Facultad
de
Ciencias

LOS TEOREMAS DE INCOMPLETITUD DE GÖDEL

(Gödel's Incompleteness Theorems)

Trabajo de Fin de Grado
para acceder al
GRADO EN MATEMÁTICAS

Autor: Jorge García López
Director: Luis Felipe Tabera
Julio - 2023

Resumen

En esta memoria, explicamos y demostramos los Teoremas de Incompletitud de Gödel, que muestran que no puede existir una teoría completa, consistente y recursiva que abarque todas las matemáticas, publicados en 1931. Para ello, primeramente damos una breve introducción a la lógica en la que están enunciados, la lógica de primer orden, y explicamos sus características. También analizaremos con cierto detalle la técnica de gödelización. También abrimos y cerramos la memoria explicando el interés de éstos en la lógica matemática.

Palabras Clave:

Lógica, Teoremas de Incompletitud, Gödel, Fundamentos de las matemáticas.

Abstract

In this report, we discuss and prove Gödel's Incompleteness Theorems, that prove that there can not be a recursive, complete and consistent theory that encompasses all mathematics, published in 1931. First, a short introduction to first order logic is provided, along with some of its properties. We will also analyze the so called gödelization technique. Also, we start and conclude the exposition discussing their interest in mathematical logic.

Keywords:

Logic, Incompleteness Theorems, Gödel, Foundations of mathematics.

Índice

1. Contexto del problema **1**
- 1.1. Characteristica universalis 1
- 1.2. Principia Mathematica 1
- 1.3. Programa de Hilbert 2
- 1.4. Lógica de Primer Orden 3

2. Sintaxis de la lógica de primer orden **4**
- 2.1. Alfabeto 4
- 2.2. Términos y fórmulas 5
- 2.3. Axiomas 6
- 2.4. Demostraciones 7
- 2.5. Consistencia 10

3. Semántica de la lógica de primer orden **11**

4. La Axiomática de Peano **15**
- 4.1. Nuestro modelo de los números naturales $\mathbb{N}$ 17
- 4.2. Resultados necesarios para los teoremas 19

5. Gödelización de la Aritmética de Peano **22**
- 5.1. La función β de Gödel 24
- 5.2. Codificación 28

6. Los Teoremas de Incompletitud **36**
- 6.1. Primer Teorema de Incompletitud 36
- 6.2. Segundo Teorema de Incompletitud 38

7. Consecuencias de los Teoremas **42**

1. Contexto del problema

En 1931, Kurt Gödel, un matemático austriaco, publica un artículo llamado *Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I* [Göd31], donde son presentados unos enunciados en la aritmética de Peano que no son demostrables, enunciando con ellos los llamados *Teoremas de Incompletitud de Gödel*. Esto responde a una pregunta que llevaba abierta varios siglos.

1.1. Characteristica universalis

En 1619 el matemático y filósofo René Descartes, inspirado por un sueño, concibe la idea de la existencia de un método, la *"science admirable"*, que abarcase todas las ciencias, tanto teóricas como empíricas, permitiendo descubrir matemáticamente todas sus verdades y falsedades.

A partir de esta idea, Descartes construye la geometría analítica y escribe su *Discurso del método* [Des37] donde es desarrollada. Con estas obras como impulso, se extiende en las ciencias y el pensamiento de la Europa moderna un afán de formalización y justificación última, que se puede observar, por ejemplo, en el desarrollo del mecanicismo o en el surgimiento del racionalismo.

Dentro de esta última corriente se encuentra el filósofo Gottfried Leibniz, quien, siguiendo esos intentos, en 1714, esboza una lógica universal llamada *Characteristica universalis*, que permitiría formalizar todo nuestro pensamiento: dada una pregunta enunciada en el lenguaje de cualquier ciencia, permitiría obtener su respuesta a partir de mero cálculo.

Ya que este lenguaje pretende poder enunciar toda idea de la razón, también los enunciados sobre dicho lenguaje deberían poderse enunciar en él. Ante lo cual, uno se podría preguntar qué ocurriría aquí con oraciones del tipo "Este enunciado es falso" o la conocida paradoja del barbero. En una ciudad hay un barbero que afeita exactamente a aquellas personas que no se afeitan a sí mismas. ¿Quién afeita al barbero?

1.2. Principia Mathematica

Como se discute en [LI], el sueño de Leibniz resonaba en las matemáticas del siglo XIX, desarrollando la formalización de sus ramas con un afán de universalización, pero sin ningún sistema que permitiera englobarlas todas hasta 1893. En ese año, el matemático alemán Gottlob Frege publica, en su

Grundgesetze der Arithmetik, la primera formalización de las matemáticas. Con este trabajo, se inicia, en el siglo XX, el proyecto de construir el sueño de Leibniz para las matemáticas. La creencia de que se puedan reducir todas las matemáticas a la lógica de esa forma se llama logicismo. A partir un conjunto de axiomas y unas reglas de inferencia, sería posible deducir todos los resultados definidos en matemáticas.

En 1901, el filósofo Bertrand Rusell encuentra una contradicción a partir de los axiomas que supone Frege en su formalización de las matemáticas. Dicha contradicción, llamada Paradoja de Rusell, es posible ya que en la axiomática de Frege se pueden construir conjuntos formados por todos los objetos que verifican una propiedad. Tomando como propiedad el no pertenecerse así mismo, se obtiene el conjunto

$$R = \{x \mid x \notin x\},$$

que cumple la propiedad de que $R \in R$ si y sólo si $R \notin R$.

Como alternativa al sistema inconsistente de Frege, Rusell construye la teoría de tipos, ya que no incluiría, en principio, esta contradicción. Publica, junto con filósofo Alfred North Whitehead, en 1913, su propuesta de formalización en el *Principia Mathematica*, [WR13]. Allí encontramos:

> The present work has two main objects. One of these, the proof that all pure mathematics deals exclusively with concepts definable in terms of a very small number of fundamental concepts, and that all its propositions are deducible from a very small number of fundamental logical principles, is undertaken in Parts II–VII of this work, and will be established by strict symbolic reasoning in Volume II. ... The other object of this work, which occupies Part I., is the explanation of the fundamental concepts which mathematics accepts as indefinable. This is a purely philosophical task...

Esta obra pasaría a ser el manual de lógica matemática de aquellos años y, después, una de las obras más influyentes de lógica.

1.3. Programa de Hilbert

Como se discute en [Zac], en el Congreso Internacional de Matemáticos de 1900, el matemático David Hilbert presenta el esquema de una demostración

de la consistencia de su axiomatización del análisis matemático. Después, tras la publicación del Principia Mathematica, Hilbert asume por válida dicha fundamentación y vuelve a trabajar en lógica matemática, esta vez en otros problemas como la *decidibilidad*, haciendo un desarrollo de la lógica de primer orden.

Este desarrollo llevó a Hilbert a rechazar la demostración de la consistencia de la aritmética encontrada en el Principia Mathematica por motivos filosóficos. Paralelamente, el logicista L.E.J. Brouwer desarrolló el *intuicionismo*. Este tipo de propuestas hizo a la comunidad matemática de aquellos años buscar alternativas a la propuesta del Principia Mathematica. Una de ellas la propone Hilbert, privilegiando la teoría elemental de números. Asume la existencia de unos objetos matemáticos llamados "signos" sin contenido semántico y que existen de forma inmediatamente empírica. Es decir, un dominio formado por secuencias de trazos que pueden ser manipulados mediante la lógica, pero que no necesitan de ella para su construcción.

Con esta base, Hilbert y sus colaboradores desarrollaron una fundamentanción de la matemática hasta el punto de asumir en 1929 que el programa estaba casi completo restando tan sólo unos casos particulares.

Dos años después, Gödel demostraría que eso no era así, negando la posibilidad del logicismo con su primer teorema y negando que se pueda probar la consistencia de la aritmética en el segundo.

1.4. Lógica de Primer Orden

Empezamos por definir nuestro entorno de trabajo. Demostraremos la incompletitud de la aritmética de Peano en la *lógica de primer orden*, un sistema formal en el que, en los enunciados, los cuantificadores alcanzan sólo a elementos individuales, pero no a otros enunciados o propiedades arbitrarias. La lógica de primer orden sería una formalización sintáctica de una teoría matemática (teoría de cuerpos, teoría de grupos, etc).

Dentro de ella, diferenciamos el nivel sintáctico y el semántico. En el nivel sintáctico definiremos las reglas de deducción y los axiomas que conforman el contenido formal de una teoría matemática. En él operaremos y tendrán lugar las demostraciones. Además, a esta sintaxis le añadimos una interpretación semántica, ya que nuestro lenguaje natural trabaja con el significado de los objetos matemáticos. El concepto de verdad es esencialmente semántico. Esta semántica la formalizamos con la *teoría de modelos* como veremos después más en detalle.

Podríamos preguntarnos por qué entonces no abordar este problema con lógicas de orden superior. La ventaja de la lógica de primer orden es que permite una relación estrecha entre la consistencia sintáctica y la consistencia semántica (existencia de modelo), bajo la forma de los teoremas de compacidad y completitud.

En esta memoria seguiremos el esquema del libro *Gödel's Theorems and Zermelo's Axioms* [HK20]. Primero definiremos la lógica en la que vamos a trabajar y su semántica en las secciones 2 y 3. En la sección 4, describiremos la teoría matemática sobre la que se enuncian los teoremas. En la sección 5, definiremos las herramientas con las que, en la sección 6, demostraremos los teoremas. Y en la última, veremos algunas de las consecuencias que tienen éstos en la lógica matemática. Sin embargo, la exposición que presentamos de la función β de Gödel es diferente, está más detallada y su hilo argumental está reordenado. Además, presentamos otras versiones de las fórmulas que expresan, en la aritmética de Peano, la sustitución de variables por términos, ya que, entendemos que, en la versión dada en esta referencia, no queda tan claro que las fórmulas sean Σ_1^0, ya que no especifica cotas concretas de algunos cuantificadores universales y deja esta demostración para secciones posteriores del documento.

2. Sintaxis de la lógica de primer orden

En esta sección construimos la sintaxis de la lógica de primer orden. Para ello definimos, primero, los elementos, y luego, las operaciones de su lenguaje.

2.1. Alfabeto

El alfabeto de nuestra lógica está compuesto por 7 tipos de símbolos:

1. **Variables**: Nuestro lenguaje consta de una cantidad infinita numerable de símbolos, que llamaremos variables y denotaremos por letras latinas: v, x, $y \ldots$

2. **Operadores lógicos**: Que son cuatro: $\neg \wedge \vee \rightarrow$. Semánticamente serán la negación, conjunción, disyunción e implicación, respectivamente.

3. **Cuantificadores lógicos**: Son $\exists$ y $\forall$. Semánticamente denotan existe y para todo.

4. **Símbolo de igualdad**: $=$, que es un tipo particular de relación.

5. **Símbolos de constantes**: Un conjunto de símbolos distinguidos: a, b, c,.... Que denotarán objetos particulares del conjunto en el que se trabaje.

6. **Símbolos de funciones**: F, G, H. Cada símbolo de función tiene asociado un número n llamado su **aridad**.

7. **Símbolos de relaciones**: R, S, T. Cada símbolo de función tiene asociado un número n llamado su **aridad**.

8. **Símbolos auxiliares**: Paréntesis y comas: $($ $)$ $,$

Los 4 primeros se llaman **símbolos lógicos** y los usaremos universalmente. Los demás, **símbolos no lógicos**. El conjunto de símbolos no lógicos tomados en cada contexto se llama **lenguaje** $\mathscr{L}$. Una **palabra** es una sucesión finita de elementos del alfabeto.

Además nos serán de utilidad los siguientes símbolos:

$$\varphi \leftrightarrow \psi :\Longleftrightarrow (\varphi \to \psi) \wedge (\psi \to \varphi)$$

$$\exists! \nu \varphi :\Longleftrightarrow \exists \nu(\varphi(\nu) \wedge \forall \mu(\varphi(\mu) \to \mu = \nu))$$

$$\tau \neq \tau' :\Longleftrightarrow \neg(\tau = \tau')$$

Donde $:\Longleftrightarrow$ es un símbolo del metalenguaje que define equidad entre símbolos. Este símbolo lo usaremos principalmente para definir en nuestro lenguaje nuevas constantes, funciones y relaciones.

Otro símbolo de nuestro metalenguaje será $\equiv$, que denotará que dos símbolos son idénticos.

2.2. Términos y fórmulas

Dados un lenguaje y una palabra de nuestro lenguaje, decimos que ésta es un $\mathscr{L}$-**término** si es el resultado de una consecución finita de los siguientes:

($T0$) Cada variable es un $\mathscr{L}$-término

($T1$) Cada constante en $\mathscr{L}$ es un $\mathscr{L}$-término.

($T2$) Dados n $\mathscr{L}$-términos τ_i y una función n-ádica F de $\mathscr{L}$, $F(\tau_1, \cdots, \tau_n)$ es un $\mathscr{L}$-término.

Los elementos de $(T0)$ y $(T1)$ se llaman **términos atómicos**. Los elementos de $(T2)$ y del lenguaje formal los denotaremos con letras griegas. Para los elementos que no dependen del lenguaje escribiremos simplemente términos.

Dados un lenguaje y una palabra de nuestro lenguaje, decimos que ésta es una $\mathscr{L}$-**fórmula** si es el resultado de una consecución finita de los siguientes:

$(F0)$ Dados n $\mathscr{L}$-términos τ_i y una relación n-ádica R de $\mathscr{L}$, $R(\tau_1, \cdots, \tau_n)$ es una $\mathscr{L}$-fórmula.

$(F1)$ Dadas φ y ψ $\mathscr{L}$-fórmulas, $\neg\varphi$, $(\varphi \vee \psi)$, $(\varphi \wedge \psi)$ y $(\varphi \to \psi)$ son $\mathscr{L}$-fórmulas.

$(F2)$ Dadas φ una $\mathscr{L}$-fórmula y ν una variable, $\exists\nu(\varphi)$ y $\forall\nu(\varphi)$ son $\mathscr{L}$-fórmulas.

Los elementos de $(F0)$ se llaman **fórmulas atómicas**. Los elementos que no dependen del lenguaje son llamados simplemente fórmulas.

Dada una fórmula φ del tipo $\exists\nu\psi$ o $\forall\nu\psi$, se dice que ν está en el **rango** del cuantificador si está en ψ. Si una variable de φ está en el rango de un cuantificador se dice **ligada**, si no, **libre**. El conjunto de variables libres de φ lo denotaremos por $libres(\varphi)$. Si $libres(\varphi) = \emptyset$, φ se dice **sentencia**.

Dada una fórmula φ con una variable ν y un término τ llamamos **sustitución** a cambiar todas las ocurrencias libres de ν por τ y lo denotamos por $\varphi(\nu/\tau)$ o simplemente $\varphi(\tau)$. Dicha sustitución es **admisible** si ninguna variable de τ se vuelve ligada al hacer la sustitución.

En caso de que ayude con la legibilidad, los paréntesis y las comas puede que los excluyamos a lo largo de la memoria, siempre que no incluyan ambigüedad. Así mismo, las relaciones o funciones a veces las escribiremos de manera infija. Por ejemplo, escribiremos $1 + 2$ en vez de $+(1, 2)$.

2.3. Axiomas

Introducimos ahora las fórmulas que conformaran los **axiomas lógicos** (de la lógica de primer orden), que daremos por válidos para cualquier teoría matemática.

