UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
FACULTAD DE FILOSOFÍA Y LETRAS
COLEGIO DE FILOSOFÍA

EL PRINCIPIO DE KREISEL: EL CASO DE LA LÓGICA DE SEGUNDO ORDEN

T E S I S
QUE PARA OPTAR POR EL GRADO DE
LICENCIADO EN FILOSOFÍA

PRESENTA:

MANUEL EDUARDO TAPIA NAVARRO

TUTOR: DR. LUIS ESTRADA GONZÁLEZ

CIUDAD UNIVERSITARIA, CIUDAD DE MÉXICO ABRIL DE 2016

Índice general

A mis padres, por todo su apoyo y cariño

a todos mis hermanos, porque gracias a ellos no perdí la consciencia

y a Natalia, por iluminarme la vida

Agradecimientos

Quiero agradecer en primer lugar a toda mi familia (en el sentido amplio de la palabra). A mis padres por todo el cariño y paciencia que me han tenido toda la vida. Les agradezco todo el apoyo que me brindaron durante la redacción y elaboración de esta tesis. Les debo mi formación y sólo gracias a ellos no soy la persona despreciable que tal vez siempre he tenido la tendencia a ser. A mis hermanas, Mariana y Paulina por darme fuerzas y felicidad en los ratos que pudieron regalarme: las salidas al cine, a tomar café, los minutos de cigarro, etc. Sin ustedes, no sé cómo hubiese podido terminar esta tesis. Le agradezco a Dominic por soportarme a mí y a mí neurosis. Creo que es ella fue la que más me sufrió en mis momentos de depresión y cansancio mental. Te amo, Dominic; gracias por acompañarme en el horrible proceso de escribir una tesis.

Quiero agradecer también a mis profesores de la licenciatura. Especialmente, quiero agradecer al profesor Cristian Gutiérrez, quien me introdujo y formó en el estudio de la lógica. Cristian me enseñó la gran mayoría de lo que sé de lógica y me contagio el gusto por esta maravillosa disciplina, y por eso le estaré siempre agradecido. También quiero agradecer a Luis Estrada, mi tutor en la elaboración de esta tesis, por toda su paciencia y disposición. Agradezco también a mis lectores Mario Gómez Torrente, Carlos Romero y Javier García Salcedo por sus comentarios y por tomarse el tiempo para leer esta tesis, a pesar de la premura.

Quiero agradecer a todos mis hermanos por regalarme paz, tranquilidad, tiempo y cariño. Gracias a todos ellos pude tener la tranquilidad suficiente para acabar esta tesis, a pesar de la inseguridad, la impotencia, el enojo y la tristeza que experimente durante este proceso. Agradezco particularmente a Jorge Serrano, Genaro Wong y Jazmín Cruz por el tiempo que me regalaron y las cosas que me han enseñado (que son más apreciables que cualquier sistema lógico o matemático). El tiempo que pasé con ustedes no lo cambiaría (y de hecho nunca lo hice) por nada del mundo, gracias por no dejarme caer.

También quiero agradecer a todos los compañeros y amigos que me acompañaron durante mi formación académica, quienes fueron muy importantes para que pudiese mejorar (en la medida de lo posible) como filósofo y lógico. En particular, quiero agradecer a Cesar Sánchez de la Rosa, Jorge Serrano, Rodrigo Valencia, Daniel Garibay, Samuel Lomelí, Jesús Granados, Francisco Aviña, María Martínez, Carlos Cesar Jiménez, David Valencia y Cesar Escobedo, por sus comentarios a mi trabajo y discusión filosófica.

Quiero dar las gracias al Programa de Estudiantes Asociados del Instituto de Investigaciones Filosóficas por facilitarme recursos para poder realizar esta investigación. También quiero agradecer al proyecto de investigación PIFFyL 2013 012

Fundamentos de las matemáticas (y su coordinador Cristian Gutiérrez) por permitirme participar en sus seminarios y sus actividades académicas, las cuales fueron muy importantes para mi formación. También quiero agradecer al proyecto PAPIIT IA401015 Tras las consecuencias. Una visión universalista de la lógica (parte I) (y a su coordinador Luis Estrada) por proporcionarme recursos académicos y formativos. Esta tesis fue escrita principalmente durante mi participación en este proyecto, así que agradezco al proyecto pues sin su apoyo no hubiese sido posible realizar esta tesis.

Finalmente, quiero agradecer a todo aquel me haya permitido servirle como asesor. No me cabe la menor duda de que lo mejor que pude hacer en mi licenciatura fue asesorar y acercar a los estudiantes a la lógica. Mis asesorados fueron los que más enseñaron sobre la lógica y algunos de ellos ahora son de mis más cercanos amigos. Gracias a ellos y a la motivación que me dieron fue que decidí realizar mi tesis de licenciatura en lógica. Les agradezco su entusiasmo y que me hayan permitido compartir con ustedes el gusto que tengo por la lógica.

Introducción

Uno de los objetivos más importantes de la lógica es caracterizar la relación de consecuencia entre las premisas y la conclusión de un argumento. Suele considerarse que una tarea importante en la lógica es evaluar argumentos, con el fin de saber si estos son "lógicamente" correctos o no, es decir, si las premisas pueden dar un apoyo adecuado a la conclusión, o no. Es por esto que es de suma importancia caracterizar la relación de consecuencia, pues la relación de consecuencia es un criterio necesario y suficiente para que un argumento sea correcto; si las premisas de un argumento ofrecen un apoyo adecuado a la conclusión, entonces parece obvio que dicho argumento es un argumento correcto. Simplemente, el tener una caracterización adecuada de la relación de consecuencia nos permite distinguir entre cuáles argumentos son lógicamente correctos y cuáles otros no lo son.

Partiendo de esto, se presentan las cuestiones de cómo garantizamos que una caracterización de la relación de consecuencia es adecuada y qué requisitos debe satisfacer una caracterización de la relación de consecuencia lógica Es común considerar que contamos con una noción *intuitiva* o *preteórica* de *consecuencia lógica*. Consideremos los siguientes argumentos:

(1) Si los filósofos son perezosos, no existe justificación alguna para su práctica. Sabemos que los filósofos son perezosos. Por lo tanto, no existe justificación para la práctica del filósofo.

(2) Existen filósofos perezosos que no saben lógica. Por lo tanto, existen filósofos perezosos que son incompetentes.

Solemos considerar *preteóricamente* o *intuitivamente* (a partir de ahora, usaré principalmente el término 'preteórico') que (1) es un argumento correcto mientras que (2) no lo es. Parece que no necesitamos apelar a propiedades relativas a una teoría lógica particular de la noción de consecuencia lógica para evaluar estos argumentos de esta manera. Una forma de considerar que una caracterización de consecuencia lógica es adecuada es afirmando que evalúa a todos los argumentos preteóricamente correctos como tales, y como incorrectos a los que no lo son; es decir, que si tenemos un argumento preteóricamente correcto, nuestra caracterización de consecuencia lógica lo evaluará como un argumento correcto, y si nuestro argumento es preteóricamente incorrecto, nuestra caracterización lo evaluará como incorrecto. Para mayor simplicidad, determina como argumento correcto únicamente y exclusivamente a los argumentos preteóricamente correctos. Si esto es así, decimos que nuestra caracterización de consecuencia lógica es *extensionalmente adecuada*:

(EXT) Una caracterización de la relación de consecuencia lógica es *extensionalmente adecuada* si evalúa como argumentos correctos única y exclusivamente aquellos argumentos *preteóricamente* correctos

En contraste con esta noción de adecuación, podemos no poner tanta atención en cuáles son los argumentos evaluados como correctos, sino en las propiedades de la noción de consecuencia lógica. Intuitivamente, cuando comparamos los argumentos (1) y (2), podemos notar que (1) parece tener ciertas características que (2) no tiene. Una forma, diferente a la anterior, de considerar una que caracterización de consecuencia es adecuada es garantizando que rescata las propiedades de la relación de consecuencia.

(INT) Una caracterización de la relación de consecuencia es *intensionalmente adecuada* si rescata las propiedades de la relación *preteórica* de consecuencia lógica

Se podría pensar que si una caracterización de consecuencia satisface (INT), debe satisfacer (EXT). Supongamos que postulamos una caracterización de consecuencia lógica, C, tal que satisface (INT) pero no (EXT). Es decir, hay al menos un argumento, digamos A, que tiene todas las propiedades que consideramos que la relación de consecuencia lógica satisface pero que no es un caso de argumento *preteóricamente* correcto. En este caso, parece natural considerar que si A no es efectivamente un caso de argumento correcto, debe de ser así porque no satisface alguna de las propiedades de la relación de consecuencia lógica, por lo cual, nuestra teoría C, en realidad no cumpliría con (INT). Por supuesto, la conversa intuitivamente no se sostiene, es decir, si C satisficiera (EXT), no necesariamente satisface (INT), pues puede evaluar todos sus argumentos preteóricamente correctos como correctos, por algún tipo de casualidad y por algún aspecto poco relevante para la noción de consecuencia lógica. Así pues, es perfectamente posible que una teoría no cumpla con (INT) pero sí con (EXT).[1]

Alfred Tarski en (Tarski 1936) presentó una caracterización de la relación de consecuencia lógica que presuntamente rescata las propiedades que preteóricamente se le atribuye a esta relación, apelando a nociones semánticas de los lenguajes formales. La teoría tarskiana de consecuencia lógica (CLT) es como sigue:

(CLT) Sea L un lenguaje cualquiera. Primero, debemos sustituir todas las constantes extra-lógicas de las oraciones de L por variables de la correspondiente

[1]No considero que el hecho de que una teoría cumpla con (EXT) pero no con (INT) sea necesariamente negativo. Creo muy plausible que existan contextos en los cuales las propiedades de la relación de consecuencia sean irrelevantes para la práctica. En concreto, Etchemendy (1990, cap. 8) argumenta que la definición tarskiana de consecuencia lógica cumple (EXT) para el caso de los lenguaje de primer orden por hechos extralógicos y que la defnción tarskiana no cumple (INT). Sin embargo, creo que el hecho de que la definición tarskian no cumpla (INT) (suponiendo que Etchemendy tenga razón) no significa que la definición tarskiana no sea útil. Es posible que busquemos evaluar el valor de verdad de la oración G de Gödel en la aritmética de Peano (PA). Intuitivamente, G es consecuencia lógica de PA, y podemos verificar que la defnción tarskiana de consecuencia lógica la evalúa de esta manera. Me parece que no se rechazaría el uso de la definición tarskiana, pues en este caso, el interés principal es mostrar formalmente que la oración G es consecuencia lógica de PA, lo cual es posible hacer mediante la definición tarskiana.

> categoría sintáctica, obteniendo un conjunto L' de funciones oracionales. Sea X un elemento de L' y K una clase de oraciones que pertenecen a L'.
>
> Decimos que X es una consecuencia lógica de K si y solo si todos los modelos de las oraciones de K son modelos de X (Tarski 1936: 417)[2]

Es fácil ver cómo esta propuesta da una caracterización de la noción de *verdad lógica*, como una instancia de la relación de consecuencia lógica. Simplemente, se considera que las verdades lógicas son aquellas que son verdaderas en todos los modelos. En este caso, las verdades lógicas serán consecuencias lógicas de la clase vacía de enunciados, es decir, donde K es vacío. Tradicionalmente, se ha considerado que (CLT) es *adecuada intensionalmente*, es decir, que cumple con (INT), y de la misma forma que cumple con (EXT), por lo mencionado en el párrafo anterior.

Como mencionamos, Tarski hace uso de ciertas propiedades semánticas para postular su caracterización de la relación de consecuencia lógica. Estas nociones incluyen las nociones de *verdad*, *satisfacción* y *modelo*, y fueron definidas por Tarski en (Tarski 1931). En particular, la noción de modelo, que es la que Tarski utiliza explícitamente en su caracterización, está definida a partir de las anteriores: para una oración φ, M es un modelo de φ si φ es verdadera en M, donde verdad se define en términos de la definición de *satisfacción*. Tradicionalmente, se ha considerado que los modelos de un lenguaje formal pueden ser entendidos como conjuntos. Esto es, que los modelos y las interpretaciones de un lenguaje formal pueden ser entendidos como objetos de la Teoría de Conjuntos.[3] Igualmente, se ha considerado que la noción de verdad para un lenguaje formal puede ser definida dentro de la Teoría de Conjuntos. En breve, las definiciones semánticas de Tarski se han presentado tradicionalmente dentro del lenguaje de la Teoría de Conjuntos, en consecuencia, la definición de consecuencia lógica para un sistema lógico se ha dado en términos de ésta. En realidad, es bastante simple la transformación: si consideramos que los modelos de un lenguaje formal son conjuntos, nuestra caracterización se parafrasea en:

(CLTC) Decimos que X es una consecuencia lógica de K si y solo si todos las estructuras conjuntistas que satisfacen a las oraciones de K son estructuras conjuntistas que satisfacen a X.

Es bastante claro que con este ajuste también es posible dar cuenta de la noción de verdad lógica. Las verdades lógicas serán aquellas que sean verdaderas en toda estructura conjuntista.

Ahora bien, es posible preguntarse si efectivamente (CLT) es adecuada, en particular, si es adecuada extensionalmente (si satisface (EXT)). John Etchemendy ha

[2]Esta traducción es propia. En este trabajo, gran parte de los textos a los cuales refiero son textos en inglés, por lo cual, (y por motivos de simplicidad y brevedad) deberá asumirse que todas las traducciones de fragmentos de textos en inglés son propias, y se invita al lector a confrontar con los textos originales. Indicaré mediante una nota al pie de página si algún énfasis incluido en una cita textual es propia o del autor.

[3]La teoría de conjuntos que usaré en este trabajo es la axiomatización de Zermelo-Fraenkel-Axioma de elección, generalmente conocida como ZFC. Regularmente me referiré explícitamente a esta teoría, aunque en caso contrario se debe considerar que me refiero a ésta, y no a alguna otra axiomatización. Puede verse un desarrollo de esta teoría en el Apéndice.

argumentado en contra de (TLC). En (Etchemendy, 1990), Etchemendy realizó una serie de críticas a la teoría tarskiana de consecuencia, defendiendo que no cumple (INT) ni (EXT). En su libro, Etchemendy defendió primeramente que la teoría de consecuencia lógica dada por Tarski era incorrecta en el sentido de que no rescataba las propiedades de la relación preteórica de consecuencia lógica, es decir, que (CLT) no era adecuada intensionalmente. A partir de esto, Etchemendy concluye que no se puede asegurar que la teoría sea adecuada extensionalmente. De hecho, en un segundo punto, Etchemendy presentó y defendió una serie de contraejemplos a (EXT) para la definición de consecuencia lógica tarskiana. La conclusión principal de Etchemendy es que (CLT) no es una buena caracterización de la noción de consecuencia lógica. En la literatura contemporánea se han presentado varias respuestas a los argumentos de Etchemendy (véanse por ejemplo, Ray (1996), Sher (1996) y Gómez Torrente (1996, 1998, 1998/99, 2000a, 2008a, 2009). Por otro lado, algunos han defendido posturas similares a las de Etchemendy, tales como Bates (1999), Mancosu (2010) y García Carpintero (2003). El propio Etchemendy responde a algunas críticas en Etchemendy (2008). Hay una gran discusión respecto a las críticas expuestas por Etchemendy, y a partir de su trabajo se han planteado varias objeciones a (CLT). En este trabajo no trataré de ofrecer nuevas respuestas a las críticas de Etchemendy, sin embargo, vale la pena notar que en la literatura reciente, la adecuación de (CLT) ha sido puesta a prueba constantemente.

Por otra parte, Kreisel (1967: 152) presentó un principio informal que afirma precisamente la adecuación extensional de (CLT) (para el caso de las verdades lógicas) y un pequeño argumento que pretende garantizar su verdad, al menos para los lenguajes de primer orden.[4] En realidad el argumento de Kreisel está planteado para la caracterización (CLTC), más que para (CLT).[5] El argumento es muy sencillo y para explicarlo hay que introducir tres predicados. Sea α^i donde α es una oración e i es el orden de cuantificación del lenguaje en que está expresada la oración, por ejemplo, α^1 indica que la oración está expresada en un lenguaje de primer orden. Definimos los siguientes predicados:

- Val(α^i) es la afirmación de que α es intuitivamente válida (o bien, que es verdadera bajo cualquier interpretación)
- V(α^i) es la afirmación de que α es verdadera en cualquier interpretación cuyo dominio es un conjunto (o bien, que α es una verdad lógica tarskiana en el sentido (CLTC))

[4]La principal característica de los lenguajes de primer orden es que el rango de sus variables son únicamente los objetos del dominio de cuantificación. En contraste, los lenguajes de segundo orden contienen, además de las variables de primer orden, incluyen variables cuyo rango no son los objetos del dominio, sino las propiedades (o subconjuntos) de los objetos del dominio.

[5]La diferencia entre ambas nociones es más problemática de lo que parece. En (Gómez Torrente 2000a) se argumenta que (CLT) satisface la condición F presentada por Tarski: “Si, en las oraciones de una clase K y en una oración X, las constantes – aparte de las constantes puramente lógicas – son remplazadas por cualesquiera otras constantes [...], y si denotamos la clase obtenida de K como 'K”, y la oración obtenida de X como ‘X”, entonces la oración X' debe ser verdadera con la única condición de que todas las oraciones de la clase K' son verdaderas” (Tarski 1936: 415). Así mismo, se argumenta también que (CLT) cumple otra noción menos fuerte de formalidad (la que llamaré (Form) en el capítulo 2), sin embargo, (CLTC) sólo satisface ésta última. En este trabajo me concentraré en (CLTC).

- $D(\alpha^i)$ es significa que α es demostrable en una sistema formal adecuado de reglas de inferencia.

El **Principio de Kreisel (PK)** es la afirmación de que el conjunto de verdades lógicas intuitivas y el conjunto de verdades lógicas tarskianas son coextensivos:

(PK) $\forall i \forall \alpha(\mathrm{Val}(\alpha^i) \leftrightarrow \mathrm{V}(\alpha^i))$

Otra forma de entender esto es que la fórmula anterior significa que las oraciones que son verdades lógicas según (CLTC) son única y exclusivamente las oraciones que efectivamente son verdades lógicas. El argumento que Kreisel presenta para mostrar que $\forall \alpha(\mathrm{Val}(\alpha^1) \leftrightarrow \mathrm{V}(\alpha^1))$ es el caso (es decir, que (PK) se sostiene en primer orden) es el siguiente:

(P1) $\forall i \forall \alpha(\mathrm{Val}(\alpha^i) \rightarrow \mathrm{V}(\alpha^i))$ [esto se acepta "cuando uno toma como garantía que la lógica se aplica a estructuras matemáticas" (cf. Kreisel 1967: 154)

(P2) $\forall i \forall \alpha(\mathrm{D}(\alpha^i) \rightarrow \mathrm{Val}(\alpha^i))$ [se sostiene pues "uno reconoce la validez de las reglas de Frege" (cf. Kreisel 1967: 153)]

(P3) $\forall \alpha(\mathrm{V}(\alpha^1) \rightarrow \mathrm{D}(\alpha^1))$[Teorema de compleción para la lógica de primer orden]

Es fácil ver que a partir de estas premisas se concluye que el Principio de Kreisel para el caso de la lógica de primer orden es el caso:

(PK1) $\forall \alpha(\mathrm{Val}(\alpha^1) \leftrightarrow \mathrm{V}(\alpha^1))$

Vale la pena mencionar que este argumento ha traído algunas reacciones sobre su corrección o legitimidad. No es generalmente aceptado, ni tampoco es claro que se sostenga una justificación uniforme para las premisas. De hecho, se han presentado versiones de este argumento con premisas diferentes entre ellos y con respecto al original. Actualmente, se pueden revisar estas diferentes versiones en Etchemendy (1990), Gómez Torrente (1998/1999, 2000b, 2008b), Shapiro (1987, 1991), Smith (2011, 2013) y Griffiths (2014).

Por supuesto, a excepción de (P3), que es un teorema formal, Kreisel ofrece justificaciones filosóficas a sus premisas. Después de sus respectivas justificaciones, es fácil ver que en efecto se puede concluir que $\forall \alpha(\mathrm{Val}(\alpha^1) \leftrightarrow \mathrm{V}(\alpha^1))$ es el caso (cf., Kreisel 1967: 154). Sin embargo, es importante notar que, dado que el argumento apela al teorema de compleción para la lógica de primer orden, este argumento no se puede usar para demostrar que (PK) se cumple para la lógica de segundo orden (ni de ningún lenguaje de orden superior), ya que no cumple con el teorema de compleción. Kreisel afirma que no se puede utilizar la prueba anterior para demostrar que (PK) se sostiene para el caso de segundo orden, pero que se debe esperar una (cf. Kreisel 1967: 157).

El argumento de Kreisel no proporciona una justificación de (PK) para el caso de la lógica de segundo orden, lo cual no es un problema menor. Shapiro plantea el problema de manera bastante clara:

> Supongamos, sin embargo, que el principio de Kreisel no se sostiene para el lenguaje en consideración. Esto es, asumamos que hay una oración que es verdadera bajo toda interpretación conjuntista, pero que es falsa en alguna otra interpretación (digamos una interpretación cuyo dominio es una clase). Entonces la semántica formal implica que esta oración es una verdad lógica, pero que es incorrecta. En otras palabras, nuestro uso de la jerarquía acumulativa como parte esencial de la semántica para los lenguajes formales presupone que para cualquier tipo de interpretación, hay una interpretación conjuntista que es equivalente a ella (en algún sentido apropiado de "equivalente"). Esto es, en efecto, el principio de Kreisel. Así, la plausibilidad de la semántica estándar está en juego aquí (Shapiro 1987: 310)

En caso de ser falso el principio de Kreisel, sería una muestra de que el uso de la Teoría de Conjuntos es inadecuado para el caso de la lógica de segundo orden.

Ahora bien, existen dentro de la práctica matemática algunos conceptos que son de suma importancia para ésta; tales como *buen orden*, *finitud*, *buena fundación* y *clausura minimal*. Por ejemplo, la noción de buen orden es fundamental para probar y utilizar el principio de inducción sobre los números naturales. Estos conceptos pueden ser caracterizados de forma adecuada utilizando las herramientas que proporcionan los lenguajes de segundo orden.[6] Es decir, continuando con el ejemplo anterior, podemos construir una fórmula de segundo orden que sea verdadera de cierta estructura si y solo si ésta es un buen orden. Sin embargo, no es posible garantizar que esto es verdad utilizando un lenguaje de primer orden. Igualmente, la lógica de segundo orden es capaz de caracterizar la estructura de los números naturales y la estructura de los números reales.

Debido a la importancia que estos conceptos tienen dentro de la práctica matemática, son de suma importancia en estudios fundacionales en matemáticas. Realizar estudios fundacionales en matemáticas consiste en reconstruir conceptos y estructuras matemáticas con el fin de poder garantizar que las teorías matemáticas satisfacen ciertas propiedades, por ejemplo, la consistencia.[7] Así mismo, se busca reconstruir las relaciones que pueden existir entre las diferentes teorías matemáticas. La lógica de segundo orden parece ser un buen candidato para realizar estudios fundacionales en matemáticas, dada su capacidad para reconstruir las estructuras matemáticas, así como para reconstruir conceptos matemáticos.

En la lógica matemática contemporánea, la teoría de consecuencia lógica de Tarski es considerada como la teoría más importante y la más aceptada dentro de la comunidad lógica. Es rara la ocasión en que un teórico de la lógica utiliza o postula una teoría diferente a la presentada por Tarski en (Tarski 1936) (aunque, como mencioné más arriba, no es unánimemente aceptada). La lógica estándar de segundo orden utiliza la teoría de consecuencia lógica tarskiana para definir consecuencia en este lenguaje. Si el Principio de Kreisel fuese falso, entonces parecería que no

[6]Shapiro desarrolla y argumenta cuidadosamente a favor de que estas nociones no pueden ser correctamente formalizadas con un lenguaje de primer orden, véase (Shapiro 1991, cáp.5)

[7]En (Shapiro 1991), Shapiro defiende que los estudios fundacionales no tienen que consistir en proporcionar certeza epistémica a las teorías matemáticas, y que existen otros objetivos que son independientes a este proyecto.

podemos probar que las teorías matemáticas cumplen con las propiedades que pretendemos que cumplan. Es decir, el hecho de haber probado la equivalencia entre el principio del buen orden y el principio de inducción, por ejemplo, no sería significativo. Esto porque es posible que este resultado sea incorrecto, y la prueba sea producto de la inadecuación de la teoría tarskiana de consecuencia. Es por esto que dar una respuesta afirmativa a la cuestión sobre la verdad del Principio de Kreisel es de suma importancia, al menos, para poder sostener un proyecto relacionado con estudios fundacionales.

Mi interés en el presente trabajo es estudiar el Principio de Kreisel para el caso de la lógica de segundo orden y me ocuparé para ello de la noción (CLTC) en lugar de (CLT). Anteriormente mencioné que este principio afirma la adecuación extensional (para el caso de las verdades lógicas) de la caracterización tarskiana de consecuencia lógica. Es en esta lectura donde (PK) puede garantizarnos que una prueba como la de la equivalencia entre el principio del buen orden y el principio de inducción es correcta, en tanto que (CLTC) sería una buena caracterización de consecuencia lógica para el caso de la lógica de segundo orden. En términos de Shapiro, la falsedad de (PK) implicaría que es posible que una prueba de este tipo no sea suficiente, pues sería posible (en principio) que la invalidez de una fórmula relevante para la matemática clásica depende de su falsedad en alguna interpretación no conjuntista. Así, la pregunta que enfrenta mi investigación es si es posible aceptar el Principio de Kreisel para el caso de la lógica de segundo orden. Defenderé en este trabajo una respuesta afirmativa a esta pregunta, aunque habrá ciertas limitaciones de mis argumentos que tienen relación con la distinción entre las nociones de *consecuencia lógica* y la de *verdad lógica*. El principal atractivo de abordar la adecuación de (CLTC) de esta manera es que (PK) se puede entender de dos formas: como la afirmación de que (CLTC) satisface la condición (EXT) y como la afirmación de que la Teoría de Conjuntos es una semántica adecuada para el lenguaje de la lógica de segundo orden. Esta última lectura es atractiva en tanto que proporciona una manera de entender el Principio (y el argumento) de Kreisel dentro del marco de una teoría matemática precisa como ZFC. Sin embargo, la relación entre estas dos manera de entender el principio no es obvia.

A partir de estas consideraciones, un análisis del argumento de Kreisel es necesario. El argumento proporcionado por Kreisel tiene varios problemas, la gran mayoría relacionados con la justificación filosófica de sus premisas. El objetivo principal de este análisis es esclarecer el argumento y la pretensión de fondo de éste. Así pues, trataré de responder a la siguiente pregunta: ¿El argumento de Kreisel realmente establece la adecuación extensional de (CLTC)? De la misma forma, pretendo analizar y defender la legitimidad de (PK), como un principio que efectivamente afirma la adecuación extensional de (PK). Una manera de responder afirmativamente esta pregunta y al mismo tiempo garantizar la relación de este argumento informal con el marco formal de la teoría de conjuntos es mediante una versión del argumento de Kreisel expuesto por Shapiro (1987, 1991) y por Gómez Torrente (2000b, 2008b). En esta versión del argumento, tanto Shapiro como Gómez Torrente equiparán el significado del predicado $\mathrm{Val}(\alpha^i)$ con una noción más precisa como es la de *validez en estructuras cuyo dominio es una clase*. En esta versión, es bastante sencillo establecer la conclusión deseada, y a partir de la distinción (formal y precisa) entre *clase* y *conjunto*[8], es posible tener un marco preciso en el cual analizar la cuestión de ZFC. Sin embargo, esto depende de poder relacionar la noción informal de consecuencia lógica *preteórica* (presente en (EXT)) con la *noción de validez en estructuras cuyo dominio es una clase*, lo cual es parte importante del presente trabajo. De hecho, ya que la corrección de la versión de Shapiro depende de esta relación, los argumentos a favor de la aceptabilidad de esta relación son argumentos a favor de que el Principio de Kreisel exprese la adecuación extensional de (CLTC). Finalmente, la aceptabilidad de esta relación implica que la versión del argumento de Shapiro garantiza la adecuación extensional de (CLTC).

La gran ventaja de utilizar esta versión del argumento de Kreisel es que el propio Shapiro demuestra que a partir de esta lectura es posible formalizar distintas instancias del Principio de Kreisel en el lenguaje de ZFC en segundo orden (ZFC2). Shapiro posteriormente demuestra que estas formulaciones son equivalentes a formulaciones del *Principio de Reflexión* para la teoría de conjuntos. El Principio de Reflexión para la teoría de conjuntos informalmente afirma que si una oración φ en el lenguaje de la ZFC2 es verdadera en el universo conjuntista V (la jerarquía acumulativa de la teoría de conjuntos, véase Apendice), existe un modelo cuyo dominio es un conjunto donde φ es verdadera.[9] De esta manera, la aceptabilidad del Principio de Kreisel dependerá de la aceptabilidad del Principio de Reflexión, la cual defenderé. Para ello, usaré un teorema probado por Peter Koellner (2009) en el que demuestra que los Principios de Reflexión (para un lenguaje más poderoso que el de ZFC2) son consistentes. Tanto la prueba de este teorema como las demostraciones de Shapiro requerirán de un conocimiento bastante sustancial de teoría de conjuntos, al punto que me será imposible exponer todas las nociones conjuntistas involucradas. Proporcionaré un (muy) breve apéndice con las definiciones y nociones más relevantes.

Este trabajo estará divido en cuatro capítulos. En el primer capítulo, realizo una exposición de la lógica de segundo orden y de su relación con la lógica de primer orden. El objetivo principal de esta exposición es esclarecer cuál es la diferencia a nivel semántico entre ambos lenguajes, ya que, será relevante para realizar el análisis del argumento de Kreisel. En particular, me concentro en hacer lo más explícito posible porqué la lógica de segundo orden no cumple con el teorema de compleción mientras que la de primer orden sí lo cumple. Debido a esto, es pertinente hacer una pequeña nota acerca de algunas propuestas para proporcionar una semántica de la lógica de segundo orden que cumplen con el teorema de compleción. Argumento que estas propuestas son insatisfactorias para proporcionar una adecuada semántica de la lógica de segundo orden. El punto que quiero dejar claro en esta parte es que la lógica de segundo orden es efectivamente incompleta, y no sólo relativo a un tipo especial de semántica. Un segundo objetivo de este capítulo es delimitar tanto la notación que usaré como ciertos resultados técnicos que usaré, así como sus demostraciones.

[8]Dicha distinción se basa en la demostración en ZFC de que la colección de todos los conjuntos no es un conjunto, por lo cual, se le denomina *clase propia*.

[9]No es del todo precisa esta enunciación del principio. Una versión más precisa sería la siguiente: sea φ una formula de la Teoría de Conjuntos en segundo orden. Si φ es verdadera, existe un modelo de la forma $\langle V_\kappa, \in \cap V_\kappa \times V_\kappa \rangle$ donde φ es verdadera. Ésta es una versión del principio expuesto por Rayo y Uzquiano (1999: 3).

Por ejemplo, la exposición del método de prueba del teorema de corrección tanto para la lógica de primer orden como la de segundo es importante, pues la versión del argumento de Kreisel de Shapiro apela a una prueba con el mismo método.

En el segundo capítulo, argumento que efectivamente hay una relación entre la noción *preteórica* de consecuencia lógica y la noción de *validez en estructuras cuyo dominio es una clase*. Dicha relación está establecida mediante una noción informal de consecuencia lógica presentada por Shapiro. A través de ella, es posible especificar más claramente la noción de *validez en estructuras cuyo dominio es una clase* de tal manera que será claro que satisface dos propiedades de la noción *preteórica* de consecuencia lógica: la propiedad de *formalidad* y la propiedad de *necesidad*. Retomo estas propiedades del trabajo de Gómez Torrente (2000b) principalmente, aunque el hecho de que la noción preteórica de consecuencia lógica cumple con ambas propiedades se puede encontrar también en (Asmus y Restall 2012). Previamente a esta argumentación, valdrá la pena explicar con mayor precisión qué entiendo por la noción *preteórica* de consecuencia lógica. Probablemente, el lector al final de leer esta sección le parezca que el término 'preteórico' es inadecuado, en tanto que la noción informal de Shapiro (y en general, las consideraciones del capítulo) están cargadas de teoría. La razón de usar el término 'preteórico' no es porque lo ejemplos o inferencias consideradas no depende de ninguna concepción particular sobre la lógica, sino que la validez o invalidez de dichas inferencias (aparentemente) no depende de ninguna teoría semántica, matemática o formal sobre consecuencia lógica. Es decir, la justificación de su validez no depende de ningún marco formal de evaluación. Me parece que no hay una noción *preteórica*, *informal* o *intuitiva* de consecuencia lógica independiente de cierta concepción sobre la lógica, y creo que cualquier noción de este tipo está expuesta a ser revisada constantemente y que depende de los intereses del teórico, al punto de que ciertas consideraciones formales pueden influir en las inferencias consideradas correctas. Con esto en mente, doy varios ejemplos de inferencia consideradas *preteóricamente* válidas (desde el punto de vista de este trabajo), además de tres ejemplos de teorías de consecuencia lógica que me parecen relevantes desde esta perspectiva.

En el tercer capítulo, presento un análisis del argumento de Kreisel así como las formulaciones del Principio de Kreisel y las pruebas de Shapiro de la equivalencia entre estas y los Principios de Reflexión. En este análisis, me concentro en las dificultades que presentan las justificaciones de las premisas del argumento de Kreisel y como estas se solucionan considerando la versión del argumento de Shapiro y Gómez Torrente. Como parte de este análisis presento una versión más general del Principio de Kreisel que considere la noción de consecuencia lógica (no sólo el caso de la noción de verdad lógica como en la versión original), y por qué esto es necesario. Finalmente, presento las formulaciones del Principio de Kreisel en el lenguaje de la teoría ZFC2 y su equivalencia a ciertas formulaciones del Principio de Reflexión. Para ello, expongo la necesidad de extender el lenguaje de ZFC2 a un lenguaje con la capacidad de *codificar* distintos conjuntos de fórmulas. Este capítulo requiere un conocimiento avanzado de teoría de conjuntos, en particular requiere familiaridad con las definiciones de cardinales grandes, ZFC. En general, no recomiendo la lectura de este capítulo ni del siguiente si no se tiene dicho conocimiento, así que ofrezco al lector un Apéndice de carácter con las definiciones necesarias para entender este trabajo.

En el último capítulo, expondré el teorema de Koellner que muestra la consistencia de los principios de reflexión, y su prueba. En este punto habrá un resultado negativo, pues explico por qué esta prueba no es suficiente para garantizar que los principios de reflexión que consideran el caso más general de consecuencia lógica son consistentes. Sin embargo, sí garantiza la consistencia para el caso de la noción de verdad lógica, bajo la presuposición de un cardinal de Erdös, por lo cual, presentaré algunas razones para aceptar la existencia de los cardinales de Erdös. Con todo esto en mente, argumento que es posible aceptar el Principio de Kreisel a partir de que éste está justificado, pues el Principio de Reflexión lo está. Como complemento, exploro brevemente otras alternativas para justificar el Principio de Reflexión basadas en cierta analogía que es posible hacer (en principio) entre estos y los principios de reflexión para la aritmética, basada en la modalización de los principios de reflexión.

Hay que mencionar que no pretendo defender ninguna clase de postura con respecto a los estudios fundacionales (bien lo mencioné anteriormente), aunque simpatizo bastante con proyectos como el de Shapiro (1991). Independientemente de esto, en vista de las observaciones hechas sobre las pruebas referidas, creo que es deseable dar una respuesta positiva a la cuestión de la verdad del Principio de Kreisel, pues una prueba (adecuada y correcta) de consistencia o de la equivalencia entre el principio del buen orden y el principio de inducción son deseables. A pesar de ello, creo que daré buenas razones para aceptar el Principio de Kreisel, y que estas razones son suficientes para poder argumentar a favor de tesis como la suficiencia de ZFC para proporcionar una semánticas adecuada de los lenguajes cuantificacionales (hablaré de esto al final del trabajo). Igualmente, no pretende abordar con mucho detalle cómo la lógica de segundo orden rescata ciertas nociones matemáticas. Es importante dejar en claro que tampoco pretendo garantizar que la teoría tarskiana está fuera de toda duda; me parece que no es así. Simplemente creo que la teoría tarskiana de consecuencia lógica cumple con ciertos objetivos de la práctica de la lógica, pues rescata algunas de las propiedades que uno espera de un argumento lógicamente correcto.

Capítulo 1

Lógica de Primer Orden y Lógica de Segundo Orden

El argumento de Kreisel ocupa el hecho de que la lógica de primer orden cumple con el teorema de compleción y es por eso que dicho argumento no se puede aplicar a la lógica de segundo orden, pues ésta no cumple dicho teorema. En este capítulo, expondré un lenguaje de primer orden y una extensión de éste de segundo orden; así como una semántica para el lenguaje de primer orden y una extensión de ésta para el lenguaje de segundo orden (la que llamaremos *semántica estándar*). Tomando en cuenta el argumento de Kreisel, el primer objetivo (y objetivo principal) del capítulo es explicar por qué la lógica de primer orden cumple el teorema de compleción mientras que la lógica de segundo orden no. Para poder realizar esto de manera satisfactoria, será necesario exponer otros teoremas como los teoremas de compacidad y de Lowenheim-Skolem. Esto permitirá entender con mayor precisión cuál es la diferencia entre la lógica de primer orden y la de segundo orden. Un segundo objetivo es clarificar el hecho de la lógica de segundo orden no cumple con el teorema de compleción sólo si usamos la semántica estándar. Expondremos dos semánticas más para la lógica de segundo orden: la *semántica de Henkin* y *la semántica de primer orden.* Para estas semánticas es posible demostrar el teorema de compleción. Daré algunas razones para no adoptar dichas semánticas. Al final, el objetivo del capítulo será exponer el contexto en el que entenderé el argumento de Kreisel, lo cual incluye adoptar una semántica particular.

1.1. Lógica de Primer Orden

En general, un lenguaje de primer orden es uno tal que el rango de sus variables son los objetos del dominio de una interpretación del lenguaje. Definiré un lenguaje de este tipo, *L1K=*. Este lenguaje está conformado por un conjunto de símbolos que constituyen la terminología lógica de nuestro lenguaje:

- Paréntesis: (,)
- Símbolos de variable de individuo: para cada n, número natural, x_n es un símbolo de variable (en ocasiones, para simplificar utilizaré simplemente x, y, z ...)

- Conectivas lógicas: $\supset$ y $\neg$
- Símbolo de cuantificador: $\forall$
- Símbolo de igualdad: $=$

Además de este conjunto de terminología lógica, incluiremos un conjunto K de terminología no lógica, constituido de la siguiente forma:

- símbolos de constante: para cada número natural n, a_n es un símbolo de constante (en ocasiones, usaré simplemente las letras a, b, c,..., por simplicidad);
- símbolos de predicado: para cada n, m, números naturales, P^n_m es un símbolo de predicado (en ocasiones, usaré simplemente las letras P, Q, R,..., por simplicidad);
- símbolos de función: para todo número natural n y m, f^n_m es un símbolo de función (en ocasiones usaré simplemente las letras f, g, h,..., por simplicidad).[1]

Cada conjunto de símbolos de terminología no lógica puede ser vacío, en nuestro caso no lo será e incluirá una cantidad numerable de cada tipo de símbolo. En el caso, de los símbolos de predicado (o función), el subíndice indica la aridad (o el número de argumentos) del símbolo de predicado (o función) que se esté considerando, mientras que el superíndice indica el número de símbolo de predicado (o función) que se está considerando. Las reglas de formación (también llamadas *operaciones de formación* o *construcción*) para dicho lenguaje son las reglas usuales:

Reglas de fórmación de términos

1. Las variables y las constantes son términos.
2. Si f^n_m es un símbolo de función y $t_1,\ldots,t_m$ son términos, entonces $f^n_m(t_1,\ldots,t_m)$ es un término.
3. Nada más es término.

Reglas de formación de fórmulas

1. Si t_1 y t_n son términos, entonces $(t_1 = t_n)$ es una fórmula.
2. Si $t_1,\ldots,t_m$ son términos y P^n_m es un símbolo de predicado, entonces $P^n_m(t_1,\ldots,t_m)$ es una fórmula.
3. Si φ y ψ son fórmulas, entonces $(\varphi \supset \psi)$ y $\neg\varphi$ son fórmulas.
4. Si φ es una fórmula y x_n es una variable, entonces $\forall x_n \varphi$ es una fórmula.

[1]Igualmente, tanto en el caso de las letras de predicado y las letras de funciones, omitiré los subíndices y los superíndices en caso de que no sean necesarios o en caso de que esté implícito en el contexto el valor de estos.

5. Nada más es fórmula.

El proceso de formación de fórmulas es un caso del teorema de recursión.[2] Otra peculiaridad de la manera en que definimos el lenguaje es que permitimos que existan fórmulas del lenguaje que contengan *variables libres*, es decir, una variable que no esté cuantificada por ningún cuantificador en una fórmula. Llamaré *enunciados* a todas las fórmulas que no tengan variables libres.

El lenguaje que considero no contiene más conectivas lógicas que $\supset$ y *neg*, y sólo contiene un cuantificador: $\forall$. Sin embargo, las demás conectivas usuales y el cuantificador existencial se pueden definir de la manera usual:

Conjunción $(\varphi \wedge \psi) \equiv_{def} \neg(\varphi \supset \neg\psi)$

Disyunción $(\varphi \vee \psi) \equiv_{def} (\neg\varphi \supset \psi)$

Equivalencia $(\varphi \equiv \psi) \equiv_{def} [\varphi \supset \psi) \wedge (\psi \supset \varphi)]$

Existencial $\exists x_n \varphi \equiv_{def} \neg\forall x_n \neg\varphi$

1.1.1. Semántica: verdad y modelos para el lenguaje *L1K=*

En esta sección, definiré una colección de interpretaciones para el lenguaje *L1K=*. Para ello, presupongo un contenido mínimo de la teoría de conjuntos. Llamaré a dichas interpretaciones, *estructuras*, y su función será determinar qué oraciones deberán ser entendidas como verdaderas y que otras como falsas, dependiendo de cada estructura. Después, de esto, definiré las nociones de satisfacción, verdad y *consecuencia lógica*.

Una *estructura*[3] $\mathfrak{A}$ es un par $\langle D^{\mathfrak{A}}, I^{\mathfrak{A}} \rangle$ tal que $D^{\mathfrak{A}}$ es un conjunto no vacío e $I^{\mathfrak{A}}$ es una función cuyo dominio es el conjunto K de terminología no lógica que cumple con las siguientes condiciones:

1. A cada a_n, símbolo de constante, $I^{\mathfrak{A}}$ le asigna un objeto $a_n^{\mathfrak{A}} \in D^{\mathfrak{A}}$

2. A cada P_m^n, símbolo de predicado, $I^{\mathfrak{A}}$ le asigna $P_m^{n\,\mathfrak{A}} \subseteq D^{\mathfrak{A}\,m}$ (donde $D^{\mathfrak{A}\,m}$ es el conjunto de todas las m-adas ordenadas de objetos de $D^{\mathfrak{A}}$)

3. A cada f_m^n, símbolo de función, $I^{\mathfrak{A}}$ le asigna una operación $f_m^{n\,\mathfrak{A}} : D^{\mathfrak{A}\,m} \rightarrow D^{\mathfrak{A}}$

[2]Enderton (2004) utiliza operaciones para la construcción de términos y operaciones para la construcción de fórmulas, muy parecidas a las que se acaban de enunciar. Con el uso de esas operaciones, se puede utilizar de manera muy sencilla el teorema de la recursión para enumerar las fórmulas y los términos del lenguaje.

[3]Seguiré a Enderton en su nomenclatura y utilizaré el término 'estructura' para referirme a una interpretación de nuestro lenguaje, por motivos de precisión. Cabe mencionar también que las estructuras son estructuras estrictamente de primer orden. Igualmente, usaré las letras alemanas-Fraktur ($\mathfrak{A}$,$\mathfrak{B}$,$\mathfrak{C}$,$\mathfrak{D}$,etc) para referirme a las estructuras del lenguaje. Sin embargo, la definición de estructura que usaré es diferente a la usada por Enderton (2004: 122). La razón de esto es que la definición de Enderton podría prestarse a considerar a los cuantificadores como terminología no lógica, lo cual, como señale arriba, no hago yo. Sin embargo, ambas definiciones son esencialmente equivalentes, en tanto que ambas consideran exactamente las mismas estructuras.

La idea de fondo es que $I^{\mathfrak{A}}$ asigna un objeto adecuado a cada elemento del conjunto K de terminología no lógica del lenguaje *L1K=*, a veces me referiré a este objeto como el referente de dicho término. Llamaremos a $D^{\mathfrak{A}}$, el dominio de $\mathfrak{A}$, y a $I^{\mathfrak{A}}$, la función de interpretación de $\mathfrak{A}$.

Ahora definiré qué significará que una fórmula sea verdadera para una estructura arbitraria. Primero, definiré el concepto de *satisfacción*. Sea $\mathfrak{A}$ una estructura y sea V conjunto de los símbolos de variables de *L1K=*. Definimos $\Sigma^{\mathfrak{A}}$ como el conjunto de todas las funciones, s: V $\rightarrow D^{\mathfrak{A}}$.

Intuitivamente, cada función s asigna un objeto del dominio de $\mathfrak{A}$ a cada variable del lenguaje.[4] Definiremos ahora una extensión $\bar{s}$, para cada $s \in \Sigma^{\mathfrak{A}}$, de tal forma que $\bar{s}$ proporcione significado a cada término del lenguaje (de la misma manera que s lo hace para el caso de las variables), de la siguiente forma:

1. $\bar{s}(x_n){=}s(x_n)$, para cada x_n, símbolo de variable
2. $\bar{s}(a_n){=}s(a_n^{\mathfrak{A}})$, para cada a_n, símbolo de constante
3. $\bar{s}(f_m^n(t_1,\ldots,t_m)){=}f_m^{n\,\mathfrak{A}}$ $(\bar{s}(t_1),\ldots,\bar{s}(t_m))$, para cualesquiera f_m^n, símbolo de función y cualesquieras $t_1,\ldots,t_m$ términos

La extensión $\bar{s}$ es otro caso del teorema de la recursión y es única para cada $s \in \Sigma^{\mathfrak{A}}$. Ahora es posible definir la noción de satisfacción.

Sean $\mathfrak{A}$ una estructura, φ una fórmula de *L1K=* y una $s \in \Sigma^{\mathfrak{A}}$. Simbolizamos que $\mathfrak{A}$ satisface φ con s mediante la siguiente expresión $\vDash^1_{\mathfrak{A}} \varphi[s]$. La definición es como sigue:

1. Si φ es $(t_1{=}t_2)$ entonces $\vDash^1_{\mathfrak{A}} \varphi[s]$ si y solo si $\bar{s}(t_1){=}\bar{s}(t_2)$
2. Si φ es $P_m^n(t_1,\ldots,t_m)$ entonces$\vDash^1_{\mathfrak{A}} \varphi[s]$ si y solo si$\langle\bar{s}(t_1),\ldots,\bar{s}(t_m)\rangle \in P_m^{n\,\mathfrak{A}}$
3. Si φ es $\neg\alpha$ entonces $\vDash^1_{\mathfrak{A}} \varphi[s]$si y solo si $\nvDash^1_{\mathfrak{A}} \alpha[s]$
4. Si φ es $(\alpha \supset \beta)$ entonces $\vDash^1_{\mathfrak{A}} \varphi[s]$ si y solo si, o bien, $\nvDash^1_{\mathfrak{A}} \alpha[s]$, o bien, $\vDash^1_{\mathfrak{A}} \beta[s]$
5. Si φ es $\forall x_k\alpha$ entonces $\vDash^1_{\mathfrak{A}} \varphi[s]$ si y solo si, para todo $d \in D^{\mathfrak{A}}$, $\vDash^1_{\mathfrak{A}} \alpha[s(x_k|d)]$

Donde α y β son fórmulas de *L1K=* y $s(x_k|d)$ se define de la siguiente manera:

$$s\,(x_k|d)\,(y) = \begin{cases} s\,(y) & \text{si } y \neq x_k \\ d & \text{si } y{=}x_k \end{cases}$$

Intuitivamente, la cláusula 5 afirma que una fórmula de la forma $\forall x_k\alpha$ es satisfecha por una estructura $\mathfrak{A}$ con una función s si la fórmula α es satisfecha por todos los objetos del dominio. Definida la noción de satisfacción, es posible dar una definición de las nociones de verdad y consecuencia lógica. Sean φ una fórmula de *L1K=* y $\mathfrak{A}$ una estructura, definimos

[4]Estas funciones en ocasiones son llamadas también *funciones de asignación*, *asignaciones* o *secuencias de asignación*. En ocasiones usaré esta terminología para referirme a estas funciones.

- φ es *verdadera en* $\mathfrak{A}$, ($\vDash^1_{\mathfrak{A}} \varphi$) si y solo si $\vDash^1_{\mathfrak{A}} \varphi[s]$, para toda $s \in \Sigma^{\mathfrak{A}}$
- φ es *falsa en* $\mathfrak{A}$,$\nvDash^1_{\mathfrak{A}} \varphi$) si y solo si$\nvDash^1_{\mathfrak{A}} \varphi[s]$, para toda $s \in \Sigma^{\mathfrak{A}}$
- $\mathfrak{A}$ es *modelo* de Γ si y solo si($\vDash^1_{\mathfrak{A}} \varphi$), para toda $\varphi \in \Gamma$
- Γ es *satisfacible* si y solo si($\vDash^1_{\mathfrak{A}} \varphi$), para toda $\varphi \in \Gamma$, y para alguna estructura $\mathfrak{A}$ y alguna $s \in \Sigma^{\mathfrak{A}}$

donde *Gamma* es un conjunto de fórmulas de *L1K=*. Sean Γ un conjunto de fórmulas de *L1K=* y φ una fórmula de *L1K=*. Recordando la definición de consecuencia lógica de Tarski, podríamos definir consecuencia lógica para *L1K=* de manera análoga:

> φ es una *consecuencia lógica* de Γ si y solo si toda estructura $\mathfrak{A}$ que sea modelo de Γ es modelo de $\{\varphi\}$

Esta definición es equivalente, en el caso de que φ y cada $\psi \in \Gamma$ sean enunciados, a la que será nuestra definición de consecuencia lógica:

> φ es una *consecuencia lógica* de Γ ($\Gamma \vDash^1 \varphi$) si y solo si, para toda estructura $\mathfrak{A}$ y para toda $s \in \Sigma^{\mathfrak{A}}$, si para toda $\psi \in \Gamma$, $\vDash^1_{\mathfrak{A}} \psi[s]$ entonces, $\vDash^1_{\mathfrak{A}} \psi[s]$

Es necesario aclarar que, dada la manera en que fue definido lo que es un *modelo*, esta definición no es un caso idéntico de la definición dada por Tarski (1936).[5] Sin embargo, se suele atribuir a Tarski esta definición en la literatura, pues es análoga a la definición de Tarski y utiliza esencialmente las mismas ideas. Siguiendo dicha práctica, será esta definición la que estudiaré en el presente trabajo.

Finalmente, definiremos la noción de *verdad lógica*. Sea φ una fórmula de *L1K=*:

> φ es una *verdad lógica* (o *válida*, o *universalmente válida*) $\vDash^1 \varphi$ si y solo si $\vDash^1_{\mathfrak{A}} \varphi[s]$ para toda $\mathfrak{A}$ y para toda $s \in \Sigma^{\mathfrak{A}}$

Para concluir este apartado, cabe señalar que esta definición de verdad valida las definiciones de las conectivas usuales así como las del cuantificador universal. Igualmente, es posible definir la noción de *equivalencia lógica* de manera sencilla.

[5]Para clarificar esto, la noción de *modelo* que Tarski consideró en (1936) no consideraba estructuras (o interpretaciones, para ser más preciso) que asignarán significado a la terminología no lógica. De ahí proviene la necesidad de sustituir todas las constantes no lógicas por variables del tipo adecuado. Es por esto que, en la nota 3, mencioné que considero al cuantificador universal como terminología lógica. Si no fuese así, el cuantificador universal sería una constante no lógica y podríamos intercambiarlo (en principio) por otra constante del tipo adecuado, lo cual podría invalidar la regla-ω, una de las motivaciones de Tarski para dar su definición. También se suele argumentar (por ejemplo, (Etchemendy 1990)) que la definición de Tarski no consideraba variar el dominio de cuantificación si no que consideraba un solo dominio universal, situación que claramente no sucede en nuestra definición, pues cada estructura puede tener un dominio diferente. Sin embargo, Gómez Torrente (1996) defendió la definición tarskiana mostrando que ambas concepciones del dominio de cuantificación son "prácticamente" equivalentes y que Tarski estaba al tanto de esta situación. Así que, de momento, dejaremos pasar esta última objeción.

1.1.2. Un sistema deductivo para el lenguaje *L1K=*

Ahora daré un sistema deductivo para nuestro lenguaje formal. Regularmente se utilizan sistemas axiomáticos para facilitar la demostración de los teoremas de corrección y compleción. Este sistema deductivo tendrá infinitos axiomas lógicos[6] y sólo una regla de inferencia. Primero daré una definición:

> Una fórmula φ es una *generalización* de ψ si y solo si para algún $n \geq 0$, las variables $x_1,\ldots,x_n$ son tales que para la fórmula ψ, $\varphi = \forall x_1, \ldots x_n \psi$

Axiomas

1. $\alpha \supset (\beta \supset \alpha)$
2. $(\alpha \supset (\beta \supset \gamma)) \supset ((\alpha \supset \beta) \supset (\alpha \supset \gamma))$
3. $(\neg\beta \supset \neg\alpha) \supset ((\neg\beta \supset \alpha) \supset \beta)$
4. $\forall x\alpha \supset \alpha^x_t$ [donde t se puede sustituir por x en α][7]
5. $\alpha \supset \forall x\alpha$ [donde x no ocurre libre en α]
6. $\forall x(\alpha \supset \beta) \supset (\forall x\alpha \supset \forall x\beta)$
7. $x{=}x$
8. $(x = y) \supset (\alpha \supset \alpha')$ [donde α es $(t_1 = t_n)$ o $P^n_m(t_1,\ldots,t_n)$ y α'es el resultado de remplazar x por y en cero o más lugares (aunque no necesariamente en todos)]

Regla de inferencia: Modus Ponens(MP)

$$\frac{\begin{array}{l}\alpha \supset \beta \\ \alpha\end{array}}{\beta}$$

Llamaremos a este sistema deductivo *D1=*. Para cualquier fórmula α, α^x_t es el resultado de sustituir todas las apariciones libre de x por el término t.

