I|\,\kappa$,

c) $\mu\sum_{i\in I}\kappa_i = \sum_{i\in I}\mu\kappa_i$.

A modo de ejemplo demostraremos la asociatividad generalizada de la suma de cardinales:

Teorema 13.42 *Si $I = \bigcup_{j\in J} I_j$ y los conjuntos I_j son disjuntos dos a dos, entonces*

$$\sum_{i\in I}\kappa_i = \sum_{j\in J}\sum_{i\in I_j}\kappa_i.$$

Demostración: En efecto:

$$\sum_{i\in I}\kappa_i = \Big|\bigcup_{i\in I}\kappa_i\times\{i\}\Big| = \Big|\bigcup_{j\in J}\bigcup_{i\in I_j}\kappa_i\times\{i\}\Big| = \sum_{j\in J}\Big|\bigcup_{i\in I_j}\kappa_i\times\{i\}\Big| = \sum_{j\in J}\sum_{i\in I_j}\kappa_i.$$

∎

Finalmente demostramos el teorema que nos permite calcular cualquier suma de cardinales. Las hipótesis excluyen el caso de una suma finita de cardinales finitos, pero esto es una suma usual de números naturales.

Teorema 13.43 *Si $\{\kappa_i\}_{i\in I}$ es una familia de cardinales no nulos de modo que I es infinito o algún κ_i lo es, entonces*

$$\sum_{i\in I}\kappa_i = |I|\ \sup_{i\in I}\kappa_i.$$

DEMOSTRACIÓN: Llamemos κ al supremo de los κ_i. Como los κ_i son no nulos tenemos que $1 \le \kappa_i \le \kappa$, luego

$$|I| = \sum_{i\in I} 1 \le \sum_{i\in I}\kappa_i \le \sum_{i\in I}\kappa = |I|\,\kappa.$$

Como cada $\kappa_i \le \sum_{i\in I}\kappa_i$, también $\kappa \le \sum_{i\in I}\kappa_i$.

Multiplicando las desigualdades (y teniendo en cuenta que, por las hipótesis, la suma es un cardinal infinito) obtenemos

$$|I|\,\kappa \le \Big(\sum_{i\in I}\kappa_i\Big)^2 = \sum_{i\in I}\kappa_i.$$

■

Pasemos ahora a estudiar los productos infinitos. No podemos obtener resultados tan concluyentes como los que hemos obtenido para las sumas debido a su proximidad a la exponenciación cardinal. Recordemos la definición del producto cartesiano de una familia de conjuntos:

$$\prod_{i\in I} X_i = \{f \mid f : I \longrightarrow \bigcup_{i\in I} X_i \wedge \bigwedge i \in I\ f(i) \in X_i\}.$$

El producto cartesiano de un conjunto de conjuntos es un conjunto porque está contenido en $\mathcal{P}(I \times \bigcup_{i\in I} X_i)$.

Definición 13.44 Llamaremos *producto* de una familia de cardinales $\{\kappa_i\}_{i\in I}$ al cardinal

$$\prod_{i\in I}\kappa_i = \Big|\prod_{i\in I}\kappa_i\Big|,$$

donde el producto de la izquierda es el que estamos definiendo y el de la derecha es el producto cartesiano.

El resultado básico sobre productos infinitos es el siguiente:

Teorema 13.45 *Si $\{X_i\}_{i\in I}$ es una familia de conjuntos infinitos, entonces*

$$\Big|\prod_{i\in I} X_i\Big| = \prod_{i\in I}|X_i|.$$

DEMOSTRACIÓN: Por el axioma de elección existe una familia de aplicaciones biyectivas $f_i : X_i \longrightarrow |X_i|$. Entonces la aplicación $f : \prod_{i\in I} X_i \longrightarrow \prod_{i\in I}|X_i|$ dada por $f(\{x_i\}_{i\in I}) = \{f(x_i)\}_{i\in I}$ es claramente biyectiva. Así pues,

$$\Big|\prod_{i\in I} X_i\Big| = \Big|\prod_{i\in I}|X_i|\Big| = \prod_{i\in I}|X_i|.$$

■

Recogemos en el teorema siguiente las propiedades sencillas de los productos:

Teorema 13.46 *Se cumple:*

a) *Si algún* $\kappa_i = 0$, *entonces* $\prod_{i\in I} \kappa_i = 0$,

b) $\prod_{i\in I} \kappa = \kappa^{|I|}$,

c) $\Big(\prod_{i\in I} \kappa_i\Big)^\mu = \prod_{i\in I} \kappa_i^\mu$,

d) $\prod_{i\in I} \kappa^{\mu_i} = \kappa^{\sum_{i\in I} \mu_i}$,

e) *Si* $\bigwedge i \in I\ \kappa_i \le \mu_i$, *entonces* $\prod_{i\in I} \kappa_i \le \prod_{i\in I} \mu_i$,

f) *Si* $I = \bigcup_{j\in J} I_j$, *donde los conjuntos* I_j *son disjuntos dos a dos, entonces*

$$\prod_{i\in I} \kappa_i = \prod_{j\in J} \prod_{i\in I_j} \kappa_i.$$

No es posible demostrar un teorema tan general como 13.43 para el cálculo de productos infinitos, pero a menudo basta el teorema siguiente:

Teorema 13.47 *Sea* $\{\kappa_\alpha\}_{\alpha<\mu}$ *una familia de cardinales no nulos (donde* μ *es un cardinal infinito) tal que si* $\alpha \le \beta < \mu$ *entonces* $\kappa_\alpha \le \kappa_\beta$. *Entonces*

$$\prod_{\alpha<\mu} \kappa_\alpha = (\sup_{\alpha<\mu} \kappa_\alpha)^\mu.$$

Demostración: Sea $\kappa = \sup_{\alpha<\mu} \kappa_\alpha$. Entonces $\prod_{\alpha<\mu} \kappa_\alpha \le \prod_{\alpha<\mu} \kappa = \kappa^\mu$.

Tomemos una aplicación biyectiva $f : \mu \times \mu \longrightarrow \mu$. Sea $A_\alpha = f[\mu \times \{\alpha\}]$. Así $\mu = \bigcup_{\alpha<\mu} A_\alpha$ y los conjuntos A_α tienen cardinal μ y son disjuntos dos a dos. En particular no están acotados en μ (o tendrían cardinal menor). Teniendo en cuenta la monotonía de la sucesión κ_α, es claro que $\sup_{\beta\in A_\alpha} \kappa_\beta = \kappa$.

Como los κ_β son no nulos, tenemos que $\kappa_\beta \le \prod_{\beta\in A_\alpha} \kappa_\beta$, luego

$$\kappa = \sup_{\beta\in A_\alpha} \kappa_\beta \le \prod_{\beta\in A_\alpha} \kappa_\beta.$$

Por consiguiente

$$\kappa^\mu = \prod_{\alpha<\mu} \kappa \le \prod_{\alpha<\mu} \prod_{\beta\in A_\alpha} \kappa_\beta = \prod_{\alpha<\mu} \kappa_\alpha \le \kappa^\mu.$$

■

Por ejemplo,

$$\prod_{n\in\omega} \aleph_n = \aleph_\omega^{\aleph_0}.$$

Ejercicio: Probar que el teorema anterior sigue siendo válido si sustituimos μ por un ordinal límite λ y suponemos que λ contiene $|\lambda|$ subconjuntos no acotados disjuntos dos a dos. En particular, probar que es cierto siempre que $\lambda < \omega_1$. En el capítulo siguiente veremos que en general estas restricciones no pueden ser eliminadas.

Veamos ahora una desigualdad entre una suma y un producto:

Teorema 13.48 *Si* $\bigwedge i \in I\ 2 \leq \kappa_i$, *entonces* $\sum_{i\in I} \kappa_i \leq \prod_{i\in I} \kappa_i$.

DEMOSTRACIÓN: Claramente $|I| \leq 2^{|I|} = \prod_{i\in I} 2 \leq \prod_{i\in I} \kappa_i$. Por otra parte, $\kappa_i \leq \prod_{i\in I} \kappa_i$, luego $\sup_{i\in I} \kappa_i \leq \prod_{i\in I} \kappa_i$. El teorema 13.43 nos da la conclusión si I es infinito o algún κ_i es infinito. El caso restante se demuestra fácilmente por inducción sobre el cardinal de I (aunque nunca vamos a necesitar este caso). ∎

Si nos fijamos en todos los teoremas sobre cardinales infinitos que hemos demostrado hasta ahora, no encontraremos más que una desigualdad estricta: el teorema de Cantor. El próximo teorema es la desigualdad estricta más general que se conoce sobre cardinales infinitos. Cualquier otra es un caso particular de ésta. Por ejemplo, el teorema de Cantor se obtiene haciendo $\kappa_i = 1$ y $\mu_i = 2$.

Teorema 13.49 (Teorema de König) *Si* $\bigwedge i \in I\ \kappa_i < \mu_i$, *entonces*

$$\sum_{i\in I} \kappa_i < \prod_{i\in I} \mu_i.$$

DEMOSTRACIÓN: Si $I = \varnothing$ el teorema se reduce a $0 < 1$. En otro caso, sea $I' = \{i \in I \mid \kappa_i > 0\}$. Para $i \in I'$ tenemos que $1 \leq \kappa_i < \mu_i$, luego $2 \leq \mu_i$ y podemos aplicar el teorema anterior:

$$\sum_{i\in I} \kappa_i = \sum_{i\in I'} \kappa_i \leq \prod_{i\in I'} \mu_i \leq \prod_{i\in I} \mu_i.$$

Supongamos que se diera la igualdad, es decir, que existe una aplicación biyectiva

$$f : \bigcup_{i\in I} \kappa_i \times \{i\} \longrightarrow \prod_{i\in I} \mu_i.$$

Sea $f_i : \kappa_i \longrightarrow \mu_i$ dada por $f_i(\alpha) = f(\alpha, i)(i)$.

Como $\kappa_i < \mu_i$ la aplicación f_i no puede ser suprayectiva, luego existe un $\alpha_i \in \mu_i \setminus f_i[\kappa_i]$. Los α_i determinan un elemento $\alpha = (\alpha_i)_{i\in I} \in \prod_{i\in I} \mu_i$. Como f es biyectiva, α tiene una antiimagen, de modo que $f(\beta, j) = \alpha$. Entonces $\beta \in \kappa_j$ y $f_j(\beta) = f(\beta, j)(j) = \alpha_j \in f_j[\kappa_j]$, en contradicción con la elección de α_j. ∎

13.5 Cofinalidad

El concepto de cofinalidad es esencial en el estudio de los cardinales infinitos, y en particular en el estudio de la exponenciación cardinal que todavía tenemos pendiente. En esta sección no usamos el axioma de elección salvo en unos pocos casos, donde lo indicaremos explícitamente.

Definición 13.50 Diremos que una aplicación $f : \alpha \longrightarrow \beta$ entre dos ordinales es *cofinal* si $f[\alpha]$ no está acotado estrictamente en β, es decir, si se cumple que $\bigwedge\gamma < \beta \bigvee\delta < \alpha\ \gamma \leq f(\delta)$.

Llamaremos *cofinalidad* de β al menor ordinal α tal que existe una aplicación cofinal $f : \alpha \longrightarrow \beta$. Lo representaremos por $\operatorname{cf}\beta$. Como la identidad en β es obviamente cofinal, vemos que $\operatorname{cf}\beta$ está bien definida y además $\operatorname{cf}\beta \leq \beta$.

Informalmente, podemos decir que $\operatorname{cf}\beta$ es el mínimo número de pasos que hay que dar para ascender completamente por β, es decir, para ascender rebasando (o, al menos, igualando) cualquier ordinal menor que β. Obviamente $\operatorname{cf} 0 = 0$ y $\operatorname{cf}\alpha + 1 = 1$. En efecto, la aplicación $f : 1 \longrightarrow \alpha + 1$ dada por $f(0) = \alpha$ es cofinal ($\alpha + 1$ tiene un máximo elemento y basta un paso para llegar hasta él).

Así pues, la cofinalidad sólo tiene interés sobre los ordinales límite, los cuales no se pueden recorrer en un paso. De hecho, siempre hacen falta infinitos pasos:

Teorema 13.51 $\bigwedge\lambda\ \omega \leq \operatorname{cf}\lambda \leq \lambda$.

Demostración: Ya sabemos que $\operatorname{cf}\lambda \leq \lambda$. Por otra parte $\operatorname{cf}\lambda$ no puede ser un número natural n, ya que si $f : n \longrightarrow \lambda$, entonces $f[n]$ es un conjunto finito, luego tiene un máximo $\alpha < \lambda$, luego $\alpha + 1 < \lambda$ es una cota estricta de $f[n]$, luego f no es cofinal. ∎

Decimos que $\operatorname{cf}\alpha$ es el mínimo número de pasos necesarios para ascender completamente por α. Este "número de pasos" es ciertamente un cardinal:

Teorema 13.52 $\bigwedge\alpha\ \operatorname{cf}\alpha \in K$.

Demostración: Si $\alpha = 0$ o $\alpha = \beta + 1$ sabemos que $\operatorname{cf}\alpha$ es 0 o 1, luego es un cardinal. Basta probar entonces que $\bigwedge\lambda\ \operatorname{cf}\lambda \in K$. Supongamos que $|\operatorname{cf}\lambda| < \operatorname{cf}\lambda$. Sea $f : |\operatorname{cf}\lambda| \longrightarrow \operatorname{cf}\lambda$ biyectiva y sea $g : \operatorname{cf}\lambda \longrightarrow \lambda$ cofinal. Entonces $f \circ g : |\operatorname{cf}\lambda| \longrightarrow \lambda$ tiene la misma imagen que g, luego es cofinal, en contra de la minimalidad de $\operatorname{cf}\lambda$. ∎

A partir de aquí trataremos únicamente con ordinales límite. Notemos que $f : \alpha \longrightarrow \lambda$ es cofinal si y sólo si $f[\alpha]$ no está acotado en λ, es decir, si y sólo si

$$\lambda = \sup f[\alpha] = \bigcup_{\delta<\alpha} f(\delta).$$

La forma más económica de ascender por un ordinal es no retrocediendo nunca. Veamos que esto siempre es posible:

Teorema 13.53 $\bigwedge\lambda\bigvee f\ f : \operatorname{cf}\lambda \longrightarrow \lambda$ *cofinal y normal.*

DEMOSTRACIÓN: Sea $g : \operatorname{cf}\lambda \longrightarrow \lambda$ cofinal. Definimos $f : \operatorname{cf}\lambda \longrightarrow \Omega$ como la única aplicación que cumple

$$\begin{aligned}&f(0) = g(0),\\ &\textstyle\bigwedge\alpha < \operatorname{cf}\lambda\ f(\alpha+1) = \max\{g(\alpha), f(\alpha)+1\},\\ &\textstyle\bigwedge\lambda' < \operatorname{cf}\lambda\ f(\lambda') = \bigcup_{\delta<\lambda'} f(\delta).\end{aligned}$$

Claramente f es normal. Veamos por inducción que $\bigwedge\alpha < \operatorname{cf}\lambda\ f(\alpha) < \lambda$. En efecto, para $\alpha = 0$ es obvio y si vale para α vale claramente para $\alpha+1$. Supongamos que $\lambda' < \operatorname{cf}\lambda$ y que $\bigwedge\delta < \lambda' f(\delta) < \lambda$. Entonces es claro que $f(\lambda') \leq \lambda$, pero no puede darse la igualdad porque entonces $f|'_\lambda$ sería cofinal en λ, en contradicción con que $\lambda' < \operatorname{cf}\lambda$. Así pues, también se cumple para λ'.

Tenemos entonces que $f : \operatorname{cf}\lambda \longrightarrow \lambda$ normal y, como $\bigwedge\alpha < \operatorname{cf}\lambda\ g(\alpha) \leq f(\alpha)$, es claro que f es cofinal. ∎

Este teorema nos permite expresar la cofinalidad de un ordinal límite en términos únicamente de sus subconjuntos acotados:

Teorema 13.54 *La cofinalidad de un ordinal límite λ es el mínimo cardinal κ tal que existe un subconjunto $a \subset \lambda$ no acotado de cardinal κ.*

DEMOSTRACIÓN: Si $f : \operatorname{cf}\lambda \longrightarrow \lambda$ es cofinal y normal, entonces $a = f[\operatorname{cf}\lambda]$ es un subconjunto no acotado de λ y, como f es inyectiva, su cardinal es $\operatorname{cf}\lambda$.

Recíprocamente, si $a \subset \lambda$ es un subconjunto no acotado, sea $f : |a| \longrightarrow a$ una biyección. Entonces es claro que $f : |a| \longrightarrow \lambda$ cofinal, luego $\operatorname{cf}\lambda \leq |a|$. ∎

En general, la composición de aplicaciones cofinales no es necesariamente cofinal (es fácil encontrar ejemplos). El teorema siguiente nos da una condición suficiente:

Teorema 13.55 *Si $f : \lambda_1 \longrightarrow \lambda_2$ y $g : \lambda_2 \longrightarrow \lambda_3$ son cofinales y además g es creciente, entonces $f \circ g : \lambda_1 \longrightarrow \lambda_3$ es cofinal.*

DEMOSTRACIÓN: Sea $\alpha < \lambda_3$. Como g es cofinal existe $\beta < \lambda_2$ tal que $\alpha \leq g(\beta)$. Como f es cofinal existe $\gamma < \lambda_1$ tal que $\beta \leq f(\gamma)$. Como g es creciente, $\alpha \leq g(\beta) \leq g(f(\gamma)) = (f \circ g)(\gamma)$, luego $f \circ g$ es cofinal. ∎

Esto tiene una consecuencia destacable:

Teorema 13.56 *Si $f : \lambda_1 \longrightarrow \lambda_2$ es cofinal y creciente, entonces* $\operatorname{cf}\lambda_1 = \operatorname{cf}\lambda_2$.

DEMOSTRACIÓN: Sea $g : \operatorname{cf}\lambda_1 \longrightarrow \lambda_1$ cofinal. Por el teorema anterior $g \circ f : \operatorname{cf}\lambda_1 \longrightarrow \lambda_2$ es cofinal, luego $\operatorname{cf}\lambda_2 \leq \operatorname{cf}\lambda_1$.

Sea ahora $h : \operatorname{cf}\lambda_2 \longrightarrow \lambda_2$ cofinal y definamos $r : \operatorname{cf}\lambda_2 \longrightarrow \lambda_1$ de modo que $r(\alpha)$ sea el menor $\beta < \lambda_1$ tal que $h(\alpha) < f(\beta)$, que existe porque f es cofinal. Entonces r es cofinal, pues si $\gamma < \lambda_1$ entonces $f(\gamma) < \lambda_2$, luego existe un

$\delta < \operatorname{cf}\lambda_2$ tal que $f(\gamma) \leq h(\delta)$. Por definición de r tenemos que $h(\delta) < f(r(\delta))$, y si fuera $r(\delta) \leq \gamma$ sería $f(r(\delta)) \leq f(\gamma) \leq h(\delta)$, contradicción, luego $\gamma \leq r(\delta)$ y r es cofinal. Por consiguiente $\operatorname{cf}\lambda_1 \leq \operatorname{cf}\lambda_2$ y tenemos la igualdad. ∎

Este teorema, además de servir para calcular cofinalidades, tiene una lectura negativa: en la prueba del teorema 13.53 hemos partido de una aplicación cofinal arbitraria y la hemos modificado para hacerla cofinal y normal, en particular creciente. Ahora vemos que esto no siempre puede hacerse: pueden darse casos en los que exista una aplicación cofinal entre dos ordinales límite y no exista ninguna aplicación cofinal y creciente, pues una condición necesaria para que esto ocurra es que ambos ordinales tengan la misma cofinalidad.

Respecto al cálculo de cofinalidades, el teorema siguiente es una consecuencia sencilla del anterior, pero más cómodo en la práctica:

Teorema 13.57 *Si $f : \lambda_1 \longrightarrow \Omega$ es normal y $\lambda < \lambda_1$, entonces* $\operatorname{cf}\lambda = \operatorname{cf} f(\lambda)$.

Demostración: Es claro que $f|_\lambda : \lambda \longrightarrow f(\lambda)$ es cofinal y creciente. Basta aplicar el teorema anterior. ∎

Por ejemplo, $\operatorname{cf}\aleph_{\omega^2} = \operatorname{cf}\omega^2 = \operatorname{cf}\omega\cdot\omega = \operatorname{cf}\omega = \aleph_0$. Hemos usado la normalidad de las funciones $\aleph$ y $\omega\cdot$. Una función cofinal de ω en $\aleph_{\omega^2}$ es $f(n) = \aleph_{\omega\cdot n}$.

Veamos un ejemplo típico de la utilidad del concepto de cofinalidad.

Definición 13.58 Sea $f : \lambda \longrightarrow \lambda$, donde λ cumple $\operatorname{cf}\lambda > \aleph_0$ o bien $\lambda = \Omega$. Para cada $\alpha \in \lambda$ definimos

$$\begin{aligned} f^0(\alpha) &= \alpha,\\ f^{n+1}(\alpha) &= f(f^n(\alpha)),\\ f^\omega(\alpha) &= \sup_{n\in\omega} f^n(\alpha). \end{aligned}$$

Una simple inducción prueba que $\bigwedge n \in \omega f^n(\alpha) \in \lambda$, y la hipótesis sobre λ asegura que el conjunto numerable $\{f^n(\alpha) \mid n \in \omega\}$ ha de estar acotado en λ (teorema 13.54), luego $f^\omega(\alpha) \in \lambda$. Así pues, tenemos definida una función $f^\omega : \lambda \longrightarrow \lambda$ a la que llamaremos *función iterada* de f.

Es inmediato a partir de esta construcción que $\bigwedge\alpha \in \lambda\ \alpha \leq f^\omega(\alpha)$.

Informalmente, $f^\omega(\alpha)$ resulta de aplicar infinitas veces f a α, lo cual hace que si aplicamos f una vez más no se nota:

Teorema 13.59 *Sea $f : \lambda \longrightarrow \lambda$ una función normal, donde* $\operatorname{cf}\lambda > \aleph_0$ *o bien $\lambda = \Omega$. Entonces $\bigwedge\alpha \in \lambda\ f(f^\omega(\alpha)) = f^\omega(\alpha)$.*

Demostración: Como f es normal, se cumple que $f^\omega(\alpha) \leq f(f^\omega(\alpha))$. Para probar la otra desigualdad distinguimos tres casos:

Si $f^\omega(\alpha) = 0$, entonces $\alpha = f(\alpha) = 0$, pues tanto α como $f(\alpha)$ están bajo $f^\omega(\alpha)$. Por consiguiente $f(f^\omega(\alpha)) = f(0) = f(\alpha) \leq f^\omega(\alpha)$.

Si $f^\omega(\alpha) = \gamma + 1$, entonces $\gamma < f^\omega(\alpha)$, luego $\gamma < f^n(\alpha)$ para cierto $n \in \omega$. Así,

$$f(f^\omega(\alpha)) = f(\gamma+1) \leq f(f^n(\alpha)) = f^{n+1}(\alpha) \leq f^\omega(\alpha).$$

Si $f^\omega(\alpha)$ es un ordinal límite, como f es normal,

$$f(f^\omega(\alpha)) = \bigcup_{\delta<f^\omega(\alpha)} f(\delta) \leq \bigcup_{n\in\omega} f(f^n(\alpha)) \leq \bigcup_{n\in\omega} f^{n+1}(\alpha) \leq f^\omega(\alpha).$$

■

En particular hemos demostrado:

Teorema 13.60 (Teorema de punto fijo para funciones normales) *Sea $f : \lambda \longrightarrow \lambda$ una función normal, donde $\operatorname{cf}\lambda > \aleph_0$ o bien $\lambda = \Omega$. Entonces*

$$\bigwedge\alpha \in \lambda \bigvee\beta \in \lambda(\alpha \leq \beta \wedge f(\beta) = \beta).$$

La función $(\omega+) : \omega^2 \longrightarrow \omega^2$ es un ejemplo de función normal sin puntos fijos. Destaquemos el papel que desempeña la hipótesis sobre la cofinalidad: para construir puntos fijos necesitamos ascender $\aleph_0$ pasos, luego necesitamos que la cofinalidad de λ sea no numerable para garantizar que con el ascenso no nos salimos de λ.

Así, por ejemplo, existen cardinales κ arbitrariamente grandes tales que $\kappa = \aleph_\kappa$.

Pasemos ahora al cálculo de la cofinalidad de los cardinales infinitos. Ello requiere el axioma de elección. En primer lugar damos una caracterización en términos de la aritmética cardinal:

Teorema 13.61 *(AE) Sea κ un cardinal infinito. Entonces $\operatorname{cf}\kappa$ es el menor cardinal μ tal que existe una familia de cardinales $\{\nu_\alpha\}_{\alpha<\mu}$ tales que*

$$\bigwedge\alpha < \mu\ \nu_\alpha < \kappa \qquad y \qquad \sum_{\alpha<\mu} \nu_\alpha = \kappa.$$

Demostración: Sea $f : \operatorname{cf}\kappa \longrightarrow \kappa$ cofinal. Entonces $\kappa = \bigcup_{\alpha<\operatorname{cf}\kappa} f(\alpha)$. Sea $\nu_\alpha = |f(\alpha)| < \kappa$. Entonces

$$\kappa = |\kappa| = \Big|\bigcup_{\alpha<\operatorname{cf}\kappa} f(\alpha)\Big| \leq \sum_{\alpha<\operatorname{cf}\kappa} \nu_\alpha \leq \sum_{\alpha<\operatorname{cf}\kappa} \kappa = \kappa\ \operatorname{cf}\kappa = \kappa.$$

Por consiguiente $\kappa = \sum_{\alpha<\operatorname{cf}\kappa} \nu_\alpha$. Ahora veamos que $\operatorname{cf}\kappa$ es el mínimo cardinal que cumple esto. Tomemos $\mu < \operatorname{cf}\kappa$ y sea $\{\nu_\alpha\}_{\alpha<\mu}$ una familia de cardinales tal que $\bigwedge\alpha < \mu\ \nu_\alpha < \kappa$.

La aplicación $f : \mu \longrightarrow \kappa$ dada por $f(\alpha) = \nu_\alpha$ no puede ser cofinal, luego existe un ordinal $\beta < \kappa$ tal que $\bigwedge\alpha < \mu\ \nu_\alpha < \beta$ y así

$$\sum_{\alpha<\mu} \nu_\alpha \leq \sum_{\alpha<\mu} |\beta| = \mu\,|\beta| < \kappa,$$

luego, en efecto, $\operatorname{cf}\kappa$ es el mínimo cardinal con la propiedad del enunciado. ■

Así pues, tenemos lo siguiente sobre las cofinalidades de los cardinales infinitos:

Teorema 13.62 *Se cumple*

a) $\operatorname{cf}\aleph_0 = \aleph_0$,

b) $\bigwedge\lambda\ \operatorname{cf}\aleph_\lambda = \operatorname{cf}\lambda$,

c) (AE) $\bigwedge\alpha\ \operatorname{cf}\aleph_{\alpha+1} = \aleph_{\alpha+1}$.

Demostración: a) es consecuencia inmediata de 13.51, b) es un caso particular de 13.57. Veamos c). En caso contrario, sería $\operatorname{cf}\aleph_{\alpha_1} \leq \aleph_\alpha$ y por el teorema anterior existirían cardinales $\{\nu_\delta\}_{\delta<\operatorname{cf}\aleph_{\alpha+1}}$ tales que $\bigwedge\delta < \operatorname{cf}\aleph_{\alpha+1}\ \nu_\delta \leq \aleph_\alpha$ y

$$\aleph_{\alpha+1} = \sum_{\delta<\operatorname{cf}\aleph_{\alpha+1}} \nu_\delta \leq \sum_{\delta<\operatorname{cf}\aleph_{\alpha+1}} \aleph_\alpha = \aleph_\alpha\ \operatorname{cf}\aleph_{\alpha+1} = \aleph_\alpha,$$

contradicción. ∎

Así pues, el hecho de que $\operatorname{cf}\aleph_0 = \aleph_0$ expresa que la unión finita de conjuntos finitos es finita e, igualmente, $\operatorname{cf}\aleph_1 = \aleph_1$ expresa que la unión de una cantidad numerable de conjuntos numerables es numerable. En cambio, podemos obtener un conjunto de cardinal $\aleph_\omega$ uniendo tan sólo una cantidad numerable de conjuntos de cardinal menor que $\aleph_\omega$, pues basta unir un conjunto de cardinal $\aleph_0$ con otro de cardinal $\aleph_1$, con otro de cardinal $\aleph_2$, etc. Por ello, $\operatorname{cf}\aleph_\omega = \aleph_0$.

