La pesadilla de Berkeley: Una breve introducción al análisis no estándar y a la teoría de modelos

Numa Grinberg
Universidad de Buenos Aires
Concurso de monografías 2025

Esta monografía tiene dos propósitos. El primero es escribir una introducción doble, a dos areas de la matemática que me gustan mucho y siento que no son tan conocidas ni apreciadas como deberían. La segunda es contar una historia, una historia magnífica, sobre como descubrimos, olvidamos y volvimos a descubrir uno de los conceptos más importantes de la matemática.

Índice

1. Introducción: Los infinitesimales y la modernidad

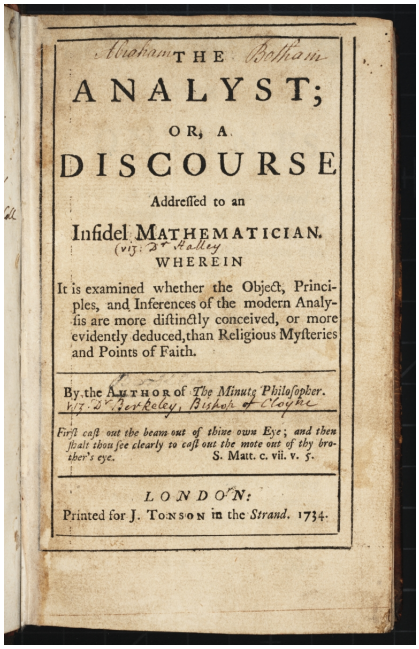
Es evidente que los períodos históricos, como el tiempo mismo, son fenómenos continuos, no discretos. No tienen ni un principio ni un fin determinado. Aun asi, se ha dedicado una gran cantidad de tiempo a debatir que fechas dan inicio o final a periodos historicos. Una de las fechas más discutidas es la que le da el inicio a la modernidad. Las propuestas abundan. La más frecuentes son 1453, año de la caída de Constantinopla, y 1492, año de la llegada de Colón a América. Otros historiadores mencionan 1439, invención de la imprenta de Gutenberg, 1517, el inicio de la reforma protestante o 1648, paz de Westphalia. Se pueden seguir proponiendo más y más fechas, y en un período de tiempo tan turbulento como aquel, acaso cada año tiene una o más razones por las que podria ser el indicado. Pero para esta monografía voy a sostener que tanto la modernidad como la matemática moderna comenzaron en el mismo año: 1687, año en el que Isaac Newton publica el *Principia Mathematica*.

Por supuesto, ni Leibniz ni Newton (que famosamente dijo "*Si he llegado a ver tan lejos es porque estoy parado en hombros de gigantes*") llegaron a estos desarrollos de la nada, sino que fueron la culminación de una larga serie de investigaciones que se pueden trazar hasta principios del renacimiento. Una generación antes, Isaac Barrow, Pierre de Fermat y René Descartes estudiaron extensamente el problema de calcular la recta tangente a una curva definida por ecuaciones algebraicas, mientras que Bonaventura Cavalieri introdujo el uso de métodos infinitesimales (llamado entonces el metodo de los indivisibles)

para calcular areas y volumenes. Pero fue a partir de la obra de ambos que todo ese conocimiento adquirió una forma completa y consistente.

A partir de ese momento, la matemática comenzó a avanzar a un ritmo vertiginoso. Matemáticos como los hermanos Bernoulli, Leonhard Euler y Joseph-Louis Lagrange extendieron y aplicaron los métodos de Newton y Leibniz, desarrollando el análisis en varias variables y sentando las base de la teoría de ecuaciones diferenciales y el calculo de variaciones. Este avance se enmarca en la Revolución Científica y más en general, en la Ilustración, movimiento que exaltaba la razón como la principal, sino única, fuente de autoridad, legitimidad y progreso. Pero todo cambio genera oposición, y pocos críticos de la ilustración fueron tan lúcidos como el obispo anglicano George Berkeley.

Un hombre profundamente devoto y de carácter conservador, Berkeley dedicó gran parte de su vida intelectual a combatir lo que consideraba las peligrosas tendencias seculares y materialistas de su época. En 1710 publicó su *Tratado sobre los principios del conocimiento humano*, un ataque brillante contra los fundamentos mismos del empirismo científico. Años más tarde, en 1732, publicó *Alciphron*, una serie de diálogos criticando a los "librepensadores", en particular a los científicos naturales que consideraban innecesaria la religión para entender el mundo. Poco despues, en 1734, extiende su último libro con un panfleto titulado ."El *Analista*"


Abraham THE Bosham
ANALYST;
OR, A
DISCOURSE
Addreſſed to an
Infidel MATHEMATICIAN.
(viz. Dr Halley)
WHEREIN
It is examined whether the Object, Principles, and Inferences of the modern Analyſis are more diſtinctly conceived, or more evidently deduced, than Religious Myſteries and Points of Faith.
By the AUTHOR of *The Minute Philoſopher*.
viz. Dr Berkeley, Biſhop of Cloyne
Firſt caſt out the beam out of thine own Eye; and then ſhalt thou ſee clearly to caſt out the mote out of thy brother's eye. S. Matt. c. vii. v. 5.
LONDON:
Printed for J. TONSON in the *Strand*. 1734.

."El *Analista*", subtitulado *Ün Discurso Dirigido al Matemático Infiel: Donde es Examinado si los principios e inferencias del analisis moderno son distintamente concebidos, o más evidentemente deducidos, que los misterios religiosos y los fundamentos de la fé*", es una critica a los principios del análisis, en particular al Principia de Newton. Para entender la argumentación de Berkeley, veamos como Newton calculaba la derivada de la función $f(x) = x^2$. El asumia la existencia de una magnitud infinitamente pequeña, un infinitesimal, que denotaba o y realizaba el siguiente cálculo:

$$\frac{(x+o)^2 - x^2}{o} = \frac{x^2 + 2o \cdot x + o^2 - x^2}{o} = \frac{2ox + o^2}{o} = 2x + o = 2x$$

Berkeley sostiene que el concepto de infinitesimal es un absurdo lógico, una contradicción. Su argumentación es la siguiente: un número representa una cantidad y una cantidad es, o bien nula o no nula, pero no hay ninguna otra posibilidad. Si o es una cantidad no nula, es evidente que no es valido descartarla al final de la ecuación. Luego debemos concluir que o es una cantidad nula, pero eso trae los conocidos problemas de dividir por 0: si asumimos que $\frac{o}{o} = 1$, debemos tambien concluir que $2 = \frac{o+o}{o} = \frac{o}{o} = 1$, lo que por supuesto es un absurdo. Berkeley concluye triunfalmente que el cálculo anterior no tiene ningún fundamento lógico, y de una forma muy similar, refuta la definición de la integral como una suma infinita de cantidades infinitesimales.

En ningún momento Berkeley niega que los resultados del análisis sean ciertos y tampoco duda que sus numerosas aplicaciónes a las ciencias naturales sean igualmente validas. Lo que Berkeley quiere demostrar es la matemática, lejos de ser un parangón de la razón pura, requiere tanta fe como la religión, aunque los cientificos naturales insistan en lo contrario. Los infinitesimales no son ni cero ni un numero positivo, pero de alguna forma funcionan. Como exactamente funcionan es un misterio que escapa la comprensión humana, como la naturaleza de la Santísima Trinidad o del destino, y el hecho de que no solo funcionen, si no que nos permitan conocer el mundo no puede ser más que una obra de gracia divina.

No es hasta 1821, casi un siglo despues de la publicación de .[El] Analista", que se desarrolla una respuesta satisfactoria a las objeciones de Berkeley, cuando Augustin-Louis Cauchy, en su *Cours d'Analyse*, define el concepto que a partir de entonces se convierte en la base del análisis: el límite de una sucesión de números reales. De esta forma, podemos definir la derivada de una funcion $f(x)$ como:

$$f'(x) = lim_{h\to\ 0}\ \frac{f(x+h) - f(x)}{h}$$

Es decir, la derivada no era el cociente con un h positivo ni con un h nulo, ambas opciones que habia refutado Berkeley, sino una tercera opción que ni él ni nadie hasta entonces habia considerado. En un espíritu muy similar, Bernharnd Rienmann formalizó el concepto de integral como el límite de sumas finitas, las llamadas sumas de Riemann.

Para principios del siglo XX, ya era claro que el viejo análisis basado en infinitesimales era algo totalmente anticuado y que todo el futuro estaba en el anásis moderno, basado en limites. Algunos matemáticos, como David Hilbert, abrazaron el nuevo formalismo, otros como Henri Poicare se mostraban mas escepticos, defendiendo la importancia de la intuición sobre el rigor formal. En Francia, muy influenciada por las ideas de Poincaré, se seguian dando los cursos introductorios con el enfoque clásico, usando infinitesimales. En 1934, un grupo de jóvenes matemáticos descontentos con la situación (entre los cuales se encontraban nombres célebres como Henri Cartan, André Weil y Jean Leray) decidieron juntarse para escribir un nuevo libro de texto de Analisis, basado en principios puramente rigurosos y deciden firmarlo bajo el pseudónimo colectivo de Nicolas Bourbaki.

El resultado no fue un solo libro, sino el monumental "*Éléments de mathématique*", una enorme serie de monografías, más tarde recopiladas en libros, que reconstruirían gran parte de la matemática conocida hasta la fecha desde fundamentos puramente logicos, empezando solo con los axiomas de teoría de conjuntos y a partir de ahi definiendo todo lo demas, sin en ningún momento apelar a la intuición geometrica o física.

La publicación de los elementos fue el triunfo definitivo del programa formalista de Hilbert y el último clavo en el ataud de los infinitesimales. Encantados por su rigor, claridad y elegancia, los matemáticos adoptaron muy rápidamente el estilo seco y preciso de Bourbaki como estandar. Los infinitesimales habian desaparecido por completo de la práctica matemática y en los escasos momentos en los que que eran recordados era solamente como una mera curiosidad historica. El único futuro que parecían tener era el olvido absoluto.

Eso era lo que parecía hasta que en 1966 Abraham Robinson, un lógico estadounidense, descubrio que herramientas de teoría de modelos permitian construir un cuerpo de números más grande que $\mathbb{R}$, que Robinson denominaría los **hiperreales**, que contiene números infinitesimales y en el cual los cálculos de Newton y Leibniz son totalmente validos. El análisis hecho en base a esta nueva estructura se denominaría **análisis no estandar**, que desde su introducción se convirtió en una area activa de investigación, con aplicaciónes a la teoría de ecuaciones diferenciales, sistemas dinámicos y procesos estocásticos.[2]

Notemos la curiosa ironía de toda la historia: El programa logicista de Hilbert termina de matar a los infinitesimales, que carecían de fundamentos rigurosos, solo para que decadas más tarde uno de los productos del mismo programa logicista, la teoría de modelos, logró revivirlos.

2. Análisis no estándar

Para definir formalmente la construcción de los hiperreales en Robinson, debemos introducir el concepto de un filtro sobre los números naturales.

2.1. Preliminares: Filtros sobre $\mathbb{N}$

Definición 2.1. Un filtro $\mathcal{F}$ es una colección de subconjuntos de $\mathbb{N}$ que cumple la siguientes propiedades:

- $\varnothing \notin \mathcal{F}$.
- Si $U, V \in \mathcal{F}$, entonces $U \cap V \in \mathcal{F}$. *(cerrado por intersección)*
- Si $U \in \mathcal{F}$ y $U \subset V$, entonces $V \in \mathcal{F}$. *(cerrado por superconjuntos)*

Observación 2.2. De la tercera propiedad se sigue que para todo filtro $\mathcal{F}$ vale que $\mathbb{N} \in \mathcal{F}$.

Veamos algunos ejemplos importantes de filtros:

Ejemplo 2.3. La colección $\mathcal{F} = \{\mathbb{N}\}$ es un filtro y es llamado el **filtro trivial**.

Ejemplo 2.4. Sea $n \in \mathbb{N}$. La colección $\mathcal{F} = \{U : n \in U\}$ es un filtro y es llamado el **filtro principal** generado por n.

Ejemplo 2.5. La colección $\mathcal{F} = \{U : U^c \text{ es finito } \}$ es un filtro y es llamado el **filtro cofinito**.

Definición 2.6. Un filtro $\mathcal{U}$ se dice **ultrafiltro** si para todo $A \subset \mathbb{N}$, o bien $A \in \mathcal{U}$ o bien $A^c \in \mathcal{U}$.

Es fácil ver que los filtros principales son siempre ultrafiltros. Por otro lado, el filtro cofinito $\mathcal{F}$ **no** es un ultrafiltro, ya que si $P \subset \mathbb{N}$ es el conjunto de los números pares, ni P ni P^c pertenecen a $\mathcal{F}$.

Definición 2.7. Un filtro $\mathcal{U}$ se dice **maximal** si no existe otro filtro $\mathcal{U}'$ tal que $\mathcal{U} \subset \mathcal{U}'$.

Proposición 2.1. *Un filtro $\mathcal{U}$ es un ultrafiltro si y solo si es maximal.*

Demostración. $\Longrightarrow$) Sea $\mathcal{U}$ un ultrafiltro y supongamos que existe otro filtro $\mathcal{U}'$ tal que $\mathcal{U} \subset \mathcal{U}'$, es decir, existe un subconjunto A de $\mathbb{N}$ tal que $A \in \mathcal{U}'$ pero $A \notin \mathcal{U}$. Como $\mathcal{U}$ es un ultrafiltro y $A \notin \mathcal{U}$, tiene que ser el caso que $A^c \in \mathcal{U}$ y como $\mathcal{U} \subset \mathcal{U}'$, se sigue que $A^c \subset \mathcal{U}'$. Pero como los filtros son cerrados bajo intersección se seguiria que $\varnothing = A \cap A^c \in \mathcal{U}'$, absurdo!. Luego no puede existir tal $\mathcal{U}$

$\Longleftarrow$) Sea $\mathcal{U}$ maximal. Supongamos que $\mathcal{U}$ no es un ultrafiltro, es decir, existe $A \subset \mathbb{N}$ tal que $A \notin U$ y $A^c \notin U$. Sea $\mathcal{U}' = \mathcal{U} \cup \{X \cap A |\ X \in \mathcal{U}\}$. Es fácil corroborar que $\mathcal{U}'$ es un un filtro y que extiende a $\mathcal{U}$, absurdo. □

Proposición 2.2. ***(Lema del ultrafiltro)*** *Todo filtro $\mathcal{F}$ se puede extender a un ultrafiltro $\mathcal{U}$.*

Demostración. Típica demostración por el lema de Zorn. Sea Ω el conjunto de todos los filtros que extienden a $\mathcal{F}$. Sea $\mathcal{F}_i$ una cadena ascendente en Ω. No es dificíl de ver que la union $\bigcup \mathcal{F}_i$ es tambien un filtro, luego se se sigue que toda cadena ascendente en Ω tiene un supremo y luego, por el lema de Zorn, se sigue que Ω tiene un elemento maximal, es decir, existe un filtro maximal $\mathcal{U}$ que extiende a $\mathcal{F}$. Por la proposición 2.7 $\mathcal{U}$ es un ultrafiltro que extiende a $\mathcal{U}$. □

Observación 2.8. La demostración anterior, como todas la que invocan el lema de Zorn, es no constructiva y depende del axioma de elección. Volveremos a esto más tarde cuando definamos los filtros en general.

Corolario 2.3. *Existe un ultrafiltro no principal sobre* $\mathbb{N}$

Demostración. Sea $\mathcal{U}$ el filtro cofinito y sea $\mathcal{U}'$ una extensión maximal de $\mathcal{U}$. Esta claro que $\mathcal{U}'$ es un ultrafiltro y que no es principal. □

Proposición 2.4. *Sea $\mathcal{U}$ un ultrafiltro que contiene a un conjunto finito. Luego $\mathcal{U}$ es un filtro principal.*

Demostración. Por inducción en el tamaño del subconjunto finito. Es claro que si $\mathcal{U}$ contiene a un singleton $\{a\}$ se sigue que $\mathcal{U}$ es el filtro principal generado por a. Para el paso inductivo, supongamos que $\mathcal{U}$ contiene a un conjunto finito A, tal que $|A| = k+1$. Sea $a \in A$. Como $\mathcal{U}$ es un ultrafiltro, o bien $\{a\} \in \mathcal{U}$ o bien $\{a\}^c \in \mathcal{U}$. Si $\{a\} \in \mathcal{U}$ se sigue que $\mathcal{U}$ es principal. Si $\{a\} \in \mathcal{U}$, como los filtros son cerrados por intersección, se sigue que $\{a\}^c \cap A = (A - \{a\}) \in \mathcal{U}$. Pero $|A - \{a\}| = k$, luego por hipótesis inductiva se sigue que $\mathcal{U}$ es principal. □

Proposición 2.5. *Sea $\mathcal{F}$ el filtro cofinito. Luego, para cualquier ultrafiltro no principal $\mathcal{U}$ vale que $\mathcal{F} \subset \mathcal{U}$*

Demostración. Sea A un conjunto cofinito, veamos que $A \in \mathcal{U}$. Como $\mathcal{U}$ es un ultrafiltro, o bien $A \in \mathcal{U}$ o bien $A^c \in \mathcal{U}$. Pero como A^c es finito, sabemos por el teorema anterior que $A^c \in \mathcal{U}$ implica que $\mathcal{U}$ es principal. Como $\mathcal{U}$ no es principal, se sigue que $A \in \mathcal{U}$ □

2.2. Los hiperreales

Fijemos ahora un cierto ultrafiltro no principal $\mathcal{U}$, veremos más tarde que la elección de cual en particular no importa.

