

Teorema del Punto Fijo de Lawvere

Alejandro Sánchez Goiz

14 de enero de 2026

Índice general

Introducción	3
1. Teoría de Categorías	15
1.1. Categorías y tipos de morfismos	15
1.2. Funtores y Transformaciones Naturales	21
1.3. Lema de Yoneda	26

Introducción

El desarrollo de la Teoría de Conjuntos de Cantor y, en general, de las matemáticas en el siglo XIX dio lugar a muchos argumentos que hoy conocemos como “diagonales”. Veamos, como ejemplo, el siguiente teorema.

Teorema (Teorema de la Diagonal de Cantor). *El intervalo real unitario $(0, 1)$ no es numerable. Es decir, no existe una biyección $\phi: \mathbb{N} \rightarrow (0, 1)$.*

Demostración. Procederemos por reducción al absurdo. Supongamos que el conjunto es numerable y que, por tanto, existe una biyección ϕ que nos permite enlistar exhaustivamente todos los elementos del intervalo como una sucesión x_1, x_2, x_3, \dots , donde $x_n = \phi(n)$.

Podemos representar cada número x_n mediante su expansión decimal única:

$$\begin{aligned}x_1 &= 0.a_{11}a_{12}a_{13}a_{14}\dots \\x_2 &= 0.a_{21}a_{22}a_{23}a_{24}\dots \\x_3 &= 0.a_{31}a_{32}a_{33}a_{34}\dots \\&\vdots \\x_n &= 0.a_{n1}a_{n2}a_{n3}a_{nn}\dots\end{aligned}$$

Nuestro objetivo es encontrar un número real $y \in (0, 1)$ que no se encuentre en dicha lista. Definimos $y = 0.d_1d_2d_3\dots$ determinando cada dígito decimal d_n en función de la diagonal de nuestra lista (los términos a_{nn}). La regla de construcción es la siguiente:

$$d_n = \begin{cases} 1 & \text{si } a_{nn} \neq 1 \\ 2 & \text{si } a_{nn} = 1 \end{cases}$$

Analicemos el resultado. Para cualquier índice $n \in \mathbb{N}$, el n -ésimo dígito de nuestro nuevo número y difiere del n -ésimo dígito del número x_n de la lista (es decir, $d_n \neq a_{nn}$).

Esto implica necesariamente que $y \neq x_n$ para todo n . Sin embargo, y es un número real bien definido con expansión decimal de unos y dos, por lo que $y \in (0, 1)$. Hemos encontrado un elemento que escapa a la enumeración ϕ , contradiciendo la hipótesis de que ϕ era sobreyectiva. \square

Podemos observar en la demostración anterior que, para construir el elemento y , recurrimos explícitamente a la diagonal principal de la matriz; es decir, nos fijamos en los términos de la forma a_{nn} .

$$\begin{array}{l}
 0.\textcolor{red}{a}_{11}a_{12}a_{13}a_{14}a_{15} \cdots a_{1n} \cdots \\
 0.a_{21}\textcolor{red}{a}_{22}a_{23}a_{24}a_{25} \cdots a_{2n} \cdots \\
 0.a_{31}a_{32}\textcolor{red}{a}_{33}a_{34}a_{35} \cdots a_{3n} \cdots \\
 0.a_{41}a_{42}a_{43}\textcolor{red}{a}_{44}a_{45} \cdots a_{4n} \cdots \\
 0.a_{51}a_{52}a_{53}a_{54}\textcolor{red}{a}_{55} \cdots a_{5n} \cdots \\
 \vdots \\
 0.a_{n1}a_{n2}a_{n3}a_{n4}a_{n5} \cdots \textcolor{red}{a}_{nn} \cdots \\
 \vdots
 \end{array}$$

La clave del argumento reside en el uso de esta información diagonal para definir un objeto que difiere sistemáticamente de todos los elementos listados. Nótese que el éxito de la construcción depende de la capacidad de “invertir” el valor encontrado en la diagonal. El apelar a esta estructura es lo que conocemos hoy como **argumento diagonal**.

Sin embargo, este esquema lógico es generalizable y trasciende la geometría de una matriz infinita. En muchos contextos, la diagonal no es tan visible, pero el mecanismo subyacente es el mismo: la **autorreferencia**.

Si trasladamos este procedimiento a la Teoría de Conjuntos, sustituyendo la indexación numérica por la relación de pertenencia, el argumento diagonal deja de producir “nuevos elementos” para revelar, en su lugar, contradicciones en la estructura misma de la teoría. Fue precisamente al aplicar este principio de diagonalización sobre la totalidad de los conjuntos que Bertrand Russell formuló, en 1901, la siguiente paradoja.

Paradoja de Russell

La paradoja de Russell concierne a la teoría de conjuntos informal, en donde cualquier propiedad define un conjunto. Para ver la paradoja, es necesario considerar a los conjuntos que se pertenecen a sí mismos, i.e., pensemos en todos los conjuntos X tales que $X \in X$. Consideremos ahora el siguiente conjunto:

$$R = \{X | X \notin X\}$$

Es decir, R es el conjunto de todos los conjuntos que no se pertenecen a sí mismos. Surge entonces la pregunta, ¿ R se pertenece a sí mismo? Si $R \in R$ entonces tenemos $R \notin R$ por definición de R . Por otro lado, si $R \notin R$ entonces es un conjunto que no se pertenece a sí mismo y por tanto $R \in R$. Podemos observar entonces que en ambos casos llegamos a

$$R \in R \iff R \notin R$$

Esta es una clara contradicción, por lo que llegamos a una paradoja.

La Paradoja de Russell nos deja ver la necesidad de formalizar la teoría de conjuntos y que no puede ser el caso de que cualquier propiedad defina a un conjunto. De hacerlo, nuestra teoría se derrumbaría.

En el corazón de la paradoja de Russell está la autoferencia. Construimos un conjunto R tal que al considerarse a sí mismo lleva a una paradoja. Esta paradoja es otro ejemplo de un argumento diagonal. Dicho esto, podemos ver que en ningún momento construimos una matriz ni apelamos a una diagonal. ¿Cómo es entonces que éste es un argumento diagonal?

Pensemos en todos los conjuntos existentes. Organizémoslos tanto de manera vertical como horizontal, i.e.,

	A	B	C	D	\dots
A					
B					
C					
D					
\vdots					

Asignemos valores como sigue: si el elemento horizontal pertenece al elemento vertical, ponemos un 1. En caso contrario, ponemos un 0. Así por ejemplo supongamos que $A \in B$ y que $A \notin A$, tendríamos entonces

	A	B	C	D	\dots
A	0				
B	1				
C					
D					
\vdots					

Nuestra matriz entonces podría lucir algo como el siguiente ejemplo,

	A	B	C	D	E	F	\dots
A	0	0	0	1	0	1	\dots
B	1	1	0	1	0	0	\dots
C	1	0	0	1	1	1	\dots
D	0	0	1	0	1	1	\dots
E	1	0	1	0	1	0	\dots
F	1	1	0	0	1	0	\dots
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\ddots

Aquí, nos interesan los valores de la diagonal principal. Las entradas 0 son los X tales que $X \notin X$ y la teoría informal de conjuntos nos dice que podemos crear el

conjunto de todos los conjuntos cuyo valor en esta diagonal principal es 0. Siendo este un conjunto, ocupa un lugar en esta matriz. ¿Qué valor tiene R en esta diagonal principal? Como vimos, vale 0 si y sólo si vale 1. La existencia de este conjunto, entonces, nos lleva a un absurdo, por lo que concluimos que tal conjunto no puede existir. Como la teoría informal de conjuntos permite la construcción de dicho conjunto, podemos ver que existe un problema fundamental con ésta.