Definición. *Una deducción de φ a partir de Γ en D1= es una sucesión finita $\langle \alpha_0, \ldots, \alpha_n \rangle$ de fórmulas de L1K=, dónde α_n es φ y para k$\geq$n, o bien:*

a. α_k *es un axioma*

b. $\alpha_k \in \Gamma$

[6]Debido a que definí un sistema deductivo para un lenguaje de primer orden que no pertenece a ninguna teoría matemática particular, todos sus axiomas serán lógicos y no tendrá axiomas propios.

[7]La razón por la cual se incluye esta restricción es sencilla. Si α es $\neg\forall y(x = y)$ entonces $\forall x\alpha \supset \alpha^x_t$ sería $\neg\forall y(x = y) \supset \neg\forall y(y = y)$ lo cual es casi siempre falso (es posible encontrar una estructura en la cual sea falso), y buscamos que los axiomas sean verdaderos.

c. α_k se obtiene por modus ponens de dos fórmulas anteriores, α_i y α_j, $i < k$ y $j < k$, donde α_j es $(\alpha_i \supset \alpha_k)$

Si para un conjunto de fórmulas Γ (posiblemente vacío) y una fórmula φ, existe una deducción de φ a partir de Γ en *D1=*, entonces decimos que φ es deducible de Γ en *D1=* y lo simbolizaremos como $\Gamma \vdash^1 \varphi$.

Definición. *Una fórmula φ es un teorema de D1= ($\vdash^1 \varphi$) si y solo si $\Gamma \vdash^1 \varphi$ es el caso, donde Γ es $\varnothing$*

Definición. *Un conjunto de fórmulas Γ de L1K= es consistente si y solo si no hay ninguna fórmula φ de L1K= tal que $\Gamma \vdash^1 \varphi$ y $\Gamma \vdash^1 \varphi$*

Los esquemas de axioma expuestos más arriba fueron seleccionados de manera cuidadosa, atendiendo a ciertos objetivos impuestos tanto por la semántica que definimos como por la necesidad de demostrar ciertos teoremas "deseables" de la lógica de primer orden. Los primeros tres axiomas (junto con la regla de modus ponens) garantizan que todo teorema de la lógica proposicional es un teorema de nuestro sistema deductivo.[8] Intuitivamente, el esquema axioma 4 garantiza que es posible "instanciar" el cuantificador universal en cualquier objeto, ya sea que nos refiramos a él mediante una constante, una variable o una función. Los esquemas 5 y 6 son esenciales para demostrar el teorema de generalización:

Teorema 1.1. ***(Generalización)*** *Si $\Gamma \vdash^1 \varphi$ y x una variable que no ocurre en ningún elemento de Γ entonces $\Gamma \vdash^1 \forall x \varphi$*

Este teorema permite derivar fórmulas cuya conectiva es un cuantificador universal. En particular, es posible universalizar una conclusión general presentada en una fórmula con una variable libre. Este teorema rescata la "intuición" de que si una oración es verdadera de un objeto arbitrario (o es verdadera de un objeto, pero su verdad no depende de ninguna propiedad particular de éste), entonces esa oración es verdadera de todos los objetos del dominio. Finalmente, los esquemas 7 y 8 nos permiten manejar fórmulas con el símbolo de identidad. El esquema 8 rescata la misma intuición detrás de la "indiscernibilidad de los idénticos", regla generalmente aceptada.

Por último, vale la pena señalar que *D1=* cumple con el teorema de la deducción "irrestricto":[9]

[8] El sistema *D1=* está definido esencialmente del mismo modo que en (Enderton 2004). Sin embargo, Enderton no considera los primeros tres axiomas de *D1=*, si no que considera a todas las generalizaciones de las tautologías como axiomas de sus sistema deductivo. La razón para optar por poner de manera explícita estos tres axiomas es para mostrar de forma explícita que *D1=* es una extensión de los cálculos de lógica proposicional, sin apelar al teorema de compleción para ésta. De cualquier forma, ambos sistemas son esencialmente el mismo.

[9] En (Mendelson 1987) se prueba una versión más débil del teorema de la deducción. Para el sistema deductivo que propone Mendelson, esta versión más débil es lo más que se puede demostrar, pues incluye dentro de las reglas de inferencia una regla de generalización, lo cual implicaría una falacia al aplicar el teorema de la deducción (cf. Mendelson 1987: 58). Como dicha regla no está incluida en *D1=*, podemos demostrar el teorema de manera irrestricta, colocando la cláusula de que la variable x no ocurra libre en ningún elemento de Γ, debilitando el teorema de generalización.

Teorema 1.2. ***(Teorema de la Deducción)***: *Si* $\Gamma \cup \{\gamma\} \vdash^1 \varphi$ *entonces* $\Gamma \vdash^1 (\gamma \supset \varphi)$

La prueba del teorema es por inducción matemática sobre la longitud de la prueba. Dado que la única regla de inferencia es el Modus Ponens, la prueba es exactamente la misma que para el caso de la lógica proposicional. El paso base es aquel en que la prueba de $\Gamma \cup \{\gamma\} \vdash^1 \varphi$ es de longitud uno, por lo cual φ debe ser un elemento de Γ, $\varphi = \gamma$ o φ es un axioma. En cualquier caso, $\Gamma \vdash^1 (\gamma \supset \varphi)$ se sigue del axioma 1 y el axioma 2. El paso inductivo considera los mismos casos, a excepción del caso en el que φ se sigue por Modus Ponens de fórmulas anteriores, de nuevo, $\Gamma \vdash^1 (\gamma \supset \varphi)$ se sigue por el axioma 1 y por el axioma 2. Para mayor profundidad en la prueba véase (Mendelson 1987: 58) o Enderton (2004: 176)

1.1.3. Teoremas de corrección y compleción

Ahora daré esbozos de las pruebas de los dos teoremas importantes de esta sección: el teorema de corrección y el teorema de compleción. De forma simplificada, el teorema de corrección afirma que todo lo que es posible demostrar dentro de *D1=* es una consecuencia lógica en *L1K=*. Por otro lado, el teorema de compleción afirma que para todo caso de consecuencia lógica de *L1K=*, es posible encontrar una deducción en *D1=*. Algunos detalles de los teoremas, aunque interesantes, no son indispensables en lo que sigue, por lo cual, daré sólo esbozos precisos de las pruebas de ambos teoremas. Mi interés se centrará en el método de prueba de ambos teoremas.

Teorema 1.3. ***Teorema de Corrección para L1K= y D1=***:*Si* $\Gamma \vdash^1 \varphi$ *entonces* $\Gamma \vDash^1 \varphi$

Demostración. En general, se conoce el método de prueba de este teorema. Basta con demostrar que todo axioma es una verdad lógica y que la regla de modus ponens "preserva verdad" para toda interpretación. Para ser más precisos hay que demostrar el siguiente lema:

Lema 1.1. *Todo axioma es universalmente válido.*

Demostración. Verificamos cada esquema garantizando que todos son universalmente válidos.

Esquemas de axioma 1-3: Los esquemas de axioma "proposicionales" son universalmente válidos a partir de las cláusulas 3 y 4 de la definición de satisfacción. *Esquema 4*: Para demostrar que $\forall x\alpha \supset \alpha^x_t$ es universalmente válida se deben demostrar antes los lemas 1.2 y :

Lema 1.2. *Sean* $\mathfrak{A}$ *una estructura y* $s \in \Sigma^{\mathfrak{A}}$. *Sean* x *y* t *términos, para cualquier término* u, u^x_t *es el resultado de remplazar la variable* x *en* u *por* t. *Entonces*

$$\bar{s}(u^x_t)=\overline{(s(x|\bar{s}(t)))}(u)$$

El lema simplemente afirma que el referente del término u (después de la sustitución de x por t) es la misma que la referencia del mismo término después de la sustitución del valor de x bajo s por el referente de t. Esto es, es lo mismo sustituir x por t en un término que sustituir sus valores bajo s en esa misma secuencia. La demostración de este lema funciona por inducción sobre el número de funciones en el término u y se aplica para demostrar el siguiente lema:

Lema 1.3. *Si t es sustituible por x en φ, entonces*

$$\models^1_{\mathfrak{A}} \varphi[s(x|\bar{s}(t))] \text{ si y solo si } \models^1_{\mathfrak{A}} \varphi^x_t[s]$$

La demostración funciona similar a los anteriores axiomas: consideramos una estructura cualquiera $\mathfrak{A}$ y una $s \in \Sigma^{\mathfrak{A}}$, tales que $\models^1_{\mathfrak{A}} \forall x\varphi[s]$. Como esto es un universal consideramos el objeto $\bar{s}(t)$ para el cual vale el universal. Por la cláusula 5 de satisfacción, $\models^1_{\mathfrak{A}} \varphi[s(x|\bar{s}(t))]$ y por el lema 1.6, $\models^1_{\mathfrak{A}} \varphi^x_t[s]$, que era lo que queríamos mostrar.

Esquema 5: Suponemos que $\models^1_{\mathfrak{A}} \alpha[s]$ (para cualesquiera $\mathfrak{A}$ y $s \in \Sigma^{\mathfrak{A}}$). Como x no ocurre libre en α, $\models^1_{\mathfrak{A}} \alpha[s(s—d)]$, para todo $d \in D^{\mathfrak{A}}$. Por lo cual, $\models_{\mathfrak{A}^1} \forall x\alpha[s]$

Esquema 6: Al igual que en los casos anteriores, suponemos $\models^1_{\mathfrak{A}} (\alpha \supset \beta)[s]$ y $\models^1_{\mathfrak{A}} \alpha[s]$, para cualesquiera $\mathfrak{A}$ y $s \in \Sigma^{\mathfrak{A}}$. Es fácil demostrar que $\models^1_{\mathfrak{A}} \beta[s]$, a partir de las cláusulas de satisfacción 3, 4 y 5.

Esquema 7: Para cualquier variable x, $\models^1_{\mathfrak{A}} x = x[s]$ se sigue directamente de la cláusula de satisfacción 1.

Esquema 8: La demostración tiene un paso intermedio:

Lema 1.4. *Sea t un término, si t' es el resultado de sustituir la variable x por y, si $x = y$, entonces $\bar{s}(t)=\bar{s}(t')$*

La prueba requiere inducción sobre t. Después, suponemos que $\models^1_{\mathfrak{A}} x = y[s]$ y $\models^1_{\mathfrak{A}} \alpha[s]$. Simplemente, verificamos el caso en que α es $(t_1 = t_n)$ y el caso en que α es $P^n_m(t_1, \ldots, t_n)$. Se concluye fácilmente, haciendo uso del lema, que $\models^1_{\mathfrak{A}} \alpha'[s]$

Para demostrar completamente el lema **??**, falta demostrar que toda generalización de una fórmula que sea instancia de estos esquemas es una verdad lógica. Esto se sigue fácilmente del siguiente lema:

Lema 1.5. *Sean varphi una fórmula y x una variable. Para cualquier $\mathfrak{A}$, $\models^1_{\mathfrak{A}} \varphi$ si y solo si $\models^1_{\mathfrak{A}} \forall x\varphi$*

La demostración es muy sencilla. El caso complicado es el caso en que el cuantificador universal cuantifica una variable libre. Sin embargo, si $\models^1_{\mathfrak{A}} \varphi$, φ ya es satisfecha por cualquier objeto, por lo tanto $\models^1_{\mathfrak{A}} \forall x\varphi$ para cualquier $\mathfrak{A}$. □

Lema 1.6. *Para cualquier $\mathfrak{A}$, y cualquier α y β fórmulas, si $\models^1_{\mathfrak{A}} (\alpha \supset \beta)$ y $\models^1_{\mathfrak{A}} \alpha$ entonces $\models^1_{\mathfrak{A}} \beta$*

Su demostración es muy sencilla y se sigue directamente de las cláusulas de satisfacción 3 y 4.

Prueba del Teorema de Corrección: Supongamos que $\Gamma \vdash^1 \varphi$. Sea $\mathfrak{A}$ y $s \in \Sigma^{\mathfrak{A}}$ cualesquiera tales que $\models^1_{\mathfrak{A}} \psi[s]$ para toda $\psi \in \Gamma$ Hay tres casos:

Caso 1: $\varphi \in \Gamma$ Por lo tanto $\models^1_{\mathfrak{A}} \varphi[s]$

Caso 2: φ es un axioma. Por lo tanto, $\vDash^1_{\mathfrak{A}} \varphi[s]$ (por el lema 1.1)

Caso 3: φ se obtiene por modus ponens de dos fórmulas anteriores, α_i y α_j, $i < k$ y $j < k$, donde α_j es $(\alpha_i \supset \alpha_j)$. Por lo tanto, $\vDash^1_{\mathfrak{A}} \varphi[s]$ por lema 1.6

Por lo tanto, $\Gamma \vDash^1_{\mathfrak{A}} \varphi$, para toda $\mathfrak{A}$.

□

Ahora daremos un esquema de prueba del teorema de compleción para la lógica de primer orden.

Teorema 1.4. ***(Teorema de Compleción para L1K= y D1=)****: Si* $\Gamma \vDash^1 \varphi$ *entonces* $\Gamma \vdash^1 \varphi$.

La demostración será considerablemente más complicada que el teorema anterior, y deberá ser de manera indirecta. Para la demostración, primero demostraremos que todo conjunto de fórmulas consistente es satisfacible, y luego demostraremos que esto es equivalente al teorema de compleción. Para mayor precisión en la prueba véase (Enderton 2004: 198-208)

Lema 1.7. *Todo conjunto* Γ *de fórmulas consistente es satisfacible*

Demostración. El esbozo de demostración es como sigue:

1) Tomaremos un conjunto Γ de fórmulas cualquiera que sea consistente

2) Consideraremos el conjunto $\Gamma \cup \Lambda$, donde Λ es el conjunto de todos los axiomas de *D1=* que será igualmente consistente, pues ambos conjuntos son consistentes.[10]

3) Extenderemos este el conjunto $\Gamma \cup \Lambda$ a un conjunto $\Gamma \cup \Lambda \cup \Theta$, agregando un conjunto C de constantes nuevas. Se introduce la fórmula θ_n que es $\neg\forall x_n \varphi_n \supset {\varphi_n}^{x_n}_{c_n}$ donde c_n es la primera $c \in C$ tal que no aparece en ninguna θ_k, $k<n$. Θ es el conjunto de todas las fórmulas, $\neg\forall x_n \varphi_n \supset {\varphi_n}^{x_n}_{c_n}$. La idea de agregar estas fórmulas es garantizar que para cada fórmula φ, haya una constante para nombrar a su contraejemplo, si es que hay alguno. El nuevo conjunto es consistente por inducción sobre el subíndice n, en θ_n

4) Extenderemos el conjunto $\Gamma \cup \Lambda \cup \Theta$ de la siguiente manera: definimos una numeración sobre las fórmulas del lenguaje[11], α_1, α_2, α_3,... Definimos nuestra extensión de manera inductiva:

$$\Delta_0 = \Gamma \cup \Lambda \cup \Theta$$

$$\Delta_{n+1} = \begin{cases} \Delta_n \cup \{\alpha_{n+1}\} & \text{si } \Delta_n \nvdash^1 \neg\alpha_{n+1} \\ \Delta_n & \text{si } \Delta_n \vdash^1 \neg\alpha_{n+1} \end{cases}$$

[10]La consistencia de los axiomas de *D1=* se sigue del teorema de corrección. De no ser así, habría una estructura $\mathfrak{A}$ tal que una fórmula φ sería verdadera y falsa en $\mathfrak{A}$ lo cual es imposible.

[11]Para esto se apela al hecho de que el lenguaje es numerable, y por lo tanto las fórmulas también.

Se considera $\Delta = \bigcup\{\Delta_n | n \in \omega\}$. Se demuestra que el conjunto Δ es consistente, por inducción sobre los subíndices de las Δ's.[12] Además, Δ es maximal en el sentido que para cualquier fórmula φ de *L1K=*, $\varphi \in \Delta$ o $\neg\varphi \in \Delta$. Finalmente, Δ es deductivamente cerrado: $\Delta \vdash^1 \varphi$ si y solo si $\varphi \in \Delta$.

5) Definiremos una estructura cuyo dominio será el conjunto de los términos del lenguaje expandido: Sea $\mathfrak{A} = \langle D^{\mathfrak{A}}, I^{\mathfrak{A}}\rangle$, donde $D^{\mathfrak{A}}$ es el conjunto de los términos de lenguaje expandido e $I^{\mathfrak{A}}$ es tal que:

   a) Para cada constante c de *L1KC=* (el lenguaje expandido), $c^{\mathfrak{A}} = c$
   b) Para cada símbolo de predicado, P^n_m
   $\langle t_1, \ldots, t_m\rangle \in P^{n\,\mathfrak{A}}_m$ si y solo si $P^n_m(t_1,\ldots,t_m) \in \Delta$
   c) Para cada símbolo de función, f^n_m, $f^{n\,\mathfrak{A}}_m$ está definida así:
   $f^{n\,\mathfrak{A}}_m(t_1,\ldots,t_m) = f^n_m(t_1,\ldots,t_m)$

   Ahora definimos un predicado binario $E^{\mathfrak{A}}$ de la siguiente forma:

   $\langle u, t\rangle \in E^{\mathfrak{A}}$ si y solo si $(u = t) \in \Delta$

   También definimos una función s: V $\to D^{\mathfrak{A}}$, simplemente como $s(x) = x$, para cualquier variable x. Además, definimos para cualquier fórmula φ, la fórmula φ^* que es el resultado de sustituir el símbolo de identidad por el predicado E. Basta con mostrar que, para cualquier fórmula φ.

   $\vDash^1_{\mathfrak{A}} \varphi*$ si y solo si $\varphi \in \Delta$

   La prueba de esto es por inducción sobre el número de conectivas y cuantificadores. En este paso, se puede ver por qué es esencial introducir el conjunto Θ, pues se requiere para poder demostrar el caso del cuantificador universal en el paso inductivo.

6) Esto no es un modelo de nuestro lenguaje con identidad, por ello se define una nueva estructura $\mathfrak{B} = \langle D^{\mathfrak{B},\ I^{\mathfrak{B}}}\rangle$, donde $D^{\mathfrak{B}}$ es el conjunto de todas las clases de equivalencia de E sobre el dominio de $\mathfrak{A}$, $\mathfrak{A}/E$, (para cada término t, definimos $[t]$ como su clase de equivalencia) e $I^{\mathfrak{B}}$ está definida de la siguiente forma:

   a) Para cada constante c de *L1KC=*, $c^{\mathfrak{B}} = [c^{\mathfrak{A}}]$
   b) Para cada símbolo de predicado, P^n_m
   $\langle [t_1], \ldots, [t_m]\rangle \in P^{n\,\mathfrak{B}}_m$ si y solo si $\langle t_1, \ldots, t_m\rangle \in P^{n\,\mathfrak{A}}_m$
   c) Para cada símbolo de función, f^n_m, $f^{n\,\mathfrak{B}}_m$ está definida así:
   $f^{n\,\mathfrak{B}}_m([t_1],\ldots,[t_m]) = [f^{n\,\mathfrak{A}}_m(t_1,\ldots,t_m)]$

   Sea h: $D^{\mathfrak{A}} \to D^{mathfrakB}$, $h(t) = [t]$. Esta función preserva la función de interpretación $I^{\mathfrak{A}}$ en $D^{\mathfrak{B}}$.[13] Por lo tanto, para la función s que definimos en el paso anterior se cumple que

[12] El caso base es Δ_0 que es consistente por el paso 3. El paso inductivo es muy sencillo: si $\Delta_n \nvdash^1 \neg\alpha_{n+1}$, agregar α_{n+1} no genera inconsistencias pues Δ_n es consistente, y el segundo caso está garantizado por la misma razón.

[13] Esta función es lo que más adelante definiremos como homomorfismo, y veremos más claramente que esta función preserva la función de interpretación.

$\models^1_{\mathfrak{A}} \varphi * [h \circ s]$ si y solo si $\varphi \in \Delta$

donde $h \circ s$ es la composición de h y s (véase A). Por lo cual, $\mathfrak{B}$ es un modelo de Δ.

7) Por último, basta señalar que podemos acotar el modelo al lenguaje original. Como $\Gamma \subseteq \Delta$, garantizamos que $\mathfrak{B}$ con $h \circ s$ satisface a todos los miembros de Γ

□

Lema 1.8. *Todo conjunto Γ de fórmulas consistente es satisfacible si y solo si, si $\Gamma \models^1 \varphi$ entonces $\Gamma \vdash^1 \varphi$*

Demostración. $\Rightarrow$ Supongamos que, $\Gamma \models^1 \varphi$ y que $\Gamma \nvdash^1 \varphi$. Por lo tanto, $\Gamma \cup \{\neg\varphi\}$ es consistente, por lo cual es satisfacible. Por lo tanto, hay una estructura donde todas las fórmulas de Γ son verdaderas y φ es falsa, lo cual contradice el supuesto. Es decir, $\Gamma \vdash^1 \varphi$

$\Leftarrow$ Sea Γ un conjunto de fórmulas consistente que no es satisfacible. Sabemos que para toda $\psi \in \Gamma$, $\Gamma \nvdash^1 \neg\psi$. Por lo cual $\Gamma \not\models^1 \neg\psi$, esto es, hay una estructura que hace verdaderas a todas las fórmulas $\psi \in \Gamma$, por lo tanto es satisfacible, pero habíamos supuesto que no. Por lo tanto, es satisfacible. □

1.1.4. Teorema de compacidad y los teoremas de Lowenheim-Skolem

En esta última parte, concluiremos la exposición de la lógica de primer orden, exponiendo los teoremas de compacidad y de Lowenheim-Skolem. De nuevo, será una exposición rápida y sólo esbozaremos las pruebas de los teoremas.

Teorema 1.5. ***(Teorema de compacidad)****:*

a) Si $\Gamma \models^1 \varphi$ entonces existe un subconjunto finito $\Gamma_0 \subseteq \Gamma$, tal que $\Gamma_0 \models^1 \varphi$

b) Si cada subconjunto finito $\Gamma_0 \subseteq \Gamma$ es satisfacible, entonces Γ es satisfacible

Ambas partes del teorema son equivalentes. Y ambas pruebas son consecuencias inmediatas del hecho de que las demostraciones son finitas y del teorema de compleción.

La demostración de los teoremas de Lowenheim-Skolem requiere una nota sobre la naturaleza del lenguaje que usamos. El lenguaje *L1K=* fue un lenguaje donde el conjunto de la terminología no lógica era un conjunto numerable. Por lo tanto nuestro lenguaje tiene una cantidad numerable de fórmulas, de términos y de expresiones, es decir, el lenguaje *L1K=* es numerable.

Consideremos un lenguaje L cuya cardinalidad es λ y supongamos que queremos demostrar el teorema de compleción para dicho lenguaje. Las modificaciones necesarias son las siguientes: en el paso 4 se agregan una cantidad λ de nuevos símbolos de constante y se construyen las θ-fórmulas con esto en mente, y se utiliza una enumeración de las fórmulas sobre λ, lo cual es posible por el axioma de elección (véase A). La demostración de que este nuevo conjunto es consistente ocupa inducción transfinita sobre λ y la construcción del conjunto Δ maximal ocupa el axioma de elección.

Finalmente, la cardinalidad del conjunto $D^{\mathfrak{A}}$ de la estructura $\mathfrak{A}$ será λ, y el conjunto de la relaciones de equivalencia sobre $D^{\mathfrak{A}}$ módulo E, $\mathfrak{A}/E$, es tal que $|D^{\mathfrak{A}/E}| \leq |D^{\mathfrak{A}}|$. Es decir, el modelo final será tal que la cardinalidad de su dominio es $\leq \lambda$. Por lo tanto:

Teorema 1.6. ***(Teorema de Lowenheim-Skolem (descendente))***

a) Sea Γ un conjunto satisfacible de fórmulas en un lenguaje de cardinalidad λ. Entonces Γ es satisfacible en alguna estructura de tamaño $\leq \lambda$.

b) Sea Σ un conjunto de enunciados de un lenguaje de cardinalidad λ, si Σ tiene modelo entonces Σ tiene un modelo de cardinalidad $\leq \lambda$.

El razonamiento anterior funciona como una prueba general de este teorema y muestra que la demostración del teorema de compleción considera el caso donde $\lambda = \aleph_0$

Teorema 1.7. ***(Teorema de Lowenheim-Skolem (ascendente))**: Sea Γ un conjunto de fórmulas en un lenguaje de cardinalidad λ y supongamos que Γ es satisfacible en alguna estructura infinita. Entonces, para todo cardinal $\kappa \geq \lambda$ hay una estructura de cardinalidad κ en la que Γ es satisfacible.*

La prueba es la siguiente. Sea $\kappa \geq \lambda$ agregamos al lenguaje un conjunto C de constantes nuevas de cardinalidad λ y consideramos el conjunto $\Gamma \cup \Sigma$, donde $\Sigma = \{c_1 \neq c_2 | c_1, c_2 \in C\}$. Todos los subconjuntos finitos de $\Gamma \cup \Sigma$ son satisfacibles en la estructura $\mathfrak{A}$ extendida con referencias de las constantes nuevas. Por el teorema de compacidad, $\Gamma \cup \Sigma$ es satisfacible y por el teorema de Lowenheim-Skolem (descendente) hay una estructura $\mathfrak{B}$ de cardinalidad $\leq \kappa$ que satisface $\Gamma \cup \Sigma$. Pero, como Σ tiene κ fórmulas, $\Gamma \cup \Sigma$ sólo es satisfacible en estructuras de cardinalidad $\geq \kappa$. Por lo tanto, la cardinalidad de $\mathfrak{B}$ es κ.

Esto es todo por lo que respecta a la lógica de primer orden, ahora expondré en qué consiste la lógica de segundo orden.

1.2. Lógica de segundo orden

Ahora hablaré de la lógica de segundo orden, haciendo énfasis en dos aspectos: (1) es una extensión de la lógica de primer orden y (2) no cumple con los teoremas de compleción, compacidad y Lowenheim-Skolem. La semántica que presentaremos es una extensión de la semántica de la lógica de primer orden, y se conoce como *semántica estándar* de la lógica de segundo orden. En vista de esto, sólo señalaré cómo extender las definiciones de la lógica de primer orden para el caso de la lógica de segundo orden.

Un lenguaje de segundo orden es aquel que, al igual que la lógica de primer orden, tiene variables cuyo rango es el dominio de una interpretación de un lenguaje y además tiene variables cuyo rango es el conjunto de las propiedades (o subconjuntos) de los objetos del dominio de una interpretación. Nuestro lenguaje es *L2K* (sin el símbolo de identidad). Para definir este lenguaje, consideraremos el lenguaje *L1K* que es simplemente *L1K=* – {'='}. Extenderemos sólo la terminología lógica de este lenguaje[14] con dos conjuntos más:

- Símbolos de variable de predicado: Para cada número natural n y m, X^n_m es un símbolo de variable de predicado.
- Símbolos de variable de función: Para cada número natural n y m, F^n_m es un símbolo de variable de función.[15]

Extenderemos las reglas de formación de *L1K=*. Solamente agregaremos una nueva regla de formación de términos:

- Si F^n_m es una variable de función y $t_1,\ldots,t_m$ son términos (que no sean variables de predicado ni de función), entonces $F^n_m(t_1,\ldots,t_m)$ es un término

Y señalaré que en el caso de la regla 1 de las reglas de formación para *L1K=* se incluyen también las variables de predicado y de función. Modificaremos las reglas de formación de fórmulas de *L1K=* quitando la regla 1 e incluyendo las siguientes:
Reglas de formación de fórmulas

1. Si $t_1,\ldots,t_m$ son términos (que no sean variables de predicado ni de función) y X^n_m es una variable de predicado, entonces X^n_m $(t_1,\ldots,t_m)$ es una fórmula
2. Si φ es una fórmula y X^n_m es una variable de predicado, entonces $\forall X^n_m\varphi$ es una fórmula
3. Si φ es una fórmula y F^n_m es una variable de función, entonces $\forall F^n_m\varphi$ es una fórmula

Al igual que en el caso anterior, permitimos tener fórmulas con variables libres (ya sea, de individuo o de predicado). Hay que indicar que es posible definir el cuantificador existencial para el caso de las variables de predicado de manera análoga, al caso de primer orden. También podemos introducir la identidad entre objetos del dominio con la fórmula, por lo cual, no es necesario introducirlo como un primitivo lógico:

$$(u = t)\equiv_{def} \forall X(X(u) \equiv X(t))$$

Se da por entendido que X es una variable de predicado de aridad 1

[14]Podríamos extender el conjunto de la terminología no lógica K, con un conjunto de constantes de predicado de segundo orden. Nosotros no lo haremos, pues no será necesario y dejaremos el conjunto K de terminología no lógica exactamente igual. En caso de necesitar un predicado para letras (o variables) de predicado, se introducirán mediante fórmulas adecuadas.

[15]Al igual que en el caso de las constantes de predicado del conjunto K de terminología no lógica, por motivos de simplicidad omitiré los subíndices y los superíndices de las variables de predicado y de función cuando no sean necesarios o quede implícito cuáles son sus valores.

1.2.1. Semántica: verdad y modelos para el lenguaje *L2K*

Ahora, ampliaremos las definiciones de satisfacción y verdad para determinar el valor de verdad de las oraciones de segundo orden. Esta semántica es conocida como *semántica estándar*. Dado que no incluimos ningún nuevo conjunto al conjunto K de terminología no lógica, no necesitamos modificar la definición de *estructura*. Así que utilizaremos exactamente las mismas estructuras que usamos para el lenguaje *L1K=* y sólo modificaremos las definiciones de satisfacción y verdad.[16] Precisaré la definición de satisfacción. Sea $\mathfrak{A}$ una estructura, sea S una función cuyo dominio es el conjunto de variables de *L2K* tal que:

- $S(x_n) = d \in D^{\mathfrak{A}}$, para toda variable de individuo
- $S(X^n_m) = P \subseteq D^{\mathfrak{A}^m}$, para toda X^n_m variable de predicado
- $S(F^n_m)) = G{:}D^{\mathfrak{A}^m} \to D^{\mathfrak{A}}$, para toda F^n_m variable de función

De nuevo, definimos $\Sigma^{\mathfrak{A}}$[17] como el conjunto de todas las funciones S. Sea $\bar{S}$una extensión para cada $S \in \Sigma^{\mathfrak{A}}$, de tal forma que $\bar{S}$ proporcione significado a cada término del lenguaje. Tomaremos las cláusulas 1-3 (con las modificaciones pertinentes) de $\bar{s}$ (en la sección 1.1.1) para el caso de los términos de la lógica de primer orden, y daremos cláusulas para el caso de las variables de segundo orden, lo cual será sencillo:

1. $\bar{S}(X^n_m) = S(X^n_m)$, para toda X^n_m variable de predicado
2. $\bar{S}(F^n_m) = S(F^n_m)$, para toda F^n_m variable de función
3. $\bar{S}(F^n_m(t_1,\ldots,t_m)) = \bar{S}(F^n_m)\ ((\bar{S}(t_1),\ldots,\bar{S}(t_m)))$

Ahora podemos extender la noción de satisfacción, para incluir las fórmulas de segundo orden: sea $\mathfrak{A}$ una estructura, φ una fórmula de *L2K* y una $S \in \Sigma^{\mathfrak{S}}$. Decimos que $\mathfrak{A}$ satisface φ con S mediante la siguiente expresión $\vDash_{\mathfrak{S}} \varphi[S]$.[18] La definición es la misma que en el caso de la lógica de primer orden (con las modificaciones apropiadas en las cláusulas 1-5), con las siguientes cláusulas nuevas:

1. Si φ es $X^n_m\ (t_1,\ldots,t_m)$, $\vDash_{\mathfrak{A}} \varphi[S]$ entonces si y solo si $\langle \bar{S}(t_1), \ldots, \bar{S}(t_m)\rangle \in \bar{S}(X^n_m)$
2. Si φ es $\forall X^n_m \psi$ entonces $\vDash_{\mathfrak{A}} \varphi[S]$ si y solo si para todo $P \subseteq D^{\mathfrak{A}^m}, \vDash_{\mathfrak{A}} \psi[S(X^n_m|P)]$

[16] En caso de incluir un conjunto de predicados no lógicos de segundo orden, sí necesitaríamos modificar la definición de estructura. Sin embargo, no implicaría un incremento demasiado significativo de estructuras.

[17] Usaremos la misma notación que en el caso de la lógica de primer orden debido a que las modificaciones en la notación son poco significativas. Sin embargo, debe tenerse en cuenta la diferencia entre ambas. Generalmente, no será ambiguo identificar a qué tipo de funciones nos referimos, en caso contrario se especificará el tipo de funciones consideradas.

[18] Hay que notar que en este caso no se especifica un superíndice, a diferencia del caso de la lógica de primer orden. La razón para omitirla es que la noción de satisfacción en segundo orden incluye el caso de la satisfacción en primer orden. Lo mismo es el caso para el resto de los superíndices omitidos.

3. Si φ es $\forall F^n_m \psi$ entonces $\models_{\mathfrak{A}} \varphi[S]$ si y solo si para toda $G : D^{\mathfrak{A}m} \to D^{\mathfrak{A}}, \models_{\mathfrak{A}} \psi[S(F^n_m|G)]$

Donde para cualquier Y variable de predicado y cualquier F variable de función:

$$S(X^n_m|P) = \begin{cases} S(Y) & \text{si } Y \neq X^n_m \\ P & \text{si } Y = X^n_m \end{cases}$$
$$S(F^n_m|G) = \begin{cases} S(F) & \text{si } F \neq F^n_m \\ G & \text{si } F = F^n_m \end{cases}$$

Vale la pena notar que los cuantificadores de las cláusulas 2 y 3 se aplican a todas las relaciones de n-lugares y a todas las funciones de n-lugares de elementos del dominio. Es decir, la cuantificación universal es efectivamente "universal", en el sentido de que el rango de las variables efectivamente es el todo el conjunto de las relaciones y de las funciones del dominio.

Las definiciones de verdad y consecuencia lógica son completamente análogas a las definiciones para el caso de primer orden: sean φ una fórmula de *L2K*, $\mathfrak{A}$ una estructura y Γ un conjunto de fórmulas de *L2K*:

1. φ es verdadera en $\mathfrak{A}$ ($\models_{\mathfrak{A}} \varphi$) si y solo si $\models_{\mathfrak{A}} \varphi[S]$, para toda $S \in \Sigma^{\mathfrak{A}}$

2. $\mathfrak{A}$ es modelo de Γ si y solo si $\models_{\mathfrak{A}} \psi$, para toda $\psi \in \Gamma$

3. φ es una consecuencia lógica de Γ ($\Gamma \models \varphi$) si y solo si, para toda estructura $\mathfrak{A}$ y para toda $S \in \Sigma^{\mathfrak{A}}$, si para toda $\psi \in \Gamma$, $\models_{\mathfrak{A}} \psi[S]$ entonces, $\models_{\mathfrak{A}} \varphi[S]$

Esta definición de consecuencia lógica para el caso de segundo orden es conservativa con respecto a la definición en primer orden. Es decir, todo caso de consecuencia lógica en primer orden es un caso de consecuencia lógica en segundo orden. Esto es natural en vista de que utilizamos exactamente las mismas estructuras y la definición de satisfacción para las fórmulas de primer orden es la misma en ambos casos. Dado esto y debido a que mi interés es la lógica de segundo orden, llamaré definición modelo teóretica a la definición de consecuencia lógica en segundo orden (con el objetivo de incluir ambas definiciones).[19]

1.2.2. Un sistema deductivo para la lógica de segundo orden

De la misma manera que lo hicimos anteriormente, sólo extenderemos el sistema *D1=*, a un sistema para la lógica de segundo orden. La extensión sólo requiere incluir esquemas para los cuantificadores universales análogas a las del caso de primer orden:

1. $\forall X^n_m \alpha \supset \alpha^{X^n_m}_T$[donde T puede ser una variable de predicado o un símbolo de predicado y se puede sustituir por X^n_m en α]

2. $\alpha \supset \forall X^n_m \alpha$[donde X^n_m no ocurre libre en α]

[19]Igualmente, cabe señalar que la noción (CLTC) expuesta en la introducción es básicamente esta definición precisa.

3. $\forall X_m^n (\alpha \supset \beta) \supset (\forall X_m^n \alpha \supset \forall X_m^n \beta)$

4. $\forall F_m^n \alpha \supset \alpha_G^{F_m^n}$[donde G puede ser una variable de función o un símbolo de función y se puede sustituir por F_m^n en α]

5. $\alpha \supset \forall F_m^n \alpha$[donde F_m^n no ocurre libre en α]

6. $\forall F_m^n (\alpha \supset \beta) \supset (\forall F_m^n \alpha \supset \forall F_m^n \beta)$

Incluiremos, dos axiomas más:

Esquema de Comprehennsión:
$\exists X_m^n \forall x_1 \ldots x_n (X_m^n (x_1 \ldots x_n) \equiv \alpha (x_1 \ldots x_n))$ [si X_m^n no ocurre libre en α]

Axioma de elección:
$\forall X_m^{n+1} (\forall x_1 \ldots x_n \exists y X_m^n (x_1 \ldots x_n y) \supset \exists F_m^n \forall x_1 \ldots x_n X_m^{n+1} (x_m \ldots x_n F_m^n (x_1 \ldots x_n)))$

Al sistema compuesto con *D1=* y estos axiomas lo llamaremos *D2*. Las razones para incluir los primeros axiomas son análogas a los casos de primer orden. Los axiomas 1 y 4 garantizan que si un universal es verdadero, la fórmula cuantificada universalmente es verdadera de cualquier objeto adecuado que caiga bajo la cuantificación. Los axiomas 2, 3, 5 y 6 son incluidos para demostrar versiones adecuadas de la regla de generalización. El axioma de comprehensión nos permite introducir propiedades a partir de fórmulas del lenguaje, mientras que el axioma de elección nos permite garantizar la existencia de funciones que se corresponden adecuadamente con cualquier relación, es decir, nos permite eliminar los símbolos de funciones por símbolos de predicado.

Al igual que en la lógica de primer orden, la única regla de inferencia es el Modus Ponens, y teniendo las correspondientes reglas de generalización, la prueba del siguiente teorema es exactamente la misma que en el caso de la lógica de segundo orden:

Teorema 1.8. ***(Teorema de la deducción)****: Si* $\Gamma \cup \{\gamma\} \vdash \varphi$ *entonces* $\Gamma \vdash \gamma \supset \varphi$

1.2.3. Otras semánticas para la lógica de segundo orden

En la sección 1.2.1 presentamos la que es conocida como la semántica estándar para la lógica de segundo orden. Ahora expondremos dos alternativas a esta semántica: la *semántica de Henkin* y la *semántica de primer orden.*

1.2.3.1. La semántica de Henkin

La idea detrás de la semántica de Henkin es que las variables de *L2K* no tienen como rango el conjunto completo de las relaciones (ni funciones) del dominio de una interpretación, si no sólo de un subconjunto fijo de él. En las semánticas de Henkin la relación de predicación es no lógica y el significado del cuantificador universal es más débil en el sentido de que no opera sobre todos los subconjuntos del dominio, sino sobre un conjunto de subconjuntos previamente determinado, lo cual hace en términos generales a este tipo de semánticas más débiles que la semántica estándar.

Definición. *Una estructura de Henkin se define como* $\mathfrak{A}^H = \langle d^H, D^H, F^H, I^H\rangle$ *donde* d^H *es un conjunto (el dominio de la estructura) e* I^H *es la función de interpretación.* D^H *y* F^H *serán dos subconjuntos de conjunto de relaciones y funciones. En general:*

- *Para cada n,* $D^H(n) \subseteq \wp(D^H)^n$, $\bigcup\{D^H(n) \mid$ *para todo número natural* $n\}=D^H$
- *Para cada n,* $D^H(n) \subseteq \{G{:}(D^H)^n \rightarrow d^H\}$, $\bigcup\{F^H(n)|$*para todo número natural* $n\}=F^H$

La función I^H asignará a la terminología no lógica un elemento adecuado de manera análoga al caso estándar, por ejemplo, a los símbolos de predicado de n-lugares les asigna elementos de $D^H(n)$. Las definiciones de satisfacción, verdad y consecuencia lógica (*consecuencia lógica de Henkin*) son exactamente las mismas, con la única salvedad de que los cuantificadores de la definición de satisfacción para las fórmulas universales correrán sobre el dominio correspondiente. Llamaré *verdad lógica de Henkin* a una verdad lógica en esta semántica. Definimos una *estructura de Henkin completa* de una estructura estándar $\mathfrak{A}$, como la estructura de Henkin $\mathfrak{A}^H$, donde $D^H(n) = \wp((D^H)^n)$ y $F^H(n) = \{G{:}(D^H)^n \rightarrow d^H\}$. Considerando las estructuras de Henkin completas, podemos ver que toda estructura estándar es una estructura de Henkin, en este sentido, podemos decir que las semánticas de Henkin tienen "más" interpretaciones y, por lo tanto:

Teorema 1.9. ***(Teorema H-E)****:*

a) Si una fórmula ϕ *es una verdad lógica de Henkin entonces es una verdad lógica estándar.*

b) Si una fórmula ϕ *es una consecuencia lógica de Henkin de* Γ *entonces* ϕ *es una consecuencia lógica estándar de* Γ

c) Si una fórmula ϕ *es satisfacible en la semántica estándar entonces es satisfacible en la semántica de Henkin.*

1.2.3.2. La semántica de primer orden

Al igual que en el caso de las semánticas de Henkin, la semántica de primer orden considera que la relación de predicación es no lógica. La esencia de esta semántica es considerar las relaciones y funciones como objetos de primer orden. Se definen dos dominios diferentes, uno para los objetos y otro para las relaciones, en los que ambos son considerados objetos de primer orden y se establece una relación no lógica entre ellos, la relación de predicación. En cada estructura habrá una relación de predicación diferente y es posible que esta sea vacía, es decir, que los dominios no estén relacionados.

Definición. *Una estructura de primer orden es una tupla* $\mathfrak{A}^1 = \langle d^1, d^1_1, d^1_2, \langle I^1, p^1, a^1\rangle\rangle$ *donde* d^1*, es un conjunto (el rango de las variables de primer orden) e* I^1 *es la función de interpretación.* d^1_1 *y* d^1_2 *serán los rangos para las variables de predicado y variables de función, respectivamente; mientras que* p^1 *y* a^1 *serán las funciones de predicación y la función de aplicación sobre funciones, respectivamente:*

- *Para cada n,* d_1^1 *(n) es el rango de las variables de predicado de n-lugares*
- *Para cada n,* d_2^1 *(n) es el rango de las variables de función de n-lugares*
- *Para cada n,* $p^1(n) \subseteq d^{1n} \times d_1^1(n)$ *es la relación de predicación entre las n-secuencias de elementos de* d^1 *y el rango de las variables de predicado de n-lugares*
- *Para cada n,* $a^1(n){:}((d^1)^n \times d_2^1(n)) \rightarrow d^1$ *es la función de aplicación entre una n-secuencia de elementos de* d^1 *y un elemento del rango de las variables de función con elementos de* d^1

Hay que notar que no se puso ninguna restricción a los dominios d_1^1 y d_2^1, es decir, pueden ser conjuntos de objetos cualesquiera. Así pues, lo que permite a estas estructuras constituir una semántica adecuada son las funciones p^1 y a^1. La función I^1, asignará a la terminología no lógica un elemento adecuado de manera análoga al caso estándar y al caso de las semánticas de Henkin, sin embargo, requieren modificaciones a la definición de satisfacción, aunque no son muy complicadas. Las definiciones de consecuencia lógica y verdad lógica son exactamente iguales, y las llamaré *consecuencia lógica de primer orden* y *verdad lógica de primer orden.*

1.2.4. Resultados metateóricos para la lógica de segundo orden

En esta sección expondré las propiedades metateóricas de la lógica de segundo orden, con las diferentes semánticas. Expondré los teoremas de categoricidad para la lógica de segundo orden con semántica estándar, y señalaremos porqué dicho resultado implica que dicha semántica es inherentemente incompleta. De la misma forma, señalaremos que no satisface los teoremas de compacidad ni Lowenheim-Skolem. Esto distinguirá la semántica estándar con las demás semánticas, pues tanto la *semántica Henkin* como la de *primer orden* son completas y cumplen los teoremas de compacidad y Lowenheim-Skolem

1.2.4.1. Metateoremas de la semántica estándar

Teorema 1.10. ***(Teorema de corrección para L2K y D2)****: Si* $\Gamma \vdash \varphi$ *entonces* $\Gamma \vDash \varphi$

Basta con extender el lema 1.1 con los axiomas aumentados de *D2*. Para el caso de los axiomas 1 y 4 se utilizan versiones adecuadas de la demostración del esquema 4 de axiomas de *D1=*, y de los lemas 1.2 y 1.3. Para el caso 2 y 5 se utilizan versiones de la demostración del esquema 5, y para los axiomas 3 y 6, corresponden al esquema 6 de *D1=*. Para el caso del axioma de comprehensión, se sigue del uso del axioma de comprensión en la metateoría. De la misma forma en el caso del axioma de elección que depende del axioma de elección en la metateoría. De esto y de una versión adecuada del lema 1.6 se sigue que este teorema es el caso.

La refutación de los teoremas de compleción, compacidad y Lowenheim-Skolem, dependen de los teoremas de categoricidad de la aritmética y del análisis. Expondremos únicamente el esquema de prueba del teorema de categoricidad para la aritmética, el caso del análisis es similar, y sólo señalaremos en qué consiste dicha teoría.

El lenguaje de la *aritmética de Peano* (en segundo orden[20]) es $L2K \cup A$, donde $A = \{0, +, \bullet, \underline{s}\}$, el lenguaje no lógico de la aritmética. La *aritmética de Peano* en segundo orden (AR) es el conjunto de los siguientes cuatro axiomas:

1. $\forall x(\underline{s}x \neq 0) \land \forall x \forall y(\underline{s}x = \underline{s}y \supset x = y)$
2. $\forall x(x + 0 = x) \land \forall x \forall y(x + \underline{s}y = \underline{s}(x + y))$
3. $\forall x(x \bullet 0 = 0) \land \forall x \forall y(x \bullet \underline{s}y = (x \bullet y) + x)$
4. $\forall X(X0 \land \forall x(Xx \supset X\underline{s}x) \supset \forall x Xx)$

Los primeros tres axiomas definen el funcionamiento de las operaciones $\underline{s}$, $+$ y $\bullet$, respectivamente, y son fórmulas de primer orden. El último axioma es el principio de inducción y caracteriza a todo conjunto cuya cardinalidad sea $\aleph_0$, y es una fórmula pura de segundo orden.

Por otro lado, el lenguaje del *análisis real* (en segundo orden) es $L2K \cup B$ donde $B = \{0, 1, \bullet, +, \leq\}$, el lenguaje no lógico del análisis real. El *análisis real* en segundo orden (AN) está compuesto por los axiomas de campo[21] (que sirven para definir el funcionamiento de las funciones $\bullet$ y $+$, y la relación *leq*), junto con el axioma de completud:

$$\forall X(\exists x \forall x(Xy \supset y \leq x) \supset \exists x(\forall y(Xy \supset y \leq x) \land \forall z(\forall y(Xy \supset y \leq x) \supset x \leq z)))$$

Que es la afirmación de que todo subconjunto acotado superiormente tiene un supremo y también es una fórmula pura de segundo orden.

Teorema 1.11. ***(Teorema de categoricidad de AR)****: Sean $\mathfrak{A}$ y $\mathfrak{B}$ dos estructuras para el lenguaje de AR. Definimos para cada $i \in \{\mathfrak{A}, \mathfrak{B}\}$, 0_i como la interpretación de 0 en D^i, y sea $\underline{s}_i$, $+_i$ y $\bullet_i$, las respectivas interpretaciones de $\underline{s}$, $+$ y $\bullet$ en la estructura i.* ***Si*** $\vDash_{\mathfrak{A}}$ ***AR y*** $\vDash_{\mathfrak{B}}$ ***AR entonces*** $\mathfrak{A}$ ***y*** $\mathfrak{B}$ ***son isomorfos****, es decir, hay una función biyectiva $f{:}D^{\mathfrak{A}} \to D^{\mathfrak{B}}$, que cumple las siguientes condiciones:*

I. $f(0_{\mathfrak{A}}) = 0_{\mathfrak{B}}$

II. $f(\underline{s}_{\mathfrak{A}} a) = \underline{s}_{\mathfrak{B}}(f(a))$, *para todo* $a \in D^{\mathfrak{A}}$

III. $f(a +_{\mathfrak{A}} b) = f(a) +_{\mathfrak{B}} f(b)$, *para cualesquiera* a, $b \in D^{\mathfrak{A}}$

IV. $f(a \bullet_{\mathfrak{A}} b) = f(a) \bullet_{\mathfrak{B}} f(b)$, *para cualesquiera* a, $b \in D^{\mathfrak{A}}$

Demostración. Sea S $\subseteq D^{\mathfrak{A}} \times D^{\mathfrak{B}}$, decimos que S es cerrado bajo la sucesión si y solo si

[20] El lenguaje de la aritmética en primer orden es el mismo pero considerando $L1K$= en lugar de $L2K$.

[21] Los axiomas de campo son demasiados y por motivo de exposición lo omitiremos.

I) $\langle 0_{\mathfrak{A}}, 0_{\mathfrak{B}} \rangle \in$ S

II) Si $\langle a, b\rangle \in$ S, entonces $\langle \underline{s}a, \underline{s}b\rangle \in$ S

Sea $f = \bigcap\{$S$\subseteq D^{\mathfrak{A}} \times D^{\mathfrak{B}}|$ S es cerrado bajo la sucesión$\}$. f no es vacío y, claramente, es cerrado bajo la sucesión. Para demostrar que f es la función que buscamos seguiremos los siguientes pasos:

1. El dominio de f es $D^{\mathfrak{A}}$ $[dom(f) = D^{\mathfrak{A}}]$. Definimos P $= \{a\in D^{\mathfrak{A}}|$hay un $b \in D^{\mathfrak{B}}$ tal que $\langle a, b\rangle \in f\}$, es decir, P $= dom(f)$. La demostración consiste en mostrar que P $= D^{\mathfrak{A}}$, para ello se demuestra que $0 \in$ P y si $a \in$ P entonces $\underline{s}_{\mathfrak{A}} a \in$ P, para cualquier $a \in D^{\mathfrak{A}}$. Como $\mathfrak{A}$ satisface el axioma de inducción y P $\subseteq D^{\mathfrak{A}}$, P $= D^{\mathfrak{A}}$

2. f es función [si $\langle a, b\rangle \in f$ y $\langle b, c\rangle \in f$, entonces $b = c$]. Definimos Q $= \{a \in D^{\mathfrak{A}}|\exists! b \in D^{\mathfrak{B}}$ tal que $\langle a, b\rangle \in f\}$, la condición del conjunto Q expresa precisamente que f es función. Al igual que en el caso anterior se demuestra que $0 \in$ Q y si $a \in$ Q entonces $\underline{s}a \in$ Q, para cualquier $a \in D^{\mathfrak{A}}$. $\mathfrak{A}$ satisface el axioma de inducción, por lo cual, Q $= D^{\mathfrak{A}}$

3. f es biyectiva. Se demuestra en dos pasos. En el primero se demuestra que fs es suprayectiva $[ran(f) = D^{\mathfrak{B}}]$, lo cual se demuestra de manera análoga al paso 1. En el segundo, se demuestra que f es inyectiva [si $\langle a, b\rangle \in f$ y $\langle c, b\rangle \in f$, entonces $a = c$], lo cual es análogo al paso 2. En ambos casos se utiliza que $\mathfrak{B}$ satisface el axioma de inducción.

4. f preserva la estructura, es decir, cumple I-IV. En esta parte, se utilizan los tres primeros axiomas de AR, sin usar el axioma de inducción, y es muy sencilla. La cláusula II, en particular, es consecuencia trivial de que f es cerrada bajo la sucesión.

Por lo tanto, $\mathfrak{A}$ y $\mathfrak{B}$ son isomorfos. □

Esto mismo se cumple también para el caso de AN:

Teorema 1.12. ***(Teorema de categoricidad de AN)*** *Sean $\mathfrak{A}$ y $\mathfrak{B}$ dos estructuras para el lenguaje de AN. Si $\vDash_{\mathfrak{A}} AN$ y $\vDash_{\mathfrak{B}} AN$ entonces $\mathfrak{A}$ y $\mathfrak{B}$ son isomorfos.*

La demostración es un poco más complicada que en el caso de AR, sin embargo, no requiere mayor profundidad ni complejidad, por lo cual, omitiré la prueba. Cabe señalar que la prueba exige el uso del axioma de completud mencionado más arriba y es fundamental para garantizar el teorema.

Estos dos teoremas implican que los teoremas que demostramos para la lógica de primer orden no se cumplen para la *semántica estándar* de la lógica de segundo orden. Para el caso del teorema de compleción, concentrémonos en el lenguaje de AR:

Teorema 1.13. ***(Refutación del teorema de compleción)*** *Sea D un sistema deductivo efectivo que es correcto para el lenguaje de AR. Entonces, D no es débilmente completo: hay una fórmula φ que es una verdad lógica que no es un teorema de D.*

Demostración. Sea T $= \{\phi | \phi$ no tiene variables de relación ni de función y $(AR \supset \phi)$ es un teorema de $D\}$. Como D es efectivo, T es efectivamente numerable. Como D es correcto, $(AR \supset \phi)$ es verdadero en los números naturales (por el teorema de categoricidad para AR). Por el teorema de incompletud de la aritmética de Gödel[22], sabemos que las oraciones de primer orden verdaderas de la aritmética no son efectivamente numerables. Por lo cual, sea φ una fórmula de primer orden verdadera de la aritmética, tal que $\varphi \notin$ T. Como $\varphi \notin$ T, $(AR \supset \varphi)$ no es teorema de D, pero es una verdad lógica (por el teorema de categoricidad para AR) □

Hay que notar que es suficiente que alguna semántica cumpla con el teorema de categoricidad para que este argumento funcione. Por lo tanto, la lógica de segundo orden con semántica estándar es inherentemente incompleta.