Definición 13.63 Un cardinal infinito κ es *regular* si $\operatorname{cf}\kappa = \kappa$ y es *singular* si $\operatorname{cf}\kappa < \kappa$.

Un cardinal infinito κ es un *cardinal sucesor* si es de la forma μ^+, para otro cardinal μ y es un *cardinal límite* en caso contrario. Es claro que los cardinales límite son $\aleph_0$ y los de la forma $\aleph_\lambda$, mientras que los cardinales sucesores son los de la forma $\aleph_{\alpha+1}$. Hemos probado que $\aleph_0$ y todos los cardinales sucesores son regulares. En cambio, $\aleph_\omega$ o $\aleph_{\omega_3}$ son ejemplos de cardinales singulares (de cofinalidades, respectivamente, $\aleph_0$ y $\aleph_3$).

De los teoremas 13.53 y 13.56 se sigue inmediatamente:

Teorema 13.64 $\bigwedge\alpha\ \operatorname{cf}\alpha$ *es un cardinal regular.*

Todo cardinal sucesor es regular y conocemos ejemplos de cardinales límite singulares. Queda abierta la cuestión de si existen cardinales límite regulares aparte de $\aleph_0$.

Definición 13.65 Un *cardinal (débilmente) inaccesible* es un cardinal límite regular distinto de $\aleph_0$.

No vamos a verlo aquí, pero puede demostrarse que

$$\underset{ZFC}{\vdash} \bigvee\kappa\ \kappa \text{ débilmente inaccesible} \to \operatorname{Consis} ZFC,$$

de donde se sigue que si ZFC es consistente no es posible demostrar la existencia de cardinales débilmente inaccesibles o, equivalentemente, es consistente

suponer que no existen. Aunque no podamos probar su existencia, podemos preguntarnos si al menos es consistente suponer que existan, es decir, resulta natural preguntarse si

$$\operatorname{Consis} ZFC \to \operatorname{Consis} ZFC + DI, \tag{13.2}$$

donde $DI \equiv \bigvee\kappa\kappa$ débilmente inaccesible. Ahora bien, si alguien encontrara un argumento (metamatemático) convincente de que esto es así, es decir, un argumento que no partiera de ningún supuesto extraño o dudoso, dicho argumento podría, sin duda, formalizarse en ZFC, con lo cual tendríamos

$$\underset{ZFC}{\vdash} \operatorname{Consis} ZFC \to \operatorname{Consis} ZFC + DI,$$

y en particular

$$\underset{ZFC+DI}{\vdash} \operatorname{Consis} ZFC \to \operatorname{Consis} ZFC + DI,$$

pero sabemos que $\underset{ZFC+DI}{\vdash} \operatorname{Consis} ZFC$, luego concluiríamos que

$$\underset{ZFC+DI}{\vdash} \operatorname{Consis} ZFC + DI.$$

Por el segundo teorema de incompletitud, ZFC+DI sería contradictorio y, por la prueba de consistencia relativa que estamos suponiendo que existe, llegaríamos a que ZFC también sería contradictorio.

Esta conclusión no procede de suponer (13.2), sino de suponer que (13.2) es demostrable. Así pues, lo que hemos probado es que si la existencia de cardinales débilmente inaccesibles es consistente con ZFC, no es posible demostrarlo ni siquiera suponiendo la consistencia de ZFC.

Por otra parte, esto no proporciona ningún indicio de que la existencia de cardinales débilmente inaccesibles sea contradictoria. Entenderemos mejor la situación en el capítulo siguiente, cuando estudiemos los cardinales fuertemente inaccesibles.

Terminamos con una última propiedad de los cardinales inaccesibles:

Teorema 13.66 *Un cardinal regular κ es débilmente inaccesible si y sólo si $\kappa = \aleph_\kappa$.*

Demostración: Una implicación es obvia. Si κ es débilmente inaccesible, entonces $\kappa = \aleph_\lambda$, para cierto λ tal que

$$\lambda \leq \aleph_\lambda = \kappa = \operatorname{cf}\kappa = \operatorname{cf}\aleph_\lambda = \operatorname{cf}\lambda \leq \lambda.$$

■

Naturalmente, la función $\aleph$ tiene infinitos puntos fijos que no son cardinales inaccesibles (son singulares).

Capítulo XIV

La exponenciación cardinal

Ya hemos comentado que la exponenciación cardinal es completamente distinta a la suma y el producto de cardinales. En efecto, estas operaciones están perfectamente determinadas por los axiomas de la teoría de conjuntos y su comportamiento es bien conocido, como ya hemos podido comprobar. Por el contrario, la exponenciación cardinal sigue siendo hoy en día objeto de investigación, pues no se sabe a ciencia cierta dónde acaba lo que se puede decir sobre ella en el seno de ZFC y qué posibilidades son consistentes aunque indemostrables. Sin entrar en pruebas de consistencia, lo cual excedería el propósito de este libro, en este capítulo trataremos de dar una idea general de la situación. Trabajaremos en ZFC o NBG, es decir, suponemos el axioma de elección, sin el cual es impensable obtener nada de valor.

Nota En este capítulo usaremos la notación ${}^\beta\alpha$ para representar al conjunto de las aplicaciones de β en α cuando la notación usual α^β pueda confundirse con la exponenciación ordinal o cardinal.

14.1 La exponenciación en ZFC

Ya hemos visto en el capítulo anterior las propiedades formales básicas de la exponenciación cardinal, que se demuestran rutinariamente como en el caso de la suma y el producto. Ahora estamos interesados en el cálculo explícito de potencias κ^μ. Sabemos que la exponenciación de números naturales se reduce a la usual, por lo que podemos centrarnos en el caso en que al menos uno de los cardinales es infinito. Más concretamente, el caso realmente interesante se da cuando el exponente es infinito, ya que si es finito la potencia se reduce a las propiedades del producto de cardinales por inducción. De hecho, en virtud del teorema 13.31, el teorema siguiente (enunciado para cardinales no necesariamente bien ordenables) es una forma equivalente del axioma de elección:

Teorema 14.1 *Si κ es un cardinal infinito y n un número natural no nulo, entonces $\kappa^n = \kappa$.*

Si la base es finita (mayor que 1, o si no el cálculo es trivial), el problema se reduce al caso en que es igual a 2. Más en general:

Teorema 14.2 *Sean κ y μ cardinales tales que $2 \leq \kappa \leq \mu$ y $\aleph_0 \leq \mu$. Entonces $\kappa^\mu = 2^\mu$.*

DEMOSTRACIÓN: $\kappa^\mu \leq (2^\kappa)^\mu = 2^\mu \leq \kappa^\mu$. ∎

Si la base es infinita podemos centrarnos en el caso en que sea un cardinal límite, en virtud de la fórmula que probamos a continuación. En la prueba hacemos uso de un argumento general que conviene destacar porque nos va a aparecer más veces:

Si $\mu < \operatorname{cf}\kappa$, entonces

$$^\mu\kappa = \bigcup_{\alpha<\kappa} {}^\mu\alpha.$$

En efecto, esto es una forma de expresar que toda función $f : \mu \longrightarrow \kappa$ está acotada.

Teorema 14.3 (Fórmula de Hausdorff) *Se cumple:*

$$\bigwedge\alpha\beta\ \aleph_{\alpha+1}^{\aleph_\beta} = \aleph_\alpha^{\aleph_\beta}\aleph_{\alpha+1}.$$

DEMOSTRACIÓN: Si $\alpha + 1 \leq \beta$, entonces $\aleph_{\alpha+1} \leq \aleph_\beta < 2^{\aleph_\beta}$, luego

$$\aleph_\alpha^{\aleph_\beta}\aleph_{\alpha+1} = 2^{\aleph_\beta}\aleph_{\alpha+1} = 2^{\aleph_\beta} = \aleph_{\alpha+1}^{\aleph_\beta}.$$

Si, por el contrario, $\beta < \alpha + 1$, entonces, como $\aleph_{\alpha+1}$ es regular,

$$^{\omega_\beta}\omega_{\alpha+1} = \bigcup_{\delta<\omega_{\alpha+1}} {}^{\omega_\beta}\delta,$$

luego

$$\aleph_{\alpha+1}^{\aleph_\beta} = |{}^{\omega_\beta}\omega_{\alpha+1}| = \Big|\bigcup_{\delta<\omega_{\alpha+1}} {}^{\omega_\beta}\delta\Big| \leq \sum_{\delta<\omega_{\alpha+1}} |\delta|^{\aleph_\beta} \leq \sum_{\delta<\omega_{\alpha+1}} \aleph_\alpha^{\aleph_\beta} = \aleph_\alpha^{\aleph_\beta}\aleph_{\alpha+1}.$$

La otra desigualdad es obvia. ∎

Al igual que 14.2, muchos de los resultados sobre exponenciación cardinal involucran la función 2^κ, la cual, ciertamente, es el esqueleto de la exponenciación cardinal. Por ello es conveniente darle un nombre:

Definición 14.4 Se llama *función del continuo* a la función $\kappa \mapsto 2^\kappa$ definida sobre los cardinales infinitos.

Así, la hipótesis del continuo generalizada no es más que una determinación de la función del continuo, en virtud de la cual $2^\kappa = \kappa^+$. Ya hemos comentado que esta hipótesis no puede ser demostrada ni refutada, lo que significa que hay otras alternativas igualmente consistentes con los axiomas de ZFC (supuesto,

claro, que éstos sean consistentes). De todos modos, no sirve cualquier determinación total o parcial de la función del continuo. Por ejemplo, es obvio que sería contradictorio suponer que

$$2^{\aleph_0} = \aleph_5 \wedge 2^{\aleph_1} = \aleph_3.$$

Más en general, la función del continuo ha de respetar la monotonía:

$$\kappa \leq \mu \to 2^\kappa \leq 2^\mu.$$

Otra restricción a la función del continuo es el teorema de Cantor: sería contradictorio suponer que $2^\kappa = \kappa$ para todo cardinal κ, a pesar de que esta función del continuo sí que es monótona. En realidad, la función del continuo está sometida a una desigualdad más fina que el teorema de Cantor, consecuencia del teorema de König 13.49 y, más concretamente, del teorema siguiente:

Teorema 14.5 (Teorema de König) *Para todo cardinal infinito κ se cumple $\kappa < \kappa^{\operatorname{cf}\kappa}$.*

Demostración: Sea $\{\mu_\alpha\}_{\alpha<\operatorname{cf}\kappa}$ una familia de cardinales menores que κ tales que $\kappa = \sum_{\alpha<\operatorname{cf}\kappa} \mu_\alpha$. Por el teorema 13.49 resulta que

$$\kappa = \sum_{\alpha<\operatorname{cf}\kappa} \mu_\alpha < \prod_{\alpha<\operatorname{cf}\kappa} \kappa = \kappa^{\operatorname{cf}\kappa}.$$

■

Ciertamente, este teorema refina al teorema de Cantor, pues en virtud de 14.2 éste puede expresarse como $\kappa < \kappa^\kappa$, y en el teorema anterior el exponente es menor o igual que κ. De todos modos, podemos expresar esta restricción en términos de la función del continuo:

Teorema 14.6 (Teorema de König) *Si κ es un cardinal infinito, entonces $\kappa < \operatorname{cf} 2^\kappa$.*

Demostración: Si $\operatorname{cf} 2^\kappa \leq \kappa$, entonces $(2^\kappa)^{\operatorname{cf} 2^\kappa} \leq (2^\kappa)^\kappa = 2^\kappa$, en contradicción con el teorema anterior. ■

Así pues, $2^{\aleph_0}$ puede ser $\aleph_1$, $\aleph_2$, $\aleph_{\omega+1}$ o $\aleph_{\omega_3}$, pero no $\aleph_\omega$. Cuando decimos "puede ser" queremos decir que es consistente suponer que lo es. En efecto, aunque no lo vamos a probar aquí, este teorema y la monotonía es todo lo que puede probarse sobre la función del continuo sobre cardinales regulares, en el sentido de que cualquier axioma que determine la función del continuo sobre cardinales regulares que sea compatible con estos dos requisitos es consistente con los axiomas de ZFC (supuesto que éstos sean consistentes).

Notemos que no exigimos que la función 2^κ sea estrictamente monótona, de modo que, por ejemplo, es consistente suponer que $2^{\aleph_0} = 2^{\aleph_1} = \aleph_5$.

Debemos resaltar la restricción a cardinales regulares. Si no fuera así podríamos decir que comprendemos completamente la función del continuo, en cuanto

que sabríamos decir exactamente qué posibilidades son consistentes y cuáles no. Sin embargo, la situación en los cardinales límite es muy confusa. Por ejemplo, es contradictorio suponer que

$$\bigwedge\alpha\ 2^{\aleph_\alpha} = \aleph_{\alpha+\omega+2},$$

a pesar de que esta (presunta) función del continuo respeta tanto la monotonía como el teorema de König. Según lo dicho, no hay inconveniente en postular este axioma únicamente para cardinales regulares, pero no puede cumplirse para todos los cardinales singulares, como muestra el teorema siguiente:

Teorema 14.7 *Si un ordinal β cumple $\bigwedge\alpha\ 2^{\aleph_\alpha} = \aleph_{\alpha+\beta}$, entonces $\beta < \omega$.*

Demostración: Supongamos que β es infinito y sea α el mínimo ordinal tal que $\beta < \alpha + \beta$. Claramente $0 < \alpha \leq \beta$. Necesariamente α es un ordinal límite, pues si $\alpha = \gamma + 1$ entonces

$$\beta < \alpha + \beta = \gamma + 1 + \beta = \gamma + \beta,$$

luego γ cumple lo mismo que α, en contra de la minimalidad de α.

$$\begin{aligned}\aleph_{\alpha+\alpha+\beta} &= 2^{\aleph_{\alpha+\alpha}} = 2^{\sum\limits_{\delta<\alpha}\aleph_{\alpha+\delta}} = \prod_{\delta<\alpha} 2^{\aleph_{\alpha+\delta}} = \prod_{\delta<\alpha}\aleph_{\alpha+\delta+\beta} = \prod_{\delta<\alpha}\aleph_{\alpha+\beta}\\ &= \prod_{\delta<\alpha} 2^{\aleph_\alpha} = (2^{\aleph_\alpha})^{|\alpha|} = 2^{\aleph_\alpha} = \aleph_{\alpha+\beta}.\end{aligned}$$

Por consiguiente, $\alpha + \alpha + \beta = \alpha + \beta$, y de aquí $\alpha + \beta = \beta$, en contra de la elección de α. ∎

No se sabe si es consistente

$$\bigwedge\alpha\ 2^{\aleph_\alpha} = \aleph_{\alpha+2}.$$

Para continuar nuestro estudio conviene introducir una operación muy relacionada con la exponenciación de cardinales:

Definición 14.8 Si β es un ordinal y A es un conjunto, definimos

$$A^{<\beta} = {}^{<\beta}A = \bigcup_{\alpha<\beta} A^\alpha,$$

es decir, $A^{<\beta}$ es el conjunto de las aplicaciones de un ordinal menor que β en A. Usaremos la segunda notación cuando pueda haber confusión con el cardinal

$$\kappa^{<\mu} = |{}^{<\mu}\kappa|.$$

El teorema siguiente nos da varias caracterizaciones interesantes de esta operación:

Teorema 14.9 *Si μ es infinito y $\kappa \geq 2$, entonces*

$$\kappa^{<\mu} = \sup_{\nu<\mu} \kappa^\nu = \sum_{\nu<\mu} \kappa^\nu,$$

donde ν recorre los cardinales menores que μ (no los ordinales).

DEMOSTRACIÓN: Si μ es un cardinal límite, $\mu = \sup_{\nu<\mu} \nu \leq \sup_{\nu<\mu} \kappa^\nu$.

Si $\mu = \nu^+$ entonces $\nu < \kappa^\nu$, pues si $\nu < \kappa$ es obvio y si $\kappa \leq \nu$ entonces $\nu < 2^\nu = \kappa^\nu$, luego $\nu < \sup_{\nu<\mu} \kappa^\nu$ y así $\mu = \nu^+ \leq \sup_{\nu<\mu} \kappa^\nu$. En cualquier caso

$$\sum_{\nu<\mu} \kappa^\nu = \sup_{\nu<\mu} \kappa^\nu.$$

Por consiguiente

$$\kappa^{<\nu} = \Big|\bigcup_{\alpha<\mu} {}^\alpha\kappa\Big| = \sum_{\alpha<\mu} \kappa^{|\alpha|} \leq \sum_{\alpha<\mu} \sup_{\nu<\mu} \kappa^\nu = \sup_{\nu<\mu} \kappa^\nu.$$

Si $\nu < \mu$, entonces ${}^\nu\kappa \subset {}^{<\mu}\kappa$, luego $\kappa^\nu \leq |{}^{<\mu}\kappa| = \kappa^{<\mu}$. Así pues, tomando el supremo, $\sup_{\nu<\mu} \kappa^\nu \leq \kappa^{<\mu}$ y tenemos la igualdad. ∎

A partir de este teorema es inmediato que si μ es infinito entonces

$$\kappa^{<\mu^+} = \kappa^\mu,$$

luego $\kappa^{<\mu}$ sólo tiene interés cuando μ es un cardinal límite.

Volviendo a la función del continuo, ahora podemos expresar la condición de monotonía como que $2^{<\kappa} \leq 2^\kappa$. El teorema siguiente es un refinamiento de esta relación que para cardinales sucesores es trivial, pero no así para cardinales límite:

Teorema 14.10 *Si κ es un cardinal infinito, entonces $2^\kappa = (2^{<\kappa})^{\operatorname{cf}\kappa}$.*

DEMOSTRACIÓN: Sea $\kappa = \sum_{\alpha<\operatorname{cf}\kappa} \nu_\alpha$, donde $\bigwedge\alpha < \operatorname{cf}\kappa\ \nu_\alpha < \kappa$. Entonces

$$2^\kappa = 2^{\sum_{\alpha<\operatorname{cf}\kappa} \nu_\alpha} = \prod_{\alpha<\operatorname{cf}\kappa} 2^{\nu_\alpha} \leq \prod_{\alpha<\operatorname{cf}\kappa} 2^{<\kappa} = (2^{<\kappa})^{\operatorname{cf}\kappa} \leq (2^\kappa)^{\operatorname{cf}\kappa} = 2^\kappa.$$

∎

Notemos que si $\kappa = \mu^+$ entonces la igualdad se reduce a $2^{\mu^+} = 2^{\mu^+}$, luego es trivial, tal y como advertíamos, pero para cardinales límite puede no serlo. Por ejemplo, si $\bigwedge n \in \omega\ 2^{\aleph_n} = 2^{\aleph_0}$ (lo cual es consistente), entonces $2^{<\aleph_\omega} = 2^{\aleph_0}$ y necesariamente $2^{\aleph_\omega} = 2^{\aleph_0}$.

Por otra parte, este teorema tampoco es definitivo pues, si tenemos, por ejemplo, $\bigwedge n \in \omega\ 2^{\aleph_n} = \aleph_{\omega+n+1}$, entonces $2^{<\aleph_\omega} = \aleph_{\omega+\omega}$ y sólo concluimos que $2^{\aleph_\omega} = \aleph_{\omega+\omega}^{\aleph_0}$, pero no sabemos qué valores puede tomar esta expresión.

Esto está relacionado con el problema de la relación que hay entre la función del continuo y la exponenciación en general κ^μ (una muestra es el teorema 14.2). Comprenderemos mejor esta relación en la sección siguiente. De momento acabamos ésta con algunos resultados técnicos de interés:

Teorema 14.11 *Si μ es un cardinal regular y $\kappa \geq 2$, entonces $(\kappa^{<\mu})^{<\mu} = \kappa^{<\mu}$.*

DEMOSTRACIÓN: Si $\mu = \xi^+$ es inmediato, así que podemos suponer que μ es un cardinal límite. Al ser regular, los subconjuntos no acotados en μ tienen cardinal μ. En particular, hay μ cardinales $\nu < \mu$, de donde

$$\mu \leq \sum_{\nu<\mu} \kappa^\nu = \kappa^{<\mu}.$$

Sea $\pi < \mu$. Como μ es regular se cumple que

$$^\pi \sup_{\nu<\mu} \kappa^\nu \subset \bigcup_{\nu<\mu} {}^\pi(\kappa^\nu).$$

En efecto, dada f en el miembro izquierdo, la aplicación $\pi \longrightarrow \mu$ que a cada $\alpha < \pi$ le asigna el mínimo $\nu < \mu$ tal que $f(\alpha) < \kappa^\nu$ no puede ser cofinal, luego ha de existir un $\nu < \mu$ tal que $f[\pi] \subset \kappa^\nu$ y f está en el miembro derecho.

Así pues, tomando cardinales,

$$(\kappa^{<\mu})^\pi \leq \sum_{\nu<\mu} \kappa^{\nu\pi} \leq \sum_{\nu<\mu} \kappa^{<\mu} = \mu\ \kappa^{<\mu} = \kappa^{<\mu}.$$

Tomando el supremo en π obtenemos $(\kappa^{<\mu})^{<\mu} \leq \kappa^{<\mu}$, y la otra desigualdad es obvia. ∎

Definición 14.12 Dado un conjunto A y un cardinal κ, llamaremos

$$\begin{aligned} [A]^\kappa &= \{x \mid x \subset A \wedge |x| = \kappa\},\\ [A]^{<\kappa} &= \{x \mid x \subset A \wedge |x| < \kappa\}. \end{aligned}$$

La exponenciación cardinal permite calcular los cardinales de estos conjuntos:

Teorema 14.13 *Sea A un conjunto infinito y κ un cardinal $\kappa \leq |A|$, Entonces*

$$|[A]^\kappa| = |A|^\kappa, \qquad |[A]^{<\kappa}| = |A|^{<\kappa}.$$

En particular A tiene $|A|$ subconjuntos finitos.

DEMOSTRACIÓN: Podemos suponer $\kappa > 0$. Sea $\mu = |A| = \kappa\mu = |\kappa \times \mu|$. Para la primera igualdad basta probar que $|[\kappa \times \mu]^\kappa| = \mu^\kappa$, pero es inmediato que ${}^\kappa\mu \subset [\kappa \times \mu]^\kappa$, de donde $\mu^\kappa \leq |[\kappa \times \mu]^\kappa|$ y, por otra parte, para cada $x \in [\kappa \times \mu]^\kappa$ podemos escoger una biyección $f_x : \kappa \longrightarrow x$, de modo que la aplicación $g : [\kappa \times \mu]^\kappa \longrightarrow {}^\kappa(\kappa \times \mu)$ dada por $g(x) = f_x$ es inyectiva, de donde $|[\kappa \times \mu]^\kappa| \leq |{}^\kappa(\kappa \times \mu)| = |\kappa \times \mu|^\kappa = \mu^\kappa$.

Respecto a la segunda igualdad,

$$|[A]^{<\kappa}| = \Big|\bigcup_{\mu<\kappa} [A]^\mu\Big| = \sum_{\mu<\kappa} |[A]^\mu| = \sum_{\mu<\kappa} |A|^\mu = |A|^{<\kappa}.$$

∎

14.2 La hipótesis de los cardinales singulares

La función del continuo más simple posible es, sin duda, la que postula la hipótesis del continuo generalizada:

$$2^\kappa = \kappa^+.$$

Sucede que esta hipótesis determina de hecho toda la exponenciación cardinal. En efecto:

Teorema 14.14 *(HCG) Si κ y μ son cardinales y μ es infinito, entonces*

$$\kappa^\mu = \begin{cases} \kappa & \text{si } \mu < \operatorname{cf}\kappa, \\ \kappa^+ & \text{si } \operatorname{cf}\kappa \le \mu \le \kappa, \\ \mu^+ & \text{si } \kappa \le \mu. \end{cases}$$

Demostración: Si $\mu < \operatorname{cf}\kappa$ tenemos la inclusión ${}^\mu\kappa \subset \bigcup_{\alpha<\kappa} {}^\mu\alpha$, de donde $\kappa^\mu \le \sum_{\alpha<\kappa} |\alpha|^\mu$. Ahora bien, dado $\alpha < \kappa$, se cumple que $\nu = \max\{|\alpha|, \mu\} < \kappa$, luego $|\alpha|^\mu \le \nu^\nu = \nu^+ \le \kappa$. Por consiguiente,

$$\kappa \le \kappa^\mu \le \sum_{\alpha<\kappa} \kappa = \kappa.$$

Si $\operatorname{cf}\kappa \le \mu \le \kappa$ entonces, por el teorema de König,

$$\kappa^+ \le \kappa^{\operatorname{cf}\kappa} \le \kappa^\mu \le \kappa^\kappa = 2^\kappa = \kappa^+.$$

Finalmente, si $\kappa \le \mu$ entonces $\kappa^\mu = 2^\mu = \mu^+$. ∎

En particular es claro que la HCG implica, para $\kappa \ge 2$ y μ un cardinal límite:

$$\kappa^{<\mu} = \begin{cases} \kappa & \text{si } \mu \le \operatorname{cf}\kappa, \\ \kappa^+ & \text{si } \operatorname{cf}\kappa < \mu \le \kappa, \\ \mu & \text{si } \kappa < \mu. \end{cases}$$

Ejemplo Suponiendo la HCG tenemos:

$$\aleph_3^{\aleph_5} = \aleph_6, \quad \aleph_7^{\aleph_2} = \aleph_7, \quad \aleph_{\omega_2}^{\aleph_1} = \aleph_{\omega_2}, \quad \aleph_{\omega_6}^{\aleph_8} = \aleph_{\omega_6}^+.$$

∎

A la vista de este resultado, es natural conjeturar que la función del continuo determina la exponenciación cardinal. En realidad existía una razón de mucho mayor peso que corroboraba esta conjetura: Durante mucho tiempo, las técnicas conocidas para construir modelos con funciones del continuo alternativas forzaban el resto de la exponenciación, es decir, se sabía cómo construir modelos con cualquier función del continuo sobre los cardinales regulares, pero, una vez determinada ésta, el resto de la exponenciación venía determinada por la construcción, además por un criterio muy simple. Esto podía deberse a que las técnicas conocidas no eran suficientemente generales o bien a un teorema desconocido que hiciese necesarias las restricciones encontradas. Además se conocía el enunciado de este hipotético teorema:

Definición 14.15 Llamaremos *hipótesis de los cardinales singulares* a la sentencia siguiente:

(HCS) *Para todo cardinal singular κ, si $2^{\mathrm{cf}\,\kappa} < \kappa$, entonces $\kappa^{\mathrm{cf}\,\kappa} = \kappa^+$.*

Notemos que la condición $2^{\mathrm{cf}\,\kappa} < \kappa$ ya implica que κ es singular. Lo expresamos explícitamente para enfatizar que la HCS sólo impone una restricción a los cardinales singulares.

Vamos a demostrar que, bajo la hipótesis de los cardinales singulares, la función del continuo sobre los cardinales regulares determina completamente la exponenciación cardinal (en particular la función del continuo sobre los cardinales singulares). En realidad la HCS no es un teorema de ZFC, pero —por razones que comentaremos más adelante— es difícil construir modelos donde no se cumpla. En particular, los modelos a los que nos referíamos antes cumplen todos esta hipótesis, y ésta es la razón de que en ellos la exponenciación cardinal esté determinada por la función del continuo. Precisamente por ello, podemos asegurar que la HCS es consistente con cualquier determinación de la función del continuo sobre los cardinales regulares compatible con la monotonía y con el teorema de König. Por otra parte, es inmediato que HCG $\to$ HCS, lo cual explica que la HCG determine la exponenciación cardinal.

Veamos ahora cómo la HCS determina la función del continuo sobre los cardinales singulares.

Ejemplo Supongamos que $\bigwedge\alpha(\aleph_\alpha \text{ regular} \to 2^{\aleph_\alpha} = \aleph_{\alpha+\omega+5})$. Entonces, usando el teorema 14.10 vemos que

$$\begin{aligned} 2^{\aleph_\omega} &= \aleph_{\omega+5}^{\aleph_0} = (2^{\aleph_0})^{\aleph_0} = \aleph_{\omega+5},\\ 2^{\aleph_{\omega_1}} &= \aleph_{\omega_1}^{\aleph_1} = \aleph_{\omega_1+1}, \end{aligned}$$

donde en la última igualdad hemos usado la HCS. Vamos a demostrar que la función del continuo en un cardinal singular puede calcularse siempre con uno de estos dos argumentos. ∎

Definición 14.16 Diremos que la función del continuo es *finalmente constante* bajo un cardinal límite κ si existe un $\mu < \kappa$ tal que si $\mu \le \nu < \kappa$ entonces $2^\nu = 2^\mu$.

