Universidad de Valladolid

Facultad de Ciencias

Los teoremas de incompletitud de Gödel

Trabajo de Fin de Grado
Grado en Matemáticas

Autor: Ándrés Infante Adrián

Tutor: Antonio Campillo López

2023

A todos los que dedican su vida al estudio,
como Kurt Gödel.

Índice general

Introducción

Este trabajo trata, como el título reza, de los teoremas de incompletitud de Gödel. El objetivo es desengranar los teoremas de forma tan completa como nos sea posible atendiendo a las exigencias de este tipo de trabajos. Especifiquemos más este objetivo. Para ello, observemos la siguiente formulación simple de los teoremas:

- **Primer teorema** Todo sistema formal de la matemática clásica que es consistente es incompleto.

- **Segundo teorema** No es posible probar la consistencia de un sistema formal de la matemática clásica.

Si reflexionamos sobre estos enunciados, posiblemente lo primero en que recalemos es que estos teoremas no son parte del Análisis, ni de la Geometría, ni de la Aritmética... ni siquiera de la Lógica. En esta primera percepción cabe pensar que estos teoremas, «stricto sensu», no son teoremas de la Matemática, sino que van más allá de la Matemática, que desbordan su campo. La *consistencia* y la *completitud* son propiedades del sistema formal, externas a él. Digamos que no *hablan* matemáticas, sino *de* matemáticas; en concreto de la lógica matemática. A este lenguaje que habla acerca de las matemáticas, Hilbert dio en llamarlo «metamatemática». Podemos decir que son teoremas de la metalógica.

Sin embargo, reparamos también en que el uso de la palabra *teorema* no puede ser casual. Los metateoremas lógicos son, con todo el derecho, resultados matemáticos. En efecto, las formulaciones metalógicas pueden demostrarse *en la lógica*, es decir, en la Matemática; su construcción es interna a las matemáticas. De alguna manera hablan, al mismo tiempo, desde dentro y fuera de las matemáticas. Aquí entra en juego la genialidad de Gödel. La demostración del teorema de Gödel se lleva a cabo en un lenguaje formal, es decir, dentro de la lógica matemática y no fuera. En concreto aparecerán tres tipos de lenguajes distintos: el lenguaje metamatemático o natural, el lenguaje formal y el lenguaje matemático relativo a un modelo del sistema formal. El cambio de lenguaje, la legitimidad para hacerlo, es la clave de la demostración. En particular, en este y en el resto de metateoremas (como el teorema de Löwenheim-Skolem-Tarski o el teorema de la Indefinibilidad de la Verdad de Trasrki) la técnica crucial radica en la Teoría de Modelos, que permite una interpretación semántica (matemática, que nos permite evaluar la *veracidad*) de las sentencias de un lenguaje formal.

Pero no ignoremos las consecuencias no *estrictamente* matemáticas (formales), en este caso, del teorema de Gödel. Admitimos su clara irrupción en el campo de la metodología matemática, esto es, en el campo de la Filosofía de la Matemática. No creemos descubrir nada con esto: la lógica matemática (no exclusivamente, véase el impacto del desarrollo de la geometría no euclidiana para el intuicionismo), en tanto en cuanto trata la cuestión metodológica (la construcción) de las matemáticas, estudia los fundamentos de la Matemática. Y esto es una cuestión filosófica (no es casualidad que los más grandes lógicos hayan sido, además, grandes filósofos). En este sentido, hacemos nuestras las siguiente palabras de Jesús Padilla Gálvez:

> Al mismo tiempo los lenguajes en los que se ha estructurado la noción de verdad y de los que habla la teoría de modelos son, por lo general, sistemas matemáticos. Las «cosas» representadas en dichos lenguajes son también sistemas matemáticos. Por esto, la teoría de modelos es una teoría semántica que pone en relación unos sistemas matemáticos con otros sistemas matemáticos. Dicha teoría nos proporciona algunas pistas con respecto a aquella semántica que pone en relación los lenguajes naturales con la realidad. Sin embargo, ha de tenerse siempre presente que no hay ningún sustituto matemático para los problemas genuinamente filosóficos. Y el problema de la verdad es un problema netamente filosófico. ([28] Padilla 2007, 229)

Nos encontramos pues con dos partes en el teorema: el significado filosófico y el formal, ambos amparados por su demostración formal. Ante el objetivo inicial del trabajo, a saber, «desengranar los teoremas de forma tan completa como nos sea posible», nos parece necesario y legítimo tratar estas dos vertientes en profundidad. Por eso hacemos este planteamiento:

Antes de nada expondremos el contexto histórico en el que se enmarca el trabajo de Gödel. A continuación , en el primer capítulo, hacemos un análisis de las corrientes clásicas en filosofía de la Matemática: platonismo, logicismo, intuicionismo y formalismo. En este repaso tratamos de centrarnos en lo relacionado con los teoremas de incompletitud de Gödel, pero se busca una visión general.

Posteriormente hacemos una aproximación formal a los teoremas. El objetivo principal es conseguir llegar a una demostración adecuada y poder comprenderla. Para ello debemos tratar las teorías formales de primer orden, a lo que dedicamos las dos primeras secciones del capítulo. La tercera versa sobre otros metateoremas lógicos que son fundamentales, como el teorema de completitu de Gödel o el teorema de Löwenheim-Skolem-Tarski. Finalmente, en la cuarta sección del capítulo dos, exponemos en profundidad los teoremas de incompletitud. Comenzamos analizando la intruducción de Gödel a su artículo original. Posteriormente tratamos en detalle los tres pilares de los teoremas: la gödelización, la teoría de la recursión y la representabilidad sintáctica. A continuación explicamos las demostraciones y analizamos el significado de los teoremas, y terminamos con dos apartados que creemos cierran el círculo, no sólo para entender los teoremas de Gödel, sino también sus consecuencias inmediatas: presentamos una sentencia puramente aritmética que no es demostrable y exponemos las ideas de Alan Turing en relación a la computabilidad.

Como no podía ser de otra manera, dedicamos unas páginas a comentar las conclusiones que hemos extraído de este trabajo. Además hemos añadido un apéndice interesantísimo. Se trata de una demostración que hizo Gödel, aunque nunca la publicó, del argumento ontológico sobre la exitencia de Dios. Dado que en este trabajo hemos hablado mucho de filosofía y de Kurt Gödel (no sólo de sus teoremas de incompletitud), nos parece adecuado incluirlo.

Queremos hacer constar los dos siguientes puntos: que pretendemos demostrar el teorema de Gödel, pero que la demostración no será realmente rigorosa debido a la complejidad de la misma, sino superficial; y que no pretendemos elaborar una teoría filosófica de la Matemática, sino simplemente exponer con claridad las interpretaciones filosóficas de la Matemática más importantes. Posteriormente explicaremos nuestra idea de filosofía, pero ya adelantamos que, a diferencia de las matemáticas, no es una ciencia, y que no caben resultados *científicos* en ella. Con esto no negamos su papel en el conjunto del saber.

Como aclaración, en la bibliografía hemos optado por escribir el año de publicación original inmediatamente después del nombre del autor. En el caso en el que hemos usado una edición de años posteriores, escribimos la fecha al final, después de la editorial y el lugar de edición.

Contexto histórico

A principios del siglo XIX, Lobachevski y Bolyai, e, independientemente, Gauss, desarrollaron la primera geometría no euclídea: la hiperbólica. A mediados vendría la geometría elíptica de la mano de Riemann. Por dos mil años, los *Elementos* de Euclides habían sido la base de la geometría, que ahora se revelaba más compleja. Otros matemáticos había explorado previamente los fallos del sistema euclideano, pero su autoridad era tal que incluso algunos, como Saccheri, acabaron tomando sus propios trabajos como absurdos y aceptando el famoso postulado de las paralelas de los *Elementos*.

Los axiomas de Euclides (mas que de axiomas en el sentido moderno, los postulados de los *Elementos* consisten en ciertos métodos válidos para la construcción geométrica) se consideraban universales y autoevidentes: eran afirmaciones verdaderas sobre el espacio que percibimos sensorialmente. Sin embargo, ahora, otros axiomas, contradictrios a los del maestro, llevaban a geometrías no euclidianas. El problema era que nuestra intuición espacial ordinaria no se ajustaba a estas nuevas geometrías: *parecían* claramente falsas. La forma de tratar los axiomas no euclideanos, de constatar su validez, o no, será plantear el problema de su consistencia interna, es decir, garantizar que no son contradictorios. La misma posibilidad de las geometrías no euclideanas dependía de este problema.

Todo esto provocó una renovación en el interés (y la necesidad) por fundamentar la Matemática. Se abría paso la concepción formalista: la tarea del matemático es deducir las consecuencias lógicas *necesarias* de los axiomas. Se trata de revisar y perfeccionar los axiomas de los distintos sistemas matemáticos. El programa formalista, que será impulsado por el prestigioso matemático David Hilbert, constará de dos cuestiones fundamentales: la primera, construir un sistema que formalice completamente la matemática clásica; la segunda, demostrar la consistencia del sistema. Los conceptos matemáticos han de ser reemplazados por signos gráficos carentes de sentido, la demostración se reduce a la deducción formal conforme a reglas mecánicas. Deducir: esta es la tarea del matemático.

En 1879 Gottlob Frege publica *Begriffsschrift* (Ideografía), iniciando de manera sorprendente la lógica moderna. Michael Dummet señala: «[la obra] es asombrosa porque no tiene precedentes: parece haber surgido del cerebro de Frege no fertilizado por influencias externas» ([9] Dummet, 1973, 17). El lenguaje formal creado por

Frege es un vehículo directo al logicismo y al estudio de los sistemas matemáticos desde una perspectiva formal. El mismo Frege publica, en 1884, los *Grundlagen der Arithmetik* (Fundamentos de la Aritmética), de nuevo, precursor, esta vez en la Filosofía de las Matemáticas según se entiende actualmente.

Entre 1878 y 1884 Georg Cantor desarrolla la Teoría de Conjuntos, que rápidamente se revela a la vez potente y enigmática. Los estudios de Cantor sobre cardinales y ordinales sumaban a la fiesta al que Borges consideraba «[el concepto] que es el corruptor y el desatinador de los otros. No hablo del Mal cuyo limitado imperio es la ética; hablo del infinito». Cantor demostró que existían diferentes tipos de infinitos inconmensurables entre sí, es decir, no todos los conjuntos infinitos eran equipotentes. Para Javier de Lorenzo, la teoría intuitiva de conjuntos supone "una ruptura epistemológica en el hacer matemático", que constituye, frente al constructivismo anterior, "una teoría auténtica del infinito actual" ([6] De Lorenzo, 1979). Así, la Teoría de Conjuntos se granjeó un enemigo notable: los matemáticos intuicionistas.

En una carta del 16 de junio de 1902, el joven matemático Bertrand Russell alerta a Frege de un error catastrófico en sus leyes lógicas básicas: la famosa *paradoja de Russell*. A pesar de ello, Russell y Alfred Whitehead se erigen como dignos sucesores para relevar a Frege en el programa logicista. Tras un trabajo monumental, que se plasmaría en los *Principia Mathematica*, publicados entre 1910 y 1913, consiguieron soslayar las antinomias y paradojas, pagando el precio de crear una teoría increíblemente abstracta y compleja. Hasta los años treinta se admite, a pesar de las duras críticas al *axioma de reducibilidad* de Russell y Whitehead, que los *Principia* (también otras teorías) conseguían formalizar completamente la matemática clásica. Sin embargo la consistencia del sistema dependía de la consistencia de la Teoría de Conjuntos.

El avance formalista continuaba: Peano axiomatiza la Aritmética en 1889 en sus *Grundlagen der Arithmetik* (Fundamentos de la Aritmética), que luego refinaría en 1897, reduciendo los axiomas a cinco: los *Axiomas de Peano*; y, en 1898, Hilbert axiomatiza la geometría en *Grundlagen der Geometrie* (Fundamentos de la geometría). Al igual que los *Principia*, la consistencia de ambos sistemas sigue dependiendo de la consistencia de la Teoría de Conjuntos. Posteriormente se axiomatiza la *bête noire*, la Teoría de Conjuntos de Cantor, que había quedado seriamente tocada por las paradojas y antinomias. La primera publicación de Zermelo, de 1908, se completa en 1922 con las contribuciones de Fraenkel, estableciendo los axiomas de Zermelo-Fraenkel (por sus iniciales ZF, o ZFC si incluye el *axioma de elección*) y logrando evitar las paradojas de Russell, Cantor, y otras tantas.

En el año 1900, en el Congreso Internacional de Matemáticos en París, Hilbert propone 23 problemas matemáticos que él consideraba claves para la matemática futura. El segundo dice así: *Probar que los axiomas de la aritmética son consistentes (esto es, que la aritmética es un sistema formal que no supone una contradicción).* Hilbert estaba convencido de que esta conjetura era cierta, de que un sistema formal tiene tantos teoremas formales como matemáticos, y que no cabe deducir contradicciones. Recordemos sus famosas palabras:

> La convicción de la resolubilidad de todo problema matemático ha sido siempre un potente estímulo para el trabajo científico; dentro de nosotros resuena siempre el lema: «aquí está el problema, busca la solución». Puedes buscarla con el pensamiento puro, porque en la matemática no cabe el Ignorabimus. ([20], 1898)

En septiembre de 1930 hubo un congreso celebrado en Königsberg, la ciudad natal de Hilbert. En una entrevista en la radio local pronunció las célebres palabras que hoy están escritas en su tumba (y de las cuales se conserva un audio): «*Wir müssen wissen, wir werden wissen*» («Debemos saber, sabremos»). Magnífica coincidencia: en el mismo congreso, dos días antes, el lógico austríaco Kurt Gödel, de sólo 23 años (los mismos que tiene quien escribe este trabajo; nótense los abismos, los oceános, las distancias cósmicas que nos separan a los mortales de los grandes genios), exponía un resultado sorprendente a una audencia más bien escasa: había demostrado que había proposiciones aritméticas verdaderas que no podían demostrarse en el sistema formal, y que la consistencia del mismo es precisamente una de ellas. Al parecer la repercusión no fue muy amplia, aunque sí sabemos que un gran matemático como von Neumann quedó sorprendido por las cualidades de Gödel. Un año después, Gödel publica en la revista «*Monatshefte für Mathematik un Physik*» sus teoremas de incompletitud en el artículo «*Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I*» («Sobre proposiciones formalmente indecidibles de Principia Mathematica y sistemas relacionados»). Es el fin del proyecto formalista como aspiración a fundamentar las Matemáticas. Gödel trabaja con el sistema formal de *Principia*, pero los teoremas abarcan de forma análoga todos los sistemas formales que sean capaces de representar aritmética básica, en particular la Teoría de Conjuntos, asegurando que, si ZFC no es una teoría contradictoria (es consistente), entonces existe una fórmula *verdadera* que no es demostrable en ZFC. Y, además, que la consistencia de ZFC no es demostrable en ZFC[1].

[1] Para un desarrollo amplio de la vida de Frege, von Neumann, Cantor, Russell y Gödel, véase ([26] Mosterín, 2000)

Capítulo 1

Aproximación filosófica

Antes de nada queremos dejar claro que el artículo ([23] Madrid Casado, 2009) ha sido fundamental para la redacción de este apartado, sirviendo como base. Agradecemos también al autor sus aclaraciones.

Nos vamos a acoger (aquí, sin otra justificación que la conveniencia, pues somos más o menos legos en la materia) a la definición de filosofía del filósofo español Gustavo Bueno. A todos nos es conocido el lema de la Academia de Platón: «No entre aquí quien no sepa geometría» (el gran filósofo griego no se quedó aquí en sus exigencias matemáticas, y llegaba a pedir la pena de muerte para los profesores que no enseñaran a sus alumnos los números irracionales). Metafóricamente Bueno define la filosofía como la "Geometría de las Ideas". La filosofía *razona*: con las leyes de la lógica y la dialéctica, y sobre las Ideas. ¿Qué son las Ideas? Consideramos dadas las ciencias positivas actuales (física, matemáticas, química, biología...[1]), que operan con conceptos. Sin embargo, hay conceptos que no se consiguen enclaustrar en ninguna de las ciencias, pues pertenecen a distintos contextos al mismo tiempo, ya sean científicos o no. Son estos conceptos a los que llamamos Ideas, que por resistirse a ser captados por un campo científico concreto, quedan *fuera*, desbordan las ciencias y precisan un tratamiento filosófico. Las Ideas tratan de conceptos, de su coordinación entre ellos y con otros. Los conceptos, por contraposición a las Ideas, son configuraciones definidas de ciertos contenidos de un campo científico[2]. Por ejemplo, la Idea de Espacio engloba los conceptos científicos de espacio vectorial (matemáticas) o de espacio newtoniano (física); no es enclasable en ninguna ciencia particular.

Según esta concepción, la Idea de Ciencia es un concepto filósofico, y no científico, que trata de las relaciones entre las distintas ciencias positivas: en qué se diferencian (y por qué), qué tienen en común, etc. La Idea de Matemática tratará, pues, de cuestiones que van más allá de las matemáticas, que quedan fuera de su campo. Qué

[1]Para una aclaración sobre esto y las siguientes ideas sobre qué es la ciencia, véase ([4] Bueno, 1995b).

[2]Para un desarrollo de todo esto véase ([4] Bueno, 1995a).

es hacer matemáticas, qué es la Matemática, cuál es su relación con las otras ciencias, son preguntas filosóficas. Reparamos pronto en que hay varias respuestas posibles, o sea, que hay varias Ideas de la Matemática. Tratamos esto en el primer apartado de esta sección, donde hacemos un amplio repaso de las cuatro corrientes clásicas en Filosofía de las Matemáticas: platonismo, logicismo, intuicionismo y formalismo. Queremos destacar también la corriente del materialismo formalista de Carlos M. Madrid Casado ([23] Madrid Casado, 2009), enmarcada en el sistema filosófico de Gustavo Bueno, el materialismo filosófico. No lo hemos incluido en este trabajo para no alargarlo, y considerando que desde un punto de vista general lo adecuado era tartar las cuatro verteintes clásicas.

1.1. Filosofías de la Matemática

Las ciencias positivas, como la física o la química, tratan con términos que refieren a ciertas cosas «materiales». Por ejemplo, en el campo de la química están el hidrógeno, el carbono, etc, cuyos referenciales son ciertos elementos naturales, «corpóreos», con ciertas propiedades. El campo de la física contiene términos como electrones o campos, que asumimos que existen en el universo. Estas teorías recurren a un lenguaje muy preciso, y además muy distinto al lenguaje natural: las matemáticas. Pero, ¿a qué refieren las matemáticas? Conjunto, función, variedad, grupo, etc, son objetos matemáticos que no tienen existencia espacio-temporal (material). Sin embargo, claramente, existen; y aún más: son parte indisociable e indispensable de nuestras mejores teorías científicas que aspiran a describir el mundo. Razón suficiente para que tratemos de responder a las grandes preguntas: ¿qué son los objetos matemáticos? ¿cuál es su estatus ontológico?

Nos proponemos realizar una exposición de las principales respuestas a estas preguntas, o sea, de las principales corrientes de la Filosofía de las Matemáticas. Tras la exposición, faltaría la crítica sistemática; pero creemos que esto sobrepasa los objetivos del trabajo. Hemos tratado de dividir cada apartado en tres puntos: (1) para las ideas principales de la corriente; (2) para los defensores de la postura, así como su desarrollo histórico; (3) para los problemas que habitualmente se esgrimen en su contra. En ocasiones los puntos se entremezclan inevitablemente.

1.1.1. Platonismo

1. El término “platonismo” nos refiere ya las ideas generales de esta corriente. Dentro de la Filosofía de las Matemáticas fue introducido por el lógico Paul Bernays en 1924, y no sin acierto.

Como es sabido, Platón, por contraposición al mundo sensible, postuló la existencia del Mundo de las Ideas. Los etéreos habitantes de este mundo metafísico no son los objetos físicos, con sus incómodas imperfecciones, sino las *Ideas* o *Formas*, ciertas entidades eternas e inmutables que representan la esencia de los objetos sensibles y cuya manera de captación humana es, estrictamente, la razón (por contraposición a los sentidos para los objetos del mundo físico). Por ejemplo, jamás podremos ni ver ni tocar un triángulo *ideal*, o sea, un verdadero triángulo, sino sólo ciertos objetos accidentales con forma aproximada de triángulo: accidentalmente triangulares. El (único) verdadero triángulo es la Idea de triángulo, cuya existencia es independiente de nosotros. En este sentido Platón considera que las figuras matemáticas y los números son Ideas, accesibles sólo mediante la razón, con existencia real y universal.

Como decíamos, Bernays introduce el término en el campo matemático. El objetivo era caracterizar el modo de pensar en matemáticas donde los objetos de una teoría matemática se toman como elementos de una totalidad *dada*. «Dada», o sea, existente independientemente de nosotros. Vemos el acierto del término.

Creemos conveniente seguir el argumento de José Ferreiros en la distinción de dos géneros de platonismo. En sus propias palabras ([13] Ferreirós Domínguez, 1999):

- Platonismo *interno* o propiamente *matemático*: es característico de las teorías de la matemática abstracta o moderna, donde se hace referencia a elementos cuya existencia se postula y se considera dada (se podría hablar de existencia ideal).
- Platonismo *externo*, ontológico, o propiamente *filosófico* (una de las posibles interpretaciones filosóficas de la matemática, en particular de la característica antes señalada de la matemática abstracta): consiste en la afirmación de que los objetos matemáticos gozan de una existencia real, análoga en algún sentido (aunque diferente) a la existencia de los objetos físicos.

Es habitual referirse al platonismo en matemáticas como la segunda acepción.

El platonismo *interno* es realmente casi imposible de evitar en cualquier teoría y en la práctica matemática: cualquier formulación admite como dados ciertos objetos (basta considerar cualquier uso del cuantificador existencial $\exists$, que asume como dada cierta totalidad). Sin embargo, evita plantarse frente al áspero problema de justificar su existencia real en algún extraño mundo más que en el que puedo imaginar (esto no debe sonar raro: en mi cabeza existe un centauro llamado Kurt Gödel, que claramente no es de carne y hueso, sino sólo como una creación de mi mente, como una imagen o una descripción con mis palabras). Esto nos lleva a otro problema: qué conceptos matematicos, fruto de la mente humana, son válidos, y cuáles no. Así, Hilbert planteó con insistencia que la única limitación debe ser la lógica. Es decir, los objetos matemáticos postulados no pueden dar lugar a teorías contradictorias (deben ser *consistentes*). Con eso es suficiente para garantizar su validez.

Bernays establece ([2] Bernays 1935, 63-64) que hay dos posturas platónicas esenciales en matemáticas con las que debemos tomar partido. La primera, la más débil, es asumir la existencia objetiva de la totalidad de los número naturales ($\mathbb{N}$). La segunda es admitir las nociones de conjuntos y funciones arbitrarias.

Notemos que la práctica matemática actual asume estas posturas dentro de lo que hemos llamado platonismo *interno*. Su planteamiento de la Matemática es el de una ciencia abstracta que investiga las relaciones entre elementos y estructuras que se dan por existentes. Y la noción de existencia no es otra que su admisibilidad (libre de contradicciones) en el pensamiento humano. No obstante, como resaltaremos luego, la metodología matemática actual se debe a la postura formalista.

En cuanto al platonismo *externo*, es la idea puramente platónica. Postula la existencia de una realidad matemática externa, accesible solamente a través de la razón, poblada por las proposiciones, que son estricatamente reales y verdaderas. Un matemático platónico piensa que está *descubriendo* verdades que ya están *ahí*, cual Magallanes llegando al Estrecho de Todos los Santos.

2. Como hemos comentado, Platón fue un platónico. Pero no fue el primero, sino que fue Pitágoras. La cosmovisión pitagórica establecía la esencia del mundo en los números: lo único real y eterno que ordenaba el mundo. Aunque Aristóteles, discípulo de Platón, rechazó la bifurcación de la realidad de su maestro, admitía la existencia, esta vez en los propios objetos, de esencias o *universales*. Las ideas platónicas y aristotélicas se mantienen durante siglos e influyen en San Agustín, que, en *La Ciudad de Dios*, asegura que la totalidad de los número naturales existe en acto en el inteleto divino, pues, ¿quién sería tan necio para decir que Dios deja de contar en un cierto número n, por grande que sea? La escolástica mediaeval hereda estas ideas.

Actualmente se suele decir que el platonismo es la fe de la mayor parte de los matemáticos, pero que es una religión que ejercitan en privado ([5] Davis & Hersh, 1982, 247). Otro dicho habitual es que los matemáticos son platónicos los días que trabajan pero formalistas los días de fiesta, como decía Dieudonné:

> En lo relativo a fundamentos, creemos en la realidad de la matemática, pero, por supuesto, cuando los filósofos nos atacan con sus paradojas corremos a escondernos tras el formalismo diciendo: «La matemática es sólo una combinación de símbolos carentes de significado», y entonces sacamos los capítulos 1 y 2 [de los Eléments de mathématiques] sobre teoría de conjuntos. Finalmente nos dejan en paz y podemos volver a nuestra matemática, haciéndola como siempre, trabajando en algo real. ([7] Dieudonné 1970)

Efectivamente, Bourbaki era un platónico convencido que veía los objetos matemáticos tan reales como los objetos físicos.

Otro gran matemático platónico fue Georg Cantor, para quien los conjuntos y los números transfinitos, que él había desarrollado, existían en un doble sentido. Primero, en un reino transuránico que llama *intellectu Divino*, y segundo, en la misma naturaleza. Así lo expresa en esta carta a Hermite el 30 de noviembre de 1895:

> Dice usted [Hermite] muy bellamente en su carta del 27 de Nov.: «Los números (enteros) me parecen constituir un mundo de realidades que existen más allá de nosotros con el mismo carácter de absoluta necesidad que las realidades de la naturaleza, cuyo conocimiento nos es dado por los sentidos, etc.» Permítame, sin embargo, el comentario de que en mi opinión la realidad y absoluta legalidad de los números enteros es mucho mayor que la del mundo sensorial. Y el que así sea, tiene una única y muy simple razón, a saber, que los números enteros existen en el grado sumo de realidad, tanto separados como en su totalidad actualmente infinita, en la forma de ideas eternas in intellectu Divino. ([25] Meschkowski 1983, 275)

Para Cantor la existencia del infinito actual no era ningún problema. De hecho, como recalca Javier De Lorenzo ([6] De Lorenzo, 1979), en su teoría "lo primero es el infinito actual escindido en una escala de infinitos y, después, se constituye lo finito, como mero rincón de lo infinito". Así, la Teoría de Conjuntos era perfectamente real, y Cantor solo hacía de escribano ante las inspiraciones, podríamos decir divinas, que «le venían». Es evidente que la filosofía platónica de Cantor es indispensable para entender su incansable desarrollo matemático, cuyas teorías fueron siempre puestas en cuestión *filosóficamente*.

Kurt Gödel, el protagonista de este trabajo, tomó fuertes posiciones platónicas en torno a las matemáticas. En ([18] Gödel, 1944a), Gödel analiza la filosofía de Bertrand Russell, en particular el *axioma de reducibilidad* que había propuesto para evitar las definiciones impredicativas (y así las paradojas). En este texto el platonismo gödeliano es absoluto (y en sintonía con el platonismo de compromiso de Quine):

> Sin embargo, también pueden concebirse las clases y los conceptos como objetos reales, a saber, como «pluralidades de cosas» o como estructuras que consistan en una pluralidad de cosas, y los conceptos como las propiedades y las relaciones de las cosas que existen independientemente de nuestras definiciones y construcciones. Me parece que la aceptación de tales objetos es tan legítima como la aceptación de los cuerpos físicos, y que hay tantas razones para creer en la existencia de aquellos como en la de éstos. Son necesarios para obtener un sistema de matemáticas satisfactorio en el mismo sentido en el que los cuerpos físicos lo son para una teoría satisfactoria de nuestras percepciones sensibles, y en ambos casos es imposible interpretar los enunciados acerca de estas entidades como enunciados acerca de «datos», es decir, en el último caso acerca de las percepciones sensibles. (325-326)

En el mismo texto Gödel comenta que la idea de Russell de comparar los axiomas matemáticos con las leyes de la naturaleza le parece del todo acertada. La idea es que la correspondencia evidencia matemática-axiomas de la matemática se asemeja a la correspondencia percepción sensorial-leyes físicas.

> Russell extendió también en otro aspecto (en un de sus primero escritos) la analogía entre las matemáticas y una ciencia natural. Compara los axiomas de la lógica y las matemáticas con las leyes de la naturaleza, y la evidencia lógica con la percepción sensible, de modo que los axiomas no tienen por qué ser necesariamente evidentes por sí mismos, sino que su justificación estriba (como en la física) en el hecho de que permiten que estas «percepciones sensibles» sean deducidas; esto no excluiría, por supuesto, que tuviesen también una suerte de plausibilidad intrínseca similar a la que se da en física. Creo que (en el supuesto de que «evidencia» se entienda de un modo suficientemente estricto) este punto de vista ha sido ampliamente justificado por posteriores desarrollos y se puede esperar que aún lo sea más en el futuro. (316)

Y posteriormente, criticando la teoría russelliana de la inexistencia de clases, dice:

> Todo esto sólo es una verificación del punto de vista antes defendido de que la lógica y las matemáticas (del mismo modo que la física) están construidas con axiomas que tienen un contenido real que no puede ser eludido. (331-332)

Al igual que en el caso de Cantor, parece claro que la “militancia” platónica de Gödel tuvo algo que ver en el desarrollo de los teoremas de Incompletitud. Rebecca Goldstein, en ([19] Goldstein, 2005, 44), tras una exposición de la vida de Gödel, dice:

> Fue precisamente la osada ambición de Gödel de llegar a una conclusión matemática que fuese al mismo tiempo un resultado metamatemático que refrendase el realismo matemático, la que propició sus teoremas de incompletitud. (...) Su orientación filosófica no era expresión de su labor matemática, sino al contrario: su labor matemática era expresión de su orientación filosófica, de su platonismo, que por tanto era la más profunda expresión del hombre propiamente dicho.

Actualmente, el físico matemático Roger Penrose se declara platónico y escribe:

> No oculto mis fuertes simpatías por el punto de vista platónico de que la verdad matemática es absoluta, externa y eterna, y no se basa en criterios hechos por el hombre; y que los objetos tienen una existencia intemporal por sí mismos, independientemente de la sociedad humana o de objetos físicos particulares. ([29] Penrose 2006, 186)

A continuación exhibe uno de los argumentos clásicos: ¿cómo explicar el hecho de que distintas personas puedan alcanzar un acuerdo tan perfecto en torno a las matemáticas, un saber objetivo sin objetos materiales? Penrose responde que la comunicación es posible porque cada matemático, a través del intelecto, entra en contacto con el mismo mundo platónico donde verdaderamente existen los objetos matemáticos. En este mismo libro (un *best seller*, por cierto, en divulgación científica) Penrose expone una filosofía de la mente que merece ser comentada en este trabajo. Penrose considera el hecho de que los matemáticos humanos pueden demostrar los resultados no probables *formalmente* que predice el teorema de Gödel. Así pues, argumenta que esta disparidad sistema formal vs humano significa que los matemáticos humanos no se pueden describir como sistemas de prueba formales y, por lo tanto, en sus procesos mentales están ejecutando un algoritmo no computable. Es decir, los procesos físicos de la mente no se rigen por las leyes físicas de carácter algorítmico computable. En otras palabras: el pensamiento humano no puede reducirse a procesos mecánicos como los de un ordenador. Así, los teoremas de Gödel, esgrimidos no pocas veces como el argumento definitivo de los límites del conocimiento humano, significarían sólo los límites de nuestros modelos computacionales. Penrose propone el colapso de la función de onda en mecánica cuántica como un proceso no computable candidato a explicar el funcionamiento de la mente humana. Dejemos este tema aquí, no sin antes añadir que la teoría de Penrose y su argumento gödeliano ha recibido duras y numerosas críticas.

También Penélope Maddy esgrime argumentos platónicos actualmente en su teoría del realismo conjuntista, en línea con el naturalismo de Willard V. Quine. Quine propuso la siguiente idea: nuestra mejores teoría físicas están totalmente comprometidas con la aceptación de entidades teóricas, como los electrones, y con la aceptación de entidades abstractas, como los números o las funciones. Así, si creemos en la existencia de las primeras, estamos obligados a aceptar la existencia de las segundas (para Quine no hay distinción entre la física y la matemática).

3. Los problemas del platonismo (al menos en su vertiente *externa*) saltan a la vista: resulta realmente misterioso que entremos en contacto con un mundo verdadero, independientemente del mundo sensible, que intuimos con el intelecto. Un problema es, como señala Carlos M. Madrid Casado ([23] Madrid Casado, 2009), que el platonismo *sobrepuebla* los cielos. Puede llegar a ser atractivo pensar que ciertas entidades matemáticas sencillas, como la Idea de triángulo, o la de números naturales, tengan existencia real. Sin embargo, a medida que ha avanzado el desarrollo matemático, la existencia de objetos matemáticos tan complejos y abstractos como los que actualmente se manejan complica el tema. ¿Existen realmente todaslas complejas estructuras abstractas usadas en la demostración del Último Teorema de Fermat? ¿Existe la distribución de los ceros de la función zeta de Riemann en el fondo del Mundo de las Ideas y aún no la hemos alcanzado? Parece más natural concebir estos objetos abstractos como existentes sólo en la mente humana, como constructos válidos para ciertas cosas.

En cuanto al argumento de Penrose, no parece excesivamente difícil de contestar. Quizás sería un error hacerlo argullendo que los seres humanos se comunican sobre muchos otros temas simplemente a través del lenguaje, sin recurrir a cielos platónicos, pues es cierto que el lenguaje natural, a diferencia del matemático, refiere, habitualmente, o a objetos reales (corpóreos), o a objetos imaginarios derivados de ellos. Pero la matemática, realmente, no se refiere a nada. Ahora bien, teniendo en cuenta que todos los seres hurmanos somos seres biológicos que compartimos la capacidad del lenguaje, no parece descabellado pensar que podamos comunicarnos sobre cualquier cosa que construyamos mentalmente, incluso sobre aquello que no tiene referenciales materiales. Y además, en cierto modo, las matemáticas puede que sí refieran a algo objetivo: a sus símbolos tipográficos.

En 1973 Paul Benacerraf planteó el siguiente dilema: Asumimos estos dos asertos: (A) Para que las proposiciones matemáticas sean verdaderas, los objetos con los que tratan y a los que se refieren deben existir necesariamente; (B) Las humanos sólo podemos conocer un objeto si interactuamos causalmente con él.

Y de aquí se deduce la siguiente paradoja: Si las matemáticas son universales, entonces, por (A), los objetos matemáticos no son contingentes. Sin embargo, como no absorven ni emiten energía, no son espacio-temporales, luego, por (B), no podemos conocerlos. De otra manera, si aceptamos que podemos conocer los objetos matemáticos, entonces, por (B), deben ser contingentes. Es decir: si la matemática es universal, no la podemos conocer, y si la podemos conocer, entonces no es universal.

El platonismo no escapa de las garras del *dilema de Benacerraf*: La matemática es universal, pues los objetos matemáticos no son contingentes sino necesarios; pero, por ser necesarios, eternos e inmutables, no son espacio-temporales, de lo que se deduce, por (B), que los humanos no podemos conocerlos. Pero los seres humanos tenemos conocimientos matemáticos, luego el platonismo es falso.