Definición 2.9. Definimos la relación $\sim_{\mathcal{U}}$ sobre $\mathbb{R}^{\mathbb{N}}$ como $(a_n) \sim_{\mathcal{U}} (b_n)$ si y solo $\{n : a_n = b_n\} \in \mathcal{U}$.

Proposición 2.6. *La relación $\sim_{\mathcal{U}}$ sobre $\mathbb{R}^{\mathbb{N}}$ es una relación de equivalencia.*

Demostración. Esta claro por la definición que $\sim_{\mathcal{U}}$ es simetrica y reflexiva. Para ver transitividad, sean $(a_n), (b_n), (c_n) \in \mathbb{R}^{\mathbb{N}}$ tal que $(a_n) \sim_{\mathcal{U}} (b_n)$ y $(b_n) \sim_{\mathcal{U}} (c_n)$. Por definición vale que $\{n : a_n = b_n\} \in \mathcal{U}$ y $\{n : b_n = c_n\} \in \mathcal{U}$. Luego, como los filtros son cerrados por intersección, $\{n : a_n = b_n\} \cap \{n : b_n = c_n\} \in \mathcal{U}$. Por otro lado, vale que $(\{n : a_n = b_n\} \cap \{n : b_n = c_n\}) \subset \{n : a_n = c_n\}$ y como los filtros son cerrados por superconjuntos se sigue que $\{n : a_n = c_n\} \in \mathcal{U}$. □

Definición 2.10. Sea $\mathcal{U}$ un ultrafiltro no principal. El conjunto de los **Hiperreales** se define como el conjunto cociente $\mathbb{R}^{\mathbb{N}}/\sim_{\mathcal{U}}$ y se denota $\mathbb{R}^*$.

Definición 2.11. Tenemos una inclusión canonica $i : \mathbb{R} \to \mathbb{R}^*$ dada por $i(a) = [(a, a, a, ...)]$

Teorema 2.7. ***(Los hiperreales son un cuerpo ordenado)*** *Definamos sobre $\mathbb{R}^*$ las operaciones binarias $(+)$ y $(\cdot)$ y la relacion binaria $(<)$ como:*

- $[(a_n)] + [(b_n)] = [(a_n + b_n)]$
- $[(a_n)] \cdot [(b_n)] = [(a_n \cdot b_n)]$
- $[(a_n)] \leq [(b_n)]$ *si y solo si* $\{n : a_n \leq b_n\} \in \mathcal{U}$

Las operaciones anteriores estan bien definidas, y más aun, la estructura $(\mathbb{R}^*, i(0), i(1), +, \cdot, <)$ es un cuerpo ordenado.

Demostración. La demostración completa del teorema es larga, no porque requiera argumentos particularmente complicados, sino unicamente por la gran cantidad de cosas a demostrar.

- (Buena definición). Veamos que la suma esta bien definida. Sean $(a_n), (b_n), (c_n), (d_n) \in \mathbb{R}^{\mathbb{N}}$ tal que $(a_n) \sim_{\mathcal{U}} (c_n)$ y $(b_n) \sim_{\mathcal{U}} (d_n)$. Queremos ver que $(a_n + b_n) \sim_{\mathcal{U}} (c_n + d_n)$. Para eso, notemos que $\{n : a_n + b_n = c_n + d_n\} \supset \{n : a_n = c_n\} \cap \{n : b_n = d_n\}$. Como $\mathcal{U}$ es cerrado por intersección, se sigue que $\{n : a_n = c_n\} \cap \{n : b_n = d_n\} \in \mathcal{U}$ y como $\mathcal{U}$ es cerrado por superconjuntos, se sigue que $\{n : a_n + b_n = c_n + d_n\} \in \mathcal{U}$, luego $(a_n + b_n) \sim_{\mathcal{U}} (c_n + d_n)$. Ver que la multiplicacion y el orden estan bien definidos es análogo.

- (Anillo conmutativo) Ver que cumple los axiomas de un anillo conmutativo es fácil porque los cumple término a termino. Para ver conmutatividad de la suma, basta notar que $[(a_n)] + [(b_n)] = [(a_n + b_n)] = [(b_n + a_n)] = [(b_n)] + [(a_n)]$. Para ver asociatividad de la suma, basta notar que $[(a_n)] + ([(b_n)] + [(c_n]) = [(a_n + (b_n + c_n))] = [((a_n + b_n) + c_n)] = ([(a_n)] + [(b_n)]) + [(c_n)]$. Para ver que el $i(0)$ es neutro basta notar que $[(a_n)] + i(0) = [(a_n + 0)] = [(a_n)]$. El resto de propiedades se verifican de forma similar.

- (Inverso multiplicativo) Para ver que es un cuerpo tenemos que ver que todo elemento no nulo tiene inverso multiplicativo. Sea $(a_n) \in \mathbb{R}^{\mathbb{N}}$ tal que $[(a_n)] \neq i(0)$. Sea $U = \{n : a_n \neq 0\}$. Por definición, $U \in \mathcal{U}$. Sea (b_n) la sucesión de naturales definida de la siguiente manera:

$$b_n = \begin{cases} \frac{1}{a_n} & \text{si } n \in U \\ 0 & \text{si } n \notin U \end{cases}$$

  Se sigue que $[(a_n)] \cdot [(b_n)] = i(1)$ ya que $\{n : a_n \cdot b_n = 1\} = U$ y sabemos que $U \in \mathcal{U}$.

- (Orden lineal) Veamos ahora que $<$ es un orden lineal. Para ver que es transitivo, sean $(a_n), (b_n), (c_n) \in \mathbb{R}^{\mathbb{N}}$ tal que $[(a_n)] < [(b_n)]$ y $[(b_n)] < [(c_n)]$. Por definición del orden, $\{n : a_n < b_n\} \in \mathcal{U}$ y $\{n : b_n < c_n\} \in \mathcal{U}$. Como vale que $\{n : a_n < c_n\} \supset \{n : a_n < b_n\} \cap \{n : b_n < c_n\}$, se sigue que $\{n : a_n < c_n\} \in \mathcal{U}$ y por tanto $[(a_n)] < [(c_n)]$. Para ver que es total, sean $(a_n), (b_n) \in \mathbb{R}^n$. Notemos que $\{n : a_n < b_n\}$, $\{n : a_n = b_n\}$ y $\{n : a_n > b_n\}$ son disjuntos y que $\mathbb{N} = \{n : a_n < b_n\} \cup \{n : a_n = b_n\} \cup \{n : a_n > b_n\}$, luego como $\mathcal{U}$ es un ultrafiltro, uno y solo uno de esos tres conjuntos pertenece a $\mathcal{U}$, es decir, una y solo una de las siguientes vale: $[(a_n)] < [(b_n)]$, $[(a_n)] = [(b_n)]$ o $[(a_n)] > [(b_n)]$.

- (Cuerpo ordenado) Veamos por último que el orden es compatible con las operaciones del cuerpo. Sean $(a_n), (b_n), (c_n)$ tal que $[(a_n)] < [(b_n)]$. Es fácil ver que $[(a_n)] + [(c_n)] = [(a_n + c_n)] < [(b_n + c_n)] = [(b_n)] + [(c_n)]$. Por otro lado, sean $[(a_n)], [(b_n)] > 0$, queremos ver que $[(a_n)] \cdot [(b_n)] > 0$. Tenemos $\{n : a_n \cdot b_n > 0\} \supset \{n : a_n > 0\} \cap \{n : b_n > 0\}$. Por defínción, ambos conjuntos de la derecha pertenecen a $\mathcal{U}$, luego su intersección pertenece a $\mathcal{U}$ y como $\mathcal{U}$ es cerrado por superconjuntos se sigue que $\{n : a_n \cdot b_n > 0\} \in \mathcal{U}$ luego $[(a_n)] \cdot [(b_n)] > 0$.

□

Proposición 2.8. *La inclusión $i : \mathbb{R} \to \mathbb{R}^*$ es un morfismo de cuerpos ordenados.*

Demostración. Basta corroborar que $i(a) + i(b) = [(a)] + [(b)] = [(a + b)] = i(a + b)$, que $i(a) \cdot i(b) = [(a)] \cdot [(b)] = [(a \cdot b)] = i(a \cdot b)$ y que $a < b$ implica que $i(a) < i(b)$. □

(Abuso de notación) Para simplificar las expresiones, a partir de ahora escribiremos $i(a)$ simplemente como a.

La siguientes definiciónes son fundamentales porque nos permite extender toda función con variable real a una con variable hiperreal.

Definición 2.12. Sea $A \subset \mathbb{R}$. Su **extensión** $A^* \subset \mathbb{R}^*$ es el conjunto de todos los hiperreales de forma $[(a_1, a_2, a_3, ...)]$ con $a_i \in A$.

Definición 2.13. Sea $f : A \to \mathbb{R}$. Definimos su extension a los hiperreales $f^* : A^* \to \mathbb{R}^*$ como la función dada por $f^*([(a_n)]) = [f(a_n)]$

Proposición 2.9. *Sean $f, g : \mathbb{R} \to \mathbb{R}$ funciones. Luego vale que:*

- $(f+g)^* = f^* + g^*$
- $(f \cdot g)^* = f^* \cdot g^*$
- $(f \circ g)^* = f^* \circ g^*$

Demostración. Ejercicio para el lector. □

(Abuso de notación) Para simplificar las expresiones, a partir de ahora denotaremos a f^* simplemente como f. En particular, $|\alpha|$ denotara $f^*(\alpha)$ donde $f(x) = |x|$.

Definición 2.14. Un hiperreal $\alpha \in \mathbb{R}^*$ no nulo se dice **infinitesimal** si $|\alpha| < \frac{1}{n}$ para todo $n \in \mathbb{N}$.

Definición 2.15. Un hiperreal $\alpha \in \mathbb{R}^*$ se dice **infinito** si $|\alpha| > n$ para todo $n \in \mathbb{N}$. Un hiperreal se dice **acotado** si no es infinito.

Antes de seguir, veamos algunos ejemplos explicitos de hiperreales:

Ejemplo 2.16. Sea ε el hiperreal dado por $\varepsilon = [(1, \frac{1}{2}, \frac{1}{3}, ...)]$. Luego ε es infinitesimal, ya que para todo k la sucesión es menor a $\frac{1}{k}$ en todos salvo finitos índices, es decir, el $\{n : \varepsilon_n < \frac{1}{k}\}$ es cofinito y por la proposición 2.5 sabemos que el ultrafiltro $\mathcal{U}$ contiene a todos los conjuntos cofinitos.

Ejemplo 2.17. Sea ω el hiperreal dado por $\omega = [(1, 2, 3, ...)]$. Luego ω es infinito.

Ejemplo 2.18. Sea α el hiperreal dado por $\alpha = [(0, 1, 0, 1, ...)]$. Luego $\alpha = 0$ o $\alpha = 1$, dependiendo si el ultrafiltro $\mathcal{U}$ contiene a los números pares o no.

Definición 2.19. Un hiperreal no nulo $\alpha \in \mathbb{R}^*$ se dice **apreciable** si no es infinitesimal ni infinito. Equivalentemente, es apreciable si existen $n, m \in \mathbb{N}$ tal que $\frac{1}{n} < |\alpha| < m$.

Teorema 2.10. *Sean $\varepsilon, \delta \in \mathbb{R}^*$ infinitesimales, $a, b \in \mathbb{R}^*$ apreciables y $\omega \in \mathbb{R}^*$ infinito. Luego:*

- *$\varepsilon + \delta$ es infinitesimal*
- *$\varepsilon + a$ es apreciable.*
- *$\varepsilon + \omega$ es infinito.*
- *$a + b$ es apreciable.*
- *$a + \omega$ es infinito.*

Demostración. Ejercicio para el lector. □

Observación 2.20. Si ω y γ son infinitos, no hay ningún criterio para decir de que tipo sera $\omega + \gamma$. Por ejemplo, si $\omega = [(1, 2, 3, ...)]$ y $\gamma = [(-1, -2, -3, ...)]$ entonces $\omega + \gamma = 0$, pero si $\gamma = [(-1, -4, -9, ...)]$ entonces $\omega + \gamma$ es infinito. Este problema no es más que la famosa indeterminación $\infty - \infty$ en análisis clásico.

Teorema 2.11. *Sean $\varepsilon, \delta \in \mathbb{R}^*$ infinitesimales, $a, b \in \mathbb{R}^*$ apreciables y $\omega \in \mathbb{R}^*$ infinito. Luego:*

- *$\varepsilon \cdot \delta$ es infinitesimal.*
- *$\varepsilon \cdot a$ es infinitesimal.*
- *$a \cdot b$ es apreciable.*
- *$a \cdot \omega$ es infinito.*
- *$\omega \cdot \theta$ es infinito.*

Demostración. Ejercicio para el lector. □

Observación 2.21. Si ω es infinito y ε es infinitesimal, no hay ningún criterio para decir de que tipo sera $\omega \cdot \varepsilon$. Por ejemplo, si $\omega = [(1, 2, 3, ...)]$ y $\varepsilon = [(1, \frac{1}{2}, \frac{1}{3}, ...)]$ entonces $\omega \cdot \varepsilon$ es apreciable, pero si $\varepsilon = [(1, \frac{1}{4}, \frac{1}{9}, ...)]$ entonces $\omega \cdot \varepsilon$ es infinitesimal. Este problema no es más que la famosa indeterminación $\infty \cdot 0$ en análisis clásico.

Teorema 2.12. *Sea $\varepsilon \in \mathbb{R}^*$ infinitesimal, $a \in \mathbb{R}^*$ apreciable y $\omega \in \mathbb{R}^*$ infinito. Luego:*

- $\frac{1}{\varepsilon}$ *es infinito.*
- $\frac{1}{a}$ *es apreciable.*
- $\frac{1}{\omega}$ *es infinitesimal.*

Demostración. Consecuencia directa del teorema anterior. □

Definición 2.22. Sean $\alpha, \beta \in \mathbb{R}^*$. Decimos que α y β estan **infinitamente cerca** (denotado $\alpha \approx \beta$) si $\alpha - \beta$ es infinitesimal o cero.

Proposición 2.13. *$\approx$ es una relación de equivalencia en $\mathbb{R}^*$.*

Demostración. Es evidente que es reflexiva y simetrica. Para ver que es transitiva, supongamos que $\alpha \approx \beta$ y $\beta \approx \gamma$. Luego, $\alpha - \gamma = (\alpha - \beta) + (\beta - \gamma)$ y por hipótesis $(\alpha - \beta)$ y $(\beta - \gamma)$ son infinitesimales o cero, luego por el teorema 2.10 se sigue que $\alpha - \gamma$ es infinitesimal o cero. □

Los siguientes resultados son muy utiles porque relacionan los números hiperreales con las sucesiones que los representan.

Teorema 2.14. *Sea (x_n) una sucesión de reales tal que $lim_{n\to\infty}\, x_n = x$ y sea α el hiperreal dado por $\alpha = [(x_1, x_2, ...)]$. Luego $\alpha \approx x$.*

Demostración. Queremos ver que $\alpha - x$ es cero o un infinitesimal, es decir, que para todo $k \in \mathbb{N}$ vale que $|\alpha - x| < \frac{1}{k}$. Por definición, $\alpha - x = [(x_1 - x, x_2 - x, ...)]$ y como $lim_{n\to\infty}\, x_n = x$ debe existir un n_0 tal que para todo $n > n_0$ vale que $|x - x_n| < \frac{1}{k}$. Luego se sigue que $\{n : |x - x_n| < \frac{1}{k}\} \supset \{n : n > n_0\}$ y como $\{n : n > n_0\} \in \mathcal{U}$ y $\mathcal{U}$ es cerrado por superconjuntos, se sigue que $\{n : |x - x_n| < \frac{1}{k}\} \in \mathcal{U}$ y por tanto $|\alpha - x| < \frac{1}{k}$.

□

Proposición 2.15. *Sean $\alpha, \beta \in \mathbb{R}^*$ tal que $\alpha < \beta$. Luego existen sucesiones $(a_n), (b_n) \in \mathbb{R}^{\mathbb{N}}$ tal que $\alpha = [(a_n)]$, $\beta = [(b_n)]$ y para todo $n \in \mathbb{N}$ vale que $a_n < b_n$.*

Demostración. Sea $(a_n), (c_n)$ sucesiones tal que $\alpha = [(a_n)]$ y $\beta = [(c_n)]$ y definamos la sucesión b_n de la siguiente manera:

$$b_n = \begin{cases} c_n & \text{si } c_n > a_n, \\ a_n + 1 & \text{si } a_n \geq c_n. \end{cases}$$

Es claro que para todo $n \in \mathbb{N}$ vale que $a_n < b_n$. Por otro lado, $\{n : b_n = c_n\} = \{n : a_n < c_n\}$ y como vale que $\alpha < \beta$ sabemos que $\{n : a_n < c_n\} \in \mathcal{U}$, luego $\{n : b_n = c_n\} \in \mathcal{U}$ y por tanto $\beta = [(b_n)]$. □

La siguiente proposición muestra que $\approx$ es compatible con las operaciones de anillo:

Proposición 2.16.