En tal caso es obvio que $2^{<\kappa} = 2^\mu$. Notemos además que si la condición se cumple para todo ν regular, entonces se cumple para todo ν, por la monotonía. Así mismo, no perdemos generalidad si suponemos que μ es regular.

Teniendo esto en cuenta, el teorema siguiente nos permite calcular 2^κ para un cardinal singular κ supuesto que sabemos calcular 2^μ para todo cardinal regular $\mu < \kappa$. Más aún, lo que probamos es que la HCS implica que 2^κ toma siempre el mínimo valor posible:

Teorema 14.17 *Sea κ un cardinal singular.*

a) Si la función del continuo es finalmente constante bajo κ, entonces

$$2^\kappa = 2^{<\kappa}.$$

b) En caso contrario $2^\kappa \geq (2^{<\kappa})^+$ y si suponemos la HCS *tenemos la igualdad.*

DEMOSTRACIÓN: En el caso a), sea $\mu < \kappa$ tal que si $\mu \leq \nu < \kappa$ entonces $2^\nu = 2^\mu$. Así, $2^{<\kappa} = 2^\mu$ y, por 14.10, tenemos que $2^\kappa = (2^\mu)^{\operatorname{cf}\kappa} = 2^{\mu\operatorname{cf}\kappa} = 2^\mu$.

En el caso b), para todo cardinal $\mu < \kappa$ se cumple que $2^\mu < 2^{<\kappa}$. Por consiguiente, la aplicación $\kappa \longrightarrow 2^{<\kappa}$ dada por $\alpha \mapsto \kappa^{|\alpha|}$ es cofinal y creciente, luego el teorema 13.56 nos da que $\operatorname{cf} 2^{<\kappa} = \operatorname{cf}\kappa < \kappa$.

Por otra parte, por el teorema de König, $\operatorname{cf} 2^\kappa > \kappa$, luego $2^\kappa \neq 2^{<\kappa}$ y, como la desigualdad $2^{<\kappa} \leq 2^\kappa$ es obvia, tenemos en realidad que $(2^{<\kappa})^+ \leq 2^\kappa$.

Respecto a la otra desigualdad, tenemos que $2^{\operatorname{cf} 2^{<\kappa}} = 2^{\operatorname{cf}\kappa} < 2^{<\kappa}$, luego podemos aplicar la HCS a $2^{<\kappa}$, lo cual nos da que $(2^{<\kappa})^{\operatorname{cf} 2^{<\kappa}} = (2^{<\kappa})^+$, es decir, $2^\kappa = (2^{<\kappa})^{\operatorname{cf}\kappa} = (2^{<\kappa})^+$. ∎

Veamos ahora que la HCS determina toda la exponenciación cardinal a partir de la función del continuo:

Teorema 14.18 (HCS) *Sean κ y μ cardinales infinitos. Entonces*

$$\kappa^\mu = \begin{cases} \kappa & \text{si } 2^\mu < \kappa \wedge \mu < \operatorname{cf}\kappa, \\ \kappa^+ & \text{si } 2^\mu < \kappa \wedge \operatorname{cf}\kappa \leq \mu, \\ 2^\mu & \text{si } \kappa \leq 2^\mu. \end{cases}$$

DEMOSTRACIÓN: Si $\kappa \leq 2^\mu$, entonces $2^\mu \leq \kappa^\mu \leq (2^\mu)^\mu = 2^\mu$.

Observemos que en esta parte no hemos usado la HCS, así como tampoco hace falta para concluir que $\kappa \leq \kappa^\mu$ y que si $\operatorname{cf}\kappa \leq \mu$ entonces $\kappa^+ \leq \kappa^{\operatorname{cf}\kappa} \leq \kappa^\mu$. Así pues, lo que vamos a probar con la ayuda de la HCS es que κ^μ toma siempre el mínimo valor posible.

El caso $2^\mu < \kappa$ lo probamos por inducción sobre κ, es decir, lo suponemos cierto para todos los cardinales menores que κ.

Si $\kappa = \nu^+$, entonces $\mu < 2^\mu < \kappa = \operatorname{cf}\kappa$. Por lo tanto hemos de probar que $\kappa^\mu = \kappa$.

Tenemos que $2^\mu \leq \nu$. Si es $2^\mu < \nu$, entonces por hipótesis de inducción tenemos que $\nu^\mu = \nu$ o bien $\nu^\mu = \nu^+$, y en cualquier caso $\nu^\mu \leq \kappa$. Si, por el contrario, $2^\mu = \nu$ entonces $\nu^\mu = 2^\mu < \kappa$.

Por consiguiente podemos afirmar que $\nu^\mu \leq \kappa$. Por la fórmula de Hausdorff

$$\kappa^\mu = (\nu^+)^\mu = \nu^\mu\nu^+ = \nu^\mu\kappa = \kappa.$$

Consideramos ahora el caso en que κ es un cardinal límite. Si $\nu < \kappa$, por hipótesis de inducción tenemos que ν^μ es ν, ν^+ o 2^μ, pero en cualquier

caso $\nu^\mu < \kappa$ (si ν es finito no podemos aplicar la hipótesis de inducción, pero $\nu^\mu = 2^\mu$).

Si $\mu < \operatorname{cf}\kappa$, entonces

$$\kappa \le \kappa^\mu = |{}^\mu\kappa| \le \Big|\bigcup_{\alpha<\kappa} {}^\mu\alpha\Big| = \sum_{\alpha<\kappa} |\alpha|^\mu \le \sum_{\alpha<\kappa} \kappa = \kappa.$$

Por lo tanto $\kappa^\mu = \kappa$.

Si $\operatorname{cf}\kappa \le \mu$, expresemos $\kappa = \sum_{\alpha<\operatorname{cf}\kappa} \nu_\alpha$, donde $\nu_\alpha < \kappa$. Entonces

$$\kappa^\mu = \Big(\sum_{\alpha<\operatorname{cf}\kappa} \nu_\alpha\Big)^\mu \le \Big(\prod_{\alpha<\operatorname{cf}\kappa} \nu_\alpha\Big)^\mu = \prod_{\alpha<\operatorname{cf}\kappa} \nu_\alpha^\mu \le \prod_{\alpha<\operatorname{cf}\kappa} \kappa = \kappa^{\operatorname{cf}\kappa} \le \kappa^\mu,$$

luego $\kappa^\mu = \kappa^{\operatorname{cf}\kappa}$. Como $2^{\operatorname{cf}\kappa} \le 2^\mu < \kappa$, la HCS nos da que $\kappa^{\operatorname{cf}\kappa} = \kappa^+$ y tenemos la conclusión. ∎

Ejemplo Si suponemos la HCS y que

$$\textstyle\bigwedge\alpha(\aleph_\alpha \text{ regular } \to 2^{\aleph_\alpha} = \aleph_{\alpha+\omega+5}),$$

entonces

$$\aleph_5^{\aleph_3} = \aleph_{\omega+5}, \quad \aleph_{\omega_1}^{\aleph_3} = \aleph_{\omega_1+1}, \quad \aleph_{\omega_1+4}^{\aleph_3} = \aleph_{\omega_1+4}.$$

∎

Así pues, la exponenciación cardinal bajo la HCS no está determinada (pues la función del continuo sobre los cardinales regulares puede ser cualquiera que no contradiga a la monotonía ni al teorema de König) pero sí que está completamente comprendida, en cuanto que sabemos exactamente cómo depende de la función del continuo. El problema es que la HCS no es un teorema de ZFC (en la sección siguiente volveremos sobre esto), y lo que no está claro en absoluto es lo que se puede decir exclusivamente en ZFC sobre la exponenciación cardinal o sobre la función del continuo sobre los cardinales singulares. Si no suponemos la HCS sólo conocemos hechos aislados, algunos sencillos y otros muy profundos. Veamos un ejemplo de los sencillos:

Teorema 14.19 *Si* $2^{\aleph_1} < \aleph_\omega$ *y* $\aleph_\omega^{\aleph_0} \ge \aleph_{\omega_1}$, *entonces* $\aleph_\omega^{\aleph_0} = \aleph_{\omega_1}^{\aleph_1}$.

Demostración: Aplicamos la fórmula de Hausdorff:

$$\aleph_\omega^{\aleph_0} \le \aleph_{\omega_1}^{\aleph_1} \le (\aleph_\omega^{\aleph_0})^{\aleph_1} = \aleph_\omega^{\aleph_1} = \Big(\sum_{n\ge1} \aleph_n\Big)^{\aleph_1} \le \Big(\prod_{n\ge1} \aleph_n\Big)^{\aleph_1}$$

$$= \prod_{n\ge1} \aleph_n^{\aleph_1} = \prod_{n\ge1} 2^{\aleph_1}\aleph_n = 2^{\aleph_1}\aleph_\omega^{\aleph_0} = \aleph_\omega^{\aleph_0}.$$

∎

Consideremos ahora el valor de $\aleph_\omega^{\aleph_1}$. Se trata de un cardinal que queda invariante al elevarlo a $\aleph_0$, luego el teorema de König nos da que ha de tener

cofinalidad no numerable. Por su parte, la monotonía exige que sea mayor que el propio $\aleph_\omega$. Así pues, estas condiciones generales no excluyen la posibilidad de que $\aleph_\omega^{\aleph_0} = \aleph_{\omega_1}$. Más aún, si suponemos que $2^{\aleph_0} = \aleph_{\omega_1}$ (lo cual es consistente) entonces

$$\aleph_{\omega_1} = 2^{\aleph_0} \leq \aleph_\omega^{\aleph_0} \leq \aleph_{\omega_1}^{\aleph_0} = (2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0} = \aleph_{\omega_1},$$

con lo que, de hecho, $\aleph_\omega^{\aleph_0} = \aleph_{\omega_1}$.

Sin embargo, si suponemos que $2^{\aleph_1} < \aleph_\omega$ (lo cual es consistente), la HCS implica que $\aleph_\omega^{\aleph_0} = \aleph_{\omega+1}$, pero sin ella aún podemos asegurar que $\aleph_\omega^{\aleph_0} \neq \aleph_{\omega_1}$, ya que en caso contrario el teorema anterior nos daría $\aleph_{\omega_1}^{\aleph_1} = \aleph_{\omega_1}$, en contradicción con el teorema de König.

Así pues, nos encontramos con una restricción en ZFC al valor que puede tomar $\aleph_\omega^{\aleph_0}$ distinta de las que imponen la monotonía y el teorema de König. Una restricción que, además, depende de forma no trivial de los valores de $2^{\aleph_0}$ y $2^{\aleph_1}$. Si queremos un ejemplo en términos de la función del continuo podemos suponer que $\bigwedge n < \omega 2^{\aleph_n} < \aleph_\omega$, en cuyo caso tenemos que $2^{\aleph_\omega} = \aleph_\omega^{\aleph_0} \neq \aleph_{\omega_1}$.

Los resultados básicos sobre la exponenciación de cardinales fueron establecidos por Hausdorff y Tarski. Éste último probó un caso particular del teorema 13.47 y conjeturó que si $\{\kappa_\alpha\}_{\alpha<\lambda}$ es una sucesión estrictamente creciente de cardinales ≥ 2 y $\kappa = \sup\limits_{\alpha<\lambda} \kappa_\alpha$, entonces

$$\prod_{\alpha<\lambda} \kappa_\alpha = \kappa^{|\lambda|}.$$

Observemos que la restricción de que λ sea un ordinal límite es necesaria. Por ejemplo, si tomáramos $\lambda = \omega_1 + 1$ y la sucesión $\{\aleph_\alpha\}_{\alpha<\omega_1} \cup \{\aleph_{\omega_1\cdot 2}\}$ (con lo que $\kappa = \aleph_{\omega_1\cdot 2}$), suponiendo la HCG y aplicando 13.47 obtenemos que

$$\prod_{\alpha<\omega_1} \aleph_\alpha \cdot \aleph_{\omega_1\cdot 2} = \aleph_{\omega_1}^{\aleph_1} \cdot \aleph_{\omega_1\cdot 2} = \aleph_{\omega_1+1} \cdot \aleph_{\omega_1\cdot 2} = \aleph_{\omega_1\cdot 2} < \aleph_{\omega_1\cdot 2+1} = \aleph_{\omega_1\cdot 2}^{\aleph_1}.$$

Más en general, es necesario exigir que $\kappa_\alpha < \kappa$ para todo α. Un contraejemplo sin esta hipótesis (siempre bajo la HCG) sería $\lambda = \omega_1 + \omega$ y

$$\kappa_\alpha = \begin{cases} \aleph_\alpha & \text{si } \alpha < \omega_1, \\ \aleph_{\omega_1\cdot 2} & \text{si } \alpha = \omega_1 + n. \end{cases}$$

En tal caso $\kappa = \aleph_{\omega_1+\omega}$ y el producto sigue valiendo $\aleph_{\omega_1\cdot 2}^{\aleph_0} = \aleph_{\omega_1\cdot 2}$. Por otra parte la HCS implica la conjetura de Tarski:

Teorema 14.20 (HCS) *Sea λ un ordinal límite y $\{\kappa_\alpha\}_{\alpha<\lambda}$ una sucesión creciente (no exigimos que lo sea estrictamente) de cardinales ≥ 2. Sea $\kappa = \sup\limits_{\alpha<\lambda} \kappa_\alpha$ y supongamos que $\bigwedge \alpha < \lambda\ \kappa_\alpha < \kappa$. Entonces*

$$\prod_{\alpha<\lambda} \kappa_\alpha = \kappa^{|\lambda|}.$$

DEMOSTRACIÓN: La desigualdad $\leq$ es inmediata por la monotonía de los productos. Si $\kappa \leq 2^{|\lambda|}$ entonces, por el teorema 14.18,

$$\kappa^{|\lambda|} = 2^{|\lambda|} = \prod_{\alpha<\lambda} 2 \leq \prod_{\alpha<\lambda} \kappa_\alpha \leq \kappa^{|\lambda|}.$$

Si $|\lambda| < 2^{|\lambda|} < \kappa$, tomemos una sucesión de ordinales $\{\alpha_\delta\}_{\delta<\operatorname{cf}\lambda}$ cofinal creciente en λ. Entonces

$$\kappa^{|\lambda|} = \Big(\sum_{\delta<\operatorname{cf}\lambda} \kappa_{\alpha_\delta}\Big)^{|\lambda|} \leq \prod_{\delta<\operatorname{cf}\lambda} \kappa_{\alpha_\delta}^{|\lambda|} \leq \prod_{\delta<\operatorname{cf}\lambda} \kappa = \kappa^{\operatorname{cf}\lambda} = \prod_{\delta<\operatorname{cf}\lambda} \kappa_{\alpha_\delta} \leq \prod_{\alpha<\lambda} \kappa_\alpha.$$

Hemos usado que $\kappa_{\alpha_\delta}^{|\lambda|} < \kappa$ por el teorema 14.18. ∎

Nota Puede probarse —aunque es muy complicado— que es consistente que $\aleph_{\omega_1\cdot 2}$ sea un límite fuerte, $\aleph_{\omega_1}^{\aleph_1} = \aleph_{\omega_1\cdot 2+\omega+2}$ y $\aleph_{\omega_1\cdot 2+\omega}^{\aleph_0} = \aleph_{\omega_1\cdot 2+\omega+1}$. En estas condiciones tenemos un contraejemplo a la conjetura de Tarski. Basta tomar $\lambda = \omega_1 + \omega$ y

$$\kappa_\alpha = \begin{cases} \aleph_\alpha & \text{si } \alpha < \omega_1, \\ \aleph_{\omega_1\cdot 2+n} & \text{si } \alpha = \omega_1 + n. \end{cases}$$

En efecto, el producto da

$$\aleph_{\omega_1}^{\aleph_1} \cdot \aleph_{\omega_1\cdot 2+\omega}^{\aleph_0} \leq 2^{\aleph_{\omega_1}} \aleph_{\omega_1\cdot 2+\omega+1} = \aleph_{\omega_1\cdot 2+\omega+1} < \aleph_{\omega_1\cdot 2+\omega}^{|\omega_1+\omega|}.$$

∎

14.3 Cardinales fuertemente inaccesibles

Vamos a estudiar ahora una versión fuerte de los cardinales inaccesibles que introdujimos en el capítulo anterior.

Definición 14.21 Un cardinal infinito κ es un *límite fuerte* si para todo cardinal $\mu < \kappa$ se cumple $2^\mu < \kappa$.

Es claro que un cardinal límite fuerte es en particular un cardinal límite, ya que si fuera $\kappa = \mu^+$, entonces tendría que ser $2^\mu < \mu^+$, lo cual es imposible. Obviamente $\aleph_0$ es un cardinal límite fuerte.

Un cardinal *fuertemente inaccesible* es un cardinal límite fuerte regular distinto de $\aleph_0$.

En particular, todo cardinal fuertemente inaccesible es débilmente inaccesible, aunque el recíproco no es necesariamente cierto. Por consiguiente todo cuanto dijimos en el capítulo anterior sobre la imposibilidad de demostrar la existencia de cardinales inaccesibles vale también en el caso que nos ocupa, la diferencia es que con los cardinales fuertemente inaccesibles podremos entender mejor por qué es así.

Nota No hay acuerdo sobre si "cardinal inaccesible" significa "cardinal débilmente inaccesible" o bien "cardinal fuertemente inaccesible". En este libro significará "débilmente inaccesible".

Conviene observar que bajo la HCG todos los cardinales límite son límites fuertes y, en particular, los cardinales débilmente inaccesibles coinciden con los fuertemente inaccesibles.

También es claro que si κ es un límite fuerte, entonces $2^{<\kappa} = \kappa$. Más aún, si μ, $\nu < \kappa$, entonces $\mu^\nu < \kappa$, pues si $\xi < \kappa$ es el máximo de μ y ν, tenemos que $\mu^\nu \leq \xi^\xi = 2^\xi < \kappa$. Si κ es fuertemente inaccesible podemos decir más:

Teorema 14.22 *Si κ es un cardinal fuertemente inaccesible entonces $\kappa^{<\kappa} = \kappa$.*

DEMOSTRACIÓN: Basta probar que $\kappa^\mu \leq \kappa$ para todo $\mu < \kappa$. En efecto, como κ es regular ${}^\mu\kappa = \bigcup_{\alpha<\kappa} {}^\mu\alpha$ luego

$$\kappa^\mu \leq \sum_{\alpha<\kappa} |\alpha|^\mu \leq \sum_{\alpha<\kappa} \kappa = \kappa.$$

■

Del mismo modo que los cardinales límite pueden caracterizarse como los de la forma $\aleph_0$ o $\aleph_\lambda$, existe una caracterización similar para los cardinales límite fuerte, en términos de la llamada función bet.[1]

Definición 14.23 Definimos $\beth : \Omega \longrightarrow K$ (función bet) como la única función que cumple:

$$\beth_0 = \aleph_0 \quad \wedge \quad \bigwedge\alpha\, \beth_{\alpha+1} = 2^{\beth_\alpha} \quad \wedge \quad \bigwedge\lambda\, \beth_\lambda = \bigcup_{\delta<\lambda} \beth_\delta.$$

Teniendo en cuenta que el supremo de un conjunto de cardinales es un cardinal, una simple inducción prueba que $\beth$ toma todos sus valores en K. Obviamente es una función normal.

Ejercicio: La HCG es equivalente a que $\beth = \aleph$.

La caracterización a la que nos referíamos es:

Teorema 14.24 *Los cardinales límite fuerte son exactamente los de la forma $\beth_0$ o $\beth_\lambda$.*

DEMOSTRACIÓN: Se cumple que $\beth_\lambda$ es un límite fuerte, pues si $\kappa < \beth_\lambda$ entonces existe un $\delta < \lambda$ tal que $\kappa < \beth_\delta$, luego

$$2^\kappa \leq 2^{\beth_\delta} = \beth_{\delta+1} < \beth_{\delta+2} \leq \beth_\lambda.$$

Recíprocamente, si κ es un límite fuerte, entonces $\kappa \leq \beth_\kappa < \beth_{\kappa+1}$, luego podemos tomar el mínimo ordinal α tal que $\kappa < \beth_\alpha$. Ciertamente α no puede ser 0 ni un cardinal límite, luego $\alpha = \gamma + 1$ y, por consiguiente,

$$\beth_\gamma \leq \kappa < \beth_{\gamma+1} = 2^{\beth_\gamma}.$$

[1] Bet ($\beth$) es la segunda letra del alfabeto hebreo.

Si la primera desigualdad fuera estricta κ no sería un límite fuerte, luego $\kappa = \beth_\gamma$. Falta probar que γ no puede ser de la forma $\delta + 1$, pero es que en tal caso sería $\beth_\delta < \kappa$ y $2^{\beth_\delta} = \kappa$, y de nuevo κ no sería un límite fuerte. Por consiguiente $\gamma = 0$ o bien es un ordinal límite. ∎

La prueba del teorema siguiente es idéntica a la de su análogo 13.66:

Teorema 14.25 *Un cardinal regular κ es fuertemente inaccesible si y sólo si $\kappa = \beth_\kappa$.*

Es claro que V_ω es una unión numerable de conjuntos finitos, luego su cardinal es $|\mathrm{V}_\omega| = \aleph_0 = \beth_0$. A partir de aquí, una simple inducción nos da el teorema siguiente:

Teorema 14.26 $\bigwedge\alpha\ |\mathrm{V}_{\omega+\alpha}| = \beth_\alpha$. *En particular,* $\bigwedge\alpha(\omega^2 \le \alpha \to |\mathrm{V}_\alpha| = \beth_\alpha)$.

(Recordemos que si $\omega^2 \le \alpha$ entonces $\alpha = \omega^2 + \beta$ y $\omega + \alpha = \omega + \omega^2 + \beta = \omega(1 + \omega) + \beta = \omega^2 + \beta = \alpha$.)

De este modo, si κ es fuertemente inaccesible tenemos que $|\mathrm{V}_\kappa| = \kappa$. Más aún:

Teorema 14.27 *Si κ es un cardinal fuertemente inaccesible se cumple que*

$$\bigwedge x(x \in V_\kappa \leftrightarrow x \subset V_\kappa \wedge |x| < \kappa).$$

Demostración: Si $x \in \mathrm{V}_\kappa$, entonces $x \in \mathrm{V}_\delta$, para cierto $\delta < \kappa$ (podemos suponer $\omega^2 \le \delta$), luego $x \subset \mathrm{V}_\delta$ y $|x| \le |\mathrm{V}_\delta| = \beth_\delta < \beth_\kappa = \kappa$. Además $x \subset V_\kappa$ porque V_κ es transitivo.

Recíprocamente, si $x \subset V_\kappa$ y $|x| < \kappa$, entonces el conjunto

$$A = \{\mathrm{rang}\, y \mid y \in x\} \subset \kappa$$

es imagen de x, luego tiene cardinal menor que κ y, como κ es regular, A está acotado. Si $\delta < \kappa$ es una cota concluimos que $x \subset V_\delta$, luego $x \in V_{\delta+1} \subset V_\kappa$. ∎

Ahora ya estamos en condiciones de probar un hecho crucial:

Teorema 14.28 *Si κ es un cardinal fuertemente inaccesible, entonces*

$$\mathrm{V}_\kappa \vDash \mathrm{ZFC}.$$

Observaciones Cuando consideramos a un conjunto como modelo del lenguaje de la teoría de conjuntos, sobrentendemos que el relator de pertenencia[2] $\ulcorner\in\urcorner$ se interpreta como la relación de pertenencia $\in$.

[2]Usaremos los ángulos de Quine únicamente cuando haya posibilidad de confusión entre expresiones metamatemáticas y expresiones de la lógica formalizada, como ocurre aquí

La prueba se basa en que, en virtud del teorema 14.27, el conjunto V_κ es cerrado para todas las operaciones conjuntistas, en el sentido de que si $x, y \in V_\kappa$ es inmediato comprobar que $\{x, y\}$, (x, y), $x \cup y$, $x \cap y$, $x \times y$, $\mathcal{P}x$, etc. están también en V_κ. Más aún, si $f : a \longrightarrow b$ con $a, b \in \mathrm{V}_\kappa$, entonces $f \in \mathrm{V}_\kappa$. Similarmente se prueban hechos análogos.

Informalmente, esto se traduce en que si a un matemático le "mostráramos" únicamente los conjuntos de V_κ no echaría en falta nada: Si un axioma de la teoría de conjuntos dice que todo conjunto ha de tener un conjunto de partes, efectivamente, para cada conjunto $x \in \mathrm{V}_\kappa$, él "vería" a $\mathcal{P}x$, pues $\mathcal{P}x \in \mathrm{V}_\kappa$. Lo mismo vale para todos los demás axiomas, y eso es justo lo que vamos a comprobar. No ocurriría lo mismo si κ no fuera fuertemente inaccesible. Por ejemplo, si le mostráramos únicamente los conjuntos de V_ω "echaría en falta" los conjuntos infinitos, es decir, "se daría cuenta" de que falla el axioma de infinitud. Si le "mostráramos" sólo los conjuntos de V_{ω_1} "vería" conjuntos infinitos como ω o $\mathbb{R}$, pero, por ejemplo, $\mathbb{R}$ "no tendría cardinal", pues todos los ordinales que "vería" serían numerables, signo inequívoco de que falla algún axioma o, equivalentemente, que los conjuntos que "ve" no son todos los conjuntos.

Demostración: Veamos que $\mathrm{V}_\kappa \vDash \bigwedge xy(\bigwedge u(u \in x \leftrightarrow u \in y) \to x = y)$. Si aplicamos la definición de satisfacción en un modelo, vemos que esto equivale a

$$\bigwedge xy \in \mathrm{V}_\kappa(\bigwedge u \in \mathrm{V}_\kappa(u \in x \leftrightarrow u \in y) \to x = y).$$

Ahora bien, según las observaciones previas, si $x \in \mathrm{V}_\kappa$, entonces $u \in x$ ya implica $u \in \mathrm{V}_\kappa$, por lo que es redundante explicitarlo y esta sentencia es equivalente a

$$\bigwedge xy \in \mathrm{V}_\kappa(\bigwedge u(u \in x \leftrightarrow u \in y) \to x = y),$$

lo cual se cumple trivialmente.[3]

Similarmente, comprobar $\mathrm{V}_\kappa \vDash \bigwedge xy \bigvee z \bigwedge u(u \in z \leftrightarrow u = x \vee u = y)$ equivale a ver que

$$\bigwedge xy \in \mathrm{V}_\kappa \bigvee z \in \mathrm{V}_\kappa \bigwedge u \in \mathrm{V}_\kappa(u \in z \leftrightarrow u = x \vee u = y).$$

De nuevo la condición $u \in \mathrm{V}_\kappa$ es redundante, luego esta sentencia equivale a

$$\bigwedge xy \in \mathrm{V}_\kappa \bigvee z \in \mathrm{V}_\kappa \bigwedge u(u \in z \leftrightarrow u = x \vee u = y),$$

o también a

$$\bigwedge xy \in \mathrm{V}_\kappa \ \{x, y\} \in \mathrm{V}_\kappa,$$

lo cual es cierto.[4]

Del mismo modo se comprueba que el axioma de la unión se cumple en V_κ si y sólo si $\bigwedge x \in \mathrm{V}_\kappa \ \bigcup_{y \in x} x \in \mathrm{V}_\kappa$, lo cual es cierto, y el axioma del conjunto vacío se reduce a $\varnothing \in \mathrm{V}_\kappa$, lo cual también es cierto.

[3]Lo que hemos probado es que si alguien que "viva" en V_κ encuentra los mismos elementos en dos conjuntos x e y, es porque, ciertamente, x e y tienen los mismos elementos (ya que no deja de ver ninguno de los elementos de x e y), luego ambos conjuntos son el mismo.

[4]Con esto hemos probado que si alguien "viva" en V_κ ve dos conjuntos x e y, también ve el par formado por ellos, luego juzgará que se cumple el axioma del par.