Dadas φ, φ_1, φ_2, φ_3, ψ, fórmulas arbitrarias de nuestro lenguaje:

L_0: $\varphi \vee \neg\varphi$
L_1: $\varphi \to (\psi \to \varphi)$
L_2: $(\psi \to (\varphi_1 \to \varphi_2)) \to ((\psi \to \varphi_1) \to (\psi \to \varphi_2))$
L_3: $(\varphi \wedge \psi) \to \varphi$
L_4: $(\varphi \wedge \psi) \to \psi$

L_5: $\varphi \to (\psi \to (\psi \wedge \varphi))$
L_6: $\varphi \to (\varphi \vee \psi)$
L_7: $\psi \to (\varphi \vee \psi)$
L_8: $(\varphi_1 \to \varphi_3) \to ((\varphi_2 \to \varphi_3) \to ((\varphi_1 \vee \varphi_2) \to \varphi_3))$
L_9: $\neg\varphi \to (\varphi \to \psi)$

Dados una variable ν y un término τ donde $\varphi(\nu/\tau)$ es admisible:

L_{10}: $\forall\nu\varphi(\nu) \to \varphi(\nu/\tau)$
L_{11}: $\varphi(\nu/\tau) \to \exists\nu\varphi(\nu)$

y, además, si ν no es una variable en ψ:

L_{12}: $\forall\nu(\psi \to \varphi(\nu)) \to (\psi \to \forall\nu\varphi(\nu))$
L_{13}: $\forall\nu(\varphi(\nu) \to \psi) \to (\exists\nu\varphi(\nu) \to \psi)$

Dados n términos τ_i, un símbolo de relación R y un símbolo de función F:

L_{14}: $\tau = \tau$
L_{15}: $(\tau_1 = \tau_1' \wedge \ldots \wedge \tau_n = \tau_n') \to (R(\tau_1, \ldots, \tau_n) \to R(\tau_1', \ldots, \tau_n'))$
L_{16}: $(\tau_1 = \tau_1' \wedge \ldots \wedge \tau_n = \tau_n') \to (F(\tau_1, \ldots, \tau_n) = F(\tau_1', \ldots, \tau_n'))$

Podemos, además de estos, tomar otros axiomas, llamados **no lógicos** para formar una teoría. Una **teoría** o **$\mathscr{L}$-teoría** es un conjunto de axiomas no lógicos en los que, además, podemos tener símbolos de un conjunto $\mathscr{L}$ llamado **signatura**.

En los axiomas L_{12} y L_{13} hemos hecho un ligero cambio con respecto a la referencia principal [HK20], ya que así simplificamos unos resultados posteriores.

2.4. Demostraciones

En esta modelización de una teoría matemática, vamos a dar una definición puramente sintáctica de demostración. Construiremos fórmulas sintácticamente válidas partiendo de otras fórmulas usando una herramienta llamada **reglas de inferencia**:

Dadas dos fórmulas φ y ψ, se llama **Modus Ponens (MP)** a la siguiente deducción:

$$\frac{\varphi \to \psi, \varphi}{\psi}$$

Donde las fórmulas de encima de la línea son las premisas y la de debajo, la conclusión. Lo denotaremos por $MP(\varphi \to \psi, \varphi)$.

Dadas una fórmula φ y una variable ν, se llama **Generalización ($\forall$)** a la siguiente deducción:

$$\frac{\varphi}{\forall \nu \varphi}$$

Y lo denotaremos por $\forall(\varphi)$.

Con estas reglas de inferencia estamos en disposición de definir demostración formal.

Dado un lenguaje $\mathscr{L}$ y un conjunto $\mathbf{\Phi}$ de $\mathscr{L}$-fórmulas, una $\mathscr{L}$-fórmula se dice **demostrable** en $\mathbf{\Phi}$, que denotaremos por $\mathbf{\Phi} \vdash \varphi$, si existe una sucesión finita $\varphi_0, \ldots, \varphi_n$ de $\mathscr{L}$-fórmulas llamada **demostración** de φ, tal que $\varphi_n \equiv \varphi$ y para todo $i \leq n$ se da uno de los siguientes:

- φ_i es un axioma.
- $\varphi_i \in \mathbf{\Phi}$.
- (MP) existen $j, k < i$ con $\varphi_j \equiv \varphi_k \to \varphi_i$.
- ($\forall$) existe $j < i$ con $\varphi_i \equiv \forall \nu \varphi_j$ para alguna variable ν

Cuando $\mathbf{\Phi}$ es el conjunto vacío escribiremos $\vdash \varphi$. Si no existe demostración de φ en $\mathbf{\Phi}$, escribiremos $\mathbf{\Phi} \nvdash \varphi$. A $\mathbf{\Phi}$ lo llamaremos **conjunto de hipótesis**.

Si una fórmula es demostrable a partir de sólo los axiomas lógicos, ésta es una **tautología**.

Dos fórmulas φ, ψ se dicen **lógicamente equivalentes** si $\vdash \varphi \leftrightarrow \psi$ y se denota por $\varphi \Leftrightarrow \psi$.

Veamos un ejemplo de cómo hacer demostraciones.

Tautología 2.1. $\vdash \varphi \to \varphi$

Demostración.

$\varphi_0 : (\varphi \to ((\varphi \to \varphi) \to \varphi)) \to ((\varphi \to (\varphi \to \varphi)) \to (\varphi \to \varphi))$	L_2
$\varphi_1 : \varphi \to ((\varphi \to \varphi) \to \varphi)$	L_1
$\varphi_2 : (\varphi \to (\varphi \to \varphi)) \to (\varphi \to \varphi)$	$MP(\varphi_0, \varphi_1)$
$\varphi_3 : \varphi \to (\varphi \to \varphi)$	L_1
$\varphi_4 : \varphi \to \varphi$	$MP(\varphi_2, \varphi_3)$

□

Normalmente en matemáticas para demostrar un enunciado de la forma $\varphi \to \psi$, se asume φ y se concluye ψ. Hasta ahora hemos tenido que demostrar la implicación en sí misma. En lógica de primer orden este proceder se justifica con el llamado teorema de la deducción. Denotamos $\mathbf{\Phi} \cup \{\varphi\}$ por $\mathbf{\Phi} + \varphi$.

Teorema 2.1 (Teorema de la deducción)**.** *Sea $\mathscr{L}$ un lenguaje, $\mathbf{\Phi}$ un conjunto de fórmulas de $\mathscr{L}$ y φ, ψ $\mathscr{L}$-sentencias. Son equivalentes*

1. $\mathbf{\Phi} \vdash \psi \to \varphi$
2. $\mathbf{\Phi} + \psi \vdash \varphi$

Demostración. Supongamos *1.*

$\varphi_0 : \psi \to \varphi$	Hipótesis
$\varphi_1 : \psi$	Hipótesis
$\varphi_2 : \varphi$	$MP(\varphi_0, \varphi_1)$

Ahora supongamos *2.* Sea $\varphi_0, \ldots, \varphi_n$ una demostración de φ de longitud n a partir de $\mathbf{\Phi} + \psi$, demostramos $\mathbf{\Phi} \vdash \psi \to \varphi$ por inducción en n.

$n = 1$) Entonces φ es, o bien, un axioma, o bien, un elemento de $\mathbf{\Phi} + \psi$:

$\varphi_n : \varphi$	Hipótesis
$\varphi_{n+1} : \varphi \to (\psi \to \varphi)$	L_1
$\varphi_{n+2} : \psi \to \varphi$	$MP(\varphi_{n+1}, \varphi_n)$

- Si $\varphi \equiv \psi$, se tiene que $\psi \to \psi$ por la Tautología 2.1.

$n > 1$) H. I.: Para todo $j < n$ se tiene que $\mathbf{\Phi} \vdash \varphi_j$. Asumiendo que φ no es como en los casos anteriores, existen dos posibilidades:

- φ se obtiene por $MP(\varphi_k, \varphi_j)$ con $j, k < n$:

$\varphi_n : \psi \to \varphi_j$	H.I.
$\varphi_{n+1} : \psi \to (\varphi_j \to \varphi)$	H.I.
$\varphi_{n+2} : \varphi_{n+1} \to ((\psi \to \varphi_j) \to (\psi \to \varphi))$	L_2
$\varphi_{n+3} : (\psi \to \varphi_j) \to (\psi \to \varphi)$	$MP(\varphi_{n+2}, \varphi_{n+1})$
$\varphi_{n+4} : \psi \to \varphi$	$MP(\varphi_{n+3}, \varphi_n)$

- φ se obtiene por generalización, es decir, existe $j < n$ con $\varphi \equiv \forall \nu \varphi_j$ y ν una variable que no está en ψ, tenemos:

$\varphi_n : \psi \to \varphi_j$	$H.I.$
$\varphi_{n+1} : \forall \nu(\psi \to \varphi_j)$	$\forall(\varphi_n)$
$\varphi_{n+2} : \forall \nu(\psi \to \varphi_j) \to (\psi \to \varphi)$	L_{12}
$\varphi_{n+3} : \psi \to \varphi$	$MP(\varphi_{n+2}, \varphi_{n+1})$

□

Esto nos permite simplificar las demostraciones, por ejemplo:

Tautología 2.2. $\vdash (\varphi \to \psi_1) \to ((\varphi \to \psi_2) \to (\varphi \to (\psi_1 \wedge \psi_2)))$

Demostración. Aplicando tres veces el teorema de la deducción, basta demostrar $\psi_1 \wedge \psi_2$ asumiendo $\varphi \to \psi_1$, $\varphi \to \psi_2$ y φ.

$\varphi_0 : \varphi \to \psi_1$	Hipótesis
$\varphi_1 : \varphi \to \psi_2$	Hipótesis
$\varphi_2 : \varphi$	Hipótesis
$\varphi_3 : \psi_1$	$MP(\varphi_0, \varphi_2)$
$\varphi_4 : \psi_2$	$MP(\varphi_1, \varphi_2)$
$\varphi_5 : \psi_1 \to (\psi_2 \to (\psi_1 \wedge \psi_2))$	L_5
$\varphi_6 : \psi_2 \to (\psi_1 \wedge \psi_2)$	$MP(\varphi_5, \varphi_3)$
$\varphi_7 : \psi_1 \wedge \psi_2$	$MP(\varphi_6, \varphi_4)$

□

Tautología 2.3. $\vdash (\varphi \to \psi) \leftrightarrow (\neg\varphi \vee \psi)$

Tautología 2.4. $\vdash \neg(\varphi \wedge \psi) \leftrightarrow (\neg\varphi \vee \neg\psi)$

Tautología 2.5. $\vdash \neg(\varphi \vee \psi) \leftrightarrow (\neg\varphi \wedge \neg\psi)$

Estas tres últimas tautologías permiten expresar toda fórmula como una secuencia de conjunciones y disyunciones de fórmulas atómicas y negaciones de fórmulas atómicas.

Tautología 2.6. $\vdash ((\varphi_1 \to \varphi_2) \to ((\varphi_2 \to \varphi_3) \to (\varphi_1 \to \varphi_3)))$

2.5. Consistencia

Una **contradicción** es una fórmula del tipo $\varphi \wedge \neg\varphi$. El siguiente resultado ilustra sus propiedades.

Proposición 2.1 (*Ex falso quodlibet*)**.** *Dado un conjunto de fórmulas* $\mathbf{\Phi}$ *y* φ *una fórmula cualquiera, son equivalentes*

- $\mathbf{\Phi} \vdash \varphi \wedge \neg\varphi$
- *Para cada fórmula* ψ*:* $\mathbf{\Phi} \vdash \psi$

Demostración. Tomando $\varphi \wedge \neg\varphi$ como hipótesis, se tiene:

$\varphi_0 : \varphi \wedge \neg\varphi$	Hipótesis
$\varphi_1 : (\varphi \wedge \neg\varphi) \to \varphi$	L_3
$\varphi_2 : \varphi$	$MP(\varphi_1, \varphi_0)$
$\varphi_3 : (\varphi \wedge \neg\varphi) \to \neg\varphi$	L_4
$\varphi_4 : \neg\varphi$	$MP(\varphi_3, \varphi_0)$
$\varphi_5 : \neg\varphi \to (\varphi \to \psi)$	L_9
$\varphi_6 : \varphi \to \psi$	$MP(\varphi_5, \varphi_4)$
$\varphi_7 : \psi$	$MP(\varphi_6, \varphi_2)$

Para la otra implicación basta tomar $\psi \equiv \varphi \wedge \neg\varphi$, obteniendo $\mathbf{\Phi} \vdash \varphi \wedge \neg\varphi$. □

En general, en otro tipo de lógicas, diríamos que algo es consistente si no deriva contradicciones. Debido al resultado anterior, podemos dar la siguiente definición equivalente en nuestro caso. Un conjunto de fórmulas $\mathbf{\Phi}$ es **inconsistente** si puede demostrar una contradicción, $\mathbf{\Phi} \vdash \varphi \wedge \neg\varphi$. En caso contrario diremos que es **consistente**.

Una teoría T se dice **completa** si para toda fórmula $\varphi \in T$ se tiene $T \vdash \varphi$ o $T \vdash \neg\varphi$.

3. Semántica de la lógica de primer orden

La semántica es la parte del metalenguaje que se encarga de dotar significados de verdad a las fórmulas y a las deducciones.

En nuestra exposición, seremos bastante flexibles con nuestro metalenguaje y los resultados que pueden probarse en él. Más específicamente, asumiremos que el metalenguaje es aquel que hemos aprendido a lo largo del grado. En particular, permitiremos en nuestro metalenguaje el uso de conjuntos infinitos, funciones y relaciones definidas sobre conjuntos infinitos y podremos aplicar razonamientos de teoría de modelos. De una manera más formal, los teoremas que aparecen en nuestro metalenguaje se pueden formalizar en la teoría de conjuntos ZFC de Zermelo-Fraenkel con el axioma de elección. El hecho de contar con una metateoría tan potente y, en particular, el uso de un modelo de números naturales $\mathbb{N}$ nos permitirá simplificar la exposición. Nuestra metateoría *sabe* que los axiomas de Peano son consistentes.

Dado un lenguaje $\mathscr{L}$, una **$\mathscr{L}$-estructura M** consiste de un conjunto no vacío A llamado **dominio** y una asignación que hace corresponder:

- A cada constante $c \in \mathscr{L}$ un elemento $c^{\mathbf{M}} \in A$,
- A cada símbolo de relación n-ario $R \in \mathscr{L}$ un conjunto de n-tuplas $R^{\mathbf{M}} \subseteq A^n$,
- A cada símbolo de función n-ario $F \in \mathscr{L}$ una función n-aria $F^{\mathbf{M}}$ de A^n en A,
- Al símbolo de igualdad $=$, la igualdad en A, $\{(a, a) \mid a \in A\} \subseteq A^2$.

Una interpretación $\mathbf{I}$ de $\mathscr{L}$ es un par $(\mathbf{M}, j)$ con $\mathbf{M}$ una $\mathscr{L}$-estructura y j una aplicación del conjunto de variables en A.