Esta construcción nos permite ver que en un problema de autoreferencia permite ser representado en forma de un elemento de la diagonal principal. En sí, entendemos que la n -ésima columna está *hablando algo* de la n -ésima fila, es decir, en cierto modo pensamos que está *hablando algo de sí mismo* al compartir este valor n en común.

De estos casos de autoreferencia uno de los más famosos y el más rodeado de interés filosófico es el primer teorema de incompletitud de Gödel. Un planteamiento y demostración formal de éste teorema ocuparía demasiado espacio y saldría mucho del interés de ésta introducción, por lo que nos limitaremos dar una explicación breve del teorema, en un esbozo que sigue al dado por Gödel en [Göd92].

La Paradoja de Russell

La Paradoja de Russell concierne a la llamada *Teoría Intuitiva de Conjuntos* (o *Naive Set Theory*), la cual se fundamentaba en el principio de que cualquier propiedad lógica define un conjunto. Para visualizar el conflicto, es necesario considerar conjuntos que puedan pertenecerse a sí mismos (por ejemplo, el conjunto de “todas las ideas” es, en sí mismo, una idea).

Consideremos el conjunto R , definido como la colección de todos los conjuntos que **no** son elementos de sí mismos:

$$R = \{X \mid X \notin X\}$$

Surge entonces la pregunta inevitable: ¿Se pertenece R a sí mismo? Analicemos las dos posibilidades:

- Si $R \in R$, entonces debe cumplir la condición que define al conjunto, es decir, $R \notin R$.
- Si $R \notin R$, entonces cumple la condición de entrada, por lo que debería estar dentro, es decir, $R \in R$.

En ambos casos llegamos a la contradicción:

$$R \in R \iff R \notin R$$

Este resultado evidencia que la teoría intuitiva es inconsistente: no puede ser el caso que cualquier propiedad defina un conjunto. De permitirlo, el edificio lógico de la teoría se derrumbaría.

Ahora bien, hemos afirmado que en el corazón de esta paradoja reside la **autorreferencia** y que se trata de un argumento diagonal. Sin embargo, en la formulación anterior no construimos ninguna matriz ni apelamos a dígitos decimales. ¿Cómo justificamos entonces que este es el mismo fenómeno que observamos con Cantor?

Imaginemos, bajo la hipótesis de la teoría intuitiva, que podemos listar “todos los conjuntos existentes”. Organicémoslos tanto vertical como horizontalmente para formar una tabla de pertenencia:

	A	B	C	D	\dots
A	
B	
C	
\vdots					\ddots

Asignemos valores de verdad a las intersecciones: pondremos un 1 si el conjunto de la fila pertenece al conjunto de la columna, y un 0 en caso contrario. Supongamos, por ejemplo, que $A \notin A$ (diagonal es 0), $A \in B$, etc. Nuestra matriz de incidencia luciría así:

	A	B	C	D	E	\dots
A	0	0	0	1	0	\dots
B	1	1	0	1	0	\dots
C	1	0	0	1	1	\dots
D	0	0	1	0	1	\dots
E	1	0	1	0	1	\dots
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\ddots

Al igual que en el teorema anterior, nuestra atención se centra en la **diagonal principal** (marcada en negritas). Las entradas con valor 0 en la diagonal corresponden a los conjuntos X tales que $X \notin X$. La definición de R es precisamente la instrucción de tomar esos casos: R se define “invirtiendo” la diagonal principal.

$$X \in R \iff \text{Diagonal}(X, X) = 0$$

Dado que R es un conjunto dentro de esta teoría, debe ocupar algún lugar en la lista (digamos, una fila R). ¿Qué valor tiene R en su propia intersección diagonal? Por construcción, debe tener el valor opuesto al que tiene en la diagonal. Es decir, el valor en (R, R) debe ser 0 si y solo si es 1.

Esta visualización confirma que la paradoja es estructuralmente idéntica al argumento de Cantor: el problema surge al intentar introducir dentro de la matriz un objeto definido por la negación de la diagonal de dicha matriz.

El Teorema de Incompletitud de Gödel

De estos casos de autorreferencia, el más célebre y con mayor peso filosófico es el **Primer Teorema de Incompletitud de Gödel**. Un planteamiento y demostración

rigurosa de este teorema ocuparía demasiado espacio y se saldría de los objetivos de esta introducción. Por ello, nos limitaremos a dar un esbozo que sigue la línea argumental presentada por el propio Gödel en [Göd92], resaltando cómo es, en el fondo, otro argumento diagonal.

La idea central y genial de Gödel fue trasladar el problema de la autorreferencia al lenguaje de la aritmética. Para ello, ideó un método de codificación (conocido como *numeración de Gödel*) que asigna a cada fórmula ϕ un número natural único, denotado por $\ulcorner \phi \urcorner$. Esto permite que la aritmética “hable” de sus propias fórmulas.

Gracias a esta codificación, la propiedad metamatemática “ y es una demostración de x ” se puede expresar como una relación aritmética, llamémosla $\text{Dem}(x, y)$. A partir de aquí, podemos definir el concepto de ser demostrable como

$$\text{Dem}(x) \iff \exists y \text{Dem}(x, y).$$

A partir de aquí aparece la diagonal como en los otros dos ejemplos. Gödel definió una función de sustitución, $s(n, m)$, que calcula el código de la fórmula resultante al tomar la fórmula con código n y reemplazar su variable libre por el número m , es decir,

$$s(\ulcorner \phi(x) \urcorner, m) = \ulcorner \phi(m) \urcorner.$$

Esta función s es la herramienta que nos permite “caminar por la diagonal”: si evaluamos $s(n, n)$, estamos pidiéndole intrínsecamente a la fórmula número n que hable del número n . Es el análogo aritmético de mirar la entrada (n, n) en la matriz del argumento de la diagonal de Cantor o en la matriz que visualizamos de pertenencia para la Paradoja de Russell.

Ahora, consideremos la siguiente fórmula:

$$P(x) := \neg \text{Dem}(s(x, x)).$$

Sea k el número de Gödel de esta fórmula, es decir, $k = \ulcorner P(x) \urcorner$. Si evaluamos ahora la diagonal en este punto (calculamos $P(k)$), obtenemos una sentencia G tal que:

$$G \iff \neg \text{Dem}(s(k, k))$$

Pero por definición de nuestra función de sustitución, $s(k, k)$ es precisamente el código de G . Es decir, $\ulcorner G \urcorner = s(k, k)$. Sustituyendo esto, llegamos a:

$$G \iff \neg \text{Dem}(\ulcorner G \urcorner)$$

Esta sentencia G afirma, literalmente: “yo no soy demostrable”.

Si el sistema es consistente, no puede demostrar falsedades, por lo que no puede demostrar G (ya que si lo hiciera, G sería falsa). Pero al no poder demostrarla, lo que G afirma es verdad. Tenemos entonces una sentencia que es verdadera pero no demostrable dentro del sistema.

Este resultado, al igual que los de Cantor y Russell, surge de la capacidad de construir un objeto que niega su propia posición en la diagonal del sistema. El propósito de este trabajo es mostrar cómo la Teoría de Categorías, y en particular el Teorema de Lawvere, unifica todos estos argumentos bajo un mismo principio estructural.