Teorema 1.14. ***(Refutación de teorema de compacidad)*** *El conjunto* $\Gamma = E \cup \{\forall F \neg (\forall x \forall y (Fx = Fy \supset x = y) \wedge \exists x \forall y (Fx \neq x))\}$ *, (donde* $E = \{\exists x_1 \exists x_2 (x_1 \neq x_2), \exists x_1 \exists x_2 \exists x_3 (x_1 \neq x_2 \wedge x_2 \neq x_3 \wedge x_1 \neq x_3), \ldots\}$*), es tal que todos sus subconjuntos finitos son satisfacibles pero* Γ *no es satisfacible*

Demostración. La fórmula $\forall F \neg (\forall x \forall y (Fx = Fy \supset x = y) \wedge \exists x \forall y (Fx \neq x))$ es verdadera en toda estructura cuyo dominio es finito. Es una fórmula propia de segundo orden y no existe ninguna fórmula de primer orden equivalente a ella. Por otro lado, las fórmulas del conjunto E afirman, intuitivamente, cada una que hay al menos n objetos, donde n es el subíndice mayor de las variables de la fórmula. Por ejemplo, $\exists x_1 \exists x_2 \exists x_3 (x_1 \neq x_2 \wedge x_2 \neq x_3 \wedge x_1 \neq x_3)$ afirma que hay al menos tres objetos en el dominio de la estructura, y es satisfecha por las estructuras que tienen al menos tres objetos en su dominio. Cada subconjunto finito es satisfacible, por una estructura finita que contenga en su dominio al menos n objetos, donde donde n es el subíndice mayor de las variables de las fórmulas del subconjunto. Sin embargo, Γ no es satisfacible pues el conjunto $\{\exists x_1 \exists x_2 (x_1 \neq x_2), \exists x_1 \exists x_2 \exists x_3 (x_1 \neq x_2 \wedge x_2 \neq x_3 \wedge x_1 \neq x_3), \ldots\}$ sólo es satisfacible por una estructura cuyo dominio es infinito y $\forall F \neg (\forall x \forall y (Fx = Fy \supset x = y) \wedge \exists x \forall y (Fx \neq x))$ sólo es satisfacible por estructuras cuyo dominio es finito. □

Los teoremas de Lowenheim-Skolem fallan también como consecuencia de los teoremas de categoricidad. Consideremos la teoría AN, por el teorema de categoricidad para AN (teorema 1.12), AN sólo es satisfecha por estructuras que satisfagan el axioma de completud, pero toda estructura que satisfaga este axioma tiene un dominio cuya cardinalidad es $2^{\aleph_0}$, por lo cual:

Teorema 1.15. ***(Refutación del teorema de Lowenheim-Skolem (descendente))*** *No existe una estructura* $\mathfrak{A}$ *que tenga cardinalidad* $\leq \aleph_0$ *que satisfaga* AN

Considerando que $\aleph_0$ es el cardinal del lenguaje de AN.

[22] Muy suscintamente, el teorema de incompletud de la aritmética demuestra que Si T es una teoría efectiva de la aritmética y es consistente, entonces hay una oración ϕ que es verdadera en T pero que no es demostrable en T. Esto implica que las oraciones verdaderas y las demostrables no coinciden, y que el conjunto de las oraciones verdaderas de T no es efectivamente numerable (pues los teoremas de T de hecho lo son).

Para el caso ascendente, consideremos *AR*. Toda estructura que satisfaga a *AR* satisface el axioma de inducción, pero toda estructura que satisface el axioma de inducción tiene un dominio cuya cardinalidad es $\aleph_0$, por lo cual:

Teorema 1.16. ***(Refutación del teorema de Lowenheim-Skolem (ascendente))*** *Para cualquier cardinal $\lambda > \aleph_0$, no existe una estructura $\mathfrak{A}$ cuya cardinalidad sea λ y que satisfaga a AR.*

En general, las pruebas funcionan con todo rigor y podemos ver que extender el lenguaje *L1K=* implica perder la mayoría de las propiedades metatéoricas que habíamos demostrado anteriormente.

1.2.4.2. Metateoremas para la *semántica de Henkin* y la *semántica de primer orden*

Ahora expondré muy brevemente algunos resultados metateóricos para el caso de estas semánticas.

Sea $\mathfrak{A}^1 = \langle d^1, d^1_1, d^1_2, \langle I^1, p^1, a^1\rangle\rangle$ una estructura de primer orden, se puede encontrar una estructura de Henkin $\mathfrak{A}^H$ con $d^1_1 = D^H$ y $d^1_2 = F^H$, y p^1 como la predicación natural y a^1 es la relación de aplicación natural (por supuesto, d^H e d^1 son los mismos). Por esto:

Teorema 1.17. ***(Teorema 1-H)***

a) Una fórmula ϕ es una verdad lógica de Henkin si y solo si es una verdad lógica en primer orden.

b) Una fórmula ϕ es una consecuencia lógica de Henkin de Γ si y solo si ϕ es una consecuencia lógica de primer orden de Γ

c) Una fórmula ϕ es satisfacible en la semántica de primer orden si y solo si es satisfacible en la semántica de Henkin.

Esto nos ayuda a considerar a ambas semánticas, básicamente idénticas, sobre la base de que son simplemente equivalentes.

Consideremos la siguiente estructura de Henkin $\mathfrak{A}$: d^H = {a,b}, $D^H(2)$ = {⟨ x, a⟩| x $\in d^H$}, $F^H(1) = \{f(x) = x$, para todo $x \in d^H\}$. Consideremos la siguiente instancia del axioma de comprehensión: $\exists X \forall x \forall y (X(x, y) \equiv x \neq x)$. Esta fórmula afirma que existe una relación binaria vacía, sin embargo, nuestro modelo no incluye dicha relación, por lo cual la fórmula es falsa. El axioma de comprehensión no es universalmente válido en la semántica de Henkin. El mismo modelo hace falso el axioma de elección en la siguiente forma: $\forall X(\forall x \exists y X(x, y) \supset \exists F \forall x X(x, F(x)))$. La única relación que hay en $D^H(2)$ satisface el antecedente del condicional sin embargo la única función en $F^H(1)$ aplicada a b, sería f(b)=b. Pero ⟨b, b⟩ $\notin X$, y siendo esta la única función posible, hace falso al consecuente, haciendo falso el axioma de elección. Este razonamiento muestra que las semánticas de Henkin son incorrectas para el sistema *D2*, y por el teorema 1.17, también la semántica de primer orden lo es.

Definición. *Una estructura de Henkin (de primer orden) es fiel a D2 si y solo si satisface el axioma de elección y todos las instancias del axioma de comprehensión.*

Ahora se puede garantizar lo siguiente, (tanto para la *semántica de primer orden* como la *de Henkin*)

Teorema 1.18. *Si* $\Gamma \vdash \varphi$ *entonces toda estructura de Henkin que es fiel a D2 que satisface a todos los elementos de* Γ *satisfacen a* φ

Teorema 1.19. *Para todo conjunto* Γ *consistente, hay una estructura de Henkin fiel que satisface a todos los elementos de* Γ

Es decir, la semántica de Henkin (y por lo tanto, la semántica de primer orden) son completas y correctas. Ambas pruebas son muy parecidas a sus correspondientes pruebas para la lógica de primer orden.

Llamaremos a una estructura de Henkin (de primer orden) *identidad estándar* si y solo si para toda S, $\vDash_{\mathfrak{A}^H} \forall X(Xx \equiv Xy)$ $[S]$,$\bar{S}(x) \neq \bar{S}(y)$. La definición es análoga para el caso de la semántica de primer orden.

Teorema 1.20. *Si cada subconjunto finito* $\Gamma_0 \subseteq \Gamma$ *es satisfacible en una estructura de Henkin (de primer orden) fiel e identidad-estándar, entonces* Γ *es satisfacible en una estructura de Henkin (de primer orden) fiel e identidad-estándar*

Teorema 1.21.

a) Sea Γ *un conjunto satisfacible en una estructura de Henkin (de primer orden) de fórmulas en un lenguaje de cardinalidad* λ *entonces* Γ *es satisfacible en alguna estructura de Henkin (de primer orden) de tamaño* $\leq \lambda$

b) Sea Σ *un conjunto de enunciados de un lenguaje de cardinalidad* λ*, si* Σ *tiene modelo de Henkin (de primer orden) entonces* Σ *tiene un modelo de Henkin (de primer orden) de cardinalidad* $\leq \lambda$

Teorema 1.22. *Sea* Γ *un conjunto de fórmulas en un lenguaje de cardinalidad* λ *y supongamos* Γ *es satisfacible en alguna estructura de Henkin (de primer orden) fiel e identidad- estándar infinita. Entonces, para todo cardinal* $\kappa \geq \lambda$*, hay una estructura de Henkin (de primer orden) fiel e identidad- estándar de cardinalidad* κ *en la que* Γ *es satisfacible*

Todo esto nos muestra que tanto la semántica de Henkin como la semántica de primer orden no cumplen con los teoremas de categoricidad, pues son incompatibles con los teoremas de Lowenheim-Skolem, como vimos en el caso de la lógica de segundo orden. También parecen insinuar que ambas semánticas no ofrecen mayor expresividad que la lógica de primer orden, y que en términos metateóricos, son esencialmente lo mismo que la lógica de primer orden.

1.3. ¿Porqué elegir la semántica estándar?

Hemos dado un lenguaje para la lógica de segundo orden, *L2K*, para el cual dimos un sistema deductivo *D2*. Dimos también tres distintas semánticas para ese lenguaje de las cuales sólo la semántica estándar es inherentemente incompleta. Sólo en el contexto de la semántica estándar tiene sentido preguntarse si el argumento de Kreisel se aplica o no a la lógica de segundo orden, pues se puede demostrar el teorema de compleción para las otras semánticas, lo cual permitiría probar el Principio de Kreisel. Entonces, ¿porqué utilizar la semántica estándar?

La semántica estándar tiene una gran "capacidad expresiva". Ésta puede caracterizar cardinales transfinitos, es decir, para un cardinal transfinito λ, hay una fórmula de segundo que es satisfecha únicamente por estructuras cuyo dominio tiene cardinalidad λ. Un ejemplo de este tipo de fórmulas es el axioma de inducción de AR, que es satisfecho sólo por estructuras cuyo dominio tiene cardinalidad $\aleph_0$. En contraste, la lógica de primer orden sólo puede caracterizar cardinales finitos. De la misma forma, la lógica de segundo orden, puede caracterizar varios conceptos matemáticos importantes: buen orden, clausura minimal o finitud. De hecho, la demostración de los metateoremas para la semántica estándar hacen uso esencial de la capacidad expresiva de la semántica estándar. Veamos cómo

> Sea P un conjunto, una estructura $\mathfrak{A}$ y una $s \in \Sigma^{\mathfrak{A}}$. Decimos que P es $\mathfrak{A}$-*definible en primer orden* si y solo si, $\mathrm{P} \subseteq D^{\mathfrak{A}}$ y hay una fórmula ϕ de primer orden tal que, para todo $d \in D^{\mathfrak{A}}$, $\vDash_{\mathfrak{A}} \phi[s\,(x|d)]$ si y solo si $d \in \mathrm{P}$

Es decir, los elementos de P son exactamente aquellos elementos de $D^{\mathfrak{A}}$ que satisfacen la fórmula ϕ. Habíamos dicho que la aritmética de Peano en segundo orden (AR) como el conjunto de los cuatro axiomas de la aritmética, donde el último, el axioma de inducción era una fórmula pura de segundo orden, mientras que los demás axiomas eran fórmulas de primer orden. Podríamos describir la aritmética de Peano en primer orden (*AR1*) como el conjunto de los tres primeros axiomas de AR, y el siguiente esquema de inducción:

> **(Esquema de inducción)** Sea ϕ un fórmula de *L1K=*, $(\phi(0) \wedge \forall x(\phi(x) \supset \phi(sx)))$ es un axioma de *AR1*

El axioma de inducción en AR consideraba todos los subconjuntos del dominio de una estructura, la pregunta es ¿qué subconjuntos considera el esquema de inducción? La respuesta es sencilla: considera sólo los subconjuntos $\mathfrak{A}$-*definibles en primer orden* para la estructura $\mathfrak{A}$. Sólo sobre estos subconjuntos se puede aplicar el esquema de inducción. En la demostración del teorema de categoricidad, el axioma de inducción se aplicó a conjuntos que de hecho no son $\mathfrak{A}$-*definible en primer orden.* Para ejemplificar, tomaremos el conjunto P que se definió para el primer paso del teorema:

$$\mathrm{P} = \{a \in D^{\mathfrak{A}} |\ \text{hay un } b \in D^{\mathfrak{B}} \text{ tal que } \langle a, b\rangle \in f\}$$

Dicho conjunto no es $\mathfrak{A}$-*definible en primer orden* pues depende de la definición de la función *f*. Dicha función la definimos como la *clausura minimal* de $\langle 0_{\mathfrak{A}}, 0_{\mathfrak{B}} \rangle$ bajo la operación que genera $\langle \underline{s}_{\mathfrak{A}}, \underline{s}_{\mathfrak{B}} \rangle$ aplicada a $\langle a, b \rangle$. Intuitivamente, la clausura minimal de un conjunto A bajo cierta(s) operación(es) es el mínimo conjunto M que incluye al conjunto A y es cerrado bajo la(s) operación(es). Es decir, es tal que para cualquier elemento x de C, C contiene al resultado de aplicar a x las operaciones. A f se le definió precisamente como una *clausura minimal*, sin embargo, dicha noción no es caracterizable en primer orden. Por lo tanto, el conjunto P tampoco lo es. Lo mismo sucede con los demás conjuntos que usan en la prueba del teorema de categoricidad para AN. Así pues, no es posible usar el esquema de inducción para demostrar el teorema de categoricidad de AR. Este problema no se presenta para usar el axioma de inducción, precisamente porque éste considera todos los subconjuntos de dominio, sin importar si son definibles o no.

En la refutación del teorema de compleción, el paso crucial apelaba al teorema de categoricidad para la aritmética. En efecto, que el sistema deductivo D sea efectivo (así como el teorema de Gödel) no es materia de discusión de ninguna clase. La refutación del teorema de compleción es un corolario del teorema de categoricidad. Tanto la semántica de Henkin como la de primer orden están construidas precisamente para incluir "más" estructuras y refutar el teorema de categoricidad, y así demostrar el teorema de compleción. Es decir, están diseñadas para dar una semántica para la lógica de segundo orden que no tenga la "capacidad expresiva" de la semántica estándar. ¿Hay algún problema con este objetivo?

Rossberg (2004) retoma una propiedad *preteórica* de la lógica: la *formalidad*. Para Rossberg, la noción de sistema formal debe ser tal que "incluya cualquier sistema axiomático [...]; en particular incluye la teoría de modelos" (Rossberg 2004: 308). El punto clave para Rossberg es considerar la consecuencia deductiva y la consecuencia semántica como definiciones igualmente adecuadas para capturar una noción preteórica de consecuencia. Desde una posición como esta, es posible defender un proyecto que rechace los teoremas de categoricidad y demuestre compleción. De la misma forma, es posible defender una posición que opte por demostrar los teoremas de categoricidad, restando importancia a la falsedad del teorema de compleción. Rossberg sostiene que desde esta posición "se deben dar argumentos independientes sobre cuál de los dos sistemas es el que captura la noción preteórica, si alguna lo hace" (Rossberg 2004: 316). Desde una posición como la de Rossberg, no parece haber ningún problema con los objetivos de la semántica de Henkin. El teorema de compleción es una propiedad importante de nuestro sistema deductivo, así como de una definición de consecuencia semántica. Es claro que, ante la ausencia del teorema de compleción, se debe ofrecer una posible solución, entre las cuales podría ser optar por una semántica más débil. La condición que se exigiría es, precisamente, que dicha semántica no valide los teoremas de categoricidad. Por supuesto, en esta perspectiva se estaría optando por restar importancia a una noción de consecuencia semántica, para rescatar una noción de consecuencia deductiva. Siguiendo a Rossberg, un proyecto así debería ofrecer razones independientes para optar por una noción de consecuencia deductiva.

Análogamente, aunque el teorema de categoricidad es un resultado que parece en principio deseable, el defensor de la semántica de Henkin (o de la semántica de primer orden) podría optar por considerar el teorema poco relevante. Meadows (2013: 526) señala que existen tres objetivos que puede buscar el filósofo de la lógica para defender un interés en los teoremas de categoricidad:

1. Demostrar que hay una única estructura que corresponde a una intuición o práctica matemática
2. Demostrar que una teoría se refiere a una única estructura
3. Clasificar diferentes tipos de teorías

Alguien podría defender que algunos o todos los objetivos no son objetivos de la lógica, aunque sería deseable tener una semántica para la lógica de segundo orden. Otra alternativa sería mostrar que con las semánticas no estándar se pueden obtener resultados similares para el caso de la lógica de primer orden. Sin embargo, esto no es muy prometedor: el propio Meadows (2013: 529) utiliza la *semántica de primer orden* para demostrar una versión del teorema de categoricidad para la aritmética en primer orden. La idea es tomar la aritmética en primer orden, y utilizar la aritmética en segundo orden con la *semántica de primer orden* como metateoría.[23] El problema es que la metateoría no es categórica, y por lo tanto, necesitaríamos demostrar que lo es para garantizar el resultado. Esto es posible utilizando una aritmética en tercer orden con la misma semántica, pero tiene exactamente el mismo problema de hecho ninguna de estas teorías será categórica y por lo tanto "jamás alcanzaremos un punto en el cual tengamos una teoría que podamos mostrar que es absolutamente categórica" (Meadows 2013: 532). Un defensor de esta posibilidad podría argumentar que no es necesaria una teoría absolutamente categórica, aunque creo que dicha posición es poco satisfactoria.[24] Aún cuando no sea posible garantizar los teoremas de categoricidad, me parece que, en principio, no hay porqué rechazar una semántica porque no los cumpla. Finalmente, no es obvio que los teoremas de categoricidad sean objetivos fundamentales de la lógica.

Sin embargo, los motivos para no adoptar la semántica de Henkin ni la de primer orden son otros. Aunque es cierto que no hay razón para rechazar inmediatamente una semántica de segundo orden por el hecho de no demostrar los teoremas de categoricidad, también es cierto que la semántica estándar proporciona una semántica "natural" para la lógica de segundo orden. La semántica estándar para la lógica de segundo orden es "natural" en el sentido de que es una extensión de la semántica para la lógica de primer orden. De hecho, ni siquiera es necesario dar una definición diferente de estructura y considerar como interpretaciones de *L2K* a las mismas estructuras que consideramos interpretaciones para el caso de *L1K=*. Como vimos, los

[23] En dicha metateoría necesitamos el esquema de comprensión, por lo cual, sus modelos deberán ser modelos fieles de *primer orden*

[24] Podemos ver que si consideramos a todas las aritméticas de n-orden, (que sirven como metateorias para demostrar la categoricidad de las aritméticas de n-1-orden) y formamos una nueva teoría como la unión de todas ellas, esta nueva teoría seguirá sin ser categórica, pues la semántica de primer orden es inherentemente no categórica. Sin embargo, una impresión inicial es que las pruebas de categoricidad efectivamente muestran que la teoría base es categórica, por lo cual, sería más adecuado optar por una teoría que muestre que ésta es categórica. Cualquier otra opción parece insatisfactoria.

requerimientos para proporcionar una semántica para la lógica de segundo orden, a partir de la semántica de primer orden son mínimos, y ofrecen una mayor simplicidad. Esto es claro, en vista de que, para las semánticas no estándar, la cuantificación parece ser no lógica, en el sentido de que es necesario seleccionar el rango de las variables de segundo orden, de manera análoga a como se selecciona la extensión de un predicado no lógico para una estructura cualquiera: la frase "todas los propiedades" en las semánticas no estándar, parece significar "todas las propiedades escogidas".

Hasta ahora, no he dado razones para rechazar las semánticas no estándar. Finalmente, el defensor del teorema de compleción podría insistir en que la "naturalidad" de la semántica estándar no es un aspecto importante. Aun así, creo que las semánticas expuestas reflejan problemas más profundos. Bueno (2010), señaló que hay criterios para rechazar estas semánticas: "Primero, nótese que si tomamos todos los modelos de Henkin para proporcionar una semántica de la lógica de segundo orden, la semántica resultante ni siquiera es correcta, pues como vimos, el esquema de comprehensión para esta lógica no es válido" (Bueno 2010: 372). Como vimos, ninguna de las dos semánticas era una semántica correcta para *D2*. Había que imponer el requisito de que las estructuras fuesen fieles para *D2*, para que fuesen correctas. Sin embargo, continúa Bueno: "es *ad hoc* imponer estas restricciones sobre los modelos de segundo orden [. . .] pues la razón por la cual esas restricciones son asumidas es simplemente proporcionar una semántica para la lógica de segundo orden que sea correcta y completa"[25] (Bueno 2010: 372). La idea de las semánticas no estándar es proporcionar una semántica para la lógica de segundo orden que sea correcta y completa, por lo cual, el reconocer la importancia de la corrección del sistema *D2* obliga a implementar un mecanismo *ad hoc*, pues la cláusula de *fidelidad* pide precisamente que valide las instancias del esquema de comprehensión y del axioma de elección.

A mi manera de ver, el problema de estas semánticas no estándar es que identificaron adecuadamente el problema de la refutación del teorema de compleción y optaron por una manera un tanto inadecuada de solucionarlo. Recordemos que su solución fue agregar ciertas interpretaciones que no validaran los teoremas de categoricidad, en particular, se agregaron aquellas que no los validan en una aritmética de primer orden. Esta estrategia garantiza que se validan los axiomas de un sistema deductivo de primer orden (por ejemplo, *D1=*), sin embargo, para el esquema de comprehensión y el axioma de elección, la estrategia no garantiza su validez. A final de cuentas, hay una estructura de la semántica de primer orden muy parecido a la estructura de Henkin considerada en la sección 1.2.4.2 que invalida el esquema de comprehensión y el axioma de elección.[26] Siguiendo la analogía anterior, no es suficiente con "escoger" las propiedades arbitrariamente, es necesario "escoger" las adecuadas. Me parece que esta clase de estrategias requiere, en muchas ocasiones, mecanismos *ad hoc*.

[25]El argumento de Bueno concluye que es necesario dar una razón independiente para aceptar los modelos de Henkin. Y, si dicha razón es porque es necesario garantizar el teorema de compleción, entonces el defensor de la semántica de Henkin pide la cuestión contra el defensor de la semántica estándar que sostiene que no es así (cf. Bueno 2010: 372).

[26]La estructura $\mathfrak{A}$, que asigna a todas las predicados binarios la misma relación asignada a $D(2)$ y a todas las funciones de un argumento la función asignada a $F(1)$. En esta estructura, todas las oraciones de la forma $\forall x \forall y P_2^n(x, y) \equiv x \neq x$ y las de la forma $\forall x \exists y P_2^n \supset \forall x P_2^n(x, f(x))$ son falsas.

Todas estas observaciones muestran que las semánticas no estándar no son semánticas adecuadas para la lógica de segundo orden. Es posible que exista alguna aplicación para ellas, aunque de momento no puedo imaginar alguna. En cualquier caso, la semántica estándar conserva el atractivo independiente de contar con una "capacidad expresiva" inmensa, y para algunos, eso parece suficiente. En el resto del trabajo, optaremos por usar una semántica estándar para la lógica de segundo orden y veremos que, aunque nos provoque una gran cantidad de problemas, refleja una gran cantidad de propiedades metateóricas interesantes.

1.4. El teorema del homomorfismo

Recordemos que la semántica para la lógica de segundo orden es una extensión de la semántica para la lógica de primer orden. A partir de ahora, reflejaremos esta idea y llamaremos *semántica tarskiana*[27] al aparato para generar ambas semánticas. Se puede notar que es posible dar una extensión para una lógica de tercer orden, con los ajustes necesarios y análogos a los ofrecidos en la semántica de segundo orden, esta semántica también sería un caso de *semántica tarskiana*. La semántica tarskiana cumple una propiedad muy interesante, para cualquier lenguaje de cualquier orden: el teorema del *homomorfismo*.[28] Dado nuestro interés, sólo hablaremos del caso para el lenguaje, *L2K*. Veamos en qué consiste:

Definición. *Sean* $\mathfrak{A}$ *y* $\mathfrak{B}$ *dos estructuras para L2K. Un homomorfismo de* $\mathfrak{A}$ *en* $\mathfrak{B}$ *es una función h:* $D^{\mathfrak{A}} \to D^{\mathfrak{B}}$*,tal que:*

1) *Para cada* P^n_m *símbolo de predicado y cada* $\langle a_1, \ldots, a_m\rangle$ *tal que* $a_1, \ldots, a_m \in D^{\mathfrak{A}}$
$\langle a_1, \ldots, a_m\rangle \in P^{n\mathfrak{A}}_m$ *si y solo si* $\langle h(a_1), \ldots, h(a_m)\rangle \in P^{n\mathfrak{B}}_m$

2) *Para cada* f^n_m *símbolo de función y cada* $\langle a_1, \ldots, a_m\rangle$ *tal que* $\langle a_1, \ldots, a_m\rangle \in P^{n\mathfrak{A}}_m$
$h(f^{n\mathfrak{A}}_m(a_1, \ldots, a_m)) = f^{n\mathfrak{B}}_m(h(a_1), \ldots, h(a_m))$

3) *Para cada* a_n*, símbolo de constante,* $h(a^{\mathfrak{A}}_n) = a^{\mathfrak{B}}_n$

De manera intuitiva, un homomorfismo es una función que "preserva" la estructura, en el sentido de que preserva las funciones de asignación de las estructuras, en la definición, h preserva $I^{\mathfrak{A}}$. Hay que notar que como el conjunto K no es diferente en *L1K=* y *L2K*, no es necesario definir más cláusulas. Y que no se exige nada de h, de hecho, h podría ser una función inyectiva o suprayectiva, o ambas o ninguna. Si h es inyectiva, llamaremos a h, *homomorfismo uno a uno*. Si h es suprayectiva, llamaremos a h, *homomorfismo sobreyectivo*. Si h es biyectiva, llamaremos a h, *homomorfismo isomórfico*.

Teorema 1.23. ***(Teorema del homomorfismo)****: Sea h un homomorfismo de* $\mathfrak{A}$ *en* $\mathfrak{B}$*, y sea* $S \in \Sigma^{\mathfrak{A}}$*.*

a) *Para todo término t,* $h(\bar{S}(t)) = \overline{(h \circ S)}(t)$

[27] Esto debido a que ambas están inspiradas en el mecanismo ofrecido por Tarski (1931) "The concept of truth in formalized languages".

[28] Utilizando el nombre que Enderton le da al teorema.

b) Para toda fórmula φ, que no tenga cuantificadores y no tenga el símbolo de igualdad
$\vDash_{\mathfrak{A}} \varphi[S]$ *si y solo si* $\vDash_{\mathfrak{B}} \varphi[h \circ S]$

c) Si h es un homomorfismo uno a uno, entonces para toda fórmula de la forma $(t_1 = t_2)$
$\vDash_{\mathfrak{A}} (t_1 = t_2)[S]$ *si y solo si* $\vDash_{\mathfrak{B}} (t_1 = t_2)[h \circ S]$

d) Si h es un homomorfismo sobreyectivo, entonces para toda fórmula de la forma $\forall x_k \alpha$*, de la forma* $\forall X^n_m \psi$ *y de la forma* $\forall F^n_m \psi$
$\vDash_{\mathfrak{A}} \forall x_k \alpha[S]$ *si y solo si* $\vDash_{\mathfrak{B}} \forall x_k \alpha[h \circ S]$
$\vDash_{\mathfrak{A}} \forall X^n_m \psi[S]$ *si y solo si* $\vDash_{\mathfrak{B}} \forall X^n_m \psi[h \circ S]$
$\vDash_{\mathfrak{A}} \forall F^n_m \psi[S]$ *si y solo si* $\vDash_{\mathfrak{B}} \forall F^n_m \psi[h \circ S]$

La prueba se realiza por inducción sobre el número de funciones (en el caso del inciso a)) y sobre el número de conectivas y cuantificadores (en el caso del inciso b) y d)). El inciso a) afirma que para toda $S \in \Sigma^A$ se puede encontrar una función $S \in \Sigma^{\mathfrak{B}}$ (a saber, $h \circ S$) que preserva las interpretaciones asignadas a los términos por la extensión $\bar{S}$ de S. Por otro lado, el inciso b) afirma que un homomorfismo implica que las dos estructuras relacionadas satisfacen las mismas fórmulas (que no tengas cuantificadores ni el símbolo de igualdad). La necesidad de que h sea un homomorfismo uno a uno en el inciso c) es porque si hubiese una igualdad en la estructura $\mathfrak{B}$, entonces ésta podría ser falsa en la estructura $\mathfrak{A}$. Supongamos que hay un homomorfismo h entre dos estructuras $\mathfrak{A}$ y $\mathfrak{B}$, tal que no es *uno a uno*. Sea una $S \in \Sigma^{\mathfrak{A}}$, tal que para dos términos t_1 y t_2, $\bar{S}(t_1){=}a$ y $\bar{S}(t_2){=}b$, para dos a, b $\in D^A$ diferentes. Supongamos también que h es tal que $h(a){=}c{=}h(b)$, para algún $c \in D^{\mathfrak{B}}$. Por lo tanto, $h(\bar{S}(t_1)){=}\overline{h \circ S}(t_1) = c = \overline{h \circ S}(t_2){=}h(\bar{S}(t_2))$, por lo tanto $\vDash_{\mathfrak{B}} (t_1 = t_2)[h \circ S]$ pero $\nvDash_{\mathfrak{A}} (t_1 = t_2)[S]$. Algo similar sucede en el caso de d), si h no fuese sobreyectiva, entonces las cuantificaciones universales podrían ser verdaderas en $\mathfrak{A}$ pero falsas en $\mathfrak{B}$.

El siguiente corolario lo usaremos más adelante:

Corolario 1.1. ***(Teorema de isomorfismo)*** *Si h es un homomorfismo isomórfico de* $\mathfrak{A}$ *en* $\mathfrak{B}$*, entonces* $\mathfrak{A}$ *y* $\mathfrak{B}$ *son equivalentes, es decir, para toda fórmula* φ*,* $\vDash_{\mathfrak{A}} \varphi$ *si y solo si* $\vDash_{\mathfrak{B}} \varphi$

Esto es natural pues un homomorfismo isomórfico establece básicamente que $I^{\mathfrak{A}}$ y $I^{\mathfrak{B}}$ son exactamente iguales (salvo "notación").

1.5. Conclusiones

He presentado un lenguaje de primer orden, *L1K=*, y expuse una semántica y un sistema deductivo para dicho lenguaje. Así mismo, expuse el lenguaje *L2K* para la lógica de segundo orden. Expandimos tanto la semántica como el sistema deductivo de *L1K=* para proporcionar una semántica (la *semántica estándar*) y un sistema deductivo para el lenguaje *L2K*. Vimos que el lenguaje *L1K=* cumple con los teoremas de compleción, corrección, compacidad y Lowenheim-Skolem. También vimos que *L2K* sólo cumple con el teorema de corrección, mientras que los demás teoremas fallan como consecuencia (directa) de los teoremas de categoricidad para la *aritmética de Peano* (AR) y el *análisis real* (AN). En la sección 1.3, explique que esto se debía a que el lenguaje *L2K* con la semántica estándar tenía una "capacidad expresiva" superior a la lógica de primer orden.

También expuse dos semánticas no estándar para la lógica de segundo orden: la *semántica de Henkin* y la *semántica de primer orden*. Vimos que ambas semánticas cumplen con los teoremas de compleción, corrección, compacidad y Lowenheim-Skolem. En la sección 1.3, argumenté que, por un lado, esto muestra que la capacidad expresiva de estas semánticas era significativamente inferior a la semántica estándar, y que esto no es necesariamente un problema para estas semánticas. Sin embargo, argumenté que había problemas independientes a eso en estas semánticas, y que estos estaban relacionados con algunos mecanismos *ad hoc* presentes en la construcción de la semántica. Por estas razones, en el resto del trabajo, adoptaré una semántica estándar para lógica de segundo orden. Así pues, dado que es en el contexto de la semántica estándar donde el problema del argumento de Kreisel tiene sentido, dicho problema sigue presente para el presente propósito y en los siguientes capítulos abordaremos ese problema en el contexto de los planteamientos presentados en este capítulo.

Capítulo 2

Validez *Preteórica*

Shapiro (1991, 1987) y Gómez Torrente (2000b, 2008b) utilizan versiones del argumento de Kreisel que equiparan la noción intuitiva de validez (o noción intuitiva de consecuencia lógica) con la noción de validez en estructuras cuyo dominio es una clase. En estas versiones, es fácil establecer que el Principio de Kreisel es verdadero para el caso de la lógica de primer orden. La presentación de dichas versiones se verá con precisión en el capítulo 3. Sin embargo, el propio Gómez Torrente (2000b) hace una observación importante:

> El interés del resultado depende en gran medida de que la noción intuitiva de consecuencia lógica sea analizable por medio de una de las nociones generalistas de consecuencia lógica expuestas en este capítulo. Si no lo es- si, por ejemplo, tenemos razones para pensar que la [noción de validez en estructuras cuyo dominio es una clase] no es una condición suficiente de la consecuencia lógica a secas-, entonces tendremos razones para pensar que no hemos mostrado (ni siquiera para los lenguajes de primer orden) que la noción de [consecuencia lógica tarskiana][29] es correcta con respecto a la noción intuitiva de consecuencia lógica. (Gómez Torrente, 2000b: 65)

En el presente capítulo, me propongo explicar la relación que existe entre ambas nociones. Para ello utilizaré dos propiedades presentadas por el propio Gómez Torrente (2000b) de la noción *preteórica* de validez (o consecuencia lógica): la *formalidad* y la *necesidad*. Para ello hay un problema importante: ¿qué naturaleza tiene la noción de validez *preteórica*? Esto podría ser un problema poco relevante si se acepta que dicha noción es poco significativa para la lógica. En la primera sección de este capítulo me propongo explicar qué entiendo por la noción de consecuencia lógica intuitiva. Después de explicar esto, presentaré una breve discusión de las propiedades de *formalidad* y *necesidad* para explicar qué relación existe entre una noción *preteórica* de consecuencia lógica y la noción de *validez en estructuras cuyo dominio es una clase*.

[29]El texto entre corchetes en esta cita es mío.

2.1. Validez *preteórica* (la noción *preteórica* de consecuencia lógica)

En esta primera sección me ocuparé de hablar un poco sobre lo que entiendo por la noción preteórica de consecuencia lógica. Esta noción ha sido considerada dudosa y problemática por ser vaga y confusa. Peter Smith, en su discusión del argumento de Kreisel, sostiene que "[...]no hay una noción preteórica intuitiva de consecuencia válida suficientemente precisa para que dicho argumento se sostenga" (Smith 2011: 26) Smith presiona en este punto afirmando que si alguien cree que dicha noción existe debe responder preguntas sobre si la noción intuitiva de consecuencia es analítica o si está limitada por restricciones de relevancia (cf. Smith 2011: 27), y estas preguntas no tienen una respuesta determinada. Sin embargo, muchas veces en los cursos introductorios sobre lógica se suele apelar a ciertas *intuiciones inferenciales* para motivar definiciones o algún otro contenido relevante del curso. Haré algunas precisiones para aclarar un poco más la clase objetivos e intereses que consideraré en este capítulo. A mi modo de ver, cuando se usa o se apela a esta noción, se busca apelar a ciertas ideas *preformales* o *preteóricas* sobre las inferencias lógicas. No creo que dichas intuiciones o ideas preformales estén fijas. Rossberg señala que "podría ser el caso que el estudio de los sistemas formales afecten nuestra intuición preteorética" (Rossberg 2004: 316). Hay algunos ejemplos de cómo las intuiciones sobre las inferencias lógicas cambian (como el caso de algunas inferencias aristótelicas o las inferencias analíticas[30]), e incluso estas modificaciones son resultado de usar algunos sistemas lógicos. A partir de esto, vale la pena señalar que las intuiciones lógicas no son ajenas a la práctica lógica misma y la tradición en la que está inscrita dicha práctica. MacFarlane apunta de manera acertada que "nuestras intuiciones sobre logicidad no son un tipo de realidad extramental: estas son artefactos históricos, un producto de nuestra *educación* lógica y filosófica" (McFarlane 2000: 28).

Con esto en mente, habrá que aclarar que la noción *preteórica* que discuto se aplica a los lenguajes formales de la lógica clásica y al lenguaje informal de la matemática clásica. El lenguaje de la matemática puede ser formalizado (en principio) con algo de facilidad y la razón de esto es porque tienen en común dos características mencionadas por Jané:

> si queremos que algunas fórmulas de nuestros lenguajes cuenten como contrapartes adecuadas de algunas oraciones matemáticas informales, entonces estas fórmulas deben tener al menos un significado rudimentario: su terminología lógica debe ser significativa, y las categorías semánticas de la terminología no lógica deben estar delimitadas (Jané 1993: 68-69)

Una ventaja de los lenguajes formales y el lenguaje (informal) de la matemática

[30]Por ejemplo, inferencias como la de 'Martín es soltero' a 'Martín es no casado' se solían considerar inferencias lógicamente correcta, lo cual no es el caso desde el punto de vista de la lógica de primer orden (como caso particular, la lógica definida en la sección 1.1 del capítulo anterior). Lo mismo sucede con la inferencia de 'Algún hombre es mortal' a partir de la oración 'Todos los hombres son mortales'. A pesar de ello, no se considera a la lógica de primer orden como un sistema lógico inadecuado.

clásica es que cumple con ambos criterios. A mi forma de ver, estos criterios precisan el contexto en que se puede evaluar la corrección de la noción de validez *preteórica* que considero y permiten acotar el tipo de inferencias que son de interés en este contexto. Cabe mencionar que a lo largo de este capítulo usaré principalmente ejemplos no formalizados y poco relacionados con la matemática, sin embargo, esto sólo será por motivos de claridad. Considerando estos lenguajes, se puede pensar que aunque se logró un avance en lo que a precisión se refiere, se pierde interés en tanto que no hay realmente *intuiciones preteóricas* en lenguajes como estos. Sin embargo, hay filósofos como Jané quién piensan que hay justificaciones *informales* para las reglas formales de inferencia tales como la eliminación del cuantificador universal[31], que dependen sólo de atribuir significado a la terminología lógica. Tarski mismo proporciona un ejemplo de esta clase de inferencias informalmente justificadas cuando considera la ω-regla: de todas las oraciones $\varphi(1)$, $\varphi(2)$, $\varphi(3)$,... se puede inferir que $\forall x\varphi(x)$, donde el rango del cuantificador es el conjunto de los números naturales (cf. Tarski 1936: 410 - 11). Es en este sentido que es posible atribuir validez *preteórica* o informal a las reglas de inferencia de un lenguaje preciso.

Vale la pena dar algunos ejemplos históricos de teorías de la consecuencia lógica para aclarar un poco más el contexto histórico en que está inmersa lo que entenderé como la noción *preteórica* de consecuencia lógica. Presentaré tres autores importantes para esta noción *preteórica* de consecuencia lógica: Aristóteles, Bolzano y Frege. Señalaré algunos puntos sobre el trabajo de cada uno en torno a la noción de consecuencia lógica que considero relevantes para la noción *preteórica* que trato de rescatar. En particular, me concentraré en tres puntos: en la naturaleza de la definición de consecuencia que proponen, el tipo de inferencias que se buscan rescatar y el lenguaje que se considera.

Aristóteles es considerado en la literatura como el primer filósofo en hablar de la lógica como disciplina y se concentraba en el uso de ésta en la argumentación racional. En el tema de la noción de consecuencia, él identifico algunas *estructuras inferenciales* como inferencias válidas, lo que Corcoran (2009: 11) llama 'Teoría de las deducciones categóricas de Aristóteles'. Para Aristóteles, el método de la deducción permitía dar una teoría sobre lo que hoy llamamos consecuencia lógica. Una *deducción* es "un discurso en el cual, dadas ciertas cosas por supuesto, algo diferente de las cosas supuestas resulta por necesidad porque dichas cosas son el caso" (*Prior Analytics*: 24b20). Con mayor precisión, una deducción es una cadena de razonamientos tales que éstos están justificados por las estructuras inferenciales o por reglas derivadas de la estructuras inferenciales. Por lo cual, "toda deducción [...] tiene una cadena de razonamientos que muestran o hacen evidente que la conclusion (final) se sigue lógicamente de las premisas-y así que la afirmación de las premisas es también una afirmación virtual de la conclusión" (Corcoran 2009: pág. 8). Hay que mencionar que el interés de Aristóteles eran las deducciones que representaban *demostraciones* de la conclusión, es decir, deducciones donde se sabía (o se podía saber) que las premisas son verdaderas. Es por este interés en las demostraciones

[31]Es decir, la regla que infiere φ(a) de $\forall x\varphi(x)$, si a está en el rango del cuantificador. Jané atribuye propiedades informales a la relación de consecuencia lógica (como la de ser neutral al contenido de las oraciones), en vista de que la justificación de esta regla depende sólo del significado del cuantificador universal (cf. Jané 1993: 67)

que, para Aristóteles, las demostraciones agregaban información nueva al conjunto de creencias y por lo cual, inferencias del tipo, *P por lo tanto P*, no son casos de consecuencia, de otra manera no se demostraría en realidad. La importancia histórica de Aristóteles, además de la identificación de ciertas estructuras inferenciales, es que "Como Tarski enfatizó, una prueba formal en el sentido moderno resulta del refinamiento y 'formalización' de la demostración tradicional de Aristóteles" (Corcoran 2009: 2). Cabe mencionar que se atribuye a Aristóteles el uso de contraejemplos para mostrar que un argumento es inválido, pues esto es suficiente para mostrar que la conclusión de un argumento no se sigue por *necesidad* de las premisas (cf. Asmus y Restall 2012: 19).[32]

Por otro lado, a Bernard Bolzano se le atribuye en los últimos años el *descubrimiento* de la definición semántica de consecuencia lógica (eg. van Benthem 1985). Bolzano fue el primero en considerar que una oración A es una *verdad lógica* si todas las oraciones resultantes de A al sustituir uniformemente los conceptos no lógicos de A por otros, son verdaderas. Esto se puede generalizar a la noción de consecuencia. Esta definición es muy cercana a la definición de Tarski. Como Gómez Torrente (2002: 6) nota, se le puede atribuir a Bolzano también la idea de que las oraciones que son verdades lógicas, lo son en virtud de su forma. Para poder proporcionar esta definición, Bolzano defiende un análisis de las proposiciones donde ciertas partículas (los conceptos no lógicos) varían mientras que otras no (los conceptos lógicos), y es mediante este análisis que la definición de consecuencia puede recuperar la noción de "seguirse por necesidad" (cf. Asmus y Restall 2012: 31). La definición de Bolzano de consecuencia lógica se encuentra en su noción de derivabilidad:

> digo que las proposiciones M, N, O,... serían *derivables* de las proposiciones A, B, C, D,... con respecto a las variables i, j,..., si todo conjunto de *ideas*[33] que hacen a A, B, C, D,... todas verdaderas cuando son sustituidas por i, j,..., también hacen a M, N, O,... todas verdaderas. [...] a veces también diré que las proposiciones M, N, O,... se siguen de o pueden ser inferidas o se concluyen del conjunto de proposiones A, B, C, D,... (Bolzano 1973: §155, 205)

Las ideas i, j,..., son aquellos conceptos variables, los conceptos no lógicos. Ahota bien, van Benthem afirma que, a partir de la teoría de Bolzano, "dado un conjunto de proposiciones, uno puede considerar cualquier conjunto de sus constituyentes (palabras, frases) y considerar 'variantes' obtenidas sustituyendo adecuada y uniformemente por otros constituyentes para obtener nuevas proposiciones" (van Benthem 1985: 389). De esta idea, van Benthem reconstruye una noción de consecuencia que permite variar cualquier conjunto de partículas lingüísticas, generando diferentes definiciones de consecuencia dependiendo del conjunto de partículas lingüísticas que se elija. Así mismo, Bolzano definía nociones de consecuencia que no cumplían con

[32] Esto quiere decir que si un argumento A es inválido es suficiente encontrar un ejemplo de un argumento con la misma *forma* de A que tenga premisas verdaderas y conclusión falsa, para mostrar que dicho argumento no se sigue por *necesidad*. Estas dos propiedades son las que discutiré más adelante.

[33] Este énfasis es mío, no del autor.

ciertas propiedades de la definición tarskiana, tales como monotonicidad (cf. van Benthem 1985: 390).

Finalmente, hablaré un poco sobre Frege, considerado el padre de la lógica moderna. Hay que recordar que el principal objetivo de Frege era reducir las nociones de la aritmética a las nociones de la lógica, en contra de lo que sostenía Kant, que defendió que la aritmética se reducía a la *intuición* kantiana del tiempo. Para ello, Frege diseñó un sistema simbólico donde lograba expresar una gran cantidad de estructuras inferenciales, así como la relación de predicación. Dicho lenguaje es lo que hoy conocemos como un lenguaje de orden superior, del cual los lenguajes de primer orden son sólo una parte. Gómez Torrente (2000b: 23) señala que en este lenguaje siempre es completamente claro cuál es la forma lógica de un argumento del lenguaje. La noción de consecuencia lógica en el lenguaje de Frege es muy parecida a la de Aristóteles, en tanto que identificó una serie de reglas de inferencia y un conjunto de axiomas que son intuitivamente correctos y, por lo cual, todo teorema de estos serán intuitivamente correctos. Cabe mencionar que los axiomas y las reglas de Frege son *necesarios* (según el propio Frege) en tanto que son leyes que rigen el razonamiento de cualquier tipo: "Para él la lógica era la ciencia sobre la realidad más general de todas" (Kneale y Kneale 1962: 628). Al igual que el lenguaje diseñado por Frege, la relación de consecuencia lógica es *precisa*; de hecho, es básicamente la relación de consecuencia deductiva que se usa actualmente.

La definición de consecuencia lógica de Tarski está diseñada para aplicarse a lenguajes similares a los creados por Frege y en consideración de las inferencias que éste consideraba intuitivamente correctas. Así pues, la noción preteórica de consecuencia lógica que usaremos se inscribe en esta tradición, por lo cual, me ocuparé de las inferencias que ésta considera intuitivamente válidas y utilizaré conceptos como 'forma', 'lenguaje lógico' o 'deducción', atendiendo al tipo de inferencias que ellos consideraban. Ahora presentaré un análisis de la noción preteórica de consecuencia lógica.

2.2. Algunas propiedades de la noción de consecuencia lógica

En la presente sección señalaré dos propiedades que generalmente se le atribuyen a la noción preteórica de consecuencia lógica. Retomo esta manera de caracterizar la noción de consecuencia lógica del trabajo de Gómez Torrente (2000, 2003), aunque también se puede encontrar referencia a estas propiedades como propiedades de la noción de consecuencia lógica en (Asmus y Restall, 2012).

En los cursos introductorios de lógica se suele presentar la noción de validez indicando que un argumento es válido si, siempre que sus premisas sean verdaderas, su conclusión deberá serlo también (o alguna frase similar o equivalente). Con ello se busca motivar la intuición de que un argumento es válido si no es posible que las premisas sean verdaderas y la conclusión falsa. A partir de esta clase de caracterizaciones informales, se suele decir que la noción intuitiva de consecuencia lógica determina una relación especialmente estricta: nunca sucederá que las premisas sean verdaderas y la conclusión falsa. Si un argumento cumple con esta propiedad se suele decir que las premisas *implican por necesidad lógica* a la conclusión, y es en este sentido que se sostiene que la noción intuitiva de consecuencia lógica determina una relación *necesaria*. Que la noción preteórica de consecuencia lógica determine una relación necesaria, es lo que llamaré la propiedad *modal* de la noción de consecuencia lógica (o bien, la propiedad de *necesidad lógica*): si un argumento es intuitivamente válido, entonces las premisas implican a la conclusión por necesidad lógica.

Algunos ejemplos de argumentos que no satisfacen esta propiedad son:

(M1) Carlos estudió filosofía

Carlos no es racista

(M2) Humberto es un ser humano

Humberto no tiene tres piernas

(M1) no cumple la propiedad pues es posible que Carlos haya estudiado filosofía y sin embargo sea racista, por lo cual el argumento tendría premisas verdaderas y conclusión falsa. En (M2), aunque sabemos que generalmente los seres humanos tienen dos piernas, en principio es posible que los seres humanos hayan tenido tres piernas, en cuyo caso Humberto sería un ser humano (siendo la premisa verdadera) pero tendría tres piernas (y la conclusión sería falsa).

Sin embargo, los siguientes son casos de argumentos en los que las premisas implican por necesidad lógica a la conclusión:

(M3) Si la teoría de la rigidez de Kripke es verdadera entonces
los enunciados de identidad son necesarios
La teoría de la rigidez de Kripke es verdadera

Los enunciados de identidad son necesarios

(M4) Algunos suegros son prejuiciosos

Algunos hombres son prejuiciosos

El argumento (M3) cumple la propiedad de necesidad lógica pues si las premisas son verdaderas, la conclusión debe serlo. Mientras que (M4) satisface la propiedad debido a que es parte del significado del término 'suegro' el ser un hombre. Sin embargo, (M4) no es aceptado generalmente como un caso de consecuencia lógica, debido a que existen argumentos de la misma forma lógica que no cumplen con la propiedad modal de la relación de consecuencia lógica. En particular, hay casos de argumentos con la misma forma lógica que (M4) en los que la premisa es verdadera y la conclusión es falsa. Por ejemplo:

(M4') Algunos humanos son vertebrados

Algunos paramecios son vertebrados

El cual, casi por definición, tiene premisa verdadera y conclusión falsa.

La segunda propiedad que se le suele atribuir a la noción preteórica de consecuencia lógica es que ésta determine una relación formal. Se suele considerar que la validez de un argumento es una cuestión de forma, o bien que ésta dependa de la forma lógica del argumento. Así pues, se suele considerar que si un argumento es un caso de consecuencia lógica entonces cumple con la siguiente propiedad de *formalidad*: si un argumento es un caso de consecuencia lógica, entonces todos los argumentos con la misma forma lógica también son casos de consecuencia lógica. De este modo, (M4) no puede ser un caso consecuencia lógica, en contraste con el argumento (M3) que sí lo es pues cumple con la propiedad de formalidad. Todo esto depende de algunos presupuestos sobre la forma lógica de los argumentos, pero más adelante explicaré con mayor precisión a que me refiero al decir que dos argumentos tienen la misma forma lógica.

Se suele atribuir estas dos propiedades a la noción preteórica de consecuencia lógica y se considera que tanto la propiedad de *necesidad* como la de *formalidad* son condiciones necesarias, más no suficientes. No pretendo defender que estas propiedades son condiciones suficientes para la noción de consecuencia lógica (aunque me inclino a creer que es el caso). En las siguientes secciones, presentaré una breve discusión de ambas propiedades, a partir de la cual pretendo establecer una relación entre la noción intuitiva de consecuencia lógica y la noción de *validez en estructuras cuyo dominio es una clase.*

2.2.1. Formalidad

Es común encontrar en la literatura la afirmación de que la lógica (al menos, la lógica clásica contemporánea) es una cuestión de forma. Por lo general, se considera que la noción de consecuencia lógica es una noción que se explica con base en las propiedades formales de los argumentos.

Una razón para atribuir esta propiedad a la noción de consecuencia lógica es que se considera que la lógica debe ser indiferente al contenido de las oraciones, es decir, debe ser neutra con respecto al discurso particular en el que se enuncian los argumentos. Una idea común dentro de la práctica usual de la lógica es que la validez de un argumento no debe depender de aspectos particulares del tipo de discurso en que se enuncia, así como de los objetos de los cuales dicho argumento habla. Así mismo, se considera que apelar a la forma lógica de un argumento permite atender a esta clase de intuiciones. Se ha considerado esto por el hecho de que la forma lógica de una oración está íntimamente ligada con su estructura gramatical en cierto lenguaje, y por esto no está ligada al contexto de enunciación (al menos, a partir de algunos ejemplos claros de la noción de forma lógica). Consideremos las siguientes oraciones:

(1) Javier es profesor

(2) Si Jorge es filósofo entonces es profesor

(3) Bertrand Russell es matemático y filósofo

Cuyas correspondientes formas lógicas serían:

(F1) a es A

(F2) Si a es A entonces a es B

(F3) a es A y a es B[34]

Donde a 'a' y 'A' deben entenderse como símbolos esquemáticos de las correspondientes categorías gramaticales ('a' funciona como un símbolo esquemático para nombres propios, mientras que 'A' lo es para predicados). Esto refleja la intuición de que la forma lógica está ligada a la estructura gramatical de las oraciones y que dichas formas lógicas no afirman nada sobre el discurso específico de sus correspondientes oraciones. Por el momento no discutiré si efectivamente cada uno de estos esquemas es la forma lógica de la oración correspondiente, simplemente busco señalar que si estos esquemas son las formas lógicas de las oraciones, parece natural suponer que dichos esquemas no reflejan características particulares ni del discurso en que está enunciada la oración, ni de los objetos de los que habla. De hecho, parece que estos esquemas pueden serlo, no solo de las oraciones (1)-(3), sino de otras oraciones similares a estas, por ejemplo:

(1') Carlos es racista

(2') Si Francisco es altanero entonces es despreciable

(3') Kant es protestante y comprensivo

Parece que tanto (1) como (1') pueden ser recuperadas esquemáticamente por la misma forma lógica, a saber (F1), caso similar a los de (2) y (2') o (3) y (3'), que se pueden recuperar por (F2) y (F3) respectivamente. Intuitivamente, si la forma lógica de una oración dependiera de características del discurso o de los objetos de los que habla la oración, habría que explicar porqué oraciones como (1) y (1') tienen la misma forma lógica sin que exista una relación aparente entre el discurso o los objetos de los que hablan (1) y (1'). A partir de los casos anteriores, parece que si aceptamos que la forma lógica de una oración depende la estructura gramatical de la oración (de la manera en la que dependía en lo ejemplos), nos lleva a aceptar intuitivamente que logramos el tipo de neutralidad que buscamos.[35]

[34] Aunque generalmente se acepta que estas son las formas lógicas de estas oraciones, existen otras teorías como las de Quine que no considerarían estas como las formas lógicas de las oraciones. La razón de esto es que Quine (1982) buscaba *eliminar* los términos singulares de las oraciones debido a que términos singulares como 'Zeus' son presuntamente vacíos y por lo tanto, existen dificultades para determinar sus condiciones de verdad. En su lugar, sustituye los términos singulares por descripciones que nos permiten identificar las referencias de los términos singulares usando únicamente predicados. Sin embargo, utilizaré una noción más tradicional de *forma lógica*.

[35] Sin embargo, existen problemas en esta clase de caracterizaciones sobre la forma lógica. MacFarlane señala que considerar esta clase de caracterizaciones tienen el problema del "chauvinismo gramatical", esto es "toma lo que es una característica meramente contingente de nuestros sistemas lógicos y la presenta como una regla general" (MacFarlane 2000: 46). Esta característica es considerar que todo lenguaje tiene categorías gramaticales bien delimitadas, de tal forma que siempre se puede identificar qué partículas constituirán la forma lógica de los argumentos. A mi parecer, es complicado encontrar una caracterización de la noción de forma lógica que no cometa un error parecido, aunque no es mayor problema tampoco en tanto que buscamos considerar lenguajes donde las categorías gramaticales están delimitadas, sin pretender dar una generalización como la que señala MacFarlane.