Dada una interpretación $\mathbf{I}$, una variable ν y un elemento $a \in A$, definimos la interpretación $\mathbf{I}^{\frac{a}{\nu}}$ como la dada por la asignación $j\frac{a}{\nu}$ siguiente:

$$j\frac{a}{\nu}(\nu') = \begin{cases} a & \text{si } \nu' \equiv \nu \\ j(\nu') & \text{si no} \end{cases}$$

de donde definimos:

- $\mathbf{I}^{\frac{a}{\nu}} := (\mathbf{M}, j\frac{a}{\nu})$
- Dada una variable ν, $\mathbf{I}(\nu) := j(\nu)$
- Dada una constante $c \in \mathscr{L}$, $\mathbf{I}(c) := c^{\mathbf{M}}$
- Dada una función n-aria $F \in \mathscr{L}$ y términos $\tau_1, \ldots, \tau_n$,

$$\mathbf{I}(F(\tau_1, \ldots, \tau_n)) := F^{\mathbf{M}}(\mathbf{I}(\tau_1), \ldots, \mathbf{I}(\tau_n))$$

Con esto podemos definir la verdad de nuestras fórmulas bajo nuestro metalenguaje. Escribimos $\mathbf{I} \vDash \varphi$ cuando φ es **verdadera** en $\mathbf{I}$ definiéndolo recursivamente de la siguiente forma:
$\mathbf{I} \vDash \tau_1 = \tau_2$ si y sólo si $\mathbf{I}(\tau_1)$ y $\mathbf{I}(\tau_2)$ son el mismo objeto
$\mathbf{I} \vDash R(\tau_1, \ldots, \tau_n)$ si y sólo si $\langle \mathbf{I}(\tau_1), \ldots, \mathbf{I}(\tau_n) \rangle \in R^{\mathbf{M}}$
$\mathbf{I} \vDash \neg\psi$ si y sólo si NO $\mathbf{I} \vDash \psi$
$\mathbf{I} \vDash \psi_1 \wedge \psi_2$ si y sólo si $\mathbf{I} \vDash \psi_1$ Y $\mathbf{I} \vDash \psi_2$
$\mathbf{I} \vDash \psi_1 \vee \psi_2$ si y sólo si $\mathbf{I} \vDash \psi_1$ O $\mathbf{I} \vDash \psi_2$
$\mathbf{I} \vDash \psi_1 \to \psi_2$ si y sólo si $\mathbf{I} \vDash \psi_1$ ENTONCES $\mathbf{I} \vDash \psi_2$
$\mathbf{I} \vDash \exists\nu\psi$ si y sólo si EXISTE a EN A CON $\mathbf{I}^{\frac{a}{\nu}} \vDash \psi$
$\mathbf{I} \vDash \forall\nu\psi$ si y sólo si PARA TODO a EN A $\mathbf{I}^{\frac{a}{\nu}} \vDash \psi$
Observamos que las interpretaciones tienen estas propiedades

- Sea φ una fórmula con $\nu \notin \mathit{libres}(\varphi)$, entonces:

$$\mathbf{I}\frac{a}{\nu} \vDash \varphi \text{ si y sólo si } \mathbf{I} \vDash \varphi$$

- Sea $\varphi(\nu)$ una fórmula con $\varphi(\nu/\tau)$ admisible, entonces:

$$\mathbf{I}\frac{\mathbf{I}(\tau)}{\nu} \vDash \varphi(\nu) \text{ si y sólo si } \mathbf{I} \vDash \varphi(\tau)$$

- Si σ es una sentencia, para todo par de interpretaciones $\mathbf{I}$, $\mathbf{J}$ sobre $\mathbf{M}$

$$\mathbf{I} \vDash \sigma \text{ si y sólo si } \mathbf{J} \vDash \sigma$$

Dado un conjunto $\mathbf{\Phi}$ de $\mathscr{L}$-fórmulas, una $\mathscr{L}$-estructura $\mathbf{M}$ es un **modelo** de $\mathbf{\Phi}$ si para toda interpretación $(\mathbf{M}, j)$ y toda fórmula $\varphi \in \mathbf{\Phi}$ se tiene $(\mathbf{M}, j) \vDash \varphi$, es decir, si toda fórmula de $\mathbf{\Phi}$ es verdadera en $\mathbf{M}$ independientemente de la asignación j. Lo denotamos por $\mathbf{M} \vDash \mathbf{\Phi}$.

Obsérvese que, por cómo hemos definido las interpretaciones, que si σ es una sentencia entonces o $\mathbf{M} \vDash \sigma$ o $\mathbf{M} \vDash \neg\sigma$. Si φ es una fórmula de variables libres $x_1, \ldots, x_n$, entonces

$$M \vDash \varphi \text{ si y solo si } M \vDash \forall x_1 \ldots \forall x_n \varphi$$

Una de las razones por la que nos decantamos por la lógica de primer orden es que los modelos de esta lógica son coherentes y completos.

El teorema de coherencia viene a decir que si tenemos un modelo en el que nuestras hipótesis son verdaderas, las conclusiones también son verdaderas.

Teorema 3.1 (Teorema de la Coherencia)**.** *Sea $\mathbf{\Phi}$ un conjunto de $\mathscr{L}$-fórmulas y $\boldsymbol{M}$ un modelo de $\mathbf{\Phi}$, entonces dada una $\mathscr{L}$-fórmula φ se tiene:*

$$\textit{Si } \mathbf{\Phi} \vdash \varphi, \textit{ entonces } \boldsymbol{M} \vDash \varphi$$

Demostración. Para comprobar el resultado tenemos que ver que ambos los axiomas $L_0 - L_{16}$ y las reglas de inferencia son válidas en $\mathbf{M}$. Las fórmulas de $\mathbf{\Phi}$ ya sabemos que son válidas porque $\mathbf{M}$ lo modela. En primer lugar tomamos las tablas de verdad de los operadores de nuestro metalenguaje según hemos definido antes la verdad de cada fórmula, tomando ahora 1 si $\mathbf{M} \vDash \Phi$ y 0 si $\mathbf{M} \nvDash \Phi$,

Φ	Ψ	NO Φ	Φ Y Ψ	Φ O Ψ	Φ ENTONCES Ψ
0	0	1	0	0	1
0	1	1	0	1	1
1	0	0	0	1	0
1	1	0	1	1	1

De esta forma se pueden construir las tablas de verdad correspondientes a los axiomas. Comprobando que el enunciado correspondiente a cada axioma es siempre verdadero independientemente de los valores de sus variables habríamos visto que son ciertos en $\mathbf{M}$. Lo comprobamos para L_5 y los demás hasta L_9 son análogos,

Φ	Ψ	Φ Y Ψ	Ψ ENT. (Φ Y Ψ)	Φ ENT. (Ψ ENT. (Φ Y Ψ))
0	0	0	1	1
0	1	0	0	1
1	0	0	1	1
1	1	1	1	1

Para los demás utilizaremos L_{11} como ejemplo: Sea A un dominio de $\mathbf{M}$, j arbitraria e $\mathbf{I} = (\mathbf{M}, j)$, tenemos:

$$\mathbf{I} \vDash \varphi(\tau) \to \exists\nu\varphi(\nu) \text{ si y sólo si } \mathbf{I} \vDash \varphi(\tau) \text{ ENTONCES } \mathbf{I} \vDash \exists\nu\varphi(\nu)$$

$$\mathbf{I} \vDash \exists\nu\varphi(\nu) \text{ si y sólo si EXISTE } a \text{ EN } A \text{ CON } \mathbf{I}\frac{a}{\nu} \vDash \varphi(\nu)$$

Entonces, tomamos φ una $\mathscr{L}$-fórmula con $\varphi(\nu/\tau)$ admisible. Hemos visto que:

$$\mathbf{I} \vDash \varphi(\tau) \text{ si y sólo si } \mathbf{I}\frac{\mathbf{I}(\tau)}{\nu} \vDash \varphi(\nu)$$

Por lo tanto, tomando $a \equiv \mathbf{I}(\tau)$,

$$\mathbf{I} \vDash \varphi(\tau) \text{ ENTONCES EXISTE } a \text{ EN } A \text{ CON } \mathbf{I}\frac{a}{\nu} \vDash \varphi(\nu)$$

o, lo que es lo mismo, ya que j es arbitraria:

$$\mathbf{M} \vDash \varphi(\tau) \to \exists\nu\varphi(\nu),$$

que es lo que queríamos comprobar.

Ahora para ver que (MP) es válido, calculamos la tabla de verdad de

$$(\mathbf{M} \vDash \varphi \to \psi \text{ Y } \mathbf{M} \vDash \varphi) \text{ ENTONCES } \mathbf{M} \vDash \psi$$

Φ	Ψ	Φ ENT. Ψ	(Φ ENT. Ψ) Y Φ	((Φ ENT. Ψ) Y Φ) ENT. Ψ
0	0	1	0	1
0	1	1	0	1
1	0	0	0	1
1	1	1	1	1

Luego, ($\forall$) es una conclusión inmediata de la definición de modelo

$$\mathbf{M} \vDash \varphi \text{ si y sólo si } \mathbf{M} \vDash \forall\nu\varphi(\nu),$$

lo que completa la demostración. □

El teorema de completitud nos dice que si φ es consecuencia semántica de $\mathbf{\Phi}$, entonces es demostrable a partir de $\mathbf{\Phi}$. Donde por consecuencia semántica queremos decir que φ sea cierta en todo modelo de $\mathbf{\Phi}$. El teorema de completitud también se suele expresar como que todo conjunto de fórmulas consistentes tiene un modelo.

Teorema 3.2 (Teorema de Completitud)**.** *Sea $\mathbf{\Phi}$ un conjunto de $\mathscr{L}$-fórmulas. Si $\mathbf{\Phi}$ es consistente, entonces tiene un modelo. Además, sea σ una sentencia, si para cada modelo $\boldsymbol{M}$ de $\mathbf{\Phi}$, $\boldsymbol{M} \vDash \sigma$, entonces, $\mathbf{\Phi} \vdash \sigma$*

Esto lo podemos interpretar como que, dada una sentencia σ, o bien es demostrable o bien existe un contraejemplo. Más precisamente

Teorema 3.3. *Sea $\mathbf{\Phi}$ un conjunto consistente de $\mathscr{L}$-fórmulas, sea σ una sentencia, entonces se cumple una y sólo una de las siguientes*

- $\mathbf{\Phi} \vdash \sigma$,
- *Existe $\boldsymbol{M}$ con $\boldsymbol{M} \vDash \mathbf{\Phi}$ y $\boldsymbol{M} \vDash \neg\sigma$.*

4. La Axiomática de Peano

Ahora que hemos construido la estructura lógica, vamos a trabajar en la aritmética de Peano. Esta teoría matemática, que nos permite construir los números naturales, está formada por una lista de axiomas que fueron descritos por primera vez por el matemático Giussepe Peano en el siglo XIX. La llamaremos axiomática de Peano **AP**.

La signatura $\mathscr{L}_{\mathbf{AP}}$ de esta teoría constará de los siguientes símbolos

0) Símbolo de constante que llamaremos cero.
s) Símbolo de una función de aridad 1, que llamaremos sucesor. Si τ es un término, en ocasiones escribiremos $s\tau$ en vez de $s(\tau)$.
$+$) Símbolo de una función de aridad 2, que llamaremos suma. Escribiremos $\tau + \sigma$ en vez de $+(\tau, \sigma)$.
$\cdot$) Símbolo de función de aridad 2, que llamaremos multiplicación. Escribiremos $\tau \cdot \sigma$ o $\tau\sigma$ en vez de $\cdot(\tau, \sigma)$.

La teoría **AP** estará constituida por los siguientes axiomas

$\mathbf{AP}_0$: $\neg\exists x(sx = 0)$
$\mathbf{AP}_1$: $\forall x\forall y(sx = sy \to x = y)$
$\mathbf{AP}_2$: $\forall x(x + 0 = x)$
$\mathbf{AP}_3$: $\forall x\forall y(x + sy = s(x + y))$
$\mathbf{AP}_4$: $\forall x(x \cdot 0 = 0)$
$\mathbf{AP}_5$: $\forall x\forall y(x \cdot sy = (x \cdot y) + x)$
$\mathbf{AP}^*_6$: $(\varphi(0) \wedge \forall x(\varphi(x) \to \varphi(sx))) \to \forall x\varphi(x)$

donde $\mathbf{AP}^*_6$ no es un axioma sino el esquema de axiomas de inducción. Para cada φ una $\mathscr{L}_{\mathbf{AP}}$-fórmula con $x \in libres(\varphi)$, tenemos un axioma distinto.

En el caso de que dos fórmulas sean equivalentes bajo estos axiomas, es decir $\mathbf{AP} \vdash \varphi \leftrightarrow \psi$, escribiremos $\varphi \Leftrightarrow_{\mathbf{AP}} \psi$.

Como es usual en matemáticas, definiremos nuevas constantes, funciones y relaciones a partir de conceptos más básicos. Esto lo haremos de manera sucesiva y cada vez con mayor complejidad. Por primer ejemplo definimos la relación de orden como

$$x \leq y :\Longleftrightarrow \exists r(x + r = y),$$

$$x < y :\Longleftrightarrow x \leq y \wedge x \neq y,$$

$$x \geq y :\Longleftrightarrow y \leq x,$$

$$x > y :\Longleftrightarrow y < x.$$

En **AP** se pueden demostrar las propiedades usuales de este orden.

Teorema 4.1.

a) $\boldsymbol{AP} \vdash \forall x(x \leq x)$

b) $\boldsymbol{AP} \vdash \forall x\forall y((x \leq y \wedge y \leq x) \to x = y)$

c) $\boldsymbol{AP} \vdash \forall x \forall y \forall z((x \leq y \wedge y \leq z) \rightarrow (x \leq z))$

d) $\boldsymbol{AP} \vdash \forall x(0 \leq x)$

e) $\boldsymbol{AP} \vdash \forall x \forall y(x \leq y \vee y \leq x)$

f) $\boldsymbol{AP} \vdash \forall x \forall y(x < y \vee x = y \vee y < x)$

g) $\boldsymbol{AP} \vdash \forall x \forall y \forall z(x \leq y \rightarrow x + z \leq y + z)$

4.1. Nuestro modelo de los números naturales $\mathbb{N}$

Como hemos avanzado al comienzo de la Sección 3, el modelo de **AP** que tomamos serán los números naturales $\mathbb{N}$, los mismos que hemos visto en la carrera y con las propiedades que hemos estudiado, que no repetiremos en esta memoria, incluida la inducción. La diferencia más significativa entre **AP** y cómo hemos estudiado los números naturales en el grado es que en el último caso la inducción es presentada como un axioma de segundo orden:

$$\forall X \subseteq \mathbb{N}((0 \in X \wedge (\forall n(n \in X \rightarrow n + 1 \in X))) \rightarrow X = \mathbb{N})$$

Cada número natural "de verdad" n, lo podemos escribir de manera explícita como $s(\ldots s(0))$. Por tanto, a cada número natural n concreto, le podemos asignar el término $\underline{n} :\Longleftrightarrow s(\ldots s(0))$. Por ejemplo, $\underline{3}$ es una abreviatura de $s(s(s(0)))$. Para diferenciar los símbolos de $\mathbb{N}$ de los de **AP** escribiremos $\mathbb{N}$ como exponente en ellos, por ejemplo, $+^{\mathbb{N}}$ en vez de $+$ si sumamos en el modelo. De esta forma, se cumple lo siguiente

Proposición 4.1. *Dados $n, m \in \mathbb{N}$, se tiene:*

$\boldsymbol{N}_0$*:* $\boldsymbol{AP} \vdash s\underline{n} = \underline{n+1}$
$\boldsymbol{N}_1$*:* $\boldsymbol{AP} \vdash \underline{m} + \underline{n} = \underline{m +^{\mathbb{N}} n}$
$\boldsymbol{N}_2$*:* $\boldsymbol{AP} \vdash \underline{m} \cdot \underline{n} = \underline{m \cdot^{\mathbb{N}} n}$
$\boldsymbol{N}_3$*: Si m y n son el mismo número, entonces $\boldsymbol{AP} \vdash \underline{m} = \underline{n}$, y si m y n son números distintos, entonces $\boldsymbol{AP} \vdash \underline{m} \neq \underline{n}$.*
$\boldsymbol{N}_4$*: Si $m < n$, entonces $\boldsymbol{AP} \vdash \underline{m} < \underline{n}$, y si $m \not< n$, entonces $\boldsymbol{AP} \vdash \underline{m} \not< \underline{n}$.*
$\boldsymbol{N}_5$*:* $\boldsymbol{AP} \vdash \forall x(x < \underline{n} \leftrightarrow \bigvee_{k=0}^{n-1} x = \underline{k})$

Donde el símbolo $\bigvee$ denota

$$\bigvee_{k=0}^{n-1} x = \underline{k} :\equiv x = \underline{0} \vee x = \underline{1} \vee \ldots \vee x = \underline{n-1},$$

y la fórmula $m < n$ está definida en **AP** en la subsección 4.2.