Primer Teorema de Incompletitud de Gödel

Los sistemas formales que proveen los Axiomas de Peano o los de Zermelo-Fraenkel, que son capaces de construir aritmética básica de los naturales, tienen limitaciones internas respecto a qué pueden demostrar. Es decir, existen proposiciones en dichos sistemas que no pueden ser demostradas ni refutadas. Añadir estas proposiciones como axiomas no resolvería el problema: en este nuevo sistema existirían también proposiciones indecidibles, i.e., ni demostrables ni refutables. Este problema no es exclusivo a los dos sistemas mencionados, como veremos en el capítulo 3, una cantidad extensiva de sistemas formales de este tipo son incompletos, en el sentido de que tendrán proposiciones indecidibles. Es importante observar que estas proposiciones indecidibles ocurren también porque suponemos que estos sistemas son consistentes, una hipótesis necesaria para más adelante.

El argumento heurístico que provee Gödel, y que nos permitirá exhibir la diagonalidad del argumento, nace con la idea de que en un sistema formal, es inconsecuente que signos son ocupados para describir el sistema. Es decir, la elección de símbolos como paréntesis “()”, conectivos lógicos “ \vee ”, “ \wedge ”, etc. es arbitraria y, por tanto, es igual de plausible usar números naturales en su lugar. De esta forma, fórmulas en nuestro sistema se convierten en sucesiones finitas de números naturales y demostraciones se vuelven sucesiones finitas de sucesiones finitas de números naturales.

Lo increíble de esta idea es que vuelve problemas metamatemáticos, es decir, problemas que nosotros observamos respecto al sistema, en problemas de números naturales. Por ejemplo, si nosotros usamos el número 0 en lugar del conectivo lógico \vee , entonces una fórmula como “ $x = 0$ ” en el sistema se puede entender también a nivel metamatemático como “ $x = \vee$ ”. Es entonces posible construir dentro del sistema una fórmula $F(v)$, donde v es una sucesión finita de números, que a nivel metamatemáticos nos diga “ v es una fórmula demostrable”. En su demostración, Gödel construye 40 funciones recursivas para llegar a ésta fórmula. Esta fórmula es el interés principal de hacer estos cambios de simbología.

Ahora, para el objetivo en mano, nos concentraremos en fórmulas bien formadas con una sola variable libre, por ejemplo, como la provista arriba “ $x = 0$ ”. Podemos imaginar estas fórmulas con una sola variable libre ordenadas en una sucesión, de modo que sean fácilmente identificables. En esta sucesión, denotamos a la n -ésima entrada por R_n .

Representamos por $\text{Sub}(R_n, m)$ a la fórmula derivada de sustituir todas las instancias de la única variable libre en R_n por el número natural m . A su vez, si una fórmula es demostrable, lo denotamos por $\text{Dem}(F)$. Observemos que esto último es la fórmula mencionada arriba por la cual surge esta necesidad de una numeración de los símbolos, fórmulas, etc. Con estas notaciones en mano, definimos el siguiente conjunto K :

$$K = \{n \in \mathbb{N} \mid \neg \text{Dem}(\text{Sub}(R_n, n))\}$$

Es decir, los elementos en K son los naturales n tales que la fórmula resultante de substituir todas las entradas de la variable libre en la fórmula R_n (que es una fórmula con una sola variable libre) no es demostrable en el sistema formal.

Las nociones que conforman a este conjunto $(R_n, \text{Dem}, \text{Sub})$ son todas definibles en el sistema, por lo que la noción de pertenencia al conjunto K es también definible dentro del sistema. Es decir, existe una fórmula de una sola variable libre, S , tal que

$$\text{Dem}(\text{Sub}(S, n)) \iff n \in K$$

Al ser una fórmula de una sola variable libre, tiene un lugar en la secuencia, por lo que hay un $g \in \mathbb{N}$ para el cual $S = R_g$. Llegados a este punto tenemos ya en nuestras manos una proposición que no es ni demostrable ni refutable en el sistema, a saber

$$\text{Sub}(R_g, g)$$

Antes de ver por qué no es demostrable ni refutable esta proposición, observemos del esbozo dado la presencia de un argumento diagonal. Por un lado, podemos acomodar en columnas a todas las fórmulas de una sola variable libre y por otro, en filas a todos los números n . De ahí podemos asignar f a la entrada con fórmula R_n y número m si la fórmula $\text{Sub}(R_n, m)$ no es demostrable en el sistema y v si, por el contrario, es demostrable. Nuestra matriz entonces podría lucir algo como

	R_1	R_2	R_3	\dots	R_{n-1}	R_n	\dots
1	f	f	f	f	v	v	\dots
2	f	v	v	v	f	v	\dots
3	v	v	f	v	f	f	\dots
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\ddots
$n-1$	v	v	v	v	v	v	\dots
n	f	v	f	v	v	f	\dots
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\ddots

Nuestro conjunto K serían los $n \in \mathbb{N}$ para los cuales el valor de la n -ésima fila en la diagonal principal son f . Entonces, ¿cuál será el valor que presente el número g en la diagonal principal? Sale a la luz la similitud con la paradoja de Russell a través de esta visualización de la matriz. Por supuesto, aquí no lidiamos con una paradoja, sino con una limitación de nuestro sistema.

Retomando el esbozo de la demostración, veamos por que $\text{Sub}(R_g, g)$ no es demostrable ni refutable.

Primero, de ser demostrable, tendríamos que $\text{Dem}(\text{Sub}(R_g, g))$ y, por tanto, $g \notin K$. Pero $S = R_g$ y S es tal que $\text{Dem}(\text{Sub}(S, g)) \iff g \in K$, por lo que tendríamos

$g \in K \iff g \notin K$. Esto es una clara contradicción, pues suponemos que nuestro sistema es consistente, por lo que no puede ser demostrable.

Segundo, de ser refutable, tendríamos que $\neg \text{Dem}(\text{Sub}(R_g, g))$ por lo que por definición de K implicaría que $g \in K$. Como $g \in K$, tenemos $\text{Dem}(\text{Sub}(S, g))$, pero S es R_g , por lo que $\text{Dem}(\text{Sub}(R_g, g))$, de lo que llegamos a que ocurre tanto $\neg \text{Dem}(\text{Sub}(R_g, g))$ como $\text{Dem}(\text{Sub}(R_g, g))$ lo cual es absurdo, porque suponemos que nuestro sistema es consistente y, por tanto, concluimos que no puede ser refutable.

Como tal, hemos construido entonces una proposición que no es ni demostrable ni refutable en nuestro sistema. Podemos ver en la construcción de ésta la noción de autoreferencia. En cierto sentido, la proposición dice a un nivel metamatemático “No soy demostrable en el sistema formal” a través del sentido semántico que le damos a los símbolos.

Por supuesto, el esbozo dado de este Teorema no es bajo ningún motivo una demostración del mismo. Muchos detalles son omitidos y la visualización de la diagonalidad en el argumento de la demostración es mucho más complicado de ver explícitamente. Sin embargo, este esbozo nos permite observar la relación que guarda con los dos ejemplos anteriores y la noción de autoreferencia que lo rodea.

Los tres ejemplos expuestos son considerablemente distintos entre sí en el contenido de sus proposiciones. El primero habla de la diferencia en cardinalidad entre el conjunto de los números naturales y el de los reales; el segundo, de los problemas irremediables en la formulación informal de la teoría de conjuntos; el tercero, de las limitaciones de algunos sistemas formales consistentes que puedan desarrollar aritmética básica. Sin embargo, llegar a vislumbrar estos tres resultados ocupa un método muy similar y, en dos de estos casos, la motivación principal surge de una noción de autoreferencia siendo “transmitida” a la demostración. Este último detalle es el que rodea de mucho misticismo y consideraciones filosóficas a los Teorema de Incompletitud de Gödel y, en general, el que rodea a muchos de estos resultados obtenidos mediante argumentos diagonales de consideraciones más allá de las netamente matemáticas. Si el germen de la autoreferencia puede ser incrustado en el corazón de las matemáticas por un método tan simple, ¿acaso hay una noción más allá de la científica que las limite?