1.1.2. Logicismo

1. Las premisas fundamental logicistas son dos: (1), que la Matemática es reducible a la Lógica; (2) que los entes lógicos tienen existencia real al margen de los sujetos pensantes.

Comparte pues, con el platonismo, una sustantivización de la *Forma* (término que tomamos prestado de la gnoseología de Gustavo Bueno, véase [[4] Bueno 1995b]), es decir, de la parte del cuerpo de las matemáticas que explica su unidad (y por consiguiente, su diferenciación de las otras ciencias), esto es, la estructura y los modelos relativos a las matemáticas (para el logicismo, los entes lógicos y sus relaciones). Así, el logicismo revela la esencia de las matemáticas como una adecuación entre los matemáticos de carne y hueso (existentes) y los entes lógicos (existentes), accesibles sólo mediante el intelecto.

2. Si decíamos que Gottlob Frege fue el creador (o descubridor) de la lógica moderna, ahora añadimos que también lo fue del programa logicista. En concreto,

en los *Fundamentos de la Artimética* (1884), donde expone sus tesis logicistas. Frege mantiene que la aritmética se subordina esencialmente a la Lógica: «la aritmética no sería más que una Lógica más desarrollada, todo teorema aritmético sería una ley lógica, aunque derivada» ([8] Dou, 1974, 62). Así, del realismo lógico deduce un realismo matemático: «Calcular es deducir» ([15] Frege 1, 128).

Los entes lógicos, dice, son las leyes de las leyes de la Naturaleza. Notemos la relación entre esta consideración y las proposiciones platónicas de Gödel en referencia a Bertrand Russell (del que pronto hablaremos). El pensamiento de Frege supone una reacción contra las teorías matemáticas del gran filósofo idealista Imanuel Kant, muy influyente en la época, sobre la calidad sintética (el predicado no está incluido en el sujeto) de los juicios Aritméticos y Geométricos. Para Frege, como estos juicios se reducen a juicios lógicos, deben ser analíticos (el predicado no añade nada nuevo al sujeto).

Las obras de Frege, donde trató, por primera vez, de reducir la aritmética a un sistema formal, utilizaban una lógica basada en ciertos principios lógicos. Uno de ellos era el *principio de comprehensión*, la *Ley V* de los dos monumentales volúmenes de los *Grundgesetze der Arithmetik* (*Leyes Básicas de la Aritmética*, 1893). Esta ley se refiere a la posibilidad de definir una clase como los elementos que cumplen cierta propiedad (a cada concepto es posible asignarle una extensión). Debemos cuidar los términos: para Frege toda propiedad define un conjunto, y todo conjunto está definido por una propiedad. En esta lógica los conjuntos se toman como elementos; esto es, el sistema habla de conjuntos. El lector ya intuirá la vulnerabilidad de esta tesis frente a la *paradoja de Russell*. Russell escribe a Frege el 16 de junio de 1902, cuya carta comenzaba así:

> Querido Colega:
>
> He sabido de tu Grundgesetze desde hace un año y medio, pero sólo ha sido ahora que he podido encontrar tiempo para el detallado estudio que pretendo dedicar a tus escritos. Estoy totalmente de acuerdo contigo en todos sus puntos principales, en particular en tu rechazo de todo elemento psicologicista en lógica, y en el valor que asignas a una notación conceptual para la los fundamentos de las matemáticas y de la lógica formal que, de paso, a duras penas pueden distinguirse. En muchos detalles, encuentro en las discusiones, distinciones y definiciones de tus escritos todo lo que uno busca en vano en otros lógicos. En particular, en lo que respecta a las funciones (sección 9 de tu Notación Conceptual) he llegado independientemente a las mismas conclusiones incluso en detalle. He encontrado una dificultad tan sólo en un punto. Afirmas (p. 17) que una función puede también constituir el elemento indeterminado. Esto es lo que solía creer yo, pero este punto de vista me parece ahora dudoso debido a la siguiente contradicción: Sea w el predicado de ser un predicado que no puede ser predicado de sí mismo. ¿Puede w ser un predicado de sí mismo? De ambas respuestas se sigue una contradicción. Debemos por tanto concluir que w no es un predicado. Igualmente, no hay una clase

> (en su totalidad) de todas las clases que, en su totalidad, no sean miembros de sí mismas. De esto concluyo que bajo ciertas circunstancias un conjunto definible no forma un conjunto completo.

La contradicción es la siguiente: consideremos la clase de Rusell $R = \{x : x \notin x\}$, donde x es un conjunto. Notemos que la definición es intensional, o sea, para Frege, R también es un conjunto. En palabras, R es el conjunto de los conjuntos que no se contienen a sí mismos. Russell se pregunta: ¿$R \in R$? Si $R \in R$, entonces, por definición, es un conjunto que cumple que no pertenece a sí mismo, esto es, $R \notin R$. Recíprocamente, si $R \notin R$, entonces es un conjunto con las propiedades necesarias para estar en R: se deduce que $R \in R$. Tenemos que $R \in R \leftrightarrow R \notin R$. Con esto el sistema fregeano se desmoronaba. Frege, con una honestidad intelectual admirable, paralizó el tercer tomo de su obra tras años de duro trabajo. Su teoría adolecía del peor error posible de un sistema formal: era contradictoria, pues se puede probar que $R \in R$ y que $R \notin R$. Y de una contradicción se sigue que *todo* es un teorema.

El continuador del programa logicista de Frege iba a ser, *paradójicamente*, Russell. En los años posteriores se investigó la razón de las paradojas y antinomias. La cuestión parecía radicar, como señaló Henri Poincaré, en la autorreferencia, en particular en el uso de las *definiciones impredicativas*, aquellas en las que el término que se desea definir forma parte del conjunto que se usa para la definición. Por ejemplo, la clase de Rusell $R = \{x : x \notin x\}$ está definida por una relación entre el objeto a ser definido (R), con todos los objetos de un cierto tipo (conjuntos) de los cuales el objeto a ser definido es, él mismo, parte (R es un conjunto). Es el *principio del círculo vicioso* de Russell, que Gödel explicaba así (para luego criticarlo, pues el tema es más peliagudo de lo que parece):

> La falacia, según se sostiene, consiste en la circunstancia de que se definen (o se asumen tácitamente) totalidades cuya existencia implica la existencia de ciertos nuevos elementos de la misma totalidad, a saber, elementos definibles únicamente en términos de la totalidad entera. Esto lleva a la formulación de un principio que dice que ninguna totalidad puede contener miembros definibles únicamente en términos de la totalidad, o miembros que involucran o presuponen esta totalidad (principio del círculo vicioso). [[18] Gödel, 1944a, 322]

Notemos que estas paradojas no son exlusivas del lenguaje formal, sino que en neustro lenguaje natural habitual también se producen paradojas semánticas debido a la autorreferencia. Por ejemplo la paradoja del mentiroso[3] o la antinomia de Richard, que Gödel cita como argumentos análogos a su demostración ([18] Gödel 1944b, 56).

[3] «Un hombre afirma que está mintiendo. ¿Lo que dice es verdadero o falso?» O en la versión más simple: «Esta oración es falsa». Si esta oración falsa, entonces es verdadera, y si es verdadera, entonces es falsa.

Rusell y su colaborador Alfred North Whitehead llevaron a cabo un esfuerzo titánico para idear un nuevo sistema formal que expresase las verdades aritméticas y evitase todas las paradojas y antinomias que se venían detectando. Es el sistema de *Principia Mathematica*, que publicaron en varios tomos (y dejaron sin concluir por agotamiento) entre 1910 y 1913. Para que nos hagamos una idea, nada más y nada menos que 700 páginas les fueron necesarias para poder demostrar que $1+1=2$. En ese momento, con magistral ironía, comentan: «*The above proposition is occasionally useful*» («esta proposición es ocasinalmente útil»).

En su obra exponen una Teoría de Tipos, donde se exige que para que $X \in Y$ sea una fórmula bien formada, X debe ser un objeto de un «tipo» inmediatamente inferior al objeto Y. Podría explicarse de la siguiente manera: los individuos son de tipo I, los conjuntos de individuos son de tipo II, los conjuntos de conjuntos son de tipo III, etc. Con esta formulación se evita la pardoja de Russell (y el resto de paradojas de autorreferencia); se convierte, de hecho, en una fórmula sin sentido que no cabe si quiera plantear. R es del mismo tipo que R, luego $R \in R$ no es una fórmula bien fromada, y no tiene sentido preguntarse si es cierta o no.

Aunque los *Principia* hablan por sí mismos de la filosofía de sus autores, con esta cita de Russell ([33] Russell 1920, 169) queda bien claro: «*Logic, I should maintain, must no more admit a unicorn than zoology can; for logic is concerned with the real world just as truly as zoology, though with its more abstract and general features*» («La lógica, debo mantener, no debe admitir un unicornio más que lo que puede hacerlo la zoología; la lógica trata del mundo real de igual manera que la zoología, aunque con rasgos más abstractos y generales»). No obstante esta cita se suprimió de las ediciones posteriores, y las posiciones logicistas de Russell se relajaron con el tiempo.

A finales del siglo XIX y principios del XX, la ciudad de Viena, capital del Imperio Austrohúngaro, era una ciudad con un panorama intelectual y cultural único. Incluso después de la Primera Guerra Mundial y la desintegraciín del Imperio, Viena continuaba siendo un brillante herbidero científico. Es en este contexto en el que Gödel realiza sus estudios en lógica.

El más destacado grupo intelectual (y el que interesa aquí) de la época es, posiblemente, el Círculo de Viena, un conjunto de pensadores que desarrollaron el movimiento conocido como «positivismo lógico» (a veces, «empirismo lógico» o «empirismo radical») dirigido por el filósofo Moritz Schlick. Algunos de sus selectos miembros (nunca llegaron a aceptar a Karl Popper) más ilustres fueron Rudolf Carnap (alumno de Frege), Hans Hahn, Otto Neurath, Olga Hahn-Neurath, Herbert Feigl y un joven Kurt Gödel. El Círculo de Viena tenía como objetivo purgar la ciencia de metodologías e ideas metafísicas. Para ello, tomaba la teoría empirista de David Hume pero rectificándola sustancialmente. En palabras de Rebecca Goldstein:

> Los positivistas lógicos transformaron la teoría empírica del conocimiento en

> una teoría del significado. Según ésta, los recursos empíricos que sirven para determinar si una proposición concreta es verdadera también nos brindan el significado mismo de la proposición. (...) los límites de la cognoscibilidad empírica coinciden con los de la significatividad. Si no se puede concebir un conjunto de experiencias posibles que corroboren una posición dada, es que tal proposición lo es tan sólo en apariencia, está vacía de significado y constituye lo que los positivistas bautizaron como «pseudoproposición». ([19] Goldstein, 2005, 71)

Nótense las similitudes con la teoría de Russell: lo que no significa nada (lo que es contradictorio), no es válido (al fin y al cabo la solución de Russell para las antinomias es no permitir lo que no es permisible).

Las ideas logicistas en torno a la matemática del Círculo de Viena quedan patentes en los planteamientos de Carnap y Hahn: las porposiciones matemáticas se reducen a proposiciones lógicas, y las proposiciones lógicas son análogas a las tautologías; carecen totalmente de sentido descriptivo. Las matemáticas son puramente sintácticas (analíticas, si se quiere), luego su verdad proviene de las reglas de los sitemas formales (aquí está la sustantivización de la lógica). El planteamiento es opuesto al platónico, donde el concepto de *verdad matemática* es similar al concepto de verdad que utilizamos habitualmente.

A estas alturas se estará preguntando el lector qué narices hacía un platónico total como Gödel en el Círculo de Viena (Gödel tenía plantamientos realistas respecto a la matemática desde 1925, y acudió a las reuniones del Círculo de Viena entre 1926 y 1928). A todas luces era un infiltrado, un disidente silencioso que jamás reveló sus opiniones filosóficas. Varios de los miembros del Círculo hablarían de él, posteriormente, en términos similares a los que hace Fiegl: «un hombre aplicado, de lo más sencillo y diligente, pero con una mente a todas luces genial». Karl Menger (hijo de Carl Menger, el ilustre economista, padre de la Escuela Austríaca), que acudía habitualmente a las reuniones del Círculo, refiere: «Jamás oí a Gödel hablar en esas reuniones ni intervenir en los debates; pero manifestaba su interés mediante leves movimientos de cabeza que indicaban conformidad, escepticismo o discrepancia».

Resaltemos la discrepancia fundamental del logicismo del Círculo de Viena con el platonismo de Gödel: para Gödel las proposiciones matemáticas son descriptivas, no así empíricas; la concepción de las matemáticas como «carente de sentido» le era inaceptable. Gödel no entro a debatir dialécticamente con los logicistas, más bien entró, directamente, a ganar; los teoremas de Incompletitud, consideraba, eran una réplica casi definitiva. Así lo expresa en una carta que iba a enviar al lógico chino Hao Wang (tuvieron una rica correspondencia) en 1971, pero que se quedó en el cajón:

> Si bien es verdad que el interés en el fundamento de las matemáticas me lo despertó el «Círculo de Viena», las consecuencias filosóficas de mis resultados,

> así como los principios heurísticos que me condujeron a ellos, son todo menos positivistas o empiristas. (...)
>
> Soy un realista conceptual y matemático desde 1925 más o menos. Jamás he sostenido que las matemáticas sean una sintaxis. Al contrario: es precisamente esa idea, concebida en un sentido razonable, la que mis resultados rebaten.

Por último, queremos tratar el logicismo del filósofo Ludwig Wittgenstein.

Wittgenstein dejó la carrera de ingeniería aeronáutica para estudiar con Russell en Inglaterra, con quien mantendría una intensa relación profesional. Russell admiraba a Wittgenstein, a quien describió como «...tal vez el ejemplo más logrado que yo haya conocido de lo que tradicionalmente se tiene por un genio: apasionado, profundo, intenso y dominante». En una carta a su amante Ottoline Morrell expresa que Wittgenstein, que por entonces aún era estudiante, llegaba a exasperarlo y le hacía dudar de todo:

> Estábamos los dos sulfurados. Le he mostrado una parte crucial de lo que estaba escribiendo y me ha dicho que era un error de cabo a rabo, que no tenía en cuenta las complicaciones que conllevaba; que él ha puesto a prueba mis tesis y ha visto que no valen. Yo no acertaba a entender sus objeciones -apenas podía hablar de lo acalorado que estaba-, pero me ha dado el pálpito de que tenía razón, de que se me había pasado por alto alguna cosa. Si yo también pudiese ver de qué se trata, no me importaría, pero la situación, tal y como está, me preocupa, y ya no me causa ningún placer escribir. Lo único que puedo hacer es seguir adelante con lo que veo, pero me da la sensación de que seguramente todo está equivocado y de que Wittgenstein va a pensar que soy un bribón embustero por seguir en mis trece.

El logicismo de Wittgenstein se encuentra en su única obra publicada en vida, el *Tractatus logico-philosophicus* (1921), terminado de escribir en las trincheras de la Primera Guerra Mundial, que fue prologado (con poco éxito para Wittgenstein) por Russell. Como el lector quizás sepa, al tratar su pensamiento se habla del primer Wittgenstein, el del *Tractatus*, y el segundo Wittgenstein, el de las *Investigaciones filosóficas*, pues sus posiciones filosóficas cambiaron radicalmente en un momento de su vida. Pues bien, Wittgenstein I era casi una deidad, un verdadero ídolo al que invocar para zanjar cualquier discusión, en el Círculo de Viena (al que nunca quiso acudir y a muchos de cuyos miembros despreció). Los positivistas lógicos interpretaron el *Tractatus* como la definitiva y pulida codificación de sus pensamientos. Wittgenstein I defendía, efectivamente, que las verdades lógicas son tautologías:

> 6.1 Las proposiciones de la lógica son tautologías ([35] Wittgenstein, 1921, 126)

> 6.11 Las proposiciones de la lógica, pues, no dicen nada. (Son las proposiciones analíticas). [126]

Aquí vemos que el logicismo puro de Wittgenstein se separa de las ideas del Círculo de Viena:

> 5.61 La lógica llena el mundo; los límites del mundo son también sus límites. No podemos, por consiguiente, decir en lógica: en el mundo hay esto, y esto, aquello no. En efecto, esto presupondría, aparentemente, que excluimos ciertas posibilidades; y ello no puede ser el caso, porque, de otro modo, la lógica tendría que rebasar los límites del mundo: si es que, efectivamente, pudiera contemplar tales límites también desde el otro lado. Lo que no podemos pensar no lo podemos pensar; así pues, tampoco podemos *decir* lo que no podemos pensar. ([35] Wittgenstein, 1921, 123)

> 6.113 Que a la sola luz del símbolo pueda reconocerse que son verdaderas, es característica peculiar de las proposiciones lógicas, y este hecho encierra en sí toda la filosofía de la lógica. Y del mismo modo, que *no* pueda reconocerse en la sola proposición la verdad o flasedad de las proposiciones no lógicas, es también uno de los hecho más importantes. (127)

> 6.124 Las proposiciones lógicas describen el armazón del mundo o, más bien, lo representan. No «tratan» de nada. Presuponen que los nombres tienen significado, y las proposiciones elementales sentido; y ésta es su conexión con el mundo. Está claro que algo tiene que indicar sobre el mundo el hecho de que ciertas conexiones de símbolos -que tienen esencialmente un carácter determinado- sean tautologías. Aquí radica lo decisivo. Decíamos que hay algo de arbitrario en los símbolos que usamos y algo hay que no lo es. En la lógica sólo ésto se expresa: Pero quiere decir que en la lógica no expresamos *nosotros* lo que queremos con ayuda de los signos, sino que en la lógica es la propia naturaleza de los signos naturalmente necesarios lo que se expresa. Si conocemos la sintaxis lógica de un lenguaje sígnico cualquiera, entonces ya están dadas todas las proposiciones de la lógica. (131-132)

> 6.2321 Y que las proposiciones de la matemática puedan ser probadas no quiere decir otra cosa sino que su corrección puede ser percibida sin necesidad de que lo que expresan sea ello mismo comparado, en orden a su corrección, con los hechos. (135)

En contra de Russell, no considera la Matemática como una parte de La Lógica, sino como un método suyo:

> 6.2 La matemática es un método lógico. Las proposiciones de la matemática son ecuaciones, es decir, pseudo-proposiciones. ([35] Wittgenstein, 1921, 134)

Wittgenstein jamás fue un positivista lógico. Según el propio Wittgenstein, la parte más importante del *Tractatus* no es lo escrito, sino lo no escrito: lo que él llama lo *ético* o lo *místico*. Para un positivista lógico, fuera de los límites de lo decible no hay absolutamente nada. Pero para Wittgenstein está aquello «de lo que no se puede hablar». Esta es, podríamos decir, la incompletitud wittgensteniana, muy distinta a la de Gödel (pero en ello coinciden contra el positivismo lógico: de alguna manera, el hombre no es la medida de todas las cosas). Lo místico es inexpresable:

> 6.522 Lo inexpresable, ciertamente, existe. Se *muestra*, es lo místico. ([35] Wittgenstein, 1921, 144)

> 6.53. El método correcto de la filosofía sería propieamente éste: no decir nada más que lo que se puede decir, o sea, proposiciones de la ciencia natural -o sea, algo que nada tiene que ver con la filosofía-, y entonces, cuantas veces alguien quisiera decir algo metafísico, probarle que en sus proposiciones no había dado significado a ciertos signos. Este método le resultaría insatisfactorio -no tendría el sentimiento de que le enseñábamos filosofía-, pero sería el único estrictamente correcto. (144-145)

Dejamos este tema aquí, no sin antes escribir la célebre última proposición del *Tractatus*:

> 7. De lo que no se puede hablar hay que callar. ([35] Wittgenstein, 1921, 145)

3. Para analizar los problemas del logicismo, plantearemos las críticas a los *Principia* y a Wittgenstein.

Primero, aunque, formalmente, los *Principia* son correctos, la Teoría de Tipos de Russell y Whitehead parece sacada de la chistera. No se ofrece ninguna explicación de por qué algunos conjuntos están permitidos y otros no. Además, se parte de unos axiomas que han sido cuestionados. En concreto el *axioma de reducibilidad* en el cálculo de funciones proposicionales. Por función proposicional consideran toda expresión que contenga una variable x tal que, cuando x tome un valor determinado, se convierta en una proposición. El axioma, entonces, dice así: cualquier función proposicional (sea del Tipo que sea), es equivalente, extensionalmente, a alguna función proposicional de un Tipo inferior. Salta a la vista que está postulado, *ad hoc*, para evitar las paradojas. Así lo reconoce Rusell:

> No veo ninguna razón para creer que el axioma de reducibilidad sea lógicamente necesario, que es lo que significaría decir que es cierto en todos los mundos posibles. La admisión de este axioma en un sistema lógico es, por tanto, un defecto ... una suposición dudosa. ([33] 1920, 155)

Y en su «Introducción» de 1927 a la segunda edición de *Principia Mathematica*:

> Un punto en el que la mejora es obviamente deseable es el axioma de reducibilidad (* 12.1.11). Este axioma tiene una justificación puramente pragmática: conduce a los resultados deseados y no a otros. Pero claramente no es el tipo de axioma con el que podemos quedarnos satisfechos.

Los matemáticos no lo aceptaron generalmente, y preferían trabajar con la Teoría de Conjuntos de Zermelo. En este sentido, otro axioma dudoso era el *axioma de infinitud*, que postulaba que «si n es un número cardinal inductivo cualquiera, existe por lo menos una clase de individuos que tiene n elementos». La controversia surge porque, en el sistema de *Principia*, sin este axioma es imposible definir un número. Su única manera de justificarlo fue suponer que en el mundo hay infinitos objetos: si sólo hubiera n cosas en el mundo, la clase de todas las $(n+1)$-uplas sería vacía, y no podrían definir el número $n+1$. Hermann Weyl expresaba sin ironía que los *Principia* ponían a prueba nuestra fe apenas algo menos que los primeros Padres de la Iglesia.

Para Gödel ([18] Gödel, 1944a, 314) los *Principia* eran «[el único sistema donde] se hizo uso completo del nuevo método [, la lógica formal iniciada por Frege,] para derivar gran parte de las matemáticas a partir de muy pocos axiomas y conceptos lógicos». Pero a continuación expresa una crítica a los mismos:

> Es una lástima que esta primera presentación amplia y detallada de una lógica matemática y la derivación de las matemáticas a partir de ella ostente una falta de precisión formal tan grande en sus fundamentos (en *1-*21 de *Principia*) que represente un paso atrás en comparación con Frege. Lo que falta, ante todo, es una presentación exacta de la sintaxis del formalismo. Se omiten las consideraciones sintácticas incluso en casos en que resultan esenciales para la corrección de las deducciones, en particular en conexión con los «signos incompletos».

Segundo, en cuanto a Wittgenstein, son los teoremas de Gödel los que constituyen la prueba de fuego. La mera posibilidad de una demostración como la de Gödel era inadmisible en el sistema de Wittgenstein, incluso en Wittgenstein II. No se puede hablar fuera de un sistema fomral, ¡es imposible! Expresaba que «no hay cálculo que pueda decidir un problema filosófico. Un cálculo no puede aportarnos información sobre los fundamentos de las matemáticas».

Recordemos el temperamento y la confianza de Wittgenstein en su sistema. En la introducción al *Tractatus* llega a decir que:

> La *verdad* de los pensamientos aquí comunicados me parece, en cambio, intocable y definitiva. Soy, pues, de la opinión de haber solucionado definitivamente, en lo esencial, los problemas. ([35] Wittgenstein, 1921, 57)

No nos sorprende pues, que, al contrario que otros filósofos cuyas ideas también eran puestas en cuestión por los teoremas de incompletitud, como Hilbert, el vienés no aceptara los resultados de Gödel. Llegó a calificarlos de «*logishe Kuntstücke*» («truquitos lógicos») y a decir que su labor «no es hablar de la demostración de Gödel, sino soslayarla». Para el filósofo todo lo que es conocimiento es formalizable, incluido el método lógico de la matemática. Para Gödel el conocimiento era expresable, pero no en nuestros sistemas.

Como último comentario, tal y como indica Javier de Lorenzo ([6] De Lorenzo, 1979, p. 17), la gödelización del sistema de la artimética que Gödel lleva a cabo en su demostración supone una identificación entre el sistema formal y la propia aritmética, haciéndolas equivalentes. De esta manera, frente a las aspiraciones del logicismo, la Lógica no es base para la Aritmética.

1.1.3. Intuicionismo

1. En 1984 se publica la última versión del diccionario soviético de filosofía en español, editado por Iván T. Frolov y traducido del ruso por O. Razinkov. Echemos un vistazo a la definición de intuicionismo matemático:

> Orientación en los fundamentos filosóficos de las matemáticas (lo mismo que el logicismo, el formalismo y el efectivismo), que surgió a comienzos del siglo 20 con motivo de la polémica en torno a sus bases teóricas. Según el intuicionismo, el pensamiento matemático exacto se asienta en la intuición racional, que incluye el proceso de la estructuración mental de todos los objetos matemáticos. De acuerdo con el intuicionismo, mediante tal intuición se crean todas las matemáticas, por lo cual los objetos matemáticos no existen al margen de las especulaciones mentales. Para evitar paradojas, la demostración matemática no debe basarse en la rigurosidad lógica, sino en la evidencia intuitiva: es verídica a condición de que se comprenda intuitivamente cada uno de sus grados, comenzando por las premisas de partida y las reglas de razonamiento. Así pues, en definitiva, la intuición debe juzgar también acerca de la aplicabilidad en las demostraciones de unas u otras leyes y reglas lógicas. Sin embargo, el intuicionismo, a diferencia del intuitivismo, no opone la intuición a la lógica. Sólo considera que las matemáticas no pueden basarse en la lógica y desarrolla su comprensión de la lógica como parte de las matemáticas, enfocando los teoremas lógicos como teorías matemáticas de generalidad máxima. ([16] 1984, 234)

En este sentido, la respuesta a: ¿cuál es el estatus ontológico de los objetos matemáticos? Es: el mismo que el de La Cenicienta, en el sentido de que son construcciones mentales, basadas en la lógica, pero nada más. No hay una sustantivización de la lógica.

El núcleo del intuicionismo es, pues, que todo objeto matemático es producto de la mente humana. De esta concepción surge su concepto de verdad matemática, a saber, que la validez (o no) de un objeto matemático depende de la posibilidad (o no) de *construirlo*. La sencillez de estos enunciados esconde consecuencias más que relevantes: por lo pronto, la prueba por reducción al absurdo queda descartada como método válido de demostración, pues la negación de la falsedad de un objeto matemático no significa que sea posible construirlo. En otras palabras, la *Ley del tercero excluido* aristotélica, por la cual cualquier proposición, o se da, o no se da ($A \vee \neg A$), no se acepta. Por esto se considera al intuicionismo una corriente particular dentro del constructivismo matemático. La diferencia es la concepción subjetiva del primero, por la cual las matemáticas surgen de la *intuición*, y son por lo tanto previas a la lógica y al mismo lenguaje. Para explicar esto tendremos que volver a Kant, que influyó notablemente en el matemático constructivista más representativo, el holandés L.E.J. Brouwer.

El *axioma de elección* que asumían las aplaudidas tesis de Zermelo se ve como una locura de cabo a rabo. Intuitivamente, desde luego, da lugar a locuras de gran calibre, como la Paradoja de Banach-Tarski. Como señala Mariano Martínez-Pérez ([24] 1991, 337), estamos ante el milagro de los panes y los peces.

Los matemáticos intuicionistan tampoco aceptan los conjuntos no numerables, tan sólo los numerables (intuidos en el tiempo, como veremos). Leopold Kroenecker, otro de los precursores del intuicionismo, exclamaba que «Dios hizo los naturales; el resto es obra del hombre».

2. En la *Crítica del juicio* ([21] 1790) Kant se ocupa de la Aritmética y la Geometría. La Geometría, dice, se ocupa de conocer las propiedades del *espacio*, formulando juicios sintéticos a priori. Es decir, los juicios geométricos son una forma a priori de la sensibilidad, pues se refieren al espacio que captamos inevitablemente para toda experiencia (son universales); y son sintéticos porque su conocimiento amplía nuestro conocimiento de la realidad. Ahora bien, los conceptos geométricos no los captamos *por negación* de su contradicción (algo necesario, pero no suficiente), sino por su efectiva construcción:

> El que un concepto semejante se halle libre de toda contradicción, es una condición lógica necesaria. Pero ello no basta, ni de lejos, en relación con la realidad objetiva del concepto, es decir, como la posibilidad de un objeto pensado a través del concepto. Así, el concepto de una figura encerrada entre dos rectas no implica contradicción alguna, ya que los conceptos de dos rectas y su cruce no implica la negación de ninguna figura. La posibilidad no descansa en el concepto como tal, sino en la construcción de tal figura en el espacio, es decir, en las condiciones del espacio y de la determinación de este. (268)

Notemos las cadenas euclídeas que, a finales del siglo XVIII, ataban inevitablemente a Kant. Es el espacio euclídeo, y no otro, el que captamos; en términos kantianos, las formas a priori del espacio son las del espacio euclídeo, luego los juicios sintéticos a priori espaciales, esto es, los jucios geométricos, sólo pueden ser relativos al espacio euclídeo: no existe ninguna otra geometría. Como dice Carlos M. Madrid Casado ([23] Madrid Casado, 2009), la matemática se libró del yugo kantiano el 10 de junio de 1854, cuando Riemann leyó su discurso de entrada en la Universidad de Gotinga.

En cuanto a la Aritmética, ¿cuál es nuestra intuición a priori? Kant responde que es el *tiempo*. En este caso se trata de una forma a priori de la sensibilidad interna, no externa como el espacio; es decir, no proviene de un sustrato empírico. Los números (el contar) son el primer juicio aritmético, intuido en el paso del tiempo: 0, 1, 2, 3, ... De nuevo, los juicios aritméticos son sintéticos a priori; sintéticos, pues, en palabras de Kant, «El concepto de 12 no está todavía pensado en modo alguno al pensar yo simplemente dicha unión de 7 y 5»; a priori porque, como el espacio, el tiempo es universal y condición de la posibilidad de toda experiencia.

Tras las geometrías no euclideanas el movimiento intuicionista se acogió al apriorismo del tiempo kantiano, abandonando el espacial; se denomina *neointuicionismo*, aquí es donde se enmarca el programa de Brouwer.

Las ideas kantianas calaron en Henri Poincaré, aunque con rectificaciones fundamentales. Para el genial matemático francés no es posible reducir las matemáticas a lógica formal. Las siguientes ideas las expone en Ciencia e Hipótesis ([31] 1902). Poincaré toma la distinción kantiana de las pruebas matemáticas entre pruebas analíticas y sintéticas. Las primeras consisten en una mera verificación mecánica. En cambio en las segundas necesitamos razonamientos por recurrencia. ¿Puede un razonamiento de este tipo (por ejemplo, el Principio de Inducción) ser analítico? La respuesta de Poincaré es negativa. En estos razonamientos tratamos, en el fondo, con infinitos silogismos, y es imposible probar infinitos silogismos en un número finito de pasos formales: «Un jugador de ajedrez puede hacer combinaciones para cuatro o cinco movimientos futuros; pero, no obstante cuán extraordinario jugador sea, no puede prepararse para más de un número finito de movimientos. Si aplica sus facultades a la aritmética, no podría concebir sus verdades generales solamente a partir de la intuición directa; para probar incluso el teorema más pequeño, debe usar el razonamiento por recurrencia, porque ese es el único instrumento que nos permite pasar de lo finito a lo infinito». Poincaré creía firmamente en la intuición como motor esencial de las matemáticas. En estas bonitas reflexiones (tanto como para ponerlas *in extenso*) queda patente:

> Si asistís a una partida de ajedrez, para comprenderla no os bastará saber las reglas del movimiento de las piezas. Esto os permitirá solamente reconocer que cada jugada ha sido hecha conforme a estas reglas, y esta ventaja tendrá verdaderamente muy poco valor. Es, sin embargo, lo que haría un lector de un libro de matemáticas, si no fuera más que lógico. Comprender la partida es enteramente otra cosa; es saber por qué el jugador avanza tal pieza más bien que tal otra que habría podido mover sin violar las reglas del juego. Es advertir la razón íntima que hace de esta serie de jugadas sucesivas una especie de todo organizado. Con mayor razón, esta facultad es necesaria al jugador mismo, es decir, al inventor.
>
> (...) Por ejemplo, veamos lo que ha ocurrido con la idea de función continua. Al principio no era más que una imagen sensible, por ejemplo, la de un trazo continuo descrito con tiza en un pizarrón. Después se ha depurado poco a poco; pronto se ha utilizado para construir un complicado sistema de desigualdades, que reproduciría, por decirlo así, todas las líneas de la imagen primitiva; cuando esta construcción estuvo terminada, se ha descimbrado, por

> decirlo así, se ha desechado esta representación grosera que momentáneamente le había servido de apoyo y que sería inútil en adelante; no ha quedado más que la construcción misma, irreprochable ante los ojos del lógico. Sin embargo, si la imagen primitiva hubiera desaparecido totalmente de nuestro recuerdo, ¿cómo adivinaríamos por qué capricho se han dispuesto estas desigualdades de esa manera, unas sobre otras?
>
> (...) Así es como las antiguas nociones intuitivas de nuestros antepasados, aun cuando las hayamos abandonado, todavía imprimen su forma a los andamiajes lógicos que hemos colocado en su lugar.
>
> Esta vista de conjunto es necesaria al inventor; es necesaria igualmente a aquél que quiere realmente comprender al inventor. ¿Puede dárnosla la lógica?
>
> No, el nombre que le dan los matemáticos bastaría para probarlo. En matemáticas, la lógica se llama análisis, y análisis significa división, disección. No puede tener, pues, otra herramienta que el escalpelo y el microscopio.
>
> De este modo, la lógica y la intuición tienen cada una un papel necesario. Ambas son indispensables. La lógica, que puede dar por sí misma la certeza, es el instrumento de la demostración; la intuición es el instrumento de la invención.

Brouwer llevará a cabo, influido por Kant y Poincaré, un ambicioso programa: quiere reconstruir efectivamente la Matemática desde la perspectiva intuicionista. En 1918 publicó *Fundamentación de la Teoría de Conjuntos independientemente del principio del Tercio Excluso*. Brouwer era un genial topólogo, un gran matemático, y comenzó a ganar influencia. Por ello Hilbert llegó a expulsarlo del consejo de redacción de los famosos *Mathematische Annalen* en 1929. Se granjeó grandes discípulos como Hermann Weyl y Arend Heyting, que continuaron el programa intuicionista. En 1918 Weyl publica *El continuo*, una fundamentación del Análisis desde bases intuicionistas. Heyting se ocupó del Álgebra y de la Aritmética, desarrollando los conocidos como *álgebra de Heyting* y *aritmética de Heyting*, que eran, respectivamente, una axiomatización intuicionista del álgebra de Boole y de la Aritmética.

Los intuicionistas también plantearon una Teoría de Conjuntos y una Lógica. La Teoría de Conjuntos, desarrollada sobre todo por Brouwer, se denominó Teoría de Especies. Las especies son los conjuntos de los intuicionistas, con la gran diferencia de que no es aceptable considerar una colección infinita de números, al menos en acto. En cuanto a la Lógica Intuicionista, rechazando la demostración por reducción al absurdo y el principio del Tercero Excluido, se basan en modelos resolutivos de Kolmogorov.