Vamos ahora a definir las fórmulas más sencillas de la llamada *jerarquía aritmética*, que nos será útil para relacionar los resultados semánticos en $\mathbb{N}$ con resultados demostrables en **AP**.

Una $\mathscr{L}_{\mathbf{AP}}$-fórmula ψ es una **$\exists$-fórmula estricta** si está constituida por fórmulas atómicas o negaciones de formulas atómicas a partir de los símbolos $\wedge$, $\vee$, $\exists\nu$ y cuantificadores universales acotados ($\forall\nu < \tau$). Una $\mathscr{L}_{\mathbf{AP}}$-fórmula φ es una **$\exists$-fórmula** si existe una $\exists$-fórmula estricta ψ tal que $\varphi \Leftrightarrow_{\mathbf{AP}} \psi$.

Las $\exists$-fórmulas se suelen llamar en la literatura, las Σ_1^0-fórmulas.

Proposición 4.2. *Dada una $\exists$-fórmula $\varphi(x_1, \ldots, x_n)$ cuyas variables libres se encuentran entre $x_1, \ldots, x_n$ y dados $a_1, \ldots, a_n \in \mathbb{N}$, si $\mathbb{N} \vDash \varphi(a_1, \ldots, a_n)$, entonces $\boldsymbol{AP} \vdash \varphi(\underline{a_1}, \ldots, \underline{a_n})$.*

Demostración. Por inducción en la construcción de φ.

Si φ es una fórmula atómica, entonces es de la forma $\tau_0(x_1, \ldots, x_n) = \tau_1(x_1, \ldots, x_n)$ para términos τ_0, τ_1 cuyas variables estén en $x_1, \ldots, x_n$. Demostramos por inducción en la construcción de los términos que para todo término $\tau(x_1, \ldots, x_n)$ y para todo $a_1, .., a_n \in \mathbb{N}$ se tiene que

$$\mathbf{AP} \vdash \underline{\tau^{\mathbb{N}}(a_1, \ldots, a_n)} = \tau(\underline{a_1}, \ldots, \underline{a_n}).$$

Si τ es una variable o 0, el resultado es trivial. Si τ es de la forma $\tau_0 + \tau_1$ con $\mathbf{AP} \vdash \underline{a} = \tau_0(\underline{a_1}, \ldots, \underline{a_n})$ y $\mathbf{AP} \vdash \underline{b} = \tau_0(\underline{b_1}, \ldots, \underline{b_n})$, entonces $\tau^{\mathbb{N}}(a_1, \ldots, a_n)$ es $a +^{\mathbb{N}} b$. Usando $\mathbf{N}_1$ obtenemos

$$\mathbf{AP} \vdash \underline{\tau^{\mathbb{N}}(a_1, \ldots, a_n)} = \underline{a +^{\mathbb{N}} b} = \underline{a} + \underline{b} = \tau_0(\underline{a_1}, \ldots, \underline{a_n}) + \tau_1(\underline{a_1}, \ldots, \underline{a_n}).$$

Similarmente se demuestran los casos $s\tau_0$, $\tau_0 \cdot \tau_1$ y $\tau_0 < \tau_1$.

Ahora, nótese que por las Tautologías 2.3, 2.4 y 2.5, queda demostrar el resultado para disyunciones, conjunciones, negaciones de fórmulas atómicas y cuantificación.

Si $\varphi(x_1, \ldots, x_n) \equiv \varphi_0(x_1, \ldots, x_n) \vee \varphi_1(x_1, \ldots, x_n)$ y $\mathbb{N} \vDash \varphi(a_1, \ldots, a_n)$, tenemos que $\mathbb{N} \vDash \varphi_0(a_1, \ldots, a_n)$ o $\mathbb{N} \vDash \varphi_1(a_1, \ldots, a_n)$. Si $\mathbb{N} \vDash \varphi_0(a_1, \ldots, a_n)$, por inducción sabemos que $\mathbf{AP} \vdash \varphi_0(\underline{a_1}, \ldots, \underline{a_n})$. Por L_6 tenemos $\mathbf{AP} \vdash \varphi_0(\underline{a_1}, \ldots, \underline{a_n}) \vee \varphi_1(\underline{a_1}, \ldots, \underline{a_n})$. Si suponemos lo contrario, obtenemos el resultado por L_7. Análogamente se demuestra la conjunción.

Si $\varphi(x_1, \ldots, x_n) \equiv \neg\varphi_0(x_1, \ldots, x_n)$ con $\mathbb{N} \vDash \varphi(a_1, \ldots, a_n)$ y $\varphi_0(x_1, \ldots, x_n)$ una fórmula atómica. Tenemos que φ es de la forma

$$\tau_0(x_1, \ldots, x_n) \neq \tau_1(x_1, \ldots, x_n)$$

para τ_0 y τ_1 términos, por la tricotomía de el orden, esto es equivalente a $(\tau_0(x_1,\dots,x_n) < \tau_1(x_1,\dots,x_n)) \vee (\tau_1(x_1,\dots,x_n) < \tau_0(x_1,\dots,x_n))$. Como el resultado ya está demostrado para las dos partes de la disyunción, la demostración es análoga a la anterior a partir de aquí.

Si $\varphi(x_1,\dots,x_n) \equiv \forall y < \tau(x_1,\dots,x_n)\psi(x_1,\dots,x_n,y)$ y $\mathbb{N} \vDash \varphi(a_1,\dots,a_n)$, sea $a \equiv \tau^{\mathbb{N}}(a_1,\dots,a_n)$, entonces para todo $b < a$, $\mathbb{N} \vDash \psi(a_1,\dots,a_n,b)$, y por hipótesis de inducción $\mathbf{AP} \vdash \psi(\underline{a_1},\dots,\underline{a_n},\underline{b})$. Como $\mathbf{AP} \vdash \underline{a} = \tau(\underline{a_1},\dots,\underline{a_n})$, por $\mathbf{N}_5$, tenemos que

$$\mathbf{AP} \vdash \varphi(\underline{a_1},\dots,\underline{a_n}) \leftrightarrow \forall y(\bigvee_{b=0}^{a-1} y = \underline{b} \rightarrow \psi(\underline{a_1},\dots,\underline{a_n},y)).$$

Si $\varphi(x_1,\dots,x_n) \equiv \exists y\psi(x_1,\dots,x_n,y)$ y $\mathbb{N} \vDash \varphi(a_1,\dots,a_n)$, existe $b \in \mathbb{N}$ con $\mathbb{N} \vDash \psi(a_1,\dots,a_n,b)$. Por hipótesis de inducción tenemos $\mathbb{N} \vDash \psi(\underline{a_1},\dots,\underline{a_n},\underline{b})$, que era lo que queríamos demostrar. □

Obsérvese que para la negación de una $\forall$-fórmula también se cumple el resultado anterior ya que equivale a una $\exists$-fórmula. A las relaciones de la forma

$$R(x_1,\dots,x_n) :\Longleftrightarrow \varphi(x_1,\dots,x_n)$$

con φ y $\neg\varphi$ verificando el resultado de la proposición anterior las llamaremos **$\mathbb{N}$-conformes**. Una función se dice $\mathbb{N}$-conforme si está definida por una relación $\mathbb{N}$-conforme.

4.2. Resultados necesarios para los teoremas

Veamos ahora algunos resultados en **AP** que necesitaremos para demostrar los teoremas de incompletitud. Como es de esperar, se cumplen propiedades de la suma y de la multiplicación, como la conmutatividad o la unicidad de elemento neutro. La demostración de éstos y otros más avanzados es técnica y de poco interés en esta memoria, por lo que hemos decidido omitirlas.

En primer lugar veamos las propiedades asociativa de la suma y distributiva.

Lema 4.1.

a) $\boldsymbol{AP} \vdash \forall x\forall y\forall z((x+y)+z = x+(y+z))$

b) $\boldsymbol{AP} \vdash \forall x\forall y\forall z(x\cdot(y+z) = (x\cdot y)+(x\cdot z))$

También nos será de utilidad la propiedad cancelativa.

Lema 4.2. $\boldsymbol{AP} \vdash \forall x \forall y \forall z (x + y = x + z \to y = z)$

Damos ahora la siguiente caracterización para las fórmulas de desigualdad.

Lema 4.3.

a) $\boldsymbol{AP} \vdash \forall x \forall y (x < sy \leftrightarrow x \leq y)$

b) $\boldsymbol{AP} \vdash \forall x \forall y (x < y \leftrightarrow sx \leq y)$

Ahora definimos la resta y la divisibilidad. Obsérvese que definimos la resta sólo para números no negativos.

$$x - y = z :\Longleftrightarrow (y \leq x \wedge y + z = x) \vee (x \leq y \wedge z = 0)$$

$$x \mid y :\Longleftrightarrow \exists r (rx = y)$$

Nos serán de utilidad los siguientes resultados de divisibilidad.

Lema 4.4.

a) $\boldsymbol{AP} \vdash \forall x \forall y \forall z \leq y (x \mid y \wedge x \mid z \to x \mid y \pm z)$

b) $\boldsymbol{AP} \vdash \forall x \forall y \forall z (x \mid y \to x \mid yz)$

En la idea de la demostración de los teoremas está implícito el concepto de coprimalidad. Así que definimos ser **coprimos** y ser **primo** en **AP**.

$$coprimos(x, y) :\Longleftrightarrow x \neq 0 \wedge y \neq 0 \wedge \forall z (x \mid yz \to x \mid z)$$

$$primo(p) :\Longleftrightarrow p \neq \underline{1} \wedge \forall y \forall z (p \mid yz \to (p \mid y \vee p \mid z))$$

Los dos siguientes resultados serán importantes en la demostración. En el segundo obtenemos una forma de construir números coprimos entre sí que utilizaremos más adelante, así que incluimos su demostración para afianzar el concepto.

Lema 4.5. $\boldsymbol{AP} \vdash \forall x \forall y \forall k (k \mid x \wedge coprimos(x, y) \to coprimos(k, y))$

Lema 4.6. $\boldsymbol{AP} \vdash \forall k \forall y \neq 0 \forall j (k \mid y \to coprimos(\underline{1} + (j + k)y, \underline{1} + jy))$

Demostración. Empezamos demostrando $\mathbf{AP} \vdash \forall y \neq 0 \forall j(coprimos(y, \underline{1} + jy))$. Dado un número z, asumamos $y \mid (\underline{1} + jy)z$. Por el Lema 4.4 b) tenemos $y \mid jyz$ y como consecuencia del Lema 4.4 a) obtenemos $y \mid z$, que prueba lo que queríamos.

Pasando ahora al lema, tomamos k, y y j como en las hipótesis y un z arbitrario. Asumimos:

$$(\underline{1} + jy) \mid (\underline{1} + (j + k)y)z$$

o, lo que es lo mismo:

$$(\underline{1} + jy) \mid (\underline{1} + jy)z + kyz$$

Por el Lema 4.4 b) se tiene $(\underline{1} + jy) \mid (\underline{1} + jy)z$ y por el Lema 4.4 a), $(\underline{1} + jy) \mid kyz$. Como hemos demostrado antes que $coprimos(\underline{1} + jy, y)$, se tiene que $(\underline{1} + jy) \mid kz$. Y, como hemos asumido $k \mid y$, por el Lema 4.5 y $coprimos(\underline{1} + jy, y)$, $coprimos(\underline{1} + jy, k)$. Consecuentemente $(\underline{1} + jy) \mid z$, que era lo que queríamos demostrar. □

También nos será de utilidad la siguiente caracterización de coprimalidad.

Proposición 4.3.

$$\boldsymbol{AP} \vdash \forall x \forall y(coprimos(x, y) \leftrightarrow (x \neq 0 \wedge y \neq 0 \wedge \forall z((z \mid x \wedge z \mid y) \rightarrow z = \underline{1})))$$

Con el principio del menor número somos capaces de encontrar en **AP** el menor número de todos los que satisfacen una fórmula φ.

Proposición 4.4. *(Principio del Menor Número) Dada $\varphi(x)$ una $\mathcal{L}_{\boldsymbol{AP}}$-fórmula, se tiene:*

$$\boldsymbol{AP} \vdash \exists x \varphi(x) \rightarrow \exists x(\varphi(x) \wedge \forall y < x \neg \varphi(y))$$

El siguiente resultado nos servirá para construir un común múltiplo de series de números:

Lema 4.7. $\boldsymbol{AP} \vdash \forall m \exists y \forall k((k \neq 0 \wedge k \leq m) \rightarrow k \mid y)$

Demostración. Por inducción en m. El caso $m = 0$ es el caso vacío. Dado y tal que para todo $0 < k \leq m$, asumimos $k \mid y$. Sea $y' = y \cdot \mathbf{s}m$, por el Lema 4.4 $k \mid y'$. □

Una $\mathscr{L}_{\mathbf{AP}}$-fórmula φ con, al menos, las variables libres $x_1, \ldots, x_n, y$ se dice **funcional** si puede demostrarse

$$\mathbf{AP} \vdash \forall x_1 \ldots \forall x_n \exists! y \varphi(x_1, \ldots, x_n, y)$$

Veamos ahora cómo podemos definir funciones en **AP** a partir de fórmulas. Una función de la forma

$$F(x_1, \ldots, x_n) = y :\Longleftrightarrow \varphi(x_1, \ldots, x_n, y)$$

se dice **definible en AP** si φ es funcional. Obsérvese que éstas últimas son $\mathbb{N}$-conformes.

El siguiente resultado permite construir el máximo en un dominio finito de una función definible en **AP**.

Lema 4.8. *Dada una función F definible en **AP**,*

$$\boldsymbol{AP} \vdash \forall k > 0 \exists! x(\exists i < k(F(i) = x) \wedge \forall j < k(F(j) \leq x))$$

5. Gödelización de la Aritmética de Peano

Hasta ahora hemos visto la sintaxis de la aritmética de Peano por un lado y por otro la hemos dotado de una semántica. El reto al que nos enfrentamos ahora es poder hablar de la propia teoría de la aritmética de Peano, no en el metalenguaje sino en el lenguaje formal. A este proceso se le suele llamar gödelización. Por ejemplo, la demostración de un teorema la podemos codificar mediante el código Latex que usamos para escribirlo. Cada uno de los caracteres lo podemos identificar por un número mediante una codificación y por tanto, en principio, fijado un código podríamos describir cualquier cadena de palabras de **AP** mediante un número.

Sin embargo, aparentemente, la aritmética de Peano no parece que pueda hablar de sucesiones finitas de números naturales arbitrarios. Por ejemplo, podemos escribir una fórmula que interpretemos como "x es el cubo de un número primo".

$$\exists p(primo(p) \wedge p \cdot p \cdot p = x).$$

De manera análoga, para cada exponente k, podemos escribir una fórmula que diga $x = p^k$ para un primo p, pero no podríamos escribir $\exists p \exists k(x = p^k)$. Otro tipo de fórmula que tampoco podemos escribir directamente es "todo

número mayor que uno se expresa como producto de números primos", ya que el número de primos en la factorización es finito pero indeterminado.