Esta noción fue muy popular en el siglo XX, tras los varios avances que ocurrieron en el desarrollo de fundamentos de la matemática. En 1969, se publicaron notas de William F. Lawvere que silenciosamente derrumbarían todo este misticismo que rodeaba a estos resultados. En su artículo, llamado simplemente *Diagonal Arguments and Cartesian Closed Categories*, observaría que todos estos argumentos son simples resultados de trabajar en tipos particulares de categorías. Citándolo,

The original aim of this article was to demystify the incompleteness theorem of Gödel and the truth-definition theory of Tarski by showing that both are consequences of some very simple algebra in the cartesian-closed setting.

El logro de Lawvere con este trabajo es importante: logra unificar en un solo resultado

algunos de los resultados matemáticos más importantes del siglo XX. Más adelante, este resultado sería conocido como el *Teorema del punto fijo de Lawvere*, uniéndose a una larga lista de teoremas de puntos fijos, de los cuales abundan en diferentes áreas de la matemática y que suelen ser de mucha importancia en el área respectiva.

A pesar de esto, este teorema es muy poco conocido. Una conjetura al respecto de por qué esto puede ser es debido a que la teoría en donde toma lugar este resultado, la teoría de categorías, es relativamente nueva y no muy conocida incluso entre matemáticos. Por ello, a pesar de la significancia del resultado, no hay mucho estudio al respecto de éste.

El presente trabajo tiene por intención presentar este teorema de forma autocontenida, por lo que no se presupone ningún conocimiento previo de teoría de categorías. Todos los resultados y conceptos relevantes para entender el teorema serán presentados en el siguiente capítulo. Este capítulo no es un curso extensivo de la teoría de categorías, solo de las herramientas necesarias para llegar al resultado deseado; el lector interesado en averiguar más a fondo esta teoría puede consultar [Rie16], [Lei16], [Mac97].

El capítulo 2 lidiará con el Teorema del Punto Fijo de Lawvere. Presentamos una versión modernizada y una demostración más detallada que la presentada por Lawvere en [Law06]. Introducimos aquí, como Lawvere hace, el concepto de categorías cartesianas cerradas para hacer énfasis en la necesidad de estar trabajando en éstas para la aplicación del teorema. Hacemos también una explicación más detallada de la contrapositiva al teorema, pues este resultado será muy usado para las aplicaciones que veremos más adelante.

El capítulo 3 abarca aplicaciones del teorema a través de diferentes áreas de la matemática. Este capítulo es el más extenso de todos porque nuestro principal interés es mostrar la utilidad, relevancia e importancia que tiene el Teorema de Lawvere como una forma de unificar distintos resultados en uno solo. Aunque todas estas aplicaciones son tratadas desde una argumentación categórica, las áreas donde nacen estos son de distinta naturaleza y por lo tanto conocimientos previos de éstas son útiles para una comprensión más adecuada. Cada problema es explicado y las demostraciones son lo más detalladas posibles sin que entremos mucho en la teoría necesaria, para permitir a los lectores la comprensión más amplia de todos los resultados. Bibliografía adecuada para entender a fondo cada problema será provista para cada uno de ellos para los interesados.

En el capítulo 4 presentamos las conclusiones de nuestro trabajo. Al final, la lista de la bibliografía en detalle es provista.

A pesar de que los tres ejemplos expuestos difieren sustancialmente en su contenido proposicional —el primero trata sobre cardinalidades infinitas, el segundo sobre inconsistencias en la teoría intuitiva de conjuntos y el tercero sobre las limitaciones deductivas de la aritmética—, todos comparten un método subyacente idéntico. En cada caso, la motivación principal surge de una noción de **autorreferencia** que es “transmitida” a la demostración matemática.

Es precisamente este rasgo autorreferencial el que ha rodeado de un halo de misticismo y especulación filosófica a los Teoremas de Incompletitud de Gödel y, en general, a muchos resultados obtenidos mediante argumentos diagonales. Surge entonces una interrogante natural: si el germen de la autorreferencia puede incrustarse en el corazón de las matemáticas mediante un mecanismo tan simple, ¿existe acaso una limitación extrateórica o metafísica que restrinja el conocimiento científico?

Esta inquietud fue popular durante gran parte del siglo XX, impulsada por la crisis de los fundamentos. Sin embargo, en 1969, William F. Lawvere publicó unas notas que, silenciosa pero contundentemente, derrumbarían este misticismo. En su artículo *Diagonal Arguments and Cartesian Closed Categories*, Lawvere observó que todos estos argumentos no son sino consecuencias naturales de trabajar en tipos particulares de estructuras categóricas. En sus propias palabras:

The original aim of this article was to demystify the incompleteness theorem of Gödel and the truth-definition theory of Tarski by showing that both are consequences of some very simple algebra in the cartesian-closed setting.

El aporte de Lawvere es fundamental: logró unificar en un solo esquema formal algunos de los resultados de imposibilidad más importantes del siglo pasado. Más adelante, este hallazgo sería conocido como el **Teorema del Punto Fijo de Lawvere**, uniéndose a la vasta familia de teoremas de punto fijo que abundan en diversas áreas de la matemática.

No obstante, y a pesar de su trascendencia unificadora, este teorema permanece sorprendentemente poco difundido. Una posible explicación radica en que el lenguaje natural del resultado, la Teoría de Categorías, es una disciplina relativamente joven y, en ocasiones, ajena a la formación matemática tradicional. Por ello, existe un vacío en la literatura divulgativa y educativa respecto a este resultado.

Organización del Trabajo

El presente trabajo tiene por objetivo exponer el Teorema del Punto Fijo de Lawvere de manera autocontenida; por consiguiente, no se presupone ningún conocimiento previo de Teoría de Categorías. La estructura de la tesis es la siguiente:

- El **Capítulo 1** introduce las herramientas categóricas necesarias. Dado que no se trata de un curso extensivo de Teoría de Categorías, nos limitaremos a los conceptos indispensables para llegar al resultado deseado. El lector interesado en profundizar en esta teoría puede consultar [Rie16], [Lei16] y [Mac97].
- El **Capítulo 2** aborda el núcleo del trabajo: el **Teorema del Punto Fijo de Lawvere**. Presentamos una versión modernizada y una demostración más detallada que la original de Lawvere en [Law06]. Introducimos aquí el concepto de *categoría cartesiana cerrada*, enfatizando su papel crucial como el entorno donde estos fenómenos diagonales ocurren naturalmente. Asimismo, se ofrece

una explicación pormenorizada de la contrapositiva del teorema, resultado que será instrumental para las aplicaciones posteriores.

- El **Capítulo 3** explora las aplicaciones del teorema a través de diversas áreas de la matemática. Este es el capítulo más extenso, pues nuestro interés principal es demostrar la utilidad y capacidad unificadora del resultado de Lawvere. Aunque todas las aplicaciones se tratan desde una argumentación categórica, los problemas originales provienen de contextos muy distintos (Lógica, Computación, Análisis), por lo que se provee el contexto necesario para una comprensión adecuada. Las demostraciones se presentan con el mayor detalle posible sin desviarse del hilo conductor.
- Finalmente, en el **Capítulo 4** presentamos las conclusiones de nuestro trabajo, seguidas de la bibliografía detallada.