Sin embargo, creo que existe una razón adicional a la neutralidad del tipo de discurso (aunque relacionada) por la cual se considera que la noción de validez debe ser formal: se busca que la validez de un argumento dependa de la forma lógica del argumento. Esta dependencia es vaga, pero lo que se pretende es que la validez de un argumento dependa de la "validez"de la forma lógica. Es decir, si un argumento es intuitivamente válido, entonces no puede ser que la forma lógica de dicho argumento no lo garantice.

Consideremos ahora la propiedad de formalidad que proporcioné en la sección 2.2:

(Form) Si un argumento es un caso de consecuencia lógica entonces todos los argumentos con la misma forma lógica también son casos de consecuencia lógica

Cabe recordar que esta propiedad es una condición necesaria de la noción intuitiva de consecuencia lógica. ¿Esta propiedad rescata las intuiciones sobre la naturaleza formal de la noción de consecuencia lógica expuestas más arriba? Consideremos un argumento intuitivamente válido cualquiera, *I*. Supongamos que existe otro argumento con la misma forma lógica de *I*, *I'*, tal que no es un caso de consecuencia lógica. Si esto fuera así, entonces intuitivamente la forma lógica no garantiza la validez del argumento *I*, pues *I'* debería ser un argumento válido ya que la forma lógica de *I* intuitivamente debería garantizar su validez. La intuición que busco señalar aquí es que *I'*, en principio, podría funcionar como un argumento contra la validez de *I*, de forma parecida a como Aristóteles utilizaba la noción de contraejemplo (véase la nota 3 del presente capítulo). Lo que afirma la condición (Form) es que la validez se preserva a través de la forma lógica, y es en este sentido que sostengo que la validez de los argumentos depende de la forma lógica: no es posible que un argumento sea válido, y que exista un argumento con la misma forma lógica que no lo sea.

Para poder dar una exposición completa de (Form), habrá que explicar cómo identificar si dos argumentos tienen la misma forma lógica o no. Recuperando los ejemplos anteriores, se suele aceptar que (1)-(3) y (1')-(3') tienen la mismas formas lógicas, respectivamente. Por lo cual, es generalmente aceptado que partículas lingüísticas como 'es' y 'si. . . entonces' caracterizan las formas lógicas (F1)-(F3). Está clase de partículas lingüísticas se conocen como constantes lógicas y nos apoyaremos en ellas para caracterizar la noción de *identidad de forma*:

(IdForm) La oración A tiene la misma forma lógica que una oración B si y solo si B puede ser obtenida a partir de A mediante sustitución uniforme de constantes no lógicas

Por ejemplo, considerando (1) y (1') arriba, (1') tiene la misma forma lógica que (1) porque la obtuvimos por sustitución uniforme de (1) sustituyendo el término 'Carlos' por 'Javier' (al igual que 'racista' por 'profesor'). La sustitución debe ser uniforme en el sentido de que cada constante no lógica que aparezca en una oración debe ser sustituida por una y sólo una constante. Por ejemplo, tomando (3) de los ejemplos anteriores:

(3) Bertrand Russell es matemático y filósofo

Para obtener su forma lógica, podemos parafrasearla del siguiente modo:

(3#) Bertrand Russell es matemático y Bertrand Russell es filósofo

A partir de la anterior paráfrasis, podemos ver qué (3#) no tiene la misma forma lógica que:

(3*) Immanuel Kant es protestante y John Locke es comprensivo

Pues, (3*) no puede ser obtenida por sustitución uniforme de (3). Debido a que el término 'Bertrand Russell' debe ser sustituido en (3) por dos términos diferentes 'Immanuel Kant' y 'John Locke' para poder obtener (3#), lo cual no puede ser pues 'Bertrand Russell' debe ser sustituida por una y solo una constante no lógica.[36] Así pues, la nueva condición de formalidad sería:

(Form') Si un argumento A es un caso de consecuencia lógica, entonces, todos los argumentos que resultan por sustitución uniforme de constantes no lógicas de A, también son casos de consecuencia lógica

Nuestro criterio de identidad de forma parece adecuado si consideramos que la forma lógica de una oración depende de las constantes lógicas del lenguaje en el que está enunciada la oración, en el sentido de que no variamos las expresiones consideradas constantes lógicas. La manera en que expusimos (Form') deja claro que esta manera de entender la noción de forma lógica es un caso de lo que MacFarlane llama "formalidad esquemática", es decir, "se refiere a patrones de inferencia, o esquemas ("formas") cuyas instancias son inferencias correctas, sin importar cuál sea la materia en la que esté instanciada" (McFarlane 2000: 37). MacFarlane señala que esta clase de noción de forma lógica tiene el problema de que plantea dos *lagunas* a la noción de formalidad:

> El problema con la formalidad esquemática es que ésta tiene dos lagunas, que deben ser cubiertas antes de poder dar un veredicto definido sobre la formalidad de inferencias particulares. Para cubrir la primera laguna, uno debe especificar qué características de las inferencias son parte de los esquemas y qué partes son remplazables o son elementos esquemáticos. Para cubrir la segunda laguna, uno debe especificar el rango de expresiones que pueden remplazar los diferentes tipos de elementos esquemáticos en los patrones (MacFarlane 2000: 38)

[36]Se puede pensar que es suficiente con que dos constantes cumplan con tener la misma referencia. Así, 'Bertrand Russell escribió Sobre la denotación y Bertrand Russell escribió Los problemas de la filosofía' tendría la misma forma lógica que 'Eric Blair escribió Rebelión en la granja y George Orwell escribió 1984' pues esta oración fue obtenida de la anterior mediante sustitución de constantes no lógicas y 'Eric Blair' y 'George Orwell' refieren al mismo hombre. Sin embargo, no es necesario que sea posible saber que esta clase de relación se cumpla. Es decir, es posible que alguien objete que no tienen la misma forma lógica en vista de no sabe que 'Eric Blair' y 'George Orwell' refieren al mismo objeto. En general, esto exige que se satisfagan condiciones poco generales y dependientes del contexto en que está enunciada una oración, por lo cual es preferible que la sustitución sea uniforme en el sentido enunciado más arriba.

La segunda laguna la trataremos más adelante, cuando discutamos la propiedad de *necesidad lógica*. La primera laguna es lo que se puede entender como la pregunta de ¿cuáles son las constantes lógicas? Así pues, habrá que decir algo sobre el problema de cuáles son las constantes lógicas.

Intuitivamente, una constante lógica es una expresión de un lenguaje que tiene ciertas propiedades (lógicas) que permiten determinar la forma lógica de una oración. Entre las constantes lógicas usuales del lenguaje cotidiano se consideran expresiones como 'y', 'si... entonces', 'o', etc. Retomando los ejemplos anteriores:

(1) Javier es profesor

(2) Si Jorge es filósofo entonces es profesor

(3) Bertrand Russell es matemático y filósofo

Y sus formas lógicas:

(F1) a es A

(F2) Si a es A entonces a es B

(F3) a es A y a es B

Se suele aceptar que (F1)-(F3) son las formas lógicas de (1)-(3) respectivamente porque las expresiones que las determinan (a saber, 'es', 'si... entonces' e 'y') son expresiones lógicas, en contraste con expresiones como 'Jorge', 'Javier', etcétera. A partir de ejemplos como estos, se considera que las constantes lógicas son expresiones cuyo significado no varía de oración en oración, y es por esto que pueden determinar la forma lógica de una oración. Por las mismas razones, (1')-(3') tienen la misma forma lógica de (1)-(3), respectivamente.

El problema consiste en que no es claro que exista un criterio que nos permita identificar eficazmente las constantes lógicas de un lenguaje de manera unívoca. Esto se debe básicamente a que tampoco es claro cuáles son las propiedades propiamente lógicas que tienen las expresiones tradicionalmente consideradas constantes lógicas. Más aún, no sabemos qué tipo de propiedades consideramos que deben satisfacer, es decir, no sabemos cómo distinguir una propiedad propiamente lógica de la que no lo es. Aunque es posible defender que una propiedad de las constantes lógicas es que su significado no varía en diferentes oraciones, esto no nos ayuda a determinar cuáles expresiones son constantes lógicas y cuáles otras no. Gómez Torrente (2002) argumentó que algunas caracterizaciones de *constancia lógica* eran inadecuadas para dar condiciones necesarias y suficientes para la constancia lógica. Gómez Torrente señala que "algunos principios poco filosóficamente cargados y en gran medida pragmáticos parecen haber guiado a los lógicos en la selección de expresiones como lógicas, explicita o (muchas veces) implícitamente" (Gómez Torrente 2002: 3). Si es el caso que la elección de las constantes lógicas ha sido guiada por principios pragmáticos, parece poco probable que una teoría sobre la constancia lógica sea adecuada si no es capaz de incorporar dichos principios. Si buscamos una caracterización adecuada de la noción de constante lógica, debemos tomar en consideración estos principios.

El propio Gómez Torrente afirma que existe un principio que se ha utilizado para delimitar las constantes lógicas de las no lógicas: "[...]un principio que del que es razonable pensar como subyacente a la elección de expresiones para el estudio lógico es el principio de que la lógica debe tratar con expresiones útiles en, y relevantes para, el razonamiento general, expresiones no específicas a ninguna de las esferas donde un argumento es empleado pero común a todas o gran número de ellos" (Gómez Torrente, 2002: 3). Otro ejemplo de principio pragmático que me parece ha llevado a los lógicos a considerar que una expresión es una constante lógica está enunciado por Sagi. Ella señala que una característica de la forma lógica debe ser que: "nos permitirá evitar consideraciones metafísicas en algún grado" (Sagi 2014: 946). Dada nuestra caracterización de forma, este principio debe cumplirse para las constantes lógicas. Tomando el ejemplo de Sagi, quisiéramos que la inferencia válida:

(M5) $\dfrac{\text{A todos les gusta el queso}}{\text{A Sarah le gusta el queso}}$

lo fuese independientemente de la naturaleza metafísica de Sarah (por ejemplo, independientemente de si existe o no). Considerar términos como 'Sarah' como constantes presumiblemente nos llevaría a comprometernos con la existencia de Sarah, por lo cual 'Sarah' no sería una constante lógica. Vale la pena notar que esto es bastante cercano a las intuiciones sobre la noción de forma lógica en las que se motivaba la idea de que ésta no afirma nada del discurso en que está enunciada una oración.

No pretendo dar una caracterización de la noción de constante lógica, sin embargo creo que al menos estas dos ideas han estado detrás de la elección común de constantes lógicas. No me concentraré en discutir más sobre este punto, me limitaré a afirmar que el conjunto de las constantes lógicas de los lenguajes que nos interesan (los lenguajes formales) contienen al menos a las conectivas lógicas usuales (como las mencionadas anteriormente: 'es', 'si... entonces', 'y', etcétera), a los cuantificadores de orden finito y al símbolo de igualdad.[37] Hay que notar que, en apariencia, estas expresiones cumplen las condiciones de Gómez Torrente y Sagi.

Para finalizar esta sección, tomaremos la noción de (Form') como la enunciación de la propiedad de formalidad de la consecuencia lógica. Es posible equiparar nuestra noción de formalidad con una noción introducida por Gómez Torrente en (1998/99:

[37]El caso del cuantificador universal podría ser problemático. En la nota 3 en el capítulo 1, mencioné que Enderton podría considerar como terminología no lógica al cuantificador universal. Esto es porque cada estructura $\mathfrak{A}$ le asigna un conjunto (en principio) diferentes. En este sentido, el significado del cuantificador universal sería "Todos los objetos de un dominio D". Por ejemplo, si evaluamos una oración de la aritmética, el significado del cuantificador universal sería "Todos los números naturales", es decir, el significado del cuantificador varía con respecto a la estructura. Sin embargo, consideremos la inferencia de la nota 31 en este capítulo (la oración infiere φ(a) de $\forall x\varphi(x)$(instanciación universal)). La justificación *preteórica* de esta no parece depender de que el cuantificador signifique "Todos los objetos de un dominio D", sino más bien de que signifique "Todos los objetos". Es decir, el significado del cuantificador universal (a este nivel) parece no depender de la estructura particular en que se evalúe. Es más, la lectura "Todos los objetos de un dominio D" parece que viola la condición de Gómez Torrente de que las constantes lógicas deben ser útiles para el razonamiento en general: la regla de instanciación universal parece cumplirlo mientras que la lectura no. Además de que la regla se usa más bien como una regla del razonamiento en general. Por estas razones consideraré al cuantificador universal como una constante lógica.

379-81): *validez en virtud del significado de las constantes lógicas*. Gómez Torrente pretende que dicha noción se aplique a oraciones y que determine una propiedad que fuese una condición necesaria de la noción preteórica de validez (o verdad) lógica. Si una oración es *válida en virtud del significado de las constantes lógicas* entonces parece natural pensar que toda oración que resulte de sustitución uniforme de dicha oración también será válida. Si no fuese así, habría una sustitución de constantes no lógicas con la cual la oración sería inválida, en cuyo caso no diríamos que la oración era un caso de validez en virtud del significado de las constantes lógicas. Más adelante veremos que no es tan sencillo establecer que si un argumento (u oración) satisface (Form) implica que es válido(a) en virtud del significado de las constantes lógicas.

Parece que las consideraciones que hicimos sobre la noción de formalidad nos llevan a pensar que, para garantizar que la validez de un argumento sea una cuestión de forma requiere (como se refleja en (Form)) que el aspecto relevante para ello sea el significado de las conectivas lógicas (que se recupera en (Form')). Si efectivamente eso es el caso, entonces los objetivos mencionados al principio de esta sección se habrían logrado, es decir, se garantizaría que la noción de validez fuera neutral con respecto al discurso de enunciación porque las constantes lógicas que estamos considerando no toman elementos propios del discurso de enunciación, al menos en apariencia. Por esto consideraré la noción (Form) como la propiedad de formalidad que se atribuye comúnmente a la noción de validez, por las razones expuestas más arriba.

2.2.2. Modalidad

Había señalado en la sección 2.2 que la noción preteórica de consecuencia lógica conlleva un componente modal: si un argumento es intuitivamente válido, entonces sus premisas implican por necesidad lógica a la conclusión. En esta formulación del componente modal de la relación de consecuencia, el término clave que debemos explicar es el de 'necesidad lógica'. Habíamos visto que dicho componente se suele explicar en términos de la noción de posibilidad: no es posible que las premisas sean verdaderas y la conclusión falsa (u otras formulaciones similares). Por ahora no tenemos una forma más precisa para caracterizar esta condición. En la siguiente subsección, señalaré una limitación de la condición (Form') para motivar la introducción de la noción de *interpretación* para rescatar el componente modal de la noción de consecuencia lógica.

2.2.2.1. Formalidad e interpretaciones

Una propiedad importante de la noción *preteórica* de consecuencia lógica es que debe cumplir con la propiedad de formalidad, la cual definimos como la condición (Form), que junto con el criterio de identidad de forma (IdForm) se obtenía la condición (Form'). Habíamos señalado que el significado de las constantes lógicas no podía cambiar de una oración a otra, por lo cual, la forma en que se aborde el componente modal de la noción de consecuencia no debe violar este aspecto.

En su artículo sobre consecuencia lógica, Tarski (1936) presentó una condición similar a (Form'), la condición (F): si una oración X es consecuencia lógica de una clase de oraciones K, entonces:

> (F) Si en las oraciones de una clase K y en una oración X las constantes – aparte de las constantes puramente lógicas – son remplazadas por cualesquiera otras constantes [...], y si denotamos la clase obtenida de K como 'K'', y la oración obtenida de X como 'X'', entonces la oración X' debe ser verdadera con la única condición de que todas las oraciones de la clase K' son verdaderas (Tarski 1936: 415)

Y señaló que no era posible utilizar la condición (F) como una definición de consecuencia lógica, pues podía satisfacerse sin garantizar que el argumento en cuestión fuera un caso de consecuencia lógica ya que puede satisfacerse simplemente por insuficiencia del lenguaje: "Esta condición puede de hecho ser satisfecha simplemente porque el lenguaje con el que estemos trabajando no posee una rango suficiente de constantes extralógicas" (Tarski, 1936, págs. 415-16). La idea de fondo es que es posible que un argumento satisfaga la condición (F) sin ser un caso de consecuencia lógica, meramente porque el lenguaje que estamos considerando no tenga suficientes constantes no lógicas. Consideremos un lenguaje muy limitado, L, en el que sólo tenemos tres constantes no lógicas: 'Jorge', 'filosofo' y 'profesor', y consideremos el siguiente argumento de este lenguaje:

(A1) Si Jorge es filósofo entonces es profesor
Jorge es profesor

Jorge es filósofo

Si consideremos todas las posibles sustituciones de constantes no lógicas, la única sustitución posible es la siguiente:

(A1') Si Jorge es profesor entonces es filósofo
Jorge es filósofo

Jorge es profesor

Si consideramos como hechos metalingüísticos (es decir, no como afirmaciones del lenguaje L) que Jorge efectivamente es tanto filósofo como profesor, veremos que en ambos casos cumplen la condición (F) pues siempre que las premisas sean verdaderas, la conclusión lo será. De hecho, en ambos casos, ambas condiciones se cumplen. Sin embargo, no se considera que estos argumentos sean válidos preteóricamente. Podemos imaginar una situación posible donde Jorge no sea filósofo pero sí profesor. Consideremos un lenguaje L' que es el mismo que L agregando otra constante no lógica: 'escritor'. En este caso, podemos considerar la siguiente sustitución:

(A1") Si Jorge es escritor entonces es profesor
Jorge es profesor

Jorge es escritor

Si de nuevo consideramos como hecho metalingüístico que Jorge no es escritor, el argumento (A1") tendrá premisas verdaderas y conclusión falsa. La razón por la cual (A1) cumple con la condición (F) es porque el lenguaje L que consideramos es muy limitado, y no contiene constantes no lógicas que nos permitan referir a situaciones en las que Mario cumpla con la propiedad que refiere 'profesor' pero no con aquella a la que refiere 'filósofo'. De hecho, el propio Tarski menciona que nunca podremos tener suficientes constantes no lógicas para evitar este problema.[38]

Es fácil ver que se puede presentar una objeción similar a (Form'), pues no es posible que los lenguajes que consideremos tengan suficientes constantes no lógicas para garantizar que un argumento cumpla con (Form') simplemente porque no tenemos suficientes constantes no lógicas.[39] En este sentido, (Form') es una condición muy limitada: (Form') no implica que se satisfaga la propiedad de *necesidad lógica*. El problema por el cual ni la condición (F) ni (Form) cumplen con esto está en la noción de sustitución; finalmente, la limitación detrás ambas nociones está en que el lenguaje es muy limitado y las sustituciones son insuficientes. Sin embargo, Tarski solucionó el problema apelando a la noción de *interpretación*, y en general se ha considerado ésta como una solución adecuada.

Ahora bien, si aceptamos (IdForm) como un buen criterio de identidad de forma, aceptamos que lo relevante para determinar la forma lógica de un argumento son las constantes lógicas que están presentes en él, en el sentido de que su significado no puede variar. Es decir, si aceptamos (IdForm), aceptamos que aquello que puede variar es el significado de las constantes no lógicas. Adoptando la práctica común de recuperar la propiedad necesidad lógica apelando a la noción de interpretación, hay que aceptar que ésta está en función de las constantes no lógicas de un lenguaje, y que lo único que es posible reinterpretar son éstas.

2.2.2.2. Interpretaciones

Una *interpretación* es una manera de asignar significado a las constantes no lógicas. Una propiedad de esta asignación es que asigna *objetos adecuados* a las constantes no lógicas de un lenguaje. La noción de interpretación, entendida de esta manera, pretende ser una noción mucho más general (a mi modo de ver) que las formas tradicionales de entender las nociones modales. Consideremos de nuevo el argumento:

(M4) Algunos suegros son prejuiciosos

Algunos hombres son prejuiciosos

[38] La razón de esto es que, considerando lenguajes que pueden usarse para hablar de dominios infinitos, el conjunto de expresiones del lenguaje será estrictamente "más pequeño"que el conjunto de propiedades de objetos del dominio, pues los lenguajes que se consideran generalmente son lenguajes numerables.

[39] Hay que notar que la condición (F) es una propiedad diferente a (Form'). En (Gómez Torrente 2000a), se argumenta que ambas nociones son diferentes en tanto que la definición presentada por Tarski (1936) cumple con ambas mientras que la usual definición modelo-teorética sólo cumple con (Form'). La razón de esto es que la definición modelo-teorética contiene el concepto técnico de 'conjunto' (Gómez Torrente 2000a: 533).

el cual mencioné como ejemplo de implicación por necesidad lógica en la sección 2.2 y sin embargo no era un caso de consecuencia lógica. Ignoremos por un momento el problema de la forma lógica. Alguien podría defender que términos como 'suegros' y 'hombres' tienen el mismo significado en todo mundo posible o al menos, que ambos términos tienen la misma relación lógica en todo mundo posible. Si esto es así, claramente no es posible que el argumento tenga premisas verdaderas y conclusión falsa; sin embargo, esta necesidad descansa en el significado lingüístico de dichos términos. La noción de interpretación modifica los objetos que designan dichos términos. Es decir, si la noción de validez lógica descansa en la noción de interpretación (entendida de la manera en que se definió más arriba), entonces deberemos considerar si existe una asignación de objetos (en algún sentido de 'objeto') a los términos 'suegro', 'hombre' y 'prejuicioso' que haga a las premisas verdaderas y a la conclusión falsa.

De manera muy parecida a la que consideramos en el caso de (M4'), asignemos al término 'suegro' la colección de los humanos, al término 'hombre' la colección de los insectos y al término 'prejuicioso' la colección de los vertebrados. Presumiblemente tendremos pues premisas verdaderas y conclusión falsa. En realidad, hay que notar que dicha interpretación podría asignar objetos de este tipo tales que no exista un término del lenguaje que refiera a ellos, los cuales representaban un problema tanto para la condición (F) como (Form'), de ahí la ventaja de la noción de interpretación. Si uno busca sostener que es inadecuado considerar la noción de interpretación para analizar argumentos como el anterior y atender exclusivamente al significado de los términos, considerará menos interpretaciones que aquel que analiza el problema sin considerar el significado de los términos en un leguaje particular. Al final, como veremos más adelante, las interpretaciones que validan el argumento (M4) considerando el significado de los términos se pueden entender en esta noción más general de interpretación.

2.2.2.3. La propiedad de necesidad lógica

De esta forma, la propiedad modal de la noción de consecuencia se reduce a:

(Mod) Si un argumento es un caso de consecuencia lógica, entonces todas las interpretaciones que hacen verdaderas a las formas lógicas de las premisas hacen verdadera a la forma lógica de la conclusión

Si aceptamos (IdForm), entonces (Form') nos compromete con sólo reinterpretar las constantes no lógicas del lenguaje en cuestión, como mencioné en 2.2.2.1. Por esto es importante distinguir las constantes lógicas de un lenguaje de las no lógicas.

Como mencionamos más arriba, el problema de la distinción entre constantes lógicas y no lógicas es un problema abierto, pero podemos dar algunas condiciones para identificarlas en algunos lenguajes. Como señalé en 2.2.1, el conjunto de constantes lógicas contendrá al menos a las conectivas lógicas usuales así como a los cuantificadores de orden finito y el símbolo de igualdad. Otro requisito que se puede incluir, y que me parece deseable en estos casos (junto con los de Gómez Torrente y Sagi), está expuesto por Shapiro: "Un desiderátum del presente marco es que la terminología lógica tenga una interpretación *sencilla* en cada modelo. Esto es, dado un modelo M y un término lógico t, debe ser claro cómo t está interpretado en M" (Shapiro 1998a: 149). Es decir, si un término se considera lógico, entonces debe ser claro cuál es su denotación en cada interpretación. Shapiro pone como ejemplo el término 'Clinton', que no tiene una interpretación clara en ninguna interpretación, por lo cual, los nombres (o las constantes de individuo) no pueden ser consideradas constantes lógicas.[40] Este requisito es muy parecido al requisito de Sagi y es compatible con el requisito de Gómez Torrente, mencionados en la sección 2.2.1, y por definición esto es satisfecho por las conectivas lógicas así como por lo cuantificadores y el símbolo de igualdad. De la misma forma, Shapiro nos señala que la relación de predicación lo cumple, lo cual me parece poco problemático.[41]

Así pues, a partir de lo expuesto anteriormente sobre la noción preteórica de consecuencia lógica, y algunas observaciones acerca de las propiedades que ésta tiene, hay que considerar que lo importante para evaluar si un argumento es válido o no, es analizar si dicho argumento es válido bajo cualquier interpretación posible de las constantes no lógicas. Shapiro propuso una noción informal que incorpora esta idea:

(NP) "Φ es una consecuencia lógica de Γ si Φ vale en todas las posibilidades bajo toda interpretación de la terminología no lógica en la que Γ vale" (Shapiro 1998a: 148)

Es decir, que una oración Φ es consecuencia lógica de Γ si todas las posibles interpretaciones de la terminología no lógica en las que Φ es el caso (o todos sus elementos valen en la interpretación), Γ es el caso. A partir de lo mencionado anteriormente, se puede establecer que (NP) cumple con (Form) y con (Mod). Supongamos que existe un argumento $\langle\Gamma, \Phi\rangle$ que cumple (NP) pero que no cumple con (Form) o con (Mod). Si no cumple (Form) entonces existe un argumento que se obtiene por sustitución de constantes no lógicas de $\langle\Gamma, \Phi\rangle$, $\langle\Gamma', \Phi'\rangle$, que es inválido. Si utilizamos las denotaciones de las constantes no lógicas de $\langle\Gamma', \Phi'\rangle$ para reinterpretar (de manera uniforme) las constantes no lógicas de $\langle\Gamma, \Phi\rangle$, veremos que existe una interpretación posible de la terminología no lógica que hace inválido a $\langle\Gamma, \Phi\rangle$, por lo cual $\langle\Gamma, \Phi\rangle$ no satisface (NP), lo cual es contradictorio. Si no cumple (Mod), entonces existe una interpretación posible de la terminología no lógica que hace inválido a $\langle\Gamma, \Phi\rangle$, por lo cual A no satisface (NP), directamente contradiciendo el supuesto. Así, tomaré a (NP) como una formulación adecuada de la noción de validez preteórica de ahora en adelante.

2.3. Validez en clases

Finalmente, me dispondré a establecer la relación entre la noción informal y preteórica (NP) y la noción de *validez en estructuras cuyo dominio es una clase*, a partir

[40] Igualmente, esto es satisfecho, me parece, por los predicados y las relaciones entre individuos del dominio de discurso.

[41] En el capítulo anterior, vimos un par de casos donde la relación de predicación es no lógica. Sin embargo, esta consideración dependía de la necesidad de dar una semántica de la lógica de segundo orden completa y correcta, no con algún problema filosófico particular con respecto a la relación de predicación. No conozco ninguna objeción especial a considerar la relación de predicación como lógica (a excepción de los casos mencionados), y es en este sentido que considero que el hecho de que la relación de predicación es lógica poco problemático.

de las propiedades mencionadas anteriormente. La relación que busco establecer es que en la noción (NP), todas las interpretaciones de un lenguaje serán la colección de todas las clases. El objetivo será mostrar cómo estos supuestos, aceptados en la práctica de la lógica, nos llevan a considerar que el componente de modalidad de la noción de consecuencia lógica depende de la cardinalidad de los dominios de las interpretaciones, tal como lo defiende Shapiro (1998a).

2.3.1. Teoría de conjuntos y teoría de clases

Hasta ahora he analizado la noción de validez preteórica, con el objetivo de defender que ella satisface una par de propiedades: (Form) y (Mod). Así mismo, argumenté que estás dos propiedades implican que la noción preteórica tiene al menos como condición necesaria la propiedad de ser válido en toda interpretación de las constantes no lógicas, y propuse entender la noción de validez preteórica como la formulación (NP), pues satisface ambas propiedades y parece rescatar esta idea. Lo que resta por ahora será establecer qué conexión existe entre (NP) y la noción de validez en estructuras cuyo dominio es una clase. Para ello, hay que señalar que la noción de *todas las interpretaciones posibles* sigue sin estar claramente determinada. Explique que la noción de interpretación nos era útil para rescatar el componente modal de la noción preteórica de consecuencia lógica, por lo cual, la pregunta es ¿qué colección de interpretaciones representa el componente modal de la noción de consecuencia lógica? Shapiro propone representar el componente modal en el universo de la teoría de modelos:

> La idea subyacente es que la naturaleza modal de la consecuencia lógica está representada en el universo de los modelos. Esto es, en efecto, un intercambio de modalidad con ontología [...] Para ponerlo negativamente, el programa de la teoría de modelos fallará si no hay suficientes modelos para representar toda 'posibilidad', de la manera apropiada. La metateoría está diseñada para asegurar que esto no suceda (Shapiro 1998a: 150)

Limitémonos a partir de ahora a la definición tarskiana de consecuencia lógica usual, (es decir, la *definición modelo teorética* como la definí en la sección 1.2.1 del capítulo anterior, a partir de ahora (DMT)). En esta definición se define como interpretación a un par constituido por un conjunto (entendido técnicamente como un objeto de la teoría de conjuntos estándar (a partir de ahora, ZFC)) y una función de interpretación de la terminología no lógica. Hay que notar que si cambiamos la función de interpretación, cambiamos la interpretación que estamos utilizando, por lo cual, tenemos diferentes interpretaciones para cada dominio de cuantificación, lo que me parece algo positivo para la propuesta que consideramos.

Shapiro apoya su idea en una propiedad que satisface la (DMT), la propiedad del isomorfismo (que es una consecuencia del teorema del homomorfismo (véase sección 1.4 del capítulo anterior))

Teorema 2.1. *Si dos modelos M, M' son isomorfos en la interpretación de los ítems no lógicos en una fórmula* Φ*, entonces M satisface* Φ *si y solo si M' satisface* Φ[42] *(Shapiro 1998a: 151)*

Es decir, si dos modelos son isomorfos en todas sus interpretaciones de la terminología no lógica, satisfacen exactamente las mismas fórmulas. A partir de esta propiedad, sabemos que cualesquiera interpretaciones, si son isomorfas entonces satisfacen exactamente las mismas fórmulas. De hecho, se cumple lo siguiente: consideremos dos conjuntos, A y B, donde el primero es el dominio de una interpretación $M=\langle A, I\rangle$, tal que ambos conjuntos tienen la misma cardinalidad. Es posible construir una interpretación isomorfa a ésta utilizando al conjunto B como dominio de esta nueva interpretación, pues si tienen la misma cardinalidad existe una biyección, f, entre ambos conjuntos, y por axioma de remplazo de f sobre las asignaciones de objetos proporcionadas por la función de interpretación I, obtenemos una interpretación nueva cuyo dominio es el segundo conjunto. De esto, Shapiro señala que se obtiene el siguiente resultado:

Teorema 2.2. *Sean N y N' dos modelos cuyos dominios tienen la misma cardinalidad. Entonces* Φ *es satisfecha por N bajo cualquier interpretación de la terminología no lógica si y solo si* Φ *es satisfecha por N' bajo cualquier interpretación de la terminología no lógica (Shapiro 1998a: 152)*

De esta manera, lo único aparentemente relevante para que una fórmula sea verdadera en una interpretación es la cardinalidad del dominio y la función de interpretación. Sin embargo, como vimos, la función de interpretación es "replicable" mediante el axioma de remplazo en cualquier otro dominio, con la misma cardinalidad. Es decir, parece que la función de interpretación no varía en realidad nada de una interpretación a otra, si la cardinalidad de sus dominios es la misma. De esto Shapiro concluye "En breve, la única cosa que puede impedir a una formula dada de ser una verdad lógica, o la única cosa que puede impedir que una fórmula sea una consecuencia lógica de un conjunto de fórmulas, es un modelo de cierto *tamaño*" (Shapiro 1998a: 152). Lo que realmente es relevante para establecer las condiciones de verdad de una oración en una interpretación es el tamaño de su dominio. La pregunta sobre el componente modal de la noción de consecuencia lógica se reduce a si ZFC es suficiente para proporcionar todos los posibles *tamaños* de un dominio de interpretación.

En ZFC, un conjunto es una colección de objetos, que sean objeto de estudio de la teoría de conjuntos (es decir, conjuntos). Sin embargo, existe un teorema de ZFC que afirma que no toda colección de conjuntos es un conjunto. Dicho teorema es consecuencia del siguiente teorema de ZFC:

Teorema 2.3. *Para todo conjunto, hay un conjunto que no le pertenece*

A partir de este teorema se concluye que la colección de todos los conjuntos no es un conjunto, y se establece una distinción entre un *conjunto* y una *clase propia*, donde estas últimas son colecciones de objetos que no pueden ser conjuntos. Una clase también es, pues, una colección de objetos, por lo cual, todo conjunto es una

[42] Este teorema es consecuencia de hecho del corolario 1.1 en la sección 4 del capítulo 1.

clase pero no toda clase es un conjunto. El término 'clase' se utiliza para designar a dichos objetos, y por lo tanto es un término técnico que depende del uso de ZFC. La consideración relevante aquí es que el tamaño de una clase propia es un *tamaño límite* que se encuentra fuera de ZFC, y sin embargo, es un tamaño en principio posible de una interpretación. De esta manera, se puede considerar a la colección de las clases como la colección de todas las interpretaciones posibles de un lenguaje. La forma en la que me parece que existe una relación entre la noción (NP) y la noción de *validez en una estructura cuyo dominio es una clase* es precisamente que podemos considerar que todas las interpretaciones posibles (de las que habla (NP)) son la colección de todas las clases. Esto depende por supuesto de asumir que la (DMT) cumple la noción de (Form), lo cual me parece que de hecho hace. Consideremos un argumento I válido evaluado con (DMT) y consideremos un argumento I' que tenga la misma forma que I (en el sentido de (IdForm)), sea $\mathfrak{A}$ una estructura tal que las premisas de I' son verdaderas en ella pero la conclusión falsa. Si utilizamos esta estructura para interpretar las constantes no lógicas de I, de tal forma que la interpretación de una constante c' en I' interprete a la constante c en I por la que fue sustituida, esta nueva estructura $\mathfrak{B}$ será tal que las premisas de I son verdaderas en ella.[43] Esto es así, porque la sustitución es uniforme y por lo tanto la interpretación de I de esta manera también lo será: si $I^{\mathfrak{A}}$ asigna un objeto o a una constante c en I, entonces se le asignará el mismo objeto[44] a la constante c' (que se sustituyó por c). Por la misma razón, la conclusión de I será falsa lo que contradice la hipótesis de que I es válido en (DMT).

2.3.2. Más sobre interpretaciones

Antes de concluir, me gustaría decir algo más sobre la noción de interpretación y la noción de *necesidad lógica*. Habíamos visto el ejemplo del argumento:

(M4) Algunos suegros son prejuiciosos

Algunos hombres son prejuiciosos

y argumentos similares donde la necesidad lógica del argumento depende del significado lingüístico de los términos, tales como 'hombre'. También habíamos visto que estos términos podrían tener la misma relación en todo mundo posible, por lo cual, este argumento sería un caso de necesidad lógica. El objetivo ahora es indicar en qué sentido la noción de interpretación general que hemos usado incluye las interpretaciones que validan este argumento. Consideremos una paráfrasis del componente modal de la relación de consecuencia:

(MP) No existe un mundo posible donde las premisas sean verdaderas y la conclusión falsa

43 Vale la pena notar que el dominio de $\mathfrak{B}$ es $D^{\mathfrak{A}}$, y que la única diferencia entre $I^{\mathfrak{A}}$ y $I^{\mathfrak{B}}$ es posiblemente las asignaciones de las constantes de I.

44 En este pequeño argumento, uso el término 'objeto' en un sentido amplio que incluye a las relaciones entre objetos del dominio.

De esta manera, se podría evaluar el componente modal de la noción de consecuencia lógica considerando que una interpretación está determinada por un mundo posible y asigna un objeto de ese mundo posible a cada constante no lógica del lenguaje en cuestión. Dicha estrategia es tomada por Gómez Torrente en (1998/99, 2000b), donde define una serie de "nociones generalistas", y a partir de ellas define una noción de consecuencia lógica para cada noción. Dichas nociones son colecciones de objetos entendidas dentro de un marco técnico. Con (MP) podríamos rescatar la intuición de que cuando afirmamos que la noción de consecuencia lógica es necesaria, significa que si un argumento es válido lo es en todo mundo posible. Cuando pensamos en las posibles interpretaciones de un lenguaje, podemos considerar cuatro posibilidades:

- (modo de interpretar)$_T$ es una estructura $\mathfrak{A}$ como las utilizadas en el capítulo anterior, donde $D^{\mathfrak{A}}$ es un un conjunto que contiene sólo objetos del mundo actual

- (modo de interpretar)$_{CoP}$ es una estructura $\mathfrak{A}$ como las utilizadas en el capítulo anterior, donde $D^{\mathfrak{A}}$ es un conjunto que contiene sólo objetos meramente posibles (no necesariamente actuales, o que de hecho existan)

- (modo de interpretar)$_{Cl}$ es una interpretación cuyo dominio es una clase compuesta de objetos actuales

- (modo de interpretar)$_{ClP}$ es una interpretación cuyo dominio es una clase compuesta de objetos meramente posibles

Cabe mencionar que cada una asigna un objeto adecuado a las constantes no lógicas del lenguaje, y en el caso de la segunda y la cuarta, dichas asignaciones están limitadas a un posible m, es decir, la interpretación asigna objetos existentes en un mundo posible m.

Existen varias relaciones entre estas nociones. Primero, es claro que toda interpretación que sea (modo de interpretar)$_T$, será (modo de interpretar)$_{Cl}$ (pues todo conjunto es una clase); lo mismo sucede con (modo de interpretar)$_{CoP}$ y (modo de interpretar)$_{ClP}$. Ahora bien, toda interpretación que sea (modo de interpretar)$_T$ es (modo de interpretar)$_{CoP}$, simplemente aceptando que los objetos actuales son objetos posibles. Por la misma razón, toda interpretación que es (modo de interpretar)$_{Cl}$ es (modo de interpretar)$_{ClP}$. De estas relaciones podemos inferir que toda interpretación de (modo de interpretar)$_T$ es (modo de interpretar)$_{ClP}$.

De forma análoga a la definición tarskiana de consecuencia lógica, Gómez Torrente define una noción de consecuencia lógica para cada modo de interpretar:

- (CLCoP) Una oración Φ es (consecuencia lógica)$_{CoP}$ de un conjunto de premisas Γ si y solo si Φ es verdadera en todo (modo de interpretar)$_{CoP}$ en los que todas las oraciones de Γ son verdaderas

- (CLCl) Una oración Φ es (consecuencia lógica)$_{Cl}$ de un conjunto de premisas Γ si y solo si Φ es verdadera en todo (modo de interpretar)$_{Cl}$ en los que todas las oraciones de Γ son verdaderas

- (CLClP) Una oración Φ es (consecuencia lógica)$_{ClP}$ de un conjunto de premisas Γ si y solo si Φ es verdadera en todo (modo de interpretar)$_{ClP}$ en los que todas las oraciones de Γ son verdaderas

Estas nociones son útiles pues nos permiten recuperar nuestra intuición sobre el término 'posible' en las definiciones intuitivas del componente modal de la relación de consecuencia, como en la noción (MP). Uno puede considerar que la propiedad de necesidad lógica sólo puede ser recuperada por la noción de interpretación, si nuestras interpretaciones pueden garantizar que los argumentos válidos son válidos en todo mundo posible. Podríamos considerar que si un argumento es un caso de (consecuencia lógica)$_{ClP}$ entonces es válido en todo mundo posible. El punto importante es que Gómez Torrente también argumenta rigurosa y brevemente que es suficiente considerar (modo de interpretar)$_{Cl}$ y (modo de interpretar)$_T$, y por lo tanto, (CLCl) y la (DMT). Retomando las relaciones señaladas más arriba, sabemos que si un argumento es un caso de (consecuencia lógica)$_{Cl}$, es un caso de (DMT), pues toda interpretación que sea (modo de interpretar)$_T$, será (modo de interpretar)$_{Cl}$. Por la misma razón, si un argumento es un caso de (consecuencia lógica)$_{ClP}$, entonces será un caso de (consecuencia lógica)$_{CoP}$. Si un argumento es un caso de (DMT), es un caso de (consecuencia lógica)$_{CoP}$, ya que toda interpretación que sea (modo de interpretar)$_T$ es (modo de interpretar)$_{CoP}$. Y finalmente, todo argumento que sea un caso de (consecuencia lógica)$_{Cl}$, es un caso de (consecuencia lógica)$_{ClP}$, por la misma razón. Por lo tanto, falta demostrar que todo caso de (consecuencia lógica)$_{CoP}$ es un caso de (DMT), y que todo caso de (consecuencia lógica)$_{ClP}$ es un caso de (consecuencia lógica)$_{Cl}$. Para ello presenta la siguiente demostración:

Teorema 2.4. *Para todo* (*modo de interpretar*)$_{Cl}$*, existe un* (*modo de interpretar*)$_{Cl}$ *isomorfo cuyo universo es una clase pura*[45]

Demostración. Sea I un (modo de interpretar)$_{Cl}$ y sea D su universo. Es una consecuencia del axioma de elección global y el axioma de regularidad que toda clase C puede ponerse en correspondencia biyectiva con un segmento incial de los ordinales (que suponemos se han construido como conjuntos puros...) Este segmento inicial será la clase de todos los ordinales, Ω [*Ord*][46], si C es una clase propia... En particular, hay una correspondencia biyectiva entre un segmento inicial de Ω y D. Luego, un isomorfismo entre I y un (modo de interpretar)$_{Cl}$ cuyo universo es ese segmento inicial inducido por esa correspondencia 1-1 (Gómez Torrente 2000b: 67) □

Y utilizando esto, muestra que:

Teorema 2.5. ***Cl-ClP*** *Todo caso de* (*consecuencia lógica*)$_{Cl}$*, es un caso de* (*consecuencia lógica*)$_{ClP}$

Demostración. Supongamos que X es (consecuencia lógica)$_{Cl}$ de K pero que X no es (consecuencia lógica)$_{ClP}$ de K. Entonces, hay un mundo posible *m* habitado por una clase no vacía de objetos de U con la que se puede construir un (modo de

[45] Una *clase pura* es una clase cuyos elementos son conjuntos puros. De la misma manera, un *conjunto puro* es un conjunto cuyos elementos son conjuntos puros.

[46] El texto entre corchetes en esta nota es mío, no del autor

interpretar)$_{Cl}$, J [...] que hace verdaderas a todas las oraciones de K y falsa a X. [El teorema 2.4][47] es verdadero en todos los mundos posibles, pues es una verdad matemática. Por tanto, m contiene un (modo de interpretar)$_{Cl}$ L isomorfo a J cuyo universo es una clase pura. Presumiblemente, L existe en todos los mundo posibles, en todos los mundo posibles las oraciones de K son verdaderas bajo L, y X es falsa bajo L. Pero entonces las oraciones de K son verdaderas bajo L y X es falsa bajo L en el mundo actual; y por tanto, X no es (consecuencia lógica)$_{Cl}$ de K, lo cual contradice nuestro supuesto (Gómez Torrente 2000b: 67) □

Y dado que todo conjunto es una clase, sabemos que para cada (modo de interpretar)$_{CoP}$ existe un modo de interpretar (modo de interpretar)$_T$, cuyo dominio es un conjunto puro. Y aplicando el mismo argumento, podemos concluir que todo caso de (consecuencia lógica)$_{CoP}$ es un caso de (DMT).

Este argumento nos muestra que es suficiente considerar las interpretaciones conjuntistas y aquellas cuyo dominio es una clase, usadas más arriba. Esto porque considerar que, en principio, nuestra noción informal de interpretación considera aquellas interpretaciones que se construyen con objetos posibles, no presenta mayor reto a la teoría de conjuntos. Así mismo, se logra sostener que es suficiente considerar la teoría de conjuntos estándar (sin urelementos), pues la naturaleza de los objetos que se consideren parece irrelevante. De igual forma, logramos rescatar la posible intuición de que la representación del componente modal de la noción de consecuencia lógica debe dar cuenta la intuición de que un argumento es válido si lo es en todo mundo posible.

Si esto es así, no parece que haya mayor dificultad para considerar las interpretaciones que validan el argumento (M4), ya que podemos considerar cualquier interpretación sobre cualquier colección de objetos posibles. Sin embargo, el argumento tal cual sigue siendo inválido en un análisis tarskiano, y uno querría poder garantizar que sólo usaremos las interpretaciones que lo validen. Una opción es indicar de forma clara y explícita que existe una relación entre los términos 'hombre' y 'suegro', agregando una premisa extra:

(M4*) Algunos suegros son prejuiciosos
(Todos los suegros son hombres)

Algunos hombres son prejuiciosos

Dicha estrategia nos garantiza que sólo consideraremos las interpretaciones que validan el argumento y podremos modelar el argumento sin mayor problema.[48] A final de cuentas, ya habíamos considerado dichas interpretaciones.

[47] El texto entre corchetes en esta nota es mío, no del autor

[48] Se puede objetar que mediante esta manera de enfrentar este problema implicaría que cualquier argumento es válido, simplemente hay que encontrar las premisas necesarias. A mi modo de ver, la manera de evaluar qué premisas se agregarán (o se aceptarán) depende del contexto. Un ejemplo similar a (M4*), supongamos que tenemos un argumento cuya única premisa es 'Algunos mexicanos son perezosos' y su conclusión es 'Algunos ladrones son perezosos', es posible agregar la premisa de que 'Todo mexicano es un ladrón', sin embargo, creo que esta premisa no tiene el mismo tipo aceptabilidad que tiene su análoga en (M4*), y dependerá del contexto su aceptabilidad (o posible defensa). La razón para modelar (M4) es que parece una inferencia correcta, lo cual no es claro que lo sea en este ejemplo.

Si todo lo mencionado anteriormente es correcto, habremos establecido una relación entre (NP) y la noción de *validez en estructuras cuyo dominio es una clase*. Quiero resaltar que dicha relación dependerá enteramente de que (DMT) cumpla con la condición (Form) y de que sea admisible usar la teoría de conjuntos como fuente de interpretaciones. Alguien podría objetar que no hay porqué utilizar la teoría de conjuntos como tal, para lo cual diré algunas cosas. Es cierto que, en la definición tarskiana de consecuencia lógica, tal cual la presenta Tarski en (1936), no se hace mención de qué clase de interpretaciones son aquellas con las cuales él entiende su definición. Sin embargo, por alguna razón se ha entendido que dichas interpretaciones son conjuntos, entendidos como objetos de estudio de la teoría de conjuntos ZFC. Más allá de las razones por las cuales se ha optado por usar la teoría de conjuntos, existen razones para considerar que dicho uso es deseable. La que me parece una razón importante es que nos permite analizar la noción de consecuencia en términos mucho más claros proporcionados por dicha teoría matemática. Como un punto aparte, me parece que la noción informal de interpretación expuesta en la sección anterior puede ser entendida sin muchos problemas como objetos de la teoría de conjuntos (como en realidad se hizo en el capítulo anterior). Mencionamos que una interpretación asignaba objetos adecuados a la terminología no lógica de un lenguaje, y presenté un ejemplo donde trato de mostrar qué tipos de objetos se asigna a qué términos y traté de motivar la noción de interpretación entendida de esta manera. Me parece que los objetos de la teoría de conjuntos han sido considerados buenos candidatos para ser estos objetos, al menos dentro de la práctica de la lógica clásica hasta ahora.

2.4. Conclusiones

Hemos establecido qué relación existe entre una noción preteórica de consecuencia lógica y la noción de validez en estructuras cuyo dominio es una clase, exponiendo sus presupuestos los cuales traté de motivar como poco controvertidos. Así mismo, intenté explicar que sólo depende de aceptar que la noción de consecuencia lógica tarskiana satisface la condición (Form) y optar por ZFC como fuente de interpretaciones. Esto último a partir de la idea de que ZFC provee el material adecuado para evaluar el componente modal de la noción de consecuencia lógica. Si esto es así, hay que notar que no he mostrado es que (DMT) cumpla (Mod). Recordemos que el objetivo del trabajo es proporcionar razones para evaluar esta cuestión a través de la noción de *validez en estructuras cuyo dominio es una clase*. Si esto es así, tanto la pregunta de si (DMT) cumple (Mod) como la de la adecuación extensional de ésta se reducen a la pregunta ¿ZFC tiene suficientes interpretaciones? Consideremos un argumento que es válido preteóricamente, si lo que mostramos de hecho es el caso, podemos concluir que será válido en la (DMT); a partir de que, en terminología de Gómez Torrente, si un argumento es un caso de (consecuencia lógica)$_{Cl}$, es un caso de consecuencia lógica tarskiana. Sin embargo, consideremos un argumento preteóricamente inválido: no es posible contestar afirmativamente que sea un caso de consecuencia lógica tarskiana.

Shapiro (1991) presenta un par de condiciones informales que se le puede pedir a una semántica:

> (1b) La semántica propuesta es fiel a las nociones pre-formales de 'interpretación posible' y 'validez'. Esto es:
>
>> (1b1) Cada modelo en la semántica es una interpretación legítima del lenguaje natural. Si un argumento en un lenguaje natural A corresponde a un argumento $\langle \Gamma, \Phi \rangle$ en L, y A es 'válido' en el sentido de que su conclusión es verdadera bajo toda interpretación del lenguaje natural en que sus premisas son verdaderas, entonces Φ es satisfecha por todo modelo de la semántica que también satisface a todos miembro de Γ (conformidad).
>> (1b2) Hay "suficientes" modelos en la semántica. Si un argumento en un lenguaje natural A corresponde a un argumento $\langle \Gamma, \Phi \rangle$ en L, y $\langle \Gamma, \Phi \rangle$ es válido en la semántica modelo teorética, entonces la conclusión de A es verdadera bajo toda interpretación del lenguaje que hacen verdaderas a las premisas de A. Esto equivale a decir que si hay una interpretación del lenguaje natural que hace verdaderas a todo miembro de un conjunto S de oraciones, y si todo miembro de Γ corresponde a un miembro de S, entonces Γ es satisfacible en la semántica (suficiencia) (Shapiro 1991: 41)

Lo que podríamos establecer hasta ahora sería la condición de conformidad (1b1), nos hace falta verificar que ZFC tiene suficientes interpretaciones para garantizar que todo argumento inválido tenga un contraejemplo, (es decir, la condición de suficiencia (1b2)).

Finalmente, me parece pertinente señalar algo que dejamos pendiente en nuestro análisis de la noción de forma lógica. Vimos que era posible argumentar que si un argumento es válido en virtud del significado de las constantes lógicas, entonces deberá cumplir con la condición (Form). Sin embargo, ahora podemos ver porqué no es fácil garantizar que si un argumento cumple la condición (Form), deba ser válido en virtud del significado de las constantes lógicas. Es posible que un argumento sea válido en la definición tarskiana de consecuencia lógica, (y por lo tanto) cumpla la condición (Form), y sin embargo su validez no dependa del significado de las constantes lógicas. Ha habido filósofos (por ejemplo, Etchemendy (1990: cap. 8) o Jane (2005: 790)) que efectivamente han defendido dicha afirmación, y han mostrado contraejemplos, apelando a que la noción de consecuencia lógica tarskiana valida argumentos que dependen de la naturaleza de ZFC. Esto va más allá de los límites de este trabajo y lo analizaré en trabajos posteriores.

Capítulo 3

El Principio de Kreisel

En el capítulo anterior argumenté que era posible formular una noción informal de validez preteórica (NP) relacionada con una *validez en estructuras cuyo dominio es una clase*. En este capítulo usaré esa noción para dar una caracterización del Principio de Kreisel que sea susceptible de ser estudiada formalmente. La razón de usar esta estrategia para estudiar el Principio de Kreisel es que entones será posible relacionarlo con otro principio de la teoría de conjuntos, ZFC: el principio de reflexión. Esto es así por dos razones: la noción (NP) ya es considerablemente más precisa que la noción de validez preteórica y es posible usar métodos formales muy parecidos a los presentados en el capítulo 1 para definir la noción de *satisfacción en estructuras cuyo dominio es una clase* (lo cual explicaré más adelante). El objetivo de este capítulo es proporcionar un marco lo más preciso posible en el cual sea posible estudiar el Principio de Kreisel y las condiciones de su satisfacción en el caso de la lógica de segundo de orden.

Este capítulo está compuesto en dos partes. En la primera (que estará compuesta de las dos primeras secciones) presentaré el Principio de Kreisel en su formulación original, así como el argumento que presentó Kreisel para defender que el principio se sostiene para el caso de la lógica de primer orden. En la primera sección, haré un análisis de dicho argumento tratando de mostrar el carácter informal y problemático del argumento, usando las justificaciones originales de Kreisel para exhibir esto. Me parece que hay evidencia textual de que Kreisel tenía en mente un argumento mucho más informal (y que apela a nociones preteóricas) de lo que algunas reconstrucciones de éste insinúan, entre ellas la de Shapiro (1987, 1991) mismo. Sin embargo, creo que la noción (NP) es una noción útil en vista de la manera en la que defendí dicha noción en el capítulo anterior. Así mismo, daré una nueva versión más general del principio en términos de la noción de consecuencia lógica en lugar de la noción de verdad lógica (que es como está planteado originalmente). En la segunda sección, presento una reconstrucción del Principio de Kreisel inspirada en la del propio Shapiro, usando la noción (NP). En la segunda parte del capítulo, precisaré la noción de satisfacción en estructuras cuyo dominio es una clase, mediante un aparato formal propuesto por Shapiro (1991). Con esto, será posible exponer el Principio de Kreisel mediante una fórmula de la Teoría de Conjuntos en segundo orden (ZFC2) que de hecho es equivalente a un Principio de Reflexión de la Teoría de Conjuntos. Expondré cuatro formulaciones del Principio de Kreisel presentadas por Shapiro, así como las demostraciones de que éstas son equivalentes a los Principios de Reflexión adecuados. Finalmente, también expondré algunas condiciones para su satisfacción. Todo esto lo retomo directamente del trabajo de Shapiro (1987, 1991), así que remito al lector a las fuentes originales para confrontar las pruebas de estos teoremas.