Notemos que hasta aquí no hemos usado que κ sea fuertemente inaccesible, sino que cualquier ordinal límite hubiera servido. La hipótesis se requiere únicamente para probar el axioma del reemplazo.[5]. En efecto, fijamos $\phi \in \mathrm{Form}(\mathcal{L})$ y hemos de comprobar que

$$\mathrm{V}_\kappa \vDash \bigwedge xyz(\phi(x,y) \wedge \phi(x,z) \to y = z) \to \bigwedge a \bigvee b \bigwedge y(y \in b \leftrightarrow \bigvee x \in a\, \phi(x,y)).$$

Fijada una valoración v, esto equivale a

$$\bigwedge pqr \in \mathrm{V}_\kappa(\mathrm{V}_\kappa \vDash \phi[v^{pq}_{xy}] \wedge \mathrm{V}_\kappa \vDash \phi[v^{pr}_{xy}] \to q = r) \to$$

$$\bigwedge a \in \mathrm{V}_\kappa \bigvee b \in \mathrm{V}_\kappa \bigwedge q \in \mathrm{V}_\kappa(q \in b \leftrightarrow \bigvee p \in \mathrm{V}_\kappa(p \in a \wedge \mathrm{V}_\kappa \vDash \phi[v^{pq}_{xy}])).$$

Tras eliminar una condición redundante obtenemos:

$$\bigwedge pqr \in \mathrm{V}_\kappa(\mathrm{V}_\kappa \vDash \phi[v^{pq}_{xy}] \wedge \mathrm{V}_\kappa \vDash \phi[v^{pr}_{xy}] \to q = r) \to$$

$$\bigwedge a \in \mathrm{V}_\kappa \bigvee b \in \mathrm{V}_\kappa \bigwedge q \in \mathrm{V}_\kappa(q \in b \leftrightarrow \bigvee p \in a\ \ \mathrm{V}_\kappa \vDash \phi[v^{pq}_{xy}]).$$

Suponemos la hipótesis y tomamos un $a \in \mathrm{V}_\kappa$. Sea

$$D = \{p \in a \mid \bigvee q \in \mathrm{V}_\kappa\ \mathrm{V}_\kappa \vDash \phi[v^{pq}_{xy}]\}.$$

Sea $F : D \longrightarrow \mathrm{V}_\kappa$ la función que a cada $p \in a$ le asigna el único $q \in \mathrm{V}_\kappa$ que cumple $\mathrm{V}_\kappa \vDash \phi[v^{pq}_{xy}]$. Notemos que la unicidad la tenemos por la primera parte del axioma que estamos suponiendo. Sea $b = F[D] \subset \mathrm{V}_\kappa$. Es claro que b cumple lo que requiere el axioma del reemplazo. Sólo falta probar que $b \in \mathrm{V}_\kappa$.

Para ello partimos de que ciertamente $D \in \mathrm{V}_\kappa$ (pues si $a \in \mathrm{V}_\alpha$ con $\alpha < \kappa$ entonces $D \in \mathrm{V}_{\alpha+1}$), luego, por la observación previa al teorema, $|D| < \kappa$ y, en consecuencia, $|b| < \kappa$. El conjunto de los rangos de los elementos de b es un subconjunto de κ de cardinal menor que κ. Como κ es regular, tiene que estar acotado, es decir, existe un ordinal $\alpha < \kappa$ tal que $b \subset \mathrm{V}_\alpha$, luego $b \in \mathrm{V}_{\alpha+1}$ y, por consiguiente $b \in \mathrm{V}_\kappa$.

Con esto tenemos que $\mathrm{V}_\kappa \vDash \mathrm{ZF}^*$. La comprobación del axioma de regularidad es tan elemental como la del axioma de extensionalidad o el axioma del par. Veamos el axioma de partes. Hemos de probar que

$$\mathrm{V}_\kappa \vDash \bigwedge x \bigvee y \bigwedge u(u \in y \leftrightarrow \bigwedge v(v \in u \to v \in x)).$$

Esto equivale a

$$\bigwedge x \in \mathrm{V}_\kappa \bigvee y \in \mathrm{V}_\kappa \bigwedge u \in \mathrm{V}_\kappa(u \in y \leftrightarrow \bigwedge v \in \mathrm{V}_\kappa(v \in u \to v \in x)).$$

La condición $v \in \mathrm{V}_\kappa$ es redundante, y al suprimirla tenemos simplemente

$$\bigwedge x \in \mathrm{V}_\kappa \bigvee y \in \mathrm{V}_\kappa \bigwedge u \in \mathrm{V}_\kappa(u \in y \leftrightarrow u \subset x).$$

[5]Por comodidad la usaremos también en el axioma de elección, aunque podríamos haberlo evitado

Ahora resulta ser redundante la condición $u \in \mathrm{V}_\kappa$, con lo que nos quedamos con $\bigwedge x \in \mathrm{V}_\kappa\, \mathcal{P}x \in \mathrm{V}_\kappa$, lo cual es cierto.[6]

La comprobación de los axiomas de infinitud y elección presenta el inconveniente de que en ellos aparecen varios términos definidos que sería tedioso reducir por completo al relator de pertenencia. En lugar de ello, vamos a analizar la interpretación en V_κ de estos conceptos. Ante todo conviene simplificar un poco la notación: sustituiremos las variables de $\ulcorner\mathfrak{L}\urcorner$ (la formalización del lenguaje de ZFC) por los conjuntos de V_κ que denotan respecto a una valoración dada. Por ejemplo, en lugar de escribir

$$\mathrm{V}_\kappa \vDash (x : y \longrightarrow z)[v^{fab}_{xyz}],$$

donde x, y, $z \in \mathrm{Var}(\mathfrak{L})$ y f, a, $b \in \mathrm{V}_\kappa$, escribiremos

$$\mathrm{V}_\kappa \vDash f : a \longrightarrow b.$$

Así mismo conviene que abreviemos $M = \mathrm{V}_\kappa$. Notemos que M satisface todas las consecuencias de los axiomas que ya hemos demostrado. En particular

$$M \vDash \bigwedge xyu(u \in \{x, y\} \leftrightarrow u = x \vee u = y).$$

Esto significa que[7]

$$\bigwedge xyu \in \mathrm{V}_\kappa(u \in M(\{x, y\}) \leftrightarrow u = x \vee u = y).$$

Puesto que el conjunto $M(\{x, y\})$ (es decir, el objeto denotado por el término $\ulcorner\{u, v\}\urcorner \in \mathrm{Term}(\ulcorner\mathfrak{L}\urcorner)$ cuando las variables u, $v \in \mathrm{Var}(\ulcorner\mathfrak{L}\urcorner)$ se interpretan como los conjuntos x, y) está en V_κ, el hecho de que u pertenezca a este conjunto ya implica que está en V_κ, y lo mismo sucede si $u = x \vee u = y$, luego la fórmula anterior se simplifica hasta

$$\bigwedge xy \in \mathrm{V}_\kappa \bigwedge u(u \in M(\{x, y\}) \leftrightarrow u = x \vee u = y),$$

luego concluimos que[8] $\bigwedge xy \in \mathrm{V}_\kappa\ M(\{x, y\}) = \{x, y\}$.

Razonando igualmente con la sentencia

$$\bigwedge xyu(u \in (x, y) \leftrightarrow u = \{x, x\} \vee u = \{x, y\})$$

obtenemos que

$$\bigwedge xy \in \mathrm{V}_\kappa \bigwedge u(u \in M((x, y)) \leftrightarrow u = M(\{x, x\}) \vee u = M(\{x, y\}),$$

[6] Hemos probado que si alguien "vive" en V_κ, entonces, para cada x que vea, verá también a $\mathcal{P}x$, luego no encontrará ningún contraejemplo al axioma de partes.

[7] Notemos que en la línea anterior x, y, u representan variables de $\ulcorner\mathfrak{L}\urcorner$ y, abusando de la notación, en la línea siguiente pasan a ser conjuntos de V_κ.

[8] Alguien que "viva" en V_κ no verá más ordinales que los menores que κ, luego al ver a κ creerá estar viendo la clase de todos los ordinales. Así, al ocultarle una parte de los conjuntos le llevamos a engaño, de modo que lo que él llama Ω no es Ω, sino κ. Sin embargo, ahora acabamos de probar que nunca podremos engañarle acerca de cuál es el par formado por dos conjuntos: lo que él llama "par desordenado formado por x e y" es realmente el par desordenado formado por x e y.

es decir,

$$\bigwedge xy \in \mathrm{V}_\kappa \bigwedge u(u \in M((x,y)) \leftrightarrow u = \{x\} \vee u = \{x,y\}),$$

de donde $\bigwedge xy \in \mathrm{V}_\kappa\ M((x,y)) = (x,y)$.

Ahora usamos que M ha de satisfacer la sentencia

$$\bigwedge fab(f : a \longrightarrow b \leftrightarrow \bigwedge x \in f \bigvee uv(u \in a \wedge v \in b \wedge x = (u,v)) \wedge$$
$$\bigwedge u \in a \bigvee v \in b\ (u,v) \in f).$$

Tras eliminar las redundancias oportunas obtenemos que

$$\bigwedge fab \in \mathrm{V}_\kappa(M \vDash (f : a \longrightarrow b) \leftrightarrow \bigwedge x \in f \bigvee uv(u \in a \wedge v \in b \wedge x = (u,v)) \wedge$$
$$\bigwedge u \in a \bigvee v \in b\ (u,v) \in f)),$$

luego en definitiva

$$\bigwedge fab \in \mathrm{V}_\kappa(M \vDash (f : a \longrightarrow b) \leftrightarrow f : a \longrightarrow b).$$

Lo mismo puede decirse de $f : a \longrightarrow b$ inyectiva, suprayectiva o biyectiva, sin más que considerar las fórmulas obvias que determinan estas propiedades.[9]

Llamemos $f : \omega \longrightarrow \omega$ a la función $f(n) = n + 1$. Es claro que $\omega,\ f \in \mathrm{V}_\kappa$, luego, según lo que acabamos de ver,[10]

$$M \vDash f : \omega \longrightarrow \omega \text{ inyectiva no suprayectiva},$$

y, por consiguiente,

$$M \vDash \bigvee xs(s : x \longrightarrow x \text{ inyectiva no suprayectiva}),$$

es decir, M cumple el axioma de infinitud.

Puesto que M satisface el axioma de regularidad, tiene que satisfacer el siguiente teorema de ZF + V = R:

$$\bigwedge x(x \text{ es un ordinal} \leftrightarrow \bigwedge uv(u \in v \wedge v \in x \to u \in x) \wedge$$
$$\bigwedge uv \in x(u \in v \vee v \in u \vee u = v).$$

(Podemos suprimir la condición de buena fundación en la definición de ordinal porque si V = R todo conjunto está bien fundado.)

Al expresar el hecho que que M satisface esta sentencia junto a las simplificaciones usuales obtenemos simplemente que

$$\bigwedge x \in \mathrm{V}_\kappa(M \vDash x \text{ es un ordinal} \leftrightarrow x \in \Omega).$$

[9]Es decir, que alguien "viva" en V_κ no puede equivocarse al juzgar si un conjunto f es o no una aplicación entre dos conjuntos dados y, de serlo, sabrá también si es inyectiva, etc.

[10]Alguien que "viva" en V_κ verá a ω y a f y se dará cuenta de que f es una aplicación inyectiva y no suprayectiva en ω, porque así es y hemos probado que en estas cosas no puede equivocarse.

Consideremos ahora la sentencia:

$$\bigwedge xy(\overline{\overline{x}} = \overline{\overline{y}} \leftrightarrow \bigvee f\ f : x \longrightarrow y \text{ biyectiva}).$$

Al expresar que M la satisface obtenemos

$$\bigwedge xy \in \mathrm{V}_\kappa(M \vDash \overline{\overline{x}} = \overline{\overline{y}} \leftrightarrow \bigvee f \in \mathrm{V}_\kappa\ f : x \longrightarrow y \text{ biyectiva}).$$

Ahora bien, como hemos observado antes de empezar la prueba, si $x, y \in \mathrm{V}_\kappa$, el hecho de que $f : x \longrightarrow y$ ya implica que $f \in \mathrm{V}_\kappa$, luego podemos eliminar esta condición y así

$$\bigwedge xy \in \mathrm{V}_\kappa(M \vDash \overline{\overline{x}} = \overline{\overline{y}} \leftrightarrow \bigvee f\ f : x \longrightarrow y \text{ biyectiva}),$$

es decir,

$$\bigwedge xy \in \mathrm{V}_\kappa(M \vDash \overline{\overline{x}} = \overline{\overline{y}} \leftrightarrow \overline{\overline{x}} = \overline{\overline{y}}).$$

Teniendo esto en cuenta ya es fácil probar que M satisface el teorema de numerabilidad y, por consiguiente, el axioma de elección: dado $x \in \mathrm{V}_\kappa$, sabemos que $\mu = |x| < \kappa$, luego $\mu \in \mathrm{V}_\kappa$ y $\overline{\overline{x}} = \overline{\mu}$. Así pues:

$$\bigwedge x \in \mathrm{V}_\kappa \bigvee y \in \mathrm{V}_\kappa(y \text{ es un ordinal} \wedge \overline{\overline{x}} = \overline{\overline{y}}),$$

lo que equivale a

$$M \vDash \bigwedge x \bigvee y(y \text{ es un ordinal} \wedge \overline{\overline{x}} = \overline{\overline{y}}),$$

y esto es el teorema de numerabilidad. ∎

Nota Para algunas observaciones que haremos después conviene ir un poco más lejos en la línea de resultados que hemos obtenido en el transcurso de la prueba anterior. Ahora sabemos que M satisface todos los teoremas de ZFC. En particular

$$M \vDash \bigwedge x(x \text{ es un cardinal} \leftrightarrow x \text{ es un ordinal} \wedge \neg \bigvee y \in x\ \overline{\overline{x}} = \overline{\overline{y}}).$$

Al desarrollar esto obtenemos que

$$\bigwedge x \in \mathrm{V}_\kappa(M \vDash x \text{ es un cardinal} \leftrightarrow x \in K).$$

Consideramos ahora que

$$M \vDash \bigwedge xu(u \in \mathcal{P}x \leftrightarrow u \subset x),$$

lo que se reduce a

$$\bigwedge xu \in \mathrm{V}_\kappa(u \in M(\mathcal{P}x) \leftrightarrow u \subset x),$$

pero si $x \in \mathrm{V}_\kappa$, el hecho de que $u \subset x$ ya implica que $u \in \mathrm{V}_\kappa$, y lo mismo vale para $u \in M(\mathcal{P}x)$, luego tenemos que

$$\bigwedge x \in \mathrm{V}_\kappa(\bigwedge u(u \in M(\mathcal{P}x) \leftrightarrow u \subset x)),$$

es decir,

$$\bigwedge x \in \mathrm{V}_\kappa \; M(\mathcal{P}x) = \mathcal{P}x.$$

Por último usamos que

$$M \vDash (x \text{ es un cardinal f.i.} \leftrightarrow x \text{ es un cardinal} \wedge \bigwedge u \in x \bigvee v \in x \; \overline{\overline{\mathcal{P}u}} = \overline{\overline{v}}$$

$$\wedge \bigwedge y((y \subset x \wedge \bigwedge u \in x \bigvee v \in y \; u \in v) \to \overline{\overline{y}} = \overline{\overline{x}})).$$

Al desarrollar esta sentencia obtenemos —volviendo a la notación habitual— que

$$\bigwedge u \in \mathrm{V}_\kappa(\mathrm{V}_\kappa \vDash (x \text{ es un cardinal f.i.})[v_x^u] \leftrightarrow u \text{ es un cardinal f.i.}). \tag{14.1}$$

■

Observaciones En particular, el teorema 14.28 implica que

$$\underset{ZFC}{\vdash} \bigvee \kappa \, \kappa \text{ es fuertemente inaccesible} \to \mathrm{Consis}\, ZFC,$$

por lo que los teoremas de incompletitud nos aseguran que si ZFC es consistente no es posible demostrar a partir de sus axiomas la existencia de cardinales fuertemente inaccesibles. No obstante, la prueba del teorema 14.28 nos da un argumento directo para llegar a esta conclusión que no requiere de los teoremas de Gödel. En efecto: si pudiera probarse la existencia de un cardinal fuertemente inaccesible κ, podríamos tomar el menor de todos ellos μ. Entonces V_μ sería un modelo de ZFC, luego debería cumplirse

$$\mathrm{V}_\mu \vDash \bigvee \kappa \, \kappa \text{ es fuertemente inaccesible},$$

es decir,

$$\bigvee \kappa \in \mathrm{V}_\mu \; \mathrm{V}_\mu \vDash x \text{ es fuertemente inaccesible } [v_x^\kappa].$$

Ahora bien, según la nota posterior al teorema, esto implica que

$$\bigvee \kappa \in \mathrm{V}_\mu \; \kappa \text{ es fuertemente inaccesible}.$$

Pero $\kappa \in \mathrm{V}_\mu$ implica $\kappa < \mu$, lo que contradice la minimalidad de μ. En resumen: no se puede probar la existencia de cardinales fuertemente inaccesibles porque el mínimo de tales cardinales proporciona un modelo donde no existen tales cardinales.

Un axioma que postule la existencia de un cardinal fuertemente inaccesible es esencialmente análogo al axioma de infinitud. En efecto: Un modelo de ZFC$-$AI es el conjunto de todos los conjuntos hereditariamente finitos. En este modelo, los conjuntos infinitos son clases propias, los únicos ordinales son los números naturales. Postular la existencia de un conjunto infinito equivale a ampliar el universo de trabajo, de modo que lo que antes era la clase de todos los ordinales pasa a ser ω, un cardinal límite fuerte regular (es decir, un

cardinal fuertemente inaccesible si no lo hubiéramos excluido arbitrariamente en la definición), lo que antes era la clase de todos los conjuntos ahora pasa a ser V_ω, el conjunto de todos los conjuntos hereditariamente finitos. Para que los axiomas de ZFC−AI se sigan cumpliendo cuando forzamos la existencia de un conjunto infinito cualquiera es necesario incorporar a los modelos conjuntos de cardinales enormes (la sucesión de los alefs). Sabemos que podemos encontrar un modelo así en el que no haya cardinales fuertemente inaccesibles (supuesta, como es habitual, la consistencia de ZFC). Si postulamos la existencia de uno de estos cardinales, lo que estamos haciendo es ampliar de nuevo el universo por un proceso idéntico, en virtud del cual lo que antes era la clase de todos los ordinales ahora pasa a ser un cardinal fuertemente inaccesible κ y lo que era la clase de todos los conjuntos pasa a ser el conjunto V_κ. Los axiomas de ZFC fuerzan ahora a que la sucesión de los alefs se prolongue mucho más allá de κ, igual que antes habíamos prolongado la sucesión de los cardinales mucho más allá de ω.

Naturalmente el proceso puede realizarse de nuevo. Llamemos

$$\mathrm{FI}_2 \equiv \bigvee \kappa_1\kappa_2(\kappa_1 < \kappa_2 \wedge \kappa_1,\ \kappa_2 \text{ son fuertemente inaccesibles}),$$

mientras que FI_1 será la existencia de un cardinal fuertemente inaccesible. Es claro que de la existencia de un cardinal fuertemente inaccesible no se deduce la existencia de dos de ellos, es decir, $\neg \underset{ZFC}{\vdash} (\mathrm{FI}_1 \to \mathrm{FI}_2)$. En efecto, si así fuera, trabajando en ZFC+FI_1 dispondríamos al menos de dos cardinales fuertemente inaccesibles, luego podríamos tomar el mínimo κ y el siguiente μ, pero entonces $V_\mu \vDash \mathrm{ZFC} + \mathrm{FI}_1 + \neg\mathrm{FI}_2$, cuando, por otra parte, el hecho de que V_μ cumpla FI_1 obligaría a que cumpliera FI_2, luego tenemos una contradicción.

Más aún, hemos probado que $\underset{\mathrm{ZFC}+\mathrm{FI}_2}{\vdash} \mathrm{Consis}\,\mathrm{ZFC} + \mathrm{FI}_1$, luego la consistencia de la existencia de dos cardinales fuertemente inaccesibles no puede probarse ni siquiera suponiendo la consistencia de que exista uno de ellos (ya que en tal caso tendríamos

$$\underset{\mathrm{ZFC}}{\vdash} \mathrm{Consis}\,\mathrm{ZFC} + \mathrm{FI}_1 \to \mathrm{Consis}\,\mathrm{ZFC} + \mathrm{FI}_2,$$

luego $\underset{\mathrm{ZFC}+\mathrm{FI}_2}{\vdash} \mathrm{Consis}\,\mathrm{ZFC} + \mathrm{FI}_2$, y el teorema de Gödel nos daría que ZFC+FI_2 (luego también ZFC+FI_1) sería contradictoria.

En general, las teorías

$$\mathrm{ZFC} - \mathrm{AI}, \quad \mathrm{ZFC}, \quad \mathrm{ZFC} + \mathrm{FI}_1, \quad \mathrm{ZFC} + \mathrm{FI}_2, \quad \mathrm{ZFC} + \mathrm{FI}_3, \quad \ldots$$

—donde FI_n se define de forma obvia— forman una sucesión con la propiedad de que la consistencia de cada una de ellas no puede demostrarse ni siquiera suponiendo la consistencia de las anteriores (y, de hecho, sólo sabemos probar la consistencia de la primera). Lo importante es que el hecho de que no podamos probar su consistencia no debe ser tomado como un indicio de que no sean consistentes. Simplemente es una cuestión técnica relacionada con el segundo teorema de incompletitud. Esto es importante porque muchos resultados de

consistencia requieren como hipótesis la consistencia de alguna de estas teorías, o incluso de otras más fuertes.

Por ejemplo, es conocida la existencia de subconjuntos de $\mathbb{R}$ que no son medibles Lebesgue, así como que la prueba depende drásticamente del axioma de elección, cuando todos los resultados básicos de la teoría de la medida pueden probarse usando tan sólo versiones débiles de este axioma (esencialmente se requieren elecciones numerables). Por ello es natural preguntarse si el uso del axioma de elección es realmente necesario o podría evitarse mediante razonamientos más finos. En otros términos, si llamamos *axioma de Solovay* AS a la sentencia "todo subconjunto de $\mathbb{R}$ es medible Lebesgue", podemos preguntarnos si $\neg$AS puede probarse en ZF (es decir, sin el axioma de elección) o si, por el contrario, ZF+AS es consistente. Pues bien, R. Solovay demostró en 1970 que

$$\operatorname{Consis}(\mathrm{ZFC} + \mathrm{FI}_1) \to \operatorname{Consis}(\mathrm{ZF} + \mathrm{AS}).$$

y en 1984 S. Shelah probó la implicación contraria, con lo que de hecho[11]

$$\operatorname{Consis}(\mathrm{ZF} + \mathrm{AS}) \leftrightarrow \operatorname{Consis}(\mathrm{ZFC} + \mathrm{FI}_1).$$

Observemos que ya sabíamos que era imposible demostrar la consistencia de ZF+AS, pues en particular tendríamos la consistencia de ZF, la cual —según demostró Gödel— es equivalente a la consistencia de ZFC. A lo máximo a lo que podíamos aspirar es a una prueba de consistencia relativa, es decir, a probar que

$$\operatorname{Consis}\mathrm{ZFC} \to \operatorname{Consis}\mathrm{ZF} + \mathrm{AS},$$

pero la equivalencia anterior nos muestra que esto tampoco es posible, que la consistencia de ZF+AS no puede demostrarse ni siquiera suponiendo la de ZFC, pero, por otra parte, al ser equivalente a la consistencia de ZFC+FI_1, lo cierto es que ya no tenemos motivos serios para desconfiar de la consistencia de ZF+AS. Ahora sabemos que podemos confiar en esta teoría en la misma medida en que podemos confiar en que "no hay peligro" en "aumentar" una vez el universo conjuntista más allá de lo estrictamente necesario para que se cumplan los axiomas de ZFC.

Observemos también que las teorías FI_n no son las únicas extensiones posibles de ZFC en esta línea. Por ejemplo, podemos postular que existen $\aleph_0$ cardinales fuertemente inaccesibles, o $\aleph_1$ o incluso una clase propia de ellos, es decir, podemos postular que

$$\bigwedge\alpha\bigvee\kappa(\alpha < \kappa \wedge \kappa \text{ es fuertemente inaccesible}).$$

Cada una de estas teorías es más fuerte que las anteriores en el sentido que ya hemos explicado.

Por último comentaremos que la técnica que empleó Gödel para demostrar la consistencia de la HCG permite construir a partir de un modelo cualquiera

[11] En realidad, en lugar de ZF deberíamos poner ZF más una versión débil del axioma de elección, suficiente para desarrollar toda la teoría de la medida básica, sin la cual no tendría sentido AS.

de ZFC otro modelo con los mismos cardinales débilmente inaccesibles y donde se satisfaga la HCG, por lo que éstos pasan a ser cardinales fuertemente inaccesibles. Por ello, a efectos de consistencia los cardinales débilmente inaccesibles son equivalentes a los fuertemente inaccesibles, es decir,

$$\underset{\text{ZFC}}{\vdash} \text{Consis}\,(\bigvee\kappa\ \kappa \text{ es d.i.}) \leftrightarrow \text{Consis}\,(\bigvee\kappa\ \kappa \text{ es f.i.}),$$

y lo mismo es válido para todos los axiomas FI_n, etc. (su consistencia es equivalente a la de las versiones con cardinales débilmente inaccesibles). ∎

Tras estas observaciones podemos formarnos una primera idea del status de la HCS en la teoría axiomática de conjuntos: sucede que a partir de ¬HCS puede probarse la consistencia de que existan infinitos cardinales inaccesibles. De hecho puede probarse la consistencia de un axioma mucho más fuerte que todos los que hemos considerado. Por ello es imposible demostrar la consistencia de ¬HCS a partir de la mera consistencia de ZFC. En otras palabras, para construir un modelo donde falle la HCS es necesario partir de un modelo que contenga muchos cardinales inaccesibles, y ésta es la razón por la que los primeros resultados de consistencia sobre la función del continuo "se encontraban" siempre con la HCS aunque no se buscara expresamente. En el capítulo siguiente estaremos en condiciones de aproximarnos algo más a la hipótesis que requiere una prueba de consistencia de ¬HCS.

Terminamos la sección con una aplicación de la función $\beth$ que no tiene nada que ver con cardinales inaccesibles. El *axioma de elección de Gödel* es la sentencia

$$(AEG) \qquad \bigvee F(F : \mathrm{V} \longrightarrow \mathrm{V} \wedge \bigwedge x(x \neq \varnothing \to F(x) \in X)).$$

Este axioma involucra esencialmente clases propias, luego no puede ser considerado como sentencia de ZFC. Sólo tiene sentido como extensión de NBG. El axioma de elección de Gödel postula la existencia de una función de elección sobre la clase universal, por lo que implica trivialmente el axioma de elección de Zermelo, que sólo postula la existencia de una función de elección (distinta) para cada conjunto.

Teorema 14.29 (NBG+AEG) *Todas las clases propias son equipotentes.*

Demostración: Basta observar que podemos descomponer V y Ω en respectivas clases de conjuntos disjuntos como sigue:

$$\mathrm{V} = \mathrm{V}_\omega \cup \bigcup_{\alpha\in\Omega} (\mathrm{V}_{\omega+\alpha+1} \setminus \mathrm{V}_{\omega+\alpha}), \qquad \Omega = \beth_0 \cup \bigcup_{\alpha\in\Omega} (\beth_{\alpha+1} \setminus \beth_\alpha).$$

Teniendo en cuenta el teorema 14.26 y la aritmética cardinal básica es claro que

$$|\mathrm{V}_{\omega+\alpha+1} \setminus \mathrm{V}_{\omega+\alpha}| = \beth_{\alpha+1} = |\beth_{\alpha+1} \setminus \beth_\alpha|.$$

El axioma de elección de Gödel nos permite elegir funciones

$$f_\alpha : \mathrm{V}_{\omega+\alpha+1} \setminus \mathrm{V}_{\omega+\alpha} \longrightarrow \beth_{\alpha+1} \setminus \beth_\alpha \text{ biyectivas.}$$

Por otra parte es claro que podemos tomar $f^* : V_\omega \longrightarrow \beth_0$ biyectiva. Con todas estas funciones podemos construir

$$F = f^* \cup \bigcup_{\alpha\in\Omega} f_\alpha : \mathrm{V} \longrightarrow \Omega \text{ biyectiva.}$$

Así, si A es cualquier clase propia, $F[A]$ es una subclase de Ω, es decir, una clase bien ordenada por una relación conjuntista. Por el teorema 11.27 concluimos que $F[A]$ es semejante a Ω (y A es equipotente a $F[A]$), luego toda clase propia es equipotente a Ω. ∎

Observemos que el axioma de regularidad —al contrario de lo que suele suceder— desempeña un papel crucial en la prueba anterior. En estas condiciones tenemos una nueva caracterización de las clases propias: una clase es propia si y sólo si su tamaño es comparable al de la clase universal.