3. Como comentábamos en la Introducción, antes del descubrimiento de las geometrías no euclídeas, la piedra angular de las matemáticas, en sentido amplio, era la geometría griega. Pues bien, punto a favor del intuicionismo, la matemática griega era, en principio y en ejercicio, *más o menos* intuicionista. Detengámonos un mo-

mento, por curiosidad, en esta cuestión; en concreto en los *Elementos* de Euclides [4].

Según la doctrina peripatética, el infinito, si bien existe, sólo lo hace *en potencia*, jamás *en acto*. Por ejemplo, el conjunto de los números naturales, como *totalidad*, no existe para Aristóteles, pero se admite que si tengo un número finito n, siempre existe uno más grande, sin límite (aquí tenemos el infinito *en potencia*, pero no llegamos a tenerlo *en acto*).

Es apasionante observar cómo en los *Elementos*, Euclides mide con un cuidado milimétrico sus definiciones para no invocar al infinito *en acto*, respetando las tesis de El Filósofo. Consideremos una recta. Desde los institutos aprendemos que las rectas no acaban. [El hecho de] *no acabar* está [incluido] en [el concepto de] *recta*: estamos invocando el infinito en acto. Suspenso aristotélico. ¿Cómo define recta Euclides? En la definición 3 del libro primero establece que «los límites de una línea son puntos». ¡Está definiendo un segmento!

Veamos qué ocurre con el *V postulado* de Euclídes, el Postulado de las paralelas. Quizás el lector sospeche que este postulado reza lo siguiente: *Por un punto exterior a una recta se puede trazar, a lo sumo, una recta paralela.* Pues bien, no es así, y la diferencia es fundamental. En este enunciado, en el término paralelas, estamos asumiendo el infinito en acto. Magistralmente (no es casualidad, por descontado) Euclides se expresa de la siguiente manera:

> **Postulado V** Si una recta secante corta a dos rectas formando a un lado ángulos interiores, la suma de los cuales sea menor que dos ángulos rectos; las dos rectas, suficientemente alargadas, se cortarán en el mismo lado. ([11] Euclides III aC)

Al asegurar que las rectas se cortarán, se evita el infinito en acto. En el fondo lo que ha evitado aquí Euclides es lo que no puede dejar de definir (parece que lo atrasa lo máximo posible, pues es la última definición del libro primero) posteriormente: la noción de rectas paralelas. He aquí:

> **EI Definición 23** Rectas paralelas son aquellas que, estando en un mismo plano y siendo prolongadas indefinidamente en ambos sentidos, no se encuentran una a otra en ninguno de ellos. ([11] Euclides III aC)

Es la única definición que, inevitablemente, asume el infinito en acto. Para comprobar que no se cortan es necesario prolongarlas infinitamente. No es la única vez que Euclides no puede evitarlo: las proposiciones 12 y 22 del libro primero también asumen el infinito en acto.

[4] Véase ([30] Pla i Carrera, 2012) para un análisis similar de Arquímides y Proclo.

Por otro lado, aunque la mayoría de demostraciones en los *Elementos* son constructivas, hay un número sustancial de proposiciones y teoremas probados por *Reductio ad absurdum*, por ejemplo las proposiciones 6, 31 y 20 de los libros primero, séptimo y noveno respectivamente.

Cerramos ya este paréntesis euclidiano y volvemos al tema que nos ocupa.

Como decíamos, la matemática de los griegos era esencialmente intuicionista. Sin embargo la irrupción de las geometrías no euclideanas era un revés importante. Desde entonces es claro el hecho del distanciamiento brutal entre los planteamientos intuicionistas y el hacer matemático cotidiano. Por eso la crítica a la intuición matemática suelen dirigirse en torno a las *construcciones* complejas a las que se ha llegado en la matemática avanzada: ¿alguien ha tenido una intuición espacial, o celestial, de la *bola de Banach-Tarski* (efectivamente construida)? [5]

En la Matemática moderna lo no intuitivo era el pan de cada día. Hahn utilizaba este argumento y comentaba: «*Because intuition turned out to be deceptive in so many instances, and because propositions that had been accounted true by intuition were repeatedly proved false by logic, mathematicians became more and more sceptical of the validity of intuition*» («Puesto que la intuición resultó ser decepcionante en tantos casos, y debido a que las proposiciones que la intuición había considerado verdaderas fueron repetidamente probadas como falsas por la lógica, los matemáticos se volvieron cada vez más escépticos sobre la validez de la intuición»). Y esto es importante pues, en principio, según esta visión, la intuición debe acompañarnos en todo nuestro *hacer matemático*. Es decir, no sirve intuir, de alguna manera, "lo más básico", y dedicarnos a extraer conclusiones no intuitivas.

El Análisis de Weyl destruye el Análisis, básicamente. El estricto constructivismo le impide llegar al continuo: si consideramos la construcción de $\mathbb{R}$ con sucesiones de Cauchy en $\mathbb{Q}$, el problema con el que se topa Weyl es que no puede considerar todas las posibles sucesiones de Cauchy de $\mathbb{Q}$, sino sólo las que puede construir, que son una cantidad numerable. Brouwer, en un intento de llegar al continuo, acabaría aceptando poder considerar todas las sucesiones *posibles* de Cauchy *en potencia*, o sea, que las sucesiones, más que sucesiones, son objetos que, por poder, pueden continuarse indefinidamente, por así decirlo. Hilbert no daba crédito, no entendía que un gran matemático como Brouwer aceptara destruir el Análisis por no dar validez el Principio del Tercio Excluso.

La Aritmética de Heyting tuvo algo más de éxito. En 1933 Gödel probó que para toda fórmula F demostrable en la Aritmética de Peano existe una fórmula equivalente G demostrable en la Aritmética de Heyting. Tampoco es para tirar cohetes: G es una fórmula de la lógica clásica, quizás no sea válida en la Lógica intuicionista. Además ocurre lo siguiente ([17] George & Velleman, 2002, 122): es

[5] En realidad este argumento podría rechazarse arguyendo que la construcción depende críticamente del axioma de elección, que no es intuitivo. El error vendría de asumir un axioma así.

posible que la demostración de F utilice reducción al absurdo. Aún así sabemos que existe una fórmula equivalente G en la Aritmética de Heyting. Pero puede que todavía no se conozca una demostracón constructiva de G, en cuyo caso G aún no es verdadera, ni falsa, para el matemático intuicionista.

Pero lo que late en el fondo del intuicionismo es la gran pregunta filosófica: ¿podemos fiarnos de nuestra intuición? ¿Y por qué? Está claro que siempre, siempre, no. La intuición no es una facultad fija e inmóvil, sino que depende de nuestros conocimientos del mundo, de nuestra experiencia acumulada y de nuestras reflexiones sobre ella. Un ejemplo es la teoría de la relatividad especial, donde la no simultaneidad, la dilatación del tiempo o el carácter absoluto de la velocidad de la luz se nos revelan extraños y antiintuitivos. A pesar ello se han corroborado por cientos de experimentos, y son aceptados desde hace décadas por la comunidad científica. Como dicen Alan Sokal y Jean Bricmont ([34] Sokal & Bricmont, 1999), «no existe ninguna contradicción entre las predicciones de la relatividad y nuestra experiencia cotidiana, sino que, una vez más, la contradicción se da entre la relatividad y una extrapolación *errónea* de nuestra experiencia de cada día». Actualmente las paradojas relacionadas con la relatividad especial no son generalmente conocidas, pero con el paso del tiempo se van asentando más y más en nuestro conocimiento a pesar de su antiintuitividad, pues el contacto de la sociedad con ellas, sobre todo en los colegios, se va generalizando (si bien es cierto que seguimos sin tener «experiencias relativistas», y su carácter antiintuitivo se mantiene y se mantendrá).

Por último, comentar que el carácter constuctivo de las demostraciones que Gödel da de sus teoremas hace que sean admitidas por los matemáticos intuicionistas. En general, salvo más o menos estrafalarias excepciones, como Wittgenstein, todos los lógicos y matemáticos las aceptan.

1.1.4. Formalismo

1. ¿Cuál es la diferencia entre las matemáticas y el ajedrez? Respuesta formalista: esencialmente, ninguna. La posición formalista plantea que las matemáticos son, por así decirlo, un juego. La Matemática es un sistema de signos sin sentido alguno para los cuales se definen ciertas reglas de manipulación válidas: los conceptos pasan a ser los signos; sus ideas a ser hileras de signos; el razonamiento y la deducción a la manipulación combinatoria y la deducción formal. Es decir, trabajamos con *lenguajes formales*. Recordando las palabras de Poincaré, vemos que las posiciones del formalismo y el intuicionismo son opuestas.

Una teoría matemática consistirá en lo siguiente: Partimos de ciertas relaciones de símbolos básicas dadas por válidas llamadas *axiomas* y, a través de las reglas válidas de deducción (lógica de primer orden), obtenemos nuevas relaciones de símbolos, los *teoremas*. *Hacer matemátcas* es *deducir*. Ahora bien, para que esto tenga sentido[6] debemos garantizar las siguientes propiedades: (1) que de los axiomas no se puedan deducir contradicciones (teoría consistente); (2) que de los axiomas se puedan deducir *todas* las proposiciones verdaderas (teoría completa). Debemos ser cuidadosos: no hemos dicho todos los *teoremas*. La completitud se refiere a si todos las proposiciones verdaderas pueden derivarse de los axiomas; esto es, si todas las proposiciones verdaderas son teoremas (o, si se quiere, teoremas formales).

Además parece necesario asumir que sólo serán válidos los procesos *finitistas*, esto es, que el número de pasos lógicos de una demostración sea finito.

2. El formalismo en matemáticas va inevitablemente asociado a dos cosas, y casi se podría decir que en relación causa-efecto: David Hilbert y los teoremas de Gödel. Pero empecemos por el principio:

El primer sistema axiomático fue el que presenta Euclides en los *Elementos*, como ya hemos comentado. Por muy genial que era los matemáticos fueron descubriendo fallas en el mismo. En 1733 Jerónimo Saccheri, negando el axioma de las paralelas, desarrolló los primeros teoremas de le geometría no euclídea. Sin embargo, como ya dijimos, acabó por no aceptar su propio trabajo. Dedekind encontró incongruencias debido a la falta de un *axioma de continuidad* en los *Elementos*: un plano sin los puntos con coordenadas irracionales sigue siendo euclidiano, pero no tiene por qué ocurrir que dos circunferencias que pasan cada una por el centro de la otra se corten. Posteriormente, a finales del siglo XIX, llega el descubrimiento de la geometría hiperbólica, que supone un terremoto en la fundamentación de la Matemática. La tabla de salvación de la Geometría la presenta David Hilbert en 1899, publicando los famosos *Grundlagen der Geometrie* ([20] Fundamentos de la geometría), un trabajo genuinamente formalista (no hay definiciones constructivas) donde axiomatiza la geometría sobre una base conjuntista (a la manera de Peano). El programa formalista quedaba claro en la *Introducción*:

> Para su consecuente construcción, la Geometría, lo mismo que la Aritmética, solamente necesita unas pocas y sencillas proposiciones elementales.
>
> Estas proposiciones elementales se llaman *axiomas* de la Geometría. El poner de manifiesto los axiomas de la Geometría y el averiguar sus conexiones es un problema que se discute desde los tiempo de Euclides en numerosos y excelentes tratados de la literatura matemática. El problema se reduce al análisis lógico de nuestras intuiciones espaciales.
>
> La presente investigación es un nuevo ensayo para construir la Geometría sobre un sistema completo de axiomas, lo más sencilo posible, deduciendo de él los más importantes teoremas, de tal manera que en ese proceso aparezcan

[6] Pretendemos que un sistema formal refleje perfectamente una teoría definida semánticamente (esto es, todas las sentencias verdaderas de un modelo determinado). Por ejemplo, buscamos una aritmética formal que refleje la aritmética natural, que, como toda teoría definida semánticamente, es completa y consistente. Es preciso entonces asegurar la completitud y la consistencia del sistema formal.

> con la máxima claridad la interpretación de los distintos grupos de axiomas y el alcance de las consecuencias que aisladamente se derivan de cada uno de ellos. (1)

Reflexionemos un momento sobre la frase «...deduciendo de él los más importantes teoremas, de tal manera que en ese proceso aparezcan con la máxima claridad la interpretación de los distintos grupos de axiomas». Creemos captar las ideas de Hilbert en torno a los conceptos de *consistencia* y *verdad* en matemáticas. Ahondaremos en el tema; y para ello recurriremos al debate que mantuvo con Frege, que quedó plasmado en una correspondencia que es oro puro. Las cuestiones son: ¿Qué axiomas son válidos? ¿Qué otorga validez a los axiomas? Y antes de nada debemos explicar cómo se comprueba la no contradicción de los mismos.

El método general para comporbar la no contradicción de unos axiomas, esto es, la consistencia interna de un sistema, es encontrar un «modelo» (el teorema de Henkin nos asegura que una teoría de primer orden es consistente si, y sólo si, tiene un modelo) donde se pueda evaluar la veracidad de esas afirmaciones abstractas. De esta manera estamos trasladando un problema «metamatemático», la consistencia de un sistema, al sistema, a las matemáticas. Pongamos un ejemplo que tomamos de ([27] Nagel & Newman, 1958, 13-14):

Consideremos las clases K y L carcaterizadas por los siguientes axiomas:

1. Dos miembros cualesquiera de K se hallan contenidos en un solo miembro de L.

2. Ningún miembro de K se halla contenido en más de dos miembros de L.

3. No todos los miembros de K se hallan contenidos en un único miembro de L.

4. Dos miembros cualesquiera de L contienen a un solo miembro de K.

5. Ningún miembro de L contiene a más de dos miembros de K.

De estos postuldos podemos deducir teoremas. Por ejemplo, resulta que K sólo contiene tres miembros. Para comprobar la consistencia introducimos el siguiente modelo:

Sea K la clase cuyos elementos son los vértices de un triángulo, y L la clase de las líneas que froman los lados del mismo. Cuando decimos «un miembro de K está contenido en un miembro de L» nos referimos a que un punto que es un vértice está contenido en los lados del triangulo de los cuales es vértice.

Y resulta que en este modelo geométrico (hemos trasladado un problema metamatemático a la geometría) los cinco postulados son verdaderos, y por tanto consistentes:

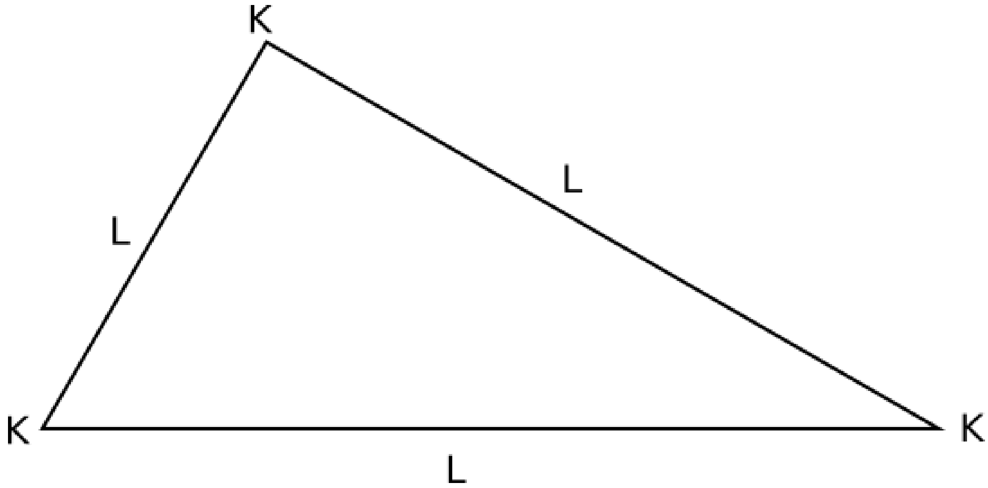

Figura 1.1.1 Un modelo para los postulados 1-5 es un triángulo.

1. Dos vértices están contenidos en un sólo lado del triángulo.
2. Ningún vértice está contenido en más de dos lados.
3. No todos los vértices están en un sólo lado.
4. Dos lados cualesquiera del triángulo contienen a un único vértice.
5. Ningún lado contiene a más de dos vértices.

Sin embargo, la realidad no suele ser tan sencilla. La geometría riemanniana, por ejemplo, puede interpretarse en un modelo euclidiano, considerando «esfera» en lugar de «plano» y «arcos de máxima longitud» en lugar de «rectas». ¿Queda probada con esto la consistencia de la geometría riemanniana? No, sólo se ha pospuesto el problema: la consistencia de la geometría riemanniana depende de la consistencia de la geometría euclidiana. Esto es lo que se conoce como *prueba relativa de consistencia*. Por contra, una *prueba absoluta de consistencia* es aquella en la que la consistencia queda probada sin necesidad de exigir la consistencia de otro sistema.

Además este método contiene otro problema espinoso: en los modelos donde interpretamos los axiomas puede ser necesario considerar un número infinito de elementos. Es imposible entonces comprobar totalmente los axiomas, pues tendríamos que hacer infinitas observaciones. Si tuviésemos un modelo finito, como el triángulo del ejemplo, entonces no hay ningún problema, pues podemos examinar todos los vértices y lados para ver que *todos* los elementos del modelo verifican los axiomas, y así poder asegurar su veracidad. Pero generalmente los axiomas de las distintas ramas de la Matemática requieren modelos infinitos. Consideremos por ejemplo el axioma aritmético por el cual, para todo número entero dado, siempre existe un inmediato sucesor. Es eviente que necesitamos un modelo infinito para comprobarlo.

Volvamos a los *Grundlagen*. La axiomatización de la Geometría que plantea Hilbert no contiene una demostración absoluta de consistencia, sino una demostración relativa. Hilbert recurre a un modelo algebraico, en particular a la geometría cartesiana. Los axiomas de Euclides son *verdades en este modelo*. Por ejemplo, el teorema geométrico de que una línea corta a un círculo en dos puntos como máximo se convierte un teorema algebraico que consiste en un sistema de dos ecuaciones con dos incógnitas, cuya solución son dos pares de números reales a lo sumo. Hemos trasladado el problema de consistencia de la Geometría al de la consistencia del álgebra:

> Conforme con esta exigencia, he demostrado en los Fundamentos de la geometría la no-contradicción de los axiomas propuestos haciendo ver que toda contradicción que se deriva lógicamente de axiomas geométricos debía necesariamente manifestarse también en la aritmética del conjunto de los números reales. ([20] 1917, 133)

Que, a su vez, se reduce a la consistencia de los números enteros:

> Asimismo, el problema de la no-contradicción de un sistema de axiomas para los números reales se puede referir a un problema que contempla los números enteros. Es mérito de Weierstrass y de Dedekind el haberlo mostrado en su teoría de los números irracionales. (133)

Y la consistencia de los número naturales, si queremos evitar una regresión infinita, nos lleva a la consistencia del que Hilbert consideraba «el paraíso» de los matemáticos: la teoría de conjuntos. Y parece que llegamos a un callejón sin salida:

> El axioma de los números enteros y las bases de la teoría de conjuntos constituyen, sin embargo, casos únicos de excepción. El camino que conduciría a un dominio científico más específico que el suyo parecería inaccesible, porque fuera de la lógica no existe ninguna disciplina a la cual apelar como último recurso. No obstante, como el deber de establecer la no-contradicción es ineluctable, es necesario, me parece, axiomatizar la lógica y probar que tanto la teoría de números como la de conjuntos son sólo partes de la lógica. (133)

Como solución Hilbert propone su Teoría de la Demostración. Como dijimos, una *prueba absoluta de consistencia* es aquella en la que la consistencia es probada sin necesidad de exigir la consistencia de otro sistema. Añadimos ahora que es necesario también que no precisemos de un número infinito de de propiedades y/o operaciones. Se definen ciertos símbolos, se dan las reglas válidas para manipularlos; el número de manipulaciones debe ser finito; se llega así a fórmulas, que a su vez estarán conectadas por un número finito de relaciones. Es lo que entendemos por procedimientos *finitistas*. Con ello, y manteniendo la idea de demostrar las cuestiones «metamatemáticas» a través de la matemática, debemos llegar a una prueba de que no podemos obtener fórmulas contradictorias a partir de los axiomas. Von

Neumann, defensor también del formalismo, describía así el programa de Hilbert en «*Die formalistische Grundlegung der Mathematik*» («Los fundamentos formalistas de las matemáticas», 1930): tenemos ciertas fórmulas no ambigüas y caracterizadas de modo finitista que damos por ciertas, los axiomas; ahora, si probamos A y probamos $A \to B$, entonces hemos probado B. Esto es la Matemática.

Ahora que queda claro la cuestión de la consistencia de los axiomas, pasamos a la pregunta sobre qué da realmente validez a los mismos. Podemos tomar, a grosso modo, una de estas posiciones: (1) Si los axiomas son verdaderos, entonces no se contradicen; (2) Si los axiomas no se contradicen, entonces son verdaderos. La primera es la posición de Frege, la segunda es la de Hilbert.

El 27 de diciembre de 1899, tras leer los *Grundlagen*, Frege escribe una carta a Hilbert en la que dice lo siguiente:

> A mi entender esto [el hecho de establecer que los axiomas establecen la definición de los signos que involucran] borra la línea divisoria entre las definiciones y los axiomas de una forma dudosa, y junto al antiguo significado de la palabra "axioma" que emerge en la proposición de que los axiomas expresan hechos fundamentales de la intuición, emerge otro significado que no comprendo suficientemente bien. ([15] Frege 2, 35-36)

Nótese el logicismo de Frege, al considerar los axiomas como proposiciones intuidas, en contraposición al formalismo, donde los axiomas son simples hileras de signos que, pudiendo ser verdad o no, al ser carentes de significado, no pueden jamás intuirse. Para Hilbert la no-contradicción es criterio de verdad y existencia. Contesta lo siguiente:

> Usted escribe: «Denomino axiomas a las proposiciones que son verdaderas pero que no se han demostrado, porque nuestro conocimiento de ellas surge de una fuente distinta de la lógica, una fuente que podría llamarse intuición espacial. De la verdad de los axiomas se sigue que no se contradicen entre sí». Encuentro muy interesante leer esta frase de su carta, ya que por lo que a mi se refiere, siempre que he pensado, escrito y dado clases o conferencias sobre estos temas, he dicho justo lo contrario: Si los axiomas dados arbitrariamente no se contradicen entre sí con todas sus consecuencias, entonces son verdaderos ([15] Frege 2, 37)

No debemos enclasar a Frege en un intuicionismo ingenuo. En esta declaración que escribe en «*Die Grundlagen der Arithmetik*» («Los fundamentos de la Aritmética», [15] 1884) vemos su compromiso con grantizar la consistencia de los sistemas axiomáticos:

> Pues bien, entonces lo que debemos hacer en primer lugar es probar que estos otros postulados nuestros no contienen contradicción. Hasta que no hayamos hecho esto, todo el rigor, por mucho que nos esforcemos, será música celestial. (121)

Frege era bien consciente de que la intuición era engañosa. La verdad intuitiva de los axiomas, bien que es previa al sistema axiomático, debe refrendarse con un modelo. La validez en el modelo implica la consistencia:

> La única forma que conozco [de establecer la consistencia de un sistema] es ésta: para ver que un objeto cumple todas esas propiedades, es preciso ofrecer un caso en el cual tales requerimientos se satisfagan. ([15] Frege 2, 43)

Como anticipábamos en la Introducción, nos enfrentamos a cuestiones que se encuentran simultánemente, como si dijéramos, entre la Matemática (ciencia) y la Filosofía de las Matemáticas (filosofía). Creemos entreverlo en Hilbert en las siguientes palabras:

> Este logro, sin embargo, requiere todavía de un trabajo nuevo y múltiple. En efecto, una reflexión más profunda muestra pronto que el problema de la no-contradicción en los conjuntos y en los números enteros no se limita a ellos mismos, sino que se relaciona con un vasto dominio de preguntas difíciles que competen a la teoría del conocimiento aunque tengan un carácter netamente matemático. Para caracterizar brevemente este conjunto de preguntas, me limitaré a una simple enumeración. *Un problema matemático, ¿supone siempre una solución?* Pregunta capital a la que se relaciona subsidiariamente la siguiente: el resultado de una investigación matemática ¿es siempre *controlable*? En el mismo orden de ideas, ¿qué se entiende por *criterio de simplicidad* de las pruebas matemáticas? ¿Cómo definir en la matemática y en la lógica la relación entre *contenido y forma*? Finalmente, ¿en qué consiste la *determinación* de un problema matemático por medio de un número finito de operaciones?
>
> La axiomatización de la lógica sólo podrá satisfacemos enteramente cuando todas las preguntas de esta naturaleza sean resueltas y su relación aclarada. ([20] 1917, 133-134)

Decíamos que el sistema axiomático de Hilbert para la Geometría se había construido sobre base conjuntista "a la manera de Peano". Y esque la primera axiomatización moderna (tras los trabajos de Frege, claro, pero aquí nos centramos ya en el programa formalista) no son los «*Grundlagen*», sino los «*Grundlagen der Arithmetik*» (1889) de Giuseppe Peano, donde el famoso matemático italiano había axiomatizado la Aritmética. Posteriormente, en 1897, refinaba su teoría presentando los célebres *Axiomas de Peano*. La consistencia de su teoría descansaba en la consistencia de la teoría de conjuntos. En sus obras, Peano propone (y emplea) una nueva simbología lógica más manejable que la de Frege. Ésta se expandió rápidamente entre los matemáticos y, con aportaciones de Whitehead y Russell en *Principia*, se convirtió en el lenguaje común de la Lógica Matemática. Hilbert es en este sentido heredero de Peano.

En 1904, el alemán Ernst Zermelo prueba el teorema del buen orden postulando el *axioma de elección*, pero para fundamentar completamente su trabajo se hacía patente la necesidad de una axiomatización de la teoría de conjuntos. Él mismo abre el melón publicando un sistema axiomático en 1908, a pesar de no haber podido probar la consistencia. Posteriormente su compatriota Adolf Fraenkel completa la teoría (la consistencia sigue sin ser probada), dando lugar a los axiomas de Zermelo-Fraenkel (ZF). Con esta axiomatización se evitaban las paradojas de autorreferencia, como la paradoja de Cantor o la de Russell. En concreto gracias al *axioma de regularidad* (también llamado de *fundamentación*), que impide que un conjunto pueda pertenecerse a sí mismo:

> **Axioma de regularidad** Si X es un conjunto no vacío ($X \neq \emptyset$), existe un $x \in X$ tal que $x \cap X = \emptyset$.

Este axioma impide que haya cadenas infinitas de pertenencias sucesivas o cadenas finitas cerradas, como $x \in x$. Por tanto, en la teoría axiomática de conjuntos, la clase de Russell $\mathcal{R} = \{x : x \notin x\}$ coincide con el universo $\Upsilon = \{x : x = x\}$ de todos los conjuntos. Así, como $\mathcal{R}$ es una clase propia (esto es, una clase que no es un conjunto, pues no es un elemento de ninguna otra clase), Υ también, y evitamos así la paradoja de Cantor[7] Posteriormente se probó que era independiente del resto de axiomas de ZF. Entre los axiomas de ZF el que quizás más dudas generaba era el *axioma de elección* (AE) heredado de Peano, que dice así:

> **Axioma de elección** Para cada conjunto X existe una función de elección: una función $f : X \longrightarrow \bigcup X$ tal que, para cada conjunto $x \in X$, con $x \neq \emptyset$, $f(x) \in x$.

Este postulado sólo presenta dificultad para conjuntos infinitos, pues para el caso finito podemos hacer elecciones explícitas. Los axiomas de Zermelo-Fraenkel, contando el axioma de elección, son diez, y la teoría se abrevia ZFC («C» de Choice, de *axiom of choice*). Actualmente podemos decir que el «universo de la Matemática» es la teoría ZFC+HGC, donde HGC designa la hipótesis general del continuo, la generalización de la Hipótesis del continuo conjeturada por Cantor:

> **Hipótesis del continuo** No existe ningún conjunto X cuyo cardinal sea mayor que el de $\mathbb{N}$ y menor que el de $\mathbb{R}$; esto es, que cumpla lo siguiente:
>
> $$\aleph_0 < |X| < 2^{\aleph_0} \tag{1.1.1}$$
>
> **Hipótesis general del continuo** Para cualquier conjunto X, no existe ningún conjunto Y cuyo cardina sea mayor que el de X y menor que el de $\mathcal{P}(X)$; esto es, que cumpla lo siguiente:
>
> $$|X| < |Y| < |\mathcal{P}(X)| \tag{1.1.2}$$

[7]La paradoja de Cantor se formula así: el conjunto $\mathcal{P}(\Upsilon)$, que está bien defnido, cumple, por definición de universo Υ, que $\mathcal{P}(\Upsilon) \subset \Upsilon$. Pero entonces $|\mathcal{P}(\Upsilon)| \leq |\Upsilon|$, lo cual contradice el teorema de Cantor, según el cual, para todo conjunto X, $|X| < |\mathcal{P}(X)|$. La solución está en que el universo Υ no es un conjunto.

Por último debemos remarcar que en la década de 1930 se presentó otra teoría axiomática conjuntista, la Teoría de Clases y Conjuntos de von Neumann-Bernays-Gödel. Es una extensión de ZFC. NBG introduce la noción de clase, que es una colección de conjuntos definidos por una fórmula cuyos cuantificadores se extienden sólo sobre conjuntos. NBG puede definir clases que son más grandes que los conjuntos, como la clase de todos los conjuntos y la clase de todos los ordinales. En cualquier caso ZFC es más popular que NBG, sobre todo debido a que trabajar en NBG es más laborioso y a que NBG es una extensión conservadora de ZFC.

3. La magna pregunta es: Dios mío, ¿es la teoría de conjuntos consistente y completa? La respuesta son los teoremas de incompletitud de Gödel: si es consistente, entonces no es completa y habrá verdades no demostrables; precisamente una de ellas es la consistencia del sistema. O sea, que no.

Los teoremas de incompletitud impiden que una teoría formal sea a la vez consistente y completa, los dos requisitos del programa formalista de construcción de la Matemática. En otras palabras, es imposible formalizar completamente la Matemática. Aunque el formalismo fracasó como fundamentación de la Matemática, su metodología nos acompaña actualmente en última instancia. Destaquemos también las incursión científica que suponen los teoremas de Gödel en la Filosofía de las Matemáticas. Esto está en concordancia con la concepción de Filosofía que expusimos: las Ideas filosóficas brotan de los conceptos científicos; y la Filosofía es por tanto, no tanto la madre de la Ciencia, sino su hija, pues toma como dadas las ciencias positivas.

Gödel y Paul Cohen probaron que el AE es independiente de los axiomas de ZF. Esto significa que, si ZF es consistente, entonces ni AE ni su negación pueden demostrarse formalmente a través de los axiomas de ZF (es una de las sentencias indecidibles en ZF). En consecuencia, asumir AE o su negación nunca llevará a una contradicción que no se pudiera obtener sin tal supuesto. Por lo tanto, la decisión de usar el AE en las demostraciones deberá tomarse al margen de los axiomas de ZF. Una razón a favor podría ser la simple conveniencia: el hecho de usarlo no provoca contradicciones y hace posible demostrar algunas proposiciones que de otro modo no se podrían probar.

Por su parte, la HGC también es independiente de ZF. Es más, es independiente de ZFC. Sin embargo, ZF + HCG necesariamente implica AE, con lo cual HCG es estrictamente más fuerte que AE, aunque ambos sean independientes de ZF.

Que la teoría de conjuntos sea incompleta (asumiendo que es consistente, pues si no, como yo hemos dicho, *todo vale*[8]) no implica que todas las teorías lo sean; depende de la lógica usada.

[8] Parafraseando la famosa expresión del filósofo de la ciencia Paul Feyereband: «Todas las metodologías tienen sus límites, y la única “regla” que sigue siendo válida es: “Todo vale”» ([14] Feyereband, 1975, 296). El vienés siempre defendió que la cita no debía interpretarse literalmente.

La Lógica es el estudio del razonamiento, que se da en el lenguaje. Dependiendo del mismo, la lógica emergente tendrá unas características u otras. En concreto, las dos cuestiones a estudiar en un lenguaje formal son: (1) su sintaxis, que caracteriza su capacidad para efectuar demostraciones, y (2) su semántica, es decir, su concepción de *verdad*. ¿Cuál es la relación entre la demostrabilidad y la verdad? ¿Todo lo que es verdadero es demostrable? ¿Todo lo demostrable es verdadero?

Generalemente un lenguaje formal rico en propiedades expresivas, esto es, capaz de incluir buena parte de las ramas de la Matemática, tiene propiedades semánticas (metateóricas) pobres. Por ejemplo, la teoría de conjuntos, que si es consistente, entonces es incompleta; o en general las lógicas de segundo orden, que tienen el poder expresivo suficiente como para ser afectadas también por los teoremas de Gödel. Y viceversa, un lenguaje formal con una semántica robusta no puede expresar sintácticamente una buena parte las matemáticas. Por ejemplo, la lógica correspondiente a los lenguajes formales de primer orden. El teorema de completitud de Gödel establece que en una lógica de primer orden toda sentencia verdadera es lógicamente demostrable, pero sólo nos permite cuantificar sobre individuos y no sobre funciones o relaciones.

En la aproximación formal precisaremos qué teorías formales caen bajo los teoremas de incompletitud.

Capítulo 2

Aproximación formal

2.1. Lenguajes formales de primer orden

Por extensión, no nos es viable abarcar con exactitud la teoría de lenguajes de primer orden, luego vamos a ceñirnos a lo imprescindible para entender la demostración de Gödel. Para un tratamiento completo véase ([13] Fernández Margarit, 2012). Este libro, «Lógica Matemática y Teoría de Conjuntos», basado en los detallados apuntes que el profesor Alejandro Fernández Margarit elaboró durante años, ha sido fundamental para la redacción de este capítulo. El fantástico libro de Josep Pla i Carrera, «El Teorema de Gödel. Un análisis de la verdad matemática» ([30] Pla i Carrera, 2012), publicado por la Real Sociedad Matemática Española en conmemoración del Año Turing en 2012, también ha sido absolutamente imprescindible y ha constituido la base de este capítulo, sin duda. Mi reconocimiento sincero a estos dos autores. Si el lector desea repasar lo lógica proposicional, así como una introducción a la lógica de primer orden, puede consultar el libro clásico de Kolmogórov ([22] Kolmogórov & Dragalin), o ([1] Aranda et al., 2000), un manuel universitario orientado a estudiantes de computación.

2.1.1. Sintaxis

Un lenguaje de primer orden **L** es un conjunto de ciertos símbolos con ciertas características. Comunes en todos ellos son los símbolos *lógicos*, y específicos de cada lenguaje son los símbolos *no lógicos*. Dependiendo de ellos, el lenguaje será uno u otro, y será de más o menos utilidad para describir una teoría matemática.

Definición 2.1.1 (Símbolos)**.** Los símbolos de un lenguaje de primer orden son los siguientes:

(1) *Lógicos*:

- Variables: $v_0, v_1, v_2, ...$ $(x, y, z...)$

- Predicado binario de igualdad: $=$
- Conectivas lógicas:
  - Proposicionales: $\begin{cases} \neg & \text{, negación} \\ \vee & \text{, disyunción} \end{cases}$
  - $\exists$, cuantificador existencial

(2) *No lógicos*:

- Constantes: $\mathbf{c}$
- Funciones: para cada $n > 0$, letras funcionales n-arias ($\mathbf{f}$, $\mathbf{g}$, $\mathbf{h}$...)
- Predicados: para cada $n > 0$, letras relacionales n-arias ($\mathbf{p}$, $\mathbf{q}$, $\mathbf{r}$...)