Capítulo 1

Teoría de Categorías

1.1. Categorías y tipos de morfismos

En el estudio moderno de las matemáticas, la mayoría de áreas empiezan a través de la teoría de conjuntos. Así, por ejemplo, un grupo es un conjunto con cierta estructura añadida; una topología es un conjunto y un subconjunto del conjunto potencia con ciertos axiomas a cumplir; incluso un sistema formal parte de conjuntos de símbolos, axiomas, reglas de formación y de inferencia. El lenguaje matemático es entonces, por lo general, uno que nace de la visión conjuntista. Sin embargo, el teorema del punto fijo de Lawvere sólo puede aplicarse en categorías cartesianas cerradas. Por tanto, todos los ejemplos que veremos en el capítulo 3 serán tratados desde la visión categórica, no la conjuntista. Es necesario entonces introducir los conceptos básicos de la teoría de categorías para poder llegar al teorema deseado y poder visualizar los ejemplos a mostrar desde una nueva percepción, i.e., la categórica. El punto natural para iniciar es con la noción de categoría, dándonos nuestra primera

Definición 1.1.1 (Categoría). Una *categoría* C consta de:

- Una colección de objetos, simbolizada por $\text{Ob}(C)$,
- Para cada A, B en $\text{Ob}(C)$, una colección de morfismos $f: A \rightarrow B$, donde A es denominado el **dominio** y B el **codominio**.
- Para cada A en $\text{Ob}(C)$, un **morfismo identidad**, denotado por 1_A , con dominio y codominio A .
- Para cada par de morfismos $f: A \rightarrow B$, $g: B \rightarrow C$ un **morfismo composición**, $gf: A \rightarrow C$, también denominado el **compuesto de f con g** .

tales que cumplen los siguientes dos axiomas:

- **Identidad:** Para cada $f: A \rightarrow B$, se cumple que $1_A f = f = f 1_B$.
- **Asociatividad:** Para cualesquiera $f: A \rightarrow B$, $g: B \rightarrow C$, $h: C \rightarrow D$, tenemos que $h(gf) = (hg)f$.

Para denotar objetos usamos intercambiamente letras mayúsculas o letras minúsculas, según sea más conveniente en el contexto que se usen.

A los morfismos $f: A \rightarrow B$ también los denominamos como **flechas** o **mapas** entre objetos. A la colección de morfismos entre dos objetos A y B de una categoría \mathcal{C} la simbolizamos por $\text{Hom}_{\mathcal{C}}(A, B)$ o bien simplemente $\text{Hom}(A, B)$ si no existe confusión. Adicionalmente, los morfismos $f: A \rightarrow B$ también los podemos representar como $A \xrightarrow{f} B$.

En el estudio de la teoría de categorías nos veremos constantemente usando *diagramas*. Someramente, un diagrama es una representación visual de objetos y flechas entre estos objetos. Pensado como una gráfica, los objetos actúan como vértices y las flechas como aristas. Existe, como veremos a lo largo de este capítulo, una ventaja en representar cuestiones de la teoría a través de diagramas. Usualmente, nos permitirá una nueva percepción del problema y puede facilitar nuestro entendimiento de éste. Así, por ejemplo, la composición $gf: A \rightarrow C$ puede ser representado como la afirmación de que el siguiente diagrama conmuta:

$$\begin{array}{ccc} & B & \\ f \nearrow & & \searrow g \\ A & \xrightarrow{gf} & C \end{array}$$

Cuando decimos que el diagrama conmuta, nos referimos a que es lo mismo tomar el camino superior que el inferior, es decir, del objeto A pasar al objeto B a través de f y del objeto B pasar a través de g al objeto C , que simplemente pasar del objeto A al objeto C a través de gf .

Para ejemplificar mejor, mostramos los axiomas de identidad y asociatividad a través de diagramas.

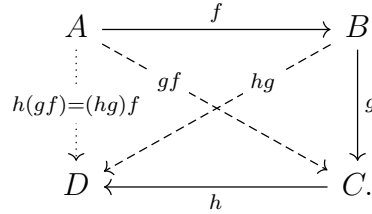
Axioma de identidad

Para toda $f: A \rightarrow B$, el siguiente diagrama es conmutativo

$$\begin{array}{ccccc} A & \xrightarrow{1_A} & A & & \\ & \searrow f & \downarrow f & \searrow f & \\ & & B & \xrightarrow{1_B} & B \end{array}$$

Axioma de asociatividad

Para cualesquiera, $f: A \rightarrow B$, $g: B \rightarrow C$ y $h: C \rightarrow D$, el siguiente diagrama es conmutativo.



Dos observaciones son importantes en la definición de categoría. La primera, la visión conjuntista habitual nos haría pensar que los objetos son en realidad conjuntos y los morfismos, funciones. Es importante observar que esta visión puede privar a uno de una visión más amplia que ofrece ésta teoría y que, en lo general, no es cierto que se pueda hacer dicha comparación.

La segunda observación, se hace mención de *colecciones*, no de conjuntos en la definición. Esto es porque en principio las categorías que se manejan pueden recolectar una cantidad tan grande de objetos y morfismos que no pueden ser conjuntos. Sin embargo, hay casos en que podemos hablar de conjuntos en vez de colecciones para los morfismos de una categoría, y la distinción es lo suficientemente importante para valer una definición:

Definición 1.1.2. Sea \mathcal{C} una categoría. Decimos que es

1. **Localmente pequeña** si para cada par de objetos A, B en $\text{Ob}(\mathcal{C})$, $\text{Hom}_{\mathcal{C}}(A, B)$ es un conjunto.
2. **Pequeña** si tanto la colección de sus objetos como la colección de todas sus flechas son conjuntos.

Exhibimos algunos ejemplos de categorías.

Ejemplos 1.1.3.

- a) **Set** es la categoría cuyos objetos son conjuntos y los morfismos son funciones entre conjuntos.
- b) **Grp** es la categoría cuyos objetos son grupos y los morfismos son homomorfismos entre grupos. Del mismo modo podemos definir categorías para grupos abelianos, **Ab**, para anillos unitarios, **Ring**, etc.
- c) Un sólo grupo G puede ser visto como una categoría \mathbf{G} , que consta de un único objeto $*$ y cuyos morfismos son todos los elementos del grupo. Así, la composición está dada por la operación del grupo y el morfismo identidad de \mathbf{G} es el elemento neutro del grupo.
- d) **Top** es la categoría cuyos objetos son espacios topológicos y los morfismos son las funciones continuas entre espacios topológicos.

- e) Top_* es la categoría cuyos objetos son espacios topológicos con un punto base y los morfismos son funciones continuas que preservan puntos base.
- f) $\text{Mat}_{\mathbb{K}}$ es la categoría cuyos objetos son números naturales y cuyos morfismos son matrices con entradas elementos de un campo \mathbb{K} , es decir, dados números naturales m, n un morfismo $A: m \rightarrow n$ es una matriz de $m \times n$ con entradas en \mathbb{K} . La composición está dada por multiplicación de matrices, y los morfismos identidad son las matrices identidad.
- g) Dado un orden parcial (P, \leq) podemos tratar éste como una categoría \mathbf{P} . Sus objetos son los elementos del orden parcial y dados dos objetos $a, b \in \text{Ob}(\mathbf{P})$, decimos que existe un morfismo $a \rightarrow b$ si y solo si $a \leq b$. Por la propiedad reflexiva del orden parcial, $a \leq a$ para todo $a \in \text{Ob}(\mathbf{P})$ por lo que existen morfismos identidad para cada objeto. La propiedad transitiva del orden define la composición de morfismos.
- h) La categoría Poset tiene como objetos órdenes parciales (P, \leq) y como morfismos funciones que preservan órdenes, es decir, tales que si (P, \leq_P) y (Q, \leq_Q) son órdenes parciales, $f: (P, \leq_P) \rightarrow (Q, \leq_Q)$ es una función tal que para todo $p, q \in P$, $fp \leq_Q fq$ si y sólo si $p \leq_P q$.