Capítulo XV

Conjuntos cerrados no acotados

Introducimos en este último capítulo uno de los conceptos más importantes que aparecen al profundizar en el estudio de la teoría de conjuntos. En el fondo se trata de la topología de orden en los ordinales, si bien se puede omitir de forma natural toda referencia explícita a la topología. Entre otras aplicaciones demostraremos un profundo teorema de Silver (1974) sobre la función del continuo en los cardinales singulares. Como en el capítulo anterior, trabajamos en NBG o, equivalentemente, en ZFC.

15.1 Conjuntos cerrados no acotados

Empezamos introduciendo la noción de subconjunto cerrado en un ordinal. El lector familiarizado con la topología de orden puede probar que la definición que damos coincide con la de conjunto cerrado para dicha topología. No obstante, la definición usual es la que damos a continuación, pues es mucho más práctica que la topológica.

Definición 15.1 Sea λ un ordinal límite. Un conjunto $C \subset \lambda$ es *cerrado* en λ si cuando un ordinal límite $\delta < \lambda$ cumple que $\delta \cap C$ no está acotado en δ, entonces $\delta \in C$.

Intuitivamente la definición exige que si C contiene puntos arbitrariamente cercanos a un límite δ (topológicamente: si δ es un punto de acumulación de C), entonces δ está en C. Una caracterización útil es la siguiente:

Teorema 15.2 *Sea λ un ordinal límite. Un subconjunto C de λ es cerrado en λ si y sólo si para todo $X \subset C$ no vacío y acotado en λ se cumple que* $\sup X \in C$. *Equivalentemente: para todo $X \subset C$ no vacío, si* $\sup X \in \lambda$, *entonces* $\sup X \in C$.

DEMOSTRACIÓN: Supongamos que C es cerrado y sea X un subconjunto en las condiciones indicadas. Llamemos $\delta = \sup X$.

Si $\delta \in X$ entonces $\delta \in C$. Supongamos que $\delta \notin X$ y veamos que igualmente $\delta \in C$. En primer lugar, δ es un ordinal límite, pues si fuera $\delta = 0$ tendría que ser $X = \{0\}$ y si $\beta < \delta$ entonces $\beta < \alpha$ para cierto $\alpha \in X$, luego $\alpha \leq \delta$, pero, como $\delta \notin X$, ha de ser $\alpha < \delta$, luego $\beta + 1 < \alpha + 1 \leq \delta$.

En realidad hemos probado también que $\delta \cap C$ no está acotado en δ, pues, dado $\beta < \delta$, hemos encontrado un $\alpha \in \delta \cap C$ mayor que β. Por definición de cerrado concluimos que $\delta \in C$.

Recíprocamente, si C tiene la propiedad indicada y $\delta < \lambda$ es un ordinal límite tal que $\delta \cap C$ no está acotado en δ, es claro que $\delta = \sup(\delta \cap C)$, luego $\delta \in C$. ∎

Definición 15.3 En lo sucesivo las iniciales c.n.a. significarán "cerrado no acotado", es decir, diremos que C es c.n.a. en λ si es cerrado en λ y no está acotado en λ.

Teorema 15.4 *Sea λ un ordinal límite de cofinalidad no numerable, $\beta < \operatorname{cf}\lambda$ y $\{C_\alpha\}_{\alpha<\beta}$ una familia de conjuntos c.n.a. en λ. Entonces se cumple que $\bigcap_{\alpha<\beta} C_\alpha$ es c.n.a. en λ.*

DEMOSTRACIÓN: Del teorema anterior se sigue inmediatamente que la intersección de cualquier familia de cerrados es cerrada. Sólo queda probar que $\bigcap_{\alpha<\beta} C_\alpha$ no está acotado.

Sea $f_\alpha : \lambda \longrightarrow \lambda$ la función dada por $f_\alpha(\delta) = \text{mín}\{\epsilon \in C_\alpha \mid \delta < \epsilon\}$. La definición es correcta porque C_α no está acotado en λ. Para todo $\delta < \lambda$ tenemos que $\delta < f_\alpha(\delta) \in C_\alpha$.

Sea ahora $g : \lambda \longrightarrow \lambda$ la función dada por $g(\delta) = \sup_{\alpha<\beta} f_\alpha(\delta)$. Notemos que $g(\delta) \in \lambda$ por la hipótesis de que $\beta < \operatorname{cf}\lambda$. Es claro que si $\delta < \lambda$ entonces $\delta < g(\delta) \leq g^\omega(\delta) < \lambda$, donde g^ω es la función iterada de g definida en 13.58.

Además $g^\omega(\delta)$ es un ordinal límite, pues si $\alpha < g^\omega(\delta)$, entonces $\alpha \in g^n(\delta)$ para cierto $n \in \omega$ y así $\alpha + 1 \leq g^n(\delta) < g(g^n(\delta)) = g^{n+1}(\delta) \leq g^\omega(\delta)$.

Se cumple que $g^\omega(\delta) \cap C_\alpha$ no está acotado en $g^\omega(\delta)$, pues si $\gamma \in g^\omega(\delta)$, entonces $\gamma \in g^n(\delta) < f_\alpha(g^n(\delta)) \in C_\alpha$ y $f_\alpha(g^n(\delta)) \leq g(g^n(\delta)) = g^{n+1}(\delta) < g^\omega(\delta)$, o sea, $\gamma < f_\alpha(g^n(\delta)) \in g^\omega(\delta) \cap C_\alpha$.

Como C_α es cerrado podemos concluir que $g^\omega(\delta) \in C_\alpha$, y esto para todo $\alpha < \beta$, luego $\delta < g^\omega(\delta) \in \bigcap_{\alpha<\beta} C_\alpha$, lo que prueba que la intersección es no acotada. ∎

Ahora veamos algunos ejemplos de cerrados no acotados. El primero es, de hecho, una caracterización de los subconjuntos cerrados no acotados de un cardinal regular no numerable.

Teorema 15.5 *Sea κ un cardinal regular no numerable. Un conjunto $C \subset \kappa$ es c.n.a. en κ si y sólo si existe una función normal $f : \kappa \longrightarrow \kappa$ tal que $f[\kappa] = C$.*

DEMOSTRACIÓN: Por el teorema 11.26, ord$C \leq \kappa$ y como $|C| = \kappa$ (porque C no está acotado en κ y κ es regular), ha de ser ord$C = \kappa$. Sea, pues, $f : \kappa \longrightarrow C$ la semejanza. Basta probar que $f : \kappa \longrightarrow \kappa$ es normal. Claramente sólo hay que ver que si $\lambda < \kappa$ entonces $f(\lambda) = \bigcup_{\delta<\lambda} f(\delta)$. Ahora bien, $\lambda = \sup_\kappa\{\delta \mid \delta < \lambda\}$, luego, al ser f una semejanza,

$$f(\lambda) = \sup_C\{f(\delta) \mid \delta < \lambda\} = \bigcup_{\delta<\lambda} f(\delta).$$

En efecto, como κ es regular tenemos que $\bigcup_{\delta<\lambda} f(\delta) \in \kappa$ y como C es cerrado tenemos que $\bigcup_{\delta<\lambda} f(\delta) \in C$, luego obviamente se trata del supremo del conjunto $\{f(\delta) \mid \delta < \lambda\}$.

Supongamos ahora que C es el rango de una función normal f. Entonces $|C| = \kappa$ y en consecuencia C no está acotado en κ. Si $\delta < \kappa$ es un ordinal límite tal que $\delta \cap C$ no está acotado en δ, entonces sea $\lambda = \{\alpha < \kappa \mid f(\alpha) < \delta\}$. Es fácil ver que λ es un ordinal límite y $f|_\lambda : \lambda \longrightarrow \delta$ es inyectiva, luego $|\lambda| \leq |\delta| < \kappa$, de donde $\lambda < \kappa$. Por consiguiente podemos calcular

$$f(\lambda) = \bigcup_{\alpha<\lambda} f(\alpha) = \sup(\delta \cap C) = \delta \in C,$$

luego C es cerrado. ∎

El teorema 13.60 prueba que una función normal en un cardinal regular no numerable tiene un conjunto no acotado de puntos fijos. Ahora probamos que dicho conjunto también es cerrado.

Teorema 15.6 *Sea κ un cardinal regular no numerable y $f : \kappa \longrightarrow \kappa$ una función normal. Entonces $F = \{\alpha < \kappa \mid f(\alpha) = \alpha\}$ es c.n.a. en κ.*

DEMOSTRACIÓN: Ya sabemos que F no está acotado en κ. Veamos que es cerrado. Para ello tomamos $\lambda < \kappa$ tal que $\lambda \cap F$ no esté acotado en λ. Entonces $f(\lambda) = \bigcup_{\delta<\lambda} f(\delta)$. Si $\delta < \lambda$, entonces existe un $\eta \in F$ tal que $\delta < \eta$, luego $f(\delta) < f(\eta) = \eta < \lambda$, y por consiguiente concluimos que $f(\lambda) \leq \lambda$. La otra desigualdad se da por ser f normal (teorema 11.22), luego $\lambda \in F$, que es, por tanto, cerrado. ∎

La mayor parte de las ocasiones en que se dice que un conjunto es evidentemente c.n.a. se está apelando tácitamente al siguiente teorema:

Teorema 15.7 *Sea κ un cardinal regular no numerable y A un conjunto de aplicaciones $f : {}^n\kappa \longrightarrow \kappa$, donde n es un número natural que depende de f. Supongamos que $|A| < \kappa$. Entonces el conjunto*

$$C = \{\alpha < \kappa \mid \bigwedge fn(f \in A \wedge f : {}^n\kappa \longrightarrow \kappa \to f[{}^n\alpha] \subset \alpha)\}$$

es c.n.a. en κ.

DEMOSTRACIÓN: Sea $\lambda < \kappa$ tal que $C \cap \lambda$ no esté acotado en λ, sea $f \in A$, $f :^n \kappa \longrightarrow \kappa$, tomemos ordinales $\epsilon_1, \ldots, \epsilon_n \in \lambda$ y sea $\beta \in C \cap \lambda$ mayor que todos ellos.

Así $f(\epsilon_1, \ldots, \epsilon_n) \in f[^n\beta] \subset \beta < \lambda$, luego $\bigwedge x \in {}^n\lambda\ f(x) \in \lambda$, es decir, $f[^n\lambda] \subset \lambda$, lo que implica que $\lambda \in C$, el cual es, por tanto, cerrado.

Sea $\alpha \in \kappa$. Definimos recurrentemente una sucesión $\{\alpha_m\}$ de ordinales en κ. Tomamos $\alpha_0 = \alpha$ y, supuesto definido α_m, para cada $f \in A$, $f : {}^n\kappa \longrightarrow \kappa$ sea β_f el mínimo ordinal tal que $f[^n\alpha_m] \subset \beta_f < \kappa$ (existe porque $|f[^n\alpha_m]| \leq |^n\alpha_m| < \kappa$, luego $f[^n\alpha_m]$ está acotado en κ).

Definimos $\alpha_{m+1} = \bigcup_{f \in A} \beta_f < \kappa$, pues $|A| < \kappa$ y κ es regular. Finalmente definimos $\alpha^* = \sup_{m \in \omega} \alpha_m \in \kappa$. Claramente $\alpha \leq \alpha^*$. Si probamos que $\alpha^* \in C$ tendremos que C es no acotado.

Tomemos una función $f \in A$, $f :^n \kappa \longrightarrow \kappa$ y ordinales $\epsilon_1, \ldots, \epsilon_n \in \alpha^*$. Entonces existe un natural m tal que $\epsilon_1, \ldots, \epsilon_n \in \alpha_m$ y así

$$f(\epsilon_1, \ldots, \epsilon_n) \in f[^n\alpha_m] \subset \beta_f \leq \alpha_{m+1} \leq \alpha^*,$$

luego $f[\alpha^*] \subset \alpha^*$ y, consecuentemente, $\alpha^* \in C$. ∎

Si κ es un cardinal regular no numerable, el teorema 15.4 nos da que la intersección de una cantidad menor que κ de subconjuntos c.n.a. es c.n.a. Obviamente esto no es cierto para familias cualesquiera de κ conjuntos, pero sí se cumple un hecho parecido y de gran utilidad. Para enunciarlo necesitamos una definición:

Definición 15.8 Sea $\{X_\alpha\}_{\alpha<\kappa}$ una familia de subconjuntos de un cardinal κ. Llamaremos *intersección diagonal* de la familia al conjunto

$$\triangle_{\alpha<\kappa} X_\alpha = \{\gamma < \kappa \mid \gamma \in \bigcap_{\alpha<\gamma} X_\alpha\}.$$

Si intentamos probar algo "razonable" y "nos gustaría" que una intersección de κ conjuntos c.n.a. fuera c.n.a. es probable que en realidad nos baste lo siguiente:

Teorema 15.9 *Sea κ un cardinal regular no numerable y $\{C_\alpha\}_{\alpha<\kappa}$ una familia de conjuntos c.n.a. en κ. Entonces $\triangle_{\alpha<\kappa} C_\alpha$ es c.n.a. en κ.*

DEMOSTRACIÓN: Por abreviar, llamaremos D a la intersección diagonal. Tomemos $\lambda < \kappa$ tal que $\lambda \cap D$ no esté acotado en λ. Hemos de probar que $\lambda \in D$, es decir, tomamos $\alpha < \lambda$ y hemos de ver que $\lambda \in C_\alpha$. A su vez, para ello basta probar que $\lambda \cap C_\alpha$ no está acotado en λ, pero si $\beta \in \lambda$ tenemos que existe un $\epsilon \in \lambda \cap D$ tal que $\alpha, \beta < \epsilon$. Como $\epsilon \in D$ se cumple que $\epsilon \in C_\alpha \cap \lambda$, luego, efectivamente, $C_\alpha \cap \lambda$ no está acotado en λ y D es cerrado.

Para cada $\beta < \kappa$, el teorema 15.4 nos permite tomar $g(\beta) \in \bigcap_{\alpha<\beta} C_\alpha$ tal que $\beta < g(\beta)$. Tenemos así una función $g : \kappa \longrightarrow \kappa$.

Por el teorema 15.7, el conjunto

$$C = \{\lambda \in \kappa \mid g[\lambda] \subset \lambda\}$$

es c.n.a. en κ (en principio tenemos que es c.n.a. el conjunto de todos los ordinales $\alpha < \kappa$ tales que $g[\alpha] \subset \alpha$, pero es claro que el conjunto de los $\lambda < \kappa$ también es c.n.a., y C es la intersección de ambos conjuntos). Si probamos que $C \subset D$ tendremos que D no está acotado.

Sea $\lambda \in C$. Tomamos $\alpha < \lambda$ y hemos de ver que $\lambda \in C_\alpha$, para lo cual se ha de cumplir que $\lambda \cap C_\alpha$ no esté acotado en λ. Ahora bien, si $\delta < \lambda$, tomamos $\epsilon \in \lambda$ tal que $\alpha, \delta < \epsilon$. Así $\delta < g(\epsilon) \in \lambda \cap C_\alpha$. ∎

15.2 Conjuntos estacionarios

Los resultados que hemos obtenido sobre los conjuntos cerrados no acotados en un ordinal límite λ se interpretan más adecuadamente con ayuda de la noción siguiente:

Definición 15.10 Sea λ un ordinal límite de cofinalidad no numerable. Definimos el *filtro de cerrados no acotados* en λ como el conjunto

$$\text{c.n.a.}(\lambda) = \{X \subset \lambda \mid \bigvee C(C \subset X \wedge C \text{ es c.n.a. en } \lambda\} \subset \mathcal{P}\lambda.$$

Es inmediato comprobar las propiedades siguientes:

a) $\lambda \in \text{c.n.a.}(\lambda)$, $\varnothing \notin \text{c.n.a.}(\lambda)$,

b) Si $X \subset Y \subset \lambda$ y $X \in \text{c.n.a.}(\lambda)$ entonces $Y \in \text{c.n.a.}(\lambda)$,

c) Si $\{X_\alpha\}_{\alpha<\beta}$ es una familia de elementos de $\text{c.n.a.}(\lambda)$ con $\beta < \text{cf}\,\lambda$, entonces

$$\bigcap_{\alpha<\beta} X_\alpha \in \text{c.n.a.}(\lambda).$$

Informalmente, podemos pensar que los elementos de $\text{c.n.a.}(\lambda)$ son subconjuntos "muy grandes" de λ. Así, la propiedad a) dice que λ es muy grande, mientras que $\varnothing$ no lo es, b) afirma que todo conjunto que contenga a otro muy grande es muy grande, y c) dice que la intersección de menos de $\text{cf}\,\lambda$ conjuntos grandes sigue siendo un conjunto grande.

Los complementarios de los conjuntos "muy grandes" son los conjuntos "muy pequeños". Definimos el *ideal de cerrados no acotados* de λ como el conjunto

$$\text{c.n.a.}(\lambda)' = \{X \subset \lambda \mid \lambda \setminus X \in \text{c.n.a.}(\lambda)\}.$$

Las propiedades siguientes se deducen inmediatamente de las del filtro:

a) $\varnothing \in \text{c.n.a.}(\lambda)'$, $\lambda \notin \text{c.n.a.}(\lambda)'$,

b) Si $X \subset Y \subset \lambda$ e $Y \in \text{c.n.a.}(\lambda)'$ entonces $X \in \text{c.n.a.}(\lambda)'$,

c) Si $\{X_\alpha\}_{\alpha<\beta}$ es una familia de elementos de c.n.a.$(\lambda)'$ con $\beta < \operatorname{cf}\lambda$, entonces

$$\bigcup_{\alpha<\beta} X_\alpha \in \text{c.n.a.}(\lambda)'.$$

Ejercicio: Probar que si λ es un ordinal de cofinalidad no numerable y $X \subset \lambda$, $|X| < \operatorname{cf}\lambda$, entonces $X \in \text{c.n.a.}(\lambda)'$.

El interés de estos conceptos se debe a que, por ejemplo, si tenemos una familia de menos de κ subconjuntos "muy grandes" de un cardinal regular κ, sabemos que la intersección será también "muy grande", y en particular será no vacía, luego podremos tomar ordinales que cumplan simultáneamente las propiedades que definen a todos los conjuntos de la familia. No obstante, es frecuente tener que trabajar con conjuntos que no son "muy grandes", pero puede ser suficiente con que no sean "muy pequeños".

Definición 15.11 Sea λ un ordinal de cofinalidad no numerable. Un conjunto $E \subset \lambda$ es *estacionario* en λ si $E \notin \text{c.n.a.}(\lambda)'$.

He aquí las propiedades elementales de los conjuntos estacionarios:

Teorema 15.12 *Sea λ un ordinal de cofinalidad no numerable y $E \subset \lambda$. Se cumple:*

a) Si E es c.n.a en λ entonces E es estacionario en λ.

b) E es estacionario en λ si y sólo si corta a todo c.n.a. en λ.

c) Si E es estacionario en λ entonces no está acotado en λ.

d) Si E es estacionario en λ y C es c.n.a. en λ entonces $E \cap C$ es estacionario en λ.

Demostración: a) es inmediato: si E no fuera estacionario entonces $\lambda \setminus E$ contendría un c.n.a. disjunto de E.

b) E es estacionario si y sólo si $\lambda \setminus E \notin \text{c.n.a.}(\lambda)$, si y sólo si no existe ningún c.n.a. C tal que $C \subset \lambda \setminus E$, si y solo si todo c.n.a. C corta a E.

c) Se sigue de b) junto con el hecho obvio de que si $\alpha \in \lambda$ entonces $\lambda \setminus \alpha$ es c.n.a.

d) Si C' es otro c.n.a. en λ, entonces $C \cap C'$ es c.n.a., luego $E \cap C \cap C' \neq \varnothing$ por b), luego, también por b), $E \cap C$ es estacionario. ∎

Veamos ahora un ejemplo de conjuntos estacionarios:

Teorema 15.13 *Sea λ un ordinal límite de cofinalidad no numerable y $\kappa < \operatorname{cf}\lambda$ un cardinal regular. Entonces el conjunto*

$$\{\alpha < \lambda \mid \operatorname{cf}\alpha = \kappa\}$$

es estacionario en λ.

DEMOSTRACIÓN: Llamemos E al conjunto del enunciado. Sea C un c.n.a. en λ y sea $\alpha = \mathrm{ord}C$. Tenemos que $|\alpha| = |C| \geq \mathrm{cf}\,\lambda > \kappa$. Por lo tanto $\kappa < \alpha$. Sea $f : \alpha \longrightarrow C$ la semejanza. Igual que en la prueba de 15.5 se ve que f es una función normal, por lo que $\mathrm{cf}\, f(\kappa) = \mathrm{cf}\,\kappa = \kappa$, de modo que $f(\kappa) \in C \cap E$. Por el teorema anterior concluimos que E es estacionario. ∎

Ejemplo Los conjuntos

$$\{\alpha < \omega_2 \mid \mathrm{cf}\,\alpha = \aleph_0\} \quad \text{y} \quad \{\alpha < \omega_2 \mid \mathrm{cf}\,\alpha = \aleph_1\}$$

son estacionarios disjuntos en ω_2, luego vemos que la intersección de conjuntos estacionarios no es necesariamente estacionaria. Más aún, de aquí se deduce que existen conjuntos estacionarios que no son cerrados. ∎

Veamos ahora una caracterización muy útil de los conjuntos estacionarios en un cardinal regular. Para ello necesitamos una definición:

Definición 15.14 Si $A \subset \kappa$, una aplicación $f : A \longrightarrow \kappa$ es *regresiva* si

$$\bigwedge \alpha \in A(\alpha \neq 0 \to f(\alpha) < \alpha).$$

Teorema 15.15 (Fodor) *Sea κ un cardinal regular no numerable y $E \subset \kappa$. Las afirmaciones siguientes son equivalentes:*

a) E es estacionario en κ,

b) Si $f : E \longrightarrow \kappa$ es regresiva, existe un $\alpha < \kappa$ tal que

$$f^{-1}[\{\alpha\}] = \{\beta \in E \mid f(\beta) = \alpha\}$$

es estacionario en κ,

c) Si $f : E \longrightarrow \kappa$ es regresiva, existe un $\alpha < \kappa$ tal que

$$f^{-1}[\{\alpha\}] = \{\beta \in E \mid f(\beta) = \alpha\}$$

no está acotado en κ.

DEMOSTRACIÓN: a) → b) Si f es regresiva pero $f^{-1}[\{\alpha\}]$ no es estacionario para ningún $\alpha < \kappa$, tomemos un c.n.a. C_α tal que $C_\alpha \cap f^{-1}[\{\alpha\}] = \varnothing$. Entonces $D = \underset{\alpha<\kappa}{\triangle} C_\alpha$ es también c.n.a. en κ. Por consiguiente $E \cap D$ es estacionario, luego podemos tomar $\gamma \in E \cap D$, $\gamma \neq 0$. En particular $\gamma \in \bigcap_{\alpha<\gamma} C_\alpha$. Sea $\delta = f(\gamma) < \gamma$. Así $\gamma \in f^{-1}[\{\delta\}] \cap C_\delta = \varnothing$.

b) → c) es obvio.

c) → a) Si E no es estacionario, sea C un c.n.a. en κ tal que $C \cap E = \varnothing$. Sea $f : E \longrightarrow \kappa$ la aplicación dada por $f(\alpha) = \sup(C \cap \alpha)$. Claramente $f(\alpha) \leq \alpha$, pero $f(\alpha) \in C$ y $\alpha \in E$, luego de hecho $f(\alpha) < \alpha$ y f es regresiva.

Por otra parte, si $\gamma < \kappa$, como C no está acotado, existe un $\alpha \in C$ tal que $\gamma < \alpha$. Vamos a probar que $f^{-1}[\{\gamma\}] \subset \alpha + 1$, es decir, que $f^{-1}[\{\gamma\}]$ está acotado en κ para todo γ, en contradicción con c). En efecto, si $\delta \in E$ y $\alpha < \delta$, entonces $f(\delta) = \sup(C \cap \delta) \geq \alpha$, pues $\alpha \in C \cap \delta$. Así pues, $f(\delta) \neq \gamma$. ∎

Terminamos la sección con un resultado nada trivial sobre conjuntos estacionarios que fue probado por Solovay en 1971. No lo vamos a usar después, pero tiene aplicaciones importantes. Necesitamos un resultado previo auxiliar.

Teorema 15.16 *Sea κ un cardinal regular no numerable y sea E un subconjunto estacionario en κ. Entonces el conjunto*

$$T = \{\lambda \in E \mid \operatorname{cf}\lambda = \aleph_0 \vee (\operatorname{cf}\lambda > \aleph_0 \wedge E \cap \lambda \text{ no es estacionario en } \lambda\}$$

es estacionario en κ.

Demostración: Tomamos un conjunto c.n.a. C en κ. Hemos de probar que $C \cap T \neq \varnothing$. Sea

$$C' = \{\lambda \in C \mid C \cap \lambda \text{ no está acotado en } \lambda\}.$$

Veamos que C' es c.n.a. Por el teorema 15.5 sabemos que existe una función normal $f : \kappa \longrightarrow \kappa$ tal que $f[\kappa] = C$. Por otra parte, el conjunto $L = \{\lambda | \lambda < \kappa\}$ es claramente c.n.a. en κ, luego existe $g : \kappa \longrightarrow \kappa$ normal tal que $g[\kappa] = L$. Sea $h = g \circ f : \kappa \longrightarrow \kappa$. Basta ver que $C' = h[\kappa]$.

Si $\alpha \in \kappa$, entonces $g(\alpha) \in \kappa$ es un ordinal límite, luego

$$h(\alpha) = f(g(\alpha)) = \bigcup_{\delta \in g(\alpha)} f(\delta),$$

donde cada $f(\delta) \in C$, luego $h(\alpha) \cap C$ no está acotado en $h(\alpha)$. Así pues, $h(\alpha) \in C'$.

Tomemos ahora $\lambda \in C'$. Sea $\alpha < \kappa$ tal que $f(\alpha) = \lambda$. Si $\alpha = 0$ entonces λ es el mínimo de C, luego $C \cap \lambda = \varnothing$ está acotado en λ, lo cual contradice que $\lambda \in C'$.

Si $\alpha = \beta + 1$ entonces $f(\beta)$ es una cota de $C \cap \lambda$ en λ, pues $f(\beta) \in f(\beta+1) = \lambda$ y, si $\delta \in C \cap \lambda$ entonces $\delta = f(\gamma)$ para un $\gamma \in \kappa$. Así, $\delta < \lambda$, $f(\gamma) < f(\alpha)$, $\gamma < \alpha = \beta + 1$, $\gamma \leq \beta$, $\delta = f(\gamma) \leq f(\beta)$, luego también $C \cap \lambda$ resulta estar acotado en λ y tenemos otra contradicción.

La única posibilidad es que α sea un límite, luego existe $\epsilon < \kappa$ tal que $\alpha = g(\epsilon)$, y así $\lambda = h(\epsilon) \in h[\kappa]$.

Como E es estacionario y C' es c.n.a. tenemos que $E \cap C' \neq \varnothing$. Sea λ el mínimo de $E \cap C'$. Si $\operatorname{cf}\lambda = \aleph_0$ entonces $\lambda \in T \cap C \neq \varnothing$. Supongamos que $\operatorname{cf}\lambda > \aleph_0$. Como $\lambda \in C'$ tenemos que $\lambda \cap C$ no está acotado en λ. Vamos a probar que, de hecho, $\lambda \cap C'$ no está acotado en λ. Para ello consideramos la aplicación $f : \lambda \longrightarrow \lambda$ que a cada $\alpha \in \lambda$ le asigna el mínimo ordinal en $\lambda \cap C$ mayor que α.

Vamos a probar que si $\alpha < \lambda$, entonces $f^\omega(\alpha) \in \lambda \cap C'$ y, desde luego, $\alpha \leq f^\omega(\alpha)$. Teniendo en cuenta que $\delta < f(\delta)$ para todo $\delta < \lambda$, es claro que

la sucesión $f^n(\alpha)$ es estrictamente creciente de ordinales de $\lambda \cap C$. De aquí deducimos que su supremo $f^\omega(\alpha)$ es un ordinal límite y $f^\omega(\alpha) \cap C$ no está acotado en $f^\omega(\alpha)$. Como C es cerrado concluimos que $f^\omega(\alpha) \in C$ y de aquí a su vez que $f^\omega(\alpha) \in \lambda \cap C'$.