Sean $\alpha, \alpha', \beta, \beta' \in \mathbb{R}^*$ tal que $\alpha \approx \alpha'$ y $\beta \approx \beta'$. Luego:

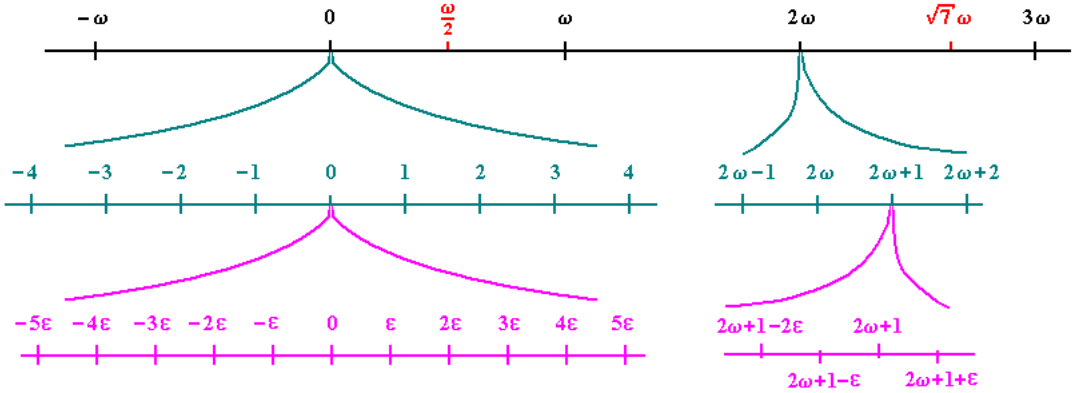
- $\alpha + \beta \approx \alpha' + \beta'$
- *Si α y β son acotados, $\alpha \cdot \beta \approx \alpha' \cdot \beta'$*

Demostración.

- Tenemos $(\alpha+\beta)-(\alpha'+\beta')=(\alpha-\alpha')+(\beta-\beta')$. Por hipótesis $(\alpha-\alpha)$ y $(\beta-\beta)$ son infinitesimales o cero, luego su suma es infinitesimal o cero y por tanto $\alpha+\beta\approx\alpha'+\beta'$.
- Tenemos $(\alpha\cdot\beta-\alpha'\cdot\beta')=(\alpha\cdot\beta-\alpha\cdot\beta')+(\alpha\cdot\beta'-\alpha'\cdot\beta')=\alpha\cdot(\beta-\beta')+\beta'\cdot(\alpha-\alpha')$. Por hipótesis, $\alpha-\alpha'$ y $\beta-\beta'$ son infinitesimales o cero. Ademas, sabemos que α y β' son acotados (esto último porque $\beta'\approx\beta$). Luego, se sigue que $\alpha\cdot(\beta-\beta')+\beta'\cdot(\alpha-\alpha')$ es infinitesimal o cero y por tanto $\alpha\cdot\beta\approx\alpha'\cdot\beta'$.

□

La siguiente imagen ilustra muy bien la estructura de los hiperreales. Tenemos un hiperreal infinitesimal ε y un hiperreal infinito ω. En la recta superior se encuentran hiperreales a distancias infinitas del orden de ω, como ω y 2ω. En las rectas inmediatamente inferiores, se encuentran hiperreales a distancias apreciables, como 1 y 2 o ω y $\omega+1$. En las rectas inferiores a estas tenemos hiperreales a distancias infinitesimales del orden de ε, como $2\omega+1$ y $2\omega+1-\varepsilon$. Por supuesto, los hiperreales no terminan ahi, sino que siguen y siguen. Si quisiera representar distancias infinitesimales del orden de ε^2 deberia ir una recta más abajo, mientras que si quisiera representar distancias infinitas del orden de ω^2 deberia ir una recta más arriba.



Los siguientes teoremas muestra que $\mathbb{R}^*$ tienen un orden extremadamente complejo, mucho más rico que el de los números reales:

Teorema 2.17. *Sea $A\subset\mathbb{R}*$ un conjunto contable. Luego existe un $\theta\in\mathbb{R}*$ tal que$\theta>\alpha$ para todo $\alpha\in A$.*

Demostración. Sea $A=\{\alpha^1,\alpha^2,\alpha^3,..\}$ tal que $\alpha^1<\alpha^2<\alpha^3<....$ Por la proposición 2.15, podemos elegir, para todo k, una sucesión (a_n^k) tal que $\alpha^k=[(a_n^k)]$ tal que para todo n y todo $k>s$ valga que $a_n^k>a_n^s$. Sea c_n la sucesión dada por $c_n=a_n^n$ y sea $\theta=[(c_n)]$. Es fácil que para todo $\alpha_n\in A$ vale que $\theta>\alpha_n$. □

Teorema 2.18. *Sea $A,B\subset\mathbb{R}^*$ subconjuntos contables tal que para todo $\alpha\in A$ y todo $\beta\in B$ vale que $\alpha<\beta$. Luego existe $\theta\in\mathbb{R}^*$ tal que para todo $\alpha\in A$ vale que $\theta>\alpha$ y para todo $\beta\in B$ vale que $\theta<\beta$.*

Demostración. Similar al teorema anterior. Sea $A=\{\alpha^1,\alpha^2,\alpha^3,..\}$ tal que $\alpha^1<\alpha^2<\alpha^3<...$ y sea $B=\{\beta1,\beta2,\beta3,..\}$ tal que $\beta1>\beta2>\beta3>....$ Por la proposición 2.15, podemos elegir, para todo k sucesiones (a_n^k) y (b_n^k) tal que $\alpha^k=[(a_n^k)]$, $\beta^k=[(b_n^k)]$ tal que para todo n y todo $k>s$ valga que $a_n^k>a_n^s$, $b_n^k<b_n^s$ y $a_n^k<b_n^k$. Sea c_n la sucesión dada por $c_n=a_n^n$ y sea $\theta=[(c_n)]$. Es fácil que θ es mayor a todo elemento de A y menor a todo elemento de B. □

Para cerrar esta sección demostraremos que, si asusimos la hipótesis del continuo, la construcción de los hiperreales es independiente de la eleccion de ultrafiltro no principal $\mathcal{U}$. La herramienta principal (y donde radica toda la dificultad de la demostración) es un teorema de Erdos.[3]:

Definición 2.23. Sea $(R,<)$ un orden lineal y κ un cardinal. Decimos que R es κ-saturado si:

- Para todo conjunto A de cardinal menor a κ, existen $c,d\in R$ tal que $d<a<c$ para todo $a\in A$
- Para todo par de conjuntos $A,B\subset R$ de cardinal menor a κ tal que para todo $a\in A$ y $b\in B$ vale que $a<b$, se sigue que existe un $c\in R$ tal que para todo $a\in A$ y $b\in B$ vale que $a<c$ y $c<b$

Teorema 2.19. ***(Erdős-Gillman-Henriksen)*** *Sean E y F dos cuerpos ordenados de cardinal κ tal que sus ordenes son κ-saturados. Luego E y F son isomorfos.*

Teorema 2.20. ***(Unicidad de los hiperreales)*** *Asumiendo la Hipótesis del Continuo, para todo par de ultrafiltros no principales $\mathcal{U}$ y $\mathcal{V}$ los cuerpos $\mathbb{R}^{\mathbb{N}}/\sim_{\mathcal{U}}$ y $\mathbb{R}^{\mathbb{N}}/\sim_{\mathcal{V}}$ son isomorfos.*

Demostración. Sabemos que los cuerpos hiperreales tienen cardinalidad c, luego si vale la hipótesis del continuo se sigue que los hiperreales tienen cardinalidad $\aleph_1$. Por los teoremas 2.17 y 2.18 sabemos que $\mathbb{R}^{\mathbb{N}}/\sim_{\mathcal{U}}$ y $\mathbb{R}^{\mathbb{N}}/\sim_{\mathcal{V}}$ son $\aleph_1$-saturados, luego por el teorema de Erdős-Gillman-Henriksen se sigue que $\mathbb{R}^{\mathbb{N}}/\sim_{\mathcal{U}}$ y $\mathbb{R}^{\mathbb{N}}/\sim_{\mathcal{V}}$ son isomorfos. □

2.3. Un nuevo analísis

"Vi un cielo nuevo y una tierra nueva; porque el primer cielo y la primera tierra pasaron, y el mar ya no existía más"
—Apocalipis 21:1

Ahora que tenemos bien definidos los hiperreales, podemos empezar volver a demostrar todos los resultados clásicos de analisis con infinitesimales. En el resto de esta sección, daremos nuevas definiciónes de los conceptos básicos de analisis (continuidad, continuidad uniforme, diferenciabilidad), demostraremos que son equivalentes a las definiciónes tradicionales y daremos demostraciones de los resultados fundamentales usando las nuevas definiciónes. Estas nuevas demostraciones tiene un caracter muy distinto a las tradicionales: carecen por completen de limites y casi no hay cotas ni sucesiones.

Definición 2.24. Decimos que una función $f : E \to \mathbb{R}$ es **continua** en $x \in E$ si para todo infinitesimal $\varepsilon \in \mathbb{R}^*$ vale que $f(x) \approx f(x+\varepsilon)$. Equivalentemente, es continua en x si para todo $\alpha \in E^*$ tal que $x \approx \alpha$ se cumple $f(x) \approx f(\alpha)$.

Definición 2.25. Decimos que $f : \mathbb{R} \to \mathbb{R}$ es **continua** si es continua en todo $x \in E$.

Ejemplo 2.26. Usemos la definición anterior para probar que $f(x) = x^2$ es continua. Sea ε un infinitesimal. Luego $f(x+\varepsilon) = (x+\varepsilon)^2 = x^2 + 2\varepsilon \cdot x + \varepsilon^2 \approx x^2 = f(x)$.

Ejemplo 2.27. Usemos la definición anterior anterior para demostrar que la función de Heaviside, dada por $H(x) = 1$ si $x > 0$ y $H(x) = 0$ si $x \leq 0$, **no** es continua. Consideremos el infinitesimal $\varepsilon = (1, \frac{1}{2}, \frac{1}{3}, ...)$. Luego $H(0) = 0 \not\approx 1 = H(0+\varepsilon)$

Proposición 2.21. *La definición 2.25 de continuidad es equivalente a la definición clásica de continuidad.*

Demostración. $\Rightarrow$) Sea $f : \mathbb{R} \to \mathbb{R}$ una función continua en x en el sentido clásico, veamos que para todo infinitesimal ε vale que $f(x+\varepsilon) \approx f(x)$, es decir, para todo k vale que $|f(x+\varepsilon) - f(x)| < \frac{1}{k}$.

Sea $\varepsilon = [(a_n)]$. Por la definición clásica de continuidad, existe un m tal que para todo y tal que $|x-y| < \frac{1}{m}$ vale que $|f(x) - f(y)| < \frac{1}{k}$. Como ε es infinitesimal, se sigue que $\{n : |a_n| < \frac{1}{m}\} \in \mathcal{U}$, luego tengo $\{n : |f(x) - f(x+a_n)| < \frac{1}{k}\} \in \mathcal{U}$, luego $|f(x+\varepsilon) - f(x)| < \frac{1}{k}$.

$\Leftarrow$) Supongamos que f no es continua en el sentido clásico. Luego existe una sucesion (x_n) tal que $lim\ x_n = x$ pero tal que para todo n vale que $|f(x) - f(x_n)| > \frac{1}{k}$. Sea s el hiperreal dado por $s = [(x - x_1, x - x_2, ...)]$. Como $lim_n\ x_n = x$ se sigue por la proposición 2.14 que s es un infinitesimal. Por otro lado, como para todo n vale que $|f(x) - f(x_n)| > \frac{1}{k}$ se sigue que $|f(x+s) - f(x)| > \frac{1}{k}$, es decir, $f(x) \not\approx f(x+s)$, luego f no es continua en el sentido de la definición 2.25. □

Proposición 2.22. ***(Álgebra de las funciones continuas)*** *Sean $f, g : \mathbb{R} \to \mathbb{R}$ funciones continuas. Luego:*

- *$f + g$ es continua*

- $f \cdot g$ *es continua*
- $f \circ g$ *es continua*

Demostración. Sea ε un infinitesimal y $x \in \mathbb{R}$.

- $(f+g)(x+\varepsilon) = f(x+\varepsilon) + g(x+\varepsilon)$. Como f y g son continuas, sabemos que $f(x+\varepsilon) \approx f(x)$ y $g(x+\varepsilon) \approx g(x)$. Ademas, por la proposición 2.16, se sigue que $(f+g)(x+\varepsilon) \approx (f+g)(x)$

- $(f \cdot g)(x+\varepsilon) - (f \cdot g)(x) = f(x+\varepsilon) \cdot g(x+\varepsilon) - f(x) \cdot g(x)$. Sumando y restando un término de la forma $f(x) \cdot g(x+\varepsilon)$ nos queda:

  $= f(x+\varepsilon) \cdot g(x+\varepsilon) - f(x) \cdot g(x+\varepsilon) \; + \; f(x) \cdot g(x+\varepsilon) - f(x)g(x)$

  Reordenando:

  $= g(x+\varepsilon) \cdot (f(x+\varepsilon) - f(x)) \; + \; f(x) \cdot (g(x+\varepsilon) - g(x))$

  Como f y g son continuas, vale que $(f(x+\varepsilon) - f(x)) \approx 0$ y $(g(x+\varepsilon) - g(x)) \approx 0$ y por tanto $(f \cdot g)(x+\varepsilon) - (f \cdot g)(x) \approx 0$.

- $(f \circ g)(x+\varepsilon) = f(g(x+\varepsilon)) = f(g(x) + [g(x+\varepsilon) - g(x)])$. Ahora, como g es continua sabemos que $g(x+\varepsilon) - g(x)$ es infinitesimal, y como f es continua vale que $f(g(x)+\varepsilon) \approx f(g(x))$ para todo infinitesimal ε, luego en particular vale que $f(g(x) + [g(x+\varepsilon) - g(x)]) \approx f(g(x))$

□

Definición 2.28. Un conjunto $K \subset \mathbb{R}$ se dice **compacto** si para todo $\alpha \in K^*$ existe $a \in K$ tal que $\alpha \approx a$.

Proposición 2.23. *Sea $K \subset \mathbb{R}$ un compacto y $f : K \to \mathbb{R}$ una funcion continua. Luego f alcanza un máximo y un mínimo en K.*

Demostración. Empecemos viendo que la función esta acotada. Supongamos que no. Se sigue para todo $n \in \mathbb{N}$ existe un $a_n \in K$ tal que $|f(a_n)| > n$. Sea $\alpha \in K^*$ dado por $\alpha = (a_1, a_2, a_3, ...)$. Es fácil ver que $f(\alpha)$ debe ser infinito. Por otro lado, como K es compacto, debe haber un $a \in K$ tal que $\alpha \approx a$ y como f es continua, se sigue que $f(a) \approx f(\alpha)$. Pero entonces $f(a)$ debe ser infinito, lo que es absurdo porque $f(a) \in \mathbb{R}$. Luego la funcion debe estar acotada.

Veamos ahora que el máximo de f efectivamente se realiza (el argumento para ver que el mínimo de f se realiza es dual). Como f esta acotada, debe existir una sucesion $(b_n) \in E$ tal que para todo $x \in K$ vale que $f(x) < f(b_n) + \frac{1}{n}$. Sea β el hiperreal dado por $[(b_n)]$. Por compacidad, $\beta \approx b \in K$ y por la definición de b_n vale que para todo $k \in \mathbb{N}$ y todo $x \in K$ vale que $f(x) < f(b) + \frac{1}{k}$. Se sigue que para todo $x \in K$ vale que $f(x) \leq f(b)$, es decir, la función f alcanza un máximo en b. □

Proposición 2.24. *Sea $f : \mathbb{R} \to \mathbb{R}$ una función continua y $K \subset \mathbb{R}$ un compacto. Luego $f(K) \subset \mathbb{R}$ es un compacto.*

Demostración. Sea $\alpha \in f(K)^*$. Existe un $\beta \in K^*$ tal que $\alpha = f(\beta)$. Ahora, como K es compacto existe un $b \in K$ tal que $\beta \approx b$ y como f es continua se sigue que $\alpha \approx f(b) \in f(K)$ y por tanto, $f(K)$ es un compacto. □

Definición 2.29. Una función $f : E \to \mathbb{R}$ es **uniformemente continua** si para todo $\alpha \in E^*$ y todo infinitesimal ε vale que $f(\alpha + \varepsilon) \approx f(\alpha)$. Equivalentemente, es uniformemente continua si para todo $\alpha, \beta \in \mathbb{R}^*$ tal que $a \approx \beta$ vale que $f(\alpha) \approx f(\beta)$.

Ejemplo 2.30. Usemos la definición anterior para ver que $f : \mathbb{R} \to \mathbb{R}$ dada por $f(x) = x^2$ **no** es uniformemente continua. Sea $\omega = [(1, 2, 3, ..)]$ y $\varepsilon = [(1, \frac{1}{2}, \frac{1}{3}, ...)]$, un infinitesimal. Por definición, $f(\omega) = [(1, 4, 9, ...)]$ pero como $(n + \frac{1}{n})^2 = n^2 + 2 + \frac{1}{n^2}$ es fácil ver que $f(\omega + \epsilon) \approx [(3, 6, 11, ..)]$, luego se sigue que $f(\omega) \not\approx f(\omega + \varepsilon)$

Proposición 2.25. *La definición anterior de continuidad uniforme y la definición clásica son equivalentes.*

Demostración. $\Rightarrow$) Sea $f : \mathbb{R} \to \mathbb{R}$ una función uniformemente continua en el sentido clásico, veamos que para todo infinitesimal ε vale que $f(\alpha+\varepsilon) \approx f(\alpha)$, es decir, para todo k vale que $|f(x+\varepsilon)-f(x)| < \frac{1}{k}$.