En contraste a la teoría de conjuntos, aquí no hablamos de elementos. La pertenencia como tal no es propia en el manejo de una categoría. Es decir, si consideramos un objeto X en la categoría Set , no podemos hablar de un elemento $a \in X$. Existe, en este sentido, una forma de hablar de elementos y pertenencia a través de morfismos. Para ello consideramos las siguientes definiciones:

Definición 1.1.4. Decimos que un objeto A en una categoría \mathbf{C} es

1. **terminal**, si para cualquier objeto X en \mathbf{C} existe un único morfismo con dominio X y codominio A ,
2. **inicial**, si para cualquier objeto X en \mathbf{C} existe un único morfismo con dominio A y codominio X .

Ejemplos 1.1.5.

- a) En la categoría Set , cualquier unitario, i.e., conjunto con un solo elemento $\{x\}$ es un objeto terminal. En particular, el conjunto $\mathbf{1}$ que definimos como $\{\emptyset\}$ es un objeto terminal. En contraste, el único objeto inicial es el conjunto vacío, \emptyset , que definimos como $\mathbf{0}$, por vacuidad sólo existe una función con dominio \emptyset y codominio X para cualquier conjunto X .
- b) En un orden parcial (P, \leq) un mínimo a es tal que para todo $b \in P$, $a \leq b$. Por otro lado, un máximo c es tal que para todo $b \in P$, $b \leq c$. Con estas nociones en mente, podemos observar que un objeto inicial en un orden parcial es un mínimo y, por el contrario, un objeto terminal es un máximo.

Con esta noción en mente, podemos definir un elemento y la pertenencia. Pensemos, por ejemplo, en un elemento a en un conjunto A . En términos categóricos, podemos

definir a a como el morfismo $\mathbf{1} \xrightarrow{a} A$. Tenemos entonces la siguiente definición.

Definición 1.1.6. En la categoría \mathbf{Set} , un **elemento global** de un objeto A es un morfismo $\mathbf{1} \xrightarrow{a} A$.

En particular, podemos entender así la noción de pertenencia, \in , en términos categóricos. Cuando expresamos $a \in A$ es lo mismo que expresar, en forma categórica, que existe un morfismo $a: \mathbf{1} \rightarrow A$.

Pensando ahora en los morfismos, hay propiedades interesantes en estos que merecen ser distinguidas. Tenemos las siguientes definiciones.

Definición 1.1.7. Decimos que un morfismo $f: A \rightarrow B$ es

1. **monomorfismo** o **mono** si para cada par de morfismos $h, g: C \rightarrow A$ tales que $fg = fh$ esto implica $g = h$. Es decir, si el siguiente diagrama conmuta

$$\begin{array}{ccc} C & \xrightarrow{g} & A \\ & \xrightarrow{h} & \\ & & A \xrightarrow{f} B \end{array}$$

entonces, $g = h$.

2. **epimorfismo** o **epi** si para cada par de morfismos $h, g: B \rightarrow C$ tales que $gf = hf$ esto implica $g = h$. Es decir, si el siguiente diagrama conmuta,

$$\begin{array}{ccc} C & \xleftarrow{g} & B \\ & \xleftarrow{h} & \\ & & B \xleftarrow{f} A \end{array}$$

entonces $g = h$.

3. **invertible** si existe $g: B \rightarrow A$ tal que $fg = 1_B$ y $gf = 1_A$. En este caso decimos que f es un **isomorfismo** y que A y B son **isomorfos**. Es decir, existe g tal que el siguiente diagrama conmuta

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ \downarrow 1_A & \nearrow g & \downarrow 1_B \\ A & \xrightarrow{f} & B \end{array}$$

Ejemplos 1.1.8.

- a) En \mathbf{Set} los monomorfismos son las funciones inyectivas, los epimorfismos son las funciones suprayectivas y los isomorfismos son las funciones biyectivas.
- b) En \mathbf{Top} los isomorfismos son los homeomorfismos.

- c) En un orden parcial (P, \leq) , todas las flechas son tanto monomorfismos como epimorfismos. Como las flechas $a \rightarrow b$ se dan si y sólo si $a \leq b$, cada flecha $a \rightarrow b$ es única. Por contraste, los únicos isomorfismos son las flechas identidades. En efecto, si tenemos $a \rightarrow b$ invertible, esto implica que existe una flecha $b \rightarrow a$. Tenemos entonces $a \leq b$ y $b \leq a$, por lo que $a = b$ por antisimetría.
- d) En un grupo G considerado como categoría, todos los morfismos son isomorfismos. Recordemos que un grupo visto como categoría consta de un único objeto $*$ y los elementos del grupo actúan como morfismos, mientras que el elemento neutro actúa como el morfismo identidad del objeto G . Recordemos que para cada elemento g del grupo G , existe un inverso h bajo la operación del grupo tal que $g \cdot h = e$, con \cdot la operación del grupo y e el elemento neutro. De esto vemos que cada morfismo tiene un inverso que da el morfismo identidad, por lo que todos son isomorfismos.

El ejemplo anterior ayuda a exhibir una noción importante que pudiese confundir. En la teoría de conjuntos, una función biyectiva es tanto suprayectiva como inyectiva. Conversamente, si una función es tanto suprayectiva como inyectiva, es biyectiva. En teoría de categorías el análogo es cierto para la primera parte: si un morfismo es isomorfismo es también epimorfismo y monomorfismo. Pero como el inciso c) del ejemplo 1.1.8 nos muestra, el converso no es cierto: un morfismo que es tanto epi como mono no implica que sea isomorfismo.

Procedemos a demostrar la afirmación hecha.

Proposición 1.1.9. *Si $f: A \rightarrow B$ es un isomorfismo, entonces es también un epimorfismo y un monomorfismo.*

Demostración. Demostremos primero que es un monomorfismo. Sean entonces $g, h: C \rightarrow A$ morfismos tales que $fg = fh$. Como f es isomorfismo, existe $f^{-1}: B \rightarrow A$ tal que $f^{-1}f = 1_A$. Tenemos entonces

$$g = 1_A g = (f^{-1}f)g = f^{-1}(fg) = f^{-1}(fh) = (f^{-1}f)h = 1_A h = h$$

por lo tanto $g = h$. Probar que f es epimorfismo es completamente análogo a como procedimos. \square

En el uso de la prueba anterior usamos la notación f^{-1} para hablar de un morfismo inverso para f . Esto se debe a que dicho morfismo inverso, de existir, es único. Demostramos esta afirmación.

Proposición 1.1.10. *Si f es un isomorfismo, entonces existe un único morfismo inverso.*

Demostración. Sea $f: A \rightarrow B$ isomorfismo y $g, h: B \rightarrow A$ inversas de f . Tenemos entonces

$$g = g1_B = g(fh) = (gf)h = 1_A h = h$$

Por lo tanto $g = h$. \square

Por tanto nuestra notación f^{-1} esta debidamente justificada.