Por otra parte, es inmediato comprobar que $\lambda \cap C'$ es cerrado en λ, luego se trata de un c.n.a. Ahora bien, $(\lambda \cap C') \cap (\lambda\mathcal{E}) \subset \lambda \cap (\mathcal{C}' \cap \mathcal{E}) = \varnothing$, porque λ es el mínimo de $C' \cap E$. Esto significa que $E \cap \lambda$ no es estacionario en λ, luego $\lambda \in T \cap C' \neq \varnothing$. ∎

No es fácil encontrar ejemplos de conjuntos estacionarios disjuntos en $\aleph_1$. Sin embargo, lo cierto es que existen, como se desprende del siguiente teorema general:

Teorema 15.17 (Solovay) *Sea κ un cardinal regular no numerable y A un conjunto estacionario en κ. Entonces existen conjuntos $\{E_\alpha\}_{\alpha<\kappa}$ estacionarios en κ y disjuntos dos a dos tales que*

$$A = \bigcup_{\alpha<\kappa} E_\kappa.$$

Demostración: Sea

$$T = \{\lambda \in A \mid \operatorname{cf}\lambda = \aleph_0 \vee (\operatorname{cf}\lambda > \aleph_0 \wedge A \cap \lambda \text{ no es estacionario en } \lambda\},$$

que según el teorema anterior es estacionario en κ. Para cada $\lambda \in T$ tomemos $f_\lambda : \operatorname{cf}\lambda \longrightarrow \lambda$ cofinal y normal. Veamos que si $\operatorname{cf}\lambda > \aleph_0$ podemos exigir que $f_\lambda[\operatorname{cf}\lambda] \cap T = \varnothing$.

En efecto, si $\operatorname{cf}\lambda > \aleph_0$ tenemos que $A \cap \lambda$ no es estacionario en λ, luego tampoco lo es $T \cap \lambda$. Por consiguiente existe un c.n.a. C en λ de manera que $C \cap T \cap \lambda = \varnothing$. Definimos $f^*_\lambda : \operatorname{cf}\lambda \longrightarrow \lambda$ mediante

$f^*_\lambda(0) = \min C$,

$f^*_\lambda(\alpha+1) =$ mínimo ordinal $\epsilon \in C$ tal que $f^*_\lambda(\alpha) < \epsilon$ y $f_\lambda(\alpha) < \epsilon$.

$f^*_\lambda(\lambda') = \bigcup_{\delta<\lambda'} f^*_\lambda(\delta)$.

Claramente f^*_λ es normal y una simple inducción (usando que C es cerrado en el caso límite) prueba que $f^*_\lambda : \operatorname{cf}\lambda \longrightarrow C$. Como $f_\lambda(\alpha) < f^*_\lambda(\alpha+1)$ para todo $\alpha < \lambda$ y f_λ es cofinal, es claro que f^*_λ también lo es, y además $f^*_\lambda[\operatorname{cf}\lambda] \cap T \subset C \cap T = \varnothing$.

En lo sucesivo suprimiremos los asteriscos. Veamos ahora que existe un $\delta < \kappa$ tal que para todo $\epsilon < \kappa$ el conjunto

$$F_\epsilon = \{\lambda \in T \mid \delta < \operatorname{cf}\lambda \wedge f_\lambda(\delta) \geq \epsilon\}$$

es estacionario en κ.

En caso contrario, para todo $\delta < \kappa$ existe un $\epsilon(\delta) < \kappa$ y un c.n.a. C_δ en κ tales que

$$\{\lambda \in T \mid \delta < \operatorname{cf}\lambda \wedge f_\lambda(\delta) \geq \epsilon(\delta)\} \cap C_\delta = \varnothing.$$

Equivalentemente, para todo $\delta < \kappa$ existe un $\epsilon(\delta) < \kappa$ y un c.n.a. C_δ en κ tales que si $\lambda \in T \cap C_\delta$ y $\delta < \operatorname{cf}\lambda$, entonces $f_\lambda(\delta) < \epsilon(\delta)$.

Sea $C = \triangle_{\alpha<\kappa} C_\alpha$, que es c.n.a. en κ. Si $\lambda \in C \cap T$, entonces

$$\bigwedge \delta < \operatorname{cf}\lambda\ f_\lambda(\delta) < \epsilon(\delta)$$

(puesto que $\lambda \in T \cap C_\delta$).

Claramente, $D^*_\delta = \{\gamma \in \kappa \mid \epsilon(\delta) < \gamma\} = \kappa \setminus (\epsilon(\delta)+1)$ es c.n.a. en κ. Por consiguiente, $D_\delta = \{\gamma \in C \mid \epsilon(\delta) < \gamma\} = C \cap D^*_\delta$ es c.n.a. en κ y, a su vez, $D = \{\gamma \in C \mid \bigwedge \delta < \gamma\ \epsilon(\delta) < \gamma\} = \triangle_{\delta<\kappa} D_\delta$ es c.n.a. en κ.

En consecuencia $T \cap D$ es estacionario y, en particular, tiene al menos dos elementos $\gamma < \lambda$. Veamos que

$(*)$ Si $\delta < \gamma$ y $\delta < \operatorname{cf}\lambda$, entonces $f_\lambda(\delta) < \epsilon(\delta) < \gamma$.

En efecto, $\lambda \in D$, $\lambda \in C \cap T$, $f_\lambda(\delta) < \epsilon(\delta)$ y, como $\gamma \in D$, también $\epsilon(\delta) < \gamma$.

Como f_λ es cofinal, existe un $\delta < \operatorname{cf}\lambda$ (podemos tomarlo infinito) tal que $\gamma \le f_\lambda(\delta)$, luego —por lo que acabamos de probar— $\gamma \le \delta < \operatorname{cf}\lambda$. En particular la condición $\delta < \operatorname{cf}\lambda$ es redundante en $(*)$, y además tenemos que $\operatorname{cf}\lambda > \aleph_0$.

Tenemos, pues, que si $\delta < \gamma$ entonces $f_\lambda(\delta) < \gamma$, luego $f_\lambda(\gamma) = \bigcup_{\delta<\gamma} f_\lambda(\delta) \le \gamma$. Como f_λ es normal tenemos, de hecho, la igualdad $f_\lambda(\gamma) = \gamma$, pero esto es imposible, pues $\gamma \in T$ y $f_\lambda(\gamma) \notin T$.

Con esto hemos encontrado un $\delta < \kappa$ tal que para todo $\epsilon < \kappa$ el conjunto F_ϵ es estacionario en κ. Sea $g : T \longrightarrow \kappa$ la función dada por $g(\lambda) = f_\lambda(\delta)$, obviamente regresiva.

Para cada $\epsilon < \kappa$ tenemos que $g|_{F_\epsilon} : F_\epsilon \longrightarrow \kappa$ es regresiva, luego por 15.15 existe un $\gamma_\epsilon < \kappa$ tal que $G_\epsilon = (g|_{F_\epsilon})^{-1}(\{\gamma_\epsilon\})$ es estacionario en κ.

Si $\lambda \in G_\epsilon$, entonces $\gamma_\epsilon = g(\lambda) = f_\lambda(\delta) \ge \epsilon$ (porque $G_\epsilon \subset F_\epsilon$). Así pues, $\bigwedge \epsilon < \kappa\ \epsilon \le \gamma_\epsilon$.

Por consiguiente, el conjunto $B = \{\gamma_\epsilon \mid \epsilon < \kappa\}$ no está acotado en κ, luego tiene cardinal κ. Sea $h : \kappa \longrightarrow B$ biyectiva y sea $E_\alpha = G_{h(\alpha)}$. Así, los conjuntos E_α son estacionarios en κ y disjuntos dos a dos, pues si $\gamma_\epsilon \ne \gamma_{\epsilon'}$ entonces $G_\epsilon \cap G_{\epsilon'} = \varnothing$. Además $E_\alpha = G_{h(\alpha)} \subset F_{h(\alpha)} \subset T \subset A$.

Sea $U = A \setminus \bigcup_{\alpha<\kappa} E_\alpha$. Podemos cambiar E_0 por $E_0 \cup U$, y así conseguimos que la unión de los E_α sea A. ∎

15.3 Un teorema de Silver

Vamos a aplicar los resultados sobre conjuntos estacionarios y cerrados no acotados para probar un importante resultado sobre la hipótesis de los cardinales singulares.

Diremos que un cardinal infinito κ cumple la HCG si $2^\kappa = \kappa^+$. Diremos que κ cumple la HCS si $2^{\operatorname{cf}\kappa} < \kappa \to \kappa^{\operatorname{cf}\kappa} = \kappa^+$.

Es claro que la HCG (resp. la HCS) equivale a que la HCG (la HCS) se cumpla sobre todos los cardinales.

Teorema 15.18 (Silver) *Se cumple:*

a) Si κ es un cardinal singular de cofinalidad no numerable y los cardinales (infinitos) menores que κ cumplen la HCG *entonces κ cumple la* HCG.

b) Si no se cumple la HCS, *entonces el mínimo cardinal que no la cumple tiene cofinalidad numerable.*

c) Si la HCS *se cumple sobre los cardinales de cofinalidad numerable, entonces se cumple sobre todos los cardinales.*

En adelante supondremos que $\aleph_0 < \mu = \operatorname{cf}\kappa < \kappa$ y que $\{\kappa_\alpha\}_{\alpha<\mu}$ es una sucesión normal de cardinales cofinal en κ.

Definición 15.19 Dos funciones f y g de dominio μ son *casi disjuntas* si el conjunto $\{\alpha < \mu \mid f(\alpha) = g(\alpha)\}$ está acotado en μ.

Una familia $\mathcal{F}$ de funciones de dominio μ es *casi disjunta* si está formada por funciones casi disjuntas dos a dos.

Teorema 15.20 *Si $\bigwedge\nu < \kappa\ \nu^\mu < \kappa$, $\mathcal{F} \subset \prod_{\alpha<\mu} A_\alpha$ es una familia casi disjunta de funciones y el conjunto $\{\alpha < \mu \mid |A_\alpha| \leq \kappa_\alpha\}$ es estacionario en μ, entonces $|\mathcal{F}| \leq \kappa$.*

Demostración: No perdemos generalidad si suponemos que los conjuntos A_α están formados por ordinales y que $\{\alpha < \mu \mid A_\alpha \subset \kappa_\alpha\}$ es estacionario en μ pues, biyectando cada A_α con su cardinal podemos construir otra $\mathcal{F}$ equipotente a la dada y en las mismas condiciones.

Sea $E_0 = \{\lambda < \mu \mid A_\lambda \subset \kappa_\lambda\}$, que es estacionario en μ, pues es la intersección del conjunto que estamos suponiendo que es estacionario con el conjunto de los ordinales límite $< \mu$, que es c.n.a.

Si $f \in \mathcal{F}$, entonces para todo $\lambda \in E_0$ tenemos que $f(\lambda) \in A_\lambda \subset \kappa_\lambda$ y como $\{\kappa_\alpha\}_{\alpha<\mu}$ es normal existe un ordinal $g(\lambda) < \lambda$ tal que $f(\lambda) \in \kappa_{g(\lambda)}$.

Como E_0 es estacionario y $g : E_0 \longrightarrow \mu$ es regresiva, el teorema 15.15 nos da un conjunto estacionario $E_f \subset E_0$ tal que g es constante en E_f: En particular f esta acotada en E_f por un $\kappa_\alpha < \kappa$.

La aplicación que a cada f le asigna $f|_{E_f}$ es inyectiva, pues si $f|_{E_f} = g|_{E_g}$ entonces $f = g$ por ser $\mathcal{F}$ casi disjunta (los conjuntos E_f y E_g son no acotados).

El número de funciones $h : E \longrightarrow \kappa_\alpha$ con $E \subset \mu$ fijo es a lo sumo (teniendo en cuenta la hipótesis)

$$\Big|\bigcup_{\alpha<\mu} \kappa_\alpha^E\Big| \leq \sum_{\alpha<\mu} \kappa_\alpha^\mu \leq \sum_{\alpha<\mu} \kappa = \kappa.$$

Como $|\mathcal{P}\mu| = 2^\mu < \kappa$, el número de funciones $h : E \longrightarrow \kappa_\alpha$ para cualquier E es a lo sumo $2^\mu \cdot \kappa = \kappa$.

Como hemos asociado a cada $f \in \mathcal{F}$ una función $h = f|_{E_f}$ distinta y a lo sumo puede haber κ funciones h, ha de ser $|\mathcal{F}| \leq \kappa$. ∎

En realidad vamos a necesitar una ligera variante de este teorema:

Teorema 15.21 *Si $\bigwedge \nu < \kappa\ \nu^\mu < \kappa$, $\mathcal{F} \subset \prod_{\alpha<\mu} A_\alpha$ es una familia casi disjunta de funciones y el conjunto $\{\alpha < \mu \mid |A_\alpha| \leq \kappa_\alpha^+\}$ es estacionario en μ, entonces $|\mathcal{F}| \leq \kappa^+$.*

Demostración: Como en el teorema anterior podemos suponer que los conjuntos A_α están formados por ordinales y que $E_0 = \{\alpha < \mu \mid A_\alpha \subset \kappa_\alpha^+\}$ es estacionario en μ.

Sea $f \in \mathcal{F}$ y $E \subset E_0$ estacionario. Definimos

$$\mathcal{F}_{f,E} = \{g \in \mathcal{F} \mid \bigwedge \alpha \in E\ g(\alpha) \leq f(\alpha)\}.$$

Claramente se trata de una familia casi disjunta contenida en $\prod_{\alpha<\mu} B_\alpha$, donde

$$B_\alpha = \begin{cases} f(\alpha)+1 & \text{si } \alpha \in E, \\ \kappa & \text{en caso contrario.} \end{cases}$$

Así, si $\alpha \in E \subset E_0$, tenemos que $f(\alpha) \in \kappa_\alpha^+$, luego $|B_\alpha| = |f(\alpha)+1| \leq \kappa$. Por consiguiente el conjunto $\{\alpha < \mu \mid |B_\alpha| \leq \kappa_\alpha\}$ es estacionario (contiene a E) y podemos aplicar el teorema anterior, según el cual $|\mathcal{F}_{f,E}| \leq \kappa$.

Ahora definimos

$$\mathcal{F}_f = \{g \in \mathcal{F} \mid \bigvee E \subset E_0(E \text{ estacionario} \wedge \bigwedge \alpha \in E\, g(\alpha) \leq f(\alpha))\} = \bigcup_E \mathcal{F}_{f,E},$$

donde E varía en los subconjuntos estacionarios de E_0. Claramente

$$|\mathcal{F}_f| \leq \sum_E \kappa \leq 2^\mu \kappa = \kappa.$$

Veamos finalmente que $|\mathcal{F}| \leq \kappa^+$. En otro caso tomemos $\{f_\alpha\}_{\alpha<\kappa^+}$ funciones distintas en $\mathcal{F}$. Tenemos que $\Big|\bigcup_{\alpha<\kappa^+} \mathcal{F}_\alpha\Big| \leq \sum_{\alpha<\kappa^+} \kappa = \kappa^+$, luego ha de existir una función $f \in \mathcal{F} \setminus \bigcup_{\alpha<\kappa^+} \mathcal{F}_\alpha$.

En tal caso el conjunto $\{\gamma \in E_0 \mid f(\gamma) \leq f_\alpha(\gamma)\}$ no es estacionario para ningún $\alpha < \kappa^+$, luego su complementario $\{\gamma \in E_0 \mid f_\alpha(\gamma) \leq f(\gamma)\}$ sí lo es, y esto significa que cada $f_\alpha \in \mathcal{F}_f$, lo cual es imposible, dado que hay κ^+ funciones f_α y $|\mathcal{F}_f| \leq \kappa$. ∎

El apartado a) del teorema de Silver es un caso particular del teorema siguiente:

Teorema 15.22 *Si el conjunto $\{\alpha < \mu \mid 2^{\kappa_\alpha} = \kappa_\alpha^+\}$ es estacionario en μ, entonces $2^\kappa = \kappa^+$.*

DEMOSTRACIÓN: Veamos que $\bigwedge\nu < \kappa\ \nu^\mu < \kappa$. En efecto, si $\nu < \kappa$ sea α tal que $\nu,\ \mu < \kappa_\alpha$ y $2^{\kappa_\alpha} = \kappa_\alpha^+$. Entonces $\nu^\mu \le \kappa_\alpha^{\kappa_\alpha} = 2^{\kappa_\alpha} = \kappa_\alpha^+ \le \kappa_{\alpha+1} < \kappa$.

Para cada $X \subset \kappa$ sea $f_X = \{X_\alpha\}_{\alpha<\mu}$, donde $X_\alpha = X \cap \kappa_\alpha$. Definimos $\mathcal{F} = \{f_X \mid X \subset \kappa\}$. Si $X \ne Y$ entonces f_X y f_Y son casi disjuntas, pues ha de existir un α tal que $X \cap \kappa_\alpha \ne Y \cap \kappa_\alpha$ y entonces $\{\delta < \mu \mid f_X(\delta) = f_Y(\delta)\} \subset \alpha$. En particular, si $X \ne Y$ entonces $f_X \ne f_Y$, luego $|\mathcal{F}| = 2^\kappa$.

Por otra parte $\mathcal{F}$ es una familia casi disjunta de funciones contenida en $\prod_{\alpha<\mu} \mathcal{P}\kappa_\alpha$ y el conjunto $\{\alpha < \mu \mid |\mathcal{P}\kappa_\alpha| = \kappa_\alpha^+\}$ es estacionario en μ. El teorema anterior nos da, entonces, que $2^\kappa = |\mathcal{F}| \le \kappa^+$. ∎

Para probar el resto del teorema de Silver necesitamos un paso más:

Teorema 15.23 *Si $\bigwedge\nu < \kappa\ \nu^\mu < \kappa$ y el conjunto $\{\alpha < \mu \mid \kappa_\alpha^{\operatorname{cf}\kappa_\alpha} = \kappa_\alpha^+\}$ es estacionario en μ, entonces $\kappa^\mu = \kappa^+$.*

DEMOSTRACIÓN: Para cada $h : \mu \longrightarrow \kappa$ sea $f_h = \{h_\alpha\}_{\alpha<\mu}$, donde las aplicaciones $h_\alpha : \mu \longrightarrow \kappa$ vienen dadas por

$$h_\alpha(\beta) = \begin{cases} h(\beta) & \text{si } h(\beta) < \kappa_\alpha, \\ 0 & \text{en otro caso.} \end{cases}$$

Sea $\mathcal{F} = \{f_h \mid h \in {}^\mu\kappa\}$. Si $h \ne g$, entonces f_g y f_h son casi disjuntas, pues si $h(\delta) \ne g(\delta)$ y ambos son menores que κ_α, entonces

$$\{\delta < \mu \mid f_h(\delta) = f_g(\delta)\} \subset \alpha + 1.$$

En particular si $h \ne g$ se cumple $f_h \ne f_g$, luego $|\mathcal{F}| = \kappa^\mu$. Además $\mathcal{F}$ es casi disjunta y está contenida en $\prod_{\alpha<\mu} {}^\mu\kappa_\alpha$.

Queremos aplicar el teorema 15.21 para concluir que $\kappa^\mu = |\mathcal{F}| \le \kappa^+$. Necesitamos, pues, probar que el conjunto $E = \{\alpha < \mu \mid \kappa_\alpha^\mu = \kappa_\alpha^+\}$ es estacionario en μ. Para ello consideramos el conjunto

$$C = \{\lambda < \mu \mid \textstyle\bigwedge\nu < \kappa_\lambda\ \nu^\mu < \kappa_\lambda\}.$$

Veamos que si $\lambda \in C$ entonces $\kappa_\lambda^{\operatorname{cf}\kappa_\lambda} = \kappa_\lambda^\mu$. De aquí se seguirá que

$$\{\alpha < \mu \mid \kappa_\alpha^{\operatorname{cf}\kappa_\alpha} = \kappa_\alpha^+\} \cap C \subset E$$

y, como el conjunto de la izquierda es estacionario por hipótesis, si probamos también que C es c.n.a., concluiremos que E es estacionario, tal y como nos hace falta.

Sea, pues, $\lambda \in C$. Entonces $\operatorname{cf}\kappa_\lambda = \operatorname{cf}\lambda \le \lambda < \mu$. Sea $\kappa_\lambda = \sum_{\alpha<\operatorname{cf}\kappa_\lambda} \nu_\alpha$, donde $\bigwedge\alpha < \operatorname{cf}\kappa_\lambda\ \nu_\alpha < \kappa_\lambda$. Así

$$\kappa_\lambda^{\operatorname{cf}\kappa_\lambda} \le \kappa_\lambda^\mu = \Big(\sum_{\alpha<\operatorname{cf}\kappa_\lambda} \nu_\alpha\Big)^\mu \le \prod_{\alpha<\operatorname{cf}\kappa_\lambda} \nu_\alpha^\mu \le \prod_{\alpha<\operatorname{cf}\kappa_\lambda} \kappa_\lambda = \kappa_\lambda^{\operatorname{cf}\kappa_\lambda}.$$

Según lo dicho, ahora sólo queda probar que C es c.n.a. en μ. Para ello definimos $l : \mu \longrightarrow \mu$ mediante

$$l(\alpha) = \text{mín}\{\beta < \mu \mid \kappa_\alpha^\mu < \kappa_\beta\}.$$

Basta probar que

$$C = \{\lambda \mid \lambda < \mu\} \cap \{\alpha < \mu \mid l[\alpha] \subset \alpha\}.$$

En efecto, si $\lambda \in C$ y $\alpha < \lambda$, entonces $\kappa_\alpha^\mu < \kappa_\lambda$, existe un $\beta < \lambda$ tal que $\kappa_\alpha^\mu < \kappa_\beta$, luego $l(\alpha) \leq \beta < \lambda$. Por lo tanto $l[\lambda] \subset \lambda$.

Recíprocamente, si $l[\lambda] \subset \lambda$ y $\nu < \kappa_\lambda$, sea $\alpha < \lambda$ tal que $\nu < \kappa_\alpha$. Entonces $\nu^\mu \leq \kappa_\alpha^\mu < \kappa_{l(\alpha)} < \kappa_\lambda$, luego $\lambda \in C$. ∎

Ahora estamos en condiciones de probar el apartado b) del teorema de Silver, y el apartado c) es una consecuencia inmediata. Sea κ el mínimo cardinal que incumple la HCS, es decir, $\kappa > \aleph_0$, $2^{\text{cf}\,\kappa} < \kappa$, pero $\kappa^{\text{cf}\,\kappa} > \kappa^+$. Supongamos que $\text{cf}\,\kappa > \aleph_0$.

Sea $\mu = \text{cf}\,\kappa$ y $\{\kappa_\alpha\}_{\alpha<\mu}$ como en los teoremas precedentes. Tenemos que la HCS se cumple bajo κ, luego el argumento del teorema 14.18 es válido en este contexto y nos permite probar que si $\nu < \kappa$ entonces ν^μ toma uno de los valores 2^μ, μ o μ^+, luego en particular $\bigwedge \nu < \kappa\ \nu^\mu < \kappa$.

Sea $E = \{\alpha < \mu \mid \text{cf}\,\kappa_\alpha = \aleph_0 \wedge 2^{\aleph_0} < \kappa_\alpha\}$. Es claro que E es estacionario en μ, pues contiene a la intersección del c.n.a. $\mu \setminus \alpha_0$, donde α_0 es el mínimo ordinal tal que $2^{\aleph_0} < \kappa_{\alpha_0}$, con el conjunto $\{\lambda < \mu \mid \text{cf}\,\lambda\ (= \text{cf}\,\kappa_\lambda) = \aleph_0\}$, el cual es estacionario por el teorema 15.13.

Si $\alpha \in E$, entonces $2^{\text{cf}\,\kappa_\alpha} < \kappa_\alpha$, con $\text{cf}\,\kappa_\alpha = \aleph_0$ y, como $\kappa_\alpha < \kappa$ cumple la HCS, $\kappa_\alpha^{\text{cf}\,\kappa_\alpha} = \kappa_\alpha^+$, de modo que $E \subset \{\alpha < \mu \mid \kappa_\alpha^{\text{cf}\,\kappa_\alpha} = \kappa_\alpha^+\}$. Concluimos que este último conjunto es estacionario y ello nos permite aplicar el teorema anterior, según el cual $\kappa^{\text{cf}\,\kappa} = \kappa^+$. ∎

Tenemos así un ejemplo no trivial de las numerosas restricciones que se conocen sobre la función del continuo en cardinales singulares. Por ejemplo, si suponemos que $\bigwedge \alpha < \omega_1\ 2^{\aleph_\alpha} = \aleph_{\alpha+1}$, entonces necesariamente $2^{\aleph_{\omega_1}} = \aleph_{\omega_1+1}$. En cambio, aunque supongamos

$$\bigwedge n \in \omega\ 2^{\aleph_n} = \aleph_{n+1}$$

no podemos demostrar —aunque no es fácil probar que así es— que $2^{\aleph_\omega} = \aleph_{\omega+1}$, es decir, la HCS no puede demostrarse ni siquiera para $\aleph_\omega$. Esto no significa que $2^{\aleph_\omega}$ esté libre de este tipo de restricciones. Por ejemplo, un profundo teorema de S. Shelah de 1982 afirma que, para todo ordinal límite λ:

$$\aleph_\lambda^{\text{cf}\,\lambda} < \aleph_{(|\lambda|^{\text{cf}\,\lambda})^+}.$$

En particular, si $\bigwedge n \in \omega\ 2^{\aleph_n} < \aleph_\omega$, entonces $2^{\aleph_\omega} = \aleph_\omega^{\aleph_0} < \aleph_{(2^{\aleph_0})^+}$.

Más sorprendente aún es otro teorema de Shelah de 1990, según el cual, si $2^{\aleph_0} < \aleph_\omega$ entonces $\aleph_\omega^{\aleph_0} < \aleph_{\omega_4}$. Estos resultados son algunas consecuencias de la llamada teoría de las cofinalidades posibles, descubierta por Shelah y que tiene muchas más consecuencias en muchas ramas de la teoría de conjuntos.

15.4 Cardinales de Mahlo

Los cardinales inaccesibles son los más pequeños de los llamados "cardinales grandes". Los siguientes en la lista son los cardinales de Mahlo, que ahora vamos a introducir.

Definición 15.24 Un cardinal κ es *(fuertemente) de Mahlo* si κ es (fuertemente) inaccesible y el conjunto $\{\mu < \kappa \mid \mu \text{ es regular}\}$ es estacionario en κ.

En realidad los cardinales de Mahlo cumplen mucho más de lo que exige la definición:

Teorema 15.25 *Si κ es un cardinal (fuertemente) de Mahlo, entonces el conjunto $\{\mu < \kappa \mid \mu$ es (fuertemente) inaccesible$\}$ es estacionario en κ.*

Demostración: Basta ver que el conjunto

$$C = \{\mu < \kappa \mid \mu \text{ es un cardinal límite (fuerte)}\}$$

es c.n.a. en κ, pues el conjunto del enunciado es la intersección con C del conjunto de la definición de cardinal de Mahlo.

El conjunto C es cerrado porque el supremo de un conjunto no acotado de cardinales es un cardinal límite, y si los cardinales son límites fuertes el supremo también lo es.

Si $\alpha < \kappa$, sea $\mu_0 = \alpha^+$ y definimos $\bigwedge n \in \omega\ \mu_{n+1} = \mu_n^+$ (respectivamente $\bigwedge n \in \omega\ \mu_{n+1} = 2^{\mu_n}$). Como κ es un cardinal límite (fuerte), se cumple que $\bigwedge n \in \omega\ \mu_n \in \kappa$ y, como κ es regular, $\mu = \sup\limits_{n\in\omega} \mu_n \in \kappa$. Claramente μ es un cardinal límite (fuerte), de modo que $\mu \in C \wedge \alpha < \mu$. Así pues, C no está acotado en κ. ∎

De este modo, si κ es un cardinal fuertemente de Mahlo, entonces V_κ es un modelo de ZFC en el que existe una clase propia de cardinales fuertemente inaccesibles (porque los ordinales de V_κ son los ordinales $< \kappa$ y por encima de cualquiera de ellos hay un cardinal inaccesible. Más aún, éstos están "bien distribuidos", en el sentido de que forman una clase estacionaria. Más aún, si $f : \kappa \longrightarrow \kappa$ es la semejanza entre κ y el conjunto de cardinales fuertemente inaccesibles bajo κ y $\mu = f(\omega_{17} + 1)$, entonces μ es un cardinal fuertemente inaccesible que tiene exactamente $\aleph_{17}$ cardinales fuertemente inaccesibles bajo sí, luego V_μ es un modelo de ZFC + existen exactamente $\aleph_{17}$ cardinales fuertemente inaccesibles.