A partir de ahora, al conjunto de las constantes lo llamaremos $\mathbf{L_C}$, al de las funciones $\mathbf{L_F}$ y al de los predicados $\mathbf{L_P}$.

Ejemplo 2.1.2. Veamos algunos ejemlos de lenguajes de primer orden que nos permiten estudiar distintas teorías (recordemos que basta con dar sus símbolo *no lógicos*):

1. **Teoría de conjuntos:** $\mathbf{L}_{con} = \{\in\}$, con $\in$ un predicado binario.
2. **Teoría de Grupos:** El lenguaje formal para describir una teoría no es único, y aquí tenemos un ejemplo. Se puede estudiar la teoría de grupos con distintos lenguajes:
   - $\mathbf{LG} = \begin{cases} + & \text{, función binaria} \\ \mathbf{0} & \text{, constante (elemento neutro)} \end{cases}$
   - $\mathbf{LG}_1 = \begin{cases} + & \text{, función binaria} \\ \mathbf{0} & \text{, constante (elemento neutro)} \\ - & \text{función 1-aria (opuesto)} \end{cases}$
   - $\mathbf{LG}_2 = \{\mathbf{p}\}$, con $\mathbf{p}$ un predicado 3-ario (grafo de la suma)
3. **Teoría de números (aritmética)**[1]**:** $\mathbf{LA} = \begin{cases} + & \text{, función binaria} \\ \cdot & \text{, función binaria} \\ \mathbf{S} & \text{, función 1-aria (siguiente)} \\ \mathbf{0} & \text{, constante} \\ < & \text{predicado binario} \end{cases}$

[1]Notemos que aún no tiene sentido de hablar de la función binaria + como la 'suma', o de la

Una expresión de un lenguaje de primer orden es una hilera finita de símbolos. En este sentido, cuando dos hileras **s** y **s'** son exactamente iguales, escribiremos $\mathbf{s} \equiv \mathbf{s}'$. Podemos distinguir dos expresiones fundamentales: *términos* y *fórmulas*.

Definición 2.1.3 (Términos)**.** Describen los elementos básicos del universo de discurso. Definimos el conjunto de términos, $\mathbf{Term}(\mathbf{L})$, recursivamente:

- Términos elementales: $\begin{cases} \text{Variables} \\ \text{Constantes} \end{cases}$
- Términos compuestos: $\mathbf{f}(\mathbf{t}_1, ..., \mathbf{t}_n)$, donde $\mathbf{t}_1, ..., \mathbf{t}_n$ son términos elementales y $\mathbf{f}$ una letra funcional n-aria, con $n > 0$.

Definición 2.1.4 (Fórmulas)**.** Determinan las propiedades sobre los elementos del universo de discurso. Definimos el conjunto de fórmulas, $\mathbf{Form}(\mathbf{L})$, recursivamente:

- Fórmulas atómicas $\mathbf{At}(\mathbf{L})$: $\mathbf{p}(\mathbf{t}_1, ..., \mathbf{t}_n)$, donde $\mathbf{t}_1, ..., \mathbf{t}_n$ son términos (elementales o compuestos) y $\mathbf{p}$ es un predicado n-ario, con $n > 0$.
- Fórmulas compuestas: Están formadas por fórmulas atómicas y conectivas lógicas. En concreto[2]:
  - Si φ es una fórmula, entonces $\neg\varphi$ es una fórmula.
  - Si φ y ψ son fórmulas, entonces $\varphi \vee \psi$ es una fórmula.
  - Si φ es una fórmula y x una variable, entonces $\exists x\varphi$ es una fórmula.

Una vez definidos estos conceptos, nos será de utilidad dar un símbolos a ciertas fórmulas habituales; esto no es sino abreviarlas, es decir, no estamos definiendo nuevos símbolos *no lógicos* de los lenguajes de primer orden. En particular, podemos definir la conjunción '$\wedge$', el cuantificador universal '$\forall$', el condicional material '$\rightarrow$' y el bicondicional '$\leftrightarrow$':

$$\text{Escribiremos} \begin{cases} \varphi \wedge \psi \\ \forall x\varphi \\ \varphi \rightarrow \psi \\ \varphi \leftrightarrow \psi \end{cases} \text{en lugar de} \begin{cases} \neg(\neg\varphi \vee \neg\psi) \\ \neg\exists x\neg\varphi \\ \neg\varphi \vee \psi \\ (\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi) \end{cases}$$

función 1-aria **S** como la función 'siguiente'. Estas nociones sólo tienen cabida en una interpretación semántica de los símbolos del lenguaje, en concreto en la interpretación de los números naturales. Por el momento nos movemos estrictamente en el campo de la sintaxis. Por otra parte, se introducen los siguientes nuevos símbolos del lenguaje redundantes (son abreviaturas): $\mathbf{S}(\mathbf{0}) = 1, \mathbf{S}(\mathbf{S}(\mathbf{0})) = \mathbf{S}(\mathbf{1}) = 2, ...$ etc.

[2] No necesitamos especificar que partimos de fórmulas atómicas, pues ya hemos indicado que estamos definiendo de manera recursiva.

Ejemplo 2.1.5. Veamos algunos ejemplos de términos y fórmulas en distintos lenguajes de primer orden:

- En la teoría de grupos, usando en lenguaje **LG**, la expresión $x + (y + \mathbf{0}) + \mathbf{0})$[3] es un término, y la expresión $\forall x \exists y(x + y = \mathbf{0})$ es una fórmula, y quiere decir *todo elemento tiene un inverso.*

- En teoría de números, la expresión $\mathbf{S}(\mathbf{S}(y)) \cdot z$ es un término (se utilizan la función 1-aria **S** dos veces y posteriormente la función binaria $\cdot$), mientras que la expresión $\mathbf{S}(\mathbf{0}) < z \wedge \forall x \forall y(x \cdot y = z \to x = \mathbf{S}(\mathbf{0}) \vee x = z)$ es una fórmula, y significa *z es primo.*

Para completar la expliación de la sintaxis de un lenguaje formal de primer orden debemos dar pautas respecto a como tratar las variables, en particular sus posibilidades de sustitución. Ya hemos comentado que las variables son términos. Podemos distinguir dos casos esenciales en las estancias de una variable en una fórmula: las variable ligadas (o mudas) y las variable libres. Intuitivamente, una variable es libre cuando no es más que una notación *en* la fórmula, y es sustituible. Por contra, una variable ligada está especificada, ella como variable, a ciertos valores o relaciones. Formalmente:

Definición 2.1.6 (Variables libres y ligadas)**.** Diremos que una estancia de la variable x en una fórmula φ es ligada cuando esa estancia de x ocurre en una parte de la fórmula de la forma $\exists x \psi$[4]. Y si no aparece de esta forma, diremos que es una estancia libre (así, una variable puede ocurrir libre y ligada en una fórmula).
Una variable es libre (resp. ligada) en una fórmula φ, cuando en ella se da al menos una estancia libre (resp. ligada) de x.

Definición 2.1.7 (Sentencias)**.** Una fórmula φ es una sentencia si es cerrada, esto es, si no tiene variables libres.

Denotamos $\mathbf{Sent}(\mathbf{L})$ al conjunto de sentencias de **L**, y $\mathbf{Vl}(\varphi)$ al conjunto de las variables libres de una fórmula φ. Si queremos indicar que $\mathbf{Vl}(\varphi) \subseteq \{x_1, ..., x_n\}$, usaremos la notación $\varphi(x_1, ..., x_n)$.

Para realizar las demostraciones en lógica de primer orden seguiremos la *prueba por innducción* de la siguiente manera:

[3]Formalmente esta fórmula sería '$\neg\exists x \neg\exists y = +xy\mathbf{0}$', pero utilizaremos siempre esta escritura informal que es más intuitiva (ya lo venimos haciendo, pues, por ejemplo, '$\mathbf{p}(\mathbf{t}_1, ..., \mathbf{t}_n)$' se escribe formalmente '$\mathbf{pt}_1...\mathbf{t}_n$').

[4]En la medida en que el cuantificador universal $\forall$ no es más que una abreviatura de una expresión del cuantificador existencial $\exists$, es claro que una variable que ocurre en una parte de la fórmula de la forma $\forall x \psi$, también está ligada.

- Para probar que todo término $\mathbf{t} \in \mathbf{L}$[5] tiene una propiedad P, basta con probar que:
  1. Toda variable x tiene la propiedad P.
  2. Toda constante $\mathbf{c}$ tiene la propiedad P.
  3. Si $\mathbf{t}_1, ..., \mathbf{t}_n \in \mathbf{Term}(\mathbf{L})$, $\mathbf{f}$ una letra funcional n-aria, entonces $\mathbf{f}(\mathbf{t}_1, ..., \mathbf{t}_n)$ tiene la propiedad P.
- Para probar que todas las fórmulas $\varphi \in \mathbf{L}$ tienen la propiedad P, basta con probar que:
  1. Todas las fórmulas atómicas, $\varphi \in \mathbf{At}(\mathbf{L})$ tienen la propiedad P.
  2. Si φ y ψ tienen la propiedad P, entonces $\begin{cases} \varphi \vee \psi \\ \exists x \varphi \\ \neg \varphi \end{cases}$ tiene la propiedad P.

Definición 2.1.8. Sean $\mathbf{a}$, $\mathbf{b} \in \mathbf{Term}(\mathbf{L})$ y $\varphi \in \mathbf{Form}(\mathbf{L})$. Entonces:

- $\mathbf{b}_x[\mathbf{a}]$ es la expresión que se obtiene al sustituir las estancias de x en $\mathbf{b}$ por el término $\mathbf{a}$. Escribiremos habitualmente $\mathbf{b}(\mathbf{a})$.
- $\varphi_x[\mathbf{a}]$ es la expresión que se obtiene al sustituir las estancias libres de x en φ por el término $\mathbf{a}$. Escribiremos habitualmente $\varphi(\mathbf{a})$.[6]

Nótese la distinción fundamental: en una fórmula sólo definimos la sustitución de un término (recordemos, variables, constantes y términos compuestos) en las estancias libres de la variable correspondiente. Veamos un ejemplo:

Ejemplo 2.1.9. En el lenguaje de la aritmética **LA** (ejemplo 2.1.2, 3)):

φ es $\exists x(\mathbf{S}(x) = y) \vee (\mathbf{S}(x) < z)$, $\quad \mathbf{a}$ es $y + \mathbf{0} \implies \varphi_{x(\mathbf{a})}$ es $\exists x(\mathbf{S}(x) = y) \vee (\mathbf{S}(y + \mathbf{0}) < z)$

Proposición 2.1.10. Sean $\mathbf{a}, \mathbf{b} \in \mathbf{Term}(\mathbf{L})$ y $\varphi \in \mathbf{Form}(\mathbf{L})$. Entonces:

1. $\mathbf{b}(\mathbf{a}) \in \mathbf{Term}(\mathbf{L})$.
2. $\varphi(\mathbf{a}) \in \mathbf{Form}(\mathbf{L})$.

[5] Escribo el signo '$\in$' para abreviar la escritura. Es decir, lo uso en el lenguaje natural en el que escribo. Así debe interpretarse durante todo el texto, al igual que otros signos habituales con significado conocido, como '$=$' o '$\implies$', y no confundirse con ningún lenguaje formal concreto. En tal caso, se especificará el lenguaje que se está usando.

[6] De manera similar, unido a un proceso de sustitución sencillo, se define la sustitución de n términos en las estancias de n variable libres en un término o una fórmula, y se denota $\mathbf{b}(\mathbf{t}_1, ..., \mathbf{t}_n)$ y $\varphi(\mathbf{t}_1, ..., \mathbf{t}_n)$ respectivamente.

Demostración: 1. Si x no ocurre en $\mathbf{b}$, entonces $\mathbf{b}(\mathbf{a})$ es $\mathbf{b}$, que es un término. Supongamos ahora que x tiene alguna estancia en $\mathbf{b}$. Por inducción:

- Si $\mathbf{b}$ es una variable, entonces, si x ocurre en $\mathbf{b}$, $\mathbf{b}$ es x, luego $\mathbf{b}(\mathbf{a})$ es $\mathbf{a}$, que es un término.
- Si $\mathbf{b}$ no es una variable, entonces existen $\mathbf{t}_1, ..., \mathbf{t}_n \in \mathbf{Term}(\mathbf{L})$, y una letra funcional n-aria $\mathbf{f}$ tal que $\mathbf{b} = \mathbf{f}(\mathbf{t}_1, ..., \mathbf{t}_n)$. Por lo tanto, $\mathbf{b}(\mathbf{a}) = \mathbf{f}(\mathbf{t}_1(\mathbf{a}), ..., \mathbf{t}_n(\mathbf{a}))$. Por el método de inducción, $\mathbf{t}_1(\mathbf{a}), ..., \mathbf{t}_n(\mathbf{a}) \in \mathbf{Term}(\mathbf{L})$, luego por definición $\mathbf{b(a)}$ es un término compuesto.

2. Ahora utilizamos el método de inducción sobre φ:

- Si $\varphi \in \mathbf{At}(\mathbf{L})$, entonces existen $\mathbf{t}_1, ..., \mathbf{t}_n \in \mathbf{Term}(\mathbf{L})$ y un predicado n-ario $\mathbf{p}$ tal que $\varphi = \mathbf{p}(\mathbf{t}_1, ..., \mathbf{t}_n)$. Entonces tenemos que $\varphi(\mathbf{a}) = \mathbf{p}(\mathbf{t}_1(\mathbf{a}), ..., \mathbf{t}_n(\mathbf{a}))$. Por el apartado 1) de la demostración, $\mathbf{t}_1(\mathbf{a}), ..., \mathbf{t}_n(\mathbf{a}) \in \mathbf{Term}(\mathbf{L})$, luego φ es una fórmula (por definición, $\varphi \in \mathbf{At}(\mathbf{L})$).
- Si φ es $\psi \vee \theta$, entonces $\varphi(\mathbf{a})$ es $\psi(\mathbf{a}) \vee \theta(\mathbf{a})$. Por inducción, $\psi(\mathbf{a})$ y $\theta(\mathbf{a})$ son fórmulas, luego $\varphi(\mathbf{a})$ también.
- Si φ es $\neg\psi$, entonces $\varphi(\mathbf{a})$ es $\neg\psi(\mathbf{a})$. Por inducción, $\neg\psi(\mathbf{a})$ es una fórmula, luego $\varphi(\mathbf{a})$ también.
- φ es $\exists z\psi$. Como x es una estancia libre en ψ, z no es x. Tenemos entonces que $\varphi(\mathbf{a})$ es $\exists z\psi(\mathbf{a})$. Por la hipótesis de inducción, $\psi(\mathbf{a})$ es una fórmula, luego $\exists z\psi(\mathbf{a})$ también.

□

Lo que hemos probado es la validez del proceso de sustitución *sintácticamente*. Sin embargo nos encontramos con que el encaje *semántico* en las fórmulas es problemático. Por ejemplo, si φ es $\exists y\neg(x = y)$ y $\mathbf{a}$ es y, entonces $\varphi_x[\mathbf{a}]$ es $\exists y\neg(y = y)$. El contenido semántico de φ es que *existen dos elementos distintos*, mientras que la sustitución $\varphi(\mathbf{a})$ quiere decir que *existe un elemento distinto de él mismo*. La contradicción es que, si bien φ es correcta en ciertas situaciones, $\varphi(\mathbf{a})$ no es válida nunca.

En el próximo apartado nos ocuparemos ampliamente de la semántica de los lenguajes de primer orden. Para finalizar ahora, damos unas pautas que limitan el uso de la sustitución para evitar anomalías como la expuesta.

Definición 2.1.11 (Subfórmula)**.** Sea $\varphi \in \mathbf{Form}(\mathbf{L})$. Una expresión $\mathbf{u}$ es una subfórmula de φ si $\mathbf{u} \in \mathbf{Form}(\mathbf{L})$ y ocurre en φ. Es decir, existen $\mathbf{s}'$ y $\mathbf{s}''$, expresiones de $\mathbf{L}$, tales que $\varphi \equiv \mathbf{s}'\mathbf{u}\mathbf{s}''$.

Definición 2.1.12. Una variable libre x es sustituible por un término $\mathbf{t}$ en una fórmula φ si, para cada variable z que ocurre en $\mathbf{t}$, todas las subfórmulas de φ de la forma $\exists z\psi$ no contienen estancias libres de x en φ.

2.1.2. Semántica

La semántica de la lógica estudia el concepto de verdad de un lenguaje formal. Hagamos un pequeño comentario sobre la lógica en el lenguaje natural (este ejemplo se inspira en el presentado en [[30] Pla i Carrera, 2012]). Para ello pensemos en la siguiente declaración:

«Esta pared es azul».

Esta frase pertenece al español, un lenguaje natural. Comprobamos que es sintácticamente correca: sujeto, verbo y predicado. Ahora bien, además del análisis sintáctico debemos dilucidar si es cierta o no, esto es, debemos analizarla semánticamente. Y en ello lo primero en que recalamos es que su veracidad no se puede establecer taxativamente *a pripori*: depende. Puede que sea cierta (si, por ejemplo, me estoy refiriendo a las paredes del dormitorio de Van Gogh), o puede que no (si me refiero a las paredes blancas de mi habitación). Los distintos escenarios en los que esto ocurre son, intuitivamente, lo que en lógica se entiende por «interpretación», o, informalmente, por «modelo»[7]. Ahora bien, ¿por qué esta ambigüedad? La respuesta es sencilla: esta frase no remite a *una* pared, que necesariamente será azul o no será; la palabra «pared» no tiene, por sí misma, significado preciso (contenido semántico). Aunque, por la peculiaridad del lenguaje natural, «pared» remite a una clase de objetos específicos, intuitivamente podemos entender su significado vacío asimilándolo a un símbolo de un lenguaje formal.[8] Todo esto no elimina su estricta validez sintáctica (de hecho, nada impide que una expresión sea sintácticamente correcta aunque sea falsa en todos las interpretaciones o modelos: esto es, básicamente, una paradoja).

Pasemos al estudio formal de la semántica de un lenguaje formal de primer orden. Primero, necesitamos de una nueva estructura más allá de la sintaxis, para poder interpretar los símbolos del lenguaje.

Definición 2.1.13 (Estructura)**.** Sea **L** un lenguaje formal de primer orden. Una **L**-estructura $\mathfrak{U}$ consiste en:

- Conjunto $|\mathfrak{U}| \neq \emptyset$ denominado **universo** o **dominio** (si no da lugar a cofusión, escribiremos $\mathfrak{U}$).

[7]En el apartado 1.1.4 introdujimos un ejemplo de «modelo» para un sistema de axiomas sencillos, y vimos que, para esa interpretación, éstos eran verdaderos. Así, hicimos un análisis semántico para cierta interpretación de los axiomas.

[8]En el primer apartado hablábamos del filósofo Ludwig Wittgenstein. En Wittgenstein II encontramos una filosofía del lenguaje donde las palabras y las proposiciones adquieren su significado únicamente en su función y su uso en el lenguaje. La validez de una proposición depende del contexto lingüístico al que pertenece, o en la términología de Wittgenstein, del «juego de lenguaje». Por ejemplo, una exposición matemática, un chiste o la lectura de un cuento a un niño son juegos de lenguaje. La validez de las proposiciones que se usen en ellos está encorsetada por la(su) función estructural del(en el) contexto; si se usan fuera de lugar, quedarán sin sentido.

- Para cada símbolo de constante $\mathbf{c} \in \mathbf{L_c}$, un elemento $\mathbf{c}_{\mathfrak{A}} \in |\mathfrak{A}|$, que denominaremos **interpretación** de $\mathbf{c}$ en $\mathfrak{A}$.

$$\mathbf{c} \in \mathbf{L_c} \quad \longrightarrow \quad \mathbf{c}_{\mathfrak{A}} \in |\mathfrak{A}|, \quad \text{interpretación de } \mathbf{c} \text{ en } \mathfrak{A}$$

- Para cada símbolo funcional n-ario $\mathbf{f} \in \mathbf{L_F}$, una aplicación n-aria $\mathbf{f}_{\mathfrak{A}} : |\mathfrak{A}|^n \longrightarrow |\mathfrak{A}|$, que denominaremos **interpretación** de $\mathbf{f}$ en $\mathfrak{A}$.

$$\mathbf{f} \in \mathbf{L_F} \quad \longrightarrow \quad \mathbf{f}_{\mathfrak{A}} : |\mathfrak{A}|^n \longrightarrow |\mathfrak{A}|, \quad \text{interpretación de } \mathbf{f} \text{ en } \mathfrak{A}.$$

- Para cada símbolo de predicado n-ario $\mathbf{p} \in \mathbf{L_P}$, un subconjunto $\mathbf{p}_{\mathfrak{A}} \in |\mathfrak{A}|^n$, que denominaremos **interpretación** de $\mathbf{p}$ en $\mathfrak{A}$.

$$\mathbf{p} \in \mathbf{L_P} \quad \longrightarrow \quad \mathbf{p}_{\mathfrak{A}} \in |\mathfrak{A}|^n, \quad \text{interpretación de } \mathbf{p} \text{ en } \mathfrak{A}.$$

- Interpretación del predicado binario de igualdad $=$: $\quad =_{\mathfrak{A}}$ es $\{(a,a) : a \in |\mathfrak{A}|\}$.

Así pues, podemos designar una $\mathbf{L}$-estructura $\mathfrak{A}$ de la siguiente manera:

$$\mathfrak{A} = \langle |\mathfrak{A}|, \{\mathbf{c}_{\mathfrak{A}} : \mathbf{c} \in \mathbf{L_C}\}, \{\mathbf{f}_{\mathfrak{A}} : \mathbf{f} \in \mathbf{L_F}\}, \{\mathbf{p}_{\mathfrak{A}} : \mathbf{p} \in \mathbf{L_P}\} \rangle$$

Por tanto, la $\mathbf{L}$-estructura fija la interpretación de todos los símbolos del lenguaje excepto de las variables y de las conectivas lógicas. A continuación, definimos la interpretación de las variables, para lo cual se requiere el dominio $|\mathfrak{A}|$.

Definición 2.1.14 (Interpretación de las variables)**.** Una interpretación de las variables[9] en la $\mathbf{L}$-estructura $\mathfrak{A}$ es una aplicación ω de $\mathbb{N}$ en el dominio $|\mathfrak{A}|$:

$$v_i, \quad \text{variable} \quad \longrightarrow \quad \omega(i) \in |\mathfrak{A}|, \quad \text{interpretación de } v_i \text{ en } \mathfrak{A}.$$

Sin entrar en detalle, con esto nos basta para definir los términos del lenguaje sin dificultad: si es una variable o una constante, su interpretación; si es un término compuesto, la sustitución de la interpretación de cada término en la interpretación de la letra funcional.

Definición 2.1.15 (Cardinal de una $\mathbf{L}$-estructura)**.** Sea $\mathbf{L}$ un lenguaje formal de primer orden, y $\mathfrak{A}$ una $\mathbf{L}$-estructura. Definimos el cardinal de $\mathfrak{A}$, $\mathrm{card}(\mathfrak{A})$, como el cardinal de su dominio.

[9] La interpretación de las variables no es estrictamente necesaria para interpretar por completo un lenguaje formal de primer orden. La diferencia entre hacerlo o no es la manera de definir la validez o verdad de las fórmulas. Para un estudio completo de la vía sin interpretar variables, véase (Fernández Margarit, 2012).

Ejemplo 2.1.16. El modelo[10] estándar de la aritmética: $\mathcal{N}$. Recordemos que $\mathbf{LA} = \{\mathbf{0}, \mathbf{S}, +, \cdot, <\}$.

- Dominio: $|\mathfrak{U}| = |\mathcal{N}| = \mathbb{N} \equiv \omega = \{0, 1, 2, ...\}$.
- Interpretaciones:
  - $\mathbf{0} \in \mathbf{LA_C} \quad \longrightarrow \quad \mathbf{0}_{\mathcal{N}} = 0 \in \mathbb{N}$
  - $+ \in \mathbf{LA_F} \quad \longrightarrow \quad +_{\mathcal{N}} : \mathbb{N}^2 \longrightarrow \mathbb{N}$ tal que, para cada $n, m \in \mathbb{N}$, $n +_{\mathcal{N}} m = n + m \in \mathbb{N}$
  - $\mathbf{S} \in \mathbf{LA_F} \quad \longrightarrow \quad \mathbf{S}_{\mathcal{N}} : \mathbb{N} \longrightarrow \mathbb{N}$ tal que, para cada $n \in \mathbb{N}$, $\mathbf{S}_{\mathcal{N}}(n) = n + 1 \in \mathbb{N}$
  - $\cdot \in \mathbf{LA_F} \quad \longrightarrow \quad \cdot_{\mathcal{N}} : \mathbb{N}^2 \longrightarrow \mathbb{N}$ tal que, para cada $(n, m) \in \mathbb{N}^2$, $n \cdot_{\mathcal{N}} m = n \cdot m \in \mathbb{N}$
  - $< \in \mathbf{LA_P} \quad \longrightarrow \quad <_{\mathcal{N}} \subset \mathbb{N}^2$ tal que, $<_{\mathcal{N}} = \{(n, m) \in \mathbb{N}^2 : n < m\}$

Relajando la notación, podemos escribir:

$$\mathcal{N} = \langle \mathbb{N}, 0, +, s, \cdot, < \rangle$$

Añadimos, además, una interpretación de las variables: $\omega : \mathbb{N} \longrightarrow \mathbb{N}$ / $\omega(i) = i$

Podríamos haber definido una interpretación de las variables diferente, como por ejemplo, $\omega(i) = i + 1$, o interpretaciones mucho más complejas. Cada una de ellas acturá diferente sobre las fórmulas. Consideremos la siguiente fórmula de **LA**:

$$v_1 + v_2 = v_3$$

Para la interpretación $\omega(i) = i$ (modelo $\mathcal{N}$), tenemos:

$$1 + 2 = 3$$

, que es verdadero. Sin embargo, para el modelo $\mathcal{N}$ pero cambiando la interpretación de las variables por $\omega_1(i) = i + 1$, nos queda:

$$2 + 3 = 4$$

, que es falso. Podemos encontrar fórmulas que son verdaderas para toda interpretación posible de las variables, como

$$v_1 + \mathbf{0} = v_1$$

[10] Especificaremos qué **L**-estructuras son un modelo más adelante.

Aunque, si cambiásemos la interpretación de la constante $\mathbf{0}$, por ejemplo, a $\mathbf{0}_{\mathcal{N}} = 1 \in \mathbb{N}$ entonces sería siempre falsa.

Nos preguntamos entonces qué criterio de interpretación debemos seguir para definir la *verdad* en una $\mathbf{L}$-estructura. En las siguientes definiciones (que deben ser completadas con los criterios de verdad de los cuantificadores) precisamos este concepto.

Definición 2.1.17 ($\mathfrak{A}$-satisfacible)**.** Decimos que $\varphi \in \mathbf{Form}(\mathbf{L})$ es satisfacible en la estructura $\mathfrak{A}$ si, y sólo si, existe una interpretación ω de las variables para la cual φ es verdadero en $\mathfrak{A}$. Lo denotamos $\mathfrak{A} \vDash \varphi[\omega]$.

Podríamos establecer un criterio formal para establecer cuándo $\varphi \in \mathbf{Sent}(\mathbf{L}(\mathfrak{A}))$, una vez interpretadas las variables, es verdadero en $\mathfrak{A}$. Efectivamente, por recursión sobre la longitud de $\varphi \in \mathbf{Sent}(\mathbf{L}(\mathfrak{A}))$, decimos que φ es verdadero en $\mathfrak{A}$ si, y sólo si

- $\mathfrak{A}(\mathbf{t_1}) = \mathfrak{A}(\mathbf{t_2})$[11], si φ es $\mathbf{t_1} = \mathbf{t_2}$,
- $(\mathfrak{A}(\mathbf{t_1}), ..., \mathfrak{A}(\mathbf{t_n})) \in \mathbf{p}_{\mathfrak{A}}$, si φ es $\mathbf{p}(\mathbf{t_1}, ..., \mathbf{t_n})$,
- ψ no es verdadero en $\mathfrak{A}$, si φ es $\neg\psi$,
- ψ es verdadero en $\mathfrak{A}$ o θ es verdadero en $\mathfrak{A}$, si φ es $\psi \vee \theta$,
- existe $a \in \mathfrak{A}$ tal que $\psi_x[\mathbf{c}_a]$[12] es verdadero en $\mathfrak{A}$, si φ es $\exists x\psi$.

La diferencia clave que debemos comprender es que, una vez interpretadas las variables, a través del método formal anterior, o a través de las ideas usuales en ciertas interpretaciones (como podría ser el modelo estándar de la aritmética, donde podemos decir que $2 + 1 = 10$ es «falso», o que $2 + 1 = 3$ es «verdadero», etc), conocemos ya el concepto de «verdad». Es a través de este concepto relativo a la *interpretación* que establecemos la $\mathfrak{A}$-satisfacibilidad: como aquellas fórmulas que son verdaderas para una interpretación de las variables.

Definición 2.1.18 ($\mathfrak{A}$-verdadera)**.** Decimos que $\varphi \in \mathbf{Form}(\mathbf{L})$ es verdadera en $\mathfrak{A}$ si, y sólo si, para toda interpretación ω de las variables, φ es verdadero en $\mathfrak{A}$. Lo denotamos $\mathfrak{A} \vDash \varphi$.

De nuevo, esta definición no es circular. A través de la interpretación de las variables sí podemos dilucidar la veracidad o falsedad de una fórmula. En consecuencia, cuando en la segunda parte de la definición decimos «verdadero en $\mathfrak{A}$», estamos tratando con una fórmula cuyas variables ya hemos interpretado, y no debemos confundirnos con la definición estricta de fórmula $\mathfrak{A}$-verdadera.

[11] $\mathfrak{A}(\mathbf{t})$ es la interpretación del término $\mathbf{t}$.

[12] A cada $a \in \mathfrak{A}$ podemos asignarle un nuevo símbolo de constante $\mathbf{c}_a$ llamado «nombre de a». Al lenguaje $\mathbf{L}$ al que añadimos estos nuevos símbolos lo denotamos por $\mathbf{L}(\mathfrak{A})$.

Definición 2.1.19 (Lógicamente verdad)**.** Decimos que $\varphi \in \mathbf{Form}(\mathbf{L})$ es verdadera, o lógicamente verdadera, si, y sólo si, es $\mathfrak{U}$-verdadera, $\mathfrak{U} \vDash \varphi$, para toda **L**-estructura $\mathfrak{U}$. Lo denotamos $\vDash \varphi$.

Ejemplo 2.1.20. Con $\mathcal{N}$ el modelo estándar de la aritmética, detallado en el ejemplo 2.1.16:

- $v_1 + \mathbf{0} = v_1$ es $\mathcal{N}$-verdadera, pero no es verdadera.
- $v_1{+}v_2 = v_3$ es $\mathcal{N}$-satisfacible, pero no es $\mathcal{N}$-verdadera. En particular, $\mathcal{N} \vDash \varphi[\omega]$. Pero para $\omega_1(i) = i{+}1$, 2+3=4 es falso, luego $\mathcal{N} \nvDash \varphi[\omega_1]$, y por lo tanto $\mathcal{N} \nvDash \varphi$.

Para completar la semántica de los lenguajes de primer orden, establecemos el criterio de verdad de las conectivas lógicas:

Definición 2.1.21. Sea $\mathfrak{U}$ una **L**-estructura, y sea $\varphi \in Sent(\mathbf{L})$. Tenemos que:

- Si φ es $\neg\psi$, entonces $\quad \mathfrak{U} \vDash \varphi \quad \Longleftrightarrow \quad \mathfrak{U} \nvDash \psi$
- Si φ es $\psi \vee \theta$, entonces $\quad \mathfrak{U} \vDash \varphi \quad \Longleftrightarrow \quad \mathfrak{U} \vDash \psi \quad$ ó $\quad \mathfrak{U} \vDash \theta$
- El criterio de verdad para los cuantificadores es un poco más laborioso. Intuitivamente, si φ es $\forall x\psi$, entonces $\quad \mathfrak{U} \vDash \varphi \quad \Longleftrightarrow \quad$ para todo $u \in \mathfrak{U}$, $\ \mathfrak{U} \vDash \varphi(u)$. Sin embargo, debemos dar una explicación más rigurosa en términos de la interpretación de las variables definida en 2.1.14:

  Buscamos establecer el valor de verdad de la sentencia $\forall v_i \varphi(v_i)$. Consideremos fijada la **L**-estructura $\mathfrak{U}$, y una interpretación de las variables ω, que atribuye el valor $\omega(i) \in |\mathfrak{U}|$ a la variable v_i. Dado que debemos establecer la verdad de $\varphi(v_i)$ *para todo* v_i, necesitamos atribuirle *todos* los valores posibles de v_i. No nos basta, pues, con la interpretación ω, sino que precisamos de sus variantes ω_i^u, definidas de la siguiente manera:

  $$\omega_i^u(j) = \begin{cases} \omega(j) & , \text{si} j \neq i \\ u & , \text{si} j = i \end{cases}$$

  Por tanto, para el índice i correspondiente a la variable v_i, fijo, tomamos todos los valores $u \in |\mathfrak{U}|$, esto es, recurrimos a todas las interpretaciones posibles ω_i^u cuando u recorre el dominio $|\mathfrak{U}|$. De esta manera las otras variables no se ven alteradas (v_j, con $j \neq i$, se interpretará $\omega_i^u(j) = \omega(j)$, como siempre). En consecuencia:

  $$\mathfrak{U} \vDash \forall v_i \varphi(v_i) \quad \Longleftrightarrow \quad \mathfrak{U} \vDash \varphi(v_i)[\omega_i^u], \quad \forall u \in \mathfrak{U}$$

Esto es: $\forall v_i \varphi(v_i)$ es verdadero en la **L**-estructura $\mathfrak{U}$ si, y sólo si, dando a la variable v_i *todos* los valores u de $\mathfrak{U}$, dejando fijos los valores de las otras variables, $\varphi(u)$ es verdadero en $\mathfrak{U}$. Y vemos que esto es, efectivamente, lo que intuitivamente habíamos expresado como criterio de verdad para el cuantificador universal.

Explicitamos también el criterio de verdad del cuantificador existencial: $\exists v_i \varphi(v_i)$ es verdadero en la **L**-estructura $\mathfrak{U}$ si, y sólo si, dando a la variable v_i *un* valor u de $\mathfrak{U}$, dejando fijos los valores de las otras variables, $\varphi(u)$ es verdadero en $\mathfrak{U}$.

Ejemplo 2.1.22. Veamos tres ejemplos para que quede claro el criterio de verdad de los cuantificadores:

1. $\mathcal{N} \models \forall v_1 \exists v_2 (v_2 = \mathbf{S}(v_1))$. Damos *un valor cualquiera* $n \in \mathbb{N}$ a v_1; entonces, como $s(n) = n + 1 \in \mathbb{N}$, dando a la variable v_2 *el valor* $n + 1$, tenemos que $\forall v_1 \exists v_2 (v_2 = \mathbf{S}(\mathbf{v_1}))$ es verdadera en $\mathcal{N}$.