Estos tres problemas podríamos resolverlos si fuéramos capaces de expresar en **AP** "una sucesión finita de números naturales". Este problema lo resuelve Gödel con la función β. En esta memoria presentamos la variante de ella que aparece en [HK20].

Una vez hayamos construido la función β, veamos cómo definir un código para identificar cada fórmula de nuestra teoría con un número natural. Obsérvese que la función

$$\begin{array}{rccc} f: & \mathbb{N}^3 & \longrightarrow & \mathbb{N} \\ & (p,q,s) & \longmapsto & 2^p 3^q 5^s \end{array}$$

es inyectiva. Por lo tanto, aplicando recursivamente esta función, podemos definir la codificación de palabras usando los siguientes principios:

ζ	$\#\zeta$	τ	$\#\tau$	φ	$\#\varphi$
0	$\mathbf{2^0}$	v_n	$\mathbf{2}\cdot n+\mathbf{3}$	$\tau_0=\tau_1$	$\mathbf{2}^{\#=}\cdot\mathbf{3}^{\#\tau_0}\cdot\mathbf{5}^{\#\tau_1}$
s	$\mathbf{2^1}$	st	$\mathbf{2}^{\#s}\cdot\mathbf{3}^{\#t}$	$\neg\psi$	$\mathbf{2}^{\#\neg}\cdot\mathbf{3}^{\#\psi}$
$+$	$\mathbf{2^2}$	t_0+t_1	$\mathbf{2}^{\#+}\cdot\mathbf{3}^{\#t_0}\cdot\mathbf{5}^{\#t_1}$	$\psi_0\wedge\psi_1$	$\mathbf{2}^{\#\wedge}\cdot\mathbf{3}^{\#\psi_0}\cdot\mathbf{5}^{\#\psi_1}$
$\cdot$	$\mathbf{2^3}$	$t_0\cdot t_1$	$\mathbf{2}^{\#\cdot}\cdot\mathbf{3}^{\#t_0}\cdot\mathbf{5}^{\#t_1}$	$\psi_0\vee\psi_1$	$\mathbf{2}^{\#\vee}\cdot\mathbf{3}^{\#\psi_0}\cdot\mathbf{5}^{\#\psi_1}$
$=$	$\mathbf{2^4}$			$\psi_0\to\psi_1$	$\mathbf{2}^{\#\to}\cdot\mathbf{3}^{\#\psi_0}\cdot\mathbf{5}^{\#\psi_1}$
$\neg$	$\mathbf{2^5}$			$\exists x\psi$	$\mathbf{2}^{\#\exists}\cdot\mathbf{3}^{\#x}\cdot\mathbf{5}^{\#\psi}$
$\wedge$	$\mathbf{2^6}$			$\forall x\psi$	$\mathbf{2}^{\#\forall}\cdot\mathbf{3}^{\#x}\cdot\mathbf{5}^{\#\psi}$
$\vee$	$\mathbf{2^7}$				
$\to$	$\mathbf{2^8}$				
$\exists$	$\mathbf{2^9}$				
$\forall$	$\mathbf{2^{10}}$				

donde dado un símbolo, término o fórmula ζ, $\#\zeta$ es su **número de Gödel**. Esto genera una inyección entre las palabras de nuestro lenguaje y $\mathbb{N}$. Por ejemplo, dada una fórmula

$$\varphi \equiv v_0 = v_1 \cdot v_2,$$

su número de Gödel es

$$\#\varphi \equiv 2^{16}\cdot 3^3\cdot 5^{2^8\cdot 3^5\cdot 5^7}.$$

Usaremos la función β para poder describir conceptos como c es el número de Gödel de una fórmula o "existe una sucesión finita de números $c_0,\ldots,c_{n-1}$

tal que cada c_i es el número de Gödel de una fórmula φ_i y la sucesión $\varphi_0, \ldots, \varphi_{n-1}$ es la demostración de una fórmula ψ que tiene número de Gödel d", es decir, que ψ es demostrable.

5.1. La función β de Gödel

En esta sección vamos a definir una función definible en **AP** de manera que en el modelo sea una inyección del conjunto de sucesiones finitas en $\mathbb{N}$. Sea

$$c_0, \ldots, c_{n-1}$$

una sucesión de números naturales. Al ser la sucesión ordenada tenemos que identificar la posición de cada uno de los números dentro de la lista. También guardaremos la longitud de la lista

$$(n, 0), (c_0, 1), \ldots, (c_{n-1}, n).$$

Necesitaremos ser capaces de escribir un par ordenado como un único número natural, $po(c_i)$. Necesitaremos una función $G_m : \{0, \ldots, m\} \longrightarrow \mathbb{N}$, para un m suficientemente grande, que sea inyectiva y que las imágenes sean coprimas dos a dos.

Después usaremos la factorización única en $\mathbb{N}$ para recuperar la sucesión original $c_0, \ldots, c_{n-1}$ a partir de

$$x = G_m(po(n, 0)) \cdot \prod_{i=0}^{n-1} G_m(po(c_i, i + 1)).$$

Definamos la función pares ordenados $po : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N}$ y veamos que es inyectiva:

$$po(x, y) = z :\Longleftrightarrow (x + y) \cdot (x + y) + x + \underline{1} = z$$

Lema 5.1.

$$\boldsymbol{AP} \vdash po(x, y) = po(x', y') \to x = x' \wedge y = y'$$

Demostración. Veamos primero que $x + y = x' + y'$. Por reducción al absurdo, supongamos que son distintos. Sin pérdida de generalidad, $x + y < x' + y'$. Entonces, por $\mathbf{AP}_3$ y Lema 4.3, se tiene que $s(x, y) = x + sy \leq x' + y'$ y por lo tanto

$$\begin{aligned} po(x',y') = po(x,y) &= (x+y)\cdot(x+y) + x + \underline{1} \\ &\leq (x+sy)\cdot(x+y) + (x+sy) \\ &= (x+sy)\cdot s(x+y) \\ &= s(x+y)\cdot s(x+y) \\ &\leq (x'+y')\cdot(x'+y') \\ &< po(x',y'). \end{aligned}$$

Que es absurdo, por lo tanto, $x+y \geq x'+y'$. Por simetría, $x'+y' \geq x+y$ y $x+y = x'+y'$. Entonces, $(x+y)\cdot(x+y) = (x'+y')\cdot(x'+y')$ y por el Lema 4.2, $x+\underline{1} = x'+\underline{1}$, es decir, $x = x'$ e $y = y'$. □

Dada una $\mathscr{L}_{\mathbf{AP}}$-fórmula φ, veamos cómo una función G_m como la descrita anteriormente nos permite expresar, en **AP**, el conjunto

$$\{j < m \ : \ \varphi(j)\}.$$

Siendo capaces así de aplicar la factorización única, en **AP**, en el x antes definido.

Lema 5.2. *Sea G una función definible en $\boldsymbol{AP}$, unaria y estrictamente creciente. Dada una $\mathscr{L}_{\boldsymbol{AP}}$-fórmula $\varphi(\nu)$ se tiene:*

$$\boldsymbol{AP} \vdash \forall m(\forall j < m \forall j' < m(j \neq j' \rightarrow \mathit{coprimos}(G(j), G(j'))) \rightarrow$$

$$\rightarrow \exists x \forall j < m(G(j) \mid x \leftrightarrow \varphi(j)))$$

Demostración. Procedemos por inducción en m. Para $m = 1$ tomamos $x := G(0)$ si $\varphi(0)$ se cumple y $x := 1$ si no.

Asumimos que para todos los $j, j' < m$ distintos, $G(j)$ y $G(j')$ son coprimos y que existe un x tal que para todo $j < m, G(j) \mid x \leftrightarrow \varphi(j)$. Sea x_0 el menor x que verifica esa propiedad, cuya existencia conocemos por el Principio del Menor Número. Definimos $x_1 := G(m)\cdot x_0$ si $\varphi(m)$ se cumple y $x_1 := x_0$ si no.

Veamos que para todo $j \leq m, G(j) \mid x \leftrightarrow \varphi(j)$.

Si $\varphi(m)$ se cumple, como $\mathit{coprimos}(G(j), G(m))$ para todo $j < m$,

$$G(j) \mid G(m)\cdot x_0 \leftrightarrow G(j) \mid x_0 \leftrightarrow \varphi(j).$$

Y, además, $G(m) \mid x_1$.

Si $\varphi(m)$ no se cumple, $x_1 = x_0$ y para todo $j < m$ $G(j) \mid x_0 \leftrightarrow \varphi(j)$. Supongamos que $G(m) \mid x_0$ y consecuentemente, existe r tal que $G(m)\cdot r = x_0$.

Como $m \geq 1$, G es estrictamente creciente, $G(0) \neq 0$ y coprimos(G(0),G(1)), $G(m) > 1$, lo cual significa que $r < x_0$. Como $coprimos(G(j), G(m))$ para todo $j < m$,

$$G(j) \mid G(m) \cdot r \leftrightarrow G(j) \mid r.$$

Que es absurdo por la minimalidad de x_0. □

Dada la sucesión $c_0, \ldots, c_{n-1}$, tomando, en el Lema 4.6, $j = po(c_i, i) + 1$, $k = po(c_{i'}, i') - po(c_i, i)$ e $y = m!$ con m el máximo de $po(c_i, i) + 1$, obtenemos que los $G_m(po(c_i, i)) = 1 + (po(c_i, i) + 1)m!$ son coprimos entre sí para $i \neq i'$. De esta forma, G es inyectiva y las imágenes de los elementos del paso 2, que son menores que m!, coprimas entre sí. Para construir G en **AP** necesitamos determinar primero la sucesión finita sobre la que actúa, así que, siendo m su máximo, la denotaremos por G_m.

El procedimiento descrito se puede esquematizar de la siguiente forma

$$\begin{array}{lcl}
 & c_0, \ldots, c_{n-1} & \\
 & \downarrow \quad \uparrow & \\
po & & po^{-1} \\
 & po(n,0), po(c_0,1), \ldots, po(c_{n-1}, n) & \\
 & \downarrow \quad \uparrow & \\
G_m & & G_m^{-1} \\
 & G_m(po(n,0)), G_m(po(c_0,1)) \ldots, G_m(po(c_{n-1}, n)) & \\
 & \downarrow \quad \uparrow & \text{Obtener factores} \\
\prod & & \text{de } x \text{ que sean} \\
 & x = G_m(po(n,0)) \cdot \prod_{i=0}^{n-1} G_m(po(c_i, i+1)) & \text{imagen de } G_m \\
 & \downarrow \quad \uparrow & \\
po & & po^{-1} \\
 & c = po(x, m!) &
\end{array}$$

donde el código c de la secuencia estaría definido por ambos x y $m!$.

Una vez calculado c, queremos una función β que nos devuelva la secuencia original, es decir:

*"Existe una función β, definible en **AP**, tal que, para toda sucesión finita $c_0, \ldots, c_{n-1}$ de números naturales, existe un $c \in \mathbb{N}$ tal que, para $0 \leq i \leq n-1$, $\beta(c, i+1) = c_i$ y $\beta(c, 0) = n$".*

Para esto seguiremos el camino contrario indicado en el esquema. En primer lugar calculamos x y $m!$ fácilmente. Después queremos comprobar si un elemento es divisor de x y si es imagen de G_m, es decir

$$G_m(po(c_i, i+1)) \mid x,$$

o, si queremos resaltar los c_i originales

$$(1 + (po(c_i, i) + 1) \cdot y) \mid x.$$

Por lo tanto, dado un i, para recuperar c_i basta con tomar el menor a tal que

$$(1 + (po(a, i) + 1) \cdot y) \mid x.$$

De esta forma, si definimos, en **AP**, la relación γ como

$$\gamma(a, i, y, x) :\equiv (\underline{1} + (po(a, i) + \underline{1}) \cdot y) \mid x,$$

la función $\beta(c, i)$ tendría como imagen el menor a que satisfaga γ y 0 si c e i no son el código de una serie de números. Es decir:

$$\beta(c,i) = \begin{cases} 0 & \text{si } \neg\exists x \exists y (po(x,y) = c) \\ 0 & \text{si } c = po(x,y) \wedge \neg\exists b \gamma(b,i,y,x) \\ a & \text{si } c = po(x,y) \wedge a = min\{b : \gamma(b,i,y,x)\} \end{cases}$$

y, en **AP**,

$$\beta(c,i) = a :\Longleftrightarrow (\neg\exists x\exists y(po(x,y) = c) \wedge a = 0)$$
$$\vee\exists x\exists y(c = po(x,y) \wedge ((\neg\exists b(\gamma(b,i,y,x)) \wedge a = 0)$$
$$\vee(\gamma(a,i,y,x) \wedge \neg\exists b < a(\gamma(b,i,y,x))))),$$

que es definible por ser la relación que la construye funcional.

A la hora de hacer esta construcción necesitamos que se cumpla en $\mathbb{N}$ pero también que sea demostrable en **AP**. Tomando $F(i)$ por c_{i-1} para $1 \leq i \leq n$ y $F(0)$ para n, se demuestra así

Teorema 5.1. *Sea F una función definible en **AP**,*

$$\boldsymbol{AP} \vdash \forall n \exists c \forall i \leq n (\beta(c,i) = F(i))$$

Demostración. Dado un número n, definimos $F'(i) := po(F(i), i) + 1$ y tomamos su máximo

$$m := \max_{i \leq n} F'(i).$$

Usando el Lema 4.7, tomamos y tal que para todo $j \leq m$, $j \mid y$. Definimos G_m y φ como hemos explicado en el esquema de la construcción de c:

$$G_m(j) = z :\Longleftrightarrow z = \underline{1} + (j + \underline{1})y,$$

$$\varphi(z) :\equiv \exists i < n(z = po(F(i), i)).$$

Por el Lema 4.6, $coprimos(G_m(j), G_m(j'))$ para j y j' distintos menores que m. Además, G_m es estrictamente creciente, por lo que estamos en condiciones de aplicar el lema anterior,

$$\forall j < m(1 + (j + \underline{1})y \mid x \leftrightarrow \exists i < n(j = po(F(i), i))).$$

Por construcción de x, tenemos $\gamma(F(i), i, y, x)$. Basta ver que que $F(i)$ es el mínimo con dicha propiedad. Asumamos que existe $a < F(i)$ con $\gamma(a, i, y, x)$. Por la fórmula φ, existe $j = po(a, i) = po(F(i'), i')$ con $i' < k$. Como po es inyectiva, por Lema 5.1 $i = i'$ y $a = F(i') = F(i)$, que contradice la hipótesis. □

5.2. Codificación

Una vez construida la función β estamos en disposición de codificar sucesiones de números. Por como la hemos definido, para un código c de una sucesión $c_0, \ldots, c_{n-1}$ dada por F y denotada por $\langle F(i) \mid i < n \rangle$, su longitud y cada uno de sus términos vienen dados por

$$lg(c) :\equiv \beta(c, 0),$$

$$c_i :\equiv \beta(c, i + \underline{1}).$$

Si s es el menor código que verifica las igualdades anteriores para una sucesión $s_0, \ldots, s_{lg(s)-1}$, es decir, si

$$sec(s) :\Longleftrightarrow \forall t < s(lg(t) = lg(s) \to \exists i < lg(s)(t_i \neq s_i)),$$

decimos que s es una **secuencia**.