Concluimos esta sección construyendo una categoría a partir de una dada \mathbf{C} . Dada una flecha $a \rightarrow b$, podemos pensar en el proceso de invertir la dirección de ésta. En este caso, tendríamos $b \rightarrow a$. El proceso de “revertir todas las flechas” produce lo que conocemos como un proceso *dual*. Este procedimiento nos permite definir la siguiente categoría:

Definición 1.1.11 (Categoría Opuesta). Sea \mathbf{C} una categoría. La categoría opuesta \mathbf{C}^{op} tiene:

- como objetos $\text{Ob}(\mathbf{C}^{\text{op}})$ los mismos objetos que $\text{Ob}(\mathbf{C})$,
- como morfismos los mismos que en \mathbf{C} pero con el orden de las flechas invertidos, dado un morfismo $f: A \rightarrow B$ en \mathbf{C} , el morfismo correspondiente en \mathbf{C}^{op} es $f^{\text{op}}: B \rightarrow A$.
- como morfismos identidad los mismos morfismos identidad que en \mathbf{C} .
- la composición en \mathbf{C}^{op} funciona igual que en \mathbf{C} . Dados morfismos $f: A \rightarrow B$, $g: B \rightarrow C$ su composición $gf: A \rightarrow C$ da lugar la composición $(gf)^{\text{op}} = f^{\text{op}}g^{\text{op}}: C \rightarrow A$.

Es evidente que los axiomas de identidad y asociatividad se siguen en \mathbf{C}^{op} si y solo si la \mathbf{C} es una categoría. Esta construcción y, en general, las construcciones duales que veremos más adelante serán de gran utilidad para el teorema del punto fijo de Lawvere.

1.2. Funtores y Transformaciones Naturales

La noción de categoría nos invita a pensar en objetos matemáticos y los morfismos que permiten preservar su estructura. Consideramos por ejemplo grupos. Los morfismos que preservan su estructura son los homomorfismos; pensamos en espacios topológicos, cuyos morfismos que preservan estructura son funciones continuas. Una categoría es también un objeto matemático, por lo que podemos preguntarnos en qué morfismos preservan la estructura de una categoría. Los funtores son precisamente estos morfismos.

Definición 1.2.1 (Funtor). Sean \mathbf{C}, \mathbf{D} categorías. Un **funtor** $\mathcal{F}: \mathbf{C} \rightarrow \mathbf{D}$ consta de

- objetos $\mathcal{F}C$ en $\text{Ob}(\mathbf{D})$ para cada C en $\text{Ob}(\mathbf{C})$,
- morfismos $\mathcal{F}f: \mathcal{F}A \rightarrow \mathcal{F}B$ en \mathbf{D} para cada morfismo $f: A \rightarrow B$ en \mathbf{C} . Es decir, para cada colección $\text{Hom}_{\mathbf{C}}(A, B)$ corresponde una colección $\text{Hom}_{\mathbf{D}}(\mathcal{F}A, \mathcal{F}B)$.

Tales que:

- para cada A en $\text{Ob}(\mathbf{C})$, $\mathcal{F}(1_A) = 1_{\mathcal{F}A}$,
- para cada $f: A \rightarrow B$, $g: B \rightarrow C$ en \mathbf{C} , $\mathcal{F}g\mathcal{F}f = \mathcal{F}(gf)$

Ejemplos 1.2.2.

1. El funtor $\mathcal{F}: \text{Grp} \rightarrow \text{Set}$ llamado *olvidadizo* pues “olvida” la estructura de grupo y deja tan solo al conjunto de elementos; olvida también que las funciones son homomorfismos y simplemente las convierte a funciones sin esa noción.
2. $\mathcal{P}: \text{Set} \rightarrow \text{Set}$ es un funtor que manda a cada conjunto X a su conjunto potencia $P(X)$ y que manda toda función $f: X \rightarrow Y$ a $\mathcal{P}f: P(X) \rightarrow P(Y)$, esta función manda a cada $A \subset X$ a la imagen $f(A) \in P(Y)$.
3. Dados órdenes parciales (P, \leq_P) , (Q, \leq_Q) un funtor $\mathcal{F}: P \rightarrow Q$ es una función que preserva órdenes, es decir, manda objetos $p, q \in P$ a objetos $\mathcal{F}p, \mathcal{F}q \in Q$ de tal manera que $p \leq_P q$ si y sólo si $\mathcal{F}p \leq_Q \mathcal{F}q$. Observemos que estos funtores son los morfismos en la categoría Poset.
4. El funtor $\pi_1: \text{Top}_* \rightarrow \text{Grp}$ que manda todo espacio topológico X con punto base x a su grupo fundamental $\pi_1(X, x)$ y a toda función continua $f: (X, x) \rightarrow (Y, y)$ a un homomorfismo de grupos $\pi_1(f): \pi_1(X, x) \rightarrow \pi_1(Y, y)$.
5. Sea \mathcal{C} una categoría, el funtor identidad $1_{\mathcal{C}}: \mathcal{C} \rightarrow \mathcal{C}$ manda a objetos C y morfismos $f: C \rightarrow D$ en sí mismos.

También denominamos a este tipo de funtor como **funtor covariante**, para distinguirlo de otro tipo de funtor.

Definición 1.2.3 (Funtor contravariante). Sean \mathcal{C}, \mathcal{D} categorías. Un **funtor contravariante** de \mathcal{C} a \mathcal{D} es un funtor $\mathcal{F}: \mathcal{C}^{\text{op}} \rightarrow \mathcal{D}$. Siendo más explícitos, la diferencia principal de éste y el funtor covariante de \mathcal{C} en \mathcal{D} es que morfismos $f: A \rightarrow B$ en \mathcal{C} son mandados a morfismos $\mathcal{F}: \mathcal{F}B \rightarrow \mathcal{F}A$ en \mathcal{D} . Además, si el codominio del funtor es Set , decimos que el funtor es una **pregavilla**.

Ejemplo 1.2.4. Sea \mathcal{C} una categoría localmente pequeña y C en $\text{Ob}(\mathcal{C})$. El funtor contravariante $\text{Hom}_{\mathcal{C}}(-, C): \mathcal{C}^{\text{op}} \rightarrow \text{Set}$ asigna objetos A a conjuntos $\text{Hom}_{\mathcal{C}}(A, C)$ y dado un morfismo $f: A \rightarrow B$ en $\text{Hom}_{\mathcal{C}}(A, B)$ le asigna a éste un morfismo $\text{Hom}_{\mathcal{C}}(-, C)f: \text{Hom}_{\mathcal{C}}(B, C) \rightarrow \text{Hom}_{\mathcal{C}}(A, C)$ dado por $g \mapsto gf$.

Podemos, en una manera similar, definir un funtor covariante de \mathcal{C} en Set , denotado $\text{Hom}_{\mathcal{C}}(C, -)$, que asigna objetos A en $\text{Ob}(\mathcal{C})$ a conjuntos $\text{Hom}_{\mathcal{C}}(C, A)$ y morfismos $f: A \rightarrow B$ a morfismos $\text{Hom}_{\mathcal{C}}(C, -): \text{Hom}_{\mathcal{C}}(C, A) \rightarrow \text{Hom}_{\mathcal{C}}(C, B)$ dado por $g \mapsto fg$.

Como mencionábamos al inicio de la sección, podemos entender a los funtores como morfismos entre categorías, por lo que naturalmente surge la pregunta de si existen nociones similares a epimorfismo, monomorfismo o isomorfismo para funtores. Tenemos nociones similares en categorías localmente pequeñas, que definimos a continuación.