2. $\mathcal{N} \models \forall v_1 \forall v_2 \exists v_3 (v_1 + v_2 = v_3)$. Damos a v_1 y a v_2 valores cualesquiera $n, m \in \mathbb{N}$. Si damos a la variable v_3 el valor $m+n = p \in \mathbb{N}$, que existe, pues es el resultado de la operación suma $+_{\mathcal{N}}(m, n) = m + n$, entonces $\forall v_1 \forall v_2 \exists v_3 (v_1 + v_2 = v_3$ es verdadero en $\mathcal{N}$.

3. $\mathcal{N} \not\models \exists v_1 \forall v_2 (v_1 = \mathbf{S}(v_2))$. No nos es posible interpretar v_1 de ninguna manera tal que, para cualquier interpretación $n \in \mathbb{N}$ de v_2, sea cierto que esa interpretación es igual a $s(n)$, pues si es igual a $s(n_1)$, entonces no lo es a $s(n_2)$, si $n_1 \neq n_2$. En resumen, no existe ningún elemento de $\mathbb{N}$ tal que sea el siguiente de *todos* los elementos de $\mathbb{N}$.

2.2. Teorías formales de primer orden

Vamos a introducir el concepto de **valoración de verdad**, que va a sustituir a las estructuras, para así poder asignar un valor de verdad, que será 0 o 1, a las fórmulas.

Definición 2.2.1 (Funciones de validez)**.** Definimos dos funciones de validez:

- $H_\neg : \{0,1\} \longrightarrow \{0,1\}$ es la aplicación definida por:

  $$H_\neg(a) = \begin{cases} 1 & \text{, si a=0} \\ 0 & \text{, si a=1.} \end{cases}$$

- $H_\vee : \{0,1\}^2 \longrightarrow \{0,1\}$ es la aplicación definida por:

  $$H_\vee(a,b) = \begin{cases} 0 & \text{, si a=b=0} \\ 1 & \text{, en caso contrario.} \end{cases}$$

Definición 2.2.2 (Valoración de verdad)**.** Diremos que una aplicación $\sigma : \mathbf{Form}(\mathbf{L}) \longrightarrow \{0,1\}$ es una valoración de verdad del lenguaje **L** si, para toda $\varphi \in \mathbf{Form}(\mathbf{L})$:

$\sigma(\neg\varphi) = H_\neg(\sigma(\varphi))$,
$\sigma(\varphi \vee \psi) = H_\vee(\sigma(\varphi), \sigma(\psi))$.

Y es sencillo probar la siguiente proposición que relaciona la veracidad en una **L**-estructura con la valoración de verdad.

Proposición 2.2.3. Sea $\mathfrak{U}$ una **L**-estructura. Existe una valoración de verdad de **L**, $\sigma_{\mathfrak{U}}$, tal que para toda $\varphi \in \mathbf{Sent}(\mathbf{L})$,

$$\sigma_{\mathfrak{U}}(\varphi) = 1 \iff \mathfrak{U} \models \varphi$$

A continuación establecemos los conceptos necesarios para definir una teoría, a saber, los axiomas y las reglas de inferencia.

2.2.1. Axiomas Lógicos y Reglas de Inferencia

Los axiomas lógicos son los axiomas comunes a todos los lenguajes de primer orden. Posteriormente veremos que cada teoría tiene sus propios axiomas *no lógicos*.

Definición 2.2.4 (Tautología)**.** Diremos que $\varphi \in \mathbf{Form}(\mathbf{L})$ es una tautología si, para toda valoración de verdad σ, $\sigma(\varphi) = 1$.

Definición 2.2.5 (Axiomas Lógicos)**.** Sea **L** un lenguaje de primer orden. Los axiomas lógicos de **L** son:

1. Axiomas proposicionales: Las tautologías son los axiomas proposicionales.

2. Axiomas predicativos: Sean $\varphi, \psi \in \mathbf{Form}(\mathbf{L})$, $\mathbf{t} \in \mathbf{Term}(\mathbf{L})$.

   - Ax. de clase 1: $\forall x(\psi \to \varphi) \to (\psi \to \forall x\varphi)$, si $x \notin \mathbf{Vl}(\psi)$.
   - Ax. de clase 2 o de sustitución: $\varphi_x[\mathbf{t}] \to \exists x\varphi$ [13].

3. Axioma de identidad: Para cualquier variable x, $x = x$.

4. Axiomas de igualdad:

   $x_1 = y_1 \wedge ... \wedge x_n = y_n \to \mathbf{f}(x_1, ..., x_n) = \mathbf{f}(y_1, ..., y_n)$
   $x_1 = y_1 \wedge ... \wedge x_n = y_n \to (\mathbf{p}(x_1, ..., x_n) = \mathbf{p}(y_1, ..., y_n))$

[13] x es sustituible por $\mathbf{t}$ en φ. Otra forma de verlo: $\forall\varphi(x) \to \varphi(\mathbf{t})$, si $x \notin \mathbf{Vl}(\mathbf{t})$.

Entre los axiomas proposicionales encontramos los axiomas de la lógica de orden cero, también llamada lógica proposicional. En particular, se heredan los siguientes tres axiomas que es conveniente explicitar:

Sean $\varphi, \psi . \gamma \in \mathbf{Forml(L)}$.

1. Ax. prop. 1: $\varphi \to (\gamma \to \varphi)$
2. Ax. prop. 2: $(\varphi \to (\gamma \to \psi)) \to ((\varphi \to \gamma) \to (\varphi \to \psi))$
3. Ax. prop. 3: $(\neg\varphi \to \neg\gamma) \to ((\neg\varphi \to \gamma) \to \varphi)$

Por otra parte, un axioma de igualdad que también conviene escribir es el siguiente:

Ax. igualdad: $\forall x \forall y (x = y \to (\varphi(x) \to \varphi(y))$

Es decir, que siempre podemos cambiar una variable por otra que sea igual que ella.

Teorema 2.2.6. Los axiomas lógicos de un lenguaje de primer orden $\mathbf{L}$ son fórmulas lógicamente válidas.

Demostración: Vamos a probarlo para los axiomas proposicionales. Sea $\mathfrak{U}$ una $\mathbf{L}$-estructura y $\psi(x_1, ..., x_n)$ una tautología. Veamos que $\mathfrak{U} \vDash \psi$. Para ello, tomemos $a_1, ...a_n \in |\mathfrak{U}|$, interpretaciones cualesquiera de las variables, y veamos que $\mathfrak{U} \vDash \psi(a_1, ..., a_n)$.

Probemos primero que si φ es una tautología, entonces $\varphi_x[\mathbf{t}]$ es una tautología. Si no lo fuera, entonces existiría una valorción de verdad σ tal que $\sigma(\varphi_x[\mathbf{t}]) = 0$. No es difícil ver que la aplicación $\sigma_{x,\mathbf{t}}$ definida por $\sigma_{x,\mathbf{t}}(\varphi) = \sigma(\varphi_x[\mathbf{t}])$ es también una valoración de verdad. En tal caso, tenemos que $\sigma_{x,\mathbf{t}}(\varphi) = \sigma(\varphi_x[\mathbf{t}]) = 0$, luego tendríamos que φ no es una tautología, lo cual es una contradicción. Visto que si φ es una tautología, entonces $\varphi_x[\mathbf{t}]$ es una tautología, tenemos por hipótesis que $\psi(a_1, ..., a_n)$ es una tautología. Entonces para la valoración de verdad $\sigma_{\mathfrak{U}}$, $\sigma_{\mathfrak{U}}(\psi(a_1, ..., a_n)) = 1$, luego por la proposición 2.2.3, $\mathfrak{U} \vDash \psi$. □

Definición 2.2.7 (Reglas de inferencia)**.** La lógica de primer orden tiene dos reglas de inferencia. La primera es el Modus Ponens, que se hereda de la lógica proposicional. La segunda es la regla de Introducción del cuantificador existencial, característica de la lógica de primer orden.

1. Modus Ponens (MP): $\dfrac{\varphi, \varphi \to \psi}{\psi}$

2. Introducción del cuantificador existencial (R∃): $\dfrac{\varphi \to \psi}{\exists x \varphi \to \psi}$, si $x \notin \mathbf{Vl}(\psi)$[14].

Teorema 2.2.8. Si las hipótesis de una regla de inferencia son $\mathfrak{U}$-verdaderas, entonces también lo es la conclusión.

2.2.2. Teorías y pruebas

Recogemos todos los conceptos expresados anteriormente para poder definir con rigor una teoría formal de primer orden.

Definición 2.2.9 (Teoría)**.** Una teoría de primer orden, $\mathbf{T}$, consta de:

1. Un lenguaje de primer orden, $\mathbf{L}(\mathbf{T})$.
2. Axiomas:
   - Axiomas lógicos (definición 2.2.5).
   - Axiomas no lógicos: $\mathbf{Ax}(\mathbf{T}) \subset \mathbf{Form}(\mathbf{L}(\mathbf{T}))(\equiv \mathbf{Form}(\mathbf{T}))$.
3. Reglas de inferencia: MP y R∃ (definición 2.2.7).

Ejemplo 2.2.10. En ejemplos anteriores ya hemos tratado el lenguaje de la aritmética, $\mathbf{LA} = \{\mathbf{0}; +, \times, \mathbf{S}; =\}$, y el modelo estándar $\mathcal{N} = \langle \mathbb{N}, 0, +, s, \cdot, < \rangle$. A continuación damos los axiomas no lógicos, con lo que ya tenemos definida la artimética de Peano de primer orden, que a partir de ahora denominaremos teoría $\mathbf{P}$. Esta es la teoría formal con la que trabaja Gödel.

$\mathbf{Ax}(\mathbf{P}) = \{P_1, P_2, P_3, P_4, P_5, P_6, P_7, P_8\}$

P1: $\forall x(\mathbf{0} \neq \mathbf{S}(x))$.

P2: $\forall x \forall y(\mathbf{S}(x) = \mathbf{S}(y) \to x = y)$.

P3: $\forall x \forall y(x = y \to \mathbf{S}(x) = \mathbf{S}(y))$.

P4: $\forall x(\mathbf{0} + x = x)$.

P5: $\forall x \forall y(x + \mathbf{S}(y) = \mathbf{S}(x + y))$.

P6: $\forall x(\mathbf{0} \cdot x = \mathbf{0})$.

P7: $\forall x \forall y(x \cdot \mathbf{S}(y) = (x \cdot y) + x)$.

[14] O, en términos del cuantificador universal, tendríamos la regla de Generalización universal R∀, $\dfrac{\varphi}{\forall x \varphi}$, si $x \notin \mathbf{Vl}(\varphi)$.

P8: $\forall\varphi(x) \in \mathbf{Form}(\mathbf{P})$, $\varphi(\mathbf{0}) \wedge \forall x(\varphi(x) \to \varphi(\mathbf{S}(x))) \longrightarrow \forall x\varphi(x)$.

El octavo axioma es el axioma de inducción. En esta formulación de la aritmética de Peano podemos entender bien el término «primer orden» para las teorías que estamos manejando. En la formulación original, Peano escribía el axioma de inducción en relación a un subconjunto de $\mathbb{N}$: para $K \subseteq \mathbb{N}$, si $1 \in K$ y $n \in K$, entonces $n+1 \in K$. Y es que Peano da por sentada la existencia del conjunto de los números naturales. Sin embargo, en lógica de primer orden solamente podemos cuantificar sobre los «objetos» descritos por el lenguaje formal (de primer orden), en el caso de **LA**, los números naturales, y no sobre colecciones de los mismos, o sea, sobre conjuntos de números naturales (que serían variables de «segundo orden»). Al escribir el axioma en lógica de primer orden necesitamos establecer una infinidad de hipótesis, una para cada fórmula de **P** con una variable libre. Además la formulación de primer orden es más débil que la original de Peano (véase [30], p. 171, para una explicación detallada).

Definición 2.2.11 (Demostración)**.** Una **T**-demostración (una demostración en la teoría **T**) es una sucesión finita de fórmulas de $\mathbf{L}(\mathbf{T})$, $\varphi_1, \varphi_2, ...\varphi_n$ tal que, $\forall i \in \{1, 2, ..., n\}$, φ_i cumple una de las siguientes condiciones

- Es un axioma, lógico o no lógico, de **T**.
- Se obtiene a partir de φ_j, φ_k, $k, j < i$, por MP, donde φ_j es $\varphi_k \to \varphi_i$.
- Se obtiene aplicando la regla del cuantificador existencial R$\exists$ a φ_j, $j < i$, que es $\psi \to \theta$, donde φ_i es $\exists x\psi \to \theta$, y donde x no es libre en θ.

Definición 2.2.12 (Teorema)**.** Diremos que $\varphi \in \mathbf{Form}(\mathbf{T})$ es un teorema de la teoría **T**, $\mathbf{T} \vdash \varphi$, si existe una demostración $\varphi_1, \varphi_2, ...\varphi_n$ en **T** tal que φ_n es φ.

En este sentido es claro que, una demostración $\mathbf{L}(\mathbf{T})$, $\varphi_1, \varphi_2, ...\varphi_n$, $\forall i \in \{1, 2, ..., n\}$, φ_i debe cumplir, o bien una de las condiciones de la definición 2.2.11, o bien ser un teorema.

Definición 2.2.13 (Modelo)**.** Una $\mathbf{L}(\mathbf{T})$-estructura $\mathfrak{U}$ es un modelo de la teoría **T**, $\mathfrak{U} \vDash \mathbf{T}$, si todo $\varphi \in \mathbf{Ax}(\mathbf{T})$ es $\mathfrak{U}$-verdadero, o sea, $\mathfrak{U} \vDash \varphi$. Lo indicaremos también como $\mathfrak{U} \in \mathbf{Mod}(\mathbf{T})$, la clase de modelos de **T**, o como $\mathfrak{U}_{\mathbf{T}}$.

Podríamos redefinir esto de la siguiente manera: Una $\mathbf{L}(\mathbf{T})$-estructura $\mathfrak{U}$ es un modelo de $\varphi \in \mathbf{Sent}(\mathbf{T})$ si φ es $\mathfrak{U}$-verdadera. Así, una $\mathbf{L}(\mathbf{T})$-estructura $\mathfrak{U}$ es un modelo de la teoría **T** si es un modelo de todo $\varphi \in \mathbf{Ax}(\mathbf{T})$, o equivalentemente, si es un modelo de toda $\varphi \in \mathbf{Sent}(\mathbf{T})$.

Definición 2.2.14 (Validez, Verdad)**.** $\varphi \in \mathbf{Form}(\mathbf{T})$ es válida en la teoría **T**, $\mathbf{T} \vDash \varphi$, si para todo modelo $\mathfrak{U}$ ($\mathfrak{U} \vDash \mathbf{T}$), φ es $\mathfrak{U}$-verdadera ($\mathfrak{U} \vDash \varphi$).

Una vez que hemos definido los conceptos fundamentales de las teorías formales, pasamos a definir las propiedades sobre los sistemas formales que Hilbert se había planteado, y que entran en el ámbito de la metalógica: la consistencia, la completitud y la decibilidad.

Definición 2.2.15 (Decibilidad)**.** Una teoría formal de primer orden $\mathbf{T}$ es decidible si el conjunto de sus teoremas es recursivo. Es decir, existe un algoritmo que determina para toda fórmula si es un teorema o no.

Definición 2.2.16 (Consistencia)**.** Una teoría formal de primer orden $\mathbf{T}$ es consistente si no existe una fórmula φ tal que $\mathbf{T} \vdash \varphi$ y $\mathbf{T} \vdash \neg\varphi$.

Podemos dar varias definiciones equivalentes para una teoría consistente, como podemos ver en el siguiente teorema.

Teorema 2.2.17. Sea $\mathbf{T}$ una teoría. Son equivalentes:

1. $\mathbf{T}$ es inconsistente.

2. Para toda fórmula φ de $\mathbf{L}(\mathbf{T})$, $\mathbf{T} \vdash \varphi$.

3. $\mathbf{T} \vdash x \neq x$.

Demostración: Es trivial, excepto la implicación 1) $\to$ 2) [15]. Sea $\phi \in \mathbf{Form}(\mathbf{L}(\mathbf{T}))$ tal que $\mathbf{T} \vdash \phi$ y $\mathbf{T} \vdash \neg\phi$. Veamos que para toda fórmula φ, $\mathbf{T} \vdash \varphi$:

ϕ	Teorema.
$\neg\phi$	Teorema.
$(\neg\varphi \to \phi) \to ((\neg\varphi \to \neg\phi) \to \varphi)$	Ax. Prop. 3.
$\neg\varphi \to \phi$	Teorema.
$(\neg\varphi \to \neg\phi) \to \varphi$	MP 4 y 3.
$\neg\varphi \to \neg\phi$	Teorema.
φ	MP 6 y 5.

En las líneas 4 y 6 tenemos que son teoremas sin más que aplicar el Ax. Prop. 1.

□

Definición 2.2.18 (Completitud)**.** Una teoría consistente $\mathbf{T}$ es completa si $\forall\varphi \in \mathbf{Sent}(\mathbf{L}(\mathbf{T}))$, se tiene que, o bien $\mathbf{T} \vdash \varphi$, o bien $\mathbf{T} \vdash \neg\varphi$.

Para la definición de teoría completa nos hemos restringido a las fórmulas cerradas del lenguaje de la teoría por la siguiente razón. Supongamos que tenemos una teoría consistente y completa, y consideremos la fórmula (no cerrada) $x = y$. Del

[15] Habitualmente se define la inconsistencia de una teoría como aquella en la que toda fórmula es un teorema.

teorema 2.2.17 obtenemos el corolario de que si $\mathbf{T} \vdash x \neq y$ entonces $\mathbf{T}$ es inconsistente sin más que aplicar la regla de sustitución para teoremas [16] a $\mathbf{T} \vdash x \neq x$. Pero, como $\mathbf{T}$ es consistente, entonces tenemos que $\mathbf{T} \vdash x = y$. Como veremos en el teorema de completitud (nada nos impediría definir esta parte con antelación a esta definición), esto implica que φ es verdadera en $\mathbf{T}$ ($\mathbf{T} \vDash \varphi$), por lo que es verdadera para todo modelo ($\mathfrak{U} \vDash \mathbf{T}$) de $\mathbf{T}$: $\mathfrak{U} \vDash x = y$. Pero esto implica que $\operatorname{card}(\mathfrak{U}) = 1$ (recordemos que esto es, por definición, el cardinal del dominio). En consecuencia, si en la definición de completitud de una teoría no nos restringimos a las sentencias, entonces los modelos de una teoría completa solamente tendrían un elemento.

2.3. Metateoremas relativos a los modelos y a la verdad

Nos planteamos ya en esta sección el verdadero problema en torno al teorema de Gödel; a saber, la relación entre la verdad (semántica) y la demostración (formal, sintáctica) en las teorías formales de primer orden.

2.3.1. Las colecciones de sentencias «verdaderas»

Hemos definido el concepto de verdad relativo a: 1) una $\mathbf{L}$-estructura $\mathfrak{U}$ en la definición 2.1.18, 2) de forma similar, a una $\mathbf{L}$-estructura $\mathfrak{U}$ que es un modelo de una teoría $\mathbf{T}$ ($\mathfrak{U} \in \mathbf{Mod}(\mathbf{T})$, $\mathfrak{U}_{\mathbf{T}}$), 3) todos los modelos $\mathbf{Mod}(\mathbf{T})$ en la definición 2.2.14, y 4) todas las $\mathbf{L}$-estructuras $\mathfrak{U}$ en la definición 2.1.19. Ahora bien, notemos que, fijada una $\mathbf{L}$-estructura, al carecer de variables libres, las sentencias $\varphi \in \mathbf{Sent}(\mathbf{L})$ siempre son verdaderas o falsas, y no queda lugar a la ambigüedad. De esta forma podemos escribir las siguientes cuatro clases de sentencias que definen los cuatro conceptos de verdad:

1. $\mathcal{V}(\mathfrak{U}) = \{\varphi \in \mathbf{Sent}(\mathbf{L}) : \mathfrak{U} \vDash \varphi\}$,

2. $\mathcal{V}(\mathfrak{U}_{\mathbf{T}}) = \{\varphi \in \mathbf{Sent}(\mathbf{L}) : \mathfrak{U}_{\mathbf{T}} \vDash \varphi\}$,

3. $\mathcal{V}(\mathbf{Mod}(\mathbf{T})) = \{\varphi \in \mathbf{Sent}(\mathbf{L}) : \mathfrak{U}_{\mathbf{T}} \vDash \varphi,\ \forall \mathfrak{U}_{\mathbf{T}}\} = \{\sigma \in \mathbf{Sent}(\mathbf{L}) : \mathbf{T} \vDash \varphi\}$,

4. $\mathcal{V} = \bigcap\{\mathcal{V}(\mathfrak{U}) : \mathfrak{U} \text{ es una } \mathbf{L}\text{-estructura }\} = \{\varphi \in \mathbf{Sent}(\mathbf{L}) : \ \vDash \varphi\}$.

[16] Aunque no lo hemos explicitado, de las reglas de inferencia definidas para fórmulas se deducen las mismas ideas para teoremas. R$\forall$: si $\mathbf{T} \vdash \varphi \to \psi$ y $x \notin \mathbf{Vl}(\varphi)$, entonces $\mathbf{T} \vdash \varphi \to \forall x \psi$. Regla de generalización: si $\mathbf{T} \vdash \varphi$, entonces $\mathbf{T} \vdash \forall x \varphi$. Regla de sustitución: Si $\mathbf{T} \vdash \varphi$, entonces $\mathbf{T} \vdash \varphi_{x_1,\ldots,x_n}[t_1 \ldots t_n]$.

Estableciendo estas colecciones de sentencias podemos indagar si cumplen o no las tesis exigidas por el programa de Hilbert. Es decir, nos preguntamos por su completitud y consistencia [17].

Tratemos primero las propiedades metamatemáticas de la colección de sentencias verdaderas en la **L**-estructura $\mathfrak{U}$, $\mathcal{V}(\mathfrak{U}) = \{\varphi \in \mathbf{Sent}(\mathbf{L}) : \mathfrak{U} \vDash \varphi\}$.

- **Completitud** de $\mathcal{V}(\mathfrak{U})$: Por el carácter peculiar de las sentencias, el no contener variables libres, es obvio que, o bien $\varphi \in \mathbf{Sent}(\mathbf{L})$ es $\mathfrak{U}$-verdadera ($\mathfrak{U} \vDash \varphi$), esto es, es verdadera para toda interpretación de las variables, o bien hay una interpretación ω para la que φ es falsa. Esto último implica que $\mathfrak{U} \nvDash \varphi$, luego $\mathfrak{U} \vDash \neg\varphi$.

- **Consistencia** de $\mathcal{V}(\mathfrak{U})$: Es obvio, pues para una sentencia no es posible que se den a la vez $\mathfrak{U} \vDash \varphi$ y $\mathfrak{U} \vDash \neg\varphi$, pues si una es cierta, la otra es falsa.

Queda claro entonces que el programa de Hilbert se cumple para las colecciones de sentencias verdaderas en una **L**-estructura (en particular en un modelo). Si $\mathcal{V}(\mathfrak{U}_\mathbf{T})$ coincidiese con el conjunto de los teoremas de **T**, el programa de Hilbert se cumpliría, pues toda verdad sería demostrable, y todo teorema sería verdadero en el modelo, y además la axiomatización sería consistente. Sin embargo, no conocemos por el momento ninguna relación entre $\mathcal{V}(\mathfrak{U}_\mathbf{T})$ y los toeremas de **T**.

Analicemos a continuación la colección de sentencias $\mathcal{V}(\mathbf{Mod}(\mathbf{T}))$:

- **Completitud** de $\mathcal{V}(\mathbf{Mod}(\mathbf{T}))$: Es claro que para cada $\mathfrak{U}_\mathbf{T} \in \mathbf{Mod}(\mathbf{T})$, $\mathcal{V}(\mathbf{Mod}(\mathbf{T})) \subseteq \mathcal{V}(\mathfrak{U}_\mathbf{T})$. Es pues una colección de sentencias menor que la anterior, pues estamos abarcando solamente aquellas que son verdaderas en *todos* los modelos. Y en consecuencia tenemos que $\mathcal{V}(\mathbf{Mod}(\mathbf{T}))$ no es completa, pues no es cierto que dada una sentencia necesariamente $\mathfrak{U}_\mathbf{T} \vDash \varphi$, $\forall \mathfrak{U}_\mathbf{T} \in \mathbf{Mod}(\mathbf{T})$, Habrás cierto modelos en los que seaverdadera y otros en los que sea falsa. Por ejemplo, en teoría de grupos, el axioma de conmutatividad no es cierto, ni tampoco lo es su negación, en la colección de sentencias verdaderas en todos los modelos posibles, pues hay grupos conmutativos y no conmutativos.

- **Consistencia** de $\mathcal{V}(\mathbf{Mod}(\mathbf{T}))$: Por contra, tenemos que la consistencia de $\mathcal{V}(\mathfrak{U})$, en particular de cada $\mathcal{V}(\mathfrak{U}_\mathbf{T})$, implica la consistencia de $\mathcal{V}(\mathbf{Mod}(\mathbf{T}))$. En efecto, si $\mathcal{V}(\mathbf{Mod}(\mathbf{T}))$ no fuese consistente, entonces exitiría una sentencia φ tal que $\varphi, \neg\varphi \in \mathcal{V}(\mathbf{Mod}(\mathbf{T}))$, y entonces para cada modelo $\mathfrak{U}_\mathbf{T}$ tendríamos

[17] Estrictamente hablando hemos definidos estos conceptos para teorías formales de primer orden, y no para colecciones de sentencias. Aquí estamos tratando con la «completitud semántica», para la que una definición adecuada es que, dada una sentencia $\varphi \in \mathbf{Sent}(\mathbf{L})$, la colección es completa si, y sólo si, o bien $\mathfrak{U} \vDash \varphi$, o bien $\mathfrak{U} \vDash \neg\varphi$. De forma similar, definimos las «consistencia semántica» de la siguiente manera: dada una sentencia $\varphi \in \mathbf{Sent}(\mathbf{L})$, la colección es consistente si, y sólo si, no es posible que se de a la vez que $\mathfrak{U} \vDash \varphi$ y que $\mathfrak{U} \vDash \neg\varphi$.

que $\varphi, \neg\varphi \in \mathcal{V}(\mathfrak{U}_{\mathbf{T}})$, por lo que $\mathcal{V}(\mathfrak{U}_{\mathbf{T}})$ no sería consistente. Ahora bien, todo depende de la consistencia de la propia teoría; si $\mathbf{T}$ no es consistente, entonces

Tratando con estas cuestiones Tarski fue capaz de llegar a un importante teorema que nos asegura que cualquier teorema de una teoría formal de primer orden es verdader en todos los modelos de la teoría.

Teorema 2.3.1. $\mathbf{T} \vdash \varphi \Longrightarrow \mathbf{T} \vDash \varphi$.

2.3.2. Metateoremas lógicos

Vamos a presentar aquí algunos metateoremas lógicos previos a los teoremas de incompletitud de Gödel. Éste, como se ha comentado en secciones anteriores, trata de la relación entre la deducción formal en la sintaxis de un lenguaje $\mathbf{L}$ y la verdad semántica expresable en una $\mathbf{L}$-estructura. Sin embargo, podemos preguntarnos por la relación entre los teoremas de una teoría formal y las fórmulas verdaderas *en todos sus modelos*. El teorema de completitud de Gödel nos dice que los teoremas de una teoría formal de primer orden coinciden exactamente con las fórmulas verdaderas (válidas) en todos los modelos de la misma. Tenemos pues que, en una teoría formal de primer orden, verdad y deducción formal coinciden, luego todo lo que es verdad es alcanzable sintácticamente.

Teorema 2.3.2 (de completitud de Gödel)**.** Sea $\mathbf{T}$ una teoría formal de primer orden. Entonces

$$\mathbf{T} \vdash \varphi \iff \mathbf{T} \vDash \varphi.$$

Demostración: Hagamos la implicación $\Longleftarrow$), resultado que se conoce como teorema de la validez. Utilizamos la definición de Prueba y trabajamos por inducción sobre teoremas. Sea $\varphi \in \mathbf{Form}(\mathbf{L})$. Para ver que $\mathbf{T} \vDash \varphi$ tenemos que probar que φ es verdadera para todo modelo de $\mathbf{T}$. El primer caso es que φ sea un axioma. Si es un axioma lógico, entonces sabemos que es una fórmula lógicamente válida ($\vDash \varphi$). Por lo tanto es $\mathfrak{U}$-verdadera ($\mathfrak{U} \vDash \varphi$) para toda $\mathbf{L}$-estructura $\mathfrak{U}$, en particular para todo modelo de la teoría. Por tanto, es válida: $\mathbf{T} \vDash \varphi$. Si es un axioma no lógico ($\varphi \in \mathbf{Ax}(\mathbf{T})$), entonces, por definición de modelo de una teoría, $\mathbf{T} \vDash \varphi$. El segundo caso es que sea una regla de inferencia. El resultado se deduce de 2.2.8 y de la hipótesis de inducción. □

El teorema de completitud de Gödel fue generalizado por Leon Henkin. Presentamos a continuación el teorema de completitud de Henkin. Nos dice que una teoría es consistente si, y sólo si, tiene un modelo.

Teorema 2.3.3 (de completitud de Henkin)**.** Sea $\mathbf{T}$ una teoría formal de primer orden. Entonces

$$\mathbf{T} \text{ es consistente} \iff \mathbf{T} \text{ tiene un modelo.}$$

La demostración es demasiado farragosa como para exponerla aquí. Puede encontrarla detalladamente en ([12] Fernández Margarit 20122, cap. III). Como decíamos, la versión de Henkin es una generalización del teorema de completitud de Gödel:

Proposición 2.3.4. 2.3.2 $\Longrightarrow$ 2.3.3.

Debemos entender bien que el teorema de completitud no se contradice con el teorema de incompletitud. El primero se aplica a la lógica de primer orden en general, no específicamente a la aritmética de Peano. Establece que una teoría en lógica de primer orden es consistente si, y sólo si, hay un modelo que satisface todas las sentencias verdaderas de la teoría. O en la formulación de Gödel, que las fórmulas válidas en *todos* los modelos de la teoría coinciden con los teoremas formales de la misma. Esto significa que no hay ninguna sentencia verdadera en la teoría (esto es, verdadera en todos los modelos de la misma) que no pueda ser demostrada dentro de esa teoría.

Por otra parte, el teorema de incompletitud de Gödel se refiere a sistemas formales específicos, como la aritmética de Peano, la teoría de conjuntos, etc. En concreto establece que aquellos sistemas suficientemente ricos como para expresar la aritmética de Peano de primer orden son (si son consistentes) incompletos. Pero en ningún caso trata con las fórmulas verdaderas de $\mathbf{T}$ (o lo que es lo mismo, con las fórmulas verdaderas en todos los modelos de una teoría $\mathbf{T}$).

Es llamativo que en la formulación del teorema de completitud recurrimos a *todos* los modelos de una teoría formal. Pensemos por ejemplo en la aritmética de Peano de primer orden, $\mathbf{P}$. Conocemos el modelo estándar (ver ejemplo 2.1.16), pero, ¿cómo son el resto de modelos? La respuesta es que no sabemos cómo pueden ser. De hecho existen los llamados modelos «no estándar» que comentaremos en breve.

En resumen, mientras que el teorema de completitud de Gödel establece la existencia de un modelo para una teoría consistente en lógica de primer orden, los teoremas de incompletitud de Gödel demuestran que ciertos sistemas formales, como la aritmética de Peano de primer orden, no pueden ser completos en sí mismos, y siempre habrá enunciados verdaderos pero no demostrables dentro de esos sistemas.

Un corolario del teorema de completitud es el teorema de compacidad, del que daremos una demostración.

Definición 2.3.5 (Teoría finita)**.** Diremos que una teoría formal de primer orden $\mathbf{T}$ es finita si $\mathbf{Ax}(\mathbf{T})$ es finito.

Teorema 2.3.6 (de compacidad)**.** Sea $\mathbf{T}$ una teoría formal de primer orden. Entonces

$$\mathbf{T} \text{ tiene un modelo} \iff \text{toda parte finita de } \mathbf{T} \text{ tiene un modelo.}$$

Demostración. $\Longrightarrow$) es trivial. Veamos que $\Longleftarrow$). Supongamos que $\mathbf{T}$ no tiene modelos. Entonces por el teorema de completitud, $\mathbf{T}$ es inconsistente. En consecuencia, por el teorema 2.2.17, $\mathbf{T} \vdash x \neq x$. Sea $\mathbf{T}'$ una parte finita de $\mathbf{T}$ tal que $\mathbf{T}' \vdash x \neq x$. De nuevo por 2.2.17, $\mathbf{T}'$ es inconsistente, y por el teorema de completitud, no tiene modelos. □

El teorema de completitud identifica la deducción formal con la verdad semántica, luego no es de extrañar que tengamos un teorema de compacidad sintáctica.

Teorema 2.3.7 (de compacidad sintáctica)**.** Sea $\mathbf{T}$ una teoría formal de primer orden y $\mathbf{T}'$ una parte finita de $\mathbf{T}$. Entonces

$$\mathbf{T} \vdash \varphi \iff \mathbf{T}' \vdash \varphi.$$

Para finalizar esta sección hagamos un breve comentarios sobre los modelos «no estándar» que comentábamos anteriormente. Como idea general, un modelo no estándar contiene elementos adicionales a los «habituales», como pueden ser números infinitos o infinitesimales. En un modelo de $\mathbf{P}$, por ejemplo, tendremos más elementos que los números naturales (que son los elementos que componen el modelo estándar $\mathcal{N}$). De alguna manera proporcionan una perspectiva más amplia, aunque también más «dudosa», sobre los sistemas lógicos.

Veamos con algo de detalle la construcción de un modelo no estándar para la aritmética de Peano de primer orden $\mathbf{P}$. Para ello vamos a usar el teorema de compacidad, que establece que una teoría tiene un modelo si, y sólo si, toda parte finita suya lo tiene, Considerando la aritmética de Peano $\mathbf{P}$, podemos añadir la siguiente serie infinita de axiomas

$$\Psi = \{\mu \in \mathbb{N}, \mu \neq \mathbf{0}, \mu \neq \mathbf{1}, ..., \mu \neq \mathbf{n}^{18}, ...\}.$$

Los axiomas de la nueva teoría serían $\mathbf{Ax}(\mathbf{P}) + \Psi$. El modelo que en que interpretamos μ como $\mathbf{n}+\mathbf{1}$ satisface todos los axiomas de la serie Ψ hasta $\mu \neq \mathbf{n}$ (toda colección de elementos naturales tiene un elemento máximo). Pero entonces es claro que todo subconjunto finito de estos axiomas tiene un modelo: por ejemplo, $\mathcal{N}$. En consecuencia, por el teorema de compacidad, la teoría aritmética modificada con axiomas no lógicos $\mathbf{Ax}(\mathbf{P}) + \Psi$ posee un modelo (luego, por el teorema de completitud, es consistente)

$$\mathcal{H} = \langle H, 0, +, s, \cdot, <, h\rangle.$$

Y este modelo es también un modelo de la aritmética de Peano en el cual se cumplen los axiomas que hemos añadido, luego «hay» un número $h \in H$ distinto de 0, distinto de 1, distinto de 2,... distinto de n,... etc. Es decir, un número h

[18] Recordemos que $\mathbf{n}$ es una abreviatura de $\mathbf{S}(\mathbf{S}(...(\mathbf{S}^{n)}(\mathbf{0})))...))^{n)}$.

mayor que todos y cada uno de los número naturales «estándar». Es lo que antes denominábamos un «número infinito». Y este es un modelo no estándar de **P**.