3.1. El principio de Kreisel

En esta sección presentaré lo que Shapiro llama 'Principio de Kreisel'. Georg Kreisel (1967: 152) presentó dicho principio que afirma que el conjunto de oraciones que son "intuitivamente" válidas es coextensivo con el conjunto de oraciones válidas en una definición tarskiana de consecuencia lógica. Es decir, que las oraciones válidas de un lenguaje formal son única y exclusivamente las oraciones que son válidas en la definición tarskiana. Para explicar el argumento y el principio hay que definir algunas nociones previas: sea α^i donde α es una oración e i es el orden de cuantificación del lenguaje en que está expresada la oración, por ejemplo, α^1 indica que la oración está expresada en un lenguaje de primer orden.

Definimos los siguientes predicados:

- $\mathrm{Val}(\alpha^i)$ es la afirmación de que α es intuitivamente válida (o bien, que es verdadera bajo cualquier interpretación)
- $\mathrm{V}(\alpha^i)$ es la afirmación de que α es verdadera en cualquier interpretación cuyo dominio es un conjunto (o bien, que α es una verdad lógica tarskiana)
- $\mathrm{D}(\alpha^i)$ significa que α es demostrable en una sistema formal adecuado de reglas de inferencia.

El Principio de Kreisel es:

(PK) $\forall i \forall \alpha (\mathrm{Val}(\alpha^i) \leftrightarrow \mathrm{V}(\alpha^i))$

Es decir, una oración es intuitivamente válida si y solo si es válida en cualquier interpretación cuyo dominio es un conjunto.[49]

Existen varios aspectos de estos predicados que hay que señalar antes de continuar. Para empezar, el principio literalmente está enunciado para la noción de verdad lógica, no para la relación de consecuencia, pues (PK) es una oración cuantificada sobre oraciones, no sobre pares compuestos de conjuntos de oraciones (como es el caso de la relación de consecuencia). De momento y por simplicidad, presentaré el principio en esta versión, sin embargo, será necesario para dar un análisis adecuado del concepto de consecuencia lógica explicar el principio en términos de la relación de ésta. Un segundo aspecto que vale la pena mencionar es que el principio está formulado para un lenguaje formal cuantificacional. Mencioné en el capítulo anterior

[49] Arriba mencioné que el Principio de Kreisel afirma que las oraciones válidas de un lenguaje son única y exclusivamente las oraciones que son válidas por la definición tarskiana. Es fácil ver que el predicado $\mathrm{V}(\alpha^i)$ tiene como extensión al conjunto de oraciones que son válidas bajo una definición tarskiana de consecuencia lógica. Basta con notar que la práctica usual en la teoría de modelos es considerar a las interpretaciones de un lenguaje o teoría son objetos de la Teoría de Conjuntos.

(en la sección 2.1) que la relación de validez intuitiva que consideré busca modelar al menos las inferencias de la matemática clásica y su lenguaje, y que este tiene dos características importantes, en particular, la característica de que las categorías semánticas estén delimitadas facilita mucho la formalización de este lenguaje. Por lo cual, este aspecto no representa un gran problema para el análisis del Principio de Kreisel.[50]

Kreisel también proporcionó un argumento para demostrar que el principio se cumple para el caso de la lógica de primer orden. Por supuesto, debido a la naturaleza informal del predicado Val(α^i), este argumento no es una prueba formal en un sentido riguroso. La naturaleza del predicado Val(α^i) presenta algunas dificultades para el análisis del argumento, pero este argumento permite al mismo tiempo extraer más información sobre el significado de Val(α^i). Presentaré el argumento deteniéndome para explicar y analizar la justificación de cada premisa:

(P1) $\forall i \forall \alpha (\mathrm{Val}(\alpha^i) \rightarrow \mathrm{V}(\alpha^i))$

La primera premisa es la afirmación de que toda oración que es intuitivamente válida es una verdad lógica en una definición tarskiana. O bien, es la afirmación de que la definición tarskiana no presenta contraejemplos a ninguna oración intuitivamente válida. Kreisel afirma que esto se acepta “cuando uno toma como garantía que la lógica se aplica a estructuras matemáticas” (Kreisel 1967: 154). Es difícil entender esta parte de la justificación, pues no es claro que del hecho de que una oración sea intuitivamente válida se siga que dicha oración deba ser verdadera en toda estructura conjuntista. Primeramente, hay que notar que la fórmula no está instanciada en las oraciones que son estrictamente de primer orden, esto quiere decir que su justificación debe ser independiente del lenguaje que se esté usando. Ahora bien, en efecto, se espera entre la comunidad filosófica que la lógica sea aplicable a estructuras matemáticas, o bien, que nuestras fórmulas sean satisfechas por estructuras matemáticas, en este caso los conjuntos.

Para esclarecer el significado del predicado Val(α^i), Kreisel hace la siguiente especificación: “El significado intuitivo de Val difiere del de V en algo particular: Vα (simplemente) afirma que α es verdadera en todas las estructuras en la jerarquía acumulativa, i.e., en todos los conjuntos en sentido preciso de *conjunto*, mientras que Valα afirma que α es verdadera en todas las estructuras” (Kreisel 1967: 153). Es decir, Kreisel consideraba que una de las características de las oraciones que cumplen con Val(α^i) es que son verdaderas en toda estructura.[51] Ahora bien, es fácil ver que

[50]Griffiths (2014) defiende que el principio de Kreisel puede ser usado de manera incorrecta cuando se busca usar para demostrar que la definición tarskiana es coextensiva con una propiedad que se aplica a los lenguajes naturales. Grifftihs afirma que o bien la definición tarskiana cumple con el principio trivialmente o es falsa, dependiendo de si el criterio de adecuación de una formalización que esté en uso depende de la noción de consecuencia o no. Creo que la noción preteórica de consecuencia que enuncié anteriormente no es una propiedad del lenguaje natural, o al menos, no pretende serlo. Para la aplicación de la noción que usé, presupongo una formalización adecuada de las oraciones que se analicen, que como mencioné dependen de los intereses del teórico. A partir de estas consideraciones, creo que el argumento de Griffiths no se aplica a la noción preteórica que uso.

[51]Es posible que esta oración sea un tanto problemática. Se puede pensar que esta afirmación de Kreisel hace que el tipo de validez considerada aquí no sea de ninguna manera la noción de

si una oración α es verdadera en todas las estructuras entonces debe serlo en todas las estructuras cuyo dominio es un conjunto, simplemente porque las estructuras conjuntistas son estructuras.[52] En este sentido, la justificación de Kreisel consiste en la afirmación trivial de que no estamos ignorando las estructuras matemáticas. Esto es, que las interpretaciones que consideramos interpretaciones adecuadas de una oración incluyen interpretaciones conjuntistas, lo cual parece poco controvertido.

La premisa uno no presenta mayor dificultad; sin embargo, existen algunas consideraciones importantes con respecto a la segunda:

(P2) $\forall i \forall \alpha (\mathrm{D}(\alpha^i) \rightarrow \mathrm{Val}(\alpha^i))$

Es decir, que toda fórmula que sea derivable dentro de un sistema formal adecuado es intuitivamente válida. La justificación que ofrece Kreisel de que esto es el caso, es que las reglas de inferencia de Frege ya eran aceptadas en su tiempo: "Es en general aceptado que en el tiempo de Frege quien formuló reglas para la lógica de primer orden, y la definición conjuntista de consecuencia de Bolzano olvidada (y que fue redescubierta por Tarski); aún entonces uno reconocía la validez de las reglas de Frege" (Kreisel 1967: 153). La idea de fondo es que en la época de Frege no existía una definición semántica de consecuencia lógica a través de la cual pudiésemos garantizar la validez de las reglas de inferencia, y sin embargo basta con analizar de manera intuitiva las reglas de Frege para convencerse de que nuestra premisa se sostiene. Esto queda aún más claro con el énfasis que pone en el hecho de que la definición de consecuencia lógica de Bolzano era ignorada en la época de Frege. De hecho, este análisis seguiría la misma idea que se sigue para el Teorema de corrección: analizaríamos los axiomas verificando que sean intuitivamente válidos y garantizaríamos que las reglas de inferencia preservan la propiedad de validez intuitiva. Con esta justificación, Kreisel subraya que garantizamos que el predicado $\mathrm{Val}(\alpha^i)$ tiene significado (cf. Kreisel 1967: 153). Al menos en un primer análisis, es claro que la premisa se sostiene, y que de hecho es natural pensar que es así. Si de alguna manera la premisa no se sostuviese, implicaría que tenemos una regla de inferencia que falla cuando se le analiza con cuidado; por lo cual, diríamos que nuestro primer análisis está equivocado, y que hay que excluir dicha regla de inferencia de nuestro sistema formal. La justificación de (P2) es un requerimiento mínimo de todo sistema de inferencia. Smith dice al respecto de ésta que "la validez [de las reglas de inferencia] en ese sentido es, después de todo, la principal razón por la cual los lógicos clásicos aceptan las reglas de los sistemas de prueba en primer lugar" (Smith 2011: 25). Si un sistema formal no cumpliese con ello, parece que habría suficientes razones para rechazarlo. En particular, y pensando en mi presente interés, me parece que los sistemas *D1* y *D2* (los sistemas de inferencia definidas en la secciones 1.1.2 y 1.2.2 del primer capítulo) lo cumplen.[53]

---

validez intuitiva (como argumenta por ejemplo Smith (2011)). Es cierto esto lleva a considerar no es un principio aplicable a cualquier concepción de la lógica o de la validez (posiblemente, sólo será aplicable a las nociones de validez que sean evaluables en términos de estructuras de algún tipo). Sin embargo, me parece poco problemático en la medida en que mi interés (y seguramente el de Kreisel también) es la lógica clásica, además de las observaciones hechas en el capítulo 2.

[52]Esta clase de explicación de esta justificación también puede encontrarse en (Griffiths 2014)

[53]Es posible, presentar ciertas dudas en el caso del esquema de comprehensión y del axioma de

Hay un aspecto importante con respecto la segunda premisa del argumento de Kreisel. Él argumenta con su segunda premisa que la condición de derivabilidad es suficiente para la validez intuitiva. Sin embargo, esto no implica que Kreisel quiera defender que la derivabilidad sea una condición necesaria. A partir de la segunda premisa, es claro que los teoremas son intuitivamente válidos ya que las reglas de inferencia lo son. Sin embargo, no necesariamente las intuitivamente válidas deben ser derivables. La segunda premisa de Kreisel establece solamente que una manera de garantizar que una oración es intuitivamente válida es su derivabilidad, lo cual no implica que sea necesaria esta garantía.

Sin embargo, parece haber una diferencia entre las justificaciones de ambas premisas. Si consideramos que el significado del predicado $\mathrm{Val}(\alpha^i)$ es "validez en toda estructura" entonces Kreisel habrá dado una afirmación interesante al tratar de justificar (P2), pero no habrá dado la misma justificación que expliqué en el párrafo anterior. La justificación que Kreisel ofrece de (P2) no apela nunca a la naturaleza de las estructuras en que la oración α es válida, sino a una propiedad de las reglas de inferencia que seleccionamos para construir nuestros sistemas lógicos. De la misma forma, si tratamos de justificar nuestra premisa (P1) apelando al hecho al que apela Kreisel en su justificación de (P2), nos enfrentaríamos a un problema similar. Consideremos que en este caso, $\mathrm{Val}(\alpha^i)$ significa "α es válido en el mismo sentido que las reglas de Frege". Si una oración α cumple con lo anterior, no es obvio que sea válida en toda estructura conjuntista, simplemente porque la noción que provee de significado al predicado $\mathrm{Val}(\alpha^i)$ sigue siendo una noción informal y poco clara, y no parece haber una forma obvia de relacionar una noción informal como esta con una noción matemática y precisa como la noción de consecuencia lógica tarskiana. Por un lado, la justificación de (P1) apela a una noción más precisa (como lo es la noción de *estructura*) y que se explica a partir de nociones derivadas de una teoría formal como ZFC, por otro lado, la justificación de (P2) apela una propiedad informal, preteórica e imprecisa. Por supuesto, ambas justificaciones no son incompatibles.

Me parece que ambas justificaciones indican dos criterios que debe cumplir una noción preteórica para validar el argumento que Kreisel tenía en mente. Si buscamos sostener que el argumento de Kreisel muestra que una noción precisa de consecuencia lógica es coextensiva con una noción preteorica, ésta última debe ser tal que considere las oraciones preteóricamente válidas tanto válidas en toda estructura como válidas en el sentido informal en que se consideraban las reglas de inferencia de Frege.[54]

Pasaré a la exposición de la tercera premisa del argumento de Kreisel:

(P3) $\forall\alpha(\mathrm{V}(\alpha^1) \to \mathrm{D}(\alpha^1))$

elección de *D2*, sin embargo, se suele aceptar dichos axiomas, debido a la utilidad que estos tienen, como poder introducir letras de predicado o de función en lugar de fórmulas. Este es un caso en el que, tradicionalmente, se ha optado por ignorar ciertas intuiciones para favorecer otras.

[54] Me parece que en la práctica tiene mucha importancia el que una oración sea válida en el sentido informal en que se consideraban las reglas de inferencia de Frege. En general, ambas nociones de validez son propiedades necesarias de la noción preteórica. Sin embargo, creo que las oraciones y reglas de inferencia que se consideran intuitivas en el sentido de las reglas de Frege lo son, aunque no lo hayan sido antes. Por ejemplo, la inferencia de 'Algún hombre es mortal' a partir de la oración 'Todos los hombres son mortales', no se considera una inferencia válida, aunque lo fue (según Aristóteles). Creo que la extensión de la noción de validez en el sentido informal en que lo eran las reglas de Frege cambia dependiendo de distintos factores (muchos de ellos históricos).

Esta última premisa es una instancia del teorema de compleción para la lógica de primer orden (teorema 1.4):

Si $\Gamma \vDash^1 \varphi$ entonces $\Gamma \vdash^1 \varphi$

Considerando el caso en que Γ es el conjunto vacío de fórmulas y que α es una verdad lógica (es decir, es verdadera en todas la estructuras), tal como se señaló en el capítulo 1, es claro que (P3) se sostiene. No hay mucho más que decir sobre (P3), por lo cual el argumento de Kreisel sería el siguiente:

1. $\forall i \forall \alpha (\mathrm{Val}(\alpha^i) \rightarrow \mathrm{V}(\alpha^i))$
2. $\forall i \forall \alpha (\mathrm{D}(\alpha^i) \rightarrow \mathrm{Val}(\alpha^i))$
3. $\forall \alpha (\mathrm{V}(\alpha^1) \rightarrow \mathrm{D}(\alpha^1))$

El argumento desglosado sería de la siguiente manera:

1.	$\forall i \forall \alpha (\mathrm{Val}(\alpha^i) \rightarrow \mathrm{V}(\alpha^i))$	Premisa
2.	$\forall i \forall \alpha (\mathrm{D}(\alpha^i) \rightarrow \mathrm{Val}(\alpha^i))$	Premisa
3.	$\forall \alpha (\mathrm{V}(\alpha^1) \rightarrow \mathrm{D}(\alpha^1))$	Premisa
4.	$\forall \alpha (\mathrm{D}(\alpha^1) \rightarrow \mathrm{Val}(\alpha^1))$	Instanciación universal en 2, poniendo 1 en lugar de i
5.	$\forall \alpha (\mathrm{Val}(\alpha^1) \rightarrow \mathrm{V}(\alpha^1))$	Instanciación universal en 1, poniendo 1 en lugar de i
6.	$(\mathrm{D}(\alpha^1) \rightarrow \mathrm{Val}(\alpha^1))$	Instanciación universal en 4, dejando α^1 libre
7.	$(\mathrm{Val}(\alpha^1) \rightarrow \mathrm{V}(\alpha^1))$	Instanciación universal en 5, dejando α^1 libre
8.	$(\mathrm{V}(\alpha^1) \rightarrow \mathrm{D}(\alpha^1))$	Instanciación universal en 3, dejando α^1 libre
9.	$(\mathrm{V}(\alpha^1) \rightarrow \mathrm{Val}(\alpha^1)$	Transitividad del condicional en 6 y 8
10.	$(\mathrm{Val}(\alpha^1) \leftrightarrow \mathrm{V}(\alpha^1)$	Introducción de la equivalencia material 7 y 9
11.	$\forall \alpha (\mathrm{Val}(\alpha^1) \leftrightarrow \mathrm{V}(\alpha^1)$	Universalización en 10

Este argumento muestra que (PK) se sostiene para el caso de la lógica de primer orden. Sin embargo, es importante notar que, el argumento apela al teorema de compleción para la lógica de primer orden. Por lo cual, este argumento no se puede usar para demostrar que (PK) se cumple para la lógica de segundo orden (ni de ningún lenguaje de orden superior), ya que no cumple con el teorema de compleción, aunque Kreisel agrega "pero se debe esperar una [prueba]" (Kreisel 1967: 157).

3.1.1. Otra presentación del Principio de Kreisel

Para la lógica de primer orden, es suficiente usar las versiones anteriores tanto del principio como del argumento de Kreisel. Como vimos, no es posible usar el argumento de Kreisel para establecer el principio para el caso de la lógica de segundo orden, pues no cumple con el teorema de compleción. Sin embargo, no es posible tampoco usar la versión del principio de Kreisel expuesta anteriormente para rescatar la idea de la adecuación extensional de la definición de consecuencia lógica tarskiana. Usaremos los lenguajes *L1K=* y *L2K* del primer capítulo para hacer las siguientes observaciones. Primero, es bien conocido que la lógica clásica (en particular la lógica de segundo orden) satisface lo siguiente:

> **Propiedad de monotonicidad:** Para cualesquiera conjuntos de fórmulas Γ y Δ, y para cualquier fórmula φ, si $\Gamma \vDash \varphi$ entonces $\Gamma \cup \Delta \vDash \varphi$

Es decir, si una fórmula φ es consecuencia lógica de un conjunto de fórmulas Γ, entonces agregar nuevas premisas no modifica la conclusión. Dado que toda fórmula de *L1K=* es una fórmula de *L2K*, se cumple también para el lenguaje *L1K=* de primer orden. Usando lo demostrado en el capítulo 1 es posible demostrar lo siguiente:

Teorema 3.1. *Sean Γ un conjunto de fórmulas de L1K= y φ una fórmula de L1K=, si $\Gamma \vDash \varphi$ entonces existe una fórmula ψ=$(\gamma_1 \supset (\gamma_2 \supset (\cdots \supset \gamma_n)\dots)) \supset \varphi$, tal que $\vDash \psi$, donde $\gamma_1, \gamma_2 \dots \gamma_n \in \Gamma$*

Demostración. Sean Γ un conjunto de fórmulas de *L1K=* y φ una fórmula de *L1K=*, tales que $\Gamma \vDash \varphi$. Sabemos que Γ es finito o Γ es infinito

Caso 1: Si Γ es finito, sea n=$|\Gamma|$, entonces $\gamma_1, \gamma_2 \dots \gamma_n \vDash \varphi$. Por el teorema de la deducción (semántico), $\gamma_1, \gamma_2 \cdots \vDash (\gamma_n \supset \varphi)$, por lo tanto aplicando el teorema de la deducción semántico n-veces, $\vDash (\gamma_1 \supset (\gamma_2 \supset (\cdots \supset \gamma_n)\dots)) \supset \varphi$ que era la fórmula que requeríamos

Caso 2: Si Γ es infinito, entonces no es posible utilizar el argumento anterior, pues toda fórmula es finita. Por el teorema de compacidad, existe un conjunto finito $\Gamma_0 \subseteq \Gamma$, tal que $\Gamma_0 \vDash \varphi$, como Γ_0 es finito y es posible aplicar el argumento del caso 1 a este conjunto Γ_0.

□

Por la propiedad de monotonicidad, aún si Γ es infinito, la oración ψ es una consecuencia lógica de Γ y es posible caracterizar todas las consecuencias de Γ mediante una oración que es una verdad lógica. Es decir, podemos reducir la noción de consecuencia lógica a la noción de verdad lógica. De esta manera, es suficiente sólo considerar el conjunto de las verdades lógicas de *L1K=* para poder enunciar el Principio de Kreisel.

Sin embargo, no es posible demostrar esto para el caso de la lógica de segundo orden, pues ésta no satisface el teorema de compacidad. De hecho, consideremos el siguiente conjunto de fórmulas de *L2K* (y de *L1K=*):

$$\Gamma = \{\exists x_1 \exists x_2 (x_1 \neq x_2), \exists x_1 \exists x_2 \exists x_3 (x_1 \neq x_2 \wedge x_2 \neq x_3 \wedge x_1 \neq x_3), \exists x_1 \exists x_2 \exists x_3 \exists x_4 (x_1 \neq x_2 \wedge x_2 \neq x_3 \wedge x_3 \neq x_4 \wedge x_1 \neq x_3 \wedge x_2 \neq x_4 \wedge x_1 \neq x_4) \ldots\}$$

El conjunto Γ tiene a todas y cada una de las fórmulas que afirman que el dominio tiene al menos n elementos: es decir, para cada $n \in \omega$, Γ tiene la fórmula que afirma "hay al menos n elementos en el dominio". Ahora consideremos la fórmula de segundo orden:

$$\varphi = \forall F \neg(\forall x \forall y (Fx = Fy \supset x = y) \wedge \exists x \forall y (Fy \neq x))$$

Como mencioné en el capítulo 1, la fórmula φ afirma que hay una infinidad de elementos en el dominio. Las estructuras que satisfacen a todos las fórmulas de Γ son aquellas cuyo dominio es infinito, por lo tanto, $\Gamma \vDash \varphi$. Sin embargo, ningún subconjunto finito de Γ, Γ_0, es tal que $\Gamma_0 \vDash \varphi$, pues cada subconjunto finito de Γ sólo tiene modelos cuyo dominio tiene una cantidad finita de elementos, es decir, cada subconjunto finito de Γ sólo tiene modelos finitos que son contraejemplos de φ. Para construir la fórmula ψ, requerimos que el conjunto Γ sea finito, para garantizar que dicha fórmula sea finita. De otra forma, no es posible construir la fórmula ψ pues las fórmulas son finitas por definición.

La noción de consecuencia lógica no se reduce a la noción de verdad lógica para el caso de la lógica de segundo orden. Por esto, necesitamos una manera de plantear el Principio de Kreisel en términos de consecuencia lógica. Simplemente, consideraremos los predicados mencionados anteriormente como relaciones. Sean Γ^i un conjunto de fórmulas y sea α^i una oración, donde de nuevo el índice i indica el orden de cuantificación de las fórmulas:

- Val(Γ^i, α^i) es la afirmación de que α^i es una consecuencia lógica intuitiva de Γ^i (o bien, que toda interpretación en donde todos los elementos de Γ^i son verdaderos, es una interpretación donde α^i es verdadera)
- V(Γ^i, α^i) es la afirmación de que α^i es verdadera en cualquier interpretación cuyo dominio es un conjunto donde todos los elementos de Γ^i son verdaderas (o bien, que $\Gamma^i \vDash \alpha^i$)
- D(Γ^i, α^i) significa que α^i es deducible de Γ^i en una sistema formal adecuado de reglas de inferencia ($\Gamma^i \vdash \alpha^i$)

Por lo tanto, el Principio de Kreisel (usando consecuencia lógica) es:

(PKC) $\forall i \forall \Gamma \forall \alpha (\text{Val}(\Gamma^i, \alpha^i) \leftrightarrow \text{V}(\Gamma^i, \alpha^i))$

Entonces, mi caso que de interés es:

(PKC2) $\forall \Gamma \forall \alpha (\text{Val}(\Gamma^2, \alpha^2) \leftrightarrow \text{V}(\Gamma^2, \alpha^2))$

3.2. La noción (NP) y el argumento de Kreisel

Usaré la noción (NP) del capítulo anterior y daré una versión del argumento de Kreisel para el caso de la lógica de primer orden. Argumentos similares, aunque no tan desglosados, son presentados por Shapiro (1987, 1991) y por Gómez Torrente (2000b, 2008b), así que me apoyaré en ellos constantemente debido a que las justificaciones son esencialmente las mismas. Recordemos la noción (NP):

(NP) Φ es una consecuencia lógica de Γ si Φ vale en todas las posibilidades bajo toda interpretación de la terminología no lógica en la que Γ vale

Habíamos visto que la relación entre esta noción preteórica y la noción de validez en estructuras cuyo dominio es una clase, consiste simplemente en considerar que todas las interpretaciones posibles son las clases propias. Así, la noción de validez con la que interpretaré el predicado $\mathrm{Val}(\alpha^i)$ será:

(NPC) Φ es una consecuencia lógica$_{NPC}$ de Γ si Φ vale en todas las estructuras $\mathfrak{A}^C$ cuyo dominio es una clase en la que Γ vale

No he dado una definición precisa de $\mathfrak{A}^C$; por el momento, consideraremos que es un par $\langle C^{\mathfrak{A}}, I^{\mathfrak{A}}\rangle$ donde C es una clase e I es un función de interpretación. Como sabemos, toda estructura tiene una función de interpretación que da una interpretación a la terminología no lógica, así que no es necesario especificarla pues es esencialmente la misma formulación explicada en el capítulo 1. Más adelante discutiré algunos problemas que produce la introducción de clases y cómo rescatar el contenido de estas afirmaciones. Así pues, se debe entender que las demostraciones con respecto a la verdad de una oración en una estructura cuyo dominio es una clase serán informales y poco precisos. Finalmente, el objetivo sólo es mostrar que es posible demostrar el argumento de Kreisel usando la noción (NPC).

La versión del argumento de Kreisel que presentaré consiste en relacionar la noción (NPC) con la definición proporcionada en el capítulo 1 de consecuencia lógica tarskiana. Definiremos (sólo por cuestión de notación) las relaciones explicadas en la sección anterior de la siguiente manera:

- $\mathrm{Val}(\Gamma^i, \alpha^i)$ si y solo si α^i es consecuencia lógica$_{NPC}$ de Γ
- $\mathrm{V}(\Gamma^i, \alpha^i)$ si y solo si $\Gamma \vDash \alpha$
- $\mathrm{D}(\Gamma^i, \alpha^i)$ si y solo si $\Gamma \vdash \alpha$

La noción (NPC) es más precisa que la noción preteórica de consecuencia lógica, sin embargo, en la discusión del capítulo anterior nos permite observar la relación entre estas dos nociones. Así mismo, vale la pena mencionar que, dado que en este trabajo me concentro en los casos de las lógicas de primer y de segundo orden, los índices i sólo serán 1 o 2.

La primera premisa del argumento de Kreisel sería entonces:

(P1) $\forall i \forall \Gamma \forall \alpha (\mathrm{Val}(\Gamma^i, \alpha^i) \rightarrow \mathrm{V}(\Gamma^i, \alpha^i))$

Para cualquier conjunto Γ de fórmulas de *L2K* y cualquier fórmula α de *L2K*. Si una oración α es consecuencia lógica$_{NPC}$ de Γ implica que en toda estructura cuyo dominio es una clase donde las oraciones de Γ son verdaderas, la oración α es verdadera. Ahora, si esto es así, en toda estructura cuyo dominio es un conjunto donde las oraciones de Γ son verdaderas, la oración α es verdadera, esto simplemente debido a que todo conjunto es una clase. El argumento es el siguiente: supongamos que $\mathrm{Val}(\Gamma^2, \alpha^2)$ y supongamos que para cada $\gamma \in \Gamma$, γ es verdadera en alguna estructura $\mathfrak{A}$, como el dominio y la función I de $\mathfrak{A}$ son conjuntos, el dominio y la función I de $\mathfrak{A}$ son clases y, por el primer supuesto, α debe ser verdadera en $\mathfrak{A}$. Por lo tanto, $\mathrm{V}(\Gamma^i, \alpha^i)$, pues $\mathfrak{A}$ fue arbitraria. Como esto los índices considerados son sólo 1 y 2, este argumento garantiza que la premisa se sostiene en general (de hecho, se sostendría considerando índices de cualquier orden finito). Este tipo de justificación de la primera premisa la presenta Gómez Torrente en (2008b: 208).

Shapiro, por otra parte, parece proporcionar dos justificaciones: por un lado, postula que la premisa es simplemente obvia, sin mayor justificación (cf. Shapiro 1987: 310), mientras que, por otro lado, asevera que esta premisa sólo hace la afirmación obvia de que "todo modelo conjuntista es una interpretación legítima del lenguaje" (Shapiro 1991: 142). Me inclino a pensar que en el primer caso, la justificación es bastante similar a la ofrecida por Gómez Torrente dado que en el artículo (Shapiro 1987), presenta el principio utilizando la noción de validez en estructuras cuyo dominio es una clase. Así mismo, cuando nos presenta los problemas que involucraría que el Principio de Kreisel fuese falso afirma que: "Esto es, asumir que hay una oración que es verdadera en toda interpretación teórico conjuntista, pero que es falsa en alguna otra interpretación (digamos una interpretación cuyo dominio es una clase propia)" (Shapiro 1987: 309-310). En el segundo caso, la idea de fondo sería (me parece) muy similar a la expuesta en la justificación original de Kreisel (explicada en la primera sección de este capítulo) pues, siguiendo la idea de Shapiro, esta premisa sería la afirmación de que dentro de las interpretaciones que consideramos adecuadas están las estructuras conjuntistas. En cualquier caso, me parece claro que Shapiro no tendría mayor problema para justificar la primera premisa del argumento de Kreisel. Esto porque si consideramos que todas las interpretaciones son precisamente todas las estructuras cuyo dominio es una clase, entonces inmediatamente se acepta que las estructuras cuyo dominio es un conjunto son interpretaciones legítimas del lenguaje; y esto es lo que afirma la premisa (P1).

La segunda premisa del argumento sería entonces:

(P2) $\forall i \forall \Gamma \forall \alpha (\mathrm{D}(\Gamma^i, \alpha^i) \rightarrow \mathrm{Val}(\Gamma^i, \alpha^i))$

La justificación de esta premisa consiste en demostrar una versión del teorema de corrección (Teorema 1.10 en el capítulo 1) para el caso de la validez en estructuras cuyo dominio es una clase. Para ello, será necesario mostrar que lemas análogos al caso del teorema de corrección se pueden mostrar. Así sólo presentaré un esbozo de estos lemas. Definiré que una oración φ es universalmente válida$_{NIC}$ si es consecuencia lógica$_{NIC}$ del conjunto vacío de premisas. Por conveniencia, y dado que sólo nos ocupamos de la lógica de primer y segundo orden, explicaremos los lemas para el sistema *D2*, pues el sistema *D1* es un subconjunto de este:

Lema 3.1. *Todo axioma es universalmente válido$_{NPC}$*

Los esquemas de axioma 1-6 del sistema *D1* son exactamente análogos al caso del lema 1.1 del capítulo 1. Simplemente hay que tomar en cuenta que el dominio de una estructura $\mathfrak{A}^C$ es una clase, y no hay mayor dificultad con ello.[55] El mismo tipo de observación se presenta para el caso de los axiomas 1-6 del sistema *D2*, pues sus demostraciones son versiones de los casos de *D1*. Los casos del axioma de comprehensión y del axioma de elección son un poco más complicados. Al igual que en el caso de la demostración normal, es necesario tener un esquema de comprehensión y un axioma de elección en la metateoría (que más adelante veremos, será la teoría de conjuntos en segundo orden ZFC2), en particular necesitaremos de un axioma de elección global que cuantifique sobre clases, al igual que el axioma de comprehensión. Sin embargo, ambos axiomas son consecuencias inmediatas de sus respectivas versiones en la metateoría.

Lema 3.2. *Para cualquier estructura $\mathfrak{A}^C$ si $\alpha \supset \beta$ es verdadera en $\mathfrak{A}^C$ y α es verdadera en $\mathfrak{A}^C$ entonces β es verdadera en $\mathfrak{A}^C$*

De nuevo, esto es muy fácil de demostrar y dependen sólo de condiciones de satisfacción análogas a las condiciones 3 y 4 de la sección 1.1.1 del capítulo 1.

La consecuencia de estas observaciones es que la premisa (P2) se sostiene. Si bien es cierto que este tipo de justificación no es la que Kreisel tenía en mente, creo que es posible justificar cada axioma de manera parecida a como Kreisel lo hizo. Finalmente, cada axioma del sistema *D2* fue seleccionado para cumplir un rol importante en el sistema deductivo, y para rescatar ciertas inferencias consideradas importantes (véase la explicación proporcionada de los axiomas en el capítulo 1). Así mismo, si lo que dije en el capítulo anterior es adecuado, existe una relación entre la noción (NPC) y la relación de validez intuitiva. A mi manera de ver, y recuperando ambos puntos, la justificación intuitiva de los axiomas (de la manera en cómo la ejemplifique en la sección 2.1 del capítulo anterior) estará dada, probablemente, en términos de las consideraciones hechas en el capítulo1. Así pues, continuaré especificando la premisa tres del argumento:

(P3) $\forall\Gamma\forall\alpha(\mathrm{V}(\Gamma^1, \alpha^1) \rightarrow \mathrm{D}(\Gamma^1, \alpha^1))$

Que es literalmente, la enunciación del teorema de compleción para la lógica de primer orden. Así pues, y dada estas justificaciones, es posible obtener la conclusión esperada:

(PKC1) $\forall\Gamma\forall\alpha(\mathrm{Val}(\Gamma^1, \alpha^1) \leftrightarrow \mathrm{V}(\Gamma^1, \alpha^1))$

Habiendo establecido esto, en el resto del capítulo me ocuparé del caso de la lógica de segundo orden.

[55]Esto es fácil de ver pues dado que las funciones I de una estructura tienen como dominio el conjunto K de terminología no lógica (y esta es numerable) I es un conjunto. La misma observación se puede hacer para el caso de las funciones $s \in \Sigma^{\mathfrak{A}}$, que también son conjuntos. En este sentido, los cambios en la demostración son simplemente terminológicos.

3.3. El Principio de Kreisel y el Principio de Reflexión

Como vimos, el caso de la lógica de segundo orden es problemático, ya que no es posible usar el argumento de Kreisel para recuperar la noción de consecuencia lógica en su contenido. Shapiro (1987) estudió el Principio de Kreisel y formuló versiones de éste dentro de un lenguaje formal, y estableció relaciones entre estas versiones y distintas formulaciones del Principio de Reflexión para la teoría de conjuntos. En adelante, estudiaré estas formulaciones presentadas por Shapiro y expondré qué relaciones existen entre éstas y el Principio de Reflexión. La ventaja de adoptar esta estrategia es que nos permite exponer el Principio de Kreisel dentro de un marco formal, preciso y matemáticamente manejable. Esto nos lleva a la necesidad de utilizar un aparato formal más fuerte y más complejo como metateoría: ZFC2. La teoría de conjuntos ZFC2 tiene los mismos axiomas de ZFC salvo que el esquema de comprehensión y el esquema de reemplazo, se sustituyen por fórmulas de segundo orden que cuantifican sobre subconjuntos y sobre funciones respectivamente.

El lenguaje de esta teoría será un lenguaje de segundo orden sin terminología no lógica (a excepción de $\in$), es decir, será el lenguaje *L2* (que es el lenguaje *L2K* sin el conjunto de terminología no lógica K) más el símbolo de pertenencia. Debido a que en estricto sentido el lenguaje *L2* contiene como conectivas sólo a la negación y al condicional habrá que especificar claramente cómo traducir las fórmulas de *L2* a fórmulas con otras conectivas, en particular con *vee* y $\wedge$ (estas fórmulas serán importantes en el siguiente capítulo 4). Para cada fórmula φ de *L2* que contiene cuantificadores, consideraremos su forma normal prenexa, la única fórmula equivalente a φ de la fórmula $\forall X \exists Y \ldots \forall x \exists y \psi$, donde $\forall X \exists Y \ldots \forall x \exists y$ es una secuencia con todos los cuantificadores en φ (ya sean de primer o de segundo orden) y ψ es una fórmula sin cuantificadores. Hay que mencionar que para toda fórmula de *L2*, existe una fórmula con esta forma y se puede obtener mediante la aplicación continua de las siguientes reglas: $\forall x[\exists x](\varphi \supset \psi) \equiv \exists x[\forall x]\varphi \supset \psi$ y $\varphi \supset \forall x[\exists x]\psi \equiv \forall x \exists x(\varphi \supset \psi)$. Para cada fórmula φ de *L2*, consideremos la siguientes clausulas:

1. Si φ es de la forma $t = t'$ o $T(t)$ (donde t y t*'* son términos cualesquiera del lenguaje y T es una constante de predicado o variable de predicado), entonces φ^P son $t = t'$ y $t \in T$ respectivamente.

2. Si φ es de la forma $\neg(t = t')$ o $\neg T(t)$, entonces φ^P son $t \neq t'$ y $t \notin T$ respectivamente.

3. Si φ es $\alpha \supset \beta$, entonces φ^P es $(\neg\alpha)^P \vee \beta^P$

4. Si φ es $\forall x[X]\psi$ o $\exists x[X]\psi$, entonces φ^P es $\forall x[X]\psi^P$ o $\exists x[X]\psi^P$

Intuitivamente, la fórmula φ^P es resultado de sustituir todos condicionales de la fórmula φ por disyunciones. Como resultado, φ^P tendrá sólo disyunciones, conjunciones y todas las negaciones se pueden reducir a átomos del tipo $t \neq t'$ y $t \notin T$. Dado que en el lenguaje de ZFC2 hay fórmulas que cuantifican sobre conjuntos de conjuntos (las fórmulas de segundo orden), se considera que ZFC2 es efectivamente una teoría de clases, por lo cual será de utilidad en lo que sigue. Propiamente, los lemas necesarios para demostrar (P2) en la sección anterior se pueden demostrar en ZFC2, donde el axioma de comprehensión es una fórmula cuyo rango considera clases.

Antes de presentar las formulaciones de Shapiro del Principio de Kreisel, es necesario ampliar ZFC2 para incluir fórmulas que explícitamente nos permitan expresar la noción de satisfacción en clases. La razón de esto es que en ZFC2 no hay una fórmula que caracteriza esta noción, sin embargo, es posible aumentar esta teoría con fórmulas que hagan este trabajo. La metateoría aumentada no será particularmente complicada de usar y de hecho tiene esencialmente los mismos modelos que ZFC2. En la siguiente sección introduciré esta nueva notación y en lo que sigue presentaré las formulaciones de Shapiro del Principio de Kreisel.

3.3.1. La teoría ZFC2+

La exposición de estas observaciones puede encontrarse en Shapiro (1991). El lenguaje que usaremos será el lenguaje *L2K*, de segundo orden. Como sabemos, el lenguaje *L1K=* es subconjunto de este lenguaje. Asumiré una aritmetización del lenguaje *L2K* mediante una numeración de Gödel.[56] Para todo $n \in \omega$, denotaremos como Φ_n a la expresión de *L2K* cuyo número de Gödel es n.

Definiremos satisfacción para esta teoría de conjuntos, por lo cual, las variables de primer orden tendrán como asignación conjuntos y las variables de segundo orden conjuntos de conjuntos. Es decir, utilizaremos las variables de primer orden para referirnos a conjuntos y las variables de segundo orden se referirán a clases. Definimos $\langle x \rangle_n$ la secuencia compuesta por las primeras n variables de nuestra enumeración de variables de primer orden, es decir, $\langle x \rangle_n = \langle x_1, \ldots, x_n \rangle$, así, para cada $s \in \Sigma^{\mathfrak{A}^C}$, $s(\langle x \rangle_n)=\langle s(x_1), \ldots, s(x_n) \rangle$. Definimos las siguientes fórmulas:

$$s_1(s,p) \equiv_{def} s \in \Sigma^{\mathfrak{A}^C} \text{y } ran(s) \subseteq p$$
$$s_1(s,P) \equiv_{def} s \in \Sigma^{\mathfrak{A}^C} \text{y } ran(s) \subseteq P$$

Siguiendo lo dicho anteriormente, p es un conjunto y P es una clase. Estas fórmulas permiten expresar en nuestro lenguaje el contenido de las funciones s distinguiendo entre aquellas cuyo rango es un conjunto y aquellas cuyo rango es una clase y permiten expresar la noción de asignación de un objeto a un término. Por supuesto, dado que el dominio de s es el conjunto de variables, por el axioma de remplazo $ran(s)$ es un conjunto, incluso en el segundo caso. Para realizar un procedimiento similar para el caso de las variables de segundo orden, definiremos las funciones $S \in \Sigma^{\mathfrak{A}}$ de otra manera. Dado que a cada variable de segundo orden se le asigna una clase, cada asignación deberá ser una clase de clases, lo cual nos comprometería con que la metateoría sea de tercer orden. Para toda $S \in \Sigma^{\mathfrak{A}^C}$, codificaremos el contenido de S de la siguiente manera: S es una colección de pares, tal que si $\langle q, r \rangle \in S$, entonces q es una variable de relación o una variable de función y

[56]Formalmente, una numeración de Gödel es una función cuyo dominio son las expresiones del lenguaje *L2K* y cuyo rango son los números naturales. Una numeración de Gödel asigna a cada expresión del lenguaje *L2K* un número natural.

- Si q es una variable de relación de aridad n, entonces r es una n-tupla y $S(q) = \{r|\langle q,r\rangle \in S\}$
- Si q es una variable de función, entonces r es una $n{+}1$-tupla y $\{r|\langle q,r\rangle \in S\}$ es la descripción de una función g, y $S(q){=}g$

La idea es "codificarçada función S de tal manera que no sea necesario agrupar en un conjunto los valores de q, es decir, nunca es necesario agrupar en un conjunto $\langle q,r\rangle \in S$. Esta codificación permite identificar el valor de las variables de segundo orden sin necesidad de usar variables de tercer orden. Análogamente al caso de las variables de primer orden, $\langle X\rangle_n = \langle X_1,\ldots,X_n\rangle$, y para cada $S \in \Sigma^{\mathfrak{A}}$, $S(\langle X\rangle_n){=}\langle S(X_1),\ldots,S(X_n)\rangle$, de la misma forma $\langle F\rangle_n = \langle F_1,\ldots,F_n\rangle$, así, para cada $S \in \Sigma^{\mathfrak{A}}$, $S(\langle F\rangle_n){=}\langle S(F_1),\ldots,S(F_n)\rangle$. Esto tiene el problema de que S es una clase, posiblemente propia, por lo cual, será necesario distinguir entre las funciones S que serán conjuntos y las que serán clases. Usaremos la notación donde s es una función que es un conjunto y S serán las funciones que de hecho son una clase. De la misma forma que antes:

$S_2(s,p) \equiv_{def} s \in \Sigma^{\mathfrak{A}}$ y

$$\begin{aligned}&s(x) = d \in p\\&s(X^n_m) = D \subseteq p\\&s(F^n_m) = G : p^m \to p\end{aligned}$$

$S_2(s,P) \equiv_{def} s \in \Sigma^{\mathfrak{A}}$ y

$$\begin{aligned}&s(x) = d \in P\\&s(X^n_m) = D \subseteq P\\&s(F^n_m) = G : P^m \to P\end{aligned}$$

$S_2(S,p) \equiv_{def} s \in \Sigma^{\mathfrak{A}}$, $S \neq \varnothing$ y

$$\begin{aligned}&S(x) = d \in p\\&S(X^n_m) = D \subseteq p\\&S(F^n_m) = G : p^m \to p\end{aligned}$$

$S_2(S,P) \equiv_{def} S \in \Sigma^{\mathfrak{A}}$, $S \neq \varnothing$ y

$$\begin{aligned}&S(x) = d \in P\\&S(X^n_m) = D \subseteq P\\&S(F^n_m) = G : P^m \to P\end{aligned}$$

Nótese que en el último caso no hay garantía de que S sea un conjunto; de hecho, estos casos engloban todos los casos en que se trabaja en clases propias. Siguiendo adelante, sea Φ una fórmula y sea t un término que refiere a una clase o a un conjunto, es decir, t es un término de primer orden que refiere a un conjunto, una variable de predicado monádico o bien es un símbolo no lógico de predicado. En los últimos dos casos, usaré las siguientes abreviaciones:

- $x \in t \equiv_{def} tx$

- $X^n \subseteq t^n \equiv_{def} \forall\langle x\rangle_n(X^n(\langle x\rangle_n) \supset (x_1 \in t \wedge \cdots \wedge x_n \in t))$
- $f^n|t \equiv_{def} \forall\langle x\rangle_n((x_1 \in t \wedge \cdots \wedge x_n \in t) \supset f^n(\langle x\rangle_n) \in t)$

Llamaremos Φ/t a la *relativización* de Φ a t y se define por recursión:

1) Si Φ es una fórmula atómica, entonces $\Phi/t = \Phi$

2) $(\neg\Phi)/t = \neg\Phi/t$

3) $(\Phi \supset \Psi)/t = (\Phi/t \supset \Psi/t)$

4) $(\forall x\Phi)/t = \forall x(x \in t \supset \Phi/t)$

5) $(\forall X^n\Phi)/t = \forall X^n(X^n \subseteq t^n \supset \Phi/t)$

6) $(\forall F^n\Phi)/t = \forall F^n(F^n|t \supset \Phi/t)$

Las fórmulas Φ/t son fórmulas en la metateoría que afirman que una fórmula Φ es el caso en t. La noción de satisfacción en estructuras cuyo dominio es un conjunto es definible en ZFC. Sea $\Phi_n(\langle X\rangle_m, \langle F\rangle_l, \langle x\rangle_k)$ una fórmula que sólo tiene como variables libres las indicadas por $\langle X\rangle_m$, $\langle F\rangle_l$, $\langle x\rangle_k$. La *fórmula de Tarski para satisfacción conjuntista* de Φ_n es

$$(s_1(s,p) \wedge S_2(s',p)) \supset (\mathbf{sats}(p,s,s',n) \equiv \Phi_n(s'\langle X\rangle_m, s'\langle F\rangle_l, s\langle x\rangle_k)/p)$$

Así, la formula $\mathbf{sats}(p,\ s,\ s',\ n)$ afirma que la fórmula con número de Gödel n es satisfecha en el conjunto p, con las funciones de asignación s y s'. Todas estas fórmulas se pueden probar en la metateoría. Aumentaremos la metateoría ZFC2 con nuevos axiomas que nos permitan tener fórmulas análogas para la satisfacción en clases. Agregaremos una nueva relación $\mathbf{SATS}(P,\ s,\ S,\ m)$ cuya interpretación pretendida es “la clase P satisface la fórmula Φ_m con las funciones de asignación s y S”. Los axiomas de esta nueva metatoría, a partir de ahora **ZFC2+**, son:

1) $\mathbf{SATS}(P,s,S,m) \supset m \in \omega$

2) $\mathbf{SATS}(P,s,S,m) \supset s_1(s,P) \wedge S_2(S,P)$

3) $(s_1(s,P) \wedge S_2(S,P)) \supset (\mathbf{SATS}(P,s,S,n) \equiv \Phi_n(S\langle X\rangle_m, S\langle F\rangle_l, s\langle x\rangle_k)/P)$ para toda fórmula Φ_n

Estas últimas las llamaremos *fórmula de Tarski para satisfacción en clases*. En esta nueva teoría ZFC2+ es posible definir una fórmula, $\mathbf{TR}(s,\ S,\ n)$, tal que:

$$\mathbf{TR}(s,S,n) \equiv_{def} \forall P(\forall x Px \supset \mathbf{SATS}(P,s,S,n))$$

que afirma que la fórmula Φ_n es satisfecha por todo el universo conjuntista. Es decir, todas las fórmulas:

$$\mathbf{TR}(s,S,n) \equiv \Phi_n(S\langle X\rangle_m, S\langle F\rangle_l, s\langle x\rangle_k)$$

son demostrables en ZFC2+. Así pues, es posible expresar en ZFC2+ la noción de verdad en el universo conjuntista. Esta teoría nos permitirá estudiar los Principios de Reflexión y su relación con el Principio de Kreisel.

3.3.2. Las formulaciones de (PK) de Shapiro

Ahora expondré las formulaciones que Shapiro (1987) presenta del Principio de Kreisel. Al igual que la anterior, esta sección será principalmente expositiva, pues el objetivo es presentar los resultados de Shapiro, y dejaré para el siguiente la investigación de su relevancia en el presente trabajo. La importancia de estas formulaciones es que son equivalentes al Principio de Reflexión en ZFC2, y por lo tanto la justificación del Principio de Kreisel dependerá de la justificación del Principio de Reflexión. Así pues, todo lo mencionado aquí puede confrontarse con Shapiro (1987). Ahora bien, aunque Shapiro demuestra una serie de resultados muy interesantes con respecto al poder deductivo de sus formulaciones, yo me concentraré en los resultados importantes relacionados con la satisfacción de estas formulaciones. Las fórmulas sobre las que definiremos el Principio de Kreisel son a lo más de segundo orden. De hecho, son todas fórmulas de *L2*, es decir, consideraremos un lenguaje que no tiene terminología no lógica.[57] Por supuesto, la metateoría que usaré será ZFC2, aunque más adelante usaré la teoría ZFC2+. Vale la pena notar que toda fórmula de *L2* es una fórmula de ZFC2, y cada fórmula de ZFC2 se puede traducir a una fórmula de *L2*: sea E una variable de relación nueva y sea Φ una fórmula de ZFC2, sea $\Phi[E]$ el resultado de sustituir en Φ todas las fórmulas $u \in t$ por Eut. El primer paso de Shapiro es reformular el principio en términos de satisfacción, por lo cual, definiremos los siguientes predicados:

- Sat(Φ) significa "hay una estructura cuyo dominio es una clase que satisface Φ"
- S(Φ) significa "hay una estructura cuyo dominio es un conjunto que satisface Φ"

De esta manera, el Principio de Kreisel se reduce a: Sat(Φ)$\leftrightarrow$S(Φ). A partir de los lemas para la demostración de (P2), podemos obtener que S(Φ)$\rightarrow$Sat(Φ); de nuevo, todo conjunto es una clase, por lo cual, si hay un conjunto que satisface Φ, hay una clase que satisface Φ. Por lo cual, el Principio de Kreisel se reduce a:

$$\forall\Phi(\mathrm{Sat}(\Phi)\rightarrow \mathrm{S}(\Phi))$$

El primer punto que hay que mencionar es que esta fórmula implica en ZFC2 la existencia de cardinales fuertemente inaccesibles. Sea Z2 la conjunción de todos los axiomas de ZFC2 y consideremos la fórmula $\exists E(\mathrm{Z2}[E])$, la generalización existencial de Z2. Nótese que ésta es una fórmula de *L2*. Esta fórmula es satisfecha por la jerarquía acumulativa de la teoría de conjuntos, es decir, Sat($\exists E(\mathrm{Z2}[E])$) es el caso. Por el Principio de Kreisel, S($\exists E(\mathrm{Z2}[E])$) es el caso. Es decir, hay un conjunto que satisface $\exists E(\mathrm{Z2}[E])$, que es equivalente a la existencia de un cardinal fuertemente

[57] No se pierde generalidad, pues toda fórmula de *L2K* puede transformarse en una fórmula de *L2* sustituyendo las constantes no lógicas por una nueva variable, tal que, la satisfacción de la fórmula original es equivalente a la satisfacción de la nueva fórmula. Dado que el conjunto K de terminología no lógica es el mismo que el del lenguaje *L1K=*, las nuevas variables serán a lo más de segundo orden. Así mismo, la función de interpretación de las estructuras serán superfluas pues no hay terminología no lógica que interpretar.

inaccesible. Por lo tanto, todas las formulaciones del Principio de Kreisel son independientes de ZFC2. Definimos, para cada α, número ordinal: $P(\alpha)$ es el mínimo cardinal inaccesible mayor que $\bigcup\{P(\beta)|\beta < \alpha\}$ y $EP(\alpha)$ es una fórmula que afirma que existe el mínimo cardinal inaccesible mayor que $\bigcup\{P(\beta)|\beta < \alpha\}$.[58] Es decir, $P(\alpha)$ denota al α-ésimo cardinal inaccesible y $EP(\alpha)$ afirma que éste existe.

3.3.2.1. La primera formulación del Principio de Kreisel

La formulación preliminar del Principio de Kreisel mencionada arriba no puede ser expresada literalmente en ZFC2, sino en la teoría ZFC2+, pues involucra la noción de satisfacción en estructuras cuyo dominio es una clase, sin embargo, es posible encontrar una fórmula de ZFC2 que rescata gran parte del contenido de esta fórmula: Φ/t. Como mencioné, para cualquier fórmula Φ, Φ/t denota a la relativización de Φ y afirma que t satisface Φ. Por lo tanto, siguiendo las convenciones hasta ahora usadas, la fórmula $\exists X(\Phi/X)$ afirma que hay una clase que satisface Φ y $\exists x(\Phi/x)$ afirma que existe un conjunto que satisface Φ. Así, la primera formulación del Principio de Kreisel (en esta versión más precisa) es:

(PK1) $\exists X(\Phi/X) \rightarrow \exists x(\Phi/x)$

Para toda fórmula Φ sin variables libres, es decir, para todo enunciado de *L2*. Llamaremos ZFC2 + PK1 a la teoría obtenida de ZFC2 agregando cada instancia de PK1.

Ahora bien, existe un principio en la teoría de conjuntos que afirma que si una fórmula es verdadera en el universo conjuntista, entonces existe un modelo estándar de la teoría de conjuntos[59] en que esa fórmula es verdadera. Para ser más precisos, el Principio de Reflexión afirma que si una fórmula φ es verdadera en V entonces existe un nivel de la jerarquía acumulativa V_α que satisface φ. Estos principios fueron formulados por Levy (1960) y pretenden rescatar la idea de que la jerarquía acumulativa de la teoría de conjuntos (V) no puede ser caracterizada por una fórmula. Los principios que Shapiro considera son principios sobre fórmulas de segundo orden, y el primero que considera es:

(PR1) $\Psi \rightarrow \exists x((Z2/x) \wedge (\Psi/x))$

Llamaremos ZFC2 + PR1 a la teoría obtenida de ZFC2 agregando todas las instancias de RP1. El primer resultado importante para mi propósito es la afirmación de que PK1 y PR1 son básicamente equivalentes en el marco de ZFC2:

Teorema 3.2. *ZFC2 + PK1 es equivalente a ZFC2 + PR1*

[58]Como mencioné en la introducción, este capítulo requiere un conocimiento avanzado en ZFC. Para revisar una breve explicación de los conceptos usados en esta última sección véase el Apendice del presente trabajo.

[59]Un modelo estándar de la teoría de conjuntos es un nivel de la jerarquía acumulativa indexado por un cardinal inaccesible. De hecho, sabemos que x satisface Z2 si y solo si x es isomorfo a un nivel indexado por un cardinal inaccesible.