Sea α un hiperreal cualquiera y $\varepsilon = [(s_n)]$ un infinitesimal. Por la definición clásica de continuidad, existe un m tal que para todo y tal que $|x-y| < \frac{1}{m}$ vale que $|f(x) - f(y)| < \frac{1}{k}$. Como ε es infinitesimal, se sigue que $\{n : |s_n| < \frac{1}{m}\} \in \mathcal{U}$, luego tengo $\{n : |f(a_n + s_n) - f(a_n)| <\} \in \mathcal{U}$ luego para todo k vale que $|f(\alpha + \varepsilon) - f(\alpha)| < \frac{1}{k}$, es decir, vale que $f(\alpha + \varepsilon) \approx f(\alpha)$.

$\Leftarrow$) Supongamos que f no es uniformemente continua en el sentido clásico, veamos que existen hipereales α, β tal que $\alpha \approx \beta$ pero $f(\alpha) \not\approx f(\beta)$. Como f no es uniformemente continua, existe un $k \in \mathbb{N}$ y sucesiones $(a_n), (b_n)$ tal que para todo n vale que $|a_n - b_n| < \frac{1}{n}$ pero $|f(a_n) - f(b_n)| > \frac{1}{k}$. Sean $\alpha = [(a_n)]$ y $\beta = [(b_n)]$. Es fácil ver que $\alpha \approx \beta$ pero $|f(\alpha) - f(\beta)| > \frac{1}{k}$. □

Una gran ventaja de las definiciónes no estándar de continuidad y continuidad uniforme es que la analogía entre ambas es mucho más clara que con la definición tradicional. Por ejemplo, podemos probar el análogo de la proposición 2.22 sobre continuidad uniforme usando la misma demostracion pero con un $\alpha \in \mathbb{R}^*$ en lugar de $x \in \mathbb{R}$.

Proposición 2.26. ***(Álgebra de las funciones uniformemente continuas)*** *Sean $f, g : \mathbb{R} \to \mathbb{R}$ funciones uniformemente continuas. Luego:*

- *$f + g$ es uniformemente continua*
- *$f \cdot g$ es uniformemente continua*
- *$f \circ g$ es uniformemente continua*

Demostración. Sea ε un infinitesimal y $\alpha \in \mathbb{R}^*$.

- $(f + g)(\alpha + \varepsilon) = f(\alpha + \varepsilon) + g(\alpha + \varepsilon)$. Como f y g son uniformemente continuas, sabemos que $f(\alpha+\varepsilon) \approx f(\alpha)$ y $g(\alpha+\varepsilon) \approx g(\alpha)$. Además, por la proposición 2.16, se sigue que $(f+g)(\alpha+\varepsilon) \approx (f + g)(\alpha)$

- $(f \cdot g)(\alpha+\varepsilon) - (f \cdot g)(\alpha) = f(\alpha+\varepsilon) \cdot g(\alpha+\varepsilon) - f(\alpha) \cdot g(\alpha)$. Sumando y restando un término de la forma nos queda:

  $= f(\alpha+\varepsilon) \cdot g(\alpha+\varepsilon) - f(\alpha) \cdot g(\alpha+\varepsilon) \; + \; f(\alpha) \cdot g(\alpha+\varepsilon) - f(\alpha)g(\alpha)$

  Reordenando:

  $= g(\alpha+\varepsilon) \cdot (f(\alpha+\varepsilon) - f(\alpha)) \; + \; f(\alpha) \cdot (g(\alpha+\varepsilon) - g(\alpha))$

  Como f y g son uniformemente continuas, vale que $(f(\alpha+\varepsilon) - f(\alpha)) \approx 0$ y $(g(\alpha+\varepsilon) - g(\alpha)) \approx 0$ y por tanto $(f \cdot g)(\alpha+\varepsilon) - (f \cdot g)(\alpha) \approx 0$

- $(f \circ g)(\alpha + \varepsilon) = f(g(\alpha + \varepsilon)) = f(g(\alpha) + [g(\alpha + \varepsilon) - g(\alpha)])$. Ahora, como g es uniformemente continua sabemos que $g(\alpha+\varepsilon) - g(\alpha)$ es infinitesimal, y como f es uniformemente continua vale que $f(g(\alpha) + [g(\alpha+\varepsilon) - g(\alpha)]) \approx f(g(\alpha))$

□

La demostración del siguiente teorema es bastante más fácil con la nueva defínción de continuidad uniforme:

Teorema 2.27. ***(Heine)*** *Sea $E \subset \mathbb{R}$ un conjunto compacto y $f : E \to \mathbb{R}$ una función continua. Luego, f es uniformemente continua.*

Demostración. Sea $\alpha \in E^*$ y sea ε un infinitesimal. Queremos ver que $f(\alpha) \approx f(\alpha+\varepsilon)$. Por compacidad de A, existe un $a \in A$ tal que $\alpha \approx a$. Observemos que como $\approx$ es transitiva, $a \approx \alpha + \varepsilon$.

Como f es continua en a, vale que $f(a) \approx f(\alpha)$ y que $f(a) \approx f(\alpha + \varepsilon)$. Como $\approx$ es transitiva, se sigue que $f(\alpha) \approx f(\alpha + \varepsilon)$. □

Definición 2.31. Sea $f : E \subset \mathbb{R} \to \mathbb{R}$. Decimos que f tiene derivada $c \in \mathbb{R}$ en x si para todo infinitesimal ε vale que:

$$\frac{f(x+\varepsilon) - f(x)}{\epsilon} \approx c \tag{1}$$

Observación 2.32. Otra forma de escribir la definición anterior es que la derivada de $f(x)$ en x es $c \in \mathbb{R}$ si para todo infinitesimal ε existe otro infinitesimal δ tal que:

$$f(x+\varepsilon) = f(x) + \ c \cdot \varepsilon + \ \varepsilon \cdot \delta$$

Definición 2.33. Decimos que una función $f : E \to \mathbb{R}$ es **diferenciable** si es derivable en todo punto. Denotamos su derivada en x como $f'(x)$.

Ejemplo 2.34. Usemos la definición anterior para calcular la derivada de la función $f(x) = \frac{1}{x}$. Sea $x \neq 0$ y $\varepsilon \in \mathbb{R}^*$ un infinitesimal no nulo. Luego tenemos:

$$\frac{1}{\varepsilon} \cdot (\frac{1}{x+\varepsilon} - \frac{1}{x}) = \frac{1}{\varepsilon} \cdot (\frac{x}{x \cdot (x+\varepsilon)} - \frac{x+\varepsilon}{x \cdot (x+\varepsilon)}) = \frac{-1}{x \cdot (x+\varepsilon)} \approx \frac{-1}{x^2}$$

Proposición 2.28. *Sea $f : \mathbb{R} \to \mathbb{R}$ una función diferenciable en x. Luego f es continua en x.*

Demostración. Supongamos que la derivada de f en x es $c \in \mathbb{R}$. Sea ε un infinitesimal. Tenemos que $\frac{f(x+\varepsilon)-f(x)}{\varepsilon} \approx c$. Como ε es infinitesimal, sabemos que $\frac{1}{\varepsilon}$ es infinito, luego por la proposición 2.11 sabemos que si $f(x+\varepsilon)-f(x)$ fuera infinito o apreciable, $\frac{1}{\varepsilon} \cdot (f(x+\varepsilon)-f(x))$ seria infinito, pero sabemos que no es el caso. Luego se sigue que $f(x+\varepsilon)-f(x)$ es cero o infinitesimal, es decir, $f(x) \approx f(x+\varepsilon)$. □

Proposición 2.29. ***(Álgebra de las funciones diferenciables)*** *Sean $f, g : E \to \mathbb{R}$ funciones diferenciables. Valen las siguientes igualdades:*

- $(f+g)' = f' + g'$ *(Linearidad de la derivada)*
- $(f \cdot g)' = f' \cdot g \ + \ f \cdot g'$ *(Regla del producto)*
- $(f \circ g)' = (f' \circ g) \cdot g'$ *(Regla de la cadena)*

Demostración. Sea $\varepsilon \in \mathbb{R}^*$ un infinitesimal no nulo.

1. Por definición tenemos que:

$$\frac{(f+g)(x+\varepsilon) - (f+g)(x)}{\epsilon} = \frac{f(x+\varepsilon) - f(x)}{\epsilon} + \frac{g(x+\varepsilon) - g(x)}{\epsilon} \approx f'(x) + g'(x)$$

2. Por definición tenemos que:

$$\frac{(f \cdot g)(x+\varepsilon) - (f \cdot g)(x)}{\varepsilon} = \frac{f(x+\varepsilon) \cdot g(x+\varepsilon) - f(x) \cdot g(x)}{\varepsilon}$$

   Sumando y restando un término de la forma $f(x+\varepsilon)g(x)$:

$$= \frac{f(x+\varepsilon) \cdot g(x+\varepsilon) \ - \ f(x+\varepsilon) \cdot g(x) + f(x+\varepsilon) \cdot g(x) - \ f(x) \cdot g(x)}{\varepsilon}$$

   Reordenando terminos y usando que $f(x)$ es continua:

$$= f(x+\varepsilon) \cdot \frac{g(x+\varepsilon) - g(x)}{\varepsilon} \ + g(x) \cdot \frac{f(x+\varepsilon) - f(x)}{\varepsilon} \approx f(x) \cdot g'(x) + f'(x) \cdot g(x)$$

3. Por definción tenemos que:

$$\frac{f(g(x+\varepsilon))-f(g(x))}{\varepsilon} = \frac{f(g(x+\varepsilon))-f(g(x))}{g(x+\varepsilon)-g(x)} \cdot \frac{g(x+\varepsilon)-g(x)}{\varepsilon}$$

Por definción, el segundo término es $g\prime(x)$. Por otro lado, como g es continua, $g(x+\varepsilon)-g(x)$ es infinitesimal y por tanto el primer término de la multiplicación es $f'(g(x))$. Luego, por la proposición 2.11 llegamos a que:

$$\frac{f(g(x+\varepsilon))-f(g(x))}{\varepsilon} \approx f'(g(x)) \cdot g'(x)$$

□

La siguiente proposición me parece importante porque da una definición muy simple de una hipótesis que aparece frecuentemente en análisis: que una funcion sea derivable y la derivada sea continua alrededor de un punto.

Proposición 2.30. *Sea $f : \mathbb{R} \to \mathbb{R}$ una funcion derivable y $x \in \mathbb{R}$. Luego, son equivalentes:*

- *f' es continua alrededor de x*
- *Para todo par de infinitesimales distintos ε y δ vale que:*

$$\frac{f(x+\varepsilon)-f(x+\delta)}{\epsilon-\delta} \approx f'(x)$$

Demostración. $\Rightarrow$) Tenemos:

$$\frac{f(x+\varepsilon)-f(x+\delta)}{\epsilon-\delta} = \frac{f(x+\delta+(\varepsilon-\delta))-f(x+\delta)}{\epsilon-\delta} \approx f'(x+\delta)$$

Si f' es continua en x, tenemos que $f'(x+\delta) \approx f'(x)$ y por tanto $\frac{f(x+\varepsilon)-f(x+\delta)}{\epsilon-\delta} \approx f'(x)$.

$\Leftarrow$) Supongamos que f' no es continua en x, es decir, existe un infinitesimal ε tal que $f'(x) \not\approx f'(x+\varepsilon)$. Luego tenemos que:

$$\frac{f(x+2\varepsilon)-f(x+\varepsilon)}{2\varepsilon-\varepsilon} \approx f'(x+\varepsilon) \not\approx f'(x)$$

□

Mas alla de $\mathbb{R}$ El desarrollo que hicimos aqui, de la teoría clásica de funciones de una sola variable, no es más que arañar la superficie de lo que se puede hacer usando infinitesimales. Para empezar, aunque definimos derivadas, no tocamos el otro gran invento de Newton y Leibniz, la integral. Hay dos maneras de definirla, la más sencilla formaliza el concepto de sumas infinitas de infinitesimales y da una teoría similar a la de las integrales de Rienmann: Tosca y limitada. La otra es con la llamada medida de Loeb, y da una teoría equivalente a la integral de Lebesgue, mucho más elegante y general.

Pero no hay razón por la que parar en $\mathbb{R}$. Si en vez de trabajar con $\mathbb{R}^{\mathbb{N}}/\sim_{\mathcal{U}}$ hubieramos trabajado con $(\mathbb{R}^k)^{\mathbb{N}}/\sim_{\mathcal{U}}$ podriamos demostrar los resultados clásicos de análisis en varias variables, más en general, si hubieramos trabajado $X^{\mathbb{N}}/\sim_{\mathcal{U}}$ para algun espacio de Banach X, podriamos demostrar los resultados espacios de dimensión infinita. Cuando empece esta monografia pensaba hablar de todo esto, pero el tiempo, a diferencia de los hiperreales, es abrumadoramente acotado.

3. Teoría de modelos

3.1. El arte de dar el mismo nombre a diferentes cosas

Definir que es la matemática no es una tarea facil. De hecho, hay una pagina de Wikipedia dedicada exclusivamente al respecto. La mayoria de las definciones que encontramos alli son anticuadas, como la de Aristoteles "*La matemática es la ciencia de la cantidad*", o simplemente poco iluminadoras, como la de Benjamin Pierce "*La matemática es la ciencia que llega a conclusiones necesarias*". Otras dejan a uno atónito, como la de Charles Darwin *Ün matemático es un hombre ciego en un cuarto oscuro buscando a un gato negro que no esta ahí*"

Pero de todas las definciones, mi favorita es la de Henri Poincare, que definio a la matemática como .*El arte de dar el mismo nombre a cosas diferentes*". Esta claro que el principio al que hace referencia Poincare, la abstracción, esta en el corazón mismo de la matemática, desde aquel lejano dia en el paleolitico en el que algun hombre o mujer penso por primera vez en el concepto de numero hasta los desarrollos más recientes en geometria algebraica o análisis funcional. Pero en mi opinion, son exactamente dos areas de la matemática las que más monstruosamente dilataron el sentido de la frase de Poincare: La teoría de categorías y la teoría de modelos.

Definir que es la teoría de modelos tampoco es una tarea facil. Parafraseando a Hodge, la definiría como el estudio de la relación entre la sintaxis y la semántica, entre las estructuras y las oraciones lógicas que las describen. Uno podria pensar que un proyecto asi es futíl, que no hay ningún resultado interesante a ese nivel de generalidad. Pero uno estaria completamente equivocado: hay muchisimos resultados interesantes, con numerosas aplicaciónes a diversas areas de la matemática.

En esta sección definiremos los conceptos basicos de teoría de modelos, empezando con la parte semántica, dando una definición muy general del concepto de estructura. El que ya esta familiarizado preferirá pasar directamente a la siguiente sección.

Definición 3.1. Un lenguaje $\mathcal{L}$ consiste en:

- Un conjunto $\mathcal{L}_c$ que llamaremos el conjunto de **símbolos de constante**.
- Para todo $k > 1$, un conjunto $\mathcal{L}_f^k$ que llamaremos el conjunto de **símbolos de función** de aridad k.
- Para todo $k > 1$, un conjunto $\mathcal{L}_r^k$ que llamaremos el conjunto de **símbolos de relación** de aridad k.

Definición 3.2. Sea $\mathcal{L}$ un lenguaje. Una **$\mathcal{L}$-estructura** consiste en un conjunto M (que llamaremos el dominio) equipado con lo siguiente:

- Un elemento $c^M \in M$ por cada símbolo de constante $c \in \mathcal{L}_c$
- Una funcion $f^M : M^k \to M$ por cada función $f \in \mathcal{L}_f$ de aridad k
- Una relación $R^M \subset M^K$ por cada relación $R \in \mathcal{L}_r$ de aridad k

Observación 3.3. Como un elemento destacado de M no es más que una función $f : \{*\} \to M$ y para todo conjunto $M^0 = \{*\}$, algunos autores deciden ser más económicos y no definir los símbolos de constante, en su lugar tratandolos como funciones de aridad 0. Nosotros decidimos seguir el *approach* clásico para evitar confusiones.

Ejemplo 3.4. Si $\mathcal{L}$ es el lenguaje con una única función binaria, las $\mathcal{L}$-estructuras son conjuntos equipados con una relación binaria, es decir, **magmas**. En particular, esta clase engloba a todos los grupos.

Ejemplo 3.5. Si $\mathcal{L}$ es el lenguaje con un único símbolo de constante, las $\mathcal{L}$-estructuras son conjuntos con un elemento destacado, es decir, **conjuntos punteados**.

Definición 3.6. Sea $\mathcal{L}$ un lenguaje y M una $\mathcal{L}$-estructura. Un subconjunto $N \subset M$ es llamado una **subestructura** si para todo símbolo de constante $c \in \mathcal{L}$, vale que $c^M \in N$ y para todo símbolo de función $f \in M$ vale que N es cerrado por f^M.

Definición 3.7. Sea $\mathcal{L}$ un lenguaje y sean M y N dos $\mathcal{L}$-estructuras. Una funcion $\pi : M \to N$ es un **morfismo** de $\mathcal{L}$-estructuras si:

- Para toda constante $c \in \mathcal{L}$, vale que $\pi(c^M) = c^N$.
- Para toda funcion $f \in \mathcal{L}$ de aridad k y toda tupla $\bar{m} \in M^k$, vale que $\pi(f^M(\bar{m})) = f^N(\pi(\bar{m}))$.
- Para toda relacion $R \in \mathcal{L}$ de aridad k y toda tupla $\bar{m} \in M^k$, vale que $M \models R^M(\bar{m})$ implica $N \models R^N(\pi(\bar{n}))$.