De hecho el teorema de Löwenheim-Skolem-Tarski nos asegura la existencia de modelos no estándar para cualquier teoría que tenga modelos infinitos (de dominio infinito), en particular para **P**. Nos dice que, en tal caso, existen modelos de cualquier tamaño «suficientemente grande» (no estándar). Esto implica que no existe un único modelo «más grande» que sea el modelo «definitivo» de una teoría formal de primer orden. Por lo tanto, el teorema de Löwenheim-Skolem muestra la existencia de múltiples modelos para una teoría dada, y corrobora lo que decíamos de que «no sabemos» cómo pueden ser los modelos de **P**.

Teorema 2.3.8 (de Löwenheim-Skolem-Tarski)**.** Sea **T** una teoría que tiene modelos infinitos. Entonces para todo cardinal $\kappa \geq \operatorname{card}(\mathbf{L}(\mathbf{T}))$ existe $\mathfrak{U} \models \mathbf{T}$ tal que $\operatorname{card}(\mathfrak{U}) = \kappa$.

2.4. Los teoremas de Gödel

El objetivo de este apartado es la demostración de los teoremas de incompletitud de Gödel. Como dijimos en la introducción no presentaremos una demostración perfecta, pues una exposición detallada y rigorosa exigiría, además de un número excesivo de páginas, un nivel lógico demasiado avanzado para las características de este trabajo. En cualquier caso intentaremos expresar con detalle lo más importante de la demostración.

Como hemos visto, una teoría definida semánticamente, esto es, como el conjunto de las sentencias verdaderas en un modelo determinado, es siempre completa y consistente. Por lo tanto, si queremos que un sistema formal refleje con total exactitud una teoría definida semánticamente, aquel debe ser completo y consistente. Es claro que, si queremos probar la consistencia de cierto sistema formal (como la teoría **P**), en la demostración no debemos utilizar razonamientos más «potentes» o «arriesgados» que los incorporados en tal sistema.

Gödel publica su famoso artículo «Sobre sentencias formalmente indecidibles de Principia Mathematica y sistemas afines» en 1931 ([18] Gödel 1931). En ese momento se pensaba que el sistema formal de Russell y Whitehead en «Principia Mathematica» cumplía con el primer requerimiento del programa formalista de Hilbert, a saber, la construcción de un sistema formal que formalizara completamente la matemática clásica. Los teoremas de incompletitud de Gödel probaban la imposibilidad de completar el programa, probando que cualquier sistema formal que cumpliera ciertas condiciones es incompleto. Esto no impide que se pueda probar la consistencia desde una teoría más fuerte o añadiendo razonamientos más avanzados o controvertidos, aunque esto es de dudosa utilidad. Por ejemplo, el matemático alemán Gerhard Gentzen demostró, relativamente, la consistencia de la aritmética de primer orden. Para ello utilizó la aritmética recursiva primitiva junto al principio de inducción transfinita.

Para entender el teorema, vamos a diferenciar cinco secciones que recogen las partes esenciales del artículo de Gödel. Primero, una introducción de la mano del propio Gödel; segundo, la parte donde se explica la gödelización; tercero, el estudio de la teoría recursiva, fundamental para entender la representabilidad, lo que constituye la sección cuarta. Por último, lo que en el artículo de Gödel constituye la cuarta parte, donde se prueban los teoremas.

2.4.1. Introducción de Gödel

En una primera parte Gödel hace un resumen de su demostración, destacando las ideas clave. Consideramos adecuado adjuntar aquí ese resumen, al que añadiremos notas al pie para resaltar y comentar algunas cosas (omitimos las notas al pie de Gödel).

> Como es bien sabido, el progreso de la matemática hacia una exactitud cada vez mayor ha llevado a la formalización de amplias partes de ella, de tal modo que las deducciones pueden llevarse a cabo según unas pocas reglas mecánicas. Los sistemas formales más amplios construidos hasta ahora son el sistema de *Principia Mathematica (PM)* y la teoría axiomática de conjuntos de Zermelo-Fraenkel (desarrollada aún más por J. von Neumann).
> Estos dos sistemas son tan amplios que todos los métodos usados hoy en la matemática pueden ser formalizados en ellos, es decir, pueden ser reducidos a unos pocos axiomas y reglas de inferencia. Resulta por tanto natural la conjetura de que estos axiomas y reglas basten para decidir *todas* las cuestiones matemáticas que puedan ser formuladas en dichos sistemas. En lo que sigue se muestra que esto no es así, sino que, por el contrario, en ambos sistemas hay problemas relativamente simples de la teoría de los número reales[19] que no pueden ser decididos con sus axiomas (y reglas). Este hecho no se debe a la especial naturaleza de los sistemas citados, sino que se da en una clase muy amplia de sistemas formales, a la que en especial pertenecen todos los que resultan de añadir un número finito de axiomas a los dos sistemas citados, suponiendo que ningún enunciado falso del tipo indicado en la nota 4 resulte deducible por el añadido de los nuevos axiomas[20].
> Antes de entrar en detalles vamos a esbozarla idea principal de la prueba, aunque naturalmente sin pretensiones de exactitud. Las fórmulas de un sistema formal (aquí nos limitamos al sistema *PM*), externamente consideradas, son

[19] En la nota 4 al pie, Gödel puntualiza, claro, que se refiere a que existen sentencias indecidibles utilizando **LA** y el modelo estándar $\mathcal{N}$.

[20] No es necesario que los axiomas sean independientes, pero, obviamente, no tiene sentido añadir axiomas que nos lleven a enunciados falsos. Además, con esto quitamos la posibilidad tentadora de añadir como axioma la sentencia indecidible (no demostrable) pero verdadera.

secuencias finitas de signos primitivos (variables, constantes lógicas y paréntesis o signos de puntuación) y se puede precisar fácilmente qué filas de signos primitivos son fórmulas y cuáles no. Análogamente, desde un punto formal las deducciones no son sino secuencias finitas de fórmulas (con ciertas propiedades explicitables). Para las considereaciones metamatemáticas resulta indiferente qué objetos usemos como signos primitivos. Usemos números naturales como tales signos. Consiguientemente, una fórmula será una secuencia finita de números naturales y una deducción será una secuencia finita de secuencias finitas de números naturales[21]. Los conceptos (o enunciados) metamatemáticos se convierten así en conceptos (respectivamente, enunciados) sobre números naturales o sucesiones de números naturales y por tanto pueden ser (al menos en parte) expresados con los símbolos del sistema *PM*. En particular se puede mostrar que los conceptos «fórmula», «deducción» y «fórmula deducible»[22] son definibles en el interior del sistema *PM*. Por ejemplo, se puede ofrecer una fórmula $\varphi(v)$ de *PM* con una variable v (del tipo de una secuencia de números), tal que, interpretada, dando a los signos de *PM* su significado intuitivo[23], dice: v es una fórmula deducible. Ahora construimos una sentencia indecidible del sistema *PM*, es decir, una sentencia α, tal que ni α ni $\neg\alpha$ es deducible, del siguiente modo:
Llamamos *signo de clase* a una fórmula de *PM* con exactamente una variable libre del tipo de los número naturales (clase de clases). Supongamos que los signos de clase están ordenados de alguna manera en una sucesión, designemos su miembro n-ésimo mediante $R(n)$ y observemos que el concepto «signo de clase» al igual que la relación ordenante R pueden ser definidos en el sistema *PM*. Sea α un signo de clase cualquiera; mediante $[\alpha; n]$ designamos *la fórmula que resulta de sustituir la variable libre por el signo que denota el número natural* n en el signo de clase α[24]. También la relación ternaria $x = [y; z]$ resulta ser definible en *PM*. Ahora definimos una clase K de números naturales del siguiente modo:

$$n \in K \longleftrightarrow \neg Bew[R(n); n], \tag{2.4.1}$$

donde $Bew\ x$ significa: x es una fórmula deducible. Puesto que todos los conceptos que aparecen en el definiens son definibles en *PM*[25], también lo es el

[21] A este proceso se lo conoce como «gödelización», que explicaremos más adelante. De esta manera se asignan biunívocamente números naturales (su número de Gödel g) a los signos primitivos, a las fórmulas y a las pruebas.

[22] Que son conceptos metamatemáticos.

[23] Se refiere al dado por el modelo estándar de la artimética, o sea, dando el significado a los signos en $\mathcal{N}$.

[24] Es decir, si el signo de clase α es $\varphi(v)$, con v una variable libre, entonces se trata de sustituir el signo asociado al número natural n (o sea, el signo cuyo número de Gödel es n) en dicha variable libre.

[25] Efectivamente ya ha comentado que los conceptos de «fórmula deducible» y la relación ternaria la relación ternaria $x = [y; z]$ que utiliza la «fórmula que resulta de sustituir la variable libre [de un signo de clase] por el signo que denota el número natural n» son definibles en *PM*. Posteriormente

concepto compuesto de ellos K, es decir, hay un signo de clase σ tal que la fórmula $[\sigma; n]$, interpretada de acuerdo con el significado intuitivo de sus signos, dice que el número nautral n pertenece a K. Puesto que σ es un signo de clase, es idéntico con cierto $R(q)$, es decir, ocurre que

$$\sigma = R(q), \tag{2.4.2}$$

para cierto número natural q. Ahora mostramos que la sentencia $[R(q); q]$[26] es indecidible en *PM*. Pues si suponemos que la sentencia $[R(q); q]$ fuera deducible, entonces sería verdadera; en ese caso, y de acuerdo con lo antes dicho, q pertenecería a K, es decir, por 2.4.1 valdría que $\neg Bew[R(q); q]$[27], en contradicción con el supuesto. Si, por el contrario, la negación de $[R(q); q]$ fuera deducible, entonces ocurriría que $q \notin K$, es decir, valdría que $Bew[R(q); q]$. Así, tanto $[R(q); q]$ como su negación serían deducibles, lo que de nuevo es imposible[28].
La analogía de esta argumentación con la antinomia de Richard salta a la vista; también está íntimamente relacionada con la paradoja del «mentiroso», pues la sentencia indecidible $[R(q); q]$ dice que q pertenece a K, es decir, según 2.4.1, que $[R(q); q]$ no es deducible. Así pues, tenemos ante nosotros una sentencia que afirma su propia indeducibilidad. Evidentemente el método de prueba que acabamos de exponer es aplicable a cualquier sistema formal que, en primer lugar, interpretado naturalmente, disponga de medios de expresión suficientes para definir los conceptos que aparecen en la argumentación anterior (especialmente el concepto de «fórmula deducible») y en el cual, en segundo lugar, cada fórmula deducible sea verdadera en la interpretación natural[29]. Uno de los propósitos del desarrollo exacto de la prueba indicada, que a continuación ofreceremos, consiste en sustituir la segunda de las citadas condiciones por otra puramente formal y mucho más débil.
De la observación de que $[R(q); q]$ dice de sí misma que no es deducible se sigue inmediatamente que $[R(q); q]$ es verdadera, pues $[R(q); q]$ no es deducible (ya que no es decidible). La sentencia indecidible *en el sistema PM* ha sido, pues, finalmente decidida mediante consideraciones metamatemáticas[30]. El análisis preciso de esta extraña situación conduce a resultados sorprendentes respecto a las pruebas de consistencia de sistemas formales, resultados que serán tratados más detenidamente en la sección 4 (teorema XI). ([18] Gödel 1931, pp. 53-57)

lo demostrará rigorosamente.

[26]Esta es la fórmula que resulta de sustituir la variable libre del signo de clase σ por el signo que denota el número natural q, y que, interpretada de acuerdo con el significado intuitivo de sus signos, dice que el número natural q pertenece a K.

[27]Es decir, $[R(q); q]$ no es una fórmula deducible.

[28]Suponiendo que el sistema es consistente.

[29]Cualquier teoría formal que contenga la aritmética de Peano de primer orden.

[30]Hemos hecho una reducción al absurdo en el «metalenguaje».

2.4.2. Gödelización

En la segunda parte del artículo se precisa el sistema formal con el que se trabaja, que es una unión entre el de *Principia Mathematica* y los axiomas de Peano: básicamente **P**. Además, se fija la interpretación de las variables, algo, como hemos dicho anteriormente, fundamental. En concreto con la teoría simple de tipos: habrá variables de tipo 1 («numéricas»), que se refieren a los números naturales, de tipo 2 («sentenciales»), las que se refieren a clases de números naturales, de tipo 3 («predicativas») para clases de clases, y así sucesivamente. De esta manera tenemos siempre una interpretación para toda fórmula de **P**, que es una cierta relación de números naturales, que será verdadera o falsa. Además, no es necesario tener variables para relaciones binarias o n-arias, pues se pueden definir las relaciones como clases de pares ordenados y los pares ordenados como clases de clases.

A continuación Gödel establece el procedimiento preciso para la identificación biunívoca entre los signos primitivos de **P** y los números naturales, y las propiedades y relaciones de aquellos en propiedades y relaciones de estos. La manera de hacer esto es a través de la numeración de Gödel: una proposición metamatemática acerca de ciertas expresiones es reflejada en el cálculo al utlizar sus números de Gödel (los asignamos a través de la aplicación g), convirtiéndose en una proposición aritmética. De esta forma los conceptos metamatemáticos quedan codificados numéricamente, estableciendo un puente entre el lenguaje formal y el lenguaje natural a través de la aritmética[31]. En otras palabras, la representación numérica de las entidades sintácticas determina una serie de relaciones y funciones numéricas que se corresponde exactamente con las relaciones y funciones metamatemáticas. Por ejemplo, a la propiedad metamatemática de ser un axioma le corresponde la propiedad numérica de ser el número de Gödel de un axioma. O, para representar la expresión metamatemática de ser inferible de dos fórmulas por MP (o sea, de $\alpha \to \beta$ y α inferimos β), tendremos la relación numérica de un número n con otros dos, m y p, cuando n es el número de Gödel de una fórmula inferible por MP de otras dos que tienen número de Gödel m y p[32]. Posteriormente analizaremos en detalle algunas propiedades metamatemáticas que Gödel utiliza. Ahora bien, sabemos que no todo lo que tiene sentido referido a $\mathbb{N}$ se puede expresar en **LA**, pues la capacidad expresiva de **LA** no alcanza a todos los subconjuntos, relaciones y funciones de $\mathbb{N}$[33]. En la sección siguiente explicaremos qué relaciones podemos expresar en **LA** (Gödel se encarga

[31] En particular, la expresión metalingüística, perteneciente al lenguaje natural, que será de importancia crucial es: «Esta sentencia es un teorema».

[32] Un ejemplo de función es la «negación». Es la función numérica que asigna a cada número de Gödel de una hilera de signos el número de Gödel de su negación.

[33] Una breve explicación es que **LA** es numerable, luego la cantidad de fórmulas formales será una unión numerable de conjuntos finitos o numerables, luego también será numerable. Pero, por el teorema de Cantor, es imposible dar nombre (entiéndase, definir intensional o extensionalmente) a todos y cada uno de los subconjuntos de $\mathbb{N}$. De esta forma, aún cuando cada fórmula $\varphi(x)$ de **LA** diese lugar a un conjunto metalingüístico de números naturales, éstos no serían todos los posibles.

de demostrar que puede representar todas las que necesita para la demostración). Pero tratemos por ahora la gödelización:

Primero asignamos unívocamente número naturales a los signos primitivos, incluidos los signos de puntuación, de **P** de la siguiente manera:

$$\text{g}(()=3, \quad \text{g}())=5, \quad \text{g}(\neg)=7, \quad \text{g}(\rightarrow)=9, \quad \text{g}(\exists)=11,$$
$$\text{g}(\mathbf{0})=13, \quad \text{g}(\mathbf{S})=15, \quad \text{g}(+)=17, \quad \text{g}(\cdot)=19, \quad \text{g}(=)=21.$$

A las variables de tipo n les asignamos los números de la forma ρ^n, donde ρ es un número primo mayor que 21. Es decir, a las variables de tipo 1 les asignamos número primos mayores que 21; a las variables de tipo 2 les asignamos como número de Gödel el cuadrado de un número primo mayor que 21; a las variables de tipo 3 el número de Gödel igual al cubo de un número primo mayor que 21, etc.

Para asignar un número Gödel a una fórmula, que es una hilera de signos, tomaremos todos los números Gödel correspondientes a cada uno de sus signos en orden, y los pondremos, uno a uno, como potencia de los primero números primos en orden de magnitud. Veamos un ejemplo:

Consideremos la fórmula del sistema $(\exists x)(x = \mathbf{S}(y))$, donde x e y son variables de tipo 1. Significa que existe un número natural tal que es el sucesor inmediato de otro número natural, que es lo mismo que decir que todo número tiene un sucesor inmediato. Tomando los números de Gödel de cada símbolo en orden, tenemos 3, 11, 23, 5, 3, 23, 21, 15, 3, 29, 5, 5. Por tanto, según la regla establecida anteriormente, el número Gödel de la fórmula $(\exists x)(x = \mathbf{S}(y))$ es

$$m = 2^3 \cdot 3^{11} \cdot 5^{23} \cdot 7^5 \cdot 11^3 \cdot 13^{23} \cdot 17^{21} \cdot 19^{15} \cdot 23^3 \cdot 29^{29} \cdot 31^5 \cdot 37^5. \tag{2.4.3}$$

Así tenemos asignado biunívocamente un número natural a cada signo primitivo y a cada secuencia finita de signos primitivos. Observamos que son números extremadamente grandes, pero lo importante es que existen. Además, por el teorema fundamental de la aritmética, podemos factorizar cualquier número y ver si proviene de alguna expresión. Es muy imporante escribir las fórmulas correctamente, ateniéndonos a los símbolos definidos y no a sus abreviaturas. Por ejemplo, si quisiéramos calcular el número de Gödel de la sentencia (falsa) $\mathbf{0} \neq \mathbf{0}$, tenemos que escribirla de la forma $(\neg(\mathbf{0} = \mathbf{0}))$[34]. Su número de Gödel es $2^3 \cdot 3^7 \cdot 5^3 \cdot 7^{13} \cdot 11^{21} \cdot 13^{13} \cdot 17^5 \cdot 19^5$.

También deseamos tener un sólo número de Gödel para una sucesión finita de fórmulas (que puede ser una demostración). En este caso, multiplicamos los primeros

[34] Y, si queremos evitar los paréntesis, de la forma $\neg = \mathbf{00}$, aunque esto no es habitual, y ni si quiera el propio Gödel lo hace.

número primos elevados, respectivamente, a la potencial igual al número Gödel de cada fórmula que compone la sucesión. Notemos que los número de Gödel de las fórmulas tienen siempre exponentes impares (luego los números son pares) y, por contra, los números Gödel de las sucesiones de fórmulas tienen siempre exponentes pares.

2.4.3. Teoría de la Recursión

La aritmetización de Gödel nos lleva directamente a tratar con el siguiente aspecto clave de la demostración: la recursividad. Antes decíamos que la capacidad expresiva de **LA** no alcanza a todos los subconjuntos, relaciones y funciones de $\mathbb{N}$, pero que Gödel demuestra que puede representar todas las relaciones que necesita para la demostración. El concepto de *representabilidad* es, pues, fundamental. Tenemos dos dificultades a estudiar:

1. Dar una definición precisa de *representabilidad* de una relación de objetos de $\mathbb{N}$ en **LA**.

2. Establecer un criterio para dicha representabilidad.

Para el estudio de estas cuestiones Gödel se vale de la teoría de la recursividad[35]. El teorema fundamental que da respuesta a la segunda pregunta es el siguiente

Teorema 2.4.1.

- Las funciones primitivas recursivas, a través de su grafo, son relaciones representables.

- Una relación R es representable si su aplicación característica 1_R es primitiva recursiva.

Es a través de la recursividad, en particular a través de las funciones primitivas recursivas[36], que Gödel transporta el problema del lenguaje formal (una relación) al lenguaje de la aritmética (una relación aritmética).

Definición 2.4.2 (Funciones aritméticas elementales)**.**

- Función cero: $z(n) = 0$, $\forall n \in \mathbb{N}$.

- Función siguiente: $s(n) = n + 1$, $\forall n \in \mathbb{N}$.

[35] Para un tratamiento perfectamente formal del tema véase ([12] Fernández Margarit 2012, cap. IV).

[36] Vamos a definir también las funciones recursivas, pero en la demostración de sus teoremas Gödel precisa únicamente de las funciones y relaciones primitivas recursivas.

- Funciones proyecciones k-arias: $\pi_i^k(n_1, ..., n_i, ..., n_k) = n_i$, para toda k-upla $(n_1, ..., n_k)$, con $1 \leq i \geq k$.

Vamos a definir las funciones primitivas recursivas por un proceso generador, a partir de las funciones aritméticas elementales y tres reglas de construcción. Estas tres reglas u operaciones son la composición, la recursión y la minimalización.

Definición 2.4.3 (Composición de aplicaciones)**.** Sea $g : \mathbb{N}^l \to \mathbb{N}$ una función l-aria y $f_1 : \mathbb{N}^k \to \mathbb{N}, ..., f_l : \mathbb{N}^k \to \mathbb{N}$, l funciones k-arias. La función k-aria $h : \mathbb{N}^k \to \mathbb{N}$, definida por

$$h(n_1, ..., n_k) = (g \circ (f_1, ..., f_l))(n_1, ..., n_k) = g(f_1(n_1, ..., n_k), ..., f_l(n_1, ..., n_k)),$$

es la función compuesta por $g, f_1, ..., f_l$.

La definición recursiva es habitual en matemáticas. Nosotros mismos la hemos utilizado en secciones anteriores. La idea de una función aritmética definida recursivamente es que el valor $n+1$ se pueda conocer siempre que se conozca el valor en n. De esta manera basta fijar el valor en 0. Por ejemplo, si queremos definir la función suma $+_m : \mathbb{N} \to \mathbb{N}$, $n \mapsto +_m(n) = m + n$, podemos dar el valor en cero $+_m(0) = m$, y el valor $(n+1)$-ésimo en función del n-ésimo haciendo uso de la función siguiente: $+_m(n+1) = s(+_m(n)) = (m+n)+1$. La definición que damos es sin embargo menos intuitiva.

Definición 2.4.4 (Recursión de funciones aritméticas)**.** Sea $g : \mathbb{N}^k \to \mathbb{N}$ una función k-aria y $f : \mathbb{N}^{k+2} \to \mathbb{N}$ una función $k+2$-aria. La función $(k+1)$-aria $h : \mathbb{N}^{k+1} \to \mathbb{N}$, definida por

$$h(n_1, ..., n_k, n) = \begin{cases} g(n_1, ..., n_k) & \text{, si } n = 0, \\ f(n_1, ..., n_k, h(n_1, ..., n_k, m), m) & \text{, si } n = s(m). \end{cases}$$

, es la función obtenida por recursión de f y g[37].

Definición 2.4.5 (Función especial)**.** Diremos que una función $f : \mathbb{N}^{k+1} \to \mathbb{N}$ es una función especial si para cada k-upla $n_1, ..., n_k$, existe un $m \in \mathbb{N}$ tal que $f(n_1, ..., n_k, m) = 0$.

Definición 2.4.6 (Minimalización de funciones aritméticas)**.** Sea $f : \mathbb{N}^{k+1} \to \mathbb{N}$ una función $(k+1)$-aria especial. La función k-aria $h : \mathbb{N}^k \to \mathbb{N}$ definida por

$$h(n_1, ..., n_k) = \text{mín}\{m : f(n_1, ..., n_k, m) = 0\}$$

es la función obtenida por minimalización de f.

[37] A g se le denomina función base, y a f función de iteración.

Definición 2.4.7 (Función primitiva recursiva)**.** Las funciones primitivas recursivas son aquellas funciones aritméticas obtenidas a partir de las elementales y por aplicación finita de las operaciones de composición y recursión.

Definición 2.4.8 (Función recursiva)**.** Las funciones recursivas son aquellas funciones aritméticas obtenidas a partir de las elementales y por aplicación finita de las operaciones de composición, recursión y minimalización.

Las funciones más usuales son funciones primitivas recursivas, por ejemplo: identidad, suma, producto, exponenciación, factorial, mínimo, máximo, cociente, sucesor... Puede encontrarse una demostración en ([12] Fernández Margarit 2012, pp. 88-92). Por otra parte, es inmediato ver que las reglas de derivación, que son funciones que aplican expresiones en expresiones, son recursivas. Por ejemplo, podemos expresar el MP de la siguiente forma

$$f(\alpha, y) = \begin{cases} \beta & \text{, si } y \text{ tiene la forma de } \alpha \rightarrow \beta, \\ \alpha & \text{, en otro caso.} \end{cases}$$

Si suprimimos la restricción de que la operación de minimalización únicamente se aplica a funciones especiales, además de que la función h definida en 2.4.6 no será una aplicación ($\mathrm{dom}(h) \neq \mathbb{N}^k$), puede ocurrir que dicha función h esté definida solamente en una parte de $\mathbb{N}^k$. Definimos así, eliminando tal restricción, las funciones parciales recursivas.

Definición 2.4.9 (Relación primitiva recursiva)**.** Una relación aritmética R es primitiva recursiva si, y sólo si, su función característica 1_R lo es.

Definición 2.4.10 (Relación recursiva)**.** Una relación aritmética R es recursiva si, y sólo si, su función característica 1_R lo es.

Con estas últimas definiciones, y recordando el teorema 2.4.1, ya podemos deducir que una relación R es representable si es primitiva recursiva. Ahora que ya tenemos una idea de la teoría de recursividad, podemos analizar en detalle todo lo relacionado con la representabilidad en la siguiente sección.

2.4.4. Representabilidad sintáctica

Hagamos un repaso rápido de lo que tenemos hasta ahora. Gödel trabaja en el lenguaje de primer orden **LA** con una teoría de tipos para la interpretación de las variables, con la interpretación estándar $\mathcal{N}$, adoptando los axiomas de Peano; lo que hemos llamado la teoría o sistema **P**. Junto a esta teoría tenemos su metateoría, donde se incluyen las expresiones «φ es una fórmula de **P**», «φ es un axioma» o «φ es demostrable», expresiones que de ninguna manera pertenecen a **P**. Ahora bien, y esta es posiblemente la genialidad de Gödel, estas expresiones pueden formularse como funciones recursivas primitivas. Al ser fórmulas de la metamatemática, por la aritmetización, son expresiones acerca de conjuntos aritméticos. Por ser de este tipo especial son formalizables, o sea, representables en **P**. Como consecuencia tenemos que la metamatemática de **P** se inserta en el mismo sistema **P**, que adquiere capacidad para «hablar de sí mismo». En resumen, los dos problemas que resuelve Gödel son

1. Demostrar que los predicados metamatemáticos pueden expresarse en términos de funciones y relaciones primitivamente recursivas.

2. Demostrar que las funciones y relaciones primitivamente recursivas son representables, es decir, son expresables sintácticamente en **P**.

Para el primer punto Gödel demuestra[38], en 46 lemas, que nociones como «x es divisible por y» o «x se deriva de la fórmula y» se pueden definir en términos de funciones y relaciones primitivamente recursivas. Sin embargo, la número 46, «x es una fórmula demostrable» $(Bew\ \ x)$[39], es la única de la que no se puede afirmar que sea primitiva recursiva. Algunos conjuntos, relaciones y funciones que es de utilidad representar en $\mathbb{N}$ son:

- Las fórmulas, con el conjunto $Form \subset \mathbb{N}$:

$$Form = \{n : n = g(\varphi),\ \text{con } \varphi \in \mathbf{Form}(\mathbf{LA})\}.$$

- Las fórmulas con una variable libre, con el conjunto $Form_1 \subset \mathbb{N}$:

$$Form_1 = \{n : n = g(\varphi(x)),\ \text{con } \varphi(x) \in \mathbf{Form}(\mathbf{LA}) \text{ y } \mathbf{Vl}(\varphi) = \{x\}\}.$$

- Los axiomas, con el conjunto $Ax \subset \mathbb{N}$:

$$Ax = \{n : n = g(\psi),\ \text{con } \psi \text{ un axioma de } \mathbf{P}\}.$$

- Los teoremas, con el conjunto $Teor \subset \mathbb{N}$:

$$Teor = \{n : n = g(\varphi),\ \text{con } \varphi \in \mathbf{Form}(\mathbf{LA}) \text{ y } \mathbf{P} \vdash \varphi\}.$$

- La función Num de $\mathbb{N}$ en $\mathbb{N}$ que gödeliza el término $\mathbf{n}$:

$$Num(n) = g(\mathbf{n}).$$

[38]Se basa en la aritmetización de la sintaxis con los números de Gödel y en consideraciones sobre las funciones primitivas recursivas, todo lo cual va más allá del alcance de este texto.

[39]Es la propiedad que tiene un número natural si, y sólo si, es el número de Gödel de una fórmula deducible.

- La función *Subst* de $\mathbb{N} \times \mathbb{N}$ en $\mathbb{N}$ que a la fórmula $\varphi(x)$, con $x \in \mathbf{Vl}(\varphi)$ de número de Gödel m, $m = g(\varphi(x))$, le asocia el número de Gödel de la sentencia $\varphi(\mathbf{n})$:

$$Subst(m, n) = g(\mathbf{n}).$$

- La relación binaria $W_1 \subset \mathbb{N}\times\mathbb{N}$ tal que $<m, n>\in W_1$ si, y sólo si, $m = g(\varphi(x))$, con $\varphi(x) \in \mathbf{Form}(\mathbf{LA})$ y $x \in \mathbf{Vl}(\varphi)$; y donde $n = g(\mathfrak{s})$, con $\mathfrak{s}$ una sucesión de símbolos que es una demostración de $\varphi(\mathbf{m})$.

Pasemos ahora al segundo punto, es decir, la cuestión de que las funciones y relaciones primitivamente recursivas son representables en $\mathbf{P}$.

Dentro del sistema formal de $\mathbf{P}$ disponemos de los símbolos formales $\mathbf{0}, \mathbf{1}, \mathbf{2}, ...$(estas son abreviaciones) para los número naturales $0, 1, 2, ...$ Y es claro que, interpretando la letra funcional 1-aria $\mathbf{S}$ como la aplicación siguiente s, la interpretación de $\mathbf{n}$ en la $\mathbf{LA}$-estructura estándar $\mathcal{N}$ es n. Pues bien, una relación k-aria $R \subset \mathbb{N}^k$ de números naturales significa que tenemos ciertas k-uplas tales que $<n_1, ..., n_k>\in R$, y otras tales que $<m_1, ..., m_k>\notin R$. La idea de la representabilidad en $\mathbf{P}$ se basa en la posibilidad de transcribir este hecho a través del lenguaje formal.

Definición 2.4.11 (Relación representable débilmente en $\mathbf{P}$)**.** Una relación k-aria $R \subset \mathbb{N}^k$ de números naturales es representable débilmente en $\mathbf{P}$ por la fórmula φ si, y sólo si, φ tiene k variables libres y para toda k-upla de números naturales verifica

$$< n_1, n_2, ...n_k >\in R \iff \mathbf{T} \vdash \varphi(\mathbf{n}_1, \mathbf{n}_2, ...\mathbf{n}_k)$$

Definición 2.4.12 (Relación representable en $\mathbf{P}$)**.** Una relación k-aria $R \subset \mathbb{N}^k$ de números naturales es representable en $\mathbf{P}$ por la fórmula φ si, y sólo si, φ tiene k variables libres y para toda k-upla de números naturales verifica

$$\begin{cases} \mathbf{1}. \quad < n_1, n_2, ...n_k >\in R \Longrightarrow \mathbf{T} \vdash \varphi(\mathbf{n}_1, \mathbf{n}_2, ...\mathbf{n}_k), \\ \mathbf{2}. \quad < n_1, n_2, ...n_k >\notin R \Longrightarrow \mathbf{T} \vdash \neg\varphi(\mathbf{n}_1, \mathbf{n}_2, ...\mathbf{n}_k). \end{cases}$$

Definición 2.4.13 (Función representable en $\mathbf{P}$)**.** Una función k-aria $f : \mathbb{N}^k \to \mathbb{N}$ es representable en $\mathbf{P}$ si, y sólo si, su grafo $G(f)$ es representable en $\mathbf{P}$.

Recordemos que el grafo de una función k-aria $f : \mathbb{N}^k \to \mathbb{N}$ es

$$G(f) = \{(n_1, n_2, ..., n_k, m) : m = f(n_1, n_2, ..., n_k)\},$$

que se puede ver como una relación $(k + 1)$-aria $G(f) \subset \mathbb{N}^{k+1}$ donde

$$< n_1, n_2, ..., n_k, m > \in G(f) \iff m = f(n_1, n_2, ..., n_k).$$

Podríamos decir entonces que una función k-aria $f : \mathbb{N}^k \to \mathbb{N}$ es representable en **P** por una fórmula φ con $n+1$ variable libres si, y sólo si,

1. $\mathbf{T} \vdash \varphi(\mathbf{n}_1, \mathbf{n}_2, ...\mathbf{n}_k, \mathbf{m})$,
2. $\mathbf{T} \vdash \forall \mathbf{v} \varphi(\mathbf{n}_1, \mathbf{n}_2, ...\mathbf{n}_k, \mathbf{v}) \to \mathbf{m} = \mathbf{v}$[40],

, donde $m = f(n_1, n_2, ..., n_k)$.

Es gracias a la fórmula $\varphi(x_1, x_2, ..., x_n)$ que podemos transferir una propiedad de la teoría de conjuntos (de $\mathbb{N}$) al sistema formal **P**. Gödel juega con los distintos lenguajes sin mezclarlos, traduciendo las expresiones del modelo $\mathcal{N}$ en expresiones formales internas a la teoría.

Con estas definiciones respondemos ya a la primera cuestión que planteábamos al inicio de la sección 2.4.3. En cuanto al segundo punto allí expuesto, a saber, establecer un criterio para la representabilidad en **P** de las relaciones en $\mathbb{N}$, ya anticipamos el resultado en el teorema 2.4.1, que ahora podemos enunciar en toda su generalidad y esbozar una demostración intuitiva. Corresponde con el teorema V en el texto de Gödel.

Teorema 2.4.14. Una función[41] es primitiva recursiva[42] si, y sólo si, es representable en **P**.

Demostración. Para la implicación de derecha a izquierda, supongamos que la fórmula φ representa a la función f en **P**. Veamos que f es primitiva recursiva. Si A es el conjunto de axiomas de **P**, podemos computar f de la siguiente manera: Se generan los A-teoremas; por ser f representable, por definición existen fórmulas de la forma $\varphi(\mathbf{n}, \mathbf{m})$ en **P**; en tal caso, cada vez que se obtenga dicha fórmula, tendremos $f(n) = m$. De esta manera se prueba la recursividad de f.

En la implicación de izquierda a derecha se trata de probar que toda función primitiva recursiva es representable. La idea es formar el conjunto X de todas las funciones representables en **P** y comprobar que las fnciones aritméticas, suma, producto, sucesor, etc., pertenecen a X[43]. □

[40] Esta condición expresa la unicidad de la fórmula φ.