Así pues, la consistencia de que exista un cardinal fuertemente de Mahlo implica la consistencia de que exista cualquier cantidad de cardinales fuertemente inaccesibles, luego por el teorema de incompletitud la consistencia de que exista un cardinal de Mahlo no puede probarse ni siquiera suponiendo consistente que exista cualquier cantidad de cardinales inaccesibles. Al igual que ocurre con los cardinales inaccesibles, todo esto sigue siendo cierto si hablamos de cardinales débilmente de Mahlo y cardinales débilmente inaccesibles. De hecho, la

consistencia de que existan cardinales débilmente de Mahlo equivale a la consistencia de que existan cardinales fuertemente de Mahlo, pero esto no estamos en condiciones de justificarlo aquí.

Se suele decir que un cardinal de Mahlo es "mas grande" que un cardinal inaccesible, pero esto no ha de ser entendido en sentido literal: pueden existir cardinales $\kappa < \mu$ de modo que κ sea de Mahlo y μ sea inaccesible. La comparación debe entenderse en dos sentidos: por una parte, el mínimo cardinal de Mahlo κ (si existe) ha de ser mucho mayor que el mínimo cardinal fuertemente inaccesible, pues κ ha de dejar bajo sí un conjunto estacionario de cardinales fuertemente inaccesibles; por otra parte, también se dice que un cardinal de Mahlo es más grande en el sentido de que implica la consistencia de que existan muchos cardinales inaccesibles, es decir, en el sentido de que suponer la existencia de un cardinal de Mahlo es "más fuerte" que suponer la consistencia de un cardinal fuertemente inaccesible.

A partir de la existencia de un cardinal fuertemente de Mahlo no puede probarse la existencia de dos de ellos, pues si $\mu < \kappa$ son dos cardinales fuertemente de Mahlo, entonces es fácil ver que V_κ es un modelo de ZFC donde sólo existe un cardinal de Mahlo. Por consiguiente, postular la existencia de dos cardinales de Mahlo es más fuerte que postular la existencia de uno solo, etc.

Todavía se puede ir más allá:

Definición 15.26 Sea γ un ordinal infinito. Definimos los conjuntos

$$\begin{aligned} M_0(\gamma) &= \{\kappa < \gamma \mid \kappa \text{ es (fuertemente) inaccesible}\},\\ M_{\alpha+1}(\gamma) &= \{\kappa \in M_\alpha(\gamma) \mid \{\mu < \kappa \mid \mu \in M_\alpha(\gamma)\} \text{ es estacionario en } \kappa\},\\ M_\lambda(\gamma) &= \bigcap_{\delta<\lambda} M_\delta(\gamma). \end{aligned}$$

Definimos las clases

$$M_\alpha = \bigcup_{\gamma\in\Omega} M_\alpha(\gamma).$$

A los elementos de M_α los llamaremos cardinales (fuertemente) α-Mahlo. El ordinal γ que aparece en la definición es un auxiliar técnico para evitar una recurrencia con clases propias que no estaría justificada, pero se comprueba inmediatamente lo siguiente:

Para todo cardinal infinito κ:

κ es (fuertemente) 0-Mahlo si y sólo si es (fuertemente) inaccesible.

κ es (fuertemente) α+1-Mahlo si y sólo si es (fuertemente) α-Mahlo y el conjunto $\{\mu < \kappa \mid \mu$ es (fuertemente) α-Mahlo$\}$ es estacionario en κ.

κ es (fuertemente) λ-Mahlo si y sólo si es (fuertemente) δ-Mahlo para todo $\delta < \lambda$.

De este modo, los cardinales (fuertemente) de Mahlo son precisamente los (fuertemente) 1-Mahlo. Es fácil ver que la situación en cuanto a consistencia de los cardinales 2-Mahlo respecto a los 1-Mahlo es la misma que la de los 1-Mahlo respecto a los inaccesibles, con lo que tenemos una escala de cardinales grandes.

Notemos que si κ es un cardinal (fuertemente) α-mahlo y para cada $\beta < \alpha$ llamamos μ_β al menor cardinal (fuertemente) β-Mahlo, entonces la aplicación $f : \alpha \longrightarrow \kappa$ dada por $f(\beta) = \mu_\beta$ es inyectiva y creciente, luego $\alpha \leq \kappa$. Así pues, un cardinal κ puede a lo sumo ser (fuertemente) κ-Mahlo, pero nunca $\kappa + 1$-Mahlo.

Es posible definir cardinales mucho mayores —en el sentido de consistencia relativa— que los cardinales κ-Mahlo. Por nombrar los más importantes citaremos —en orden de magnitud— los cardinales débilmente compactos, los cardinales medibles, los fuertemente compactos, los supercompactos y los cardinales enormes, aunque hay muchos más.

El interés principal de los cardinales grandes es que son necesarios en muchas pruebas de consistencia. Por ejemplo, la negación de la hipótesis cardinales singulares implica la consistencia de que existan infinitos cardinales débilmente compactos, y cada cardinal débilmente compacto κ es fuertemente κ-Mahlo. Esto significa que no podemos demostrar la consistencia de $\neg$HCS si no suponemos —como mínimo— la consistencia de que existan cardinales débilmente compactos.

Apéndice A

Conceptos elementales de la teoría de conjuntos

En este apéndice recogemos por completitud los conceptos que conforman el vocabulario básico en torno a los conjuntos (aplicaciones, relaciones de equivalencia, relaciones de orden, etc.) Todo cuanto sigue puede interpretarse de tres formas distintas:

- Como definiciones y razonamientos metamatemáticos aplicables a colecciones de objetos cualesquiera, con tal de que estén bien definidas en el sentido discutido, por ejemplo, en la primera sección del capítulo III.
- Como definiciones y teoremas de la teoría de conjuntos (básica) de von Neumann-Bernays-Gödel (NBG*), en cuyo caso hemos de sustituir la noción metamatemática de colección de objetos por el concepto técnico de "clase".
- Como definiciones y teoremas de la teoría de conjuntos (básica) de Zermelo-Fraenkel (ZF*), en cuyo caso la noción básica no definida es la de "conjunto".[1]

Ante la necesidad de tomar una opción, hemos formulado los resultados en términos de clases y conjuntos, es decir, de acuerdo con la teoría axiomática NBG*, de modo que la "traducción" a los otros dos casos se reduce a sustituir "clase" y "conjunto" por "colección" (bien definida) o por "conjunto" en cada ocasión.

No entramos en problemas de existencia de las *colecciones/clases/conjuntos* de los que hablamos, pues estas cuestiones se resuelven de forma distinta en cada contexto: a nivel metamatemático son inmediatas, mientras que en las dos teorías de conjuntos están cuidadosamente discutidas en el capítulo VIII.

[1] Aunque también tiene sentido aplicar todo cuanto sigue a clases arbitrarias en el sentido discutido en la pág. 239.

Desde el punto de vista metamatemático, una expresión de la forma $u \in X$ significará que u es un elemento de la colección X, mientras que desde el punto de vista de las teorías axiomáticas $\in$ es un relator diádico del lenguaje formal, por lo que no tiene definición. La expresión $\{x \mid \phi(x)\}$ representará a la *colección/clase/conjunto* de los *objetos/conjuntos* x que cumplan la *propiedad/fórmula* $\phi(x)$, donde $\phi(x)$ puede *hacer referencia a otros objetos / tener otras variables libres* aparte de x. A menudo emplearemos variantes como

$$\{x \in X \mid \phi(x)\} \equiv \{x \mid x \in X \wedge \phi(x)\}.$$

Representaremos por $\{x, y\}$ a la *colección/clase/conjunto* formada exactamente por los *objetos/conjuntos* x, y. El *par ordenado* de componentes x, y (en este orden) puede ser definido como $(x, y) = \{\{x\}, \{x, y\}\}$, aunque metamatemáticamente esto es superfluo. No haremos referencia a conceptos carentes de contenido metamatemático, como pueda ser la clase universal (la clase de todos los conjuntos).

El álgebra de clases Dadas dos clases A y B, se define su *unión* y su *intersección* respectivamente como

$$A \cup B \equiv \{x \mid x \in A \vee x \in B\}, \qquad A \cap B \equiv \{x \mid x \in A \wedge x \in B\}.$$

Su *diferencia* es la clase $A \setminus B \equiv \{x \mid x \in A \wedge x \notin B\}$. La clase vacía se define como $\varnothing \equiv \{x \mid x \neq x\}$. Se dice que una clase A *es una subclase de* B o que *está incluida en* una clase B si cumple

$$A \subset B \equiv \bigwedge x(x \in A \to x \in B).$$

Aplicaciones Se define el *producto cartesiano* de dos clases A y B como la clase

$$A \times B \equiv \{(a, b) \mid a \in A \wedge b \in B\}.$$

Una clase F es una *función* si sus elementos son todos pares ordenados y un mismo conjunto x no aparece como primera componente de dos pares distintos en F, es decir,

$$F \text{ es una función} \equiv \bigwedge x \in F \bigvee uv\, x = (u, v)$$

$$\wedge \bigwedge uvw((u, w) \in F \wedge (v, w) \in F \to u = v).$$

Se define el *dominio* (rango) de una clase A como la clase de las primeras (segundas) componentes de los pares ordenados que pertenezcan a A, es decir,

$$\mathcal{D}A \equiv \{x \mid \bigvee y\,(x, y) \in A\}, \qquad \mathcal{R}A \equiv \{x \mid \bigvee y\,(y, x) \in A\}.$$

De este modo, si F es una función y $x \in \mathcal{D}F$, existe un único conjunto y tal que $(x, y) \in F$. Lo representaremos por $y = F(x)$ y lo llamaremos *imagen* de x por F. Formalmente, definimos

$$F(x) \equiv y \mid (x, y) \in F,$$

teniendo en cuenta que $F(x)$ sólo es una descripción propia cuando F es una función y $x \in \mathcal{D}F$.

Una clase F es una *aplicación* de una clase A en una clase B si cumple

$$F : A \longrightarrow B \equiv F \text{ es una función} \wedge \mathcal{D}F = A \wedge \mathcal{R}F \subset B.$$

De este modo, una aplicación $F : A \longrightarrow B$ asigna a cada $x \in A$ una única imagen $F(x) \in B$.

Una aplicación $F : A \longrightarrow B$ es *inyectiva* si $\bigwedge xy \in A(F(x) = F(y) \to x = y)$, es decir, si elementos distintos en A tienen imágenes distintas en B.

Cuando F es una función y $F(x) = y$, se dice también que x es una *antiimagen* de y por F. Si $F : A \longrightarrow B$, cada elemento de b puede tener varias antiimágenes en A o no tener ninguna. Se dice que F es *suprayectiva* si $\mathcal{R}F = B$, es decir, si cada elemento de B tiene al menos una antiimagen en A.

Una aplicación $F : A \longrightarrow B$ es *biyectiva* si es a la vez inyectiva y suprayectiva, es decir, si cada elemento de A se corresponde con un único elemento de B y viceversa.

Notas A veces se define la *gráfica* de una aplicación $F : A \longrightarrow B$ como la clase $\{(x, F(x)) \mid x \in A\}$, pero conviene tener presente que, de acuerdo con las definiciones que hemos dado, la gráfica de F coincide con F.

Así mismo conviene observar que si $F : A \longrightarrow B$ y $B \subset C$, entonces también $F : A \longrightarrow C$, por lo que la noción de suprayectividad no depende únicamente de F, sino de F y de B. ∎

Si $F : A \longrightarrow B$ y $C \subset A$, se define $F[C] = \{F(x) \mid x \in C\} \subset B$. Formalmente,

$$F[C] \equiv \{b \in B \mid \bigvee c \in C\, b = F(c)\}.$$

Similarmente, si $D \subset B$ se define $F^{-1}[D] = \{x \in A \mid F(x) \in D\}$.

Es fácil probar que

$$F^{-1}[D_1 \cup D_2] = F^{-1}[D_1] \cup F^{-1}[D_2], \qquad F^{-1}[D_1 \cap D_2] = F^{-1}[D_1] \cap F^{-1}[D_2],$$

Así mismo $F[C_1 \cup C_2] = F[C_1] \cup F[C_2]$, pero $F[C_1 \cap C_2] \subset F[C_1] \cap F[C_2]$ y en general no se da la igualdad.

En general, para una clase arbitraria A definimos $A^{-1} \equiv \{(x, y) \mid (y, x) \in A\}$. De este modo, si $F : A \longrightarrow B$ biyectiva, se cumple que $F^{-1} : B \longrightarrow A$ biyectiva, y se dice que F^{-1} es la *aplicación inversa* de F.

Notemos que en este contexto tenemos dos definiciones distintas de $F^{-1}[D]$, pero ambas son equivalentes.

Dada una clase A, la aplicación $I_A : A \longrightarrow A$ dada por $\bigwedge x \in A\, I_A(x) = x$, se llama *identidad* en A. Si $A \subset B$, la identidad en A considerada como aplicación $A \longrightarrow B$ recibe el nombre de *inclusión* de A en B.

Si $F : A \longrightarrow B$ y $C \subset A$, se define la *restricción* de F a B como la clase $F|_C = F \cap C \times B$, de modo que $F|_C : C \longrightarrow B$ y $\bigwedge c \in C\, F|_C(c) = F(c)$.

En general, dadas dos clases A y B, definimos su *composición* como la clase

$$A \circ B \equiv \{(x, y) \mid \bigvee z((x, z) \in A \wedge (z, y) \in B)\}.$$

Es fácil ver que $(A \circ B) \circ C = A \circ (B \circ C)$. Si $F : A \longrightarrow B$ y $G : B \longrightarrow C$, entonces $F \circ G : A \longrightarrow C$ y se cumple[2] que $\bigwedge x \in A\, (F \circ G)(x) = G(F(x))$.

Relaciones Una clase R es una *relación* en una clase A si $R \subset A \times A$. En tal caso, en lugar de escribir $(x, y) \in R$, se escribe $x\, R\, y$ y se lee "x está relacionado con y". En estas condiciones:

a) R es *reflexiva* si $\bigwedge x \in A\, x\, R\, x$.

b) R es *irreflexiva* si $\bigwedge x \in A \neg x\, R\, x$.

c) R es *simétrica* si $\bigwedge xy \in A(x\, R\, y \to y\, R\, x)$.

d) R es *antisimétrica* si $\bigwedge xy \in A(x\, R\, y \wedge y\, R\, x \to x = y)$.

e) R es *asimétrica* si $\bigwedge xy \in A(x\, R\, y \to \neg y\, R\, x)$.

f) R es *transitiva* si $\bigwedge xyz \in A(x\, R\, y \wedge y\, R\, z \to x\, R\, z)$.

g) R es *conexa* si $\bigwedge xy \in A(x\, R\, y \vee y\, R\, x)$.

Relaciones de equivalencia Una *relación de equivalencia* en una clase A es una relación reflexiva, simétrica y transitiva en A.

Si R es una relación de equivalencia en A y $a \in A$, se define la *clase de equivalencia* de a respecto de R como la clase $[a]_R \equiv \{x \in A \mid x\, R\, a\}$. Si no hay confusión suprimiremos el subíndice R.

De la reflexividad se sigue que $a \in [a]$, por lo que $[a] \neq \varnothing$. Así mismo es fácil probar que

$$\bigwedge xy \in A(x\, R\, y \leftrightarrow [x] = [y]),$$
$$\bigwedge xy \in A(\neg x\, R\, y \leftrightarrow [x] \cap [y] = \varnothing).$$

En el resto de este apartado (si trabajamos en NBG*) hemos de suponer que A es un conjunto, de modo que por el teorema 8.8 las clases de equivalencia son también conjuntos.

Definimos el *conjunto cociente* del conjunto A respecto a la relación de equivalencia R en A como el conjunto A/R de todas las clases de equivalencia de R, es decir,

$$A/R \equiv \{[x]_R \mid x \in A\}.$$

En general, un conjunto C es una *partición* de un conjunto A si cumple

[2]Es frecuente definir $F \circ G$ de modo que $(F \circ G)(x) = F(G(x))$, pero, cuando se trabaja con muchas aplicaciones, es mucho más fácil invertir el orden cuando se deshace una composición que cuando se ha de plasmar por escrito una composición cuyo esquema está claro mentalmente.

a) $\bigwedge x \in C(x \subset A \wedge x \neq \varnothing)$,

b) $\bigwedge a \in A \bigvee x \in C\, a \in x$,

c) $\bigwedge xy \in C(x = y \vee x \cap y = \varnothing)$.

Tenemos, pues, que si R es una relación de equivalencia en un conjunto A, entonces el conjunto cociente A/R es una partición de A.

Ejercicio: Probar que si C es una partición de un conjunto A, entonces existe una relación de equivalencia R en A tal que $C = A/R$.

Relaciones de orden Una *relación de orden* en una clase A es una relación reflexiva, antisimétrica y transitiva en A. A veces se dice también que R es una *relación de orden parcial,* mientras que una *relación de orden total* es una relación de orden conexa.

Es habitual representar las relaciones de orden mediante el signo $\leq$, entendiendo que éste hace referencia a relaciones distintas según el contexto.

Una *relación de orden estricto* en una clase A es una relación irreflexiva, asimétrica y transitiva. Es claro que si $\leq$ es una relación de orden no estricto en una clase A, entonces la relación dada por $x < y \equiv (x \leq y \wedge x \neq y)$ es una relación de orden estricto en A y, recíprocamente, si $<$ es una relación de orden estricto en A, entonces la relación dada por $x \leq y \equiv (x < y \vee x = y)$ es una relación de orden no estricto en A, por lo que ambos conceptos son equivalentes.

Sea A una clase ordenada por la relación $\leq$ y sea $B \subset A$. Entonces:

a) $M \in A$ es una *cota superior* de B si $\bigwedge x \in B\, x \leq M$,

b) $m \in A$ es una *cota inferior* de B si $\bigwedge x \in B\, m \leq x$,

c) $M \in A$ es un *maximal* de B si $M \in B$ y $\bigwedge x \in B(M \leq x \to M = x)$.

d) $m \in A$ es un *minimal* de B si $m \in B$ y $\bigwedge x \in B(x \leq m \to x = m)$.

e) $M \in A$ es el *supremo* de B si M es una cota superior de B y $\bigwedge x \in A(x$ es una cota superior de $B \to M \leq x)$.

f) $m \in A$ es el *ínfimo* de B si m es una cota inferior de B y $\bigwedge x \in A(x$ es una cota inferior de $B \to x \leq m)$.

g) $M \in A$ es el *máximo* de B si $M \in B$ y M es una cota superior de B.

h) $m \in A$ es el *mínimo* de B si $m \in B$ y m es una cota inferior de B.

Es fácil ver que en un conjunto totalmente ordenado todo maximal es máximo y todo minimal es mínimo. Si un conjunto tiene máximo o mínimo, supremo o ínfimo, entonces éstos son únicos. El supremo (ínfimo) de una clase es máximo (mínimo) si y sólo si pertenece a la clase.

Un *buen orden* en una clase A es una relación de orden $\leq$ en A respecto a la cual toda subclase de A tiene un mínimo elemento (también se dice que A está *bien ordenada* por $\leq$). Todo buen orden es un orden total, pues si x, $y \in A$, tendremos $x \leq y$ o $y \leq x$ según quién sea el mínimo de $\{x, y\}$.

Cuando tenemos una clase A ordenada por una relación $\leq$ y una subclase $B \subset A$, consideramos, aunque no se indique explícitamente, que B está ordenada por la restricción de $\leq$ a B, es decir, con la intersección de $\leq$ con $B \times B$, de modo que si x, $y \in B$, se cumple $x \leq y$ como elementos de B si y sólo si se cumple como elementos de A.

Es inmediato comprobar que esta restricción es un orden en B. Más aún, B está totalmente ordenada o bien ordenada si A lo está.

Una aplicación $F : A \longrightarrow B$ entre dos clases ordenadas por respectivas relaciones de orden $\leq_1$ y $\leq_2$ es *monótona creciente* o, simplemente, *creciente* (respecto a dichas relaciones), si

$$\bigwedge xy \in A(x \leq_1 y \to F(x) \leq_2 F(y)).$$

Se dice que F es *monótona decreciente* o *decreciente* si cumple

$$\bigwedge xy \in A(x \leq_1 y \to F(y) \leq_2 F(x)).$$

Se dice que F es *estrictamente* monótona creciente o decreciente si se cumple esto mismo cambiando las desigualdades no estrictas $\leq$ por desigualdades estrictas $<$.

Apéndice B

Complementos sobre aritmética

Recogemos aquí los hechos aritméticos básicos que hemos usado en el libro, especialmente en la primera parte. El lector debe convencerse de que todas las afirmaciones que haremos aquí son verdaderas en sentido metamatemático. Después de leer el capítulo VIII no debería tener ninguna dificultad en convencerse de que todas ellas se corresponden además con teoremas de la teoría de conjuntos.

B.1 Hechos elementales

El lector puede saltarse sin duda alguna esta sección. De todos modos, la incluimos para hacer hincapié en que la aritmética elemental puede exponerse informalmente (es decir, sin basarse en una teoría axiomática) y al mismo tiempo con toda precisión y exactitud.[1]

Suponemos conocidos los números naturales, así como su ordenación, es decir, suponemos que el lector es capaz de continuar indefinidamente la sucesión

$$0,\ 1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8,\ 9,\ 10,\ 11,\ 12,\ 13,\ 14,\ \ldots$$

y que sabe responder a preguntas tales como ¿quién aparece antes en la sucesión, 15 o 27? El 15 aparece antes que el 27 y eso lo expresamos así: $15 < 27$. La expresión $m \leq n$ significará que m es menor o igual que n.

También suponemos que el lector sabe contar, en el sentido de que sabe responder a preguntas tales como ¿cuántos puntos hay aquí?

$$\circ\quad \circ\quad \circ\quad \circ\quad \circ\quad \circ$$

[1] Si el lector detecta algunos de los numerosos saltos lógicos que contienen las líneas siguientes, debe pensar que obedecen únicamente al deseo de no aburrirle con disquisiciones y matizaciones sobre hechos que conoce sobradamente.

No todas las colecciones de cosas se pueden contar. Las colecciones que no se pueden contar en el sentido usual de la palabra (porque no se acaban nunca) se llaman infinitas.

Suma También es seguro que el lector sabe sumar. Decir que $7 + 5 = 12$ significa que si a 7 cosas les añadimos 5 cosas más, tenemos 12 cosas. Está claro que si hemos de sumar varios números no importará el orden en que lo hagamos: $7 + 3 + 8 = 18$ y da igual si hacemos $7 + 3 = 10$ y $10 + 8 = 18$, que si sumamos primero $7 + 8 = 15$ y $15 + 3 = 18$, etc. En cualquier caso estamos juntando 7 cosas y 3 cosas y 8 cosas y contando cuántas tenemos. Sumar 0 es no añadir nada, luego $n + 0 = n$ siempre. Naturalmente $m \leq m + n$ y si un número $n \neq 0$ entonces $m < m + n$.

Si $m \leq n$ y tenemos n cosas podemos deshacernos de m de ellas. El número de cosas que nos quedan lo llamamos $n - m$. Podemos alterar el orden de las sumas y las restas siempre que no nos veamos en situación de quitarle a un número otro mayor.

Si $m \leq n$ resulta que $m + (n - m) = n$. Esto nos relaciona como sigue la suma y la relación de orden: $m \leq n$ si y sólo si existe un número r tal que $m + r = n$.

Otro hecho claro es que si $m + r = n + r$, entonces $m = n$. (Si al añadir r cosas a dos grupos terminan teniendo la misma cantidad de cosas, es que al principio ya tenían la misma cantidad de cosas). Similarmente se justifican otros hechos similares, tales como que si $m + n = 0$ entonces $m = n = 0$.

Multiplicación Multiplicar es sumar varias veces un mismo número. La fórmula $4 \cdot 3 = 12$ significa que $4 + 4 + 4 = 12$. Es claro entonces que $0 \cdot n = 0$ y convenimos que $n \cdot 0 = 0$. Basta pensar en un rectángulo dividido en cuadrados para convencerse de que la multiplicación es conmutativa (es decir, $mn = nm$) y basta pensar en un prisma dividido en cubos para convencerse de que es asociativa (o sea, $(mn)r = m(nr)$).

Similarmente se justifican otras propiedades elementales, como la distributividad del producto respecto de la suma, etc.

Exponenciación Finalmente tenemos la exponenciación de números naturales:

$$m^n = \overbrace{m \cdots m}^{n \text{ veces}}.$$

Convenimos en que $m^0 = 1$. No es difícil convencerse de hechos tales como que $m^{n+r} = m^n \cdot m^r$ y $m^{nr} = (m^n)^r$. Por ejemplo, si multiplicamos m^n por sí mismo r veces, como cada m^n se obtiene multiplicando n veces m, en total hemos multiplicado nr veces m.

Similarmente se prueban otros hechos conocidos, como que $1n = n$ y que si $n > 1$ entonces $1 = n^0 < n^1 < n^2 < n^3 < \cdots$

Inducción También es conocido el principio de inducción, según el cual:

> *Si* 0 *tiene una propiedad y podemos asegurar que si un número* n *cualquiera la tiene, también la tiene* $n+1$*, entonces todo número tiene esa propiedad.*

Está claro: el cero tiene la propiedad, según lo dicho podemos justificar que el 1 la tiene, de donde podemos justificar que 2 la tiene, y así sucesivamente. En el capítulo VI pueden encontrarse pruebas por inducción de la mayoría de los hechos referidos en esta sección. Las pruebas están formalizadas en teorías aritméticas, pero pueden tomarse también como pruebas metamatemáticas finitistas.

A este respecto hay que notar un hecho: dar una demostración (metamatemática) por inducción es dar una prueba de que 0 cumple algo junto con un argumento que nos garantice que $n+1$ lo cumple supuesto que n lo cumpla. A partir de estos datos es posible construir una prueba explícita de que cada número natural cumple lo pedido. Por ejemplo, una prueba de que 2 cumple lo pedido consistirá en

- La prueba de que 0 lo cumple, seguida de
- La prueba de que 1 lo cumple, basada en el argumento general y en que 0 lo cumple, seguida de
- La prueba de que 2 lo cumple, basada en el argumento general y en que 1 lo cumple.

De este modo una prueba por inducción puede entenderse como un esquema de prueba que da lugar a una prueba particular para cada número natural. Si se cumplen determinados requisitos las pruebas pueden ser constructivas, en el sentido de que si en ellas se afirma la existencia de determinados objetos, de la demostración puede extraerse un método para obtenerlos explícitamente.

Una variante del principio de inducción es como sigue:

> *Si podemos probar que un número natural cualquiera* n *cumple una propiedad supuesto que los números menores que* n *la cumplen, entonces todo número natural la cumple.*

(En particular estamos afirmando que 0 tiene la propiedad, pues "los números menores que 0" la cumplen, ya que no hay).

Buen orden Otro hecho elemental sobre números naturales es el siguiente principio de buena ordenación:

> *Si un número natural tiene una propiedad, entonces existe un número natural que es el menor con dicha propiedad.*

En efecto: no tenemos más que ir recorriendo en orden los números naturales hasta encontrarnos con el primero que tenga la propiedad y ése será el mínimo buscado.

División euclídea En el estudio de la aritmética natural juega un papel importante el siguiente resultado:

> Si D y d son números naturales con $d \neq 0$, entonces existen unos únicos c y r tales que $r < d$ y $D = dc + r$.

En este contexto D se llama *dividendo, d divisor, c cociente* y r *resto.* Está claro: si queremos repartir D cosas en d grupos, podemos ir asignando una a cada grupo tantas veces como sea posible hasta que nos quede una cantidad $r < d$ de cosas. Si en cada grupo han quedado c cosas, los números c y r cumplen la relación indicada. No es difícil justificar la unicidad.

B.2 Divisibilidad

Dados dos números naturales m y n, diremos que n es un *múltiplo* de m, que n es *divisible* entre m o que m es un *divisor* de n, y lo representaremos $m \mid n$, si existe un número r tal que $n = mr$. En caso contrario escribiremos $m \nmid n$.