Observación 3.8. La definición anterior generaliza las nociones de morfismo entre estructuras algebraicas que conocemos, como el morfismo de grupos y de anillos.

Definición 3.9. Sea $\mathcal{L}$ un lenguaje. Dos $\mathcal{L}$-estructuras M y N son **isomorfas** (denotado $M \cong N$) si existen morfismos de $\mathcal{L}$-estructuras $\pi : M \to N$ y $\theta : N \to M$ tal que $\pi \circ \theta = id_M$ y $\theta \circ \pi = id_N$.

Al igual que con todas las estructuras algebraicas que conocemos, se puede definir el producto de cualquier familia de $\mathcal{L}$-estructuras.

Definición 3.10. Sea $\mathcal{L}$ un lenguaje, I un conjunto, $(M_i)_{i\in I}$ una colección de $\mathcal{L}$-estructuras y Definimos su **producto** como la $\mathcal{L}$-estructura $\bar{M}$ con dominio $\prod_{i\in I} M_i$ dada por

- Para cada símbolo de constante $c \in \mathcal{L}$, definimos $c^{\bar{M}} = (c^{M_i})_i$
- Para cada símbolo de función $f \in \mathcal{L}$, definimos $f^{\bar{M}}(\bar{m}^1, \ ..., \bar{m}^k) = (f^{M_i}(m_i^1, \ ... \, m_i^k))_i$
- Para cada símbolo de relación $R \in \mathcal{L}$, definimos $R^{\bar{M}}(\bar{m}^1, \ \dots \ , \bar{m}^k)$ si y solo para todo $i \in I$ vale que $M_i \models R(m_i^1, ..., m_i^k)$

Por supuesto, la proyección a la k-esima coordenada, $\pi_k : \prod M_i \to M_k$ es un morfismo de $\mathcal{L}$-estructuras. El lector amante de las categorías podra divertirse comprobando que cumple la propiedad universal del producto y más aun, que la categoría de $\mathcal{L}$-estructuras tiene todos los límites finitos, pero que en general no tiene colímites finitos.

Ahora tenemos que definir la parte sintáctica, las fórmulas de primer orden sobre $\mathcal{L}$. Las fórmulas sobre $\mathcal{L}$ son una combinación formal de los símbolos de $\mathcal{L}$, los símbolos lógicos $=, \neg, \wedge, \exists$ y un conjunto infinito $\{x, y, z, ...\}$ de variables. Empecemos definiendo los términos y luego definiremos las fórmulas.

Definición 3.11. Sea $\mathcal{L}$ un lenguaje. Los **términos** sobre $\mathcal{L}$ se definen recursivamente:

- Si x es una variable, es un término.
- Si $c \in \mathcal{L}$ es un símbolo de constante, c es un termino.
- Si $f \in \mathcal{L}$ es una función de aridad k y $t_1, ..., t_k$ son terminos, $f(t_1, ..., t_k)$ es un termino.

Definición 3.12. Sea $\mathcal{L}$ un lenguaje. Las **fórmulas** sobre $\mathcal{L}$ se definen recursivamente:

- Si t_1, t_2 son términos, $t_1 = t_2$ es una fórmula.
- Si $R \in \mathcal{L}$ es una relación de aridad k y $t_1, ..., t_k$ son términos, $R(t_1, ..., t_k)$ es una fórmula.
- Si ϕ es una fórmula, $\neg\phi$ es una fórmula.
- Si ϕ_1 y ϕ_2 son fórmulas, $\phi_1 \ \wedge \ \phi_2$ es una fórmula.

- Si ϕ es una fórmula y x es una variable, $\exists x : \phi$ es una variable.

Definición 3.13. Una fórmula formada de forma i) o ii) es llamada **atómica**.

Definición 3.14. Una variable que aparece en una fórmula bajo el dominio de un cuantificador se dice que esta **ligada**, mientras que si no aparece bajo el dominio de ningún cuantificador se dice que esta **libre**.

Definición 3.15. Una fórmula sin variables libres es llamada una **oración**.

Ejemplo 3.16. Sea $\mathcal{L} = \{f\}$ el lenguaje con una unica función unaria. $\phi_1 = [\forall x : f(f(x)) = x]$ es una oración, mientras que $\phi_2 = [\exists y : f(y) = x]$ no, ya que tiene a la variable x libre.

Ahora viene la definición que conecta el mundo la semantico con el sintactico: la definición de **evaluación**, de cuando una fórmula es cierta en una $\mathcal{L}$-estructura dada. Dada una $\mathcal{L}$-estructura M y una fórmula ϕ con variables libres $x_1, ..., x_k$, para determinar si M cumple ϕ necesito una **valuación**, una función $v : \{x_1, ...x_n\} \to M$ que le asigna un valor a cada variable libre. Las valuaciones tambien se pueden escribir explicitamente como $v = \{x = m_1, ..., x = m_k\}$.

Definición 3.17. Sea $\mathcal{L}$ un lenguaje y v una valuación.

- $M[v] \models t_1 = t_2$ si y solo si $v(t_1) = v(t_2)$
- $M[v] \models R(t_1, ..., t_k)$ si y solo si $(v(t_1), ..., v(t_k)) \in R$
- $M[v] \models \neg\phi$ si y solo si $M[v] \not\models \phi$
- $M[v] \models \phi_1 \ \wedge \ \phi_2$ si y solo si $M[v] \models \phi_1$ y $M[v] \models \phi_2$
- $M[v] \models \exists x \ \phi$ si y solo si hay un $m \in M$ tal que $M[v, x = m] \models \phi$

A partir de ahora denotaremos a las oraciones como ϕ y a las fórmulas con variables libres como $\phi(x_1, .., x_k) = \phi(\bar{x})$. Ademas denotaremos $M[x_1 = m_1, ..., x_k = m_k] \models \phi(x_1, ..., x_k)$ como $M \models \phi(m_1, ..., m_k)$.

Observación 3.18. El lector atento habra notado que en nuestra definición de lógica de primer orden no incluimos ni el signo $\vee$ ni el cuantificador $\forall$. La razón es puramente económica: como muchas demostraciones son por inducción estructural, es conveniente tener la definción más corta posible. En su lugar, definimos $\phi_1 \ \vee \ \phi_2$ como $\neg(\neg\phi_1 \ \wedge \ \neg\phi_2)$ y $\forall x \ \ \phi$ como $\neg(\exists x \ \neg\phi)$. Por las leyes de Morgan, se cumple la semántica que esperamos:

- $M[v] \models \phi_1 \ \vee \ \phi_2$ si y solo si $M[v] \models \phi_1$ o $M[v] \models \phi_2$
- $M[v] \models \forall x \ \phi$ si y solo si hay para todo $m \in M$ vale que $M[v, x = m] \models \phi$

Proposición 3.1. *Sea $\mathcal{L}$ un lenguaje y $FORM(\mathcal{L})$ el conjunto de todas las fórmulas sobre lenguaje $\mathcal{L}$. Luego $|FORM(\mathcal{L})| = max(|\mathcal{L}|, \aleph_0)$*

Demostración. Ejercicio para el lector. □

Definición 3.19. Sea $\mathcal{L}$ un lenguaje y M y N dos $\mathcal{L}$-estructuras. Decimos que M y N son **elementariamente equivalentes** (denotado $M \equiv N$) si si para toda oración ϕ en lenguaje $\mathcal{L}$ vale que $M \models \phi$ si y solo si $N \models \phi$.

Teorema 3.2. *Sea $\mathcal{L}$ un lenguaje, M y N dos $\mathcal{L}$-estructuras y $\pi : M \to N$ un isomorfismo. Luego, para toda fórmula $\phi(\bar{x})$ sobre $\mathcal{L}$ y toda tupla $\bar{m} \in M$ vale:*
$M \models \phi(\bar{m})$ si y solo si $N \models \phi(\pi(\bar{m}))$.
En particular, $M \equiv N$.

Demostración. Por inducción en la estructura de $\phi(\bar{x})$:

- Como π es un isomorfismo de $\mathcal{L}$-estructuras debe preservar la igualdad de terminos y todo símbolo de relación $R \in \mathcal{L}$ luego es fácil ver que si ϕ es una fórmula atomica vale que $M \models \phi(\bar{m})$ si y solo si $N \models \phi(\pi(\bar{m}))$.
- Supongamos que $\phi(\bar{x}) = \neg\psi(\bar{x})$. Por hipótesis inductiva, vale que $M \models \psi(\bar{m})$ si y solo si $N \models \psi(\pi(\bar{m}))$. Luego tenemos la siguiente cadena de equivalencias:

  $M \models \phi(\bar{m})$ sii $M \not\models \psi(\bar{m})$ sii $N \not\models \psi(\pi(\bar{m}))$ sii $N \models \phi(\pi(\bar{m}))$
- Supongamos que $\phi(\bar{x}) = \phi_1(\bar{x}) \ \wedge \ \phi_2(\bar{x})$. Por hipótesis inductiva tengo que vale que:

  $M \models \phi_1(\bar{m})$ si y solo si $N \models \phi_1(\pi(\bar{m}))$.

  $M \models \phi_2(\bar{m})$ si y solo si $N \models \phi_2(\pi(\bar{m}))$.

  Intersecando ambas llegamos a que:

  $M \models \phi_1(\bar{m}) \ \wedge \ \phi_2(\bar{m})$ si y solo si $N \models \phi_1(\pi(\bar{m})) \ \wedge \ \phi_2(\pi(\bar{m}))$.
- Supongamos que $\phi(\bar{x}) = \exists y : \psi(y, \bar{x})$. Si $M \models \phi(\bar{m})$ se sigue que existe un m_o tal que $M \models \psi(m_0, \bar{m})$, luego por hipótesis inductiva se sigue que $N \models \psi(\pi(m_0), \pi(\bar{m}))$ y por tanto $N \models \phi(\bar{m})$. Conversamente, si $N \models \phi(\pi(\bar{m}))$ se sigue que existe $n_0 \in N$ tal que $N \models \psi(n_0, \pi(\bar{m}))$. Como π es una biyección se sigue que $n_0 = \pi(m_0)$ luego por hipótesis inductiva vale que $M \models \psi(m_0, \bar{m})$ y por tanto $M \models \phi(\bar{m})$.

□

Observación 3.20. El teorema anterior muestra que las estructuras isomorfas son elementariamente equivalentes, lo que abre la pregunta de si vale la vuelta. La respuesta es no y es facil verlo por argumentos de cardinalidad: Sea $\mathcal{L}$ el lenguaje de una unica función binaria. Por la proposición 3.1, hay una cantidad contable de oraciones sobre $\mathcal{L}$, luego la cantidad de $\mathcal{L}$-estructuras hasta equivalencia elemental no puede ser mas que $2^{\aleph_0} = c$. Por otro lado, observamos que las $\mathcal{L}$-estructuras son precisamente los magmas y sabemos que hay magmas de cualquier cardinalidad, es decir, la colección de todos los magmas ni siquiera un conjunto propiamente dicho, sino una clase. Se sigue que hay $\mathcal{L}$-estructuras elementariamente equivalentes pero no isomorfas.

La argumentación anterior es poco satisfactoria, sin embargo, porque no nos da ningún ejemplo concreto de dos estructuras elementariamente equivalentes pero no isomorfas. Veremos en la siguiente sección un metodo muy general para construir ejemplos de dicha naturaleza.

Definición 3.21. Sea $\mathcal{L}$ un lenguaje. Un conjunto T de oraciones en lenguaje $\mathcal{L}$ es llamado una **teoría** sobre $\mathcal{L}$.

Definición 3.22. Sea $\mathcal{L}$ un lenguaje, M una $\mathcal{L}$-estructura y T una teoría sobre $\mathcal{L}$. Decimos que M es **modelo** de T (denotado $M \models T$) si para todo $\phi \in T$ vale que $M \models \phi$.

Ejemplo 3.23. Sea $\mathcal{L}$ el lenguaje dado un símbolo de constante e, una función unaria i y una función binaria $*$ y consideremos las siguientes oraciones:

- $\phi_1 = \forall x \ \forall y \ \forall z : ((x * y) * z = x * (y * z))$
- $\phi_2 = \forall x \forall y \ (x * y = y * x)$
- $\phi_2 = \forall x \ ((x * e) = x)$
- $\phi_3 = \forall x \ (x * i(x) = e)$

Luego los modelos de $T = \{\phi_1, \phi_3, \phi_4\}$ son los grupos mientras que los modelos de $S = \{\phi_1, \phi_2, \phi_3, \phi_4\}$ son los grupos abelianos.

Definición 3.24. Una teoría T es **satisfacible** si tiene algun modelo.

Definición 3.25. Una teoría T se dice **completa** si es satisfacible y para toda oración ϕ o bien $T \cup \{\phi\}$ es insatisfacible o bien $T \cup \{\neg\phi\}$ es insatisfacible. Equivalentemente, T completa si para toda teoría satisfacible S tal que $T \subset S$ se cumple que T y S tienen los mismos modelos.

3.2. Filtros y Ultraproductos

En esta sección vamos a generalizar la construcción que hicimos anteriormente de los hiperreales a partir de los números reales a $\mathcal{L}$-estructuras para **cualquier** lenguaje $\mathcal{L}$ y probaremos un teorema muy general (el teorema de Łoś) que caracterizara sus propiedades. Antes de empezar la construcción, debemos generalizar el trabajo que hicimos sobre filtros sobre $\mathbb{N}$ a filtros sobre cualquier conjunto.

Definición 3.26. Un filtro $\mathcal{F}$ sobre un conjunto I es una colección de subconjuntos de I tal que:

- $\varnothing \notin \mathcal{F}$.
- Si $U, V \in \mathcal{F}$, entonces $U \cap V \in \mathcal{F}$. (cerrado por intersección)
- Si $U \in \mathcal{U}$ y $U \subset V$, entonces $V \in \mathcal{F}$. (cerrado por superconjuntos)

Observación 3.27. Por la propiedad iii), si $\mathcal{F}$ es un filtro sobre I, entonces $I \in \mathcal{F}$

Definición 3.28. Dado $a \in I$, la colección $\mathcal{F} = \{U : a \in U\}$ es un filtro y es llamado el **filtro principal** generado por a

Definición 3.29. Sea I un conjunto infinito. La colección $\mathcal{F} = \{A : A^c \text{ es finito}\}$ es un filtro y es llamado el filtro cofinito sobre I

Definición 3.30. Un filtro $\mathcal{U}$ sobre I se dice **ultrafiltro** si para todo $A \subset I$, o bien $A \in \mathcal{U}$ o bien $A^c \in \mathcal{U}$.

Observación 3.31. Una definición equivalente de ultrafiltro es la siguiente: Un ultrafiltro es una medida σ sobre $\mathcal{P}(I)$ finitamente aditiva tal que para todo $A \subset I$ vale que $\sigma(A) = 0$ O $\sigma(A) = 1$.

Definición 3.32. Un filtro $\mathcal{F}$ sobre I se dice **maximal** si no existe otro filtro $\mathcal{F}'$ tal que $\mathcal{F} \subset \mathcal{F}'$.

Proposición 3.3. *Un filtro $\mathcal{U}$ es un ultrafiltro si y solo si es maximal.*

Demostración. Análoga a la demostración de 2.7. □

Proposición 3.4. *Sea I un conjunto y $\mathcal{U}$ un ultrafiltro sobre I. Si $\mathcal{U}$ contiene a un subconjunto finito de I, se sigue que $\mathcal{F}$ es un filtro principal.*

Demostración. Por inducción en el tamaño del subconjunto finito. Es claro que si $\mathcal{U}$ contiene a un singleton $\{a\}$ se sigue que $\mathcal{U}$ es el filtro principal generado por a. Para el paso inductivo, supongamos que $\mathcal{U}$ contiene a un conjunto finito A, tal que $|A| = k + 1$. Sea $a \in A$. Como $\mathcal{U}$ es un ultrafiltro, o bien $\{a\} \in \mathcal{U}$ o bien $I - \{a\} \in \mathcal{U}$. Si $\{a\} \in \mathcal{U}$ se sigue que $\mathcal{U}$ es principal. Si $I - \{a\} \in \mathcal{U}$, como los filtros son cerrados por intersección, se sigue que $(I - \{a\}) \cap A = (A - \{a\}) \in \mathcal{U}$. Pero $|A| = k$, luego por hipótesis inductiva se sigue que $\mathcal{U}$ es principal. □

Proposición 3.5. *Sea I un conjunto finito y $\mathcal{U}$ un ultrafiltro sobre I. Luego $\mathcal{U}$ es principal.*

Demostración. Por la observación 3.27, sabemos que $I \in \mathcal{U}$ y como I es finito por la proposición 3.5 se sigue que $\mathcal{U}$ es principal. □

Teorema 3.6. ***(Lema del ultrafiltro)*** *Sea I un conjunto y $\mathcal{F}$ un filtro sobre I. Luego, existe un ultrafiltro sobre I que extiende a $\mathcal{F}$.*

Demostración. Analogo a 2.2. Sea Ω el conjunto de todos los filtros que extienden a $\mathcal{F}$. Sea $\mathcal{F}_i$ una cadena ascendente en Ω. No es dificil de ver que la union $\bigcup \mathcal{F}_i$ es tambien un filtro, luego se se sigue que toda cadena ascendente en Ω tiene un supremo y luego por el lema de Zorn se sigue que Ω tiene un elemento maximal, es decir, existe un filtro maximal $\mathcal{U}$ que extiende a $\mathcal{F}$. Por la proposición 2.7 $\mathcal{U}$ es un ultrafiltro que extiende a $\mathcal{F}$. □

(El lema del ultrafiltro y el axioma de elección) Como mencionamos anteriormente, la demostración anterior no es constructiva y depende del axioma de elección. Eso abre la pregunta ¿Es realmente necesario usar el axioma de elección? La respuesta es si: se ha demostrado que el lema del ultrafiltro es independiente de ZF (Zermelo-Frankel sin elección). Por otro lado, uno podría preguntarse si el lema del ultrafiltro es equivalente al axioma de elección, como ocurre frecuentemente con enunciados del estilo. La respuesta es negativa: El lema del ultrafiltro es estrictamente más debil que el axioma de elección y de hecho, es incluso mas débil que el axioma de elección contable..