Nótese que el cuantificador $\forall$ está acotado y que la implicación es equivalente a $lg(t) \neq lg(s) \vee \exists i < lg(s)(t_i \neq s_i)$ de manera que $sec(s)$ es una $\exists$-fórmula. Queremos que las fórmulas utilizadas en la codificación en **AP** sean ciertas ambas en **AP** y en $\mathbb{N}$, es decir, necesitamos que sean $\mathbb{N}$-conformes. De forma que, utilizaremos sólo $\exists$-fórmulas o fórmulas en las que todo cuantificador esté acotado.

Para manipular las secuencias, necesitaremos una función que las concatene

$$s \frown s' = t :\Longleftrightarrow sec(t) \wedge lg(t) = lg(s) + lg(s') \wedge$$

$$\forall i < lg(s)(t_i = s_i) \wedge \forall i < lg(s')(t_{lg(s)+i} = s'_i).$$

Así podremos añadirles elementos o subdividirlas en otras. Veamos que en efecto dicha concatenación es posible en **AP**:

Proposición 5.1.

$$\boldsymbol{AP} \vdash \forall s \forall s' \exists t(t = s \frown s')$$

Demostración. Basta tomar la siguiente F definible en **AP**

$$F(i) = \begin{cases} lg(s) + lg(s') & \text{si } i = 0 \\ \beta(s, i) & \text{si } 0 < i \leq lg(s) \\ \beta(s', i - lg(s)) & \text{si } lg(s) < i \end{cases}$$

por el teorema 5.1 sabemos que existe un código c asociado y por el Principio del Menor Número sabemos que existe una secuencia t correspondiente. $\square$

Nótese que la concatenación es asociativa, así que podemos escribir $s \frown s' \frown s''$ en vez de $s \frown (s' \frown s'')$.

Como esperábamos resolver al principio del capítulo, la función β nos permite construir potencias de números. Comprobando que la siguiente función es funcional podemos construir secuencias correspondientes,

$$x^k = y :\Longleftrightarrow \exists t(sec(t) \wedge lg(t) = sk \wedge t_0 = \underline{1} \wedge \forall i < k(t_{si} = x \cdot t_i) \wedge t_{sk} = y).$$

Para verlo, tomamos una secuencia $\langle 1, x, \ldots, x^k \rangle$ asociada a r. Poder construir el siguiente término dentro de ella significa que éste es único. Basta tomar $t = r \frown \langle x \cdot r_k \rangle$ para obtener una fórmula para él.

Ahora podemos comprobar que la factorización única es posible en **AP** para la inyección f del principio del capítulo, es decir

$$\mathbf{AP} \vdash \underline{2}^x \cdot \underline{3}^y \cdot \underline{5}^z = \underline{2}^{x'} \cdot \underline{3}^{y'} \cdot \underline{5}^{z'} \rightarrow x = x' \wedge y = y' \wedge z = z'.$$

Estos resultados nos permitirán construir números de Gödel en **AP**, y como $\mathbb{N}$ es un modelo de ésta, también en $\mathbb{N}$. Usando la siguiente notación:

$$\ulcorner \zeta \urcorner :\equiv \underline{\#\zeta},$$

para ζ un símbolo, término o fórmula de **AP** y con los números de Gödel de cada símbolo definidos en su tabla, estamos listos para introducir

$$suc(n) :\equiv \underline{2}^{\ulcorner s \urcorner} \cdot \underline{3}^{n},$$

$$mult(n,m) :\equiv \underline{2}^{\ulcorner \cdot \urcorner} \cdot \underline{3}^{n} \cdot \underline{5}^{m},$$
$$not(f) :\equiv \underline{2}^{\ulcorner \neg \urcorner} \cdot \underline{3}^{f},$$
$$or(f,f') :\equiv \underline{2}^{\ulcorner \vee \urcorner} \cdot \underline{3}^{f} \cdot \underline{5}^{f'},$$
$$ex(v,f) :\equiv \underline{2}^{\ulcorner \exists \urcorner} \cdot \underline{3}^{v} \cdot \underline{5}^{f},$$
$$sum(n,m) :\equiv \underline{2}^{\ulcorner + \urcorner} \cdot \underline{3}^{n} \cdot \underline{5}^{m}$$
$$ig(t,t') :\equiv \underline{2}^{\ulcorner = \urcorner} \cdot \underline{3}^{t} \cdot 5^{t'},$$
$$and(f,f') :\equiv \underline{2}^{\ulcorner \wedge \urcorner} \cdot \underline{3}^{f} \cdot \underline{5}^{f'},$$
$$imp(f,f') :\equiv \underline{2}^{\ulcorner \to \urcorner} \cdot \underline{3}^{f} \cdot \underline{5}^{f'},$$
$$pt(v,f) :\equiv \underline{2}^{\ulcorner \forall \urcorner} \cdot \underline{3}^{v} \cdot \underline{5}^{f}.$$

Con las siguientes relaciones podemos identificar el tipo de cada objeto de $\mathscr{L}_{\mathbf{AP}}$. Si un número es, por ejemplo, el número de Gödel de un término τ, existe una sucesión en la que sus elementos son números de Gödel de variables o números de Gödel de sucesores, sumas o multiplicaciones de elementos anteriores, de forma que su último elemento es $\ulcorner \tau \urcorner$.

En algunas de éstas nos hemos alejado de las que aparecen en la referencia principal [HK20], ya que, a nuestro parecer, en ellas, o no aparece tan claro que sean $\mathbb{N}$-conformes, o la demostración de que lo sean es innecesariamente compleja.

Definimos variables

$$var(v) :\Longleftrightarrow v \neq \underline{1} \wedge \exists n < v(v = \underline{2} \cdot n + \underline{1}),$$

términos, $c\_term(c,t)$ expresa que t es el número de un término que se construye con la sucesión codificada por c mediante la función β,

$$c\_term(c,t) :\Longleftrightarrow sec(c) \wedge c_{lg(c)-1} = t \wedge \forall k < lg(c)$$

$$(var(c_k) \vee c_k = 0 \vee \exists i < k \exists j < k(c_k = suc(c_i) \vee$$
$$c_k = sum(c_i, c_j) \vee c_k = mult(c_i, c_j))).$$

$term(t)$ expresa que t es el código de un término,

$$term(t) :\Longleftrightarrow \exists c(c\_term(c, t)).$$

Y fórmulas, $c\_fml(c, b, f)$ expresa que f es el número de una fórmula que se construye mediante la sucesión c con cota para sus elementos b

$$c\_fml(c, b, f) :\Longleftrightarrow sec(c) \wedge c_{lg(c)-1} = f \wedge \forall k < lg(c)$$
$$(\exists t < b \exists t' < b \exists d < b \exists d' < b(c\_term(d, t) \wedge c\_term(d', t') \wedge c_k = ig(t, t')) \vee$$
$$\exists i < k \exists j < k(c_k = no(c_i) \vee c_k = and(c_i, c_j) \vee c_k = or(c_i, c_j) \vee c_k = imp(c_i, c_j)) \vee$$
$$\exists i < k \exists v < b(var(v) \wedge (c_k = ex(v, c_i) \vee c_k = pt(v, c_i)))).$$

Y $fml(f)$ expresa que f es el código de una fórmula,

$$fml(f) :\Longleftrightarrow \exists c \exists b(c\_fml(c, b, f)).$$

Veamos ahora que estas definiciones funcionan, es decir:

Lema 5.3. *Dado $n \in \mathbb{N}$,*

a) $\mathbb{N} \vDash var(n)$ si y sólo si $n \equiv \#\nu$ para alguna variable ν.

b) $\mathbb{N} \vDash term(n)$ si y sólo si $n \equiv \#\tau$ para algún $\mathscr{L}_{\boldsymbol{AP}}$-término τ.

c) $\mathbb{N} \vDash fml(n)$ si y sólo si $n \equiv \#\varphi$ para alguna $\mathscr{L}_{\boldsymbol{AP}}$-fórmula φ.

Demostración. El caso a) es trivial, demostramos un caso particular de c) por inducción en la construcción de una fórmula φ y los demás casos son análogos.

Para el caso básico, sea $\varphi \equiv \tau_1 \vee \tau_2$ una fórmula con τ_1, τ_2 términos, tenemos que $\mathbb{N} \vDash c\_term(\langle \#\tau_1, \#\tau_2, \#\tau_1 \vee \tau_2 \rangle, \#\tau_1 \vee \tau_2)$. Ahora asumimos que $\mathbb{N} \vDash fml(\#\varphi')$ para una fórmula φ' de mayor longitud y veamos que el resultado se cumple para $\varphi \equiv \neg\varphi'$. Sea $c \in \mathbb{N}$ tal que $\mathbb{N} \vDash c\_term(c, \#\varphi')$, sabemos que $\mathbb{N} \vDash no(\#\varphi')$ y por lo tanto $\mathbb{N} \vDash c\_term(c \frown \langle \#\neg\varphi' \rangle, \#\neg\varphi')$.

Para el recíproco, asumamos que se cumple para todo $m < n \in \mathbb{N}$ y sea $c \in \mathbb{N}$ con $\mathbb{N} \vDash c\_term(c, n)$. Si $n \equiv no^{\mathbb{N}}(c_i)$ con $c_i < lg(c)$, tomamos una fórmula φ' con $c_i \equiv \#\varphi'$. Por $\mathbb{N}$-conformidad, $n \equiv no^{\mathbb{N}}(c_i) \equiv (\underline{2}^{\ulcorner\neg\urcorner} \cdot \underline{3}^{\underline{c_i}})^{\mathbb{N}} \equiv \#\neg\varphi'$.

□

Veamos ahora las reglas de sustitución. Empezamos por relaciones que nos indiquen si una variable con código v aparece en un término con código t

$$var\_en\_term(v,t) :\Longleftrightarrow var(v)\wedge$$

$$\exists c(c\_term(c,t)\wedge\forall c'<c\neg c\_term(c',t)\wedge\exists i<lg(c)(c_i=v)),$$

o en una fórmula de código f

$$var\_en\_fml(v,f) :\Longleftrightarrow \exists c\exists b(c\_fml(c,b,f)\wedge\forall c'<c\forall b'<b\neg c\_fml(c',b',f)\wedge$$

$$\exists i<lg(c)\exists t_0\exists t_1(term(t_0)\wedge term(t_1)\wedge c_i=ig(t_0,t_1)\wedge$$
$$(var\_en\_term(v,t_0)\vee var\_en\_term(v,t_1)))).$$

Para dar la relación que indica si una variable aparece libre en una fórmula, usamos una definición más sencilla, ya que la original es excesivamente compleja de codificar. Escribimos que una variable es libre en una fórmula cuando aparece en la fórmula pero no aparece cuantificada, ya que, en las apariencias cuantificadas, basta tomar otra variable cualquiera que no se haya utilizado y la fórmula sería equivalente. Por lo tanto, dado un código v de una variable y dado un código f de una fórmula,

$$libre(v,f) :\Longleftrightarrow \exists c\exists b(c\_fml(c,b,f)\wedge var\_en\_fml(v,f)\wedge$$

$$\forall i<lg(c)\forall j<i(c_i\neq ex(v,c_j)\wedge c_i\neq pt(v,c_j))).$$

Introducimos también una relación nueva $novent(v,t)$, que indica cuándo una variable v no está en un término t:

$$novent(v,t) :\Longleftrightarrow var(v)\wedge\exists c(c\_term(c,t)\wedge\forall i<lg(c)(v\neq c_i)),$$

que nos sirve para definir la relación $novenf(v,f)$, similar, pero para fórmulas:

$$novenf(v,f) :\Longleftrightarrow var(v)\wedge\exists c\exists b(c\_fml(c,b,f)\wedge$$

$$\forall c'<c\forall b'<b\neg c\_fml(c',b',f))\wedge\forall i<lg(c)\forall t_1<b\forall t_2<b(c_i\neq ig(t_1,t_2)\vee$$
$$(novent(v,t_1)\wedge novent(v,t_2))).$$

La siguiente relación indica cuándo la sustitución de una variable v en una fórmula f por un término t es admisible.

$$st\_adm(v,t,f) :\Longleftrightarrow var(v)\wedge var\_en\_fml(v,f)\wedge libre(v,f)\wedge$$

$$\exists c \exists c' \exists b' (c\_term(c,t) \wedge c\_fml(c',b',f) \wedge$$
$$\forall i < lg(c) \exists n (c_i = n + n \vee c_i = \ulcorner 0 \urcorner \vee libre(c_i, f)))$$

Ésta expresa que t'' (con secuencia c'') es el código resultante de sustituir v por t (con secuencia c) en el término t' (con secuencia c'):

$$c\_st\_term(c,c',c'',v,t,t',t'') :\Longleftrightarrow var(v) \wedge$$
$$c\_term(c,t) \wedge c\_term(c',t') \wedge c\_term(c'',t'') \wedge lg(c'') = lg(c) + lg(c') \wedge$$
$$\forall k < lg(c)(c''_k = c_k) \wedge \forall k < lg(c')\big($$
$$(var(c'_k) \wedge (c'_k \neq v) \wedge c''_{lg(c)+k} = c'_k) \vee$$
$$(c'_k = v \wedge c''_{lg(c)+k} = t) \vee$$
$$(c'_k = \ulcorner 0 \urcorner \wedge c''_{lg(c)+k} = \ulcorner 0 \urcorner) \vee$$
$$\exists i < k (c'_k = suc(c'_i) \wedge c''_{lg(c)+k} = suc(c''_{lg(c)+i})) \vee$$
$$\exists i < k \exists j < k (c'_k = sum(c'_i, c'_j) \wedge c''_{lg(c)+k} = sum(c''_{lg(c)+i}, c''_{lg(c)+j})) \vee$$
$$\exists i < k \exists j < k (c'_k = mult(c'_i, c'_j) \wedge c''_{lg(c)+k} = mult(c''_{lg(c)+i}, c''_{lg(c)+j}))\big).$$

Eliminando las secuencias de la fórmula:

$$st\_term(v,t,t',t'') :\Longleftrightarrow \exists c \exists c' \exists c'' (c\_st\_term(c,c',c'',v,t,t',t'')).$$

Y similarmente en fórmulas,

$$c\_st\_fml(c,b,c',b',v,t,f,f') :\Longleftrightarrow c\_fml(c,b,f) \wedge c\_fml(c',b',f') \wedge$$
$$st\_adm(v,t,f) \wedge lg(c') = lg(c) \wedge \forall k < lg(c)\big($$
$$\exists t \exists s \exists t' \exists s' (term(t) \wedge term(s) \wedge term(t') \wedge term(s') \wedge$$
$$st\_term(v,t,t,t') \wedge st\_term(v,t,s,s') \wedge c_k = ig(t,s) \wedge c'_k = ig(t',s')) \vee$$
$$\exists i < k (c_k = not(c_i) \wedge c'_k = not(c'_i)) \vee$$
$$\exists i < k \exists j < k (c_k = or(c_i, c_j) \wedge c'_k = or(c'_i, c'_j)) \vee$$
$$\exists i < k \exists j < k (c_k = and(c_i, c_j) \wedge c'_k = and(c'_i, c'_j)) \vee$$
$$\exists i < k \exists j < k (c_k = imp(c_i, c_j) \wedge c'_k = imp(c'_i, c'_j)) \vee$$
$$\exists v' \exists j < k (var(v') \wedge c_k = ex(v', c_j) \wedge c'_k = ex(v', c'_j)) \vee$$

$$\exists v' \exists j < k(var(v') \wedge c_k = pt(v', c_j) \wedge c'_k = pt(v', c'_j))\big).$$

Eliminando las secuencias de la fórmula

$$st\_fml(v, t, f, f') :\Longleftrightarrow \exists c \exists b \exists b' \exists c' (c\_st\_fml(c, b, c', b', v, t, f, f')).$$

Como habíamos mencionado, las relaciones definidas son $\mathbb{N}$-conformes, por lo que tenemos los siguientes lemas.