Demostración. $\Leftarrow$ Probaremos primero el segundo condicional y trabajaremos en ZFC2 + PR1. Supongamos que $\exists X(\Phi/X)$ para alguna Φ de *L2*. Por PR1, obtenemos que $\exists x((Z2/x) \wedge (\exists X(\Phi/X)/x)$ es decir, hay un conjunto a tal que $(\mathrm{Z2}/a) \wedge (\exists X(\Phi/X)$ $/x)$. Como $(\mathrm{Z2}/a)$ sabemos que a es un nivel inaccesible y por $\exists X(\Phi/X)/a$ que hay un $X \subseteq a$ tal que $(\Phi/X)/a$ pues a satisface $\exists X(\Phi/X)$ (nótese que X debe ser un conjunto por el axioma de separación). Como a es inaccesible, $((\Phi/X)/a) \leftrightarrow (\Phi/X)$, pues es modelo de V por lo cual, las oraciones verdaderas en a son oraciones verdaderas en V, y como X es un conjunto, $\exists x(\Phi/x)$, que era lo que queríamos probar

$\Rightarrow$ Ahora trabajaremos en ZFC2 + PK1, y supondremos que una fórmula Ψ de ZFC2 es verdadera, por lo tanto la oración $\exists E((Z2 \wedge \Psi)[E])$ de *L2* es verdadera en V, es decir, $\exists X(\exists E((Z2 \wedge \Psi)[E])/X)$. Por PK1, tenemos $\exists x(\exists E((Z2 \wedge \Psi)[E])/x)$, es decir, hay un conjunto a y un conjunto e, tales que $((Z2 \wedge \Psi)[e])/a$, es decir, la estructura $\langle a, e\rangle$ satisface a $\mathrm{Z2}[e]$ y a Ψ[e]. Como $\langle a, e\rangle$ satisface a Z2, hay un nivel inaccesible $\langle b, \in\rangle$ isomorfo a $\langle a, e\rangle$ que satisface a Z2 y Ψ (nótese que estas son las fórmulas originales), por lo cual $\mathrm{Z2}/b$ y Ψ/b, por lo cual $\exists x((Z2/x) \wedge (\Psi/x))$, que era lo que queríamos mostrar. □

Como mencioné más arriba, hay varios resultados muy interesantes acerca del poder deductivo de ZFC2 + PK1. Por ejemplo, es posible mostrar que ZFC2 + PK1 es ω-incompleta, pues puede demostrar $\mathrm{EP}(n)$, para todo $n \in \omega$, pero no puede demostrar la fórmula $\forall n \in \omega EP(n)$. Sin embargo, remito al lector interesado a revisar la fuente original para comprender estos resultados, pues aquí me concentraré principalmente en presentar algunos resultados con respecto a la satisfacción de PK1.

Definición. *Definición: Sea α un ordinal, decimos que α es i-definible si y solo si, hay una fórmula Φ(x) (del lenguaje de ZFC2) con x como su única variable libre, tal que cumpla las siguientes condiciones:*

1) $\Phi(\alpha)$ es satisfecha por el universo y por todos los V_κ tal que κ es inaccesible y $\kappa \geq P(\alpha)$ y $\kappa > \alpha$

2) $\forall x(\Phi(x) \supset x = \alpha)$ es satisfecha por el universo y por todo V_κ tal que κ es inaccesible que contenga a α

La idea detrás de esta definición es que un ordinal es *i-definible* si es definible mediante su relación con el inaccesible indexado por ese ordinal. La lectura más "natural" de la definición es que un ordinal α es *i-definible* cuando es definido en todos los modelos de Z2 en los que se sabe que está el α-ésimo inaccesible. A partir de esta definición, es posible caracterizar el modelo mínimo de ZFC2 + PK1:

Teorema 3.3. *Sea α un ordinal y asumamos $EP(\alpha)$. Si α no es i-definible, entonces $V_{P(\alpha)} \vDash$ZFC2+PK1*

Demostración. Supongamos que α no es *i-definible*. Sabemos que $V_{P(\alpha)} \vDash ZFC2$. Por lo tanto, bastará con mostrar que $V_{P(\alpha)} \vDash PK1$ o equivalentemente, por el teorema anterior $V_{P(\alpha)} \vDash PR1$; supongamos que hay una fórmula Φ tal que $V_{P(\alpha)} \vDash$ $\neg(\Phi \to \exists x((Z2/x) \wedge (\Phi/x)))$. Por lo tanto, $V_{P(\alpha)}$ satisface tanto Φ como la fórmula $\forall x((Z2/x) \supset \neg(\Phi/x)))$, de esto se sigue que la fórmula

$$Ord(x) \wedge \forall\beta \in x(EP(\beta) \supset \neg(\Phi/V_{P(\beta)})) \wedge (\neg EP(x) \supset \Phi) \wedge (EP(x) \supset (\Phi/V_{P(x)}))$$

i-define a α, lo cual habíamos supuesto que no es caso. □

De esta manera, el mínimo modelo de PK1 es el nivel indexado por el δ-ésimo inaccesible, tal que δ es el mínimo ordinal no *i-definible*.

3.3.2.2. La segunda formulación del Principio de Kreisel

La segunda formulación del Principio de Kreisel, a diferencia de la primera, rescata todas las fórmulas de *L2*, no sólo los enunciados; es decir, en esta formulación consideraremos fórmulas con variables libres. Esta formulación es una extensión de la primera en tanto que se ocupa de fórmulas con variables aunque no necesariamente debe tenerlas, así, la teoría que llamaremos ZFC2 + PK2 es una extensión conservativa de ZFC2 + PK1.[60] Por simplicidad, sólo consideraremos fórmulas que tengan a lo más una variable libre de primer orden y una variable libre de segundo orden.[61] Sea $\Phi(Z, z)$ una fórmula de *L2* con a lo más Z y z variables libres, la segunda formulación de (PK) es:

$$\textbf{(PK2)}\ \forall X \forall Z \forall z \in X((\Phi(Z, z)/X) \supset \exists x \subseteq X(z \in x \wedge (\Phi(x \cap Z, z)/x)))$$

Informalmente, la fórmula afirma que si una clase satisface la fórmula $\Phi(Z, z)$ entonces hay un subconjunto de esa clase que satisface la fórmula. Llamaremos ZFC2 + PK2 a la teoría resultante de ZFC2 agregando todas las instancias de (PK2). Hay que notar que si Φ es un enunciado, entonces una instancia de (PK2) es una instancia de (PK1) así que ZFC2 + PK2 es una extensión conservativa de ZFC2 + PK1. El Principio de Reflexión correspondiente es entonces el siguiente:

$$\textbf{(PR2)}\ \forall X \forall y(\Psi(Y, y) \supset \exists x(y \in x \wedge ((Z2/x) \wedge \Psi(x \cap Y, y)/x)))$$

El principio afirma que si una fórmula Ψ es verdadera en el universo, hay un conjunto en él que la satisface. Llamaremos ZFC2 + PR2 a la teoría obtenida de ZFC2 agregando cada instancia de PR2. Como es de esperarse, ambos principios son esencialmente equivalentes:

Teorema 3.4. *ZFC2 + PK2 es equivalente a ZFC2 + PR2*

Demostración. La prueba es análoga a la del teorema 1

$\Rightarrow$ Trabajando en ZFC2 + PK2, y supongamos que una fórmula $\Psi(Y, y)$ de ZFC2 es verdadera para alguna clase Y y un conjunto y cualesquiera. La oración $\exists E(((Z2 \wedge \Psi)(Y, y))[E])$ de *L2* es satisfecha por V, es decir, $\exists X(\exists E(((Z2 \wedge \Psi)(Y, y))[E])/X)$. Por PK2, tenemos $\exists x \subseteq X(y \in x \wedge (\exists E(((Z2 \wedge \Psi)(Y, y))[E])/x))$, es decir, hay un conjunto a y un conjunto e, tales que $((Z2 \wedge \Psi)(a \cap Y, y))[e])/a$, es decir, la estructura $\langle a, e\rangle$ satisface a Z2$[e]$ y a $\Psi(a \cap Y, y)[e]$. Como $\langle a, e\rangle$ satisface a Z2, hay

[60] Es decir, pretendemos que en ZFC2 + PK2 se puedan demostrar todas las instancias de PK1.

[61] Shapiro hace la nota de que no se pierde generalidad pues “las variables libres pueden ser combinadas mediante el uso de funciones que las emparejen” (Shapiro 1987: 316)

un nivel inaccesible $\langle b, \in\rangle$ isomorfo a $\langle a, e\rangle$ que satisface a Z2 y $\Psi(a \cap Y, y)$ (nótese que estas son las fórmulas originales), por lo cual, Z2/b y $\Psi(b \cap Y, y)/b$, por lo cual $\exists x((Z2/x) \wedge (\Psi(x \cap Y, y)/x))$, que era lo que queríamos mostrar, simplemente se cuantifica universalmente sobre Y y y, pues fueron arbitrarias.

$\Leftarrow$ Trabajaremos en ZFC2 + PR2. Supongamos que $(\Phi(Z, z)/X)$ para alguna Φ de *L2*, para alguna clase X, $Z \subseteq X$ y $z \in X$ (consideraremos a X fija por simplicidad). $(\Phi/X)(Z, z)$, como es satisfecha por el universo. Por PR2, obtenemos que $\exists x(z \in x \wedge ((Z2/x) \wedge ((\Phi/X)(x \cap z, z)/x)))$ es decir, hay un conjunto a tal que $z \in a$ y $(Z2/a) \wedge ((\Phi/X)(a \cap Z, z)/a)$. Como $(Z2/a)$ sabemos que a es un nivel inaccesible y por $(\Phi/X)(a \cap Z, z)/a$ que $a \cap Z \subseteq a$ es un conjunto, y como $(\Phi/X)(a \cap Z, z) \leftrightarrow \Phi(a \cap Z, z)/X$ pues consideramos a X fija, $((\Phi/X)(a \cap Z, z)/a \leftrightarrow (\Phi(a \cap Z, z)/X)/a$. Ahora bien, $X \subset a$, por lo tanto, X es un conjunto, y como a es inaccesible, $(((\Phi(a \cap Z, z)/X)/a) \leftrightarrow (\Phi(a \cap Z, z)/X)$, pues es modelo de V. Por lo tanto, $(\Phi(a \cap Z, z)/X)$, y como X es un conjunto, $z \in X$ y trivialmente $X \subseteq X$, es decir, $\exists x \subseteq X(z \in x \wedge (\Phi(x \cap Z, z)/x))$ es el caso, que era lo que buscábamos probar □

De nuevo, no me detendré a estudiar la capacidad deductiva de ZFC2 + PK2, pero es importante mencionar que agregar PK2 a ZFC2 implica la existencia de cardinales Mahlo.[62] El principio PR2 es llamado *Principio de Reflexión parcial.* Existen otros principios que son conocidos como Principios de Reflexión completa, tales como los siguientes:

(PR2.1) $\exists x((Z2/x) \wedge \forall Y \forall y \in x(\Psi(Y, y) \leftrightarrow (\Psi(x \cap Y, y)/x)))$

(PR2.2) $\forall Y \exists x((Z2/x) \wedge \forall y \in x(\Psi(Y, y) \leftrightarrow (\Psi(x \cap Y, y)/x)))$

Ambos afirman, informalmente, que la satisfacción en el universo y la satisfacción en algún conjunto son equivalentes, la única diferencia es que el segundo considera la clase asignada a la variable Y como fija. La diferencia es relevante, pues se puede obtener los siguientes resultados:

Proposición 3.1. *(PR2.1) es inconsistente con ZFC2*

Proposición 3.2. *ZFC2 + PR2.2 es equivalente a ZFC2 + PR2*

Las demostraciones de estas proposiciones no son muy relevantes. Sin embargo, nos permiten observar que un Principio de Reflexión completa o es inconsistente, o no implica mayor capacidad deductiva que un principio parcial. Hablando de lo que nos interesa, trataremos de caracterizar el modelo mínimo de PK2.

Definición. *Sean $m \geq 1$, $j \in \omega$, el conjunto de fórmulas Π^m_j se define como el conjunto de fórmulas en forma normal prenexa cuyos cuantificadores son de orden m tal que si $\Phi \in \Pi^m_j$ entonces $\Phi = \forall X^m_1 \ldots \forall X^m_k \Psi$ tal que $\Psi \in \Sigma^m_{j-1}$*

[62]La razón de esto es que ya que PK2 y PR2 son esencialmente equivalentes, y que de PR2 se sigue que todo funcional F definido sobre la clase de ordinales (*Ord*) creciente que para cualquier ordinal límite λ, $F(\lambda) = \bigcup\{F(\beta) | \beta < \lambda\}$, tiene un punto fijo inaccesible. El conjunto de todos esos puntos fijos es un conjunto estacionario lo cual implica la existencia de un cardinal de Mahlo.

Definición. *Sean $m \geq 1$, $j \in \omega$, el conjunto de fórmulas Σ_j^m se define como el conjunto de fórmulas en forma normal prenexa cuyos cuantificadores son de orden m tal que si $\Phi \in \Sigma_j^m$ entonces $\Phi = \exists X_1^m \dots \exists X_k^m \Psi$ tal que $\Psi \in \Pi_{j-1}^m$*

La idea es que el conjunto $\Sigma_0^1 = \Pi_0^1$ es el conjunto de fórmulas de *L2* que no tienen cuantificadores. Y, por ejemplo, si una fórmula $\Psi \in \Pi_3^2$ los cuantificadores de Ψ son de tercer orden, y Ψ es una fórmula de la forma $\forall X_{1_1}^2 \dots \forall X_{1_k}^2 \exists X_{2_1}^2 \dots \exists X_{2_l}^2 \forall X_{3_1}^2 \dots \forall X_{3_m}^2 \phi$. El subíndice j indica el número de bloque de cuantificadores ya sean universales o existenciales.

Definición. *Sea x un cardinal. Decimos que x es Π_j^m-describible con los parámetros $U_1 \dots U_n$ si hay una fórmula $\phi \in \Pi_i^m$ tal que*

1) $V_\alpha \vDash \phi(U_1 \dots U_n)$

2) $\forall \beta < \alpha, V_\beta \nvDash \phi(U_1 \cap V_\beta \dots U_n \cap V_\beta)$

Una instancia de esta definición es cuando no consideramos ningún parámetro. A partir de ahora, a menos que se indique lo contrario, trabajaremos con fórmulas sin parámetros. Sea $In(n, x)$ la fórmula que afirma que $n \in \omega$ y que x es un cardinal que no es Π_n^1- *describible*. Es posible mostrar que no existe un conjunto de los x tales que $\forall n \in \omega In(n, x)$, es decir, que esta colección es una clase propia. Los modelos de ZFC2 + PK2 están descritos en el siguiente teorema:

Teorema 3.5. *$V_{P(\alpha)} \vDash ZFC2 + PK2$ si y solo si α no es Π_0^2-describible*

3.3.2.3. La tercera formulación del Principio de Kreisel

Hasta ahora sólo nos hemos preocupado por la satisfacción de fórmulas. Sin embargo, recordando que la lógica de segundo orden no es compacta, es conveniente tener una formulación del Principio de Kreisel que hable sobre la satisfacción de conjuntos de fórmulas. En este caso, deberemos usar la teoría ZFC2+, pues el esquema Φ/t es insuficiente para este propósito. A partir de ahora presupondré de nuevo una numeración de Gödel y usaremos la notación introducida en la sección anterior. Definimos el predicado $L(x)$ como la afirmación de que x es el número de Gödel de una fórmula de *L2*, es decir, una fórmula sin terminología no lógica. La tercera formulación del Principio de Kreisel es:

(PK3) $\forall x \subseteq \omega[(\forall z \in x \wedge L(z) \exists Y \exists Z \exists w \forall m \in xSATS(Y, w, Z, m)) \supset \exists y \exists z \exists w \forall m \in xsats(y, w, z, m)]$

Aunque la fórmula es muy compleja, simplemente afirma que si existe una clase que satisface a un conjunto de fórmulas de *L2*, existe un conjunto que satisface a ese mismo conjunto de fórmulas. El cuantificador universal principal, al cuantificar sobre subconjuntos de ω, cuantifica sobre conjuntos (posiblemente infinitos) de expresiones de *L2*, ya que nuestra numeración de Gödel codifica cada expresión de *L2* con un número. Esta cuantificación muestra por qué es necesario introducir el predicado $L(x)$, pues buscamos restringirnos a conjuntos de fórmulas. Al igual que en el caso de PK1, estas fórmulas son fórmulas sin variables libres, es decir, son oraciones, por lo cual (PK3) es una extensión de (PK1) y no de (PK2). El Principio de Reflexión correspondiente a (PK3) es:

(PR3) $\forall x \subseteq \omega[(\exists Z \exists w \forall m \in xTR(w, Z, m)) \supset \exists y((Z2/y) \wedge \exists z \exists w \forall m \in xsats(y, w, z, m))]$

Análogamente a los casos anteriores, definimos ZFC2 + PK3 y ZFC2 + PR3 como las teorías obtenidas de ZFC2+ agregando las fórmulas PK3 y PR3 respectivamente. Las generalizaciones de los teoremas de 1 y 3, se siguen como se esperaba:

Teorema 3.6. *ZFC2 + PK3 es equivalente a ZFC2 + PR3*

La prueba es una generalización de la prueba del teorema 3.2. Por supuesto, para probar que en ZFC2 + PK3 es verdad PR3 se consideran traducciones de todas las fórmulas de ZFC2 consideradas en el subconjunto deseado. Sin embargo, la idea de la prueba es exactamente la misma, así que la omitiré para simplificar.

De nuevo, para estudiar la satisfacción de PK3 necesitaremos una definición adicional para caracterizar los cardinales que satisfacen estas fórmulas.

Definición. *Sea α un ordinal, $\Phi(q, x)$ una fórmula de ZFC2 con sólo q y x libres y $t \subseteq \omega$. Decimos que Φ define α con el parámetro t si y solo si $\forall q(\Phi(q, t) \supset \alpha = q)$ es verdadera en el universo.*

Definición. *Sea α un ordinal. α es definible en segundo orden con un parámetro real (2R-definible) si y solo si hay una fórmula de ZFC2 $\Phi(q, x)$ con sólo q y x libres y $t \subseteq \omega$ tales que Φ define α con el parámetro t*

Teorema 3.7. *Sea α un ordinal y asumamos EP(α). Si α no es 2R-definible, entonces $V_{P(\alpha)} \models$ZFC2+PK3*

Demostración. Supongamos que α no es 2R-definible. Sabemos que $V_{P(\alpha)} \models ZFC2$. Por lo tanto, bastará con mostrar que $V_{P(\alpha)} \models PK3$ o equivalentemente, por el teorema anterior $V_{P(\alpha)} \models PR3$. Sea $t \subseteq \omega$ y supongamos que $V_{P(\alpha)} \models \exists Z \exists w \forall m \in xTR(w, Z, m)$. Como $P(\alpha)$, es inaccesible, $\exists z \exists w \forall m \in xsats(V_{P(\alpha)}, w, z, m)$ es verdadera supongamos que $V_{P(\alpha)}$ no satisface $\exists y((Z2/y) \wedge \exists z \exists w \forall m \in xsats(y, w, z, m)$. Esto es equivalente a que $\forall \beta < \alpha \forall z \forall w \exists m \in x \neg sats(V_{P(\beta)}, w, z, m)$ sea verdadera. Por lo tanto, la fórmula:

$$Ord(x) \wedge EP(x) \wedge \exists z \exists w \forall m \in qsats(V_{P(x)}, w, z, m) \wedge \forall \beta \in x \forall z \forall w \exists m \in q \neg sats(V_{P(\beta)}, w, z, m)$$

2R-define al ordinal α con parámetro t, que es una contradicción. $\square$

De esta manera, el mínimo modelo de ZFC2 + PK3 es el nivel indexado por el mínimo ordinal no *2R-definible.*

3.3.2.4. La cuarta formulación del Principio de Kreisel

La última formulación de (PK) que presentaré tiene relación con el anterior de la misma forma en que (PK1) y (PK2) estaban relacionados: esta última formulación considera fórmulas con variables libres, aunque no necesariamente. De nuevo, estaremos trabajando en la teoría ZFC2+. La formulación es:

(PK4) $\forall x \subseteq \omega[(\forall z \in x \wedge L(z)) \supset (\forall Y \forall w \forall Z(\forall m \in xSATS(Y, w, Z, m)) \supset \exists y \subseteq Y(s_1(w, y) \wedge \forall m \in xsats(y, w, Z \cap y, m)))]$

Esta fórmula afirma que si una clase satisface todas las fórmulas de un conjunto de fórmulas (posiblemente con variables libres) entonces hay un subconjunto de esa clase que satisface las fórmulas con las mismas funciones de asignación que la clase consideraba. Así pues, la teoría ZFC2 + PK4 es una extensión de todas las formulaciones anteriores. El Principio de Reflexión correspondiente es:

(PR4) $\forall x \subseteq \forall w \forall Z[\forall m \in xTR(w, Z, m) \supset \exists y((Z2/y) \wedge s_1(w, y) \wedge \forall m \in xsats(y, w, Z \cap y, m)]$

De nueva, cuenta es posible generalizar el teorema 3.4 para demostrar:

Teorema 3.8. *ZFC2 + PK4 es equivalente a ZFC2 + PR4*

La prueba es esencialmente la misma que la del teorema 3.4 y las mimas notas hechas para el caso del teorema 3.6, aplican aquí también.

Debido a que ZFC2 + PK4 es una extensión de ZFC2 + PK2 es claro que ZFC2 + PK4 implica que hay un modelo de ZFC2 + PK2, es decir, implica la existencia de un ordinal α que no es $\Pi^2_0 - describible$. Al igual que en caso de PK2, hay un principio de reflexión completa correspondiente a PK4 que es equivalente a éste

(PR4.1) $\forall Y \exists x((Z2/x) \wedge \forall w \forall n \in \omega(s_1(w, x) \supset (TR(w, Y, n) \leftrightarrow sats(x, w, x \cap Y, n)/x)))$

Teorema 3.9. *ZFC2 + PR4.1 es equivalente a ZFC2 + PK4*

De nuevo, introduciremos algunas definiciones para caracterizar los modelos de ZFC2 + PK4.

Definición. *Sea x un cardinal y Γ un conjunto de fórmulas de ZFC2, decimos que x es Π^2_0- conjunto describible Γ con el parámetro p si $p = s \in \Sigma$ y para toda $\Phi \in \Gamma$*

1) $V_\alpha \models \Phi(p\langle X\rangle_i)$

2) $\forall \beta < \alpha, V_\beta \not\models \Phi(p \cap V_\beta\langle X\rangle_i)$

Teorema 3.10. $V_{P(\alpha)} \models ZFC2 + PK4$ *si y solo si α no es Π^2_0-conjunto describible*

De nuevo, esto muestra que el mínimo ordinal α que no sea Π^2_0-*conjunto describible* es el mínimo modelo de ZFC2 + PK4. Aunque parece bastante informativo, aún no sabemos en qué medida pueden ser satisfechas todas estas teorías, debido a que desconocemos la naturaleza de los ordinales Π^2_0-*conjunto describible.*

3.4. Conclusiones

En este capítulo, presenté un análisis del principio y el argumento de Kreisel que nos permitió formular dicho principio en el marco de la teoría formal ZFC2. Argumenté que la noción (NPC) que si lo que defendí en el capítulo anterior es correcto, establece una relación entre una noción preteórica de validez y una noción de validez en estructuras cuyo dominio es una clase, y nos permite dar una versión del argumento de Kreisel adecuada y nos permite plantear formalmente el Principio de Kreisel. Así mismo, expuse varias formulaciones del Principio de Kreisel (presentadas originalmente por Shapiro) como fórmulas de ZFC2 y mostré que existe una relación entre esas fórmulas y diversos principios de reflexión para la teoría de conjuntos, presentados por Levy. Para cada una de estas formulaciones, se dieron caracterizaciones de los modelos mínimos de la teoría de conjuntos que satisficieran estas fórmulas.

De los análisis realizados en este capítulo podemos obtener por lo menos dos conclusiones importantes. La primera es que si lo que hice en las primeras dos secciones de este capítulo es correcto, entonces la noción de consecuencia lógica para la lógica de segundo orden efectivamente es muy poderosa. Esto debido a que todas las formulaciones del Principio de Kreisel son independientes de la teoría ZFC2. El hecho de que la lógica de segundo orden no cumpla con el teorema de compleción, ahora adquiere una relevancia mayor. A partir del análisis que hice del argumento de Kreisel se puede ver que el hecho de que la lógica de primer orden cumpla con el teorema de compleción permite garantizar que las fórmulas de primer orden pueden satisfacerse siempre en estructuras cuyo dominio sabemos que es un conjunto. Recordando la demostración de los teoremas de Lowenheim-Skolem del capítulo 1, esto se ve claramente, pues todo conjunto de fórmulas que tenga un modelo infinito tiene un modelo numerable. En la lógica de segundo orden, no tenemos esta garantía, pues de las demostraciones expuestas en la segunda parte de este capítulo podemos concluir que esta garantía es independiente de ZFC2. La segunda conclusión que podemos obtener de lo hasta ahora expuesto es que no hay hasta ahora una manera clara de garantizar los Principios de Reflexión. Si bien es cierto que expuse algunas condiciones que debe cumplir ZFC2 para garantizarlos, no es claro cómo cumplir esas condiciones. Es decir, por ejemplo, en el caso PK4, no es clara cuál es la naturaleza de un cardinal que no es Π^2_0-*conjunto describible*. Si no tenemos claro qué se requiere para garantizar la existencia de estos cardinales, no es claro que podamos garantizar su existencia. En el siguiente capítulo, presentaré un resultado reciente que permitirá precisar de manera más adecuada en qué condiciones es posible satisfacer los principios de reflexión.

Capítulo 4

El Principio de Reflexión

Teniendo a la mano las formulaciones de Shapiro del Principio de Reflexión (expuestas en la secciones 3.3.2) y que estas son equivalentes al Principio de Kreisel, en este capítulo me concentraré en argumentar a favor de que podemos aceptar estas formulaciones. El argumento dependerá de un teorema demostrado por Peter Koellner (2009) que muestra la consistencia de los principios de reflexión en segundo orden son consistentes relativamente a la existencia de un cardinal de Erdös. Habiendo expuesto esto, argumentaré que hay buenas razones para defender la existencia de un cardinal de Erdös. Sin embargo, mostraré también una limitante muy grande de la prueba: ésta no es suficiente para garantizar la validez de los principios PR3 y PR4, sino sólo de PR1 y PR2. Así mismo, daré algunas razones para considerar que el Principio de Kreisel está justificado en virtud de la justificación de los principios de reflexión. Finalmente, presentaré una posible estrategia para justificar el Principio de Reflexión mostrando la relación entre éste y el Principio de Reflexión para la aritmética. Así mismo, haré un análisis crítico de esta estrategia. Si lo que expondré a continuación es correcto, me parece que habré dado buenas razones para aceptar el Principio de Kreisel (en su formulación original), o al menos, habré expuesto bajo qué presupuestos descansa la verdad de este.

4.1. La prueba de Koellner

Koellner (2009) presentó una prueba de que los principios de reflexión (para el caso de la cuantificación universal en segundo orden) son consistentes. El interés principal de Koellner es estudiar las diferentes maneras de justificar axiomas nuevos que permitan resolver problemas de indecibilidad, tales como la hipótesis del continuo. Informalmente, el principio de reflexión afirma que si una oración es verdadera en el universo conjuntista, es verdadera en algún nivel de la jerarquía acumulativa de la teoría de conjuntos. La razón por la cual se considera que los principios de reflexión pueden justificar la inclusión de nuevos axiomas es porque formalizan la intuición de que el universo conjuntista no es definible mediante ninguna fórmula: "los principios de reflexión pretenden articular la idea informal de que el tamaño del universo es "absolutamente infinito" y por lo tanto no puede ser "caracterizado desde abajo"" (Koellner 2009: 208). Koellner argumenta en su trabajo que los principios de reflexión son insuficientes para resolver problemas importantes de indecibilidad, pues o bien son muy débiles para este propósito o bien son inconsistentes. Mi interés no es discutir la capacidad de los principios de reflexión para justificar axiomas, sino utilizar la prueba de consistencia de los principios de reflexión que presenta Koellner para argumentar que las fórmulas PR1-PR4 del capítulo anterior son consistentes y, por lo tanto, las formulaciones PK1-PK4 del Principio de Kreisel. En esta sección me ocuparé únicamente de presentar esta demostración, así que invito al lector no familiarizado con estos temas a revisar el Apendice al final de este trabajo y el trabajo original de Koellner.

La formulación usual de los principios de reflexión, en contraste con la presentada anteriormente, es:

(PR) $V \vDash \Phi(A) \rightarrow \exists \alpha V_\alpha \vDash \Phi^\alpha(A^\alpha)$

donde Φ^α es el resultado de relativizar los cuantificadores de Φ a V_α y A^α es el resultado de relativizar A a V_α. Cabe mencionar que estos principios son demostrables para el caso en que Φ es una fórmula de primer orden, por lo cual, la prueba que expondré tiene como fin evaluar el caso de las fórmulas de lenguajes más expresivos (y por lo cual, con mayor capacidad demostrativa). A partir de ahora, seguiré a Koellner en los detalles y la notación para facilitar la exposición de la prueba. El lenguaje en el que se presenta la prueba es el lenguaje de la teoría de conjuntos aumentado con variables de todos los órdenes finitos. Usaré las letras minúsculas x, y, z,... como variables de primer orden y las letras mayúsculas $X^{(m)}$, $Y^{(m)}$, $Z^{(m)}$,... como variables de orden m, donde $m > 2$. Relativo a un nivel V_α de la jerarquía acumulativa de la teoría de conjuntos, el rango de las variables de primer orden es el propio V_α y el rango de la variables de orden $m > 1$ es una copia isomorfa de $V_{\alpha+(m-1)}$. [63] Si $A^{(2)}$ es un parámetro de segundo orden sobre V_α (o la interpretación de una constante de predicado), entonces la relativización de $A^{(2)}$ a V_β, ($A^{(2),\beta}$), es $A \cap V_\beta$. La relativización se define formalmente de manera inductiva: para cada $m > 1$, $A^{(m+1),\beta} = \{B^{(m)} | B^{(m)} \in A^{(m+1)}\}$.

Definición. *Una fórmula en el lenguaje de órdenes finitos es positiva si y solo si se construye mediante operaciones de construcción de los símbolos* $\wedge$, $\vee$, $\forall$ *y* $\exists$ *a partir de átomos de la forma* $x = y$, $x \neq y$, $x \in y$, $x \notin y$, $x \in Y^{(2)}$, $x \notin Y^{(2)}$, $X^{(m)} = X'^{(m)}$ *y* $X^{(m)} \in Y^{(m+1)}$, *donde* $m \geq 2$

Las operaciones que mencioné arriba son de un tipo muy parecido a las operaciones de formación de fórmulas que expuse en el capítulo 1 (sección 1.1), donde las fórmulas se construían a partir de átomos y aplicaciones de operaciones para construir fórmulas condicionales, negaciones y cuantificaciones universales. La definición anterior permite por otro lado construir conjunciones, disyunciones y cuantificaciones universales y existenciales. Así mismo, vale la pena notar que las fórmulas positivas son aquellas que sólo tienen negaciones en oraciones que involucran una relación entre variables de primer orden y de segundo orden. Esto implica que todas las

[63]Koellner señala que la razón para usar una copia isomorfa de $V_{\alpha+(m-1)}$ y no $V_{\alpha+(m-1)}$ mismo es que se busca mantener la distinción entre conjunto y clase (colección de conjuntos) en caso de que un conjunto y una clase tengan la misma extensión (Koellner 2009: 208). De esta manera, en el caso de $m = 2$, usaremos $\alpha \times V_{\alpha+1}$ e interpretaremos la pertenencia a clases "$y \in \langle \alpha, x \rangle$" como $y \in x$, análogamente en los casos $m > 2$.

fórmulas del lenguaje ZFC2 (que es básicamente L2 más el símbolo de pertenencia) son equivalente a una fórmula positiva, pues no contiene variables de orden mayor a 2, y por lo tanto, no contienen negaciones sobre fórmulas atómicas con variables de orden mayor a 2.[64]

Definición. *Para* $0 < n < \omega$, $\Gamma_n^{(2)}$ *es la clase de fórmulas de la forma:*

$$\forall X_1^{(2)} \exists Y_1^{(k_1)} \dots \forall X_n^{(2)} \exists Y_n^{(k_n)} \Phi(X_1^{(2)} Y_1^{(k_1)}, \dots, X_n^{(2)} Y_n^{(k_n)}, A^{(l_1)}, \dots, A^{(l_{n'})})$$

Donde Φ *no tiene cuantificadores de orden superior y* $k_1, \dots, k_n$, $l_1, \dots, l_{n'}$ *son números naturales*

La idea de las fórmulas $\Gamma_n^{(2)}$ es que son fórmulas prenex de orden superior (de hecho, de cualquier orden), siempre que las cuantificaciones universales sean a lo más de segundo orden (aunque sí es posible que las fórmulas tengan cuantificaciones existenciales de orden mayor a 2). Hay que notar también que el número de cuantificaciones existenciales y el número de parámetros podría no ser el mismo.

Definición. *Para cada* $0 < n < \omega$, $\Gamma_n^{(2)}$***-reflexión*** *es el esquema que afirma que para cada* $\Phi \in \Gamma_n^{(2)}$, *si* $V \models \Phi$ *entonces existe un* $\delta \in Ord$ *tal que* $V_\delta \models \Phi^\delta$

La prueba de Koellner es sobre estos principios de reflexión exactamente.

Teorema 4.1. *Supongamos que existe el cardinal* $\kappa = \kappa(\omega)$. *Entonces hay un ordinal* $\delta < \kappa$ *tal que* V_δ *satisface* $\Gamma_n^{(2)}$*-reflexión, para cada* $n < \omega$

Demostración. La idea de la prueba es usar el nivel V_κ para obtener un modelo contable y una inmersión elemental sobre dicho modelo, y examinar el nivel indexado por el punto de crítico y probar que satisface $\Gamma_n^{(2)}$-reflexión.

Consideremos la estructura $N = \langle V_\kappa, \in, < \rangle$, donde $<$ es un buen orden sobre . Sea $I' = \iota'_k$ el conjunto de los indiscernibles de N obtenidos por el lema de Silver (véase Apendice). Sea $Hull^N(I')$, la *cascara* de Skollem de este conjunto de indiscernibles y sea

$$\pi : M \to Hull^N(I')$$

la inversa del colapso transitivo de $Hull^N(I')$ sobre M. Hay que notar que M es numerable dado que $Hull^N(I')$ es numerable. Sea I la imagen de I' bajo el colapso. Hay que notar que todos los elementos de I cumplen con ser indiscernibles de M, de otra manera, habría forma de encontrar una fórmula φ que diferencia a algunos de los elementos de I, y así sería posible discernir entre algunos de los elementos de

[64]Vale la pena notar que no es posible definir en *L2* (ni en el lenguaje de ZFC2) las fórmulas del tipo $X = Y$ ni del tipo $X \neq Y$ (donde X y Y son constantes o variables de predicado), aunque se pueden *definir* como $\forall x(X(x) \equiv Y(x))$ y $\exists x(\neg X(x) \equiv Y(x))$. Koellner (2009: 210) señala la importancia de que las fórmulas consideradas no pueden tener átomos de la forma $X^{(m)} \neq Y^{(m)}$ ni de la forma $X^{(m)} \notin Y^{(m+1)}$. Hay que notar que el lenguaje de ZFC2 como se está considerando no permite entender a las variables de segundo orden como objetos (sólo las variables de primer orden pueden entenderse en este sentido), y esto es fundamental para poder desarrollar el tipo de contraejemplos que Koellner presenta, por lo cual el poder definir $X \neq Y$ como mencioné más arriba no representa mayor problema.

I'. Si incluimos la relación de buen orden $<$ entre ellos, estos objetos cumplen con todas las propiedades de los indiscernibles de Silver (con respecto a L).[65]

Sea $\rho : I \to I$, una función que preserva el orden sobre I tal que mueva el primer indiscernible ι_0. Se extiende esta función a una única inmersión elementa $j : M \to M$ tal que el punto crítico de j es el primer indiscernible, $crit(j) = \iota_0$. Demostraremos que $V^M_{\iota_0}$ satisface $\Gamma^{(2)}_n$-reflexión.

Lema 4.1. *Supongamos que* $\Phi(A_1, \ldots, A_m) \in \Gamma^{(2)}_n$. *Si* $V^M_{\iota_0} \vDash \Phi(A_1, \ldots, A_m)$ *entonces* $V^M_{\iota_0} \vDash \Phi(j(A_1)^{\iota_0}, \ldots, j(A_m)^{\iota_0})$

Demostración. La demostración es por inducción sobre n.

Paso Base: Será conveniente separar las variables de segundo orden de las variables de orden superior. Así $\Phi(A_1, \ldots, A_m)$ es:

$$\Phi(A^{(2)}_1, \ldots, A^{(2)}_j, B^{(n_{j+1})}_{j+1}, \ldots, B^{(n_m)}_m)$$

Supongamos que $V^M_{\iota_0} \vDash \Phi(A^{(2)}_1, \ldots, A^{(2)}_j, B^{(n_{j+1})}_{j+1}, \ldots, B^{(n_m)}_m)$. Como ι_0 es inaccesible en M, existe un conjunto cerrado y no acotado (un club):

$$C = \{\alpha < \iota_0 | \langle V^M_\alpha, \in, A^{(2),\alpha}_1, \ldots, A^{(2),\alpha}_j\rangle \prec \langle V^M_{\iota_0}, \in, A^{(2),\ldots,A^{(2)}_j}_1\rangle\}$$

Supongamos que $A^{(2)} \in B^{(3)}$ o $A^{(m)} \in B^{(m+1)}$ son constituyentes de Φ. Si son falsos entonces dichas partículas no influyen en el hecho de que $V^M_{\iota_0} \vDash \Phi(A^{(2)}_1, \ldots, A^{(2)}_j, B^{(n_{j+1})}_{j+1}, \ldots, B^{(n_m)}_m)$, pues están enunciadas afirmativamente. Si son verdaderos, entonces $V^M_\alpha \vDash \Phi(A^{(2)}_1, \ldots, A^{(2)}_j, B^{(n_{j+1})}_{j+1}, \ldots, B^{(n_m)}_m)$, pues V^M_α es un submodelo elemental de $V^M_{\iota_0}$. Por lo tanto, para todo $\alpha \in C$

$$V^M_\alpha \vDash \Phi(A^{(2)}_1, \ldots, A^{(2)}_j, B^{(n_{j+1})}_{j+1}, \ldots, B^{(n_m)}_m)$$

Ahora esto es equivalente a una afirmación en primer orden sobre los parámetros ι_0, C, $A^{(2)}_1, \ldots, A^{(2)}_j, B^{(n_{j+1})}_{j+1}, \ldots, B^{(n_m)}_m$. Como j es una inmersión elemental, dicho enunciado con los parámetros $j(\iota_0)$, $j(C)$, $j(A^{(2)}_1)$,...,$j(A^{(2)}_j)$,$j(B^{(n_{j+1})}_{j+1})$,...,$j(B^{(n_m)}_m)$ es verdadero en M. Por lo cual, $V^M_\alpha \vDash \Phi(j(A^{(2)}_1), \ldots, j(A^{(2)}_j), j(B^{(n_{j+1})}_{j+1}), \ldots, j(B^{(n_m)}_m))$, para todo $\alpha \in j(C)$.

Ahora como C es un club en ι_0, $j(C)$ es un club en $j(\iota_0)$. De hecho, $j(C)$ contiene todos los $\alpha < \iota_0$ (pues ι_0 es el punto crítico). Por lo tanto $j(C) \cap C = C$, como no está acotado en ι_0 y $j(C)$ es un club, y así $\iota_0 \in j(C)$. Por lo tanto,

$$V^M_{\iota_0} \vDash \Phi(j(A^{(2)}_1), \ldots, j(A^{(2)}_j), j(B^{(n_{j+1})}_{j+1}), \ldots, j(B^{(n_m)}_m))$$

que era lo que buscábamos mostrar

[65] Esto implica que es posible construir un modelo de L con estos indiscernibles.

Paso Inductivo: Supongamos que el lema se cumple para una fórmula cualquiera $\Phi \in \Gamma_n^{(2)}$. Demostraremos que se cumple para $\forall X^{(2)} \exists Y^{(k)} \Phi(X^{(2)}, Y^{(k)}, \vec{A})$ donde $\vec{A}$ es $A^{(l_1)}, \ldots, A^{(l_{n'})}$

Supongamos que $V_{\iota_0}^M \models \forall X^{(2)} \exists Y^{(k)} \Phi(X^{(2)}, Y^{(k)}, \vec{A})$

Sabemos que, $V_{\iota_0}^M \models \forall X^{(2)} \exists Y^{(k)} \Phi(X^{(2)}, Y^{(k)}, \vec{A}) \leftrightarrow \forall B \subseteq V_{\iota_0}^M [V_{\iota_0}^M \models \Phi(B^{(2)}, f(B)^{(k)}, \vec{A})]$ por la definición de $\forall$ y por una función de Skolem definida para cada $B \subseteq V_{\iota_0}^M$, como $f : \wp(V_{\iota_0}^M) \to \wp^k(V_{\iota_0}^M)$, $B \longmapsto f(B)^{(k)}$

Entonces, $\forall B \subseteq V_{\iota_0}^M [V_{\iota_0}^M \models \Phi(j(B)^{(2),\iota_0}, j(f(B))^{(k),\iota_0}, j(\vec{A})^{\iota_0})]$ por el paso base y entonces $\forall B \subseteq V_{\iota_0}^M [V_{\iota_0}^M \models \Phi(B^{(2)}, f'(B)^{(k)}, j(\vec{A})^{\iota_0})]$ por la función de Skolem obtenida de la siguiente manera:

A partir de la función de Skolem definida anteriormente se puede saber que para cada $B \subseteq V_{\iota_0}^M, j(B)^{(2),\iota_0} \longmapsto j(f(B))^{(k),\iota_0}$. Como sabemos que para cada $B \subseteq V_{\iota_0}^M, j(B)^{(2),\iota_0} = B$ pues ι_0 es el punto crítico, tenemos que existe la función de Skolem $f' : \wp(V_{\iota_0}^M) \to \wp^k(V_{\iota_0}^M)$,

$$B \longmapsto j(f(B))^{(k),\iota_0}$$

Es por esto que es posible tener cuantificaciones existenciales sobre variables de orden superior a 2.

Por último, $\forall B \subseteq V_{\iota_0}^M [V_{\iota_0}^M \models \Phi(B^{(2)}, f'(B)^{(k)}, j(\vec{A})^{(\iota_0)}] \leftrightarrow V_{\iota_0}^M \models \forall X^{(2)} \exists Y^{(k)} \Phi(X^{(2)}, Y^{(k)}, j(\vec{A}))$ que se sigue por definición de los cuantificadores.

Y esto era lo que queríamos probar.

□

Lema 4.2. $V_{\iota_0}^M \models \Gamma_n^{(2)}$*-reflexión*

Demostración. Supongamos que $V_{\iota_0}^M \models \Phi(A_1, \ldots, A_m)$. Por el lema 4.1, $V_{\iota_0}^M \models \Phi(j(A_1)^{\iota_0}, \ldots, j(A_m)^{\iota_0})$ entonces, como j movió a ι_0 y este es el punto crítico, sabemos que $j(\iota_0) > \iota_0$ y por lo tanto

$$V_{j(\iota_o}^M \models \exists \alpha < j(\iota_0) \Phi^\alpha(j(A_1)^\alpha, \ldots, j(A_m)^\alpha)$$

Por último, como j es elemental se sigue que

$$V_{\iota_0}^M \models \exists \alpha < \iota_0 \Phi^\alpha((A_1)^\alpha, \ldots, (A_m)^\alpha)$$

Que era lo que queríamos mostrar

□

Por último, sea $\delta = \pi(\iota_0)$. Aplicando π a M, tenemos que $V_\delta \models \Gamma_n^{(2)}$- reflexión, para todo $n < \omega$, donde $\delta < \kappa$, completando la prueba.

□

Esta es la prueba proporcionada por Koellner, donde se presenta un modelo de cada $\Gamma_n^{(2)}$- reflexión.

4.2. ¿Podemos aceptar el Principio de Kreisel?

La prueba anterior muestra que ciertas formulaciones del Principio de Reflexión son consistentes, en tanto que existe un modelo en dónde todas son el caso. El objetivo de esta sección es evaluar dicho resultado con respecto a las fórmulas expuestas en el capítulo anterior. En primer lugar, hay que señalar que las fórmulas consideradas en los principios PK1–PK4 y en PR1-PR4 son parte de las fórmulas consideradas por Koellner. Recordemos que las oraciones que estamos considerando en estas fórmulas son las fórmulas de ZFC2, que incluyen todas las fórmulas de *L2*. Al considerar las fórmulas de ZFC2, nos estamos concentrando únicamente sólo con fórmulas de segundo orden. El lenguaje en el que está expuesto el teorema 4.1 es un lenguaje mucho más poderoso pues considera cuantificaciones de variables de orden mayor a 2. De hecho, toda las fórmulas que se consideran pertenecen a algún conjunto $\Gamma_n^{(2)}$ pues siempre es posible que una fórmula φ tenga variables libres y la cuantificación existencial considerada en la fórmula $\forall X^{(2)} \exists Y^{(k_n)} \varphi$ es a lo más de segundo orden, es decir, $k_n \leq 2$. Así pues, cada fórmula que se considera en los principios PR1-PR4 pertenece a algún conjunto $\Gamma_n^{(2)}$, dependiendo del número de bloques de cuantificadores que contenga la fórmula.

Esto es suficiente para afirmar que los principios PK1 y PK2 son consistentes en tanto que son equivalentes a PR1 y PR2, y estos son instancias de los principios de reflexión considerados por Koellner. Recordemos PR1 y PR2:

(PR1) $\Psi \to \exists x((Z2/x) \wedge (\Psi/x))$

(PR2) $\forall X \forall y(\Psi(Y, y) \supset \exists x(y \in x \wedge ((Z2/x) \wedge \Psi(x \cap Y, y)/x)))$

Hay que recordar que la diferencia entre PR1 y PR2 es que PR1 sólo considera las fórmulas Ψ que de hecho son enunciados (o fórmulas cerradas), mientras PR2 considera además fórmulas con variables libres. De cualquier forma, ambas están enunciadas en el lenguaje de ZFC2 (cuyo lenguaje es menos poderoso que el que Koellner considera[66]). Con esto en mente, se puede ver que una formulación equivalente de PR1 sería el esquema:

(PR1*) $V \models \Psi \to \exists \alpha V_\alpha \models Z2^\alpha \wedge \Psi^\alpha$ [donde Ψ es un enunciado de ZFC2]

El cambio fue meramente de notación, y de hecho, sabemos que el ordinal α es inaccesible. Por supuesto, este esquema no es ninguno de los esquemas asociados a las diferentes clases $\Gamma_n^{(2)}$, pero PR1* sigue siendo consistente en tanto que todas las instancias de este esquema son instancias de algún esquema asociado a alguna $\Gamma_n^{(2)}$. Observaciones similares pueden hacerse sobre PR2 y el esquema:

(PR2*) $V \models \Psi(Z, z) \to \exists \alpha V_\alpha \models Z2^\alpha \wedge \Psi^\alpha(Z, z)$ [donde $\Psi(Z, z)$ es una fórmula con a lo más Z y z variables libres]

[66]Hay que recordar que el lenguaje de ZFC2 es simplemente el lenguaje *L2* (sin terminología no lógica), más el símbolo de pertenencia. Es decir, no contiene variables de orden mayor a segundo orden.

La única observación pertinente es que las variables de segundo orden son interpretadas por Shapiro y Koellner de la misma manera: como clases entendidas como colecciones de conjuntos.[67] Sin embargo, tanto PR2 como PR2* son esquemas en tanto que no hay cuantificadores sobre las fórmulas que consideran. Sin embargo, existen problemas en los casos de los principios PR3 y PR4.

4.2.1. El problema de los principios PR3 y PR4

Las observaciones hechas en los casos anteriores no se pueden replicar para los casos de PR3 y PR4, ya que estos principios no son esquemas. Esto es un problema más grande de lo que parece pues, como explicaré, implica que garantizar la consistencia de los principios de PR3 y PR4 requiere de una prueba más fuerte. Recordemos los principios PR3 y PR4:

(PR3) $\forall x \subseteq \omega[(\exists Z \exists w \forall m \in xTR(w, Z, m)) \supset \exists y((Z2/y) \wedge \exists z \exists w \forall m \in xsats(y, w, z, m))]$

(PR4) $\forall x \subseteq \forall w \forall Z[\forall m \in xTR(w, Z, m) \supset \exists y((Z2/y) \wedge s_1(w, y) \wedge \forall m \in xsats(y, w, Z \cap y, m)]$

De nuevo, la diferencia entre ambos principios es que PR3 considera sólo fórmulas cerradas mientras que PR4 considera también fórmulas con variables libres. Existen al menos dos problemas con estos principios. El primero es que éstos consideran conjuntos de fórmulas tal que sus elementos pertenezcan a diferentes clases $\Gamma_n^{(2)}$. Por ejemplo, consideremos dos fórmulas $\forall X \exists y \varphi$ y $\forall X \exists Y \forall Z \exists w \psi$ de *L2*, tal que φ y ψ tienen sólo cuantificadores de primer orden y no tienen parámetros. El conjunto de fórmulas $A = \forall X \exists y \varphi, \forall X \exists Y \forall Z \exists w \psi$ es considerada por PR3, pues para ello sólo es necesario considerar el subconjunto de ω constituido por los números de Gödel de estas fórmulas. Sin embargo, $\forall X \exists y \varphi \in \Gamma_1^{(2)}$ mientras que $\forall X \exists Y \forall Z \exists w \psi \in \Gamma_2^{(2)}$, y por lo tanto, dependen de diferentes esquemas. Esto implica que la satisfacción del Principio de Reflexión correspondiente a estas clases no garantiza que sea posible satisfacer ambas fórmulas en un mismo modelo. El segundo problema tiene esta misma conclusión: los principios asociados a cada $\Gamma_n^{(2)}$ son esquemas, mientras que PR3 y PR4 son fórmulas que permiten cuantificar sobre las fórmulas (codificadas mediante una numeración de Gödel), es decir, PR3 y PR4 no son esquemas. El teorema 4.1 garantiza que todos los esquemas de las distintas clases $\Gamma_n^{(2)}$ se pueden satisfacer en un V_δ, es decir, que para cada fórmula φ, hay un V_α que satisface φ, sin embargo, no hay garantía que un conjunto de fórmulas sea satisfecho en un V_α particular. Estas observaciones implican que el teorema 4.1 no sólo no es suficiente, sino que no hay ninguna manera de garantizar la consistencia de PR3 y PR4 a partir de la prueba. El contenido de PR3 y PR4 requiere mucho más que lo requerido que la prueba de Koellner.

---

[67] Más precisamente, ambos asignan a las variables de segundo orden colecciones de conjuntos. La única diferencia es que Koellner mantiene la distinción entre las colecciones de conjuntos que son conjuntos y aquellas que no lo son. Sin embargo, las asignaciones posibles son exactamente las mismas.

Esto es un problema aún más grande si consideramos la razón por la cual se introdujeron los principios PR3 y PR4. Como señalé en la sección 3.1.1 del capítulo 3, no es posible reducir la noción de consecuencia lógica en segundo orden a la noción de verdad lógica (contrario al caso de la lógica de primer orden). Por ello era necesario considerar una nueva versión diferente del Principio de Kreisel para el caso de la lógica de segundo orden:

(PKC2) $\forall\Gamma\forall\alpha(\mathrm{Val}(\Gamma^2, \alpha^2) \leftrightarrow \mathrm{V}(\Gamma^2, \alpha^2))$

Para poder encontrar una manera de expresar este contenido, era necesario poder caracterizar la satisfacción de conjuntos de fórmulas de segundo orden, para lo cual se introdujeron los principios PK3 y PK4. El hecho de que PR3 y PR4 no se pueden caracterizar en términos de los principios de reflexión trabajados por Koellner, implica que no es posible garantizar la consistencia de los principios PK3 y PK4 mediante esta prueba. Y por lo cual no es posible garantizar la consistencia del principio PKC2 a partir de la prueba del teorema 4.1.

Estas observaciones muestran que no tenemos una prueba de la consistencia del Principio de Kreisel que nos permita justificar éste para el caso general de la noción de consecuencia lógica en general. Como expliqué en la observación hecha en la sección 3.1.1 del capítulo anterior, el problema radica en que la lógica de segundo orden no satisface el teorema de compacidad, es decir, la dificultad es que hay consecuencia lógicas de un conjunto infinito Γ que no son consecuencias lógicas de ningún subconjunto finito de Γ. Aún es posible tener la consistencia antes mencionada para el caso donde una fórmula es consecuencia lógica de un conjunto infinito y también lo es algún subconjunto finito de este. El mismo caso se presenta cuando el conjunto considerado es finito. Sin embargo, estos casos son casos poco interesantes (y poco útiles) dentro del contexto de la lógica de segundo orden. De momento me concentraré en las formulaciones esquemáticas hasta ahora presentadas. Sin embargo, habrá que decir que una prueba similar a la de Koellner requeriría de mucho más de lo que la prueba de Koellner misma. Un argumento a favor de los principios PR3 y PR4 implicaría mucho mayor compromiso que el teorema 4.1 mismo (que expondré más adelante). Para ver esto, consideremos las siguientes definiciones. Sea LK un lenguaje con su correspondiente conjunto K de terminología no lógica.

Definición. *El número de Lowenheim de LK es el mínimo cardinal κ tal que si una fórmula Φ de LK es satisfascible, entonces Φ tiene un modelo de cardinalidad a lo más κ*

Definición. *El número de Lowenheim - conjunto de LK es el mínimo cardinal κ tal que si un conjunto Γ de fórmulas de LK es satisfascible, entonces hay un modelo $\mathfrak{A}$ (y una $s \in \Sigma^{\mathfrak{A}}$) cuya cardinalidad es a lo más κ y que satisface a todas las fórmulas Γ*

El *número de Lowenheim* (*Lowenheim–conjunto*) es el cardinal más grande en que se puede satisfacer una fórmula (conjunto de fórmulas). El problema es que no hay mucha información sobre cuál es el *número de Lowenheim–conjunto* de *L2*, el lenguaje que consideramos. Lo que sí es posible afirmar es que estos números son menores al primer cardinal supercompacto:

Definición. *Decimos que una propiedad de conjuntos $P(x)$ es una propiedad local si hay una fórmula $\psi(x)$ del lenguaje de la teoría de conjuntos en primer orden tal que para cada x, $P(x)$ es verdad si y solo si $\exists\delta(V_\delta \vDash \psi(x))$*

Definición. *Un cardinal λ es local – minimal si hay una propiedad local P de cardinales tal que λ es el mínimo cardinal que tiene la propiedad P*

Es decir, λ es el mínimo cardinal tal que $\exists\delta(V_\delta \vDash \psi(\lambda))$, donde $\psi(x)$ es la fórmula correspondiente a P.