Sin embargo, el lema del ultrafiltro no es nada débil: es capaz de probar el teorema de compacidad y como veremos en la siguiente sección, el teorema de compacidad se puede usar para deducir una amplia gama de teoremas.

Teorema 3.7. *Sea I un conjunto infinito. Luego existe un ultrafiltro no principal sobre I.*

Demostración. Analoga a la demostracion 3.6. □

Definición 3.33. Sea I un conjunto y $\mathcal{A}$ una colección de subconjuntos de I. Decimos que $\mathcal{A}$ tiene la **propiedad de la intersección finita** (PIF) si para todo conjunto finito $A_1, ..., A_k \in \mathcal{A}$ vale que $A_1 \cap (...) \cap A_k \neq \varnothing$

Proposición 3.8. *Sea I un conjunto y $\mathcal{A}$ una colección de subconjuntos de I con la PIF. Luego existe un filtro $\mathcal{F}$ sobre I tal que $\mathcal{A} \subset \mathcal{F}$*

Demostración. Sea $\mathcal{F}$ la colección dada por $\{S : (A_1 \cap (...) \cap A_k) \subset S \; con \; A_1, ..., A_k \in \mathcal{A}\}$, es decir, los superconjuntos de intersecciones finitas de $\mathcal{A}$. Esta claro que $\mathcal{F}$ es cerrado por superconjuntos y como $\mathcal{A}$ tiene la PIF, se sigue que $\varnothing \notin \mathcal{F}$. Ademas, es fácil corroborar que es cerrado por intersecciones, luego es efectivamente un filtro. □

Juntando las proposiciones 3.6 y 3.8 llegamos a la siguiente, que es extremadamente útil para construir ultrafiltros con caracteristicas deseadas.

Proposición 3.9. *Sea I un conjunto, $\mathcal{A}$ una colección de subconjuntos de I con la PIF. Luego existe un ultrafiltro $\mathcal{U}$ sobre I tal que $\mathcal{A} \subset \mathcal{U}$*

Demostración. Por la proposición 3.8, existe un filtro $\mathcal{F}$ tal que $\mathcal{A} \subset \mathcal{F}$. Por la proposición 3.6, existe un ultrafiltro $\mathcal{U}$ tal que $\mathcal{F} \subset \mathcal{U}$. □

Definición 3.34. Sea I un conjunto, $\mathcal{F}$ un filtro sobre I y $(M_i)_{i \in I}$ una familia de conjuntos. Definimos la relación $\sim_{\mathcal{F}}$ sobre $\prod A_i$ como $(a_i) \sim_{\mathcal{F}} (b_i)$ si y solo si $\{i : a_i = b_i\} \in \mathcal{F}$

Proposición 3.10. *La relación $\sim_{\mathcal{F}}$ es una relación de equivalencia.*

Ahora finalmente podemos definir una construcción que generaliza el producto de $\mathcal{L}$-estructuras, conocida como **producto reducido o $\mathcal{F}$-producto**.

Definición 3.35. Sea $\mathcal{L}$ un lenguaje, I un conjunto, $(M_i)_{i\in I}$ una colección de $\mathcal{L}$-estructuras y $\mathcal{F}$ un filtro sobre I. Definimos el **$\mathcal{F}$-producto** como la $\mathcal{L}$-estructura $\bar{M}$ con dominio $\prod M_i/\sim_{\mathcal{F}}$ definida de la siguiente manera:

- Para cada símbolo de constante $c \in \mathcal{L}$, definimos $c^{\bar{M}} = [(c^{M_i})_i]$
- Para cada símbolo de función $f \in \mathcal{L}$, definimos $f^{\bar{M}}(\bar{m}^1, \ ..., \bar{m}^k) = [(f^{M_i}(m_i^1, \ ... m_i^k))]$
- Para cada símbolo de relacion $R \in \mathcal{L}$, definimos $R^{\bar{M}}(\bar{m}^1, \ ..., \bar{m}^k)$ si y solo $\{i : M_i \models R^{M_i}(m_i^1, \ ..., m_i^k)\} \in \mathcal{F}$

Observación 3.36. Si $\mathcal{F}$ es el filtro trivial el $\mathcal{F}$-producto $\prod M_i/\sim_{\mathcal{F}}$ no es más que el producto usual $\prod M_i$.

Definición 3.37. El caso particular de un $\mathcal{F}$-producto donde todos los M_i son isomorfos es llamado una **$\mathcal{F}$-potencia** y se denota $M^I/\sim_{\mathcal{F}}$

Definición 3.38. Sea $\mathcal{L}$ un lenguaje, I un conjunto, $\mathcal{F}$ un filtro sobre I y M una $\mathcal{L}$-estructura. Luego hay una **inclusión canonica** $j : M \to M^I/\sim_{\mathcal{F}}$ dada por:

$j(m)_i = m$ para todo $i \in I$

Es fácil ver que j es inyectiva y un morfismo de $\mathcal{L}$-estructuras.

Definición 3.39. Un caso de particular importancia es cuando el filtro en cuestion es un ultrafiltro $\mathcal{U}$, en cuyo caso denominaremos al como **ultraproducto** como **ultrapotencias**.

Finalemente podemos enunciar el siguiente teorema, conocido como el teorema de Łoś o el teorema fundamental de los ultraproductos, que nos da una criterio muy simple para determinar cuando un ultraproducto cumple una fórmula dada.

Teorema 3.11. ***(Łoś)*** *Sea $\mathcal{L}$ un lenguaje, I un conjunto, $(M_i)_{i\in I}$ una colección de $\mathcal{L}$-estructuras y $\mathcal{U}$ un ultrafiltro sobre I. Sea $\bar{M} = \prod M_i/\sim_{\mathcal{U}}$ su ultraproducto. Para toda fórmula $\phi(x_1, ..., x_k)$ en lenguaje $\mathcal{L}$ y todo conjunto $\bar{m}^1, ..., \bar{m}^k$ vale que:*

$\bar{M} \models \phi(\bar{m}^1, \ ..., \bar{m}^k)$ *si y solo si* $\{i \in I : M_i \models \phi(m_i^1, \ ..., m_i^k)\} \in \mathcal{U}$.

En particular, si ϕ es una oración vale que:

$\bar{M} \models \phi$ *si y solo si* $\{i \in I : M_i \models \phi\} \in \mathcal{U}$.

Demostración. Por inducción en la estructura de $\phi(x_1, \ ..., x_k)$. Es fácil ver que vale para oraciones atómicas:

- Supongamos que $\phi(x_1, ..., x_k) = \neg\psi(x_1, ..., x_k)$. Por hipótesis inductiva sabemos que vale que $\bar{M} \models \psi$ sii $\{i : M_i \models \phi\} \in \mathcal{U}$. Luego vale que $\bar{M} \models \phi$ sii $\bar{M} \not\models \psi$ sii $\{i : M_i \models \phi\} \notin \mathcal{U}$. Como $\mathcal{U}$ es un ultrafiltro, eso último es equivalente a $\{i : M_i \models \psi\}^c \in \mathcal{U}$ sii $\{i : M_i \models \phi\} \in \mathcal{U}$
- Supongamos que $\phi = \phi_1 \ \wedge \ \phi_2$. Por hipótesis inductiva sabemos que vale que:

  $\bar{M} \models \phi_1$ si y solo si $\{i : M_i \models \phi_1\} \in \mathcal{U}$

  $\bar{M} \models \phi_2$ si y solo si $\{i : M_i \models \phi_2\} \in \mathcal{U}$

  Intersecando ambas llegamos a que:

  $\bar{M} \models \phi_1 \ \wedge \ \phi_2$ si y solo si $\{i : M_i \models \phi_1\} \in \mathcal{U}$ y $\{i : M_i \models \phi_1\} \in \mathcal{U}$

- Supongamos que $\phi(x_1,\ ...,x_k) = \exists y\ \psi(y,x_1,\ ...,x_k)$. Si $\bar{M} \models \exists y : \psi(y,\bar{m}^1,\ ...\bar{m}^k)$, debe existir un $\bar{m}^0 \in \bar{M}$ tal que $\bar{M} \models \psi(\bar{m}^0,\bar{m}^1,\ ...\ ,\bar{m}^k)$. Por hipótesis inductiva, se sigue que $\{i : M_i \models \psi(m_i^0,\ ...,\ m_i^k)\} \in \mathcal{U}$. Como vale que $\{i : M_i \models \exists y : \psi(m_i^0,m_i^1,...,m_i^k)\} \subset \{i : M_i \models \exists y\ \psi(y,)\} = \{i : M_i \models \phi(m_i^1,...,m_i^k)\}$ y los filtros son cerrados por superconjuntos, se sigue que $\{i : M_i \models \phi(m_i^1,...,m_i^k)\} \in \mathcal{U}$, como queriamos. Conversamente, supongamos que $\{i : M_i \models \phi(m_i^1,...,m_i^k)\} \in \mathcal{U}$, es decir, hay un subconjunto $U \in \mathcal{U}$ tal que para $i \in U$ existe un $c_i \in M_i$ tal que $M_i \models \psi(c_i,m_i^1,...,m_i^k)$. Sea $\bar{c} \in \prod M_i$ tal que $\bar{c}_i = i$. Es fácil ver que $\bar{M} \models \psi(\bar{c},\bar{m}^1,\ ...,\bar{m}^k)$ luego $\bar{M} \models \exists y : \psi(y,\bar{m}^1,\ ...,\bar{m}^k)$ y por tanto $\bar{M} \models \phi(\bar{m}^1,\ ...,\bar{m}^k)$.

□

En el caso particular de una ultrapotencia el teorema de Łoś se simplifica al siguiente resultado, conocido como el principio de transferencia:

Teorema 3.12. ***(Principio de transferencia)*** *Sea $\mathcal{L}$ un lenguaje, I un conjunto y M una $\mathcal{L}$-estructura. Sea $\bar{M} = M^I/\sim_{\mathcal{U}}$ su ultraproducto. Luego para toda fórmula $\phi(x_1,...,x_k)$ en lenuaje vale que:*

$M \models \phi(m_1,...,m_k)$ si y solo si $\bar{M} \models \phi(i(m_1),...,i(m_k))$.

En particular, si ϕ es una oración vale que:

$M \models \phi$ si y solo si $\bar{M} \models \phi$.

Es decir, $M \equiv \bar{M}$.

(Aplicación al análisis no estandar) En la sección anterior, antes de empezar a realizar análisis no estándar propiamente dicho, tuvimos que demostrar que los hiperreales son un cuerpo ordenado. La demostración, aunque no particularmente complicada, era larga y laboriosa, porque requeria que verificaramos cada uno de los axiomas de un cuerpo ordenado. Ahora, con el principio de transferencia, podemos dar una demostración extremadamente rapida y elegante de un principio mucho más general.

Teorema 3.13. *Sea $\mathcal{L} = \{0,1,+,\cdot\}$ el lenguaje usual de los cuerpos y sea $A = (A,0,1,+,<)$ un cuerpo ordenado. Luego, para todo conjunto I y todo ultrafiltro $\mathcal{U}$ sobre I, el ultraproducto $A^I/\sim_{\mathcal{U}}$ es un cuerpo ordenado.*

Demostración. La clave de la demostración es que la propiedad de que una $\mathcal{L}$-estructura sea un cuerpo ordenado es expresable como una oración de primer orden, especificamente, la conjunción de las siguientes oraciones:

- $\forall x\ \forall y\ \forall z\ \ x + (y + z) = (x + y) + z$
- $\forall x\ \forall y\ \ x + y = y + x$
- $\forall x\ x + 0 = 0 + x = x$
- $\forall x\ \exists y : x + y = 0$
- $\forall x\ \forall y\ \forall z\ \ x \cdot (y \cdot z) = (x \cdot y) \cdot z$
- $\forall x\ \forall y\ \ x \cdot y = y \cdot x$
- $\forall x\ x \cdot 1 = 1 \cdot x = x$
- $\forall x\ \exists y : x \cdot y = 1$
- $\forall x\ \forall y\ \forall z\ ((x < y) \wedge (y < z)) \to (x < z))$
- $\forall x\ \forall y\ \forall z\ ((x < y) \to (x + z < y + z))$
- $\forall x\ \forall y\ ((x > 0) \wedge (y > 0) \to (x \cdot y > 0))$

Luego, como por el principio de transferencia sabemos que $A^I/\sim_{\mathcal{U}}$ y A cumplen las mismas oraciones de primer orden, y A es un cuerpo ordenado, se sigue que $A^I/\sim_{\mathcal{U}}$ es un cuerpo ordenado. □

En general, el principio de transferencia es muy útil para demostrar que propiedades de los reales pasan a los hiperreales de forma automatica, es decir, sin tener que invocar la construcción explicita de los hiperreales. Por ejemplo:

Proposición 3.14. *Sea $f : \mathbb{R} \to \mathbb{R}$ una función creciente. Luego su extensión $f^* : \mathbb{R}^* \to \mathbb{R}^*$ es creciente.*

Demostración. Sea $\mathcal{L} = \{0, 1, +, \cdot, <, f\}$ el lenguaje usual de los cuerpos más un símbolo de función unaria f. El hecho de que f sea creciente es capturado por la fórmula de primer orden $\forall x\ \forall y\ ((x < y) \Rightarrow (f(x) < f(y)))$ Por el principio de transferencia, como $\mathbb{R} \models \phi$ se sigue que $\mathbb{R}^* \models \phi$ y por tanto $f : \mathbb{R}^* \to \mathbb{R}^*$ es creciente. □

El principio de transferencia tambien puede ser usado para construir modelos no estándar de la aritmetica, es decir, estructuras que cumplen las mismas oraciones de primer orden que $\mathbb{N}$ pero que no son isomorfos.

Proposición 3.15. ***(Modelos no estándar de la aritmetica)*** *Sea $\mathcal{L} = \{0, 1, +, \cdot\}$ el lenguaje con símbolos de constante 0 y 1 y funciones binarias $(+)$ y $(\cdot)$. Existe una $\mathcal{L}$-estructura A tal que $A \equiv (\mathbb{N}, 0, 1, +, \cdot)$ pero $A \ncong \mathbb{N}$.*

Demostración. Sea $\mathcal{U}$ un ultrafiltro no trivial sobre $\mathbb{N}$ y sea A el ultraproducto $\mathbb{N}^{\mathbb{N}}/\sim_{\mathcal{U}}$. Por el principio de transferencia, sabemos que $A \equiv \mathbb{N}$. Por otro lado, es fácil ver que A y $\mathbb{N}$ **no** son isomorfos: un isomorfismo $\sigma : \mathbb{N} \to A$ debe cumplir que $\sigma(1) = 1_A = [(1, 1, 1, ...)]$ luego $\sigma(n) = [(n, n, n, ...)]$. Por otro lado como $\mathcal{U}$ es no principal para todo n vale que $[(n, n, n, ...)] \neq [(1, 2, 3, ...)]$, luego se sigue que $[(1, 2, 3, ...)] \notin Im(\sigma)$ y por tanto σ no puede ser una biyección. □

Terminamos esta sección con dos maravillosos teoremas sobre ultraproductos. Invitamos al lector interesado a buscar sus demostraciones al libro de Chang.

Definición 3.40. Sea $\mathcal{L}$ un lenguaje. Sea $\mathcal{K}$ una colección de $\mathcal{L}$-estructuras. Decimos que $\mathcal{K}$ es **elemental** si y solo si $\mathcal{K}$ es de la forma $\{M : M \models T\}$ para alguna teoría T, en cuyo caso denotamos a $\mathcal{K}$ como $Mod(T)$.

Teorema 3.16. ***(Keisler)*** *Sea $\mathcal{L}$ un lenguaje y $\mathcal{K}$ una colección de $\mathcal{L}$-estructuras. Luego, $\mathcal{K}$ es elemental si y solo si es cerrada por $\equiv$ y ultraproductos.*

Teorema 3.17. ***(Keisler-Shelah)*** *Sea $\mathcal{L}$ un lenguaje y sean M y N dos $\mathcal{L}$-estructuras. Luego son equivalentes:*

- $M \equiv N$.
- $M^I/\sim_{\mathcal{U}} \cong N^I/\sim_{\mathcal{U}}$ *para algun conjunto I y algun ultrafiltro $\mathcal{U}$ sobre I.*

3.3. El Teorema de Compacidad y sus aplicaciónes

En esta sección vamos a usar los resultados anteriores de ultrafiltros para demostrar uno de los resultados más importantes de teoría de modelos, y probablemente mi teorema favorito de toda la matemática: el teorema de Compacidad. La demostración original, que es la que se encuentra en la mayoria de libros de texto[4][5], es de naturaleza sintáctica, mediante el teorema de Completitud de Gödel y la llamada costrucción de Henkin. Nosotros daremos una demostración de naturaleza más semántica, basada en ultraproductos y que, ahora que tenemos el teorema de Łoś a nuestra disposición, sera extremadamente rápida.