Todo número tiene entre sus divisores a 1 y a sí mismo. Diremos que 1 y n son los *divisores impropios* de n. Un divisor *propio* es un divisor que no es impropio. 0 es múltiplo de todos los números, pues $n \cdot 0 = 0$. Sin embargo, los divisores de un número $n \neq 0$ son menores o iguales que n, luego cada número no nulo tiene sólo un número finito de divisores.

Es inmediato comprobar algunos hechos elementales, como que si $m \mid n$ y $n \mid r$ entonces $m \mid r$, o que si $d \mid m$ y $d \mid n$ entonces $d \mid m \pm n$.

Dados dos números m y n no simultáneamente nulos, definimos el *máximo común divisor* de m y n como el mayor de sus divisores comunes. Lo representaremos por (m, n).

Es inmediato comprobar que si $m > n$ entonces m y n tienen los mismos divisores comunes que $m - n$ y n. Por consiguiente $(m, n) = (m - n, n)$. Esta observación proporciona un método para calcular el máximo común divisor de cualquier par de números. Por ejemplo:

$$(60, 42) = (18, 42) = (18, 24) = (18, 6) = (12, 6) = (6, 6) = (6, 0) = 6.$$

En general, cada vez que reducimos un par restando sus componentes pasamos a otro cuya máxima componente es estrictamente menor que la del par anterior (salvo que una sea 0). Como el proceso no puede continuar indefinidamente siempre hemos de acabar con un par de la forma $(d, 0)$ o $(0, d)$, precedido del par (d, d), y entonces d es el máximo común divisor buscado. Más aún, vemos que d se obtiene a partir de m y n mediante un cálculo sucesivo de diferencias:

$$\begin{aligned}
(18, 42) &= (60 - 42,\ 42),\\
(18, 24) &= (60 - 42,\ 2 \cdot 42 - 60),\\
(18, 6) &= (60 - 42,\ 3 \cdot 42 - 2 \cdot 60),\\
(12, 6) &= (3 \cdot 60 - 4 \cdot 42,\ 3 \cdot 42 - 2 \cdot 60),\\
(6, 6) &= (5 \cdot 60 - 7 \cdot 42,\ 3 \cdot 42 - 2 \cdot 60).
\end{aligned}$$

Es claro que este proceso es general, por lo que tenemos probado el teorema siguiente:

Teorema (Bezout) *Si m y n son números naturales no simultáneamente nulos, entonces existen números naturales r y s tales que $(m,n) = rm - sn$ o bien $(m,n) = rn - sm$.*

De aquí se sigue que (m,n) no sólo es mayor que cualquier otro divisor común de m y n, sino que de hecho es múltiplo de todos ellos.

Un número p es *primo* si $p > 1$ y no tiene divisores propios. Por ejemplo los primeros primos son: 2, 3, 5, 7, 11, 13, 17, 19, 23, ...

Una propiedad fundamental de los primos es la siguiente: si p es primo y $p \mid mn$, entonces $p \mid m$ o bien $p \mid n$. En efecto, podemos suponer que m y n son ambos no nulos. Si $p \nmid m$, entonces, como p no tiene más divisores que 1 y p, necesariamente $(p,m) = 1$. Sean r y s tales que $1 = rm - sp$. Entonces $n = rmn - spn$. Como $p \mid mn$, es claro que $p \mid n$. ∎

Con esto podemos probar:

Teorema fundamental de la aritmética *Todo número natural mayor que* 1 *se descompone de forma única (salvo el orden) en producto de números primos.*

DEMOSTRACIÓN: Sea $n > 1$. Es claro que si p_1 es el menor divisor de n distinto de 1 entonces p_1 es primo, pues un divisor de p_1 distinto de 1 lo ha de ser también de n, luego ha de ser p_1. Así pues, $n = p_1 n_1$, para cierto número natural n_1. Si $n_1 > 1$ podemos aplicar el mismo razonamiento, con lo que $n = p_1 p_2 n_2$, para un cierto primo p_2. El proceso continúa mientras $n_i > 1$, pero como $n_1 > n_2 > n_3 > \cdots$ tras un número finito de pasos ha de ser $n_i = 1$, con lo que llegamos a una descomposición de n en producto de primos.

Supongamos ahora que $p_1 \cdots p_r = p'_1 \cdots p'_s$ son dos descomposiciones en primos de un mismo número. Digamos que $r \le s$. Como p_1 divide al miembro derecho, ha de dividir a alguno de los factores, pero como éstos son primos, ha de ser $p_1 = p'_i$, para algún i. No perdemos generalidad si suponemos que $p_1 = p'_1$. Simplificando pasamos a la igualdad $p_2 \cdots p_r = p'_2 \cdots p'_s$. Repitiendo el proceso podemos cancelar los r primos del miembro izquierdo original. Vemos entonces que ha de ser $r = s$ o, de lo contrario, llegaríamos a que $1 = p'_{s-r} \cdots p_s$, lo cual es imposible, ya que entonces p_s dividiría a 1.

Siendo $r = s$, lo que hemos probado es que, reordenado adecuadamente los factores de la derecha $p_i = p'_i$ para todo i, es decir, ambas factorizaciones son la misma. ∎

Llamaremos *exponente* de un primo p en un número n al número de veces que aparece p en la descomposición en primos de n. Lo representaremos por $e_p(n)$. Es fácil probar hechos como que $e_p(mn) = e_p(m) + e_p(n)$.

Otro hecho básico sobre números primos es el siguiente:

Teorema (Euclides) *Hay infinitos números primos.*

DEMOSTRACIÓN: Dado cualquier número natural n, consideremos $n!$, es decir, el producto de todos los números menores o iguales que n. Sea p un divisor primo de $n!+1$. Si fuera $p \leq n$ tendríamos que $p \mid n!$ y $p \mid n!+1$, luego $p \mid 1$, lo cual es absurdo. Consecuentemente $p > n$. Hemos probado que hay primos arbitrariamente grandes, luego hay infinitos. ∎

B.3 Congruencias

Diremos que dos números naturales m y n son *congruentes* módulo un tercer número d, y lo representaremos $m \equiv n \pmod{d}$ si su diferencia $m-n$ o $n-m$ es múltiplo de d.

Se cumple:

a) $m \equiv m \pmod{d}$.

b) Si $m \equiv n \pmod{d}$, entonces $n \equiv m \pmod{d}$.

c) Si $m \equiv n \pmod{d}$ y $n \equiv r \pmod{d}$, entonces $m \equiv r \pmod{d}$.

d) Dados m y $d \neq 0$, existe un único natural $r < d$ tal que $m \equiv r \pmod{d}$.

Las pruebas de estos resultados se hallan (formalizadas en una teoría aritmética arbitraria) en el capítulo VI.

Nótese que dados m, n y $d \neq 0$, si $m = dc + r$ y $n = dc' + r'$ con $r, r' < d$, entonces se cumple que $m \equiv r \pmod{d}$ y $n \equiv c' \pmod{d}$, luego se tendrá que $m \equiv n \pmod{d}$ si y sólo si $r \equiv r' \pmod{d}$, pero, por la unicidad, esto es si y sólo si $r = r'$. Hemos probado, pues, que dos números son congruentes módulo d si y sólo si el resto al dividirlos por d es el mismo.

Una última propiedad:

Si $m \equiv m' \pmod{d}$ *y* $n \equiv n' \pmod{d}$*, entonces*

$$m + n \equiv m' + n' \pmod{d} \qquad y \qquad mn \equiv m'n' \pmod{d}.$$

En efecto: Podemos suponer $d \neq 0$, o si no es evidente. Sea $m = dc + r$, $n = dc' + s$, $m' = de + r$, $n' = de' + s$, con $r, s < d$. Veamos el caso del producto. El de la suma es más simple. Tenemos que $mn = d(cdc' + cs + rc') + rs$, luego $mn \equiv rs \pmod{d}$ e igualmente $m'n' \equiv rs \pmod{d}$, de donde concluimos que $mn \equiv m'n' \pmod{d}$. ∎

Vamos a probar un resultado clásico sobre congruencias, conocido como teorema chino del resto. Para ello necesitamos un resultado previo:

Teorema *Si $(m,n) = 1$, para todo número c existe un número x tal que $xm \equiv c \pmod{n}$.*

DEMOSTRACIÓN: Podemos suponer $n \geq 2$, pues los otros casos son triviales. Entonces $m \neq 0$. Por el teorema de Bezout existen r y s tales que $rm - sn = 1$, $crm - c = csn$, luego $crm \equiv 1 \pmod{n}$. Basta tomar $x = cr$. ∎

Se dice que dos números son *primos entre sí* si su máximo común divisor es igual a 1.

Teorema chino del resto *Sean $n_1, \ldots, n_k$ números primos entre sí dos a dos. Sean $c_1, \ldots, c_k$ números cualesquiera. Entonces existe un número x tal que $x \equiv c_j \pmod{n_j}$ para todo $j = 1, \ldots, k$.*

DEMOSTRACIÓN: Sea m_j el producto de todos los n_i excepto n_j. Se cumple que $(m_j, n_j) = 1$, pues si un primo $p \mid (m_j, n_j)$, entonces $p \mid m_j$ y $p \mid n_j$. Por dividir a m_j resulta que $p \mid n_i$ para un índice $i \neq j$, luego $p \mid (n_i, n_j)$, cuando estamos suponiendo que son primos entre sí. Tenemos entonces que (m_j, n_j) no es divisible entre primos, por lo que ha de ser $(m_j, n_j) = 1$, como queríamos probar.

Sea x_j tal que $m_j x_j \equiv c_j \pmod{n_j}$, que existe por el teorema anterior. Sea $x = m_1 x_1 + \cdots + m_k x_k$. Veamos que cumple lo pedido. Dado j, notemos que si $i \neq j$, entonces $m_i \equiv 0 \pmod{n_j}$, pues $n_j \mid m_i$. Por tanto $m_i x_i \equiv 0 \pmod{n_j}$, luego $x \equiv m_j x_j \pmod{n_j}$, de donde $x \equiv c_j \pmod{n_j}$. ∎

Los resultados que siguen son más avanzados, pero sólo se usan en la última sección del capítulo VII. En ellos supondremos que el lector está familiarizado con los números enteros y racionales. Desde un punto de vista metamatemático, los números enteros son simplemente los números naturales con un signo positivo o negativo, con el convenio de que $+0 = -0 = 0$, es decir, los números enteros son:

$$\cdots\ -3,\ -2,\ -1,\ 0,\ +1,\ +2,\ +3,\ \cdots$$

No obstante, para deducir sus propiedades algebraicas a partir de las de los números naturales conviene introducirlos como es habitual en álgebra:

En el conjunto de los pares de números naturales, se define la relación dada por $(a,b)\ R\ (c,d)$ syss $a + d = b + c$. Es fácil ver que se trata de una relación de equivalencia. Se definen los números enteros como las clases de equivalencia determinadas por R. Representaremos por $[m,n]$ la clase de equivalencia del par (m,n).

Observamos que si $m > n$ entonces $[m,n] = [m-n, 0]$, mientras que si $m < n$ entonces $[m,n] = [0, n-m]$. Así pues, si llamamos $+m = [m,0]$ y $-m = [0,m]$, acabamos de probar que todo número entero es de la forma $\pm m$, para un cierto número natural m, es decir, que los números enteros son exactamente lo que hemos dicho antes.

En realidad, para que todo sea correcto hemos de probar que $\pm m = \pm n$ si y sólo si $m = n$ y el signo es el mismo (salvo si $m = n = 0$, en cuyo caso el signo no tiene por qué coincidir). La comprobación es muy simple.

La representación de los números enteros como clases de equivalencia nos permite definir las operaciones

$$[m,n]+[m',n']=[m+m',n+n'], \qquad [m,n][m',n']=[mm'+nn',mn'+nm'].$$

Es simple rutina comprobar que estas operaciones tienen las propiedades que cabe esperar. La relación de orden se define mediante $x \le y$ syss $y-x$ es positivo.

Similarmente, para definir los números racionales consideramos el conjunto de los pares de números enteros (x,y) tales que la segunda componente sea no nula. Sobre estos pares, definimos la relación dada por $(x,y)\,R\,(x',y')$ syss $xy'=yx'$. Es fácil probar que se trata de una relación de equivalencia. Representaremos por x/y a la clase de equivalencia del par (x,y). Llamaremos *números racionales* a estas clases. Por definición

$$\frac{x}{y}=\frac{x'}{y'} \qquad \text{syss} \qquad xy'=yx'.$$

La suma, el producto y la relación de orden en los números racionales se definen como es habitual. Con esto podemos probar un teorema clásico:

Teorema (Langrange) *Todo número natural es suma de cuatro cuadrados.*

Demostración: El problema se reduce a estudiar los números primos gracias a la identidad:

$$\begin{aligned}&(x^2+y^2+z^2+w^2)(x'^2+y'^2+z'^2+w'^2)\\ &=(xx'-yy'-zz'-ww')^2+(xy'-yx'+zw'-wz')^2\\ &+(xz'+zx'+wy'-yw')^2+(xw'+wx'+yz'-zy')^2. \qquad \text{(B.1)}\end{aligned}$$

Esta fórmula surge de la teoría de los cuaternios, pero no vamos a entrar en ello. El lector puede comprobarla directamente. De ella se sigue que el producto de números expresables como suma de cuatro cuadrados es también expresable como suma de cuatro cuadrados, luego basta probar que todo número primo es expresable como suma de cuatro cuadrados.

Como $2=1^2+1^2+0^2+0^2$, basta verlo para primos impares. Sea, pues, p un primo impar.

Los números x^2 con $0\le x\le \frac{1}{2}(p-1)$ son incongruentes módulo p, e igualmente ocurre con $-1-y^2$ con $0\le y\le \frac{1}{2}(p-1)$.

Como en total son $p+1$, existen x, y en estas condiciones tales que

$$x^2\equiv -1-y^2 \ (\text{mód } p). \qquad \text{(B.2)}$$

Entonces $x^2<(\frac{1}{2}p)^2$, $y^2<(\frac{1}{2}p)^2$, luego $x^2+y^2<2(\frac{1}{2}p)^2$,

$$x^2+y^2+1<1+2\Big(\frac{1}{2}p\Big)^2<p^2, \qquad \text{(B.3)}$$

Por (B.2) $x^2+y^2+1 \equiv 0$ (mód p), luego $x^2+y^2+1 = mp < p^2$ (por (B.3)), de donde $0 < m < p$.

Sea r el menor natural no nulo tal que existen números enteros x, y, z, w que cumplan $rp = x^2+y^2+z^2+w^2$. Como m cumple esto, será $r \le m < p$.

Necesariamente r es impar, pues si fuera par, 0, 2 o 4 de los x, y, z, w serían pares y, reordenándolos, podríamos exigir que $x+y$, $x-y$, $z+w$ y $z-w$ fueran pares.

Entonces $\frac{1}{2}rp = \left(\frac{1}{2}(x+y)\right)^2 + \left(\frac{1}{2}(x-y)\right)^2 + \left(\frac{1}{2}(z+w)\right)^2 + \left(\frac{1}{2}(z-w)\right)^2$, en contradicción con la minimalidad de r.

Nuestro objetivo es probar que $r = 1$. Supongamos que $r > 1$.

Sean x', y', z', w' los restos módulo r de x, y, z, w entre $-r/2$ y $r/2$ (es posible ya que r es impar).

Claramente $n = x'^2+y'^2+z'^2+w'^2 \equiv x^2+y^2+z^2+w^2 = rp \equiv 0$ (mód r), pero $n > 0$, pues en otro caso $x' = y' = z' = w' = 0$, r dividiría x, y, z, w, luego $r^2 \mid x^2+y^2+z^2+w^2 = rp$, de donde $r \mid p$ y en consecuencia $r = 1$, contra lo supuesto. También es claro que $n < 4(\frac{1}{2}r)^2 = r^2$.

Sea $0 < k < r$ tal que $n = kr$. Por la identidad (B.1), $krpr = z_1^2+z_2^2+z_3^2+z_4^2$ para ciertos naturales z_1, z_2, z_3, z_4 y, teniendo en cuenta cómo se obtienen a partir de x, y, z, w, x', y', z', w', es claro que los cuatro son múltiplos de r (por ejemplo $z_1 = xx' - yy' - zz' - ww' \equiv x^2-y^2-z^2-w^2 = rp \equiv 0$ (mód r)).

Así pues $z_i = rt_i$ y por tanto $r^2kp = r^2t_1^2 + r^2t_2^2 + r^2t_3^2 + r^2t_4^2$, con lo que $kp = t_1^2+t_2^2+t_3^2+t_4^2$, en contra de la minimalidad de r. ∎

B.4 Cuerpos cuadráticos

En la última sección del capítulo VII usaremos cuerpos cuadráticos $\mathbb{Q}(\sqrt{d}\,)$. Si trabajamos en teoría de conjuntos, podemos definir simplemente, para cada número natural d que no sea un cuadrado perfecto,

$$\mathbb{Q}(\sqrt{d}\,) = \{a + b\sqrt{d} \mid a,\ b \in \mathbb{Q}\}.$$

Ahora bien, esta definición presupone el conjunto de los números reales, del cual no podemos hablar metamatemáticamente porque involucra la totalidad de las sucesiones de números racionales o la totalidad de los subconjuntos de $\mathbb{Q}$, y no tenemos una intuición completa de estas nociones (sabemos dar significado a cualquier afirmación sobre la totalidad de los números enteros o racionales, pero no si involucra a la totalidad de los números reales). Dado el contexto en el que vamos a aplicar estos cuerpos, resulta interesante constatar que podemos hablar de ellos metamatemáticamente, por lo que vamos a definirlos aquí sin hacer referencia a los números reales. El lector reconocerá sin duda técnicas habituales en el trato de los números reales, pero lo importante es que las aplicamos sin involucrar en ningún momento nociones metamatemáticamente ambiguas.

Fijemos un número natural d que no sea un cuadrado perfecto. Es fácil ver entonces que d tampoco es el cuadrado de ningún número racional, pues si

$d = a^2/b^2$ entonces $a^2 = db^2$, pero d es divisible entre un primo p con exponente impar, y entonces el exponente de p en el miembro izquierdo es par, mientras que en el miembro derecho es impar, lo cual es una contradicción.

Definimos $\mathbb{Q}(\sqrt{d}\,)$ como el conjunto de todos los pares de números racionales (a, b). Definimos las operaciones

$$(a, b) + (a', b') = (a + a', b + b'), \qquad (a, b)(a', b') = (aa' + bb'd, ab' + ba').$$

Podemos identificar cada número racional a con el par $(a, 0)$, y entonces la suma y el producto de $\mathbb{Q}(\sqrt{d}\,)$ extienden a los de $\mathbb{Q}$. Si definimos $\sqrt{d} = (0, 1)$ se cumple que $(\sqrt{d}\,)^2 = d$. Además

$$(a, b) = (a, 0) + (b, 0)(0, 1) = a + b\sqrt{d}.$$

De este modo, todo elemento de $\mathbb{Q}(\sqrt{d}\,)$ se expresa de forma única como $a+b\sqrt{d}$, con a y b racionales. Es fácil comprobar que $\mathbb{Q}(\sqrt{d}\,)$ tiene estructura de anillo conmutativo y unitario. De hecho es un cuerpo. Para probar esto último definimos el *conjugado* de $\alpha = a + b\sqrt{d}$ como $\overline{\alpha} = a - b\sqrt{d}$. Se comprueba que

$$\overline{\alpha + \beta} = \overline{\alpha} + \overline{\beta}, \qquad \overline{\alpha\beta} = \overline{\alpha}\,\overline{\beta}.$$

Definimos la *norma* de $\alpha = a+b\sqrt{d}$ como $\mathrm{N}(\alpha) = \alpha\,\overline{\alpha} = a^2-db^2$. Claramente $\mathrm{N}(\alpha\beta) = \mathrm{N}(\alpha)\,\mathrm{N}(\beta)$. Además, si $\alpha \neq 0$ entonces $\mathrm{N}(\alpha) \neq 0$, pues si $a^2 - db^2 = 0$, entonces $d = a^2/b^2$, lo cual ya hemos visto que es imposible. Esto nos permite escribir

$$\frac{\alpha\overline{\alpha}}{\mathrm{N}(\alpha)} = 1,$$

luego

$$\alpha^{-1} = \frac{\overline{\alpha}}{\mathrm{N}(\alpha)}, \qquad \text{para } \alpha \neq 0.$$

El punto más delicado es definir la relación de orden en $\mathbb{Q}(\sqrt{d}\,)$. Se trata de la relación inducida por el orden usual de los números reales, pero queremos definirlo sin hacer referencia a éstos.

Fijemos un número natural $n > 0$ y consideremos la sucesión $(k/n)^2$, donde k recorre los números naturales. Es claro que para $k = nd$ toma el valor $d^2 > d$. Sea $k_n + 1$ el mínimo natural tal que $(k_n + 1/n)^2 > d$. Entonces

$$\left(\frac{k_n}{n}\right)^2 < d < \left(\frac{k_n + 1}{n}\right)^2.$$

Definimos $r_n = k_n/n$. Así, $r_n^2 < d < d^2$, luego $r_n < d$. Más aún, observemos que

$$d - r_n^2 < \left(\frac{k_n + 1}{n}\right)^2 - \left(\frac{k_n}{n}\right)^2 = \frac{2k_n + 1}{n^2} = \frac{2r_n}{n} + \frac{1}{n^2} < \frac{2d}{n} + \frac{1}{n^2}.$$

De aquí se sigue que el término izquierdo se hace arbitrariamente pequeño cuando n se toma suficientemente grande. Por consiguiente, si $c < d$, existe un n_0 tal que si $n \geq n_0$ entonces $c < r_n^2 < d$.

Para cada $\alpha = a + b\sqrt{d}$, definimos $\alpha_n = a + br_n$. Se cumple:

Teorema Dado $\alpha \neq 0$ en $\mathbb{Q}(\sqrt{d}\,)$, existen naturales k y n tales que o bien para todo $n \geq n_0$ se cumple $\alpha_n \geq 1/k$, o bien para todo $n \geq n_0$ se cumple $\alpha_n \leq -1/k$.

DEMOSTRACIÓN: Podemos expresar $\alpha = \pm a \pm b\sqrt{d}$, con a, $b \geq 0$. Distinguimos las cuatro posibilidades para los signos.

Si $\alpha = a + b\sqrt{d}$, notamos que $d > 1$, luego para n grande $r_n^2 > 1$, luego $r_n > 1$, luego $a + br_n > a + b$. Basta tomar k tal que $a + b > 1/k$. El caso $\alpha = -a - b\sqrt{d}$ es similar.

Si $\alpha = a - b\sqrt{d}$, podemos suponer $b \neq 0$. Distinguimos dos casos, si $a^2 < b^2 d$ queremos conseguir $a - br_n \leq -1/k$, lo que equivale a que $r_n \geq a/b + 1/(kb)$. Tendremos esto si conseguimos que $r_n^2 \geq a^2/b^2 + 1/(kb)^2 + 2a/(kb^2)$. Ahora bien, tenemos que $d > a^2/b^2$, luego, para un cierto k suficientemente grande, se cumple $d > a^2/b^2 + 1/(kb)^2 + 2a/(kb^2)$. Como los r_n^2 aproximan a d, para n suficientemente grande se tiene la desigualdad que necesitamos.

Si $a^2 > b^2 d$ hemos de ver que $a - br_n \geq 1/k$, lo que a su vez equivale a que $r_n \leq a/b - 1/(bk)$, o a que $r_n^2 \leq a^2/b^2 + 1/(bk)^2 - 2a/(b^2k)$. Tenemos que $d < a^2/b^2$, luego, tomando k grande, $r_n^2 \leq d < a^2/b^2 + 1/(bk)^2 - 2a/(b^2k)$. El caso que falta es análogo. ∎

Diremos que $\alpha > 0$ o que $\alpha < 0$ según si los números α_n son finalmente $\geq 1/k$ o $\leq -1/k$. Teniendo en cuenta que $(\alpha + \beta)_n = \alpha_n + \beta_n$, es claro que la suma de números positivos es positiva.

Esto nos permite definir $\alpha \leq \beta$ como $\beta - \alpha \geq 0$, y es inmediato comprobar que se trata de una relación de orden total que extiende a la relación de orden usual en $\mathbb{Q}$.

Finalmente observamos que si $\alpha = a + b\sqrt{d} > 0$ y $\beta = a' + b'\sqrt{d} > 0$, entonces, para n suficientemente grande,

$$(\alpha\beta)_n = \alpha_n\beta_n + bb'(d - r_n^2) \geq \frac{1}{k^2} - \frac{1}{2k^2} = \frac{1}{2k^2},$$

con lo que $\alpha\beta > 0$. De aquí se siguen inmediatamente todas las propiedades que cabe esperar. Por ejemplo, si $\alpha \leq \beta$ y $\gamma \geq 0$ entonces $\alpha\gamma \leq \beta\gamma$, ya que $\gamma(\beta - \alpha) \geq 0$.

Notemos también que $a + b\sqrt{d} \leq |a| + |b|d$, es decir, que todo número cuadrático está acotado por un número natural. Esto se suele expresar diciendo que $\mathbb{Q}(\sqrt{d}\,)$ es un cuerpo arquimediano. Así, si $\alpha > 1$ la sucesión α^n es monótona creciente y tiende a infinito, en el sentido de que supera a cualquier número natural. Basta considerar el segundo término del desarrollo binomial de $\alpha^n = (1 + (\alpha - 1))^n$, cuyos términos son todos positivos, para concluir que $n(\alpha - 1) \leq \alpha^n$. Para conseguir $\alpha^n > m$ basta tomar $n > m/(\alpha - 1)$.

Bibliografía

[1] Baker, A. *Breve Introducción a la Teoría de Números,* Alianza Universidad, Madrid, 1986.

[2] Barwise, J. (editor), *Handbook of Mathematical Logic,* North Holland, Amsterdam, 1977.

[3] Bernays, P. y Fraenkel, A. *Axiomatic Set Theory,* North Holland, Amsterdam, 1958.

[4] Burke, M.R. y Magidor, M. *Shelah's pcf Theory and its Applications,* Ann. Pure and Appl. Logic **50** (1990) pp. 207-254.

[5] Cantor, G. *Contributions to the Founding of the Theory of Transfinite Numbers,* New York, (1955).

[6] Cohen, P. *Set Theory and the Continuum Hypothesis,* W.A.Benjamin inc. reading, New York, 1966.

[7] Davis, M. *Hilbert's Tenth Problem is Unsolvable,* Am. Math. Monthly **80** (1973) pp. 233–269.

[8] Devlin, K. J. *Fundamentals of Contemporary Set Theory.* Springer, New York.

[9] Enderton, H. B. *Elements of Recursion Theory,* (en Barwise).

[10] Gödel, K. *Obras completas,* Alianza Universidad, Madrid, 1981.

[11] — *Sobre Sentencias Formalmente Indecidibles de Principia Mathematica y Sistemas Afines,* (1931).

[12] — *Sobre Sentencias Indecidibles de Sistemas Formales Matemáticos,* (1934).

[13] — *La Consistencia del Axioma de Elección y de la Hipótesis Generalizada del Continuo con los Axiomas de la Teoría de Conjuntos,* (1940).

[14] Hamilton, A. G. *Lógica para Matemáticos,* Paraninfo, Madrid, 1981.

[15] Jech, T.J. *The Axiom of Choice,* North Holand, Amsterdam, 1973.

[16] — *Set Theory.* Academic Press, New York, 1978.

[17] — y Shelah, S. *On a conjecture of Tarski on products of cardinals,* Proc. Amer. Math. Soc. **112**, 4 (1991) pp. 1117–1124.

[18] Kleene, S. C. *Introducción a la Metamatemática,* Tecnos, Madrid, 1974.

[19] Kunen, K. *Set Theory. An Introduction to Independence Proofs,* North Holland, Amsterdam, 1985.

[20] Mosterín, J. *Lógica de Primer Orden,* Ariel, Barcelona, 1970.

[21] — *Teoría Axiomática de Conjuntos,* Ariel, Barcelona, 1971.

[22] Smorynski, C. *The Incompleteness Theorems,* (en Barwise).

[23] Specker, E. *Verallgemeinerte Kontinuumshypothese und Auswahlaxiom.* Archiv. Math. (Basel) **5** (1954) pp. 332–337.

Índice de Materias