[41] Es equivalente escribir «relación» en lugar de «función», pues sabemos que una relación es primitiva recursiva si, y sólo si, su función característica lo es.

[42] Es válido también para funciones recursivas.

[43] Puede encontrarse una demostración en ([12] Fernández Margarit 2012, pp. 88-92). Esta no es estrictamente la manera de proceder de Gödel, sino que la introdujo J. Robinson.

Este teorema (y este esbozo de demostración) no es válido únicamente para la teoría **P**, sino para cualquiera que satisfaga las condiciones de las definiciones 2.4.11, 2.4.12 y 2.4.13 y que sea recursivamente numerable[44] (axiomatizable).

Ya hemos dicho que Gödel se ocupar de demostrar que todas las relaciones de las que precisa en su demostración son representables en **P**. Y, en la página 75, dábamos algunos ejemplos. Para construir la fórmula indecidible vamos a utilizar la relación binaria W_1. Recordemos que $<m,n>\in W_1$ si, y sólo si, $m = g(\varphi(x))$, con $\varphi(x) \in \mathbf{Form}(\mathbf{LA})$ y $x \in \mathbf{Vl}(\varphi)$; y donde $n = g(\mathfrak{s})$, con $\mathfrak{s}$ una sucesión de símbolos que es una demostración de $\varphi(\mathbf{m})$.

2.4.5. Demostraciones

Hemos puesto de manifiesto las tres herramientas esenciales de la demostración de Gödel: la gödelización, la representabilidad y la recursividad primitiva. La demstración es de tipo constructivo; es decir, se construye efectivamente una sentencia indecidible, indemostrable en el sistema **P**. Como vimos en el texto introductorio de Gödel, esta sentencia es la que afirma de sí misma que no es demostrable. La idea es la siguiente:

1. Construcción de una fórmula $\Phi \in \mathbf{Form}(\mathbf{LA})$ que represente la declaración metamatemática «la fórmula Φ no es deducible».

$\downarrow$

2. Prueba de que Φ no es formalmente demostrable.

$\downarrow$

3. En consecuencia, Φ es verdadera.

$\downarrow$

4. Si la teoría es consistente, entonces es incompleta.

Nuestra demostración (vamos a plasmar la presentada en ([30] Pla i Carrera, 2012)) no es, como ya se dijo, plenamente rigorosa, ni mucho menos. Sin embargo, con ella y con los apartados anteriores, creemos haber dado una imagen general bastante completa de los teoremas de incompletitud.

Construyamos la fórmula Φ que diga de si misma que no es deducible. Vamos a usar la relación binaria $W_1 \subset \mathbb{N} \times \mathbb{N}$ anteriormente definida:

$$W_1 = \{< m, n >: m = g(\varphi(x)), \varphi(x) \in \mathbf{Form}(\mathbf{LA}) \text{ y } n = g(\mathfrak{s})\},$$

con $\mathfrak{s}$ una sucesión de símbolos que es una demostración de $\varphi(\mathbf{m})$. Con esto queda claro que el significado de $<m,n>\in W_1$ ($W_1(m,n)$) es: n es el número de Gödel de una prueba de la fórmula $\varphi(\mathbf{m})$, donde hemos sustituido el símbolo $\mathbf{m}$ en la variable libre de la fórmula $\varphi(x)$, la cual tiene número de Gödel m. En otras palabras

[44]Trataremos esta cuestión en la siguiente sección sobre Computabilidad.

$$W_1(m,n) \iff \text{n es el número de Gödel de una prueba de } \varphi(\mathbf{m}) \;/\; m = g(\varphi(x)).$$

Nos preguntamos ahora si es posible representar la relación W_1 en el lenguaje formal, que sabemos que ocurre si su función característica 1_{W_1} es primitiva recursiva. En este caso, sí lo es (no hacemos la demostración). Por ser la relación representable en **P**, como hemos visto en 2.4.12, tendremos una fórmula $\omega_1(x,y) \in \mathbf{Form}(\mathbf{LA})$, con $\mathbf{Vl}(\omega_1) = \{x,y\}$, que cumpla lo siguiente

$$\begin{cases} \mathbf{1.} \quad < m,n > \in W_1 \Longrightarrow \mathbf{T} \vdash \omega_1(\mathbf{m},\mathbf{n}), \\ \mathbf{2.} \quad < m,n > \notin W_1 \Longrightarrow \mathbf{T} \vdash \neg\omega_1(\mathbf{m},\mathbf{n}). \end{cases}$$

Hemos conseguido trasladar $W_1 \subset \mathbb{N} \times \mathbb{N}$, una relación en $\mathbb{N}$, a una fórmula ω_1 en **LA**: si n es el número de Gödel de una prueba $\mathfrak{s}$ (n=g($\mathfrak{s}$)) de la fórmula $\varphi(\mathbf{m})$, donde hemos sustituido el símbolo **m** en la variable libre de la fórmula $\varphi(x)$, la cual tiene número de Gödel m, entonces $\omega_1(\mathbf{m},\mathbf{n})$ es un teorema de **P**:

$$\mathbf{T} \vdash \omega_1(\mathbf{m},\mathbf{n}).$$

Pero la relación aritmética $W_1(m,n)$ entre los números de Gödel expresa, a través de su formalización en la fórmula ω_1, una relación metamatemática:

$$W_1(m,n),$$
$$\downarrow$$
$$\text{n es el número de Gödel de una prueba de la fórmula } \varphi(\mathbf{m}), \text{ con } m = g(\varphi(x)),$$
$$\downarrow$$
$$\mathbf{T} \vdash \omega_1(\mathbf{m},\mathbf{n}),$$
$$\downarrow$$
«La fórmula $\varphi(\mathbf{m})$, con $m = g(\varphi(x))$, es deducible».

Tratamos entonces de encontrar una fórmula que afirme su propia indecibilidad. Esto es, una fórmula $\gamma(x)$ con número de Gödel q que diga:

«La fórmula $\gamma(\mathbf{q})$, con $q = g(\gamma(x))$, no es deducible».

Para ello, usamos el la fórmula $\omega_1(x,y)$ de esta manera:

$$\gamma(x) := \forall y \neg\omega_1(x,y). \tag{2.4.4}$$

Tendrá un número de Gödel asociado, pongamos que es q: $q = g(\gamma(x))$. Y ahora sustituimos en la variable libre x el símbolo **q** (cuya interpretación es q), obteniendo la *sentencia de Gödel*:

$$\Phi := \gamma(\mathbf{q}) = \forall y \neg\omega_1(\mathbf{q},y). \tag{2.4.5}$$

Esta sentencia está expresando que no existe ningún número de Gödel de una prueba de la fórmula $\gamma(\mathbf{q})$, con q el número de Gödel de $\gamma(x)$. O sea, que «la fórmula $\gamma(\mathbf{q})$, con $q = g(\gamma(x))$, no es deducible». Pero esa fórmula es ella misma, luego está expresando su propia indecibilidad. Y esta es la sentencia indecidible Φ que vamos a usar en la demostración de los teoremas.

De esta forma, $W_1(q, n)$ es verdadero, o lo que es lo mismo, $\mathcal{N} \vDash \omega_1(\mathbf{q}, \mathbf{n})$ si, y sólo si, n es el número de Gödel de una demostración de $\gamma(\mathbf{q})$.

Una vez construida la que va a ser nuestra sentencia indecidible Φ, podemos pasar a la demostración de los teoremas. Antes, sin embargo, para poder entenderlos en toda su generalidad, tenemos que explicar los conceptos de clase recursiva primitiva y ω-consistencia.

Gödel prueba que si $\mathbf{P}$ es consistente, entonces la sentencia Φ no es un teorema, pero, por contra, para probar que $\neg\Phi$ tampoco es un teorema, toma como hipótesis que $\mathbf{P}$ es ω-consistente. Esta es una exigencia más fuerte que la consistencia, es decir, ω-consistencia $\Rightarrow$ consistencia, pero consistencia $\nRightarrow$ ω-consistencia. Fue el lógico estadounidense John Barkley Rosser quien, en 1936, suprimió este tecnicismo de la demostración gödeliana, reduciendo la exigencia de la ω-consistencia a la mera consistencia (eso sí, construyendo una sentencia indecidible más compleja, aunque siguiendo el hilo de Gödel). En esencia nos dice que no es posible probar que cada número natural, por separado, tiene una cierta propiedad, y a la vez probar que no la tienen todos los números naturales.

Definición 2.4.15 (ω-consistencia)**.** La teoría $\mathbf{P}$ es ω-consistente si, y sólo si, no existe $\varphi \in \mathbf{Form}(\mathbf{LA})$ tal que $\mathbf{P} \vdash \varphi(\mathbf{n})$, $\forall n \in \mathbb{N}$, y a la vez $\mathbf{P} \vdash \exists x \neg\varphi(x)$.

Proposición 2.4.16. Si $\mathbf{P}$ es ω-consistente, entonces es consistente.

Demostración. Por el teorema 2.2.17, con probar que $\exists\varphi \in \mathbf{Form}(\mathbf{LA})$ tal que $\mathbf{P} \nvdash \varphi$, ya tendremos que $\mathbf{P}$ es consistente. Tenemos que, $\forall n \in \mathbb{N}$, $\mathbf{P} \vdash (((\mathbf{n} = \mathbf{n})) \to (\mathbf{n} = \mathbf{n}))$. Esto implica, por ser $\mathbf{P}$ ω-consistente, que $\mathbf{P} \nvdash \exists x \neg(((x = x)) \to (x = x))$. □

Por otra parte, Gödel trabaja con la teoría $\mathbf{P}$. Sin embargo justifica que sus teoremas son válidos para toda teoría $\mathbf{T}$ más potente que $\mathbf{P}$. Con esto nos referimos a que, siempre que añadamos a $\mathbf{P}$ un conjunto de nuevos axiomas que sea primitivo recursivo, valdrán también los teoremas para esta nueva teoría. La exigencia de que los axiomas de la teoría deben ser siempre primitivos recursivos se debe a la necesidad de que sean representables en $\mathbf{P}$.

Definición 2.4.17 (Conjunto primitivo recursivo)**.** Un conjunto A de número naturales es primitivo recursivo si A posee una función característica primitiva recursiva. Es decir, una función χ_A tal que

$$\chi_A(n) = \begin{cases} 1 & \text{, si } n \in A, \\ 0 & \text{, si } n \notin A. \end{cases}$$

El primer teorema de incompletitud nos dice que si la teoría **P**, o cualquier extensión suya obtenida añadiendo una clase **K** recursiva primitiva y ω-consistente de nuevos axiomas, es consistente, entonces siempre hay alguna sentencia tal que ni ella ni su negación son deducibles en el sistema. Esto es, es incompleta. Hemos probado que si una teoría es inconsistente, entonces todo es un teorema, y se puede probar una fórmula y su negación, lo cual convierte en absolutamente inútil todo el sistema. Es una propiedad más que deseable.

Teorema 2.4.18 (Primer teorema de incompletitud)**.** $\exists \Phi \in \mathbf{Sent}(\mathbf{LA})$ tal que

1. $\mathbf{P}$ consistente $\Longrightarrow \mathbf{P} \nvdash \Phi$.
2. $\mathbf{P}$ ω-consistente $\Longrightarrow \mathbf{P} \nvdash \neg\Phi$.

Demostración. Sea $\Phi := \forall y \neg \omega_1(\mathbf{q}, y)$.

1. Supongamos que $\mathbf{P}$ es consistente.

   Veamos el contrarrecíproco: supongamos que $\mathbf{P} \vdash \forall y \neg \omega_1(\mathbf{q}, y)$.
   $\exists$ una prueba $\mathfrak{s}$ de Φ.
   Pongamos que dicha prueba tiene número de Gödel n: $n = g(\mathfrak{s})$.
   En consecuencia, $W_1(q, n)$.
   Por la representabilidad de W_1, $\mathbf{P} \vdash \omega_1(\mathbf{q}, \mathbf{n})$.
   Por el axioma lógico de sustitución, y por hipótesis, $\mathbf{P} \vdash \neg\omega_1(\mathbf{q}, \mathbf{n})$.
   Resulta que $\mathbf{P} \vdash \omega_1(\mathbf{q}, \mathbf{n})$ y $\mathbf{P} \vdash \neg\omega_1(\mathbf{q}, \mathbf{n})$.
   $\mathbf{P}$ no es consistente.

2. Supongamos que $\mathbf{P}$ es ω-consistente.

   Por reducción al absurdo: supongamos que $\mathbf{P} \vdash \neg\forall y \neg \omega_1(\mathbf{q}, y)$.
   O lo que es lo mismo, $\mathbf{P} \vdash \exists y \omega_1(\mathbf{q}, y)$.
   $\mathbf{P}$ ω-consistente $\Longrightarrow$ $\mathbf{P}$ consistente.
   Luego $\mathbf{P} \nvdash \forall y \neg \omega_1(\mathbf{q}, y)$.
   Por tanto, $\nexists n \in \mathbb{N}$ que sea el número de Gödel de una prueba de Φ.
   En consecuencia, $<q, n> \notin W_1$, $\forall n \in \mathbb{N}$.
   Por la representabilidad de W_1, $\mathbf{P} \vdash \neg\omega_1(\mathbf{q}, \mathbf{n})$, $\forall n \in \mathbb{N}$.
   Aplicando la ω-consistencia, $\mathbf{P} \nvdash \exists y \neg\neg\omega_1(\mathbf{q}, y)$.
   O lo que es lo mismo, $\mathbf{P} \nvdash \exists y \omega_1(\mathbf{q}, y)$.
   Resulta que $\mathbf{P} \vdash \exists y \omega_1(\mathbf{q}, y)$ y $\mathbf{P} \nvdash \exists y \omega_1(\mathbf{q}, y)$.
   Contradicción, luego $\mathbf{P} \nvdash \neg\Phi$.

$\square$

Es interesante comentar algunas cuestiones sobre el teorema, para que se entienda lo más completamente posible:

- Por la naturaleza de las sentencias, que fijada una **L**-estructura, en este caso la **LA**-estructura $\mathcal{N}$, siempre son o verdaderas o falsas, resulta que o $\mathcal{N} \vDash \Phi$ o $\mathcal{N} \vDash \neg\Phi$. Es decir, alguna de las dos representa una proposición aritmética verdadera, y sin embargo ninguna es demostrable formalmente. En conclusión, la teoría formal de primer orden **P** no es capaz de abarcar todas las sentencias verdaderas en $\mathcal{N}$.

- Una sugerencia sería que, si **P** no es capaz de demostrar todas las sentencias que son verdaderas en $\mathcal{N}$, considerásemos la teoría $\mathbf{P}^* = \{\sigma \in \mathbf{Sent}(\mathbf{LA}) : \mathcal{N} \vDash \sigma\}$, la formada por todas las sentencias que son $\mathcal{N}$-verdaderas. Pero esta teoría es completa, como vimos en la sección 2.3.1. ¿Cómo es posible? Pudiera parecer que hemos resuelto el problema, pero la realidad es que esta teoría es inútil. Dado que no se ve afectada por el primer teorema de incompletitud, entonces necesariamente el conjunto de sus axiomas no es primitivo recursivo, y esto es catastrófico, pues significa que no es representable. Es una teoría «demasiado grande», inexpresable en **LA**.

- De manera similar, se puede demostrar que el conjunto de los teoremas de **P**, $\{\varphi \in \mathbf{Form}(\mathbf{LA}) : \mathbf{P} \vdash \varphi\}$ tampoco es representables, y por tanto no es primitivo recursivo. Se puede ver una demostración intuitiva en ([30] Pla i Carrera 2012, p. 217).

- Tal y como dijimos en la sección 2.3.2, el teorema de incompletitud no contradice el teorema de compleitud de Gödel (2.3.2), que recordemos que expresa que $\mathbf{T} \vdash \varphi \iff \mathfrak{U} \vDash \varphi, \forall \mathbf{L}(\mathbf{T})$-estructura $\mathfrak{U} \in \mathbf{Mod}(\mathbf{T})$. Esto es, que lo que es verdadero en todos los modelos de una teoría formal de primer orden es deducible, y viceversa. En el teorema de incompletitud, sin embargo, no nos estamos refiriendo a *todos* los modelos posibles de **P**, sino a uno específico, el modelo estándar de la aritmética $\mathcal{N} = \langle \mathbb{N}, 0, +, s, \cdot, < \rangle$. Y aquí existen, como hemos probado, sentencias verdaderas que no son deducibles.
  Además, dado que $\mathbf{P} \nvdash \Phi$ y $\mathbf{P} \nvdash \neg\Phi$, ni Φ ni $\neg\Phi$ son verdaderas en todos los modelos, luego debe haber al menos un modelo en el que Φ es falsa, y uno en el que $\neg\Phi$.

- Podemos ver ahora el significado de la sentencia de Gödel Φ con algo más de rigor. Tenemos que $\mathcal{N} \vDash \Phi \iff \forall n \in \mathbb{N}, \mathcal{N} \vDash \neg\omega_1(\mathbf{q}, \mathbf{n})$. Por lo tanto, no hay ningún $n \in \mathbb{N}$ que sea el número de Gödel de una demostración de Φ, pues si lo hubiera, por construcción de W_1, entonces $W_1(q, n)$. Pero, por ser representable en **P**, entonces $\mathbf{P} \vdash \omega_1(\mathbf{q}, \mathbf{n})$ y $\mathcal{N} \vDash \omega_1(\mathbf{q}, \mathbf{n})$, lo cual es una contradicción. La interpretación de $\mathcal{N} \vDash \Phi$ es que $\mathbf{P} \nvdash \Phi$. Es decir, enuncia en $\mathcal{N}$ su propia indemostrabilidad en **P**. Y resulta que esta sentencia es verdadera, pues Φ es indecidible: hay sentencias verdaderas que no se pueden deducir, y la que da Gödel es la sentencia que se interpreta por «no soy deducible».

- Podríamos pensar en añadir esta sentencia Φ, que no deja de ser una sentencia un tanto «especial», como axioma, pues es «verdadera». Con ello tendríamos la teoría extendida $\mathbf{P} \cup \Phi$, en la cual Φ es trivialmente decidible. Sin embargo, como ya dijimos, al añadir una nueva clase de axiomas que sea recursiva primitiva, nada cambia, y el método de prueba de Gödel sigue siendo válido. O sea, se podría volver a construir $\Phi^* \in \mathbf{Sent}(\mathbf{LA})$ indecidible en $\mathbf{P} \cup \Phi$.

El segundo teorema de incompletitud es un corolario del primero, en el sentido de que aporta una sentencia que no es deducible. Ahora bien, esta sentencia es muy particular, y es que la interpretación natural de la misma es que el sistema es consistente. Es decir, si se cumple la hipótesis del primer teorema, esto es, que $\mathbf{P}$ es consistente, entonces no podemos deducir que lo es, no podemos probar su consistencia.

La idea es construir la sentencia κ que exprese en la interpretación natural que el sistema formal es consistente (propiedad metamatemática). Bastará con una fórmula que exprese que $x \neq x$ no es deducible. Fijándonos en $\gamma(x)$ (2.4.4), ponemos

$$\kappa := \gamma(\mathbf{u}) = \forall y \neg \omega_1(\mathbf{u}, y), \tag{2.4.6}$$

donde u es el número de Gödel de la sentencia $\mathbf{0} \neq \mathbf{0}$, $u = g(\mathbf{0} \neq \mathbf{0})$. De manera similar a los argumentos usados en la sentencia de Gödel Φ, esta fórmula está expresando que no existe ninguna demostración de la sentencia con número de Gödel u, esto es, de $\mathbf{0} \neq \mathbf{0}$. O lo que es lo mismo, que $\mathbf{P}$ es consistente. Todos los razonamientos aritméticos usados por Gödel pueden formalizarse en $\mathbf{LA}$ (todos los conjuntos, relaciones y funciones que usa, son representables en $\mathbf{P}$), y se puede demostrar que

$$\kappa \to \Phi$$

es deducible. Es decir, que es un teorema:

$$\mathbf{P} \vdash \kappa \to \Phi.$$

Este es el paso crítico. A partir de aquí, el razonamiento es simple. Recordemos que Φ expresa la proposición metamatemática que afirma su indecibilidad, luego podemos decir que la fórmula $\kappa \to \Phi$ expresa que si $\mathbf{P}$ es consistente, entonces es incompleta (Φ se cumple, luego se cumple una fórmula indecidible, luego el sistema es incompleto). Vemos por tanto que $\kappa \to \Phi$ no es sino la formalización (que Gödel efectúa) de la proposición metamatemática que es el primer teorema de incompletitud. Pero si podemos deducir la fórmula κ, o sea, si $\mathbf{P} \vdash \kappa$, entonces, aplicando MP, tendríamos que la fórmula Φ es deducible. Pero en el primer teorema de incompletitud hemos demostrado que si **P** es consistente, entonces $\mathbf{P} \nvdash \Phi$. En consecuencia, no es posible que κ sea deducible. Es decir, no podemos probar formalmente la consistencia de **P**.

Corolario 2.4.19 (Segundo teorema de incompletitud)**.** Sea κ la sentencia $\gamma(\mathbf{u}) = \forall y \neg \omega_1(\mathbf{u}, y)$.

$$\mathbf{P} \text{ consistente } \Longrightarrow \mathbf{P} \nvdash \kappa.$$

Demostración.

Por reducción al absurdo: supongamos que $\mathbf{P} \vdash \kappa$.
Se tiene que $\mathbf{P} \vdash \kappa \to \Phi$.
Por MP a κ y $\kappa \to \Phi$, $\mathbf{P} \vdash \Phi$.
Contradicción con el primer teorema de incompletitud, luego $\mathbf{P} \nvdash \kappa$.

□

2.4.6. Sentencias indecidibles de la aritmética

Gödel ha probado que hay al menos dos sentencias formales que son indecidibles: Φ y κ. De aquí se deduce que hay enunciados aritméticos verdaderos que no son demostrables formalmente. Ahora bien, Φ y κ son sentencias un tanto especiales, que relacionan numéricamente (W_1) algunos números de Gödel involucrando directamente conceptos lógicos. Nos preguntamos, ¿Existen enunciados genuinamente aritméticos que sean indecidibles en **P**? La respuesta es que sí, y vamos a dar un ejemplo (para conocer más ejemplos, como el teorema de Goodstein, véase [3] Bovykin 2006).

Uno puramente aritmético indecidible es el llamado *teorema reforzado de Ramsey* para el caso finito. La teoría de Ramsey es una interesante rama de la combinatoria que estudia, a grandes rasgos, las propiedades estructurales de ciertas estructuras matemáticas, típicamente de los grafos, a través de ciertas características específicas de sus subestructuras.

Un resultado central es el teorema de Ramsey, que tiene varias versiones (véase [36] Woods 1979, cap. 9). En este caso damos la «versión aritmética», o sea, en forma de subconjuntos de $\mathbb{N}$.

Teorema 2.4.20 (de Ramsey (caso infinito))**.** Consideremos $\mathbb{N}$ y coloreemos los subconjuntos de $\mathbb{N}$ de tamaño m con c colores diferentes. Entonces existe algún subconjunto infinito B de $\mathbb{N}$ tal que todos sus subconjuntos de tamaño m tienen todos el mismo color.

Teorema 2.4.21 (de Ramsey (caso finito))**.** Para todo $k \in \mathbb{Z}^+$ y cualquier $r \in \mathbb{Z}^+$, existe $R(k,r) \in \mathbb{Z}^+$ tal que para cualquier coloración de los subconjuntos de tamaño k de $\{1,2,...,R(k,r)\}$ con r colores, siempre existe al menos un subconjunto B de $\{1,2,...,R(k,r)\}$ tamaño k en el que todos sus elementos tienen el mismo color.

El teorema reforzado de Ramsey para el caso finito es una extensión del teorema de Ramsey para conjuntos finitos. En pocas palabras, afirma que para cualquier conjunto finito de colores y cualquier secuencia finita de números, siempre existe un número lo suficientemente grande tal que en cualquier coloración de los subconjuntos de ese número con los colores dados, siempre hay un subconjunto de la secuencia en el que todos sus elementos tienen el mismo color.

Pues bien, el teorema de Paris-Harrington prueba que el teorema reforzado de Ramsey para el caso finito no es deducible en la aritmética de Peano de primer orden[45] (en **P**). Paris y Harrington probaron que mostrando que si existe una prueba en la aritmética de Peano, entonces el teorema reforzado de Ramsey para el caso finito implica la consistencia de **P**. Pero por el segundo teorema de incompletitud de Gödel esto es imposible, luego necesariamente no existe una prueba formal en la aritmética de Peano de primer orden del teorema de Ramsey finito reforzado.

Teorema 2.4.22 (de Paris Harrington)**.** El enunciado «para todo $k \in \mathbb{Z}^+$ y cualquier $r \in \mathbb{Z}^+$, existe $R(k,r) \in \mathbb{Z}^+$ tal que para cualquier coloración de los subconjuntos de tamaño k de $A = \{k+1, k+2, ..., R(k,r)\}$ con r colores, siempre existe al menos un subconjunto B de A con card(B) >min(B) en el que todos sus elementos tienen el mismo color» es indecidible en **P**[46].

2.4.7. Computabilidad

A modo de consecuencias del teorema de Gödel vamos a dedicar un breve apartado a la computabilidad de la mano del genial matemático británico Alan Turing (para ampliar información sin perder de vista los teoremas de Gödel, véase [[30] Pla i Carrera, 2012, cap.11]; y para una introducción formal y rigorosa a la computabilidad puede consultarse [[32] Rayward-Smith 1986]).

Hilbert y Ackermann se plantearon el famoso *problema de decisión* en el año 1928 (si bien Hilbert ya lo había involucrado en torno a la resolubilidad de ciertas ecuaciones diofánticas en el décimo problema en el Congreso de París de 1900). Un problema de decisión o decibilidad consiste en la posibilidad de idear un «proceso» o un «algoritmo» para la comprobación de la validez de una fórmula (es decir, que decidiera si es, o no, un teorema). En la lógica proposicional la respuesta es afirmativa, pero en lógica de primer orden es imposible, tal y como demostraron Alan Turing y Alonzo Church.

[45] Si bien es fácilmente demostrable en aritmética de segundo orden.

[46] Puede encontrarse una demostración en ([3] Bovykin 2006).

La relación con los teoremas de incompletitud va a venir con la tesis de Church, que identifica las funciones recursivas con las funciones computables. Primero veamos las definiciones básicas que podemos manejar, por ahora, en torno a la idea matemática de algoritmo.

Definición 2.4.23 (Conjunto recursivo/decidible)**.** Un conjunto A de número naturales es recursivo si para cualquier número natural, se sabe (o sea, existe un algoritmo que determina) si pertenece, o no, a A.

Con lo comentado en la sección 2.4.3 es claro que una definición equivalente es si A posee una función característica recursiva. Es decir, una función 1_A tal que

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

Definición 2.4.24 (Conjunto recursivamente numerable/listable)**.** Un conjunto A de número naturales es recursivamente numerable si es el conjunto vacío o si se pueden generar todos los elementos de A. Esto es, si existe una función $f : \mathbb{N} \to \mathbb{N}$ tal que $\mathrm{Im}(f) = A$.

La diferencia es sutil. Mientras que para un conjunto recursivo, disponemos de un algoritmo que nos dice si un elemento pertenece o no al conjunto, para uno recursivamente numerable disponemos de un algoritmo que va efectivamente construyendo, uno tras otro, los elementos del conjunto. Es claro que todo conjunto recursivo (decidible) es recursivamente numerable (listable), pero no a la inversa, y esto va a ser clave para resolver el problema de decisión.

Teorema 2.4.25. Todo conjunto de números naturales A decidible es listale.

Demostración. Si es el conjunto vacío, es trivial. Si no lo es, y 1_A es su función característica, que es recursiva, definimos $f(n) = 1_A(n) \cdot n + (1 - 1_A) \cdot a$, con $a \in A$ fijo, y $\mathrm{Im}(f) = A$. □

Para demostrar que a la inversa es falso, nos preguntamos si, conociendo una función f tal que $\mathrm{Im}(f) = A$, podemos saber si un n natural dado pertenece a ese conjunto. A es construible haciendo $f(1)$, $f(2)$, , .. etc. Pero, a pesar de ser construible, jamás será efectivamente construido debido a la infinitud de $\mathbb{N}$.

Para formalizar matemáticamente las nociones de algoritmo y computación, Alan Turing introdujo lo que hoy se conoce como máquinas de Turing. Vamos a proceder de la manera escrita en ([30] Pla i Carrera 2012, pp. 224-233), aunque para un tratamiento verdaderamente formal véase ([32] Rayward-Smith 1986, cap. 2). Por ejemplo, nuestra definición de máquina de Turing será heurística, y no formal.

Definición 2.4.26 (Máquina de Turing)**.** Consideraremos una máquina de Turing $\mathcal{M}$ como una cinta ilimitada, dividida en celdas individuales tal que:

- Cada celda individual, o bien está vacía, lo cual denotaremos por $\wedge$, o bien contiene un «palo» o marca, que denotamos por $|$.
- Consta de un lector que lee el contenido de una única celda tras cada cómputo.
- Tiene una colección de estados internos posibles, $q_0, q_1, q_2, ..., q_s$, de manera que tiene una «caja negra» que indica en cuál de ellos se halla la máquina.
- Para una entrada de k número naturales, en la cinta se escriben esos k números $(n_1, n_2, ..., n_k)$ con una celda vacía de separación entre cada uno de ellos. El resto de celdas están vacías. Un número n se representa en la cinta con $n+1$ marcas $|$ consecutivas[47].
- Se coloca el cabezal del lector a la derecha del último $|$ (el que esá más a la derecha) escrito en la cinta.
- El estado interno inicial es q_0.

Además, la máquina actúa a través de un programa (que podremos escribir en forma de matriz) ejecutando, en cada iteración, una de las siguientes acciones dependiendo del estado interno y de la lectura del cabezal:

- Borra lo que esté escrito en la celda que lee el cabezal, dejándola vacía ($\wedge$), y se coloca en un estado interno.
- Borra lo que esté escrito en la celda que lee el cabezal, dejando escrito un palo $|$, y se coloca en un estado interno.
- El cabezal se desplaza a la celda de la derecha (D), y se coloca en un estado interno.
- El cabezal se desplaza a la celda de la derecha (I), y se coloca en un estado interno.
- Cuando no haya ninguna instrucción para el estado interno, se para.

Ejemplo 2.4.27. Las instrucciones de operación para una máquina Turing $\mathcal{M}_s$ que describe la función de paso al siguiente $s : \mathbb{N} \to \mathbb{N}$, $s(n) = n+1$, se expresan a través de la matriz

$$\mathcal{M}_s = \begin{bmatrix} q_0 & \wedge & | & q_1 \\ q_1 & | & D & q_2 \end{bmatrix},$$

donde cada fila reresenta las iteraciones que se deben hacer según el estado inicial y las columnas representan, por orden de izquierda a derecha: el estado inicial, la

[47]Una celda vacía no es «nada», una marca $|$ es un 0, dos marcas $|$ es un 1, etc.

lectura del cabezal, la acción a ejecutar y el estado interno al que se llega en la misma. Es decir, esta matriz nos está diciendo que en el estado inicial el cabezal empieza en una celda vacía $\wedge$, y que entonces se debe escribir $|$ en la celda, pasando a estar en el estado interno q_1. La segunda fila nos indica qué se debe ejecutar cuando estamos en este estado interno y con el cabezal leyendo una celda con $|$, que es mover el cabezal una celda a la derecha. Tras ello el estado interno es q_2, un estado en el que $\mathcal{M}_s$ no actúa, pues no hay instrucciones.

La máquina funcionaría de la siguiente manera. Supongamos que tenemos $n+1$ marcas $|$, es decir, que la entrada de la máquina es el número natural n. Nos colocamos inicialmente (estado inicial q_0) a la derecha de la marca $(+1)$-ésima, esto es, en una celda vacía $\wedge$. Siguiendo la primera fila de la matriz, al leer $\wedge$, la maquina escribe $|$ y entra en el estado q_1. Ahora hay escritas $n+2$ marcas $|$, que representan el número natural $n+1$. Al estar en el estado interno q_1, siguiendo las instrucciones de la segunda fila de $\mathcal{M}_s$, al leer $|$ se desplaza a la derecha, que será una celda vacía $\wedge$, y entra en el estado q_2. Como para este estado interno no hay instrucciones, se para. Obtenemos entonces

$$\mathcal{M}_s(n) \downarrow n+1,$$

que quiere decir que ante la entrada n, escribe $n+1$ y se detiene. De esta forma se puede decir que la máquina de Turind $\mathcal{M}_s$ *computa* la función de paso al siguiente (pronto daremos la definición formal).

Es claro que para una misma función puede haber varias máquinas de Turing que la computen. Otros ejemplos sería el caso de la función binaria $+ : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, $+(m,n) = m+n$: $\mathcal{M}_+(m,n) \downarrow m+n$, y aunque la matriz no es compleja, es bastante amplia y no es necesario escribirla aquí. También puede darse el caso de una máquina de Turing que no se detiene nunca, como vemos en el siguiente ejemplo:

Ejemplo 2.4.28. Unas instrucciones de operación para una máquina Turing $\mathcal{M}_{bucle}$ que entra en un bucle infinito y no para nunca serían:

$$\mathcal{M}_s = \begin{bmatrix} q_0 & \wedge & \wedge & q_0 \end{bmatrix}.$$

En estos casos en que una máquina Turing, sea cual sea su entrada n, no computa nada, escribiremos $\mathcal{M}_{bucle}(n) \downarrow \infty$. En caso de que sepamos que en algún momento se para, pero no conozcamos lo que escribe en la cinta a su salida, pondremos simplemente $\mathcal{M}(n) \downarrow$.

También es posible dar ejemplos de máquinas de Turing que en algunos casos, dependiendo del valor que se introduce en la entrada, computa, y en otros no. Puede verse un ejemplo en ([30] Pla i Carrera 2012, p. 228).

Es claro que toda máquina de Turing $\mathcal{M}$ definine, para cada k, una función k-aria recursiva (parcial[48]) $f_{\mathcal{M}}^k$ de la siguiente forma:

$$f_{\mathcal{M}}^k(n_1,...,n_k) = \begin{cases} p & \text{, si } \mathcal{M}(n_1,...,n_k) \downarrow p, \\ \text{no está definida} & \text{, si } \mathcal{M}_{bucle}(n_1,...,n_k) \downarrow \infty. \end{cases}$$

Recordemos la importancia crucial de estas funciones en la demostración de Gödel. Este es el primer punto de contacto. Sin embargo, quizás no sea tan relevante como parece, pues existen muchos otros algoritmos o métodos, además de las máquinas de Turing, para computar las funciones parciales recursivas.

Otro punto de contacto interesante es lo siguiente. Al fin y al cabo, las máquinas de Turing, como algoritmo que son, no son más que hileras de signos. Podemos por tanto hacer una analogía con la gödelización de los sistemas formales que detallábamos en la sección 2.4.2, y «gödelizar» o «aritmetizar» las propias máquinas de Turing a través de la aplicación g: $\mathcal{M} \to g(\mathcal{M})$. De esta manera estamos numerando de forma unívoca las máquinas de Turing: $\mathcal{M}_1, \mathcal{M}_2, ..., \mathcal{M}_k, ...$. Les asignamos así símbolos del lenguaje formal, en concreto funciones monarias $\varphi_1, \varphi_2, ..., \varphi_k, ...$, que son $\varphi_i = f_{\mathcal{M}_i}^1$.