Teorema 3.18. ***(Teorema de Compacidad, versión contable)*** *Sea $\mathcal{L}$ un lenguaje y $T = \{\phi_1, \phi_2, ...\}$ una teoría sobre $\mathcal{L}$. Si cada subconjunto finito de oraciones de T es satisfacible, se sigue que T es satisfacible.*

Demostración. Sabemos que todo subconjunto finito de oraciones de T es satisfacible, luego para todo $i \in \mathbb{N}$ debe haber una $\mathcal{L}$-estructura M_i tal que $M_i \models \phi_1 \ \wedge (...) \wedge \phi_i$. Sea $\mathcal{U}$ un ultrafiltro no principal sobre $\mathbb{N}$ y consideremos el ultraproducto $\bar{M} = \prod M_i / \sim_{\mathcal{U}}$. Veamos que $\bar{M} \models T$ y por lo tanto, T es satisfacible.

Por el teorema Łoś, $\bar{M} \models \phi_n$ si y solo si $\{i \in I : M_i \models \phi_n\} \in \mathcal{U}$. Pero por la construcción de los M_i, sabemos que para todo $i \geq n$ vale que $M_i \models \phi_n$, luego $\{i \in I : M_i \models \phi_n\}$ es cofinito y por tanto pertence a $\mathcal{U}$. Luego, se sigue que para todo n vale que $\bar{M} \models \phi_n$. □

Con un poco más de trabajo, podemos llegar al mismo resultado sin la hipótesis de contabilidad:

Teorema 3.19. ***(Teorema de Compacidad, versión general)*** *Sea $\mathcal{L}$ un lenguaje y T una teoría sobre $\mathcal{L}$. Si cada subconjunto finito de oraciones de T es satisfacible, se sigue que T es satisfacible*

Demostración. Sea $T^* = \{F \subset T : F \ finito\}$ el conjunto de todos los subconjuntos finitos de I. Para cada $\phi \in T$, sea $S_\phi \subset T^*$ dado por $\{F : \phi \in F\}$. Sea $\mathcal{A}$ la colección de subconjuntos de T^* dada por los S_ϕ para todo $\phi \in T$. Esta claro que $\mathcal{A}$ tiene la PIF, ya que $\{\phi_1, ..., \phi_k\} \in S_{\phi_1} \cap (...) \cap S_{\phi_k}$, luego por el teorema 3.9 existe un ultrafiltro $\mathcal{U}$ sobre T tal que $\mathcal{A} \subset \mathcal{U}$

Como por hipótesis todo subconjunto finito de T es satisfacible, se sigue que para todo $S \in T^*$ existe una $\mathcal{L}$-estructura M_S tal que $M_S \models S$. Sea $\bar{M}$ el ultraproducto $\prod M_S / \sim_{\mathcal{U}}$. Veamos que $\bar{M} \models T$.

Sea $\phi \in T$. Por el teorema de Łoś, sabemos que $\bar{M} \models T$ si y solo si $\{S : M_S \models \phi\} \in \mathcal{U}$. Ahora, nosotros construimos los M_S de forma que si $\phi \in S$ vale $M_S \models \phi$. Luego tenemos $S_\phi = \{S : \phi \in S\} \subset \{S : M_S \models \phi\}$. Sabemos que $S_\phi \in \mathcal{U}$, ya que $S_\phi \in \mathcal{A}$ y $\mathcal{U}$ extiende a $\mathcal{A}$, y como los filtros son cerrados por superconjuntos, se sigue que $\{S : M_S \models T\} \in \mathcal{U}$, luego $\bar{M} \models T$.

□

Proposición 3.20. *Sea $\mathcal{L}$ un lenguaje y T una teoría sobre $\mathcal{L}$ con modelos finitos arbitrariamente grandes. Luego T tiene un modelo infinito.*

Demostración. Sea $\phi_n = \exists x_1 \ (...) \ \exists x_n : (\bigwedge_{i \neq j} x_i \neq x_j)$, la oración que afirma que hay al menos n elementos distintos. Sea $T' = T \cup \{\phi_1, \phi_2, ...\}$. Notemos que un modelo de T' es un modelo de T de cardinal infinito. Todo subcojunto finito de T esta contenido en uno de la forma $T \cup \{\phi_1, ..., \phi_n\}$ y ese conjunto de oraciones es satisfacible, ya que como T tiene modelos finitos arbitrariamente grandes en particular tiene un modelo de más de n elementos. Luego, todo subconjunto finito de T' es satisfacible y por el teorema de compacidad se sigue que T' es satisfacible, es decir, existen modelos de T infinitos. □

El teorema de compacidad es extremadamente útil para demostrar resultados de **inexpresabilidad**, es decir, demostrar que una cierta propiedad no puede ser expresada por una oración en lógica de primer orden. Por ejemplo:

Proposición 3.21. *No existe ninguna oración ϕ en el lenguaje de grafos que exprese la propiedad ."El grafo G es conexo".*

Demostración. Supongamos que existe tal oración ϕ. Sea $\mathcal{L}$ es lenguaje que consta de un símbolo de relación $\sim$ y dos símbolos de constante a y b, que podemo interpetar como el lenguaje de grafos con dos elementos destacados. Sea ϕ_n la oración en $\mathcal{L}$ dada por $\phi_n = [\exists x_1 \ ... \ \exists x_n : (a \sim x_1) \ \wedge \ (x_1 \sim x_2) \wedge \ ... \ \wedge (x_n \sim b)]$, la oración que afirma que hay un camino de n pasos de a a b.

Consideremos la teoría $T = \{\phi, \ \neg\phi_1, \neg\phi_2, \ ...\}$. Todo subconjunto finito de T es satisfacible, ya que $\{\phi, \ \neg\phi_1, \ ... \ , \neg\phi_n\}$ es satisfacido por un grafo en el que a y b esten a distancia $n+1$. Luego, por el Teorema de Compacidad, T es satisfacible. Pero eso absurdo: Si hay un camino entre a y b, debe haber un camino de a a b de n pasos para algun n. Luego, no puede existir tal fórmula ϕ. □

El problema de la lógica de segundo orden La proposición anterior muestra que el teorema de Compacidad **no** es válido si usamos lógica de segundo orden, ya que la propiedad de que un grafo sea conexo si se puede expresar con una oración de segundo orden (monádica), de la forma:

$$\forall X\ \forall Y\ (\forall z\ z \in X\ \vee\ z \in Y) \Rightarrow (\exists x \in X\ \exists y \in Y : x \sim y)$$

Lo mismo pasa con otras extensiones de la lógica de primer orden, como la lógica infinitaria. Mas en general, el matemático sueco Per Lindström demostro en 1969 que entre una gran familia de lógicas con cuantificadores generalizados, la lógica de primer orden es la mas general que cumple tanto el teorema de Compacidad como el teorema de Löwenheim–Skolem, del que hablaremos más adelante.

Dicho resultado se puede interpretar de la siguiente manera: asi como la lógica proposicional es demasiado debil para obtener resultados interesantes, la lógica de segundo orden y las lógicas infinitarias son demasiado fuertes para obtener resultados interesantes: solo en la lógica de primer orden encontramos el punto medio.

Dejamos como ejercicio al lector demostrar los siguientes reultados de inexpresabilidad, analogos a la proposición anterior.

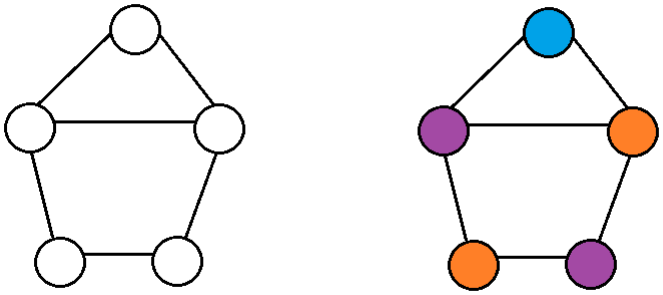
Proposición 3.22. *No hay ninguna oración ϕ en el lenguaje de los cuerpos que exprese la propiedad ."El cuerpo tiene caracteristica 0".*

Proposición 3.23. *No hay ninguna oración ϕ en el lenguaje de los grupos que exprese la propiedad ."El grupo es libre de torsión".*

Ahora veamos algunas aplicaciónes de caracter combinatorio:

Definición 3.41. Una k-coloración de un grafo $(G, \sim)$ es una partición de G en conjuntos disjuntos $A_1, ..., A_k$ tal que todo par de vertices vecinos pertenecen a distintos A_i. Un grafo G se dice k-coloreable si tiene una k coloración.

Ejemplo 3.42. Sea G el grafo representado por la figura de la izquierda. La figura de la derecha muestra que G es 3-coloreable. Por otro lado, es fácil ver que G no es 2-coloreable.



El siguiente celebre teorema demostrado por Paul Erdős y Nicolaas De Brujin muestra que la k-coloración de un grafo es una propiedad estrictamente local, en el sentido de que puede determinarse viendo solo sus subgrafos finitos.

Teorema 3.24. ***(De Brujin-Erdős)*** *Un grafo G es k-coloreable si y solo si todo subgrafo finito de G es k-coloreable.*

Demostración. Es evidente que si un grafo es k-coloreable, todo subgrafo es k-coloreable. Veamos que vale la vuelta. Sea G un grafo dado. La idea de la demostración es construir una teoría T sobre un cierto lenguaje $\mathcal{L}$ de forma que G sea k-coloreable si y solo si la teoría T es satisfacible, y luego aplicar el teorema de Compacidad a dicha teoría T para obtener el resultado enunciado.

Sea $\mathcal{L}$ el lenguaje con un símbolo de constante c_v para todo vertice de G, k predicados unarios $P_1, ..., P_k$ y una relacion binaria $\sim$. Sea T la siguiente teoría sobre $\mathcal{L}$:

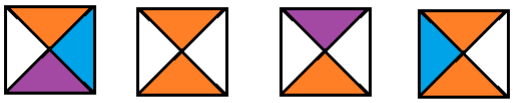
- $c_v \sim c_w$ para todo par $v, w \in G$ tal que $v \sim w$
- $(\bigvee P_i(c_v)) \ \wedge \ (\bigwedge \neg(P_i(c_v) \ \wedge \ P_j(v)))$ Para todo $v \in G$
- $(c_v \sim c_w) \rightarrow (\bigwedge \neg(P_i(v) \ \wedge P_i(w))$ para todo par $v, w \in G$

Cualquier $\mathcal{L}$-estructura que satisfaga T codifica un k-coloreo de G: el primer tipo de oraciones garantiza que los elementos c_v respetan la estructura del grafo, el segundo tipo de oraciones garantiza que todo vertice tiene un color y solo uno, y el tercer tipo garantiza que todo par de vertices adyacentes tengan distintos colores.

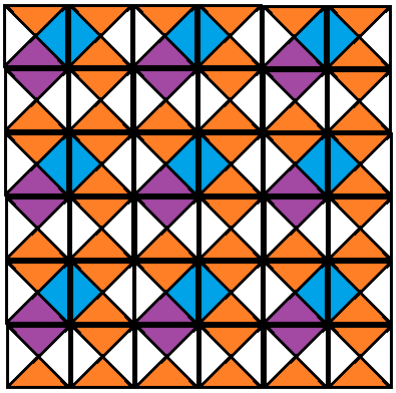
Luego, el grafo G es coloreable si y solo si la teoría T es satisfacible. Por el teorema de compacidad, la teoría T es satistacible si y solo si todo subconjunto finito de T es satisfacible, que pasa si y solo si todo subconjunto finito de G es k-coroleable. Luego, G es k-coroleable si y solo si todo subconjunto finito de G es k-coroleable.

□

Otra aplicación de espíritu muy similar es el siguiente teorema sobre el teselado con azueljos de Wang. Un conjunto de azulejos de Wang, nombrados en honor al lógico Hao Wang, es una coleccíon de cuadrados con colores asignados en cada uno de sus cuatro lados, como el siguiente.



Una teselación con un conjuntos de azulejos de Wang es una asignacion de un azulejo cada elemento de la grilla infinita $\mathbb{Z}^2$, de forma que los lados contiguos sean siempre del mismo color. Por ejemplo, el conjunto anterior de azulejos da lugar a la siguiente teselación:



Las definiciónes formales serian las siguientes:

Definición 3.43. Un conjunto de Azulejos de Wang es un conjunto Γ equipado con funciones $up, down, right, left : \Gamma \rightarrow \{1, ...k\}$ para $k \in \mathbb{N}$.

Definición 3.44. Sea Γ un conjunto de azulejos de Wang. Una teselación del plano por Γ es una asignación $\theta : \mathbb{Z}^2 \rightarrow \Gamma$ tal que para todo $i, j \in \mathbb{Z}$ vale que $up(\theta(i, j)) = down(\theta(i + 1, j))$ y $right(\theta(i, j)) = left(\theta(i, j + 1))$

Wang y Berger demostraron en 1968 el siguiente resultado asombroso: El problema de determinar si es posible teselar el plano un conjunto dado de azulejos de Wang es **indecidible**, es decir, no existe ningún algoritmo capaz de resolverlo. Su demostración consiste en reducir el problema de determinar si una maquina de Turing llega a un estado de aceptación, el ejemplo canonico de problema indecidible. Nosotros demostraremos que el problema se puede reducir a determinar si es posible teselar regiones arbitrariamente grandes del plano:

Proposición 3.25. *Sea Γ un conjunto de azulejos de Wang. Es posible teselar el plano con Γ si y solo para todo $n \in \mathbb{N}$ es posible teselar un cuadrado de $n \times n$ con Γ*

Demostración. Sea $\mathcal{L}$ el lenguaje consistente en símbolos de constante $c_{i,j}$ para todo par $i, j \in \mathbb{Z}$ y en relaciones unarias P_θ para todo $\theta \in \Gamma$. Sea T la siguiente teoría sobre $\mathcal{L}$:

- $(\bigvee P_\theta(c_{i,j})) \wedge (\bigwedge \neg(P_\theta(c_{i,j}) \wedge P_\alpha(i,j)))$ Para todo $i, j \in \mathbb{Z}$.
- $\neg(P_r(c_{i,j}) \wedge P_s(c_{i,j+1}))$ para todo $i, j \in \mathbb{Z}$ y todo $r, s \in \Gamma$ tal que $up(r) \neq down(s)$.
- $\neg(P_r(c_{i,j}) \wedge P_s(c_{i+1,j}))$ para todo $i, j \in \mathbb{Z}$ y todo $s, r \in \Gamma$ tal que $right(r) \neq left(s)$.

Cualquier modelo de T codifica una teselación del plano con Γ: el primer tipo de oraciones se asegura de que a cada (i, j) se le asigna un único azulejo de Wang y las oraciones del segundo y tercer tipo se asegura de que sean compatibles. Luego, existe un teselado del plano con Γ si y solo si T es satisfacible. Por otro lado, por el teorema de compacidad T es satisfacible si y solo si todo subconjunto finito de T es satisfacible y eso pasa si para si para todo $n \in \mathbb{N}$ es posible teselar un cuadrado de $n \times n$ con Γ. Luego es posible teselar el plano con Γ si y solo para todo $n \in \mathbb{N}$ es posible teselar un cuadrado de $n \times n$ con Γ. □

Veamos ahora algunas aplicaciónes del teorema de compacidad al álgebra, empezando con un resultado fundacional de teoría de cuerpos:

Teorema 3.26. ***(Existencia de la clausura algebraica)*** *Sea E un cuerpo. Luego existe un cuerpo $\bar{E}$ algebraicamente cerrado tal que $E \subset \bar{E}$.*

Demostración. Sea $\mathcal{L} = \{1, 0, +, \cdot\}$ el lenguaje usual de los cuerpos y sea $\mathcal{L}'$ la extensión de $\mathcal{L}$ con un símbolo de constante c_e por cada $e \in E$. Sea T la teoría sobre $\mathcal{L}'$ dada por:

- Los axiomas de un cuerpo para $\{1, 0, +, \cdot\}$
- $\neg(c_{e_1} \neq c_{e_2})$ para todo $e_1, e_2 \in E$ tal que $e_1 \neq e_2$
- $c_{e_1} + c_{e_2} = c_{e_3}$ para todo $e_1, e_2, e_3 \in E$ tal que $e_1 + e_2 = e_3$
- $c_{e_1} \cdot c_{e_2} = c_{e_3}$ para todo $e_1, e_2, e_3 \in E$ tal que $e_1 \cdot e_2 = e_3$
- $\exists x : (c_{e_0} + c_{e_1} \cdot x + (...) + c_{e_n} \cdot x^n = 0)$ Para toda colección de $e_0, e_1, ..., e_n \in E$ con $e_n \neq 0$

Cualquier modelo de T es un cuerpo (por i), que extiende a E (por ii, iii y iv) y en el cual todo polinomio no nulo de E tiene una raiz (por v). Es un resultado basico de álgebra que dados polinomios irreducibles $P_1, ..., P_n \in E[x]$ existe un cuerpo que extiende a E en el que $P_1, ..., P_n$ tienen raiz, el dado por $E[x_1, ..., x_n]/(P_1(x_1), ..., P(x_n))$. Se sigue que todo subcojunto finito de oraciones de T es satisfacible y luego, por el teorema de compacidad, que T es satisfacible: es decir, existe un cuerpo E_1 que extiende a E y en el cual todo polinomio con coeficientes en E tiene raiz.

Si iteramos la construcción anterior, llegamos a una cadena de extensiones de cuerpos $E = E_0 \subset E_1 \subset E_2 \subset ...$ con la propiedad de que todo polinomio no nulo con coeficientes en E_i tiene una raiz en E_{i+1}. Luego, el cuerpo $\bar{E} = \bigcup E_i$ es una extensión de E algebraicamente cerrada. □

Veamos ahora un resultado de teoría de grupos.