Lema 5.4. *Sea ν es una variable, τ un término y φ una fórmula, entonces:*

$$\boldsymbol{AP} \vdash var(\ulcorner \nu \urcorner)$$

$$\boldsymbol{AP} \vdash term(\ulcorner \tau \urcorner)$$

$$\boldsymbol{AP} \vdash fml(\ulcorner \varphi \urcorner)$$

Lema 5.5. *Sean τ y τ_0 dos términos, φ una fórmula y ν una variable tal que la sustitución $\varphi(\nu/\tau_0)$ es admisible, entonces:*

$$\boldsymbol{AP} \vdash st\_term(\ulcorner \nu \urcorner, \ulcorner \tau_0 \urcorner, \ulcorner \tau \urcorner, t) \leftrightarrow t = \ulcorner \tau(\nu/\tau_0) \urcorner$$

$$\boldsymbol{AP} \vdash st\_fml(\ulcorner \nu \urcorner, \ulcorner \tau_0 \urcorner, \ulcorner \varphi \urcorner, f) \leftrightarrow f = \ulcorner \varphi(\nu/\tau_0) \urcorner$$

Ahora que podemos identificar cada fórmula, basta obtener unas relaciones que nos identifiquen los axiomas y las reglas de inferencia para saber cuándo una sucesión de fórmulas es una demostración bien formada. Empezamos con los axiomas lógicos

$$ax\_L_0(f) :\Longleftrightarrow \exists f'(fml(f') \wedge f = or(f', not(f'))),$$

$$ax\_L_1(f) :\Longleftrightarrow \exists f' \exists f''(fml(f') \wedge fml(f'') \wedge f = imp(f', imp(f'', f'))),$$

$$ax\_L_{10}(f) :\Longleftrightarrow \exists t \exists v \exists f' \exists f''(st\_fml(v, t, f', f'') \wedge f = imp(pt(v, f'), f'')),$$

$$ax\_L_{11}(f) :\Longleftrightarrow \exists t \exists v \exists f' \exists f''(st\_fml(v, t, f', f'') \wedge f = imp(f'', ex(v, f'))),$$

$$ax\_L_{12}(f) :\Longleftrightarrow \exists v \exists f' \exists f''(novenf(v, f') \wedge fml(f') \wedge fml(f'') \wedge$$
$$f = imp(pt(v, imp(f', f'')), imp(f', pt(v, f'')))),$$

$$ax\_L_{13}(f) :\Longleftrightarrow \exists v \exists f' \exists f''(novenf(v, f') \wedge fml(f') \wedge fml(f'') \wedge$$
$$f = imp(pt(v, imp(f'', f')), imp(ex(v, f''), f'))).$$

Los demás se construyen similarmente. Ahora igualmente dos ejemplos de los axiomas de Peano,

$$ax\_\mathbf{AP}_1(f) :\Longleftrightarrow \exists f'\exists f''\exists v_0\exists v_1\exists g\exists g'(term(v_0)\wedge$$

$$term(v_1)\wedge g = ig(suc(v_0), suc(v_1))\wedge g' = ig(v_0, v_1)\wedge$$

$$f'' = imp(g, g')\wedge f' = pt(v_1, f'')\wedge f = pt(v_0, f)),$$

$$ax\_\mathbf{AP}_6(f) :\Longleftrightarrow \exists f'\exists f''\exists f'''\exists v\exists g(libre(v, f')\wedge$$

$$st\_fml(v, \ulcorner 0\urcorner, f', f'')\wedge st\_fml(v, suc(v), f', f''')\wedge$$

$$g = pt(v, imp(f', f'''))\wedge f = imp(and(f'', g), pt(v, f'))).$$

Con las siguientes relaciones identificamos si f es un axioma,

$$log\_ax(f) :\Longleftrightarrow ax\_L_0(f)\vee\ldots\vee ax\_L_{16}(f),$$

$$peano\_ax(f) :\Longleftrightarrow ax\_\mathbf{AP}_0(f)\vee\ldots\vee ax\_\mathbf{AP}_6(f),$$

$$ax(f) :\Longleftrightarrow log\_ax(f)\vee peano\_ax(f).$$

Y ahora, las reglas de inducción se codifican como sigue,

$$mp(f'', f', f) :\Longleftrightarrow fml(f')\wedge fml(f)\wedge f'' = imp(f', f),$$

$$gen(v, f', f) :\Longleftrightarrow var(v)\wedge fml(f')\wedge f = pt(v, f').$$

Con esto tenemos todas las herramientas necesarias para enunciar el **predicado de demostrabilidad** *dem*, que se cumple cuando los elementos de una secuencia son axiomas o aplicaciones de reglas de inferencia.

$$c\_dem(c, f) :\Longleftrightarrow sec(c)\wedge c_{lg(c)-1} = f\wedge$$

$$\forall k < lg(c)(ax(c_k)\vee\exists i < k\exists j < k(mp(c_i, c_j, c_k)\vee\exists v(gen(v, c_i, c_k))))$$

$$dem(f) :\Longleftrightarrow \exists c(c\_dem(c, f))$$

Estos dos resultados definen las propiedades del demostrado de demostrabilidad. Son esencialmente las propiedades que tiene nuestra teoría para construir los enunciados recursivos que mencionábamos en la introducción.

Lema 5.6. *Sea φ una $\mathcal{L}_{\boldsymbol{AP}}$-fórmula donde ninguna variable aparece ligada y libre a la vez, entonces:*

$$\mathbb{N}\vDash dem(\#\varphi) \text{ si y sólo si } \boldsymbol{AP}\vdash\varphi$$

Demostración. Sea c la secuencia de asociada a $dem(\#\varphi)$. Como en el Lema 5.3, $\mathbb{N} \vDash ax(m)$ si y sólo si m es un código de un axioma, y se demuestra de la misma forma, por inducción en la longitud $lg(c)$ en una dirección y en la longitud de la demostración de φ en la otra. □

Como para toda $\mathscr{L}_{\mathbf{AP}}$-fórmula en la que existan variables ligadas y libres a la vez se puede encontrar una fórmula equivalente, añadiendo nuevas variables, en la que se satisfacen las condiciones del lema, dicho resultado es válido para cualquier $\mathscr{L}_{\mathbf{AP}}$-fórmula en general.

Corolario 5.1. *Sea φ una $\mathscr{L}_{\boldsymbol{AP}}$-fórmula, entonces:*

$$\textit{Si } \boldsymbol{AP} \vdash \varphi, \textit{ entonces } \boldsymbol{AP} \vdash dem(\ulcorner \varphi \urcorner)$$

Demostración. Asumiendo que $\mathbf{AP} \vdash \varphi$, por el lema anterior sabemos que $\mathbb{N} \vDash c\_dem(c, \#\varphi)$ con $c \in \mathbb{N}$. Como c_dem y todas las fórmulas que hemos utilizado para construirlo están definidas por $\exists$-fórmulas, por la proposición 4.2, $\mathbf{AP} \vdash c\_dem(\underline{c}, \ulcorner \varphi \urcorner)$. □

6. Los Teoremas de Incompletitud

6.1. Primer Teorema de Incompletitud

Primero definimos unas relaciones que nos indican cuándo un número n es el código de un número natural. Primero, escribimos la relación que expresa que c es la sucesión que muestra que x es el código de n.

$$c\_nat(c, n, x) :\Longleftrightarrow sec(c) \wedge lg(c) = sn \wedge c_0 = \ulcorner 0 \urcorner \wedge$$

$$\forall i < n(c_{si} = suc(c_i) \wedge c_n = x)$$

Y, una fórmula que exprese que x es el código de algún número natural.

$$nat(n, x) :\Longleftrightarrow \exists c(c\_nat(c, n, x))$$

Y eliminando la secuencia,

Lema 6.1. *Para todo $n \in \mathbb{N}$ se tiene que $\boldsymbol{AP} \vdash nat(\underline{n}, \ulcorner \underline{n} \urcorner)$.*

Por lo tanto, damos la siguiente definición, el **número de Gödel** de un número.

$$ng(n) = x :\Longleftrightarrow nat(n, x) \vee (\neg\exists y(nat(n, y)) \wedge x = 0)$$

Particularmente, por el lema anterior, tenemos:

$$\mathbf{AP} \vdash ng(\ulcorner\varphi\urcorner) = \ulcorner\ulcorner\varphi\urcorner\urcorner$$

Con todo esto estamos en condiciones de demostrar que es posible construir los enunciados autorreferentes que habíamos mencionado en la introducción. El siguiente teorema indica que la aritmética de Peano es autorreferente, ya que sirve para construir sentencias que se refieren a sí mismas. Es decir, informalmente, sentencias que digan "tengo la propiedad φ".

Teorema 6.1 (Lema de la diagonalización)**.** *Dada una $\mathscr{L}_{\boldsymbol{AP}}$-fórmula $\varphi(\nu)$ con una variable libre ν que no está ligada en φ, existe una $\mathscr{L}_{\boldsymbol{AP}}$-sentencia σ_φ tal que:*

$$\sigma_\varphi \Leftrightarrow_{\boldsymbol{AP}} \varphi(\ulcorner\sigma_\varphi\urcorner)$$

Demostración. Definimos:

$$\psi(v_0) :\equiv \forall v_1(st\_fml(\ulcorner v_0\urcorner, ng(v_0), v_0, v_1) \to \varphi(\nu/v_1)).$$

La sentencia buscada es

$$\sigma_\varphi :\equiv \psi(v_0/\ulcorner\psi\urcorner).$$

Damos una demostración en base a los resultados previos:

$$\begin{array}{lll}
\sigma_\varphi & \equiv & \forall v_1(st_fml(\ulcorner v_0\urcorner, ng(\ulcorner\psi\urcorner), \ulcorner\psi(v_0)\urcorner, v_1) \to \varphi(\nu/v_1)) \\
 & & \text{Lema 6.1} \\
 & \Leftrightarrow_{\mathbf{AP}} & \forall v_1(st_fml(\underline{3}, \ulcorner\ulcorner\psi\urcorner\urcorner, \ulcorner\psi(v_0)\urcorner, v_1) \to \varphi(v_1)) \\
 & & \text{Lema 5.5} \\
 & \Leftrightarrow_{\mathbf{AP}} & \forall v_1(v_1 = \ulcorner\psi(v_0/\ulcorner\psi\urcorner)\urcorner \to \varphi(v_1)) \\
 & & L_{12} \text{ y } L_{10} \\
 & \Leftrightarrow_{\mathbf{AP}} & \varphi(\ulcorner\psi(v_0/\ulcorner\psi\urcorner)\urcorner) \\
 & \equiv & \varphi(\ulcorner\psi(\ulcorner\psi\urcorner)\urcorner) \\
 & \equiv & \varphi(\ulcorner\sigma_\varphi\urcorner).
\end{array}$$

□

Bastaría tomar como $\varphi(\nu)$ "este enunciado no es demostrable" para probar que **AP** es incompleta, $\varphi(\nu)$ no es demostrable en **AP**. Ése es el primer teorema de incompletitud.

Teorema 6.2 (Primer Teorema de Incompletitud)**.** *La axiomática de Peano es incompleta.*

Demostración. Tomando $\varphi(v_0) :\equiv \neg dem(v_0)$ en el lema de diagonalización sabemos que existe σ tal que:

$$\sigma \Leftrightarrow_{\mathbf{AP}} \neg dem(\ulcorner \sigma \urcorner)$$

Ahora, si asumimos que **AP** es completa, tenemos dos posibilidades:

Si $\mathbf{AP} \vdash \sigma$, esto implica que $\mathbf{AP} \vdash \neg dem(\ulcorner \sigma \urcorner)$. Además, por el Colorario 5.1, $\mathbf{AP} \vdash dem(\ulcorner \sigma \urcorner)$, es decir: $\mathbf{AP} \vdash dem(\ulcorner \sigma \urcorner) \wedge \neg dem(\ulcorner \sigma \urcorner)$.

Si $\mathbf{AP} \vdash \neg\sigma$, esto implica, por hipótesis, que $\mathbf{AP} \vdash \neg(\neg dem(\ulcorner \sigma \urcorner))$, es decir, $\mathbf{AP} \vdash dem(\ulcorner \sigma \urcorner)$ y $\mathbb{N} \vDash dem(\#\sigma)$. Por el Lema 5.6, tenemos $\mathbf{AP} \vdash \sigma$, es decir: $\mathbf{AP} \vdash \sigma \wedge \neg\sigma$.

Por lo tanto, si **AP** es consistente, $\mathbf{AP} \nvdash \sigma$ y $\mathbf{AP} \nvdash \neg\sigma$. **AP** es incompleta. □

6.2. Segundo Teorema de Incompletitud

Ahora entremos con el Segundo Teorema de Incompletitud, que afirma que no se puede demostrar la consistencia de la axiomática de Peano en la propia axiomática de Peano. Hemos visto un conjunto de fórmulas es consistente si no puede demostrar contradicciones. En el caso de una teoría, decimos que es consistente cuando no se pueden demostrar contradicciones a partir de su conjunto de axiomas. En primer lugar veamos los siguientes resultados.

Proposición 6.1. *Dados x,y arbitrarios, se tiene*

a) $\boldsymbol{AP} \vdash (dem(x) \wedge dem(imp(x, y))) \to dem(y)$
b) $\boldsymbol{AP} \vdash (dem(x) \wedge dem(y)) \to dem(and(x, y))$

Demostración. Para a), por las hipótesis tenemos $mp(imp(x, y), x, y)$. Sean c, c' tales que $c\_dem(c, x)$ y $c\_dem(c', imp(x, y))$. Tenemos que $c\_dem(c \frown c' \frown \langle y \rangle, y)$, por lo que $dem(y)$.

Para b), por las hipótesis tenemos $fml(x)$ y $fml(y)$. Usando L_5 formalizado

$$\mathbf{AP} \vdash dem(imp(y, imp(x, and(x, y)))).$$

Usando el apartado a) dos veces se obtiene $dem(and(x, y))$, como se deseaba. □

Tomando fórmulas en vez de números arbitrarios, se obtiene el siguiente corolario.

Proposición 6.2. *Dadas φ y ψ dos $\mathscr{L}_{\boldsymbol{AP}}$-fórmulas se cumple:*

a) $\boldsymbol{AP} \vdash dem(\ulcorner \varphi \to \psi \urcorner) \to (dem(\ulcorner \varphi \urcorner) \to dem(\ulcorner \psi \urcorner))$
b) $\boldsymbol{AP} \vdash (dem(\ulcorner \varphi \urcorner) \wedge dem(\ulcorner \psi \urcorner)) \to dem(\ulcorner \varphi \wedge \psi \urcorner)$

Estos resultados los utilizamos para demostrar las condiciones de derivabilidad, que establecen las propiedades de transitividad de implicaciones en el predicado de demostrabilidad. Estas condiciones las necesitaremos en la demostración del segundo teorema.

Proposición 6.3 (Condiciones de derivabilidad de Hilbert-Bernays-Löb)**.** *Dada φ una $\mathscr{L}_{\boldsymbol{AP}}$-fórmula se cumple:*

D_0*: Si* $\boldsymbol{AP} \vdash \varphi$*, entonces* $\boldsymbol{AP} \vdash dem(\ulcorner \varphi \urcorner)$
D_1*:* $\boldsymbol{AP} \vdash dem(\ulcorner \varphi \to \psi \urcorner) \to (dem(\ulcorner \varphi \urcorner) \to dem(\ulcorner \psi \urcorner))$
D_2*:* $\boldsymbol{AP} \vdash dem(\ulcorner \varphi \urcorner) \to dem(\ulcorner dem(\ulcorner \varphi \urcorner) \urcorner)$

Demostración. D_0 es el Corolario 5.1 y D_1 es una implicación de la Proposición 6.2 a).