Teorema 4.2. *El número de Lowenheim de L2 es la mínima cota superior del conjunto de todos los cardinales λ, tal que λ es local – minimal*

Shapiro (1991: 156-57) señala (a partir de las observaciones de Solovay) que si un cardinal κ es supercompacto, P es una propiedad local y algún cardinal tiene la propiedad P, entonces hay un cardinal $\lambda < \kappa$, tal que λ tiene la propiedad P. De esto se sigue que el número de Lowenheim de *L2* es menor que el primer cardinal supercompacto, al igual que el número de Lowenheim – conjunto de *L2*. Una prueba de PR3 y PR4 tendría como consecuencia tener una mayor precisión de cuál es el número de Lowenheim de *L2*. Así mismo, un posible argumento a favor de ellos podría consistir en defender los axiomas de cardinales supercompactos para la teoría de conjuntos. No haré esto de momento, me concentraré en los principios considerados por Koellner.

4.2.2. El Principio de Reflexión y el concepto de conjunto

En esta sección, discutiré brevemente la relevancia del teorema expuesto en la sección 4.1 con relación a la posibilidad de aceptar el Principio de Kreisel. Cómo mencioné antes, la principal preocupación de Koellner en su investigación es la justificación de nuevo axiomas. Más precisamente, Koellner investiga si los principios de reflexión que impliquen una reducción significativa en el fenómeno de la incompletud pueden ser justificados por el concepto de conjunto. El concepto de conjunto en el sentido investigado por Koellner es el concepto iterativo de conjunto, es decir, el concepto de conjunto detrás de la idea de que un conjunto es obtenido de objetos bien definidos (obtenidos anteriormente) mediante la operación “conjunto de”. Maddy (1988), realiza una presentación histórica y filosófica de este concepto de conjunto, capturado mediante una variedad de justificaciones conocidas como j*ustificaciones intrínsecas al concepto de conjunto*. Una de estas justificaciones es el Principio de Reflexión, cuya justificación recae en la idea informal de que el universo conjuntista es “absolutamente indefinible”, esto es, en la idea informal de que el universo conjuntista no puede ser definido por ninguna fórmula. Así, el Principio de Reflexión está justificado a partir del concepto de conjunto en tanto que articula esta idea informal inherente al concepto de conjunto. La pregunta es pues si esta idea es suficiente para justificar el Principio de Reflexión de esta manera.

Esto es el caso para las fórmulas de primer orden. En ZFC, es posible demostrar que el Principio de Reflexión es verdadero donde la fórmula considerada es de primer orden. El problema es que no hay una prueba de estos para el caso donde la fórmula es de segundo orden (o de cualquier orden superior). Koellner señala que hay al menos una dificultad filosófica importante para poder responder afirmativamente a esta pregunta: cómo entender la cuantificación en orden superior. La dificultad consiste en que si consideramos a V como una auténtica totalidad, es posible formular la idea de que este no es definible mediante una fórmula y formular principios de reflexión (posición *actualista*). El problema con esto es que no es claro cómo entender la cuantificación en orden superior dado que las variables de orden superior deberían correr sobre la "potencia" de V, que no es claro que sea siquiera concebible. Una opción es considerar a V no como una totalidad sino como una colección acotada a un V_α (posición *potencialista*), lo que permite entender la cuantificación de orden superior. Sin embargo, motivar el Principio de Reflexión se dificulta, en vista de que no es obvio como justificar que V_α sea V, de alguna manera. De esta dificultad, Koellner concluye que no hay una fuerte justificación para el Principio de Reflexión.

Creo que esta es conclusión es correcta, sin embargo es poco probable obtener una posible respuesta a esta cuestión y la manera en que se hace uso del Principio de Reflexión parece mostrar poca preocupación por esta cuestión. Por ejemplo, el uso del principio en la justificación de la existencia de cardinales inaccesibles: sabemos que V es cerrado bajo remplazo y potencia, por lo tanto, debe existir un nivel de la jerarquía que sea cerrado bajo remplazo y potencia, es decir, un cardinal fuertemente inaccesible. Parece que en este ejemplo se hace uso de una posición actualista y es posible entender la cuantificación de orden superior de manera informal (incluso si se tiene en mente la versión de remplazo de ZFC2). De nuevo, el Principio de Reflexión sigue sin tener una justificación fuerte, sin embargo parece que es suficiente en principio con una justificación más débil e informal, a saber, la justificación informal mencionada anteriormente. En este punto, la prueba de consistencia de Koellner provee de mayor fuerza justificatoria a esta idea informal. Finalmente, el propio Koellner considera que hay dos propiedades que tiene la noción de justificación intrínseca (o interna): el enunciado en cuestión no tiene porqué ser autoevidente y la justificación de éste debe ser más fuerte que la mera plausibilidad (cf. Koellner 2009: 207). Efectivamente el Principio de Reflexión no es evidente y las razones para sostenerlo implican más que mera plausibilidad, incluyendo el hecho de que es verdadero para el caso de primer orden y ahora la prueba de consistencia Koellner. La prueba de consistencia efectivamente es más evidencia a favor de la justificación del principio. En este sentido, la equivalencia entre el Principio de Reflexión y el Principio de Kreisel permite ver en qué sentido el Principio de Kreisel tiene cierta justificación. Esta justificación consiste en la propia justificación del Principio de Reflexión, es decir, el Principio de Kreisel tiene justificación intrínseca al concepto de conjunto y es consistente con este concepto (en vista de la prueba del teorema de Koellner), al menos para el caso de PK1 y PK2. A partir de estas obsevaciones, la justificación del Principio de Reflexión nos ofrece buenas razones para aceptar el Principio de Kreisel, en tanto que éste está igualmente justificado a partir del concepto de conjunto.

4.2.3. El cardinal de Erdös

La prueba que presenta Koellner depende estrictamente de la existencia de un cardinal de Erdös, el cardinal de Erdös, $\kappa(\omega)$. La prueba depende de manera importante en esto pues a partir de la existencia del cardinal $\kappa(\omega)$ es posible encontrar el conjunto de indiscernibles adecuado. Sin embargo, la existencia de un cardinal de Erdös es independiente de ZFC, por lo cual, no tenemos certeza de que esta prueba nos garantice la consistencia (en un sentido fuerte) de los principios de reflexión.

Sin embargo, existen varias razones que podemos dar para aceptar la existencia de un cardinal de Erdös. Penelope Maddy (1988) presenta una serie de criterios para defender la inclusión de nuevos axiomas a la axiomatización usual de la teoría de conjuntos. De hecho, Maddy explica también cómo estos criterios permiten justificar los axiomas estándar de ZFC. Estos criterios pueden dividirse en criterios internos y criterios externos. Los criterios internos son justificaciones de los axiomas que recaen directamente en el concepto de conjunto. La idea es que un axioma que cumple con un criterio interno está justificado, en tanto que se "sigue" del concepto de conjunto. Por otro lado, los criterios externos justifican la inclusión de un axioma por sus consecuencias, sus relaciones con otras áreas de la matemática o su poder explicativo. Maddy pone como ejemplo de estos últimos la indispensabilidad del axioma de elección para demostrar principios importantes e intuitivos como el teorema del buen orden o el lema de Zorn (cf. Maddy 1988: 488). Como mencioné, Koellner tenía como principal interés el papel de los principios de reflexión como un principio que permite justificar nuevos axiomas. En efecto, se considera que cumplir con el Principio de Reflexión es un criterio interno muy importante para justificar nuevos axiomas. Hay razones (independientes del Principio de Reflexión) para justificar la existencia de los cardinales de Erdös, en particular $\kappa(\omega)$ que recaen en algunos criterios internos expuestos por Maddy.

La definición de un cardinal de Erdös es

Definición. *Un cardinal $\kappa(\alpha)$ es un cardinal de Erdös si y solo si es el mínimo cardinal tal que $\kappa \to (\alpha)^{<\omega}$*

Informalmente, un cardinal de Erdös, $\kappa(\alpha)$, es el mínimo cardinal κ tal que todas las particiones (en 2 "pedazos") del conjunto de todos los subconjuntos de cardinalidad finita de κ tiene un conjunto homogéneo cuyo tipo de orden es α. Consideremos el siguiente cardinal:

Definición. *Un cardinal κ es débilmente compacto si y solo si es no contable y $\kappa \to (\kappa)^2$*

La introducción de este cardinal se da a partir del teorema de Ramsey:

Teorema 4.3. ***Teorema de Ramsey:*** $\omega \to (\omega)^n_m$

El teorema de Ramsey afirma la existencia de un conjunto homogéneo cuyo tipo de orden es ω para cada partición en m pedazos del conjunto de todos los subconjuntos de ω de cardinalidad n. Mientras que un axioma que afirme la existencia de un cardinal débilmente compacto afirma la existencia un conjunto homogéneo cuyo cardinalidad es κ para cada partición en 2 pedazos del conjunto de todos los subconjuntos de κ de cardinalidad 2. La definición de un cardinal débilmente compacto es una generalización del teorema de Ramsey, en tanto que el teorema de Ramsey es un caso particular de la definición de un cardinal débilmente compacto (el caso en que κ es $\aleph_0$, y tanto n como m son 2).

Es posible considerar la definición de un cardinal de Erdös como una generalización del teorema de Ramsey. Esto no es del todo preciso, pues no es una generalización en sentido estricto; los cardinales débilmente compactos no son cardinales de Erdös. Pero los cardinales débilmente compactos consideran que el correspondiente conjunto homogéneo debe tener cierta cardinalidad (a saber, κ mismo), mientras que un cardinal de Erdös sólo necesita que dicho conjunto tenga un tipo de orden (a saber, α); un cardinal débilmente compacto considera el caso en que el tipo de orden correspondiente es κ mismo. Además, el cardinal de Erdös considera las particiones del conjunto de todos los subconjuntos finitos, que incluyen aquellos cuya cardinalidad es 2. De hecho, existe un teorema que expresa esto:

Teorema 4.4. *Si existe $\kappa(\omega)$ cardinal de Erdös, entonces existe un cardinal débilmente compacto menor a $\kappa(\omega)$ (Jech 2002: 304)*

Además

Teorema 4.5. *Todo cardinal de Erdös $\kappa(\alpha)$ es inaccesible, y si $\alpha < \beta$ entonces $\kappa(\alpha) < \kappa(\beta)$*

En vista de esto, suponiendo la existencia de un cardinal de Erdös, en particular el cardinal $\kappa(\omega)$, se puede demostrar la existencia de cardinales débilmente compactos. Sin embargo, esto es natural pues como vimos un cardinal de Erdös tiene propiedades que abarcan los casos de los cardinales débilmente compactos. Por otro lado, los cardinales de Ramsey son casos particulares de los cardinales de Erdös:

Definición. *Un cardinal κ es un cardinal de Ramsey si y solo si cumple con $\kappa \to (\kappa)^{<\omega}$*

La definición de los cardinales de Ramsey es un caso particular de la definición de los cardinales de Erdös (donde $\kappa = \kappa(\kappa)$ o bien donde $\alpha = \kappa$). Sin embargo, la existencia del cardinal $\kappa(\omega)$ no implica la existencia de un cardinal de Ramsey, pues la existencia de los cardinales de Ramsey implica que $V \neq L$, mientras que la existencia de $\kappa(\omega)$ es consistente con $V = L$. Para que la existencia de algún cardinal $\kappa(\alpha)$ de Erdös implique la existencia de un cardinal de Ramsey, α debe ser un cardinal al menos tan grande como κ.

Ahora daré algunas razones para aceptar el cardinal de Erdös $\kappa(\omega)$ en las siguientes subsecciones. Para mayor generalidad, consideraré un axioma de cardinal de Erdös enunciado de la siguiente forma:

(CE) Para cada α, existe un cardinal $\kappa(\alpha)$ de Erdös

Esto implica un compromiso muy grande en vista de lo mencionado anteriormente con respecto a los cardinales de Ramsey, sin embargo, nos permite mayor generalidad.

4.2.3.1. Maximalidad y exhaustividad

Uno de los criterios de Maddy es el requisito de *maximalidad*. Este criterio consiste en que parte del concepto de conjunto (como parte de la jerarquía acumulativa) es que se debe maximizar el contenido de cada nivel de la jerarquía acumulativa. En este sentido, este criterio permite maximizar el contenido del conjunto potencia de cada conjunto agregando todos los subconjuntos posibles de éste[68]; el conjunto potencia de un conjunto debe ser mucho más "ancho". Además, permite garantizar que la colección de los ordinales será "absolutamente infinita". Un axioma que cumpla este criterio está justificado por el concepto de conjunto en tanto que *ensancha* el conjunto potencia y *alarga* la clase de los ordinales.

El axioma que se defiende está de acuerdo con esto en vista de que garantiza tanto que la operación potencia se efectúa de manera irrestricta y de que asegura que la colección es absolutamente infinita, pues todo cardinal de Erdös es inaccesible. En primer lugar, un cardinal inaccesible es cerrado bajo la operación potencia, aplicada completamente. Es decir, hay una garantía de que todos los niveles de la jerarquía menores a él son tales que todos los inmediatos siguientes son más "anchos" en tanto que tienen muchos más elementos que él; de hecho, tanto como es posible que tenga. Sin embargo, Maddy da una razón (generalmente planteada) mucho más fuerte en lo que respecta a la segunda condición:

> Depende de la postura comúnmente aceptada de que el universo conjuntista es demasiado complejo para ser agotado por cualquier puñado de operaciones, en particular por el conjunto potencia y remplazo, ambos dados por los axiomas de Zermelo y Fraenkel. Así, debe haber un número ordinal después de todos los ordinales generados por remplazo y conjunto potencia. Esto es un inaccesible (Maddy 1988: 502)

Un cardinal inaccesible permite incrementar la "cantidad" de ordinales. Este va de la mano con lo que se conoce como *inexhaustividad* que consiste en la afirmación de que el universo no puede ser agotado por un "puñado" de operaciones. Y dado que es posible aplicar tanto potencia como reemplazo al primer inaccesible, por *inexhaustividad*, debe existir otro inaccesible, y así sucesivamente. El CE garantiza todo esto al implicar la existencia de cardinales inaccesibles. Además, se garantiza que el universo no se agota por la existencia de conjuntos homogéneos bajo ciertas propiedades combinatorias de las particiones de los cardinales. Finalmente, las definiciones de todos estos cardinales (incluyendo el cardinal de Erdös) describen ciertas operaciones combinatorias de los cardinales, y la existencia de cada una expresa la proposición de que el universo no se agota por estas operaciones.

4.2.3.2. Uniformidad e Identidad caprichosa

El criterio de *uniformidad* es enunciado por Maddy de la siguiente forma:

> Supongamos que una condición interesante sucede en un nivel bajo de la jerarquía. Si ninguna situación similar ocurre en el resto de la jerarquía, sería como si el universo hubiese perdido su complejidad en

[68]Por esto, Maddy señala que la afirmación $V = L$ viola este criterio, pues afirma que los subconjuntos que se agregan mediante la operación de potencia son sólo aquellos que son construibles mediante fórmulas de la teoría de conjuntos y no se consideran todos los subconjuntos posibles.

> los niveles más altos, como si se hubiera vuelto más plano, homogéneo y aburrido. *Uniformidad* dice que esto no sucede (Maddy 1988: 502)

El criterio de uniformidad afirma que el universo de los conjuntos conserva su uniformidad mientras se va *escalando* la jerarquía. Es decir, el universo conserva su uniformidad al considerar los niveles de la jerarquía acumulativa indexados por ordinales cada vez más grandes. La idea es que el universo conjuntista no cambia sus propiedades de forma substancial en los niveles más altos de la jerarquía, o cuando consideramos los conjuntos más grandes: las mismas condiciones que se presentaban en un nivel bajo se presentan en algún otro nivel más alto en la jerarquía. Por ejemplo, se sabe que si no existen cardinales inaccesibles (no numerables), entonces el único cardinal inaccesible será $\aleph_0$. La propiedad de ser inaccesible es muy particular, tanto que no puede ser que ω sea el único conjunto que la tenga. Por lo tanto, debe haber inaccesibles no contables, "de otra manera, el universo estaría disperso arriba de $\aleph_0$, o cambiaría su carácter de una manera inobjetable" (Maddy 1988: 502). Si así fuese, entonces sería posible definir el conjunto $\aleph_0$ como el único cardinal inaccesible, pero parece que esto es accidental y que no es precisamente una propiedad de exclusiva de ω el ser inaccesible. A esto es lo que Maddy llama *identidad caprichosa*, en el sentido de que, si fuese el caso que el único inaccesible sea ω, entonces sería posible identificar a ω con esta propiedad de manera aparentemente accidental, por lo cual deben existir inaccesibles no contables.

El CE cumpliría este criterio de manera similar a como Maddy argumenta que lo cumple un cardinal débilmente compacto. Como mencioné, la definición de un cardinal débilmente compacto se extrae del teorema de Ramsey. Este teorema afirma la existencia de un conjunto muy particular (un conjunto homogéneo) para cada partición del conjunto de subconjuntos de una cardinalidad finita determinada de ω. Es decir, afirma que existe un conjunto muy particular para el caso de omega. Por uniformidad, debe existir otro conjunto que cumpla con esta característica, pues de otra manera sería posible definir ω como el único conjunto que cumple con esta propiedad. La definición de un cardinal débilmente compacto es una generalización parcial de esta propiedad. Es parcial en tanto que los subconjuntos que se consideran en la definición no son subconjuntos de una cardinalidad finita cualquiera, si no específicamente de cardinalidad 2, al igual que el número de elementos de las particiones no son cualquier número natural. Sin embargo, el teorema de Ramsey garantiza que $\aleph_0$ es un cardinal débilmente compacto, y es en este sentido que la definición de cardinal débilmente compacto cumple con el criterio de *uniformidad* y garantiza que no exista *identidad caprichosa*.

Por otro lado, el cardinal de Erdös podría ser una generalización de la definición de un cardinal débilmente compacto, aunque en un sentido más débil porque ω no es un cardinal de Erdös. Sin embargo, CE garantiza que el fenómeno presente en el teorema de Ramsey no se presenta únicamente en ω. Esto porque a partir del teorema de Ramsey sabemos que $\omega \to (\omega)^n_2$ es el caso para todo n, es decir, informalmente $\omega \to (\omega)^{<\omega}_2$ es el caso (aunque no sea demostrable). Esto es, podemos aceptar que el cardinal de Erdös $\aleph_0(\omega)$ existe. Del teorema de Ramsey se sigue (informalmente) que existe el cardinal de Erdös $\aleph_0(\omega)$ y CE garantiza que esto no sucede únicamente para el caso de ω. Además, un cardinal de Erdös considera propiedades combinatorias muy generales: no depende de considerar subconjuntos de una cardinalidad finita determinada de un cardinal κ, sino para todos los subconjuntos finitos. De hecho, garantiza que no es un fenómeno exclusivo de los números cardinales, sino que es posible reproducir el fenómeno del teorema de Ramsey de manera uniforme para el caso en que el conjunto homogéneo que se usa simplemente tiene un tipo de orden en general, no sólo un cardinal. Por supuesto, esto difiere de la noción de *uniformidad* que mencioné más arriba, pues ésta considera a los fenómenos que se reproducen en ciertos niveles de la jerarquía acumulativa y se busca garantizar que ese fenómeno se reproduce en otro nivel de la jerarquía. En este caso, el cardinal de Erdös garantiza que este fenómeno se reproduce en otros casos diferentes, lo cual de cualquier forma garantiza lo que uniformidad busca: que no se presente una situación como la de *identidad caprichosa*.

Una razón más para defender la existencia de un cardinal de Erdös es que aumenta la capacidad deductiva de ZFC. En particular, podemos mencionar que aceptar CE implica la existencia de cardinales inaccesibles, no sólo porque ellos mismos son cardinales inaccesibles, sino porque PR1 garantiza la existencia de un cardinal inaccesible, y como vimos la existencia de un cardinal de Erdös permite encontrar un modelo de este principio. Un razonamiento similar se puede dar para notar que también implica la existencia de los cardinales de Mahlo. También implica la existencia de un cardinal débilmente compacto y un cardinal de Ramsey. Todo esto es deseable en vista de lo mencionado más arriba, en particular, parece que el criterio de maximalidad hace deseable la existencia de un cardinal inaccesible (de hecho, hace deseable la existencia de un cardinal de Mahlo en vista del criterio de inexhaustibilidad) y el criterio de uniformidad hace deseable que exista un cardinal débilmente compacto. A partir de esto, parece que es deseable defender el axioma CE. Con ello, garantizamos que existe el cardinal $\kappa(\omega)$ y con ello es posible fortalecer la prueba de consistencia expuesta por Koellner.

En general, es posible dar una justificación general de todos los axiomas de cardinales grandes. Maddy introduce otro criterio para la inclusión de axiomas: *un paso antes del desastre* (Maddy 1988: 485). Un ejemplo de la aplicación de este criterio es la introducción del *Axioma de Comprehensión*. La primera versión de este axioma (el Axioma de Comprehensión ingenuo) es un axioma paradójico por lo cual, era necesario introducir un axioma que tuviera una capacidad deductiva lo más cercana posible al Axioma de Comprehensión pero evitando paradojas. La idea detrás del criterio de *un paso antes del desastre* “es que los principios de construcción de conjuntos tan fuertes como sea posible acercándose a la inconsistencia”[69] (Maddy 1988: 485). Así, el Axioma de Comprehensión tiene la mayor capacidad expresiva y deductiva posible evitando las contradicciones que implicaba el Axioma de Comprehensión ingenuo. Los cardinales grandes mayores a los cardinales medibles son generalizaciones de estos últimos a partir de que si un cardinal κ es medible, entonces hay una inmersión elemental entre V y un modelo clase M cuyo punto crítico es κ. Así las definiciones de los demás cardinales grandes grandes son simplemente generalizaciones sobre la longitud de las secuencias sobre las que el modelo M es cerrado, y al mismo tiempo se garantiza la existencia de una inmersión elemental no trivial entre V y M. Sin embargo, esto tiene un límite:

[69]El énfasis es mío

Teorema 4.6. *Si $j : V \to M$ es una inmersión elemental, entonces $M \neq V$*

Así, cada axioma de cardinales grandes grandes postulan que los modelos M son cada vez *más parecidos* a V, en la medida que M es cerrado bajo secuencias de mayor longitud. Por supuesto, la clave está en que ninguno de los modelos M cerrados bajo secuencias de longitud mayor son el universo V, pues por el teorema 4.6 sería contradictorio. De esta manera, cada axioma de cardinales grandes grandes son cada vez más fuertes evitando contradecir el teorema de Kunen, y a partir del criterio de *un paso antes del desastre*, están justificados. Esto permite dar justificación a los cardinales fuertemente compactos, a los supercompactos, a los cardinales enormes, etc. Como mencioné en general a todos los cardinales grandes. Esto es relevante no sólo porque la existencia de estos cardinales garantiza la existencia de los cardinales de Erdös, sino porque permite dar una justificación parcial a los principios PR3 y PR4. Esto es así ya que si la existencia de un cardinal supercompacto está justificada, entonces eso garantiza que existe el número de Lowenheim-conjunto de *L2K*. Es decir, existirá un cardinal κ tal que cada conjunto de fórmulas de L2K será satisfecho en algún cardinal menor a κ, si es satisfacible del todo. Esto representaría evidencia a favor de PR3 y PR4 en la medida en que mostraría que es suficiente considerar las estructuras conjuntistas para satisfacer los conjuntos de fórmulas que de hecho son satisfacibles. De hecho, creo que es probable que suponer la existencia de un cardinal supercompacto permite dar una prueba de la consistencia de PR3 y PR4. Sin embargo no tengo idea de cómo podría ser esa prueba y hasta ahora esto es lo mejor que tenemos para justificar PR3 y PR4.

Para concluir esta sección discutiré una posible defensa del principio de reflexión basada la modelización de los principios de reflexión.

4.2.4. La modalización de los principios de reflexión

Existen ciertos principios en la aritmética que, al igual que los considerados aquí, son llamados *principios de reflexión*. Estos principios tienen la siguiente forma esquemática:

(PRA) $Prov(n) \to \Phi_n$

donde, siguiendo la notación introducida en la sección 3.3.1, Φ_n denota a la expresión del lenguaje de la aritmética de Peano (PA) cuyo número de Gödel es n y $Prov(n)$ es la afirmación de que Φ_n es demostrable en PA en primer orden.[70] Junto con este principio se suele asociar un *principio de reflexión global*:

(PRAG) $\forall x(Prov(x) \to T(x))$

Donde el rango de las variables de la fórmula son números naturales y $T(x)$ es un predicado de verdad introducido primitivamente.[71] Estos principios de reflexión

[70]Es decir, donde los axiomas de PA están expresados en un lenguaje de primer orden. La característica principal de estos axiomas es que el principio de inducción es un esquema expresado en primer orden, véase capítulo 1, sección 1.3.

[71]Para considerar el principio de reflexión global se utiliza una teoría aritmética como PA extendida con un predicado de verdad $T(x)$, junto con una serie de axiomas que permiten caracterizar el comportamiento de este predicado, veáse Ketland (2005: 78)

buscan expresar la idea de que los teoremas de la aritmética son verdaderos. Cabe mencionar que estos principios de reflexión pueden demostrarse en una metateoría como la utilizada en el capítulo 1, siendo consecuencias directas del teorema de corrección para la lógica de primer orden. Algunas de las instancias del esquema (PRAG) no se pueden demostrar dentro de PA. Por ejemplo, consideremos el símbolo '$\perp$' (que se suele usar para denotar la contradicción o se considera como la constante de falsedad). Se suele definir el símbolo de negación como $\neg\varphi \equiv_{def} \varphi \supset\perp$. Así consideremos la instancia de PRA:

$$Prov(\perp) \to\perp$$

Esto sería equivalente entonces a $\neg Prov(\perp)$, que es la oración canónica que expresa la consistencia de PA, que no es demostrable en PA por el segundo teorema de incompleción de Gödel.

Nogina (2014) demuestra que es posible modalizar estos principios, garantizando que todos los principios de reflexión son equivalentes a una fórmula modal del tipo:

$$\Box\Phi \to \Phi$$

donde el operador $\Box$ (generalmente entendido como el operador de *necesidad*) se interpreta intuitivamente como un predicado de derivabilidad en PA. La estrategia de Nogina es definir dos operadores de derivabilidad: el operador $\Box$ que tiene el significado mencionado anteriormente y un operador $u : \Phi$, cuyo significado es "u es el número de Gödel de una prueba de la oración Φ". En el contexto del lenguaje de la lógica GLA (que es la unión de los lenguajes de la lógicas de Gödel – Löb, (GL), y la Artemov, (A)), Nogina demuestra que con esta interpretación de los operadores, todos los principios de reflexión para la aritmética son equivalentes a fórmulas o bien del tipo $\Box\Phi \to \Phi$ o bien $\Box^k u : \Phi \to \Phi$ (donde $\Box^k$ es una k-secuencia de operadores $\Box$). Nogina también muestra que existe una jerarquía de principios de reflexión determinada por la cantidad de iteraciones de los operadores modales en dichos principios.

Existen varios intentos de modalizar ciertos predicados de la teoría de conjuntos de manera análoga a como lo hace Nogina. Sin embargo, estos intentos son de momento sólo exploraciones de esta posibilidad. Joosten (2010) da la que me parece la mejor manera hasta ahora en que se podría garantizar el contenido que tienen los principios de reflexión conjuntista. Oosten trata de expresar mediante operadores modales ciertas cuestiones sobre cardinales grandes. Oosten utiliza una lógica modal de interpretabilidad donde se tienen dos operadores modales: $\Box$ y $\rhd$. El objetivo de Oosten es poder interpretar el operador modal $\Box$ como "verdad en todos los V_κ", donde V_κ es un nivel de la jerarquía acumulativa.[72] Por otro lado, el significado de $\rhd$ es una relación binaria entre oraciones tal que para una teoría base T, $A \rhd B$ si $T + A$ puede interpretar $T + B$, en el sentido de que la teoría $T + A$ puede dar un modelo de $T + B$. Por ejemplo, la oración "si una fórmula φ es verdadera en todos los V_κ, entonces es verdadera" se formalizaría en esta interpretación:

$$\Box\varphi \to \Phi$$

[72]De hecho, κ es un cardinal inaccesible, por lo cual V_κ es un modelo de ZFC, véase el Apendice.

Por otro lado, el teorema 4.1 puede tener una traducción adecuada:

$$\text{ZFC+CE} \rhd \text{ZFC+}\Gamma_n^{(2)}\text{-reflexión}$$

Oosten utiliza una lógica bimodal, en tanto que contiene dos operadores modales: [1] y [ ]. Ambos operadores son operadores de necesidad donde estos significan respectivamente "verdad en todos los V_κ" y "verdad en todos los V_λ", donde κ y λ son distintos cardinales, y κ es de tipo 1 y λ es de tipo 0. Oosten no discute mucho sobre esta distinción de tipos; sin embargo, pone como posible caso uno en el que el tipo 1 son cardinales de Mahlo y el tipo 0 son cardinales inaccesibles. Con esto puede caracterizar:

> $A \rhd B \leftrightarrow$"En todos los modelos de V_κ donde A es verdadero, B es verdadero o hay un V_λ donde B es verdadero" (Joosten 2010: 11)

En términos modales:

$$A \rhd B \leftrightarrow [1](A \to (B \vee \langle\,\rangle B))$$

Donde $\langle\,\rangle$ es un operador de posibilidad para [ ], y cuyo significado es "verdad en algún V_λ de tipo 0". Esta podría ser una modalización interesante y aparentemente adecuada del principio de reflexión, siempre que se interprete el operador de necesidad estrictamente sin hacer distinción entre los tipos que se están considerando. Sin embargo, explorar cómo podría hacerse esto excedería los intereses de la presente tesis.

Asumiendo que tenemos dicha lectura a la mano, es posible formalizar el principio de reflexión conjuntista de la siguiente manera:

$$\Phi \to \Diamond\Phi$$

Donde $\Diamond$ es un operador de posibilidad, cuyo signifado intuitivo sería "verdad en algún V_κ", dando una lectura muy natural del Principio de Reflexión. Cabe notar que, dentro de la lógica modal (véase Garson 2006: 97 - 98) esta fórmula es equivalente a la mencionada anteriormente:

$$\Box\Phi \to \Phi$$

que, como vimos, es el principio de reflexión para la aritmética, donde se interpreta el operador $\Box$ como un predicado de probabilidad (o demostrabilidad). De esta manera, es posible observar que lo se solicita es que se cumpla una relación de reflexividad entre los oraciones demostrables en la aritmética por un lado, y los modelos de la teoría de conjuntos por el otro.

Ahora bien, existen distintos argumentos para los principios de reflexión para la aritmética. Hay una discusión sobre la justificación de los principios de reflexión en el marco del debate sobre el deflacionismo para el caso de la noción de verdad. Por un lado, Shapiro (1998b) y Ketland (1999) prueban los principios de reflexión en PA en la teoría que resulta de añadir a PA un predicado de verdad y axiomas que describen cómo se comporta dicho predicado.[73] De esta manera, Ketland sostiene que "Shapiro y yo [Ketland] argumentamos que la justificación para los principios de reflexión involucra la noción de verdad" (Ketland, 2005, pág. 76). Los principios de reflexión no son demostrables en PA, de esta manera, tanto Shapiro como Ketland afirman que su justificación depende de la noción de verdad, y de la definición de esta. De hecho, Ketland sostiene lo que él llama *obligación condicional epistémica*: "Si alguien acepta una teoría matemática base S, entonces está comprometido a aceptar un grupo de *otros enunciados* del lenguaje de la teoría base (uno de estos es la oración G de Gödel)" (Ketland 2005: 79). Los principios de reflexión son enunciados como estos. Por otro lado, Tennant defiende que estos principios no requieren una justificación como la de Ketland y Shapiro, pues afirma que alguien que use una teoría matemática S "simplemente reflexiona en los métodos comunes de prueba, y le gusta lo que ve" (Tennant 2005: 91). La justificación de los principios para Tennant radica en algo parecido a la justificación que enuncia Kreisel de la premisa 2 del argumento de Kreisel, pues sólo depende de reflexionar sobre las reglas y métodos de prueba y tener confianza en que son correctos. Por lo cual, Ketland concluye:

> No es necesaria más justificación para el nuevo compromiso hecho por expresar un compromiso anterior. Tan pronto como alguien aprecia el proceso de reflexión, y cómo este está expresado por el principio de reflexión, uno ya tiene una explicación de porqué alguien que acepta S debe aceptar también todas las instancias del principio de reflexión (Ketland 2005: 92).

Sin embargo, no es posible usar justificaciones como las expuestas para justificar el principio de reflexión para la aritmética, para justificar los principios de reflexión para la teoría de conjuntos. En primer lugar, no hay razón para esperar que incluir un predicado de verdad (y axiomas para este) garantice la demostración del principio de reflexión, a diferencia del principio para el caso de PA; éste no relaciona un predicado de probabilidad con la noción de verdad. Finalmente, por esto el principio de reflexión no es trivial, en tanto que su demostración no es tan sencilla como un teorema de corrección para los axiomas de ZFC. La relación que describe el principio de reflexión es más cercana a la que le atribuyen Rayo y Uzquiano:

> los axiomas de ZFC en segundo orden no implican un principio de reflexión que asegure que si una oración es verdadera, entonces es verdadera en algún modelo estándar. Así puede haber oraciones del lenguaje de la teoría de conjuntos en segundo orden que son verdaderas pero insatisfascibles, u oraciones que son válidas pero son falsas (Rayo y Uzquiano 1999: 317).

Así, el principio afirma una relación entre verdad y satisfacción, y es una muy particular, pues lo que afirma es que la verdad de una fórmula en lo que podríamos

[73]Estos axiomas son las condiciones de verdad proporcionadas por una definición estándar de verdad, parecida a la expuesta, por ejemplo, en el capítulo 1. Simplemente se señala (en forma de equivalencias) las condiciones de verdad para oraciones de identidad, para fórmulas cuya conectiva principal es la conjunción, para fórmulas cuya conectiva principal es la negación y para fórmulas cuya conectiva principal es el cuantificador universal (cf. Ketland 2005: 78).

(de manera retórica) decir un “modelo” particular (el “modelo pretendido”) implica la existencia de otro modelo donde es verdadera. Así, no se puede esperar una prueba del principio de este tipo. Por otro lado, el argumento de Tennant depende aún más de que el principio sea una relación entre un predicado de probabilidad y uno de verdad. Pues el proceso de reflexión al que refiere funciona para el caso de las reglas de inferencia. De nuevo, el principio de reflexión no es trivial ni un requisito que los axiomas de ZFC deba satisfacer, al menos no evidentemente, y la mera inspección de los axiomas no puede darnos la garantía de su verdad. En contraste, el principio de reflexión para la aritmética parece un requisito de los axiomas, parecido al requisito que representa la premisa P2 del argumento de Kreisel. Sin embargo, aunque no haya manera de presentar alguna analogía entre estos argumentos y algunos posibles para los principios de reflexión en la teoría de conjuntos, me parece que la relación que existe entre ellos se puede utilizar. Un posible argumento útil usando esta vía podría incluir una defensa de porqué es natural o adecuada la relación reflexiva expresada por el principio de reflexión, más allá, por supuesto, de la justificación más tradicional que apela a la intuición de que el universo conjuntista no es describible por una fórmula.

4.3. Conclusiones

En este capítulo, expuse una prueba presentada por Koellner para garantizar que los Principios de Reflexión son consistentes relativo a un cardinal de Erdös. Con ello, y las formulaciones del Principio de Reflexión PR1 y PR2 expuestas en el capítulo anterior, garantizan que el Principio de Kreisel es igualmente consistente, para el caso de la formulación original de este (en términos de oraciones y fórmulas). Sin embargo, la prueba no era suficiente para el caso más general de la segunda formulación del principio expuesta en la sección 3.1.1 del capítulo 3.

(PKC) $\forall i \forall \Gamma \forall \alpha (\mathrm{Val}(\Gamma^i, \alpha^i) \leftrightarrow \mathrm{V}(\Gamma^i, \alpha^i))$

Argumenté que hay buenas razones para considerar como aceptable el Principio de Reflexión en vista de dos puntos: la prueba de consistencia de Koellner dota de una gran justificación al Principio de Reflexión y hay buenas razones para aceptar la existencia de un cardinal de Erdös. También exploré brevemente la posibilidad de establecer una analogía entre la discusión de los principios de reflexión para la aritmética y los considerados aquí, y concluí que dicha analogía sería poco útil en vista de las características particulares de ambos principios. Sin embargo, creo que es posible (en principio) establecer un argumento global a favor de ambos en vista de la relación de estos con ciertas fórmulas modales de lo que podrían ser instancias.

Me parece que la conclusión más importante de todo esto, es que es posible notar qué tipo de presupuestos hay que tomar para poder argumentar a favor del Principio de Kreisel y que un posible rechazó de este requeriría una respuesta que rechace lo que me parece una clara prueba matemática. No pretendo dar una respuesta final a la pregunta sobre la verdad del Principio de Kreisel, sin embargo, creo que de lo anterior es posible saber qué es necesario para sostener que estos se cumplen. Pero, creo que también he dado algunas buenas razones para aceptar estos supuestos, y con ello la satisfacción del principio. Sin embargo, esto tiene la limitante de que no es posible dar un argumento (en términos de una prueba como la de Koellner) para el caso de la noción de consecuencia lógica y para la segunda formulación más general del Principio de Kreisel, y que estos argumentos funcionan sólo para la formulación original.

Conclusiones

En este trabajo, he presentado el Principio de Kreisel y el argumento con el cual Kreisel defiende que el principio se sostiene para la lógica de primer orden. De nuevo, el Principio de Kreisel es la afirmación:

(PK) $\forall i \forall \alpha(\mathrm{Val}(\alpha^i) \leftrightarrow \mathrm{V}(\alpha^i))$

donde $Val(\alpha^i)$ es la afirmación de que α es intuitivamente válida (o bien, que es verdadera bajo cualquier interpretación) y $V(\alpha^i)$ es la afirmación de que α es verdadera en cualquier interpretación cuyo dominio es un conjunto (o bien, que α es una verdad lógica tarskiana), e i es el orden de cuantificación del lenguaje en que está expresada α. El argumento que Kreisel proporciona para defender que el principio es verdadero para la lógica de primer orden consiste en las siguientes premisas defendidas por Kreisel:

(P1) $\forall i \forall \alpha(\mathrm{Val}(\alpha^i) \rightarrow \mathrm{V}(\alpha^i))$ [esto se acepta "cuando uno toma como garantía que la lógica se aplica a estructuras matemáticas" (cf. Kreisel 1967: 154)

(P2) $\forall i \forall \alpha(\mathrm{D}(\alpha^i) \rightarrow \mathrm{Val}(\alpha^i))$ [se sostiene pues "uno reconoce la validez de las reglas de Frege" (cf. Kreisel 1967: 153)]

(P3) $\forall \alpha(\mathrm{V}(\alpha^1) \rightarrow \mathrm{D}(\alpha^1))$[Teorema de compleción para la lógica de primer orden]

De estas premisas es posible obtener la conclusión deseada: $\forall \alpha(\mathrm{Val}(\alpha^1) \leftrightarrow \mathrm{V}(\alpha^1))$. Hay que recordar también que este argumento no funciona en el caso en que α es una fórmula de segundo orden, debido a que en la lógica de segundo orden (P3) no es el caso.

Así mismo, presenté una prueba que es a mi parecer la mejor evidencia que hay hasta ahora a favor de éste para el caso de la lógica de segundo orden. Esta evidencia es la prueba que proporciona Koellner de que los principios de reflexión que actúan sobre fórmulas con cuantificación universal en segundo orden son consistentes relativo a un cardinal de Erdös. El hecho de que esta prueba represente evidencia a favor del Principio de Kreisel depende de dos aspectos: primero, que el Principio de Kreisel sea equivalente al Principio de Reflexión y, segundo, de que exista una relación entre la noción de *validez en estructuras cuyo dominio es una clase* y una noción de *validez preteórica.* En sentido estricto, la equivalencia entre el Principio de Reflexión y el Principio de Kreisel depende de la relación entre la noción de *validez en estructuras cuyo dominio es una clase* y una noción de *validez preteórica.*

En el primer capítulo, expuse el marco formal en que se inscribe el problema del Principio de Kreisel para el caso de la lógica de segundo orden. El argumento de Kreisel mostraba que el Principio de Kreisel es verdadero para el caso de la lógica de primer orden, y este argumento dependía del teorema de compleción para ella. En el primer capítulo también expliqué en qué consiste la lógica de primer orden y la prueba del teorema de compleción para ésta. Así mismo, expliqué en virtud de qué razones esto no era el caso para la lógica de segundo orden (con la semántica estándar). El primer capítulo tiene dos propósitos: explicar porqué la lógica de primer orden cumple con el teorema de compleción y explicar las razones por las cuales la lógica de segundo orden no. Una de estas razones es que la semántica que se considera es la semántica estándar, por lo cual debía explicar porqué uso esta semántica y no otra. El propio Kreisel dice de su argumento que "todo esto es el caso para fórmulas de primer orden. Para fórmulas de orden superior no tenemos una prueba convincente de e.g. $\forall\alpha(Val(\alpha^2) \leftrightarrow V(\alpha^2))$, aunque uno esperaría una" (Kreisel, 1967, pág. 157). La razón por la cual no hay una prueba del Principio de Kreisel para el caso de la lógica de segundo orden es que la premisa (P3) del argumento de Kreisel apela al teorema de compleción. También, expuse que la razón por la cual la lógica de segundo orden con semántica estándar no cumple con el teorema de compleción es que posee una gran capacidad expresiva. Una manera en que se podría sostener que el Principio de Kreisel era el caso para la lógica de segundo orden es argumentar la lógica de segundo orden satisface el teorema de compleción usando una semántica no estándar ya sea la semántica de *Henkin* o la semántica de *primer orden*. Sin embargo, argumenté a partir de una observación de Bueno (2010) que dichas semánticas utilizan mecanismos *ad hoc* para garantizar que se cumplía el teorema de compleción. Después de todo, la lógica de segundo orden con semánticas no estándar no parece tener mayor capacidad para caracterizar estructuras matemáticas que la lógica de primer orden, lo cual insinúa que éstas no son semánticas adecuadas para un lenguaje de segundo orden. En vista de esto, opté por utilizar la semántica estándar para estudiar el Principio de Kreisel en la lógica de segundo orden.

En el segundo capítulo me ocupé de explicar qué relación existe entre la noción de *validez en estructuras cuyo dominio es una clase* y una noción de *validez preteórica*. La necesidad de explicar esta relación radica en que tanto el principio como el argumento de Kreisel apelan a una noción *intuitiva* de validez (la cual llamé *validez preteórica*). Explicar qué relación existe entre ambas nociones permite estudiar el Principio de Kreisel dentro de una teoría formal y con conceptos matemáticos, por lo cual, era importante establecer esta relación para poder demostrar la equivalencia entre el Principio de Kreisel y el Principio de Reflexión para la teoría de conjuntos. En este capítulo expliqué con mayor detalle a qué me refiero con la noción de validez *preteórica* (o *intuitiva*) señalando que considera las inferencias tradicionalmente usadas en el contexto de la matemática clásica y presentando (brevemente) tres teorías de la consecuencia lógica para ejemplificar la clase de propiedades que se ha considerado que tiene esta relación de consecuencia lógica. Formulé una noción de *validez preteórica* (NP) y argumenté que satisfacía dos propiedades que se suelen atribuir a la relación de consecuencia lógica: *Formalidad* y *Necesidad*. Discutí ambas propiedades y di formulaciones informales pero precisas de ellas. Recordemos la noción (NP):

(NP) Φ es una consecuencia lógica de Γ si Φ vale en todas las posibilidades bajo toda interpretación de la terminología no lógica en la Γ vale

Defendí que la relación que existía entre esta noción y la noción de *validez en estructuras cuyo dominio es una clase* es considerar que todas las posibles interpretaciones de la terminología no lógica era la colección de todas las clases (entendidas como colecciones de conjuntos). En la sección 2.2.2.3, mostré que esta noción satisface las dos propiedades que se suelen atribuir a la relación de consecuencia lógica. Vale la pena mencionar que Peter Smith (2011: 24) plantea el significado del predicado $Val(\alpha^i)$ (del argumento de Kreisel) en términos muy similares a la noción (NP).

En el tercer capítulo presenté un análisis del argumento de Kreisel, así como la equivalencia entre el Principio de Kreisel y el Principio de Reflexión para la teoría de conjuntos ZFC. El objetivo principal de este capítulo es mostrar que el Principio de Kreisel (usando la noción (NPC) como interpretación del predicado $Val(\alpha^i)$) es equivalente al Principio de Reflexión en la teoría de conjuntos. La razón para analizar el argumento de Kreisel era entender las razones por las cuales éste parece tan convincente y correcto. A partir de este análisis, mostré que había una ambigüedad en el significado del predicado $Val(\alpha^i)$: la premisa 1 del argumento de Kreisel apelaba a que el significado de $Val(\alpha^i)$ era que α es válida en toda estructura, mientras que la premisa 2 apelaba a la validez informal que se suele atribuir a las reglas de inferencia. Señale que esta ambigüedad exponía la existencia de dos criterios que debía cumplir una noción preteórica de validez para poder caracterizar el tipo de inferencias que probablemente Kreisel tenía en mente. El análisis del argumento de Kreisel incluye la observación de que la formulación original del Principio de Kreisel estaba en términos de la noción de verdad lógica. Expliqué que una consecuencia del teorema de compacidad es que es posible reducir la noción de consecuencia lógica a la de verdad lógica y, por lo tanto, la formulación original del Principio de Kreisel era suficiente para el caso de la lógica de primer orden. Sin embargo, esto no era el caso en la lógica de segundo orden, pues ésta no cumple con el teorema de compacidad. Debido a esto, fue necesario plantear el Principio de Kreisel en términos de consecuencia lógica. El último punto del análisis fue presentar el argumento de Kreisel, interpretando el predicado $Val(\alpha^i)$ como la noción (NPC)

(NPC) Φ es una consecuencia lógica$_{NIC}$ de Γ si Φ vale en todas las estructuras $\mathfrak{A}^C$ cuyo dominio es una clase en la que Γ vale

que era la noción (NP), con la precisión de que incluye una noción de estructura $\mathfrak{A}^C$ que es mucho más precisa y recupera la idea de entender la colección de todas las interpretaciones (presente en (NP)) con la colección de todas las clases. En la última parte del capítulo presenté cuatro formulaciones del Principio de Kreisel (PK1-PK4) y cuatro formulaciones del Principio de Reflexión (PR1-PR4), y mostré que cada formulación del Principio de Kreisel era equivalente a una formulación del Principio de Reflexión, así como algunos teoremas sobre la satisfacción de éstos. Estas formulaciones dependen estrictamente de entender el predicado $Val(\alpha^i)$ como la afirmación de que α es una verdad lógica$_{NPC}$.

Finalmente, en el último capítulo presenté una prueba de Peter Koellner de la consistencia de los principios de reflexión. Esta prueba representa una prueba de la consistencia del Principio de Kreisel a partir de su equivalencia con el Principio de Reflexión. Expliqué que la prueba de Koellner garantiza la consistencia de las formulaciones PR1 y PR2 (y por lo tanto PK1 y PK2) del capítulo 3, pues éstos consideran fórmulas la prueba de Koellner también. Sin embargo, esta prueba no puede establecer la consistencia de las formulaciones PR3 y PR4 (ni tampoco de PK3 ni PK4), debido a que éstas no son esquemas sino fórmulas. Hay que hacer énfasis en que esto no es un problema menor: recordando que la lógica de segundo orden no cumple con el teorema de compacidad, es necesario formular el Principio de Kreisel en términos de consecuencia lógica. La instancia del Principio de Kreisel que es importante para mi interés es:

$$\forall\Gamma\forall\alpha(Val(\Gamma^2,\alpha^2) \leftrightarrow V(\Gamma^2,\alpha^2))$$

A partir de la versión del argumento de Kreisel que presenté en la sección 3.2, la oración $\forall\Gamma\forall\alpha(Val(\Gamma^2,\alpha^2) \to V(\Gamma^2,\alpha^2))$ es verdadera (pues es la premisa 1 del argumento de Kreisel). Sin embargo, la oración $\forall\Gamma\forall\alpha(V(\Gamma^2,\alpha^2) \to Val(\Gamma^2,\alpha^2))$ está en duda. Consideremos un argumento $A = \langle\Gamma^2,\alpha^2\rangle$ que es *preteóricamente inválido* (es decir, supongamos que $\neg Val(\Gamma^2,\alpha^2)$). Por lo tanto, hay una $\mathfrak{A}^C$ donde todas las oraciones de Γ^2 son verdaderas mientras que α^2 es falsa. Consideremos el conjunto $B = \gamma_1^2, \gamma_2^2, \gamma_3^2, \ldots, \neg\alpha^2$ donde $\gamma_1^2, \gamma_2^2, \gamma_3^2, \cdots \in \Gamma^2$ (por simplicidad, supondremos que $\gamma_1^2, \gamma_2^2, \gamma_3^2, \ldots, \alpha^2$ son enunciados). Consideremos $x \subseteq \omega$, el conjunto de los números de Gödel de las fórmulas de este conjunto, y supongamos PK3. Como existe una $\mathfrak{A}^C$ que es un modelo de B, sabemos que $\forall m \in xSATS(Y,w,Z,m)$ con algunas asignaciones w y Z. Por PK3, $\exists y\exists z\exists w\forall m \in xsats(y,w,z,m)$, es decir, hay un conjunto que satisface al conjunto B. Por lo tanto, $\neg V(\Gamma^2,\alpha^2)$. Entonces es posible afirmar que $\neg Val(\Gamma^2,\alpha^2) \to \neg V(\Gamma^2,\alpha^2)$, por contraposición y generalización $\forall\Gamma\forall\alpha(V(\Gamma^2,\alpha^2) \to Val(\Gamma^2,\alpha^2))$, que era lo que queríamos mostrar. Sin embargo, esto sólo funciona si Γ^2 es finito o existe un subconjunto finito de Γ^2 que implique α^2lo cual no es necesariamente cierto debido a que la lógica de segundo orden no cumple con el teorema de compacidad. Después de todo, es posible que $|x| = \aleph_0$, y este caso es el que no está garantizado por la falsedad del teorema de compacidad. Ni siquiera es posible afirmar que es consistente la afirmación de PK3. En principio, podría ser inconsistente suponer PK3, y como mencioné algún argumento que sirva como evidencia para PK3 y PK4 implica un compromiso con la existencia de un cardinal supercompacto (al menos, tomando en cuenta la evidencia que existe hasta ahora).

Alguien podría sugerir una defensa de los principios PR3 y PR4 en vista de que hay una justificación informal del Principio de Reflexión: "los principios de reflexión pretenden articular la idea informal de que el tamaño del universo es "absolutamente infinito" y por lo tanto no puede ser "caracterizado desde abajo"" (Koellner 2009: 208). Maddy también señala la existencia de esta justificación: "el universo de los conjuntos es tan complejo que no puede ser completamente descrito; por lo tanto, cualquier verdad del universo entero debe ser verdad en un segmento inicial del universo" (Maddy 1988: 503). Alguien podría argüir que el universo conjuntista (así como no puede ser descrito por ninguna fórmula) no puede ser descrito por un conjunto de fórmulas, de manera parecida a como argumenté en la sección 4.2.2. Finalmente, esta es una de las razones (ofrecidas en la sección 4.2.2) para aceptar el Principio de Kreisel en su presentación original. Sin embargo, el propio Koellner muestra que esta intuición puede tener problemas:

Teorema 4.7. ***Con1*** $\Gamma_1^{(3)}$*- reflexión es inconsistente (Koellner 2009: 213)*

Es decir, los principios de reflexión con cuantificaciones universales de tercer orden son inconsistentes; de hecho, son los principios inmediatamente más fuertes que los que considera en su prueba. No busco afirmar que PR3 y PR4 son inconsistentes, sino que los principios de reflexión no son inmunes a la inconsistencia. Finalmente, es posible que PR3 y PR4 sean inconsistentes, y parece que su justificación informal sería insuficiente.

En la sección 4.2.2 argumenté que la prueba de consistencia de Koellner proporcionaba una justificación al Principio de Reflexión a partir del concepto de conjunto. En realidad, la prueba de Koellner permite afirmar que la justificación informal del Principio de Reflexión de hecho puede justificar el este principio. Como he mencionado repetidamente la razón que usualmente se proporciona para afirmar que el Principio de Reflexión está justificado a partir del concepto de conjunto es que captura la idea informal de que el universo conjuntista es indefinible. Sin embargo, como mencioné esta idea informal puede generar inconsistencias, a partir del teorema Con1. Es por esto, que la justificación del Principio de Reflexión recae de manera importante en la prueba de Koellner. De ella, es posible concluir que el Principio de Reflexión está justificado en el concepto de conjunto, y por lo tanto, hay buenas razones para aceptar el Principio de Kreisel en su formulación original.

Ahora bien, el teorema 4.1 depende de la existencia de un cardinal de Erdös. En la sección 4.2.3 discutí algunos criterios presentes en la discusión sobre la aceptación de nuevos axiomas que permiten dar razones para aceptar la existencia de los cardinales Erdös. Creo que existen buenas razones para aceptar la existencia de los cardinales de Erdös. En general, hay buenas razones para aceptar todos los axiomas de cardinales grandes, a pesar de que estos sólo cuentan con pruebas de consistencia relativa a otros axiomas de cardinales grandes. Finalmente, se suele aceptar su existencia a pesar de esto y se suelen obtener resultados importantes e interesantes a partir de estos axiomas. De momento, creo que he presentado algunas buenas razones para aceptar la existencia del cardinal de Erdös $\kappa(\omega)$, lo cual es suficiente para mis propósitos presentes. Por último, en la sección 4.2.4 presenté una posible manera de abordar el problema de la validez del Principio de Reflexión apelando a una posible modalización de estos. No me detendré más en esta discusión; sin embargo, creo que esta podría ser una manera de abordar el problema muy fructífera filosóficamente hablando, siempre que dicho argumento defienda la idea de fondo de un Principio de Reflexión: la exigencia de una relación reflexiva entre los modelos de la teoría de conjuntos.