Así, gracias a este concepto de máquina de Turing, podemos definir una función computable (si bien, como hemos dicho, hay otros algoritmos válidos).

Definición 2.4.29 (Función computable)**.** Una función parcial $f : A \subseteq \mathbb{N}^k \to \mathbb{N}$ es computable Turing si existe una máquina de Turing $\mathcal{M}$ tal que, si su entrada es $(n_1, n_2, ..., n_k)$, entonces se para cuando en la cinta está escrito el número $f(n_1, n_2, ..., n_k)$, si $(n_1, n_2, ..., n_k) \in A$, y no se detiene nunca si $(n_1, n_2, ..., n_k) \notin A$. O sea

$$\begin{cases} \mathcal{M}(n_1,...,n_k) \downarrow f(n_1,n_2,...,n_k) & \text{, si } (n_1,n_2,...,n_k) \in A, \\ \mathcal{M}(n_1,...,n_k) \downarrow \infty & \text{, si } (n_1,n_2,...,n_k) \notin A. \end{cases}$$

Y podemos redefinir más rigorosamente algunos de los conceptos que hemos introducido antes y que son clave en la demostración de Gödel.

Definición 2.4.30 (Conjunto recursivo/decidible)**.** Un conjunto A de número naturales es recursivo si su función característica 1_A es computable a través de una máquina de Turing.

De esta manera ligamos perfectamente la decibilidad con la computabilidad a través de las máquinas de Turing. Esto será fundamental para resolver el problema de la parada.

[48] Recordemos que las funciones recursivas parciales son similares a las funciones recursivas pero donde se ha suprimido la restricción de que la operación de minimalización únicamente se aplica a funciones especiales.

Definición 2.4.31 (Conjunto recursivamente numerable/listable)**.** Un conjunto A de número naturales es recursivamente numerable si existe una función $f : \mathbb{N} \to \mathbb{N}$ tal que $\mathrm{Im}(f) = A$ que es computable a través de una máquina de Turing.

Fue utilizando estas ideas que Turing pudo resolver el problema de decisión. Gracias a la enumeración de Gödel tenemos un método para enumerar recursivamente (función de gödelización g) los teoremas de $\mathbf{P}$. De esta forma tenemos, en la aritmética, el conjunto de «teoremas» como el conjunto de números naturales que son el número de Gödel de un teorema:

$$\text{Teoremas}(\mathbf{P}) = \{n \in \mathbb{N} : n = g(\varphi), \text{ con } \mathbf{T} \vdash \varphi\}.$$

Y este conjunto es listable a través de la función g. De esta forma podemos ir enumerando, uno tras otro, los teoremas de $\mathbf{P}$. AHora bien, seguimos planteándonos el problema de decisión: ¿es decidible? Es decir, dado una fñormula φ, ¿podemos saber si $\mathbf{T} \vdash \varphi$ o no? La respuesta es negativa. No existe ninguna máquina de Turing que nos diga si un cierto $n \in \mathbb{N}$ corresponde al número de Gödel de un teorema de $\mathbf{P}$. O dicho de otro modo, la función característica $1_{\text{Teoremas}(\mathbf{P})}$ no es computable a travás de una máquina de Turing. Y con esto se resolvía finalmente el problema de decisión.

Decíamos que la demostración de Gödel tenía en su raíz la idea de las paradojas de autorreferencia. Turing se planteó el problema en el que una máquina de Turing se simula a sí misma. Para ello planteó la máquina *universal* de Turing, que podemos denominar $\mathcal{U}$. La entrada de esta máquina universal es el par de números naturales (m, n), y actúa siguiendo esta instrucción: «si m es el número de Gödel de la máquina de Turing $\mathcal{M}$, $m = g(\mathcal{M})$, entonces la actuación de $\mathcal{U}$ es $\mathcal{M}(n)$, y en caso contrario devuelve el valor 0». Esto es

$$\mathcal{U}(m,n) \downarrow \begin{cases} \mathcal{M}(n) & \text{, si } m = g(\mathcal{M}) \text{ para una cierta máquina de Turing } \mathcal{M}, \\ 0 & \text{, si para toda máquina de Turing } \mathcal{M}, m \neq g(\mathcal{M}). \end{cases}$$

La máquina universal, si tuviese tiempo suficiente, podría efectivamente computar lo que computa cualquier máquina de Turing (por la aritmetización de las máquinas de Turing).

Y relacionado con la idea de la autorreferencia está también el problema de la parada o «halting problem». En término generales este problema se refiere a la cuestión de determinar si es posible, a partir de un algoritmo y una entrada determinada, predecir si llegará un momento en que ese algoritmo se detendrá o, por contra, continuará ejecutándose indefinidamente. En otras palabras, el problema busca encontrar un algoritmo general que pueda decidir si cualquier otro algoritmo dado terminará su ejecución o no. A través de las máquinas de Turing podemos encontrar una solución.

Ya hemos comentado que se pueden numerar recursivamente las máquinas de Turing: $\mathcal{M}_1, \mathcal{M}_2, ..., \mathcal{M}_k, ...$ Consideremos ahora la función $P : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ definida de la siguiente forma:

$$P(m,n) = \begin{cases} 1 & \text{, si } \mathcal{M}_m \downarrow, \\ 0 & \text{, si } \mathcal{M}_m \downarrow \infty. \end{cases}$$

Se podría decir que es la función característica de la expresión «la m-ésima máquina de Turing con entrada n se detiene». La función P es la *función de parada.*

Y nos preguntamos: ¿es la función de parada computable Turing? Esto es lo mismo que preguntarse si el problema de la parada es decidible o no, o sea, si se puede encontrar un algoritmo que nos diga si, para un problema (por ejemplo, la pertenencia de un elemento a un conjunto), se puede encontrar un algoritmo que nos diga si el problema tiene solución. O en otras palabras: nos estamos preguntando por la decibilidad de la decibilidad.

Teorema 2.4.32 (de indecibilidad del problema de la parada)**.** La función de parada P no es computable.

Demostración. Primero consideremos $W_n = \{n : \varphi_n(n) \downarrow\}$, y veamos que no es decidible (recursivo). Supongamos que sí. Entonces, por definición, su función característica

$$1_{W_m}(n) = \begin{cases} 1 & \text{, si } \mathcal{M}_m \downarrow, \\ 0 & \text{, si } \mathcal{M}_m \downarrow \infty, \end{cases}$$

es computable Turing. Entonces la siguiente función también es computable

$$g(n) = \begin{cases} 0 & \text{, si } 1_{W_m}(n) = 1, \\ \text{no está definida} & \text{, si } 1_{W_m}(n) = 0. \end{cases}$$

En consecuencia, $\exists m \in \mathbb{N}$ tal que $g = \varphi_m$. Pero entonces

$$m \in W_m \iff \mathcal{M}_m \downarrow \iff g(m) = 0 \iff m \notin W_m,$$

que es una contradicción. Por tanto, W_n es decidible. Consideramos el conjunto

$$\mathfrak{H} = \{(m,n) : \mathcal{M}_m(n) \downarrow\}.$$

Y el problema de la parada queda planteado de la siguiente manera: $\mathfrak{H}$ es indecidible. Por reducción al absurdo: si fuese decidible, entonces su función característica $1_{\mathfrak{H}}$ sería computable Turing. También lo sería entonces la función $f(n) = 1_{\mathfrak{H}}(n,n)$.

Pero $f \equiv 1_{W_m}$, que hemos visto que no es computable. Hemos llegado a una contradicción. $\square$

Por último, tras haber comentado ya la utilidad de la computabilidad en la clasificación de los conjuntos, relaciones y funciones que Gödel utiliza en su prueba, plasmamos el teorema de Church, que finalmente identifica las funciones recursivas con las funciones computables.

Teorema 2.4.33 (de Church-Turing)**.** Las funciones computables Turing coinciden con las funciones parciales recursivas.

Corolario 2.4.34. Si f es una función total de número enteros positivos, entonces f es recursiva si, y sólo si, f es computable Turing.

Corolario 2.4.35. Una función f es representable en $\mathbf{P}$ si, y sólo si, f es computable Turing y total.

Conclusiones

Para concluir con este Trabajo Fin de Grado, sintetizaremos en las conclusiones los resultados más importantes. Comentaremos también la relación que guarda con las materias estudiadas en el grado en Matemáticas, y las posibilidades de continuar ahondando en estos temas.

A lo largo del texto hemos trabajado en dos ámbitos: la aproximación filosófica y la aproximación formal a los teoremas.

En el primero hemos repasado las corrientes clásicas más importante en filosofía de la Matemática. En primer lugar hemos justificado, tomando como base la idea de Filosofía de Gustavo Bueno, la pertinencia de este apartado filosófico dentro de un trabajo de matemáticas. A lo largo de las cuatro secciones de este capítulo hemos citado abundantemente a los matemáticos y lógicos que representan dichas filosofías. Y es que, como decíamos, la metodología de la Matemática entra a la vez en el terreno de la filosofía y en el de la lógica. Creemos que en esta síntesis hemos conseguido enlazar las distintas ideas filosóficas con los momentos históricos que las representan, desenbocando al final en los teoremas de incompletitud de Gödel. Así hemos podido comprender bien las posibles respuestas, esencialmente, a la magna pregunta: ¿qué es la Matemática? A modo de conclusión resumamos las cuatro posturas fundamentales:

- **Platonismo.** Hemos distinguido, en la línea de José Ferreirós Domínguez, entre lo que podríamos llamar un platonismo fuerte y uno débil. El primero recoge la vertiente clásica que considera que los objetos matemáticos tienen existencia por sí mismos, independientemente de nosotros. Así, cuando llegamos a un resultado nuevo en matemáticas, estamos descubriendo algo que ya existía. Defensores de esta postura, y de forma crucial para sus trabajos en matemáticas, fueron Bourbaki, Cantor y Gödel. Por contra, el platonismo débil es inevitable en las matemáticas modernas, en las que damos por existentes todo lo que sea admisible (no contradictorio). Se podría resumir en la asunción de la existencia, como totalidad, del conjunto de los número naturales.

- **Logicismo.** Esta corriente reduce estrictamente las matemáticas a la lógica, que existe primariamente. En palabras de Frege, los entes lógicos existían al ser las leyes de las leyes de la Naturaleza. Así, se entienden las verdades matemáticas como verdades analíticas (son tautologías, en la línea de Wittgenstein), pues dependen únicamente de usar correctamente las reglas lógicas. Fue el mismo Frege quien dio inicio a la lógica moderna y al programa logicista. Tras el escándalo de las paradojas de autorreferencia, Russell y Whitehead redactaron los *Principia Mathematica*, sistema con el que trabaja Gödel en la demostación de sus teoremas. A pesar de ser una fundamentación de la Matemática, los *Principia* adolecían de parecer sospechosamente *ad hoc* para evitar las antinomias, además de contener en su base una teoría de Clases que posponía el problema: o bien el logicismo es una tesis errónea, o bien se fundamenta en una Teoría de Conjuntos.

- **Intuicionismo.** Las tesis intuicionistas, que se basan en el constructivismo matemático, defienden que la esencia de las matemáticas reside en nuestra intuición, que ésta es el motor primario de las mismas. Tras la llegada de las geometrías no euclídeas y el tratamiento cada vez más abstracto en la aproximación moderna a las matemáticas, el intuicionismo no ha podido sino retroceder. A pesar de ello, casi nadie se atreverá a negar un papel importante a lo que llamamos «intuición matemática» en el Hacer Matemático. En este apartado hemos tratado con los matemáticos que desarrollaron una nueva matemática intuicionista, como Brouwer, Heyting o Poincaré, así como los orígenes kantianos de estas ideas. Posteriormente hemos analizado la base constructivista y aristotélica de los *Elementos* de Euclides, entre otras cosas.

- **Formalismo.** Mientras que el platonismo y el logicismo sustantivizan la verdad matemática como algo externo a nosotros, el formalismo vuelve al boli y el papel para decir que la Matemática no es sino el juego de signos carentes de significado que realizamos los matemáticos de carne y hueso. Estos juegos de signos se ordenan en los sistemas formales, cada uno con sus reglas de transformación. Hemos hablado *in extenso* del Programa Formalista de Hilbert para fundamentar las matemáticas, clave en el desarrollo de los teoremas de incompletitud. Además hemos introducido la noción de modelo a través de un ejemplo, y hemos analizado los conceptos de axiomas e independencia, para terminar con una primera arpoximación a los teoremas de Gödel.

En el segundo capítulo tratamos una aproximación puramente formal a los teoremas de incompetitud, cuyo objetivo principal ha sido la demostración de los mismos. Hemos definido las teorías formales de primer orden rigorosamente. Primero explicando la sintaxis y la semántica de los lenguaje formales de primer orden, y la importancia de ambas partes y su relación. Hemos logrado clarificar en detalle qué se entiende por «verdad» y «deducción» en matemáticas, sin duda una de las claves del trabajo. Tras ello hemos podido definir formalmente la decibilidad, la consistencia y la completitud de una teoría formal, conceptos esenciales en los teoremas de incompletitud.

A continuación hemos abordados ideas introductorias en torno a estos términos, como comprobar la completitud y consistencia de la colección de sentencias verdaderas en un modelo de la toería. También dedicamos un apartado a establecer algunos de los metateoremas lógicos más importantes previos a los teoremas de incompletitud. Por ejemplo, el teorema de completitud de Gödel, del que damos una demostración parcial, que identifica las verdades en todos los modelos de una teoría formal de primer orden con los teoremas de la misma. También la generalización del mismo en el teorema de Henkin, los teoremas de compacidad y el teorema de Löwenheim-Skolem-Tarski. En relación a este último, hemos dado un ejemplo de construcción de un modelo no estándar para la aritmética de Peano de primer orden, construyendo lo que hemos llamado un «número infinito», esto es, un número que es más grande que cualquier número natural.

Dentro de este segundo capítulo podemos distinguir dos partes: la primera, la anteriormente comentada, y la segunda la relativa al macroapartado de los teoremas de Gödel. En esta segunda parte, que corresponde a la sección 2.4, hemos desarrollado ampliamente las distintas ideas que Gödel utiliza en la demostración de sus teoremas de incompletitud. Posteriormente detallamos los teoremas y sus demostraciones, y dedicamos dos apartados a modo de consecuencias de los mismos. Sinteticemos estas ideas fundamentales:

- **Introducción de Gödel.** Comentamos la introducción original de Gödel, muy útil para entender el planteamiento general de demostración.

- **Gödelización.** Estudiamos en detalle el sistema creado por Gödel para aritmetizar el sistema formal **P** a través de la numeración de Gödel. De esta forma todo símbolo, fórmula o sucesión de fórmulas del sistema formal tiene asociado un número de Gödel, estableciendo un puente entre proposiciones metamatemáticas y propiedades aritméticas aritmetizando el lenguaje formal.

- **Teoría de la Recursión.** Gödel desarrolla la teoría de la recursión, que se convierte en el paso previo fundamental para clarificar qué relaciones, funciones y conjuntos de $\mathbb{N}$ (que, como hemos dicho, representan proposiciones metamatemáticas) son representables en la teoría **P**.

- **Representabilidad sintáctica.** Para la demostración de los teoremas de incompletitud, Gödel precisa que expresiones como «φ es una fórmula de **P**», «φ es un axioma» o «φ es demostrable», que de ninguna manera pertenecen a **P**, puedan, sin embargo, representarse en el sistema. Para ello debe formularlas como relaciones recursivas primitivas, y demuestra que son precisamente estas relaciones las que son representables en **P**.

- **Demostraciones.** En este apartado presentamos unas demostraciones adecuadas de los dos teoremas de incompletitud, además de una explicación detallada de los mismos: algunos conceptos previos, como la ω-consistencia y los conjuntos primitivo recursivos, y algunas conclusiones y aclaraciones.

- **Sentencias indecidibles de la aritmética.** Exponemos un ejemplo de sentencia puramente aritmética que, ni ella ni su negación, pueden demostrarse formalmente en **P**. Se trata del teorema de Paris-Harrington, que establece la indecibilidad del teorema reforzado de Ramsey para el caso finito.

- **Computabilidad.** Como apartado relativo a consecuencias del teorema de Gödel, hablamos de computabilidad. Para ello estudiamos el concepto de decibilidad y el inicio de la computación. A través de las máquinas de Turing somos capaces de definir formalmente la computabilidad y resolvemos el problema de la parada.

Un texto que trata un tema tan ambicioso como este siempre adolece de no ser del todo completo (siempre se puede tirar de algún hilo colindante) y rigoroso. En cuanto al nivel de detalle, creemos que se ajusta a las exigencias para un Trabajo de Fin de Grado, y que no trata los teoremas de incompleitud desde una perspectiva divulgativa, que es lo único que los estudiantes de matemáticas suelen conocer, sino formal. Además, hemos tratado las dos vertientes del teorema *in extenso*, algo poco habitual. Los libros estrictamente de lógica matemática ignoran la aproximación fislosófica, y raramente los manueales de filosofía universitaria estudian el tema en profundidad. Sí que hemos de decir que, en la literatura especializada en filosofía de las matemáticas, el alto nivel de la aproximación lógica es habitual, y en España tenemos grandes pensadores que así lo han demostrado, como Javier de Lorenzo o Josep Pla i Carrera.

Podríamos continuar trabajando en el texto ampliando el tratamiento relativo a la Teoría de Conjuntos. Por ejemplo estudiando la independencia axiomática de la misma, sus modelos o sus sentencias indecidibles (como la hipótesis general del continuo).

A lo largo de este Trabajo Fin de Grado no hemos utilizado, realmente, muchos conceptos de los estudiados en las asignaturas del grado en Matematicas, pues no hay posibilidad de cursar ninguna asignatura de lógica matemática, y menos aún de filosofía. Aún así, estos teoremas son, como hemos comentados, con todo derecho, teoremas matemáticos. Acaso se ha comentado algo en la asignatura de Topología relativo a la definición de conjunto, la paradoja de Russell, los conjuntos infinitos o independencia de axiomas en la teoría axiomática de conjuntos, entre otras cosas. Pero no fue parte importante de ningún tema de la asignatura. Sin embargo, es fundamental conocer las matemáticas si se pretende, tanto hablar de filosofía de las matemáticas, y más en un tema que abarca cuestiones como la «verdad» y la «demostración», como estudiar lógica matemática, en particular los teoremas de incompletitud de Gödel. Por ello, si bien estos temas no se han tratado a lo largo del grado, es todo el grado, en cierto modo, el que se trata en estos temas.

Vale.

Apéndice A

La prueba de la existencia de Dios: el argumento ontológico

El argumento ontológico es un argumento clásico sobre la existencia de Dios que se remonta al teólogo del siglo XI San Anselmo de Canterbury. Como el propio nombre (dado por Kant) indica, se basa en la ontología de Dios, es decir, en «lo que Dios es», en su categoría existencial o esencial. Más precisamente trata de probar que la ontología necesaria que debería tener aquello que llamamos Dios implica necesariamente su existencia, pues el no existir iría en contra de sus atributos ontológicos.

La idea general de San Anselmo es la siguiente: Dios es, por definición, el ser más perfecto que puede ser pensado. Este ser existe en nuestra mente, podemos concebirlo. Ahora bien, si el ser más perfecto existe en nuestra mente y no existe en la realidad, entonces es que no es el ser más perfecto, o que no es del todo perfecto, pues el más perfecto debe tener el atributo de la existencia[1]. Por lo tanto, ese ser perfecto que llamamos Dios y existe en nuestra mente debe existir también fuera de ella necesariamente.

El argumento ha sido discutido durante siglos. Una crítica breve es argüir que la existencia no es una cualidad, o sea, no es un predicado, sino un «cuantificador», de forma similar a la manera en que se trabaja en lógica formal (para decir que existe una variable x que cumple la fórmula φ, escribimos $\exists x \varphi(x)$, y no $E(x)$, $\varphi(x)$ con E un predicado de x que le otorga la propieda de existir). Por otra parte Kant critica, en consonancia creemos con su constructivismo matemático (véase p. 30), que la mera definición sin contradicción, a través de ciertas propiedades, no implica la existencia. De hecho, podemos pensar en ejecutar el argumento, en lugar de para un ser perfecto, para una «isla perfecta», o cualquier otra cosa, y demostrar su existencia. Otra cuestión sería, además, como planteaba Santo Tomás, si los seres humanos

[1] En la prueba de Gödel es el axioma A.0.11, que dice que la propiedad de existir necesariamente es una propiedad positiva.

pueden conocer la naturaleza de Dios. Esto nos recuerda al dilema de Benacerraf sobre la posibilidad de que los humanos (contingentes) conozcan las proposiciones matemáticas si estas son de naturaleza platónica (necesarias) que vimos en la p. 20[2].

En cualquier caso este tema, que se sigue debatiendo hoy, es tan amplio como interesante, y el lector puede encontrar abundante bibliografía. Lo que nos interesa aquí es la prueba lógica que Gödel redactó en 1941, un trabajo que nunca llegó a publicar. En ella trabaja con la lógica modal (de segundo orden), una lógica avanzada en la que hay dos *modos* de verdad: la verdad necesaria y la verdad contingente. Con ello trata de recoger las expresiones «es necesario que» y «es posible que», o dicho de otro modo, «verdad en todos los mundos posibles» y «verdad en algún mundo posible». Esto último no es sino una interpretación de la lógica modal. Resumamos algunas características esenciales de la lógica modal para poder entender la prueba gödeliana:

- Si una afirmación p es cierta en todos los mundos posibles, entonces es necesaria, y lo denotamos: $\Box p$.
- Si una afirmación p es cierta en algún mundo, entonces es posible, y lo denotamos: $\Diamond p$.
- No todos los mundos tienen los mismos «objetos».
- Una propiedad $\varphi(x)$ asigna a *cada* «objeto» x, en *cada* mundo posible, un valor lógico: verdadero o falso.
- La «implicación» en lógica modal se refiere a cualquier mundo posible, y se denomina «implicación modal». O sea, que la propiedad φ implica la propiedad ψ quiere decir que, en cualquier mundo posible donde un objeto tenga la propiedad φ, entonces en ese mundo posible también tiene la propiedad ψ.

Un ejemplo sencillo es la propiedad de existir un objeto $x > 0$ tal que $x^2 - 2 = 0$. Para un objeto x, tener esta propiedad es posible (si es un número real, un mundo posible) pero no la tiene necesariamente (si es un número racional, otro mundo posible). Por tanto, la afirmación $\varphi(x) :=$«existe un objeto $x > 0$ tal que $x^2 - 2 = 0$» es posible, pero no necesaria, $\Diamond p$, $\neg\Box p$. Si consideramos la propiedad $\psi(x) :=$«x no es racional», tenemos que $\varphi(x) \to \psi(x)$.

En la demostración va a ser fundamental la definición de propiedad positiva P, una propiedad que se aplica a las propiedades φ. Si una propiedad es positiva escribiremos $P(\varphi)$. Además, asumiremos que la positividad de un propiedad es independiente de los objetos que la satisfacen, luego su es positiva para un objeto, lo es para todos. Esta propiedad es un punto delicado, pues en principio parece haber

[2] Ya que vamos a dar una «prueba» del argumento ontológico, nos hemos permitido indicar en este párrafo algunas de las contraargumentaciones habituales.

propiedades «positivas» contradictorias entre sí, como por ejemplo la «justicia» y la «misericordia» la «libertad» y el «bien» (al menos en cuanto al hecho de que la libertad es también la libertad para hacer el mal; ¿es preferible que no haya mal pero no seamos libres, o que seamos libres pero haya mal?).

En cualquier caso, estamos ya en condiciones de abordar la demostración de Gödel (la tomamos de [[30] Pla i Carrera 2012, cap. 1]). Consta de cinco axiomas, 3 definiciones y 5 teoremas.

Como axioma previo que ya hemos adelantado antes, suponemos que del conjunto de propiedades podemos distinguir y separar las propiedades positivas.

Axioma A.0.1. Si φ es una propiedad positiva, y φ implica ψ, entonces ψ también es una propiedad positiva.

Axioma A.0.2. Si φ es una propiedad positiva, entonces o bien φ es cierta, o bien $\neg\varphi$ es cierta, pero no ambas.

Teorema A.0.3. Si φ es una propiedad positiva, entonces existe un objeto x en algún mundo posible que la satisface: $\varphi(x)$.

Demostración.

Por reducción al absurdo: supongamos que φ es positiva y no hay objeto que la satisfaga en ningún mundo.
Esto es, en todos los mundos posibles, no hay objetos con la propiedad φ.
O lo que es lo mismo, $\Box\neg\varphi$.
Por el axioma proposicional 1, $\varphi \to \neg\varphi$.
Trivialmente, $\varphi \to \varphi$.
Como φ es positiva, entonces $\neg\varphi$ también lo es.
Contradicción con el axioma A.0.2.

$\square$

Definición A.0.4 (Propiedad divina G, Dios)**.** Sea x un objeto en algún mundo posible. $G(x)$ es cierto si, y sólo si, en ese mundo, x satisface todas las propiedades positivas. A un objeto que tiene la propiedad divina lo denominamos Dios[3].

Axioma A.0.5. La propiedad divina G es positiva ($P(G)$).

Teorema A.0.6. En algún mundo hay un objeto que es Dios.

Demostración. Resulta de aplicar el teorema A.0.3 a la propiedad divina G, que es positiva. $\square$

[3]De esta forma ninguna propiedad negativa ψ puede ser divina, pues en tal caso la propiedad positiva $\neg\psi$ no se podría satisfacer. Dios no puede tener propiedades negativas.

Definición A.0.7 (Esencia)**.** Una propiedad φ es una esencia del objeto x si, y sólo si,

1. El objeto tiene la propiedad φ, $\varphi(x)$.
2. Para toda cualquier otra propiedad satisfecha por x, $\psi(x)$, φ implica necesariamente ψ.

Axioma A.0.8. Una propiedad positiva es una propiedad positiva en todos los mundos posibles.

Teorema A.0.9. Si x es Dios, esto es, $G(x)$, entonces la propiedad divina G (la propiedad de ser Dios) es una esencia de x.

Demostración.

Por hipótesis, $G(x)$.
Sea φ otra propiedad de x, $\varphi(x)$. Veamos que G implica necesariamente φ.
Por definición, Dios no puede tener propiedades negativas, luego φ es una propiedad positiva.
Por el axioma A.0.8, φ es una propiedad positiva en todos los mundos.
Así, $\varphi(x)$ en todos los mundos.
En cualquier mundo posible, $G(x)$ implica $\varphi(x)$.
$G(x)$ implica necesariamente $\varphi(x)$.

□

Definición A.0.10 (Necesariamente Existente NE)**.** Un objeto x es necesariamente existente (NE) si, y sólo si, satisface la siguiente propiedad: Para cada esencia φ de x, en todos los mundos posibles existe un objeto y que satisface la propiedad φ[4].

Axioma A.0.11. La propiedad NE es una propiedad positiva.

Teorema A.0.12. Existe un Dios en todos los mundos posibles.

Demostración.

Por el axioma A.0.11, NE es positiva.
Por el axioma A.0.8, NE es positiva en todos los mundos posibles.
Por el teorema A.0.6, hay un Dios (x tal que $G(x)$) en algún mundo.
Por tanto, en algún mundo Dios es NE (x es NE).
Por el teorema A.0.9, la propiedad divina G es una esencia de x.
Por la definición de NE, en todos los mundos posibles existe un objeto que satisface la propiedad G.
Esto es, en todos los mundos existe un Dios.

□

[4]Es decir, un objeto existe necesariamente si, y sólo si, todas sus esencias se satisfacen en todos los mundos posibles.

Corolario A.0.13. No puede haber dos propiedades que distingan dos dioses.

Demostración.

Por reducción al absurdo: supongamos que hay dos dioses, uno de ellos satisface la propiedad φ y el otro no.
Por ser la propiedad de un Dios, φ debe ser positiva.
Por el axioma A.0.8, φ es una propiedad positiva en todos los mundos.
Pero entonces hay un Dios que no tiene una propiedad positiva, pues no satisface φ.
Contradicción con la definición de Dios.

$\square$

Hasta quí la prueba. A continuación lo presentamos todo (sin demostraciones) formalmente (escribimos φ ess x para expresar que la propiedad φ es una esencia del objeto x):

Ax. 1. $P(\varphi) \wedge \Box\forall x(\varphi(x) \to \psi(x)) \to P(\psi)$.
Ax. 2. $P(\neg\varphi) \leftrightarrow \neg P(\varphi)$.
T. 1. $P(\varphi) \to \Diamond\exists x \varphi(x)$.
Def. 1. $G(x) \Leftrightarrow \forall(P(\varphi) \to \varphi(x))$.
Ax. 3. $P(G)$.
T. 2. $\Diamond\exists x G(x)$.
Def. 2. $\varphi \text{ ess } x \iff \varphi(x) \wedge \forall\psi(\psi(x) \to \Box\forall x[\varphi(x) \to \psi(x)])$.
Ax. 4. $P(\varphi) \to \Box P(\varphi)$.
T. 3. $G(x) \to G \text{ ess } x$.
Def. 3. $\mathrm{NE}(x) \iff \forall\varphi\ \varphi \text{ ess } x \to \Box\exists y \varphi(y)$.
Ax. 5. $P(\mathrm{NE})$.
T. 4. $\Box\exists x G(x)$.
Cor. 1. $G(x) \wedge G(y) \to \forall\varphi\ \varphi(x) \iff \varphi(y)$.

Para otra prueba de la existencia de Dios véase ([10] Einstein, 1915).

Bibliografía

[1] ARANDA, Joaquín, et al. (2000): *Fundamentos de lógica matemática*, Sanz y Torres, Madrid.

[2] BERNAYS, Paul (1935): «Sur le platonisme dans les mathématiques», *L'Enseignement Mathématique*, 34, pp. 52-69.

[3] BOVYKIN, Andrey, (2006): «Brief introduction to unprovability», Logic Colloquium 32, pp. 38-64.

[4] BUENO, Gustavo (1995a): *¿Qué es la filosofía?*, Pentalfa Ediciones, Oviedo.

(1995b): *¿Qué es la ciencia?*, Pentalfa Ediciones, Oviedo.

[5] DAVIS, P. J. & HERSH, R (1982): *Experiencia matemática*, Labor, Barcelona.

[6] DE LORENZO, Javier (1979): *Lógica y Matemática en K. Gödel*, Estudios Filosóficos, N°79, Vol. XXVIII.

[7] DIEUDONNÉ, Jean (1970): «The Work of Nicolas Bourbaki», *American Mathematical Monthly*, 77, pp. 134-145.

[8] DOU, Alberto (1974): *Fundamentos de la matemática*, Labor, Barcelona.

[9] DUMMETT, Michael (19743): *FREGE: Philosophy of Language*, Harper & Row Publishers, Nueva York. Edición online: https://es.scribd.com/document/239470905/Michael-Dummett-Frege-Philosophy-of-Language

[10] EINSTEIN, Albert (1915): «Las ecuaciones de campo de la gravitación», Academia de ciencias de Prusia, informes de sesión (parte 2), pp. 844-847. Original disponible online en: http://www.jp-petit.org/papers/cosmo/1915-Einstein-25-nov-de.pdf

[11] EUCLIDES (III aC): *Elementos*, Edición online en: https://euclides.org/los-elementos/.

[12] FERNÁNDEZ MARGARIT, Alejandro (2012): *Lógica Matemática y Teoría de Conjuntos*, editado por Andrés Cordón Franco y F. Félix Lara Martín, Fénix Editora, Sevilla.

[13] FERREIRÓS DOMÍNGUEZ, José (1999): «Matemáticas y platonismo(s)», *Gaceta de la Real Sociedad Matemática Española*, 2/2, pp. 446-473.

[14] FEYERABEND, Paul (1975): *Against Method*, New Left Books, Londres.

[15] FREGE, Gottlob (1): *Escritos filosóficos*, De Crítica, Barcelona,1996.

(2): *Philosophical and mathematical correspondence*, University of Chicago Press, Chicago, 1980.

(1884): *Fundamentos de la Aritmética*, Laia, Barcelona, 1972.

[16] FROLOV, Iván T. (1987): *Diccionario de filosofía*, Progreso, Moscú.

[17] GEORGE, A. & VELLEMAN, D. (1982): *Philosophies of mathematics*, Blackwell, Oxford.

[18] GÖDEL, Kurt (1944): «La lógica matemática de Russell», *Obras Completas*, Alianza, Madrid, 2018.

(1931): «Sobre sentencias formalmente indecibles de Principia Matematica y sistemas afines», *Obras Completas*, Alianza, Madrid, 2018.

[19] GOLDSTEIN, Rebecca (2005): *Gödel. Paradoja y vida*, Antoni Bosch, Barcelona.

[20] HILBERT, David (1898): *Fundamentos de la geometría*, CSIC, Madrid, 1991.

(1899): *Fundamentos de Geometría*, CSIC, Madrid 1953.

(1917): *Pensamiento axiomático*, Educación Matemática Vol. 11 No. 2 Agosto 1999 pp. 128-136. Disponible en: http://www.revista-educacion-matematica.org.mx/descargas/Vol11/2/10Hilbert.pdf

[21] KANT, Immanuel (1790): *Crítica del juicio*, Austral, Madrid, 1970.

[22] KOLMOGÖROV, A. & DRAGALIN, A. G. (1970): *Introducción a la lógica matemática*, Ediciones URSS, Moscú, 2013.

[23] MADRID CASADO, Carlos M. (2009): «Filosofía de las Matemáticas. El cierre de la Topología y la Teoría del Caos», *El Basilisco*, 41, pp. 1-48.

[24] MARTÍNEZ PÉREZ, Mariano (1991): «Algo más sobre la historia de la Teoría de la Medida(1904-1930): La paradoja de Banach-Tarski», *Seminario de historia matemática I*, Universidad Complutense, Madrid, pp. 335-430.

[25] MESCHKOWSKI, Herbert (1983): *Georg Cantor. Leben, Werk und Wirkung*, Bibliographisches Institut, Mannheim.

[26] MOSTERÍN, Jesús (2000): *Los lógicos*, Espasa Calpe, Madrid.

[27] NAGEL, E. & NEWMAN, J. R.. (1958): *El teorema de Gödel*, Epublibre Titivillus, 2020.

[28] PADILLA, Jesús (2007): *Verdad y demostración*, Plaza y Valdes, Madrid.

[29] PENROSE, Roger (2006): *La nueva mente del Emperador* , De Bolsillo, Barcelona.

[30] PLA I CARRERA, Josep (2012): *El Teorema de Gödel. Un análisis de la verdad matemática.*, Real Sociedad Matemática Española.

[31] POINCARÉ, Henri (1902): *Ciencia e Hipótesis*, Edición digital para la Biblioteca Digital del ILCE: file:///C:/Users/User/Downloads/ciencia-e-hipotesis.pdf

[32] RAYWARD-SMITH, V. J. (1986): *A first course in computability*, Blackwell Scientific Publications.

[33] RUSSELL, Bertrand (1920): *Introduction to mathematical philosophy*, George Allen & Unwin, Ltd., London.

[34] SOKAL, Alan & BRICMONT, Jean (1999): *Imposturas intelectuales*, Ediciones Paidós, Madrid.

[35] WITTGENSTEIN, Ludwig (1921): *Tractatus logico-philosophicus*, Alianza, 2021.

[36] WOODS, Donald R. (1979): *Notes on introductory combinatorics*, School of Humanities and Sciences, Stanford University.