La demostración de D_2 es más complicada y la damos esquemáticamente ya que los detalles son técnicos. Obsérvese que D_2 es como D_0 sólo que la implicación se encuentra en **AP** y no en el metalenguaje. Esto nos traerá algunas complicaciones técnicas. Veamos el siguiente ejemplo.

Si $\mathbf{AP} \vdash \varphi \to dem(\ulcorner \varphi \urcorner)$ fuera cierto, tomamos $\varphi \equiv (v_0 = v_1)$, obteniendo $\mathbf{AP} \vdash v_0 = v_1 \to dem(\ulcorner v_0 = v_1 \urcorner)$. En esta última expresión las variables v_0 y v_1 son libres en φ, pero $dem(\ulcorner v_0 = v_1 \urcorner)$ es una sentencia. Por lo tanto, tomando la sustitución $v_0 \equiv v_1 \equiv 0$, obtenemos $\mathbf{AP} \vdash 0 = 0 \to dem(\ulcorner v_0 = v_1 \urcorner)$ y usando (MP), $\mathbf{AP} \vdash dem(\ulcorner v_0 = v_1 \urcorner)$, que es falso porque no existe demostración de $v_0 = v_1$.

Este problema que nos encontramos con las variables libres lo podemos resolver con la siguiente definición. Dada una variable ν y un conjunto de variables V,

$$\lceil \nu \rceil_V :\equiv \begin{cases} \nu & \text{si } \nu \in V \\ \lceil \nu \rceil & \text{si no} \end{cases}$$

cuya idea es conservar las variables de V y convertir en números de Gödel las demás. Lo hacemos también para términos y fórmulas

$$\lceil 0 \rceil_V :\equiv \lceil 0 \rceil$$

$$\lceil s\tau \rceil_V :\equiv suc(\ulcorner \tau \urcorner_V)$$
$$\lceil \tau_1 + \tau_2 \rceil_V :\equiv sum(\lceil \tau_1 \rceil_V, \lceil \tau_2 \rceil_V)$$
$$\lceil \tau_1 \cdot \tau_2 \rceil_V :\equiv mult(\lceil \tau_1 \rceil_V, \lceil \tau_2 \rceil_V)$$
$$\lceil \varphi \wedge \psi \rceil_V :\equiv and(\varphi, \psi)$$
$$\lceil \forall \nu \varphi \rceil_V :\equiv pt(\ulcorner \nu \urcorner, \lceil \varphi \rceil_{V \setminus \{\nu\}})$$

El resto de fórmulas son análogas. En el caso de que todas las variables libres de φ estén en V, escribimos $\lceil \varphi \rceil$ en vez de $\lceil \varphi \rceil_V$. Para que el proceso que estamos construyendo sea igual que la codificación original, que sustituye las variables libres restantes en $\lceil \varphi \rceil$ por su número de Gödel

$$\lceil \tau \rceil_V^{ng} :\equiv \lceil \tau \rceil_V (x_1/ng(x_1), ..., x_n/(ng(x_n))$$

$$\lceil \varphi \rceil_V^{ng} :\equiv \lceil \varphi \rceil_V (x_1/ng(x_1), ..., x_n/(ng(x_n))$$

Y con estas definiciones obtenemos que la sustitución es efectivamente equivalente. Dada una variable $x \in V$ con V un conjunto finito de variables, τ un término y φ una fórmula con $x \in libres(\varphi)$, se tiene

$$\mathbf{AP} \vdash st\_term(\ulcorner x \urcorner, ng(x), \lceil \tau \rceil_{V \setminus \{x\}}, \lceil \tau \rceil_V)$$
$$\mathbf{AP} \vdash st\_fml(\ulcorner x \urcorner, ng(x), \lceil \varphi \rceil_{V \setminus \{x\}}, \lceil \varphi \rceil_V)$$

La demostración se hace por inducción en la construcción del término o fórmula de modo mecánico.

Usando ese resultado, obtenemos lo siguiente

$$\mathbf{AP} \vdash dem(\ulcorner \varphi \urcorner) \to dem(\lceil \varphi \rceil^{ng}).$$

Como corolario, tenemos que $\mathbf{AP} \vdash \varphi$ implica $\mathbf{AP} \vdash dem(\lceil \varphi \rceil^{ng})$ ya que $\mathbf{AP} \vdash \varphi$ implica $\mathbf{AP} \vdash dem(\ulcorner \varphi \urcorner)$ por D_0.

Por lo tanto, dadas dos $\mathscr{L}_{\mathbf{AP}}$-fórmulas φ y ψ equivalentes en $\mathbf{AP}$, se obtiene que

$$\mathbf{AP} \vdash dem(\lceil \varphi \to \psi \rceil^{ng}),$$

donde $\ulcorner \varphi \to \psi \urcorner^{ng}$ equivale a $imp(\ulcorner \varphi \urcorner^{ng}, \ulcorner \psi \urcorner^{ng})$ y, aplicando el Lema 6.1, obtenemos $\mathbf{AP} \vdash dem(\ulcorner \varphi \urcorner^{ng}) \to dem(\ulcorner \psi \urcorner^{ng})$. Por simetría, esto implica que $dem(\ulcorner \varphi \urcorner^{ng})$ y $dem(\ulcorner \psi \urcorner^{ng})$ son equivalentes en $\mathbf{AP}$.

Ahora obsérvese que basta probar el siguiente resultado para toda $\exists$-fórmula φ_0

$$\mathbf{AP} \vdash \varphi_0 \to dem(\ulcorner \varphi_0 \urcorner^{ng})$$

ya que el predicado de demostrabilidad es una $\exists$-fórmula. Como las $\exists$-fórmulas son equivalentes en $\mathbf{AP}$ a $\exists$-fórmulas estrictas, por el corolario anterior, con probar el resultado para las estrictas habríamos terminado.

Sea φ_0 una $\exists$-fórmula estricta, esta demostración se hace por inducción en la construcción de φ_0. Es decir, para fórmulas atómicas, para disyunción de éstas, para conjunción de éstas, para su cuantificación existencial y para su cuantificación universal acotada. Lo demostramos para el caso $\varphi_0 \equiv \exists v_i \varphi$ como ejemplo. Por L_{11} y el corolario anterior

$$\mathbf{AP} \vdash dem(\ulcorner \varphi \urcorner^{ng}) \to dem(\ulcorner \exists v_i \varphi \urcorner^{ng}).$$

Usando la Tautología 2.6,

$$\mathbf{AP} \vdash \varphi \to dem(\ulcorner \exists\ v_i \varphi \urcorner^{ng}).$$

Aplicando generalización,

$$\mathbf{AP} \vdash \forall v_i (\varphi \to dem(\ulcorner \exists v_i \varphi \urcorner^{ng})),$$

y, como v_i no es libre en $dem(\ulcorner \exists v_i \varphi \urcorner^{ng})$, aplicando L_{13} y (MP),

$$\mathbf{AP} \vdash \exists v_i \varphi \to dem(\ulcorner \exists v_i \varphi \urcorner^{ng}).$$

□

Con esto llegamos al segundo teorema y final de la memoria, que niega lo que se había estado intentado varias décadas en la lógica matemática: demostrar la consistencia de la aritmética a partir de ella misma.

Teorema 6.3 (Segundo Teorema de Incompletitud)**.** *La axiomática de Peano no puede probar su propia consistencia.*

Demostración. Por el lema de diagonalización, existe σ tal que

$$\sigma \Leftrightarrow_{\mathbf{AP}} \neg dem(\ulcorner \sigma \urcorner)$$

Asumimos que $\mathbf{AP} \vdash \neg dem(\ulcorner \sigma \wedge \neg\sigma \urcorner)$, es decir, que **AP** puede demostrar que no existe una demostración de una contradicción. Para llegar a una contradicción, veamos que por la Proposición 6.2

$$\begin{array}{rll} dem(\ulcorner \sigma \wedge \neg\sigma \urcorner) & \Leftrightarrow_{\mathbf{AP}} & dem(\ulcorner \sigma \urcorner) \wedge dem(\ulcorner \neg\sigma \urcorner) \\ & \Leftrightarrow_{\mathbf{AP}} & dem(\ulcorner \sigma \urcorner) \wedge dem(\ulcorner dem(\ulcorner \sigma \urcorner) \urcorner) \end{array}$$

Usando D_2 y la Tautología 2.2 obtenemos

$$\mathbf{AP} \vdash dem(\ulcorner \sigma \urcorner) \rightarrow (dem(\ulcorner \sigma \urcorner) \wedge dem(\ulcorner dem(\ulcorner \sigma \urcorner) \urcorner))$$

que junto a L_3 demuestra

$$dem(\ulcorner \sigma \urcorner) \Leftrightarrow_{\mathbf{AP}} dem(\ulcorner \sigma \urcorner) \wedge dem(\ulcorner dem(\ulcorner \sigma \urcorner) \urcorner)$$

De esta forma, volviendo a añadir la negación, obtenemos

$$\neg dem(\ulcorner \sigma \wedge \neg\sigma \urcorner) \Leftrightarrow_{\mathbf{AP}} \neg dem(\ulcorner \sigma \urcorner) \Leftrightarrow_{\mathbf{AP}} \sigma.$$

Por lo tanto, lo que hemos asumido implica que $\mathbf{AP} \vdash \sigma$. Hemos demostrado en el primer teorema de incompletitud que $\mathbf{AP} \nvdash \sigma$. Absurdo. □

7. Consecuencias de los Teoremas

La objeción más inmediata que podría surgir es que la axiomática de Peano no es la más adecuada para describir el modelo $\mathbb{N}$ y deberíamos buscar otra axiomática más adecuada. Por ejemplo, podemos definir el mismo lenguaje de Peano, el conjunto de axiomas

$$\{\varphi \mid \mathbb{N} \vDash \varphi\}.$$

Este conjunto será consistente, pues por construcción $\mathbb{N}$ es un modelo y es completa, pues para cada sentencia σ, se tiene que $\mathbb{N} \vDash \sigma$ o $\mathbb{N} \vDash \neg\sigma$. En última instancia, lo que afirman los teoremas de incompletitud, es que la pertenencia a este conjunto va a ser indecidible.

Más generalmente, supongamos que tenemos un lenguaje de primer orden $\mathscr{L}$ tal que:

Sea un lenguaje numerable, el conjunto de constantes, símbolos de funciones y símbolos de relaciones es, a lo sumo, numerable. Esto es razonable, pues el propio metalenguaje solo puede consistir en una cantidad finita de

símbolos. Sea T una teoría, en este lenguaje, que permita interpretar la aritmética de Peano. Es decir existe una relación unaria, definible en $\mathscr{L}$, $N(x)$ que pretenda expresar x *es un número natural*. Una fórmula $cero(x)$, de manera que $T \vdash \exists! x(N(x) \wedge cero(x))$ y añadimos una constante C que verifique $N(C) \wedge cero(C)$. Además, hay fórmulas funcionales s_T, $+_T$, $\cdot_T$ de manera que los axiomas de **AP** sean teoremas en T cuando sólo se refieran a números naturales. Por ejemplo

TAP_0: $T \vdash \neg\exists x(N(x) \wedge s_T(x) = C)$,

TAP_1: $T \vdash \forall x \forall y((N(x) \wedge N(y) \wedge s_T(x) = s_T(y)) \to x = y)$,

TAP_2: $T \vdash \forall x(N(x) \to x +_T C = x)$.

Además, T debe ser **recursivamente axiomatizable**. Es decir, existe un algoritmo que, en tiempo finito, decida si una fórmula φ en L es un axioma de T o no. Equivalentemente, por el Teorema de Craig [Cra53], que el conjunto de teoremas demostrables en T sea un conjunto recursivamente enumerable. Un conjunto de fórmulas F en $\mathscr{L}$ es recursivamente enumerable si existe un algoritmo que, ejecutándose de manera indefinida, devuelve, sucesivamente, los elementos de F.

En estas condiciones, podemos crear un código para las fórmulas en $\mathscr{L}$ y crear fórmulas $ax_T(f)$ similares a las relaciones que afirman "m es el número de Gödel de un axioma de **AP**" construidas en 5.2 que permitan expresar, f es el número de Gödel de un axioma de T. A partir de aquí, siguiendo la misma estrategia que hemos visto, podemos probar

Teorema 7.1 (Primer Teorema de Incompletitud de Gödel Generalizado)**.** *Sea T una teoría que interprete la aritmética de Peano. Entonces T no puede ser, a la vez, recursivamente axiomatizable, consistente y completa.*

En particular, el conjunto

$$\{\varphi \mid \mathbb{N} \vDash \varphi\}$$

no es recursivamente enumerable y no puede ser recursivamente axiomatizable.

El siguiente teorema nos dice que este conjunto de *verdades en* $\mathbb{N}$ ni siquiera es expresable en PA. Más concretamente, con nuestro código, podemos definir el subconjunto de $\mathbb{N}$

$$\{\#\varphi \mid \mathbb{N} \vDash \varphi\}$$

Teorema 7.2 (Teorema de indefinibilidad de la verdad de Tarski)**.** *No existe una $\mathscr{L}_{\boldsymbol{AP}}$-fórmula $verdad(x)$ con una variable libre x tal que, para cada fórmula φ*

$$\mathbb{N} \vDash verdad(\#\varphi) \leftrightarrow \varphi.$$

Demostración. Si existe dicha fórmula *verdad*, tomamos por el Lema de Diagonalización una $\mathscr{L}_{\mathbf{AP}}$ sentencia σ con

$$\mathbf{AP} \vdash \sigma \leftrightarrow \neg verdad(\ulcorner \sigma \urcorner).$$

Y por lo tanto, usando la definicion de *verdad*,

$$\mathbb{N} \vDash verdad(\#\sigma) \text{ si y sólo si } \mathbb{N} \vDash \sigma,$$

o, lo que es lo mismo,

$$\mathbb{N} \vDash verdad(\#\sigma) \text{ si y sólo si } \mathbb{N} \vDash \neg verdad(\#\sigma).$$

□

Este teorema responde a la pregunta de Leibniz. Hemos demostrado la imposibilidad de reducir la razón a la lógica, en tanto que ésta tiene un contenido semántico irrescatable de paradojas como ésta en el nivel sintáctico. Esto no quiere decir que hayamos separado absolutamente ambas disciplinas, pero sí que un proyecto tan ambicioso como identificarlas no sea posible, como era de esperar.

Referencias

[Cra53] William Craig. "On axiomatizability within a system". En: *J. Symbolic Logic* 18 (1953), págs. 30-32. ISSN: 0022-4812. DOI: 10.2307/2266324. URL: https://doi.org/10.2307/2266324.

[Des37] René Descartes. *Discours de la méthode*. 1637.

[Göd31] Kurt Gödel. "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I". En: *Monatshefte für Mathematik* (1931).

[HK20] Lorenz Halbeisen y Regula Krapf. *Gödel's Theorems and Zermelo's Axioms, A Firm Foundation of Mathematics*. Birkhäuser, 2020. ISBN: 9783030522780.

[LI] Bernard Linsky y Andrew David Irvine. *Principia Mathematica*. URL: https://plato.stanford.edu/entries/principia-mathematica/.

[WR13] Alfred North Whitehead y Bertrand Rusell. *Principia Mathematica*. Cambridge University Press, 1913.

[Zac] Richard Zach. *Hilbert's Program*. URL: https://plato.stanford.edu/entries/hilbert-program/.