Definición 3.45. Un **grupo ordenado** es un grupo G equipado con un orden total $\leq$ compatible con la operación de grupo, es decir, tal que para todo $g_1, g_2, h \in G$:

- $g_1 \leq g_2$ implica $h \cdot g_1 \leq h \cdot g_2$
- $g_1 \leq g_2$ implica $g_1 \cdot h \leq g_2 \cdot h$

Definición 3.46. Un grupo $(G, \cdot)$ se dice **ordenable** si existe un orden total $\leq$ sobre G tal que $(G, \cdot, \leq)$ es un grupo ordenado.

Ejemplo 3.47. El grupo $(\mathbb{Z}, +)$ es ordenable porque el orden usual es invariante por traslación. Más en general, el grupo $(\mathbb{Z}^n, +)$ es ordenable porque el orden lexicografico es invariante por traslación.

Observación 3.48. Si G es ordenable y H es un subgrupo de G, se sigue que H es ordenable, ya que hereda el orden de G.

Observación 3.49. Ningún grupo con torsión puede ser ordenable. Para ver esto, supongamos que existe un $g \in G$ no nulo tal que $g^n = e$. Se sigue $g^{2n} = e$, pero por las propiedades del orden se sigue tambien que $g^{2n} = (g^n)^2 > e$, lo que es un absurdo.

Teorema 3.27. ***(Levi)*** *Sea G un grupo abeliano libre de torsión. Luego G es ordenable.*

Demostración. Sea $\mathcal{L}$ el lenguaje consistente en una funcion binaria $\cdot$, una relacion binaria R y símbolos de constante c_g para cada $g \in G$. Sea T la teoría sobre $\mathcal{L}$ consistente en:

- Los axiomas de un grupo ordenado para $\cdot$ y R.
- $c_{g_1} \neq c_{g_2}$ para todo par de elementos distintos $g_1, g_2 \in G$.
- $c_{g_1} \cdot c_{g_2} = c_{g_3}$ para todo $g_1, g_2, g_3 \in G$ tal que $g_1 \cdot g_2 = g_3$.

Cualquier modelo de T es un grupo ordenado (por i)) que tiene a G como subconjunto (por ii) y iii). Ahora, dados $g_1, ..., g_k \in G$, el subgrupo $< g_1, ..., g_k >$ es ordenable, ya que como G es libre de torsión por la clasificación de grupos abelianos finitamente generados se sigue que $< g_1, ..., g_k > \cong \mathbb{Z}^n$ para algun $n \leq k$ y observamos anteriormente que esos grupos son ordenables. Se sigue que todo subconjunto finito de T es satisfacible, luego por el teorema de compacidad T es satisfacible, es decir, existe un grupo ordenado H que tiene a G como subgrupo. Por la observación 3.48 se sigue que G es ordenable. □

Como comentario, es importante notar que el teorema de compacidad no es estrictamente necesario para probar ninguna de las aplicaciónes que mencionamos y que de hecho, todas fueron originalmente demostradas por otros metodos. Dicho esto, hay dos ventajas en demostrar los resultados de esta manera. La primera es que en muchos casos el teorema de compacidad da lugar a demostraciones mas directas y rápidas. Por ejemplo, el teorema de De Brujin-Erdos se puede demostrar usando solamente el lema de Zorn, pero la demostración es considerablemente más complicada y requiere usar el Lema de Kőnig, que no es trivial de demostrar. Lo mismo ocurre con el teorema de Levi de que todo grupo abeliano sin torsión es ordenable. La segunda es que el teorema de compacidad requiere solo del lema del ultrafiltro para ser demostrado y como mencionamos anteriormente, el lema del ultrafiltro es una hipótesis estrictamente mas débil que el axioma de elección o el lema de Zorn.

3.4. El Teorema de Löwenheim–Skolem

En esta seccion demostraremos otro de los resultados fundacionales de la teoría de modelos, el teorema de Löwenheim–Skolem, que nos dice que una teoría con un modelo infinito tiene modelos de **cualquier** cardinal infinito. Demostraremos dos versiones del teorema, la ascendente y la descendente, que combinadas nos daran el teorema completo.

Teorema 3.28. *(Löwenheim–Skolem Ascendente) Sea $\mathcal{L}$ un lenguaje y T una teoría en lenguaje $\mathcal{L}$ con un modelo infinito. Luego, para todo cardinal κ existe una $\mathcal{L}$-estructura M tal que $M \models T$ y $|M| \geq \kappa$*

Demostración. Definamos un nuevo lenguaje $\mathcal{L}'$ que consiste en extender $\mathcal{L}$ con κ símbolos de constante c_i. Notemos que una $\mathcal{L}'$-estructura no es más que una $\mathcal{L}$-estructura con κ elementos destacados.

Sea T' la teoría sobre $\mathcal{L}$ dada por $T' = T \cup \{c_i \neq c_j\}_{i \neq j}$. Es fácil ver que todo subconjunto finito de T es satisfacible: Un subconjunto finito de T' consiste en oraciones $\{\phi_1, ..., \phi_n\}$ de T y una cantidad finita de oraciones de forma $c_i \neq c_j$, luego, si M es un modelo infinito de T, se sigue que M con una cantidad de elementos destacados es modelo del subconjunto finito de T.

Luego por el teorema de compacidad T' debe ser satisfacible. Luego, existe un modelo de T con κ elementos distintos destacados, es decir, existe un modelo de T de cardinal mayor o igual a κ. □

Para demostrar la versión descendente necesitaremos el siguiente lema, que da un criterio muy útil para determinar cuando una subestructura es elementariamente equivalente a su superestructura.

Teorema 3.29. ***(Test de Tarski–Vaught)*** *Sea M una $\mathcal{L}$-estructura y $N \subset M$ una subestructura con la siguiente propiedad: para toda fórmula $\phi(x, \bar{y})$ en lenguaje $\mathcal{L}$ si $M \models \exists x : \phi(x, \bar{n})$ vale que $N \models \exists x : \phi(x, \bar{n})$.*

Luego se sigue que para toda fórmula $\phi(\bar{x})$ en lenguaje $\mathcal{L}$ y tupla $\bar{n} \in N$ vale que $M \models \phi(\bar{n})$ si y solo si $N \models \phi(\bar{n})$. En particular, $M \equiv N$.

Demostración. Sea $\bar{n}$ una tupla en N. La demostración es por inducción en la estructura de $\phi(\bar{x})$:

- Como N es una subestructura de M, se sigue que para toda fórmula atomica $\phi(\bar{x})$ vale que $M \models \phi(\bar{n})$ si y solo si $N \models \phi(\bar{n})$
- Supongamos que $\phi(x) = \neg\psi(x)$. Por hipótesis inductiva, para todo $\bar{n} \in N$ vale que $M \models \phi(\bar{n})$ si y solo $N \models \psi(\bar{n})$. Luego tenemos:
  $M \models \phi(\bar{n})$ sii $M \not\models \psi(\bar{n})$ sii $N \not\models \psi(\bar{n})$ sii $N \models \phi(\bar{n})$
- Supongamos que $\phi(x) = \phi_1(x) \ \wedge \ \phi_2(x)$. Por hipótesis inductiva, para todo $\bar{n} \in N$ vale que:
  $M \models \phi_1(\bar{n})$ si y solo si $N \models \phi_1(\bar{n})$
  $M \models \phi_2(\bar{n})$ si y solo si $N \models \phi_2(\bar{n})$
  Intersecando ambos llegamos a que:
  $M \models \phi_1(\bar{n}) \wedge \ \phi_2(\bar{n})$ si y solo si $N \models \phi_1(\bar{n}) \ \wedge \phi_2(\bar{n})$
- Supongamos que $\phi(x) = \exists y : \psi(y, x)$. Sea $\bar{n}$ una tupla en N. Por nuestra hipótesis, sabemos que si $M \models \exists y : \psi(y, \bar{n})$ se sigue que $N \models \exists y : \psi(y, \bar{n})$. Por otro lado, si $N \models \exists y : \psi(y, \bar{n})$ debe existir un $n_0 \in N$ tal que $N \models \psi(n_0, \bar{n})$ luego por hipótesis inductiva vale que $M \models \exists y : \phi(y, \bar{n})$

□

Teorema 3.30. ***(Löwenheim–Skolem Descendente)*** *Sea $\mathcal{L}$ un lenguaje contable, T una teoría sobre $\mathcal{L}$ y M una $\mathcal{L}$-estructura infinita tal que $M \models T$. Luego para todo cardinal $\kappa < |M|$ existe una $\mathcal{L}$-estructura N tal que $N \models T$ y $|N| = \kappa$.*

Demostración. La siguiente técnica fue introducida por Gödel, para construir modelos de ZFC. Construiremos una subestructura $N \subset M$ de cardinal κ tal que $N \equiv M$. Se sigue que en particular $N \models T$. Construiremos N de forma recursiva, como la union de conjuntos $N_0 \subset N_1 \subset \ ...$ definidos de la siguiente manera:

- N_0 es cualquier subconjunto de M de cardinal κ que contenga a todas las constantes c_M.
- N_{k+1} se define agregando ciertos elementos a N_k, de la siguiente manera: Para toda fórmula $\phi(x, \bar{y})$ en $\mathcal{L}$ y toda tupla $\bar{n} \in N_k$ tal que $M \models \exists x : \phi(x, \bar{n})$, agregamos un n tal que $M \models \phi(m, \bar{n})$.

Observemos que la construcción depende del axioma de elección, ya que en cada paso y para cada fórmula ϕ debemos elegir de forma arbitraria un n tal que $M \models \phi(n, \bar{n})$ entre muchos posibles. Corroboremos que $N = \bigcup N_i$ cumple las propiedades que queremos:

- *(Cardinalidad)* Vamos a ver que para todo k, el conjunto N_k es de cardinal κ, de lo que se sigue que $N = \cup N_k$ es de cardinal κ. Por inducción en k. Sabemos por hipótesis que se cumple para $k = 0$. Para el caso inductivo, notemos que N_{k+1} se construye agregando a N_k, como mucho, un elemento por cada fórmula $\phi(\bar{x})$ en lenguaje $\mathcal{L}$ y por cada tupla $\bar{n} \in N_k$. Por 3.1, hay un número contable de fórmulas en $\mathcal{L}$ y por hipótesis inductiva hay κ tuplas en N_k. Luego:

$$\kappa = |N_k| \leq |N_{k+1}| = |N_k| + \kappa \cdot \aleph_0 = \kappa + \kappa \cdot \aleph_0 = \kappa$$

- *(Subestructura)* Por definición, N_0 contiene todas las constantes c_i^M, luego como $N_0 \subset N$ se sigue que N contiene todas las constantes c_i^M. Ademas, N es cerrado bajo la aplicación de funciones, ya que si f es una funcion en $\mathcal{L}$ y $\bar{n} \in N$, como $N = \bigcup N_i$ se sigue que $\bar{n} \in N_k$ para algun k. Luego, $M \models \exists x : x = f(\bar{n})$ y por la construcción de N_{k+1}, se sigue que $f(\bar{n}) \in N_{k+1} \subset N$.

- *(Equivalencia elemental)* Por ultimo, veamos que $N \equiv M$ mediante el test de Tarski-Vaugh. Sea $\phi(x, \bar{y})$ una fórmula en lenguaje $\mathcal{L}$ y $\bar{n} \in N$ una tupla tal que $M \models \exists x : \phi(x, \bar{n})$. Como $N = \bigcup N_i$, debe haber un k tal que todos los elementos de N pertenecen a N_k. Por la construcción de N_{k+1} debe haber un $n_0 \in N_{k+1}$ tal que $N_{k+1} \models \phi(n_0, \bar{n})$, luego $N \models \phi(n_0, \bar{n})$ y por tanto $N \models \exists x : \phi(x, \bar{n})$.

□

Juntando la dos versiones de Lowenstein-Skolem, podemos demostrar la version completa:

Teorema 3.31. ***(Löwenheim–Skolem)*** *Sea $\mathcal{L}$ un lenguaje contable y T una teoría sobre $\mathcal{L}$. Luego son equivalentes:*

- *T tiene un modelo de cardinal infinito.*
- *T tiene un modelo de **todo** cardinal infinito κ.*

Demostración. La unica dirección que necesita demostración es $(i) \to (ii)$. Por Löwenheim–Skolem ascendente, sabemos que existe una $\mathcal{L}$-estructura M con $|M| \geq \kappa$ tal que $M \models T$. Por Löwenheim–Skolem descendente, se sigue que existe un N con $|N| = \kappa$ tal que $N \models T$. □

La paradoja de Skolem A primera vista el teorema anterior parece simplemente un resultado simpático, pero esconde una paradoja que desconcertó a los lógicos por un tiempo. El problema es el siguiente: La teoría de conjuntos de Zermelo-Frankel es una teoría de primer orden, sobre el lenguaje consistente en una relación binaria $\in$. Luego de ser satisfacible, por el teorema de Lowenheim-Skolem descendente, debe tener un modelo de cardinal contable $(V, \in)$. Pero por otro lado sabemos que ZFC demuestra la existencia de conjuntos incontables, entonces ¿Como puede tener un modelo contable?. La paradoja se resuelve notando que la cardinalidad es una propiedad interna a cada modelo, no externa. Es decir, que el modelo contable $(V, \in)$ demuestre que existen conjuntos no contables significa que existe un conjunto $A \in V$ tal que no hay ninguna biyección entre ω y A **dentro de** V, lo que no descarta que exista una biyección entre A y ω en otro universo de conjuntos.

Quizas el siguiente ejemplo ayude a ejemplificar el espíritu de la construcción de Skolem. Fijemos una teoría axiomatica T suficientemente poderosa, digamos ZFC, y llamemos $\mathbb{R}_T$ al conjunto de los números reales definibles por una fórmula en T. Por supuesto, $\mathbb{R}_T$ contiene a todos los enteros y a todos los racionales. Contiene tambien a π y a e, que pueden definirse como ciertas sumatorias infinitas asi como a todos los números algebraicos. Contiene al resultado de toda integral definida. Contiene la solución de toda ecuación diferencial y de todo problema de probabilidad. Contiene, incluso, a monstruos no computables como la constante Ω de Chaitin. Aunque $\mathbb{R}_T$ es un conjunto muy distinto a $\mathbb{R}$ (¡Tienen distinta cardinalidad!) para toda práctica matemática real $\mathbb{R}_T$ y $\mathbb{R}$ son indistinguibles.

Por ultimo, veamos el siguiente corolario del teorema de Löwenheim–Skolem.

Definición 3.50. Una teoría $\mathcal{L}$ sobre lenguaje $\mathcal{L}$ se dice κ-categorica si tiene un solo modelo de cardinal κ hasta isomorfismo.

Teorema 3.32. ***(Test de Łoś-Vaught)*** *Sea $\mathcal{L}$ un lenguaje y T una teoría sobre $\mathcal{L}$ satisfacible, sin modelos finitos y κ-categorica para algun cardinal infinito κ. Luego T es una teoría completa.*

Demostración. Supongamos que no. Luego existe una oración ϕ tal que $T \cup \{\phi\}$ como $T \cup \{\neg\phi\}$ son satisfacibles, es decir, tienen modelo. Como T no tiene modelos finitos, se sigue que $T \cup \{\phi\}$ y $T \cup \{\neg\phi\}$ tienen modelos infinitos. Luego, por el teorema de Löwenheim–Skolem se sigue que ambas teorias tienen modelos de toda cardinalidad infinita, en particular, existen $\mathcal{L}$-estructuras M y N de cardinal κ tal que $M \models T \cup \{\phi\}$ y $N \models T \cup \{\neg\phi\}$. Como T es una teoría κ-categorica se sigue que $M \cong N$, pero eso es un absurdo porque $M \models \phi$ y $N \not\models \phi$. □

4. Conclusiones

Visto con la claridad que nos da el tiempo, debemos admitir que Berkeley tenía razón en casi todo. Los infinitesimales, tal como se concebían en sus inicios, carecian de todo fundamento lógico. Los infinitesimales, con los que Newton y tantos otros empezaron a descubir los secretos del cosmos, eran un misterio, un artilugio que funcionaba sin que se supiera explicar porque. En lo único que se equivocó Berkeley fue en decretar que era un misterio divino, un misterio que la razón humana nunca sería capaz de resolver.

No solo resolvimos el misterio, los resolvimos dos veces. Una vez con limites y otra vez con ultrafiltros. Aca termina la historia hasta donde sabemos, pero yo creo firmemente que no es el final. Supongo que no voy a presenciarlo, pero estoy convencido de que el misterio de los infinitesimales va ser resuelto no dos veces, sino diez, cincuenta o mil veces, con matemática que nos parecería tan alienigena a nosotros como los ultrafiltros a Newton. Confio ciegamente en la capacidad de los humanos para crear nuevos universos.

Referencias

[1] R. Goldblatt *Lectures on the hyperreals* (1998)

[2] F. Diener, M. Diene *Nonstandard Analysis in Practice* (1995)

[3] P. Erdös, L. Gillman, M. Henriksen *An Isomorphism Theorem for Real-Closed Fields* Annals of Mathematics Vol 61, No 3 (1955)

[4] W. Hodges *Model Theory* (1996)

[5] C.C. Chang, H.J. Keisler *Model Theory* (1977)

[6] J. Rigard *The ultrafilter lemma and its place in mathematics* Bachelor's thesis (2024)

[7] M. Kline *Mathematical Thought From Ancient to Modern Times, volume 1* (1972)