Hay que notar que en la prueba de Koellner no demuestra que el Principio de Reflexión sea verdadero (y por lo tanto, tampoco el Principio de Kreisel), así que en sentido estricto no está establecido que la definición tarskiana sea adecuada (aun considerando que la relación establecida en el capítulo 2 es correcta). Poder afirmar que el Principio de Kreisel es verdadero requeriría una prueba de que el Principio de Reflexión es verdadero. En este sentido, no hay hasta ahora una garantía de que la definición tarskiana de consecuencia lógica sea *extensionalmente adecuada.*

Sin embargo, la prueba del teorema 4.1 sí ofrece buenas razones para considerar al Principio de Reflexión como justificado, y en vista de esto podemos considerar que el Principio de Kreisel es aceptable. De nuevo, esto se puede afirmar a partir de la prueba de Koellner de las fórmulas por separado, pero no se puede sostener para el caso más general de consecuencia lógica. Para éste último no hay manera de afirmar ni siquiera en qué condiciones una definición tarskiana es suficiente para evaluar correctamente la satisfacción de diferentes conjuntos de fórmulas. Esto puede parecer una conclusión débil, sin embargo, las observaciones hechas en el capítulo 2 pueden mostrar cómo esta conclusión es aún más fuerte de lo que parece.

Shapiro considera que el Principio de Kreisel es la afirmación de una *presuposición* del uso de la jerarquía acumulativa de la teoría de conjuntos para proporcionar una semántica de los lenguajes formales (Shapiro, 1991: 142). En el capítulo 2 mencioné que Shapiro enuncia dos condiciones informales que se le pueden pedir a una semántica: *conformidad* y *suficiencia*. Shapiro explica que el Principio de Kreisel es la afirmación de que ambas condiciones se cumplen cuando se considera a la teoría de conjuntos como semántica de los lenguajes formales. Por un lado, $\forall\alpha(Val(\alpha^2) \to V(\alpha^2))$ afirma la conformidad de la teoría de conjuntos como semántica, en tanto que es la afirmación de que cada modelo de la semántica (definida en la teoría de conjuntos) de un lenguaje formal es efectivamente una interpretación legítima del lenguaje. Y esto es precisamente la justificación que explique en el capítulo 3 de esta premisa; después de todo, confiamos en que los objetos de la jerarquía acumulativa de la teoría de conjuntos sean interpretaciones legítimas de los lenguajes formales. Así, la oración $\forall\alpha(V(\alpha^2) \to Val(\alpha^2))$ (la afirmación cuya justificación depende del teorema de compleción en el argumento de Kreisel para el caso de la lógica de primer orden) es la afirmación de la suficiencia de la teoría de conjuntos. Esta es la afirmación de que hay suficientes interpretaciones en la semántica (en este caso la teoría de conjuntos) para abarcar todas las posibles interpretaciones consideradas. Si el Principio de Kreisel es verdadero, tendremos la garantía de que si una oración es inválida, habrá al menos una interpretación en la teoría de conjuntos que mostrará que es inválida, es decir, hay suficientes interpretaciones para proporcionar contraejemplos a todas los oraciones inválidas (análogamente, como lo explique en los argumentos de la sección 4.3 del capítulo 4 y en el argumento de arriba).

La pregunta es si es posible garantizar la suficiencia de la teoría de conjuntos a partir de la prueba de Koellner. La respuesta es afirmativa, si consideramos que la colección de interpretaciones se encuentra acotada al ordinal $\delta < \kappa$ (donde $\kappa = \kappa(\omega)$). Después de todo, el nivel V_δ es el nivel donde se satisfacen los principios de reflexión. El problema es que no hay razones claras para acotar de esta manera la colección de las interpretaciones. Sin embargo, si hay un cardinal de Erdös, sabemos que hay una interpretación que mostrará que una oración inválida de hecho lo es (bajo el supuesto de que es inválida en V_δ). Tal vez, la mejor conclusión que podemos obtener a este respecto es que si una oración es inválida, no es necesario apelar a una interpretación no conjuntista para mostrarlo. Esto ciertamente no garantiza la condición de suficiencia, aunque sí puede evitar algunas preocupaciones sobre el papel de la jerarquía acumulativa como semántica de los lenguajes formales. Entre ellas, la posible necesidad de apelar a clases propias para la satisfacción de estas fórmulas. La clave está en la precisamente en el término 'garantía'. No es posible tener una garantía; sin embargo, sí es posible tener una justificación. Abordar el Principio de Kreisel desde la perspectiva de Shapiro nos permite tener una conexión mucho más cercana con el Principio de Reflexión. Consideremos el principio PK1 y el principio PR1:

(PK1) $\exists X(\Phi/X) \to \exists x(\Phi/x)$

(PR1) $\Psi \to \exists x((Z2/x) \wedge (\Psi/x))$

Es posible notar informalmente que de hecho son equivalentes. Supongamos que una oración Φ es verdadera en alguna clase propia y que de hecho es verdadero PR1. Sabemos que toda clase propia es isomorfa a V, por lo cual Φ es verdadera en V (por una versión global del corolario 1.27 del capítulo 1 o Teorema del isomorfismo), y por PR1, Φ es verdadera en un conjunto que es el consecuente de PK1. Informalmente, PR1 implica PK1, y un argumento similar se puede proporcionar para argumentar que PK1 implica PR1. La suposición importante está en notar que toda clase propia es isomorfa a V. El contenido informal de PK1 no es más que el de PR1, es decir, PK1 es la afirmación de que el universo es indefinible. En este sentido, el Principio de Kreisel está tan justificado como el Principio de Reflexión.

Ahora bien, alguien puede argumentar que el Principio de Kreisel tiene un contenido muy diferente al Principio de Reflexión, y que el argumento anterior lo muestra. PK1 no es el Principio de Kreisel, pues PK1 tiene el mismo contenido que el Principio de Reflexión, pero el Principio de Kreisel es una afirmación informal sobre consecuencia lógica. Después de todo, PK1 sí es una presuposición sobre la semántica y las interpretaciones consideradas (como afirma Shapiro) pero el Principio de Kreisel es una afirmación sobre consecuencia lógica. Aquí es donde las observaciones hechas en el capítulo 2 tienen gran importancia. Si lo que argumenté en el capítulo 2 es correcto, la conexión entre que el Principio de Kreisel y PK1 recae justamente en que la propiedad de modalidad de la noción preteórica de consecuencia lógica recae en entenderla dentro de la colección de todas las clases, y por lo tanto, de entenderla en V mismo. El capítulo 2 efectivamente muestra cómo una cuestión sobre la *adecuación extensional* de la definición tarskiana de consecuencia se puede reducir a un problema semántico. Con estas observaciones en mente, es que es posible afirmar que el Principio de Kreisel está efectivamente justificado a partir de la justificación del Principio de Reflexión.

Hay otros enfoques de evaluación del Principio de Kreisel para los cuales la consistencia de éste podría ser interesante, algunos de los cuales, ya he mencionado. Rayo y Uzquiano consideran al Principio de Kreisel, al igual que el Principio de Reflexión, como una garantía de que existe una relación entre las nociones de *verdad*, *satisfacción* y *validez* dentro de la teoría de conjuntos. Esto no lo enuncian con claridad, sin embargo, utilizan una versión del argumento de Kreisel para garantizar que esto es el caso para la lógica de primer orden (cf. Rayo y Uzquiano 1999: 317). La idea es que, dado que sabemos que las oraciones que son satisfechas en todos los modelos son universalmente válidas, hay que garantizar que las oraciones universalmente válidas son verdaderas en la teoría de conjuntos. De nuevo, el argumento de Kreisel garantiza esto para la lógica de primer orden, pues si una oración α^1 es tal que $V(\alpha^1)$, sabemos que $Val(\alpha^1)$, lo que garantiza que α^1 es verdadera en el universo de la teoría de conjuntos. Sin embargo, no hay garantía de esto para el caso de la lógica de segundo orden. En este respecto, me parece que lograr una respuesta satisfactoria a esto apelando a la prueba de Koellner es poco probable. La razón principal es que es necesario garantizar que $V(\alpha^2)$ implica $Val(\alpha^2)$, lo cual no hace la prueba de Koellner. A final de cuentas, lo más que se puede afirmar a partir de la consistencia del Principio de Kreisel es que si una oración es universalmente válida entonces es verdadera en el nivel V_δ, pero esto es insuficiente pues la noción de verdad que está en juego aquí es la noción de *verdad en la jerarquía acumulativa.* De nuevo, no hay razones para creer que la jerarquía acumulativa es el nivel V_δ.

La consistencia del Principio de Kreisel también permite dar una respuesta a quiénes consideran a ZFC inadecuada para proporcionar interpretaciones de un lenguaje. Priest (2006: 36) argumenta que ZFC no es adecuada para proporcionar la semántica de la propia ZFC pues es los modelos de una teoría deben tener como dominio un conjunto, sin embargo V (el modelo pretendido de ZFC) no es un conjunto. Priest desecha posibles soluciones como la de que hay que entender como modelos de ZFC a los modelos generados por cardinales inaccesibles, a partir de que de esta manera no se están considerando todas las posibles interpretaciones (pues obviamente no se considera a V como interpretación). En esta misma manera de argumentar, Priest rechaza apelar al teorema de Lowenheim – Skolem (descendente) y considerar sólo los modelos numerables del lenguaje ZFC (en primer orden, por supuesto), pues para poder demostrarlo es necesario cuantificar sobre todas las estructuras (que era precisamente lo que se quería explicar). Incluso si nos reducimos al Principio de Kreisel en primer orden, podemos dar una respuesta a esta clase de objeciones a partir de que el Principio de Kreisel muestra algo más fuerte que el teorema de Lowenheim – Skolem. El Principio de Kreisel garantiza que es posible considerar únicamente las estructuras conjuntistas incluso si se insiste en considerar todas las estructuras (incluyendo aquellas cuyo dominio es una clase).

Además de todo esto, la prueba del teorema 4.1 representa evidencia para aceptar el Principio de Kreisel y, así como mencioné en el caso de la condición de suficiencia, nos señala que algunas preocupaciones sobre el uso de la jerarquía acumulativa como semántica son menos problemáticas de lo que parecen. La principal conclusión es que el uso de la jerarquía acumulativa, así como una definición tarskiana de verdad lógica (en consideración del problema de la noción de consecuencia lógica) no genera inconsistencias. Para los lenguajes de la lógica clásica, tradicionalmente se ha usado la definición tarskiana de consecuencia lógica dentro del marco de la teoría de conjuntos. Una preocupación presente en este contexto, y señalada por ejemplo en (Rayo y Uzquiano 1999: 316) es que la lógica de segundo orden parece comprometernos con la existencia de clases. A partir de la prueba de Koellner, se puede ver que la validez de las fórmulas de segundo orden puede definirse mediante el concepto de conjunto sin necesidad del concepto de clase. Esto parece un punto muy importante a favor de la definición tarskiana de consecuencia lógica y de su semántica definida dentro de la teoría de conjuntos. Finalmente, la naturaleza de la jerarquía acumulativa de la teoría de conjuntos es algo que sigue en investigación, y probablemente la inclusión de nuevos axiomas a la axiomatización usual proporcione más información sobre la verdad de los Principios de Reflexión y, por lo tanto, del Principio de Kreisel.

Es cierto que estas parecen pocas razones para aceptar el Principio de Kreisel, sin embargo, es lo mejor que hay hasta ahora. Después de todo, la prueba de Koellner es la justificación más precisa que se ha ofrecido para los principios de reflexión. De hecho, Koellner muestra que el fenómeno de la reflexión no está necesariamente justificado al mostrar que los principios de reflexión en tercer orden son inconsistentes. Antes de estas pruebas, el principio de reflexión, más que requerir una justificación, se consideraba como un criterio que debía cumplir un axioma para estar justificado a partir del concepto de conjunto. Por ejemplo, Maddy señala que "en cualquier caso, la reflexión es probablemente la regla de oro más universalmente aceptada en la teoría de conjuntos de orden superior" (Maddy, 1988: 503). Y la justificación de este principio como criterio descansaba en la intuición de que el universo conjuntista es indefinible. Las pruebas de Koellner representan el mejor argumento a favor de los principios de reflexión en segundo orden (así como el más desastroso para el caso de los principios en órdenes superiores). La verdad de los principios de reflexión es algo todavía en cuestión, y los avances en las pruebas de su verdad serán avances en la verdad del Principio de Kreisel. De nuevo, el desarrollo de ZFC determinará en qué medida el Principio de Kreisel es verdad.

Para concluir este trabajo simplemente quiero señalar que si bien es cierto que no tenemos la garantía de que el Principio de Kreisel sea verdadero, creo que hay buenas razones para aceptarlo a partir de su consistencia relativa. Después de todo, la que es tal vez la principal preocupación del uso de una definición de consecuencia lógica es que exista la garantía de que ésta sea inadecuada. La prueba de consistencia proporcionada anteriormente muestra que esto no es así, y que el asumir que es adecuada no implicará una inconsistencia. Los posibles contraejemplos a esta afirmación recaerán en asumir ciertas propiedades de la jerarquía acumulativa de la teoría de conjuntos, por ejemplo, que no existe el cardinal de Erdös $\kappa(\omega)$. En este sentido, estos contraejemplos serán controvertidos y requerirán de cierta defensa. Sin embargo, ninguno de estos comentarios se aplica a la noción más general de consecuencia lógica, lo cual es una preocupación importante. De nuevo, hay que señalar que la prueba de Koellner es la justificación más precisa de los principios de reflexión hasta ahora, y una respuesta sobre la satisfacción de PR3 y PR4 dependerá de posibles avances en la justificación de los principios de reflexión. De momento, lo mejor que se puede esperar es un escepticismo sobre la satisfacción del Principio de Kreisel en su versión más general: $\forall i \forall \Gamma \forall \alpha (\mathrm{Val}(\Gamma^i, \alpha^i) \leftrightarrow \mathrm{V}(\Gamma^i, \alpha^i))$. No hay una prueba de la consistencia de este principio, aunque, al igual que como Kreisel señala con respecto a $\forall \alpha (Val(\alpha^2) \leftrightarrow V(\alpha^2))$, se debe esperar una. Hasta ahora, dicha prueba parece exigir un avance mayor en la axiomatización ZFC, por lo cual dejaré esto para algún trabajo posterior.

Apéndice A

Teoría de Conjuntos, ZFC

En este breve apéndice expondré algunos conceptos de la teoría de conjuntos importantes para este trabajo. La teoría que presentaré es la axiomatización estándar de Zermelo-Fraenkel y el axioma de elección (*Axiom of Choice*) (ZFC). Presupondré los conceptos básicos del álgebra de conjuntos como los conceptos de *relación*, *función*, *buen orden*, etcétera. Para mayor detalle de esto véase (Amor y Montaño 2011).

A.1. Axiomas de ZFC

A.1.1. ZFC en primer orden

Está conformado por los siguientes axiomas:

(Axioma de Vacío) $\exists x \forall y (y \in x)$ [x se denota usualmente como $\varnothing$, el conjunto vacío]

(Axioma de Extensionalidad) $\forall x \forall y (\forall z (z \in x \leftrightarrow z \in y) \rightarrow x = y)$

(Axioma de Par) $\forall x \forall y \exists z \forall w (w \in z \leftrightarrow (w = x \vee w = y))$

(Axioma de Unión) $\forall x \exists y \forall z (z \in y \leftrightarrow \exists w (w \in x \wedge z \in w))$ [y se denota como $\bigcup x$]

(Axioma de Potencia) $\forall x \exists y \forall z (z \in y \leftrightarrow \forall w (w \in z \rightarrow w \in x))$ [y se denota como $\wp(x)$]

(Axioma de Infinito) $\exists x[\varnothing \in x \wedge \forall y (y \in x \rightarrow \exists z (z \in x \wedge \forall w (w \in x \leftrightarrow (w \in y \vee w = y))))]$

(Axioma de Regularidad) $\forall x (\varnothing \neq x \rightarrow \exists y (y \in x \wedge \forall w (w \in y \rightarrow w \notin x)))$

(Axioma de Elección) $\forall x((\varnothing \neq x \wedge \varnothing \notin x \forall y \forall z (y \in x \wedge z \in x \wedge y \neq z \rightarrow \forall w (w \in y \rightarrow w \notin z))) \rightarrow \exists u \forall v (v \in x \rightarrow \exists p (p \in u \wedge p \in v \wedge \forall o (o \in u \wedge o \in v \rightarrow o = p))))$ [para todo conjunto no vacío x de conjuntos no vacíos, sin elementos en común, existe un conjunto que contiene a uno y solo un elemento de cada elemento de x]]

Y los siguientes dos esquemas de axioma:

E1 (Esquema de Comprehensión) $\forall x \exists y \forall z(z \in y \leftrightarrow (z \in x \wedge \varphi(z)))$ (para cualquier $\varphi(x)$ fórmula con sólo x variable libre)

E2 (Esquema de Remplazo) $\forall x \forall y \forall z((\varphi(x, y) \wedge \varphi(x, z) \rightarrow y = z) \rightarrow \forall w \exists u \forall v(v \in u \leftrightarrow \exists p(p \in w \wedge \varphi(p, v))))$ para cualquier fórmula $\varphi(x, y)$ con sólo x y y variables libres] [para cualquier función expresada por la fórmula $\varphi(x, y)$, existe el conjunto que es la rango de esa función bajo un conjunto cualquiera w]

A.1.2. ZFC en segundo orden (ZFC2)

Llamaré ZFC2 a la axiomatización de la teoría estándar de conjuntos compuesta por los axiomas 1-8 y los siguientes dos axiomas (en lugar de los esquemas E1 y E2):

(Axioma de Comprensión) $\forall X \forall x \exists y \forall z(z \in y \leftrightarrow (z \in x \wedge X(z)))$

(Axioma de Remplazo) $\forall F \forall x \forall y \forall z((F(x, y) \wedge F(x, z) \rightarrow y = z) \rightarrow \forall w \exists u \forall v(v \in u \leftrightarrow \exists p(p \in w \wedge F(p, v))))$

La diferencia entre la axiomatización de ZFC en primer orden y la de segundo orden es que los esquemas de ZFC se transforman en fórmulas puras de segundo orden. Llamaré Z2 a la conjunción de los axiomas 1-10.

Las operaciones sobre conjuntos usuales se definen de manera usual. Por lo cual, $A \cup B$ (la *unión* de A y B) es el conjunto que tiene únicamente a todos los elementos de A y a todos los elementos de B, $A \cap B$ (la *intersección* de A y B) es el conjunto que tiene a los elementos que pertenecen tanto a A como a B y A–B (la *diferencia* de A y B) es el conjunto que tiene a todos los elementos que pertenecen a A y no pertenecen a B. La *unión* (o *unión generalizada*) de un conjunto A ($\bigcup A$) es el conjunto cuyos elementos son únicamente los elementos de los elementos de A. La *intersección* (o *intersección generalizada*) de A ($\bigcap A$) es el conjunto cuyos elementos son únicamente aquellos que pertenecen a todos los elementos de A.

Llamo *n-ada ordenada* a las secuencias de objetos de longitud n simbolizadas como $\langle x_1, x_2, \ldots, x_n \rangle$, llamaré A^n al conjunto de n-adas ordenadas. Una relación R sobre un conjunto A es un subconjunto de A^n (se usar la expresión $A \times A$ para denotar a A^2). Una función de un conjunto A a un conjunto B ($f : A \rightarrow B$) es una relación entre A y B (un subconjunto del conjunto de pares ordenados $\langle a, b \rangle$, donde $a \in A$ y $b \in B$) tal que cumple las condiciones a y b:

a. $\forall x \in A \exists y \in B \langle x, y \rangle \in f$

b. $\forall x \in A \forall y, z \in B(\langle x, y \rangle \in f \wedge \langle x, z \rangle \in f \rightarrow y = z)$

c. $\forall x, y \in A \forall z \in B(\langle x, z \rangle \in f \wedge \langle y, z \rangle \in f \rightarrow x = y)$ **(inyectividad)**

d. $\forall y \in B \exists x \in A \langle x, y \rangle \in f$ **(sobreyectividad)**

Si una función cumple con la condición c entonces decimos que esa función es *inyectiva* (o *uno a uno*). Si cumple con la condición d, la llamamos *sobreyectiva* (o *suprayectiva*). Si una función cumple con ambas condiciones, decimos que esa función es *biyectiva.* Sean A, B y C conjuntos, $f : A \to B$ y $g : B \to C$ dos funciones, $f \circ g$ es la función definida como $f \circ g = \{\langle a, c\rangle \in A \times C | \exists b \in B(\langle a, b\rangle \in gy\langle b, c\rangle \in f)\}$. A esta función se le conoce como la composición de f y g, y de esta definición se sigue que $f \circ g(x) = g(f(x))$, para todo $x \in A$. El dominio de una función f $(dom(f))$ es el conjunto de elementos de A a los que se aplica la función $(dom(f) = \{x \in A|\exists y \in B\langle x, y\rangle \in f\})$, por la definición de función $dom(f) = A$. El rango de f $(ran(f))$ es el conjunto $\{y \in B|\exists x \in A\langle x, y\rangle \in f\}$.

A.2. Ordinales y cardinalidad

A.2.1. Números ordinales

Definición. *Un número ordinal (denotados por las letras griegas minúsculas α, β, δ , etc.) es un conjunto transitivo y bien ordenado por la pertenencia, es decir,* $\forall x(x \in \alpha \to x \subseteq \alpha)$ *y* $\langle \alpha, \in\rangle$[74] *es un buen orden*

La definición de número ordinal es una generalización de la definición de número natural. La única diferencia entre ambas es que estos últimos deben ser finitos mientras que los números ordinales no necesariamente lo son. La colección de los ordinales determina una relación $<$ de buen orden sobre ellos, definida de la siguiente manera: $\alpha < \beta$ si y solo si $\alpha \in \beta$. Existen dos tipos de números ordinales:

Definición. *Un ordinal α es un ordinal sucesor si y solo si $\alpha = \beta + 1$*

Definición. *Un ordinal α es un ordinal límite si y solo si $\alpha = sup\{\beta : \beta < \alpha\}$[75]$=\bigcup \alpha$, si α no es sucesor*

Es posible definir 'inducción' y probar un teorema de recursión sobre los números ordinales sin mayor apuro (véase Jech 2002: 21-22). A partir del teorema de recursión sobre ordinales se pueden definir las operaciones usuales sobre los ordinales (adición, producto y exponenciación). El límite de la aplicación de dichas operaciones genera un ordinal subindicado con el siguiente ordinal del subíndice anterior. Por ejemplo, llamamos ω al conjunto de todos los ordinales finitos (o números naturales), el primer ordinal transfinito; y el límite de las operaciones sobre ω genera el ordinal ω_1, a partir del cual se genera ω_2, etc. Así mismo, el límite de todos estos ordinales es el ordinal ω_ω. A partir de la paradoja de Burali-Forti (véase Amor, Campero y Miranda 2011: 60), sabemos que la colección de todos los ordinales no es un conjunto y la denotaremos con el término '*Ord*'.

[74]Un buen orden $\langle A, r\rangle$ es una relación sobre un conjunto A tal que para todo subconjunto de A existe un elemento a de dicho subconjunto que cualquier otro elemento b del subconjunto está "por encima de a": $\langle a, b\rangle \in r$. A a se le conoce como el *mínimo* de A.

[75]El supremo de un conjunto $A \subseteq B$ (sup A) es el mínimo elemento de B que está "por encima" de todos los elementos de A, es decir, es el mínimo elemnto b de B tal que para cualquier a de A, $\langle a, b\rangle \in r$.

A.2.2. Cardinalidad

La cardinalidad de un conjunto, intuitivamente, es la cantidad de elementos que tiene un conjunto. Denotaremos la cardinalidad de un conjunto A mediante la expresión '$|A|$', y definimos para dos conjuntos A y B:

$$|A| = |B| \text{ si y solo si existe una función } f \text{ biyectiva entre } A \text{ y } B$$

De igual forma, $|A| \leq |B|$ si y solo si existe una función uno a uno entre A y B. Todos los números naturales tendrán cardinalidad finita y el resto de los ordinales tienen cardinalidad infinita. Por el teorema de Cantor sabemos que, para cualquier conjunto A, el conjunto potencia de A, $\wp(A)$, es tal que $|A| \leq |\wp(A)|$, por lo tanto, $\omega \leq \wp(\omega)$. Definimos para todo $\alpha \in Ord$:

Definición. *α es un número cardinal si y solo si $|\alpha| \neq |\beta|$, para todo $\beta < \alpha$ (denotaré a los números cardinales mediante las letras griegas κ, λ, μ, etcétera)*

Definición. $|A|$= *el mínimo ordinal α, tal que $|A| = |\alpha|$*

Un número cardinal es el mínimo ordinal cuya cardinalidad es mayor a todos los anteriores. La jerarquía de los números cardinales transfinitos (o jerarquía de los $\aleph$'s) está definida de la siguiente manera:

1) $\aleph_0 = \omega_0 = \omega$

2) $\aleph_{\alpha+1} = \omega_{\alpha+1} = \aleph_\alpha^+$

3) $\aleph_\alpha = \omega_\alpha = sup\{\omega_\beta | \beta < \alpha\}$ si α es un ordinal límite

Si la cardinalidad de un conjunto es menor o igual a $\aleph_0$, decimos que dicho conjunto es *contable*; si su cardinalidad es igual a $\aleph_0$, decimos que dicho conjunto es *numerable.*

Definición. *Un número cardinal $\aleph_\alpha$ es un cardinal sucesor si y solo si α es un ordinal sucesor*

Definición. *Un número cardinal $\aleph_\alpha$ es un cardinal límite si y solo si α es un ordinal límite*

Cada conjunto tiene un número cardinal asignado como consecuencia del Axioma de Elección pues éste es equivalente al Teorema del Buen Orden, que garantiza que existe una relación de buen orden para cada conjunto.

Para todo conjunto A, $|\wp(A)| = 2^{|A|}$ (y esta última expresión representa al conjunto de funciones cuyo dominio es A y codominio es 2). Por definición, $|\omega| = \aleph_0$; sin embargo, la cuestión cuál es el número $|\wp(\omega)|$ es independiente de ZFC. Se conoce como la *hipótesis del continuo* a la afirmación $|\wp(\omega)| = \aleph_1$.

A.2.3. Cofinalidad

Sea $\alpha > 0$ un ordinal límite. Sea una β-secuencia $\langle \alpha_\xi | \xi < \beta \rangle$ donde β es un ordinal límite y $\alpha_{xi} \in \alpha$, decimos que β es *cofinal* en α si $\lim_{\xi \to \beta} \alpha_\xi = \alpha$. Un $B \subseteq \alpha$ es cofinal en α si $supB = \alpha$.

Definición. *La cofinalidad de α, $cf(\alpha)$, es el mínimo ordinal β tal que hay β-secuencia $\langle \alpha_\xi | \xi < \beta \rangle$ donde $\lim_{\xi \to \beta} \alpha_\xi = \alpha$*

Para todo $\alpha > 0$ un ordinal límite, $cf(\alpha)$ es un ordinal límite. La cofinalidad de un ordinal α es la mínima "cantidad" de elementos de α que se requieren para "llegar" a α. Por ejemplo, $cf(\aleph_\omega) = \omega$, pues considerando la secuencia $\langle \aleph_0, \aleph_1, \aleph_2, \aleph_3, \dots \rangle$ es una secuencia de "tamaño" ω, mediante la cual se "llega" a $\aleph_\omega$. Para cualquier $\alpha > 0$ un ordinal límite, $cf(\alpha) \leq \alpha$.

Definición. *Un número cardinal infinito $\aleph_\alpha$ es regular si y solo si $cf(\aleph_\alpha) = cf(\omega_\alpha) = \omega_\alpha$*

Definición. *Un número cardinal infinito $\aleph_\alpha$ es singular si y solo si $cf(\aleph_\alpha) = cf(\omega_\alpha) < \omega_\alpha$*

Es decir, un cardinal es regular si su cofinalidad es el ordinal mismo, y es singular si su cofinalidad es estrictamente menor. Es posible demostrar en ZFC que todo cardinal singular es límite y que todo cardinal sucesor es regular. Sin embargo, $\aleph_0$ es un cardinal límite que no es singular (pues es regular, ya que su cofinalidad es el mismo).

A.3. La jerarquía acumulativa de la teoría de conjuntos

La jerarquía acumulativa de la teoría de conjuntos se define por inducción sobre *Ord*:

1) $V_0 = \varnothing$

2) $V_{\alpha+1} = \wp(V_\alpha)$

3) $V_\alpha = \bigcup_{\beta<\alpha} V_\beta$ si α es un ordinal límite

Llamaré V a la colección de todos los conjuntos, donde $V = \bigcup_{\beta \in Ord} V_\beta$. La colección V es una clase propia, no un conjunto, a partir del teorema 2.3 mencionado en el capítulo 2 (este teorema se sigue de la conocida paradoja de Russell). Hay algunas definiciones que hay que mencionar para enunciar un teorema importante para la prueba fundamental del capítulo 4.

Definición. *Sea E una relación sobre una clase P. Para cada $x \in P$, sea $ext_E(x) = \{z \in P | zEx\}$, llamada la extensión de x.*

Definición. *Una relación E sobre una clase P es bien fundada, si:*

a) *cada conjunto* $x \subseteq P \neq \varnothing$ *tiene un elemento E-minimal*

b) $ext_E(x)$ *es un conjunto, para todo* $x \in P$

Definición. *Una relación E sobre una clase P es extensional si para todas* $x, y \in P$, *si* $x \neq y$ *entonces* $ext_E(x) \neq ext_E(y)$.

Una clase M es *extensional* si la relación $\in$ sobre M es extensional. Así mismo, una clase M es *transitiva* si contiene a todos los elementos de sus elementos. V es una clase transitiva y extensional, por ejemplo.

Teorema A.1. ***Teorema del Colapso de Mostowski:***

a) *Si E es una relación bien fundada y extensional sobre una clase P, entonces existe una clase transitiva M y un isomorfismo* π *entre* $\langle P, E\rangle$ *y* $\langle M, \in\rangle$. *La clase transitiva M y el isomorfismo* π *son únicos.*

b) *En particular, toda clase extensional P es isomorfa a una clase transitiva*

c) *En b), si* $T \subseteq P$ *es transitivo,* $\pi(x) = x$, *para todo* $x \in T$

En particular, para cada $x \in P$ se define por inducción $\pi(x) = \{\pi(z) | zEx\}$.

A.4. Pequeños Grandes Cardinales

Existen varios cardinales cuya existencia es independiente de ZFC. Presentaré sólo algunos de ellos.

A.4.1. Cardinales inaccesibles

Definición. *Un cardinal* κ *es* ***inaccesible*** *si y solo si* κ *es un cardinal límite, regular y no contable*

Mencioné que $\aleph_0$ era un cardinal límite regular. Un cardinal inaccesible es aquel cumple con esta propiedad y que no es $\aleph_0$.

Definición. *Un cardinal* κ *es fuertemente inaccesible si y solo si* κ *es inaccesible y para todo* $\lambda < \kappa$, $2^\lambda < \kappa$

Es decir, κ es inaccesible y la cardinalidad de la potencia de cualquier cardinal menor que κ es menor que κ. A partir de estos, existe una jerarquía de cardinales inaccesibles que incluyen a estos y a cardinales hiperinaccesibles.

A.4.2. Cardinales de Mahlo

Definición. *Si X es un conjunto de ordinales y* $\alpha > 0$ *es un ordinal límite entonces* α *es un punto límite de X si* $sup(X \cap \alpha) = \alpha$

Definición. *Sea* κ *un cardinal regular no contable. Un conjunto* $C \subseteq \kappa$ *es cerrado no acotado si C es no acotado en* κ *y contiene a todos sus puntos límite menores que* κ

Definición. *Un conjunto $S \subseteq \kappa$ es estacionario si $(S \cap C) \neq \varnothing$ para todo conjunto $C \subseteq \kappa$ cerrado no acotado*

Definición. *Un cardinal κ es un cardinal **de Mahlo** si y solo si el conjunto de todos los cardinales regulares menores a κ es estacionario*

Al igual que en el caso de los cardinales inaccesibles, existe toda una jerarquía de cardinales de Mahlo.

A.4.3. Cardinales débilmente compactos y cardinales de Erdös

Decimos que una *partición* de un conjunto A es (informalmente) un conjunto de subconjuntos de A, tal que cada elemento de A pertenece a uno y solo un elemento de la partición.

Definición. *Para cada conjunto A y todo $n \in \omega$, $[A]^n = \{X \subseteq A | |X| = n\}$*

Definición. *Si $\{X_i | i \in I\}$ una partición de $[A]^n$, entonces $H \subseteq A$ es homogéneo para la partición, si para algún i, $[H]^n \subseteq X_i$.*

Sean α y β dos números ordinales y sean $n, m \in \omega$, se denota con la expresión '$\alpha \to (\beta)^n_m$'[76], a la afirmación de que toda partición de $[\alpha]^n$ de cardinalidad m tiene un conjunto homogéneo cuyo tipo de orden es β[77]. Una instancia es donde α y β son números cardinales y, si se omite el subíndice m, se asume que su valor es 2.

Teorema A.2. ***Teorema de Ramsey:*** $\omega \to (\omega)^n_m$

Es decir, toda partición de $[\omega]^n$ de cardinalidad m tiene un conjunto homogéneo cuyo tipo de orden es ω.

Definición. *Un cardinal κ es **débilmente compacto** si y solo si es no contable y $\kappa \to (\kappa)^2$*

Todo cardinal débilmente compacto es inaccesible.

Definición. *Un cardinal $\kappa(\alpha)$ es un cardinal **de Erdös** si y solo si es el mínimo cardinal tal que $\kappa \to (\alpha)^{<\omega}$*

En esta definición, la afirmación $\kappa \to (\alpha)^{<\omega}$ es que toda partición del conjunto de subconjuntos finitos de κ de cardinalidad 2 tiene un conjunto homogéneo cuyo tipo de orden es α. Es decir, se están considerando todos los subconjuntos finitos de κ, no se acota a un número particular.

Definición. *Un cardinal κ es un cardinal **de Ramsey** si y solo si cumple con $\kappa \to (\kappa)^{<\omega}$*

---

[76] No debe confundirse este uso del símbolo '$\to$' con el uso de éste en fórmulas lógicas, por ejemplo, el uso en la presentación de los axiomas en la sección 1 de este Apéndice. Generalmente, el uso no será ambiguo.

[77] El tipo de orden de un conjunto X es α si existe una relación de buen orden r sobre X tal que $\langle X, r\rangle$ es isomorfo a $\langle \alpha, \in\rangle$, esto es que exista una función biyectiva $f : X \to \alpha$, tal que si $\langle a, b\rangle \in r$, entonces $f(a) < f(b)$. Si esto es así, se dice que f preserva la estructura y $\langle X, r\rangle$ y $\langle \alpha, \in\rangle$ tienen la misma estructura.

A.5. Teoría de modelos

Para finalizar daré algunos conceptos de la teoría de modelos, importantes para el capítulo 4 principalmente. Al igual que en el capítulo 1, un **modelo** $\mathfrak{A}$ es un par ordenado conformado por un dominio y una función de interpretación. Un modelo de ZFC es una estructura M que consta de un dominio y una relación binaria E (que interpreta a la única constante no lógica del lenguaje, el símbolo de pertenencia, $\in$). Dos modelos son isomorfos si existe un homomorfismo isomórfico entre ellos (véase sección 1.4 del capítulo 1). Si esto es el caso, ambos modelos satisfarán las mismas fórmulas.

Un **submodelo** de $\mathfrak{A}$ es un subconjunto $B \subseteq A$ (donde A es el dominio de $\mathfrak{A}$) junto con las relaciones $P^{\mathfrak{A}} \cap B$ (para cada P letra de relación), las funciones $F^{\mathfrak{A}} \restriction B$ [la restricción de $F^{\mathfrak{A}}$ a B] (para cada F letra de función) y $c^{\mathfrak{A}}$ (para cada c constante de individuo). Todos los $c^{\mathfrak{A}}$ pertenecen a B, y B es cerrado bajo todas las $F^{\mathfrak{A}}$ (si $\langle x_1, \ldots, x_m \rangle \in B^m$, entonces $F^{\mathfrak{A}}(x_1, \ldots, x_m) \in B^m$).

Un submodelo $\mathfrak{A}' \subseteq \mathfrak{A}$ es **elemental** ($\mathfrak{A}' \prec \mathfrak{A}$) si para toda fórmula φ y cualesquiera $a_1, \ldots, a_m$ que pertenezcan al dominio de $\mathfrak{A}'$, $\mathfrak{A}' \vDash \varphi[a_1, \ldots, a_m]$ si y solo si $\mathfrak{A} \vDash \varphi[a_1, \ldots, a_m]$

Una **inmersión** de $\mathfrak{B}$ en $\mathfrak{A}$ es un homomorfismo isomórfico entre el dominio de $\mathfrak{B}$ y un submodelo $\mathfrak{A}' \subseteq \mathfrak{A}$. Una inmersión es **elemental** si $\mathfrak{A}'$ es un submodelo elemental de $\mathfrak{A}$. Sea j una inmersión, llamamos **punto crítico ($\boldsymbol{crit(j)}$)** de j al mínimo ordinal movido por j (es decir, es el mínimo ordinal α, $j(\alpha) \neq \alpha$).

Una función $h : A^n \to A$ es una **función de Skolem** para φ si se cumple que

$$\exists a \in A, \mathfrak{A} \vDash \varphi[a, a_1, \ldots, a_m] \Rightarrow \mathfrak{A} \vDash \varphi[h(a_1, \ldots, a_m), a_1, \ldots, a_m]$$

Para cualesquiera $a_1, \ldots, a_m$. Usando el Axioma de elección es posible construir una función de Skolem para cada fórmula φ. La clausura de un conjunto $X \subseteq A$ bajo todas las funciones de Skolem se conoce como la **cascara de Skolem** (**Skolem Hull**) de X ($Hull(X)$). La cascara de Skolem es un submodelo elemental de un modelo $\mathfrak{A}$, y su cardinalidad es $|X| \bullet |\mathcal{L}| \bullet \aleph_0$, donde $\mathcal{L}$ es el lenguaje considerado. Es decir la cardinalidad de $Hull(X)$ es el menor cardinal entre $|X|$, $|\mathcal{L}|$ y $\aleph_0$.

Ahora bien, en ZFC es posible obtener el siguiente resultado:

Teorema A.3. *Si κ es un cardinal fuertemente inaccesible, entonces V_κ es un modelo de ZFC*

De hecho, en ZFC2 es posible mostrar que un modelo M es un modelo de ZFC2 si y solo si es isomorfo a un modelo $\langle V_\kappa, \in \cap V_\kappa \times V_\kappa \rangle$, donde κ es un cardinal fuertemente inaccesible (esto quiere decir que ZFC2 es *cuasi-categórica*). Usualmente, usaré este hecho implícitamente.

Un modelo particular de ZFC es el universo constructible L, constituido por los conjuntos que son definibles por una fórmula del lenguaje de ZFC, y se construye por inducción transfinita. Sea un conjunto X, X es *definible* en un modelo M si hay una fórmula φ de ZFC y algunos $a_1, \ldots, a_m \in M$ tales que $X = \{x \in M | M \vDash \varphi[a, a_1, \ldots, a_m]\}$. Llamamos $def(M)$ al conjunto $\{X \subseteq M | X \text{ es definible en } M\}$. El universo L se define de la siguiente manera:

1. $L_0 = \varnothing$

2. $L_{\alpha+1} = def(L_\alpha)$

3. $L_\alpha = \bigcup_{\beta<\alpha} L_\beta$, si α es un ordinal límite

4. $L = \bigcup_{\alpha\in Ord} L_\alpha$

Gödel usó el modelo L para demostrar la consistencia de la hipótesis del continuo. Se conoce como **axioma de constructibilidad** a la afirmación $V = L$. La existencia del cardinal de Erdös $\kappa(\omega)$ es consistente con el axioma de constructibilidad así como la existencia de cardinales inaccesibles, de Mahlo y débilmente compactos. Se le llaman grandes grandes cardinales a aquellos cuya existencia no es consistente con el axioma de constructibilidad.

Por último, en la prueba fundamental del capítulo 4 se utiliza el siguiente lema:

Lema A.1. ***(Lema de Silver)*** *Sea $\alpha \geq \omega$ un ordinal límite. Entonces las siguientes afirmaciones son equivalentes:*

1) $\kappa \to (\alpha)^{<\omega}$

2) Para toda estructura M tal que $|\mathcal{L}(M)| = \aleph_0$ y $\kappa \subseteq |M|$, hay un $X \in [\kappa]^\alpha$ que es un conjunto de indiscernibles para M

donde

Definición. *Un conjunto $I \subseteq \kappa$ es un* ***conjunto de indiscernibles*** *para el modelo $\mathfrak{A}$ si para todo $n \in \omega$, y toda fórmula $\varphi(x_1, \ldots, x_n)$*

$$\mathfrak{A} \models \varphi[\alpha_1, \ldots, \alpha_n] \text{ si y solo si } \mathfrak{A} \models \varphi[\beta_1, \ldots, \beta_n]$$

Cuando $\alpha_1 < \ldots < \alpha_n$ y $\beta_1 < \ldots < \beta_n$ son dos secuencias crecientes de elementos de I

Este conjunto de indiscernibles son el conjunto sobre cual se construirá el modelo de los principios de reflexión en el capítulo 4.

A.6. Grandes Grandes Cardinales

Como mencioné, se le conocen como grandes grandes cardinales (large large cardinals) a aquellos cardinales que no son consistentes con el axioma de constructibilidad. Los primeros de ellos son los cardinales *medibles*.

Definición. *Un* ***filtro*** *F sobre un conjunto no vacío S es un subconjunto de la potencia de S ($F \subseteq \wp(S)$) tal que:*

a) $S \in F$ *y* $\varnothing \notin F$

b) Si $X \in F$ *y* $Y \in F$ *entonces* $X \cap Y \in F$

c) Si $X, Y \subseteq S$, $X \in F$ *y* $X \subseteq Y$ *entonces* $Y \in F$

Definición. *Sea $X_0 \subseteq S \neq \varnothing$, el filtro $F = \{X \subseteq S | X_0 \subseteq X\}$ se le conoce como filtro principal.*

Definición. *Un filtro U sobre S es un* ***ultrafiltro*** *si y solo si $\forall X \subseteq S(X \in S \vee X \in S–X)$*

Definición. *Un filtro F sobre S es* ***κ-completo*** *si y solo si para toda familia de conjuntos $\{X_\alpha | \alpha \in \kappa\}$ de subconjuntos de S, si para toda $\alpha \in \kappa$, $X_\alpha \in F$ entonces $\bigcap_{\alpha=0}^{\kappa} X_\alpha \in F$*

Definición. *Un cardinal κ no numerable es un cardinal* ***medible*** *si y solo si existe un ultrafiltro U no principal κ-completo sobre κ.*

Definición. *Un cardinal κ no numerable es un cardinal* ***fuertemente compacto*** *si y solo si para todo conjunto S se cumple que todo filtro κ-completo, se puede extender a un ultrafiltro κ-completo.*

Como mencioné, la existencia de un cardinal medible es inconsistente con el axioma de constructibilidad.

Teorema A.4. ***(Teorema de Scott)*** *Si existe un cardinal medible, entonces $V \neq L$*

La propiedad interesante de los cardinales medibles es que pueden definirse mediante inmersiones elementales en la medida que κ es un cardinal medible si y solo si existe una inmersión elemental $j : V \to M$, tal que $crit(j) = \kappa$. O bien, si existe una inmersión elemental $j : V \to M$, donde M es un modelo transitivo y $crit(j) = \kappa$, entonces $U = \{X \subseteq \kappa | \kappa \in j(X)\}$ es un ultrafiltro no principal κ-completo sobre κ, y de hecho, $M^\kappa \subseteq M$ (donde M^κ es el conjunto de todas las secuencias de elementos del dominio de M de tamaño κ[78]), es decir, M es cerrado bajo todas las secuencias de longitud κ. Entonces, el modelo M es muy "parecido" al universo sin embargo no puede ser V mismo por el teorema de Kunen:

Teorema A.5. ***(Teorema de Kunen)*** *Si $j : V \to M$ es una inmersión elemental no trivial, entonces $M \neq V$*

A partir de esto, se ha buscado formular nociones de cardinales grandes tal que sea posible definir inmersiones elementales de V en algún modelo transitivo M, donde el punto crítico sea un cardinal más grande. Así se definen algunos cardinales grandes grandes como los siguientes:

Definición. *κ es* ***γ-supercompacto*** *si y solo si hay una inmersión $j : V \to M$*

I. *$crit(j) = \kappa$ y $\gamma < j(\kappa)$ y*

II. *$M^\kappa \subseteq M$*

Definición. *κ es* ***supercompacto*** *si y solo si κ es γ-supercompacto para todo $\gamma \geq \kappa$.*[79]

[78]Otra manera de definir M^κ es como el conjunto de funciones cuyo dominio es κ y cuyo rango es un subconjunto de M.

[79]Esta definición se puede encontrar en Kanamori (2003: 298). En Jech (2002: 136) se puede encontrar una definición equivalente en términos de la noción de *medida*.

Definición. κ *es* ***γ-fuerte*** *si y solo si hay una inmersión* $j : V \to M$

i. $crit(j) = \kappa$ *y* $\gamma \geq j(\kappa)$ *y*

ii. $V_{\kappa+\gamma} \subseteq M$

Definición. κ *es* ***fuerte*** *si y solo si* κ *es* γ*-fuerte para todo* γ

Definición. κ *es* ***superfuerte*** *si y solo si hay una inmersión* $j : V \kappa M$

i. $crit(j) = \kappa$ *y* $\gamma \geq j(\kappa)$ *y*

ii. $V_{j(\kappa)} \subseteq M$

Definición. κ *es* ***enorme*** *si y solo si hay una inmersión* $j : V \to M$, $crit(j) = \kappa$ *y* $M^{j(\kappa)} \subseteq M$

Bibliografía

[1] Amor Montaño, J. A. (2011), *Teoría de conjuntos para estudiantes de ciencias*, 2da edición, México: Facultad de Ciencias - UNAM

[2] Amor Montaño, J. A., Campero Arena, G. y Miranda Perea, F.E. (2010), *Teoría de conjuntos. Curso intermedio*, México: Facultad de Ciencias - UNAM

[3] Aristóteles, *Prior Analytics*, trad, intr. y notas de Robin Smith, 1989, Indianapolis: Hackett Publishing Company

[4] Asmus, C. y Restall, G. (2012), "A History of the Consequence Relation" en *Handbook of the History of Logic, vol. 11. Logic: A History of Its Central Concepts*, D. M. Gabbay, F. J. Pelletier y J. Woods (eds.), Amsterdam: North – Holland

[5] Bates, J. (1999), "Etchemendy, Tarski and Logical Consequence", *Southwest Philosophy Review*, 15(1): 47 – 54

[6] Bolzano, B. (1973), *Theory of Science*, Jan Berg (ed. e intr.), Burnham Terrell (trad.), Dordrecht: D. Reidel Publishing Company

[7] Bueno, O. (2010), "A Defense of Second Order Logic", *Axiomathes*, 20(2): 365 – 383

[8] Corcoran, J. (2009), "Aristotle's Demostrative Logic", *History and Philosophy of Logic*, 30(1): 1 – 20

[9] Enderton, H. (2004), *Una introducción matemática a la lógica*, José Alfredo Amor y Montaño (trad.), 2da edición, México: UNAM – IIF

[10] Etchemendy, J. (1990), *The Concept of Logical Consequence*, Cambridge, Massachusetts: Harvard University Press

[11] ______________(2008), "Reflections on Consequence" en *New Essays on Tarski and Philosophy*, D. Patterson (comp.), Oxford: Oxford University Press, págs. 263 – 299

[12] García Carpintero, M. (2003), "Gómez Torrente on Modality and Logical Consequence", *Theoria*, 18(2): 159 – 170

[13] Garson, J. W. (2006), *Modal Logic for Philosophers*, New York: Cambridge University Press

[14] Gómez Torrente, M. (1996), "Tarski on Logical Consequence", *Notre Dame Journal of Formal Logic*, 37(1): 125 – 151

[15] ________________(1998), "On the Fallacy Attributed to Tarski", *History and Philosophy of Logic*, 19(4): 227 – 234

[16] ________________(1998/99), "Logical Truth and Tarskian Logical Truth", *Synthèse*, 117(3): 375 – 408

[17] ________________(2000a), "A Note on Formality and Logical Consequence", *Journal of Philosophical Logic*, 29(5): 529 – 539

[18] ________________(2000b), *Forma y modalidad: Una introducción al concepto de consecuencia lógica*, Buenos Aires: Eudeba

[19] ________________(2002), "The Problem of Logical Constants", *Bulletin of Symbolic Logic*, 8(1): 1 – 37

[20] ________________(2003), "Logical Consequence and Logical Expressions", *Theoria*, 18(2): 131 – 144

[21] ________________(2008a), "Are There Model - Theoretic Logical Truths that Are not Logical Truths" en *New Essays on Tarski and Philosophy*, D. Patterson (comp.), Oxford: Oxford University Press, págs. 340 – 368

[22] ________________(2008b), "Interpretaciones y conjuntos" en *Reflexiones sobre la Paradoja de Orayen*, Adolfo García de la Sienra (comp.), México: UNAM – IFF, págs. 207 - 221

[23] ________________(2009), "Reading Tarski on Logical Consequence", *Review of Symbolic Logic*, 2(2): 249 – 297

[24] Griffiths, O. (2014), "Formal and Informal Consequence", *Thought: A Journal of Philosophy*, 3(1): 9 – 20

[25] Jané, I. (1993), "A Critical Appraisal of Second – Order Logic", *History and Philosophy of Logic*, 14(1): 67 – 86

[26] ______(2005), "Higher-Order Logic Reconsidered" en *The Oxford Handbook of Philosophy of Mathematics and Logic*, S. Shapiro (ed.), New York: Oxford University Press, págs. 781 – 810

[27] Jech, T. (2002), *Set Theory*, 3da edición, Berlin: Springer – Verlag

[28] Joosten, J. J. (2010), "Interpretability Logics and Large Cardinals in Set Theory" en *Estudios de lógica, lenguaje y epistemología -IV Jornadas Ibéricas*, D. Fernández Duque, E. Gómez-Caminero Parejo e I. Hernández Antón, España: Universidad de Sevilla y Fénix Editora, págs. 157 – 176

[29] Kanamori, A. (2003), *The Higher Infinite. Large Cardinals in Set Theory from Their Beginnings*, 2da edición, Berlin: Springer-Verlag

[30] Ketland, J. (1999), “Deflationism and Tarski’s Paradise”, *Mind*, 108(429): 69 – 94

[31] __________(2005), “Deflationism and the Gödel Phenomena: Reply to Tennant”, *Mind*, 114(453): 75 – 88

[32] Kneale, W. y Kneale, M. (1962), *The Development of Logic*, New York: Oxford University Press

[33] Koellner, P. (2009), “On Reflection Principles”, *Annals of Pure and Applied Logic*, 157(2-3): 206 – 217

[34] Kreisel, G. (1967), “Informal Rigour and Completeness Proofs” en *Problems in the Philosophy of Mathematics*, I. Lakatos (ed.), Amsterdam: North - Holland Publishing Company

[35] Levy, A. (1960), “Principles of Reflection in Axiomatic Set Theory”, *Fundamenta Mathematicae*, 49(): 1 – 10

[36] Macfarlane, J. (2000), *What Does It Mean to Say that Logic Is Formal?* (tesis doctoral), Pittsburgh: University of Pittsburgh

[37] Maddy, P. (1988), “Believing the axioms I”, *The Journal of Symbolic Logic*, 53(2): 481 – 511

[38] Mancosu, P. (2010), “Fixed Versus Variable Domain Interpretations of Tarski’s Account of Logical Consequence”, *Philosophy Compass*, 5(9): 745 – 759

[39] Meadows, T. (2013), “What Can a Categoricity Theorem Tell us?”, *The Review of Symbolic Logic*, 6(3): 524 – 544

[40] Mendelson, E. (1987), *Introduction to Mathematical Logic*, 3ra. edición, California: Wadsworth & Brooks/Cole Advanced Books & Software

[41] Nogina, E. (2014), “On the Hierarchy of Reflection Principles in Peano Arithmetic”, descargado de http://arxiv.org/abs/1405.2558, 26 de enero de 2016, 14:18

[42] Priest, G. (2006), *In Contradiction. A Study of the Transconsistent*, 2da. edición, New York: Oxford Claredon Press

[43] Quine, W. V. O. (1982), *Methods of Logic*, 4ta. edición, Cambridge, Massachussets: Harvard University Press

[44] Ray, G. (1996), “Logical Consequence: A Defense of Tarski”, *Journal of Philosophical Logic*, 25(6): 617 – 677

[45] Rayo, A. y Uzquiano, G. (1999), “Toward a Theory of Second – Order Consequence”, *Notre Dame Journal of Formal Logic*, 40(3): 315 – 325

[46] Rossberg, M. (2004), "First Order Logic, Second Order Logic and Completness" en *First Order Logic Revisted*, V. Hendricks, F. Neuhaus, S. A. Pedersen, U. Scheffler y H. Wansing (eds.), Berlin: Logos – Verlag, págs. 302 – 321

[47] Sagi, G. (2014), "Models and Logical Consequence", *Journal of Philosophical Logic*, 43(5): 943 – 964

[48] Shapiro, S. (1987), "Principle of Reflection and Second Order Logic", *Journal of Philosophical Logic*, 16(3): 309 – 333

[49] ———— (1991), *Foundations without Foundationalism: A Case of Second Order Logic*, New York: Oxford University Press

[50] ———— (1998a), "Logical Consequence: Models and Modality" en *The Philosophy of Mathematics Today*, M. Schirn (ed.), Oxford: Oxford University Press, págs. 131 – 156

[51] ———— (1998b), "Proof and Truth: Through Thick and Thin", *The Journal of Philosophy*, 95(10): 493 – 521

[52] Sher, G. (1996), "Did Tarski Commit the Tarski's Fallacy?", *The Journal of Symbolic Logic*, 61(2): 653 – 686

[53] Smith, P. (2011), "Squeezing Arguments", *Analysis*, 71(1): 22 – 30

[54] ———— (2013), *An Introduction to Gödel's Theorems*, Segunda Edición, New York: Cambridge University Press

[55] Tarski, A. (1931), "The Concept of Truth in Formalized Language" en *Logic, Semantics, Metamathematics. Papers from 1923 to 1938* (1956), J. H. Woodger (trad.), Gran Bretaña: Oxford Clarendon Press, págs. 152 – 278

[56] ———— (1936), "On The Concept of Logical Consequence" en *Logic, Semantics, Metamathematics. Papers from 1923 to 1938* (1956), J. H. Woodger (trad.), Gran Bretaña: Oxford Clarendon Press, págs. 409 – 420

[57] Tennant, N. (2005), "Deflationism and the Gödel Phenomena: Reply to Ketland", *Mind*, 114(453): 89 – 96

[58] Van Benthem, J. (1985), "The Variety of Consequence, According to Bolzano", *Studia Logica: An International Journal for Symbolic Logic*, 44(4): 389 – 403