Definición 1.2.5. Sean \mathcal{C}, \mathcal{D} categorías localmente pequeñas y $\mathcal{F}: \mathcal{C} \rightarrow \mathcal{D}$ un funtor entre ellas. Decimos que el funtor es:

1. **fiel** si para cada A, B en $\text{Ob}(\mathcal{C})$, la función $\text{Hom}_{\mathcal{C}}(A, B) \rightarrow \text{Hom}_{\mathcal{D}}(\mathcal{F}A, \mathcal{F}B)$ es inyectiva;
2. **pleno** si para cada A, B en $\text{Ob}(\mathcal{C})$, la función $\text{Hom}_{\mathcal{C}}(A, B) \rightarrow \text{Hom}_{\mathcal{D}}(\mathcal{F}A, \mathcal{F}B)$ es suprayectiva;
3. **esencialmente suprayectivo** si para cada A en $\text{Ob}(\mathcal{C})$ existe una D en $\text{Ob}(\mathcal{D})$ tal que D es isomórfica a $\mathcal{F}A$.

Notemos aquí que estamos hablando de funciones, inyectividad y suprayectividad, por ello es que requerimos que las categorías en cuestión sean localmente pequeñas. Por otro lado, la noción de isomorfismo es idéntica a como vimos con morfismos.

Definición 1.2.6. Sean \mathcal{C}, \mathcal{D} categorías. Un funtor $\mathcal{F}: \mathcal{C} \rightarrow \mathcal{D}$ es un **isomorfismo** si existe un funtor $\mathcal{G}: \mathcal{D} \rightarrow \mathcal{C}$ tal que $\mathcal{F}\mathcal{G} = 1_{\mathcal{D}}$ y $\mathcal{G}\mathcal{F} = 1_{\mathcal{C}}$. En este caso decimos que \mathcal{F} y \mathcal{G} son **isomorfos** y lo denotamos por $\mathcal{F} \cong \mathcal{G}$.

Si los funtores los podemos entender como morfismos entre categorías, es igual de válido preguntarse sobre morfismos entre funtores. Históricamente, este tipo de morfismos fueron los primeros en ser estudiados y que derivaron en la creación de la teoría de categorías. Antes de definirlos propiamente, veamos un ejemplo en donde su presencia es útil.

Ejemplo 1.2.7. Sea $n \in \mathbb{N}$ y consideremos cualquier anillo conmutativo unitario R . Consideremos ahora las matrices $n \times n$ con entradas en R . Junto con la multiplicación de matrices, este conjunto forma un monoide, que denotamos por $M_n(R)$. Por otro lado, si conservamos sólo la multiplicación del anillo R y olvidamos la suma, obtenemos también un monoide que podemos denotar por $U(R)$. A partir de un anillo R entonces tenemos dos maneras de llegar a un monoide, una definida por $M_n(R)$ y otra por $U(R)$.

Recordemos ahora que cualquier matriz cuadrada A en cualquier anillo R posee un determinante $\det(A)$ con valores en R y que, en particular, cumple las siguientes dos propiedades:

$$\det(A)\det(B) = \det(AB) \quad \det(I_m) = 1_R$$

donde I_m es la matriz identidad de las matrices de $m \times m$ con $m \in \mathbb{N}$ y 1_R es la identidad bajo multiplicación de Q . Observemos que estas mismas dos propiedades son las que determinan un homomorfismo de monoides y que \det toma matrices cuadradas con entradas en R y devuelve valores en R . Por lo tanto podemos considerar el homomorfismo $\det: M_n(R) \rightarrow U(R)$, para cada anillo R .

Consideremos un homomorfismo de anillos conmutativos $f: R \rightarrow Q$. Si preservamos únicamente la estructura de multiplicación, la función en cuestión es ahora un homomorfismo de monoides. Podemos denotar esta nueva función como $U(f): U(R) \rightarrow U(Q)$. Es decir, para cada anillo conmutativo R tenemos un monoide $U(R)$ y para cada homomorfismo de anillos conmutativos $f: R \rightarrow Q$ tenemos un homomorfismo de monoides $U(f): U(R) \rightarrow U(Q)$. Del mismo modo, todo homomorfismo de anillos $f: R \rightarrow Q$ induce un homomorfismo de monoides

$M_n(f): M_n(R) \rightarrow M_n(Q)$. Del estudio que llevamos al momento salta a la luz la noción de funtor. En efecto, podemos considerar la categoría \mathbf{CRing} cuyos objetos son anillos conmutativos unitarios y sus morfismos son homomorfismos de anillos conmutativos y la categoría \mathbf{Mon} cuyos objetos son monoides y cuyos morfismos son homomorfismos de monoides. Tenemos entonces dos funtores $M_n: \mathbf{CRing} \rightarrow \mathbf{Mon}$ y $U: \mathbf{CRing} \rightarrow \mathbf{Mon}$.

Con estas nociones en mente, consideremos el siguiente diagrama

$$\begin{array}{ccc} M_n(R) & \xrightarrow{M_n(f)} & M_n(Q) \\ \downarrow \det(M_n(R)) & & \downarrow \det(M_n(Q)) \\ U(R) & \xrightarrow{U(f)} & U(Q) \end{array}$$

Podemos observar que \det actúa como un morfismo entre estos dos funtores. La primera pregunta que podría salirnos en mente es si el diagrama es conmutativo, i.e., tenemos que $U(f) \det(M_n(R)) = \det(M_n(Q)) M_n(f)$. Podemos verificar rápidamente que el diagrama conmuta considerando los elementos neutros de los monoides y recordando que todas estas funciones son homomorfismos de monoides.

La elección de los valores aquí es arbitraria, por lo que vemos que para la familia de morfismos $(\det: M_n(R) \rightarrow U(R))_{R \in \mathbf{CRing}}$ cualquier diagrama como el mostrado arriba será conmutativo. Si consideramos entonces a $\det: M_n \rightarrow U$, con \det la familia mencionada $(\det: M_n(R) \rightarrow U(R))_{R \in \mathbf{CRing}}$ vemos que para cada homomorfismos de anillos $f: R \rightarrow Q$, el diagrama mostrado arriba conmutará. Esto refleja que la noción de determinante es uniforme para todos los anillos, i.e., no es dependiente de cómo está determinado un anillo en particular. Esta es la idea fundamental detrás de una transformación natural, la cual procedemos a definir formalmente.

Definición 1.2.8 (Transformación Natural). Sean \mathbf{C} y \mathbf{D} categorías, y $\mathbf{C} \xRightarrow[\mathcal{G}]{\mathcal{F}} \mathbf{D}$ funtores. Una **transformación natural** $\alpha: \mathcal{F} \Rightarrow \mathcal{G}$ consiste en una colección de morfismos $\alpha_c: \mathcal{F}c \rightarrow \mathcal{G}c$ para cada objeto c en $\mathbf{Ob}(\mathbf{C})$ tales que para cada morfismo $f: c \rightarrow d$ en \mathbf{C} , el siguiente diagrama de morfismos en \mathbf{D} es conmutativo:

$$\begin{array}{ccc} \mathcal{F}c & \xrightarrow{\mathcal{F}f} & \mathcal{F}d \\ \downarrow \alpha_c & & \downarrow \alpha_d \\ \mathcal{G}c & \xrightarrow{\mathcal{G}f} & \mathcal{G}d \end{array}$$

Si además cada uno de los morfismos α_c es un isomorfismo, entonces decimos que es un **isomorfismo natural** y lo denotamos por $\alpha: \mathcal{F} \cong \mathcal{G}$. A los morfismos α_c también los denotamos como los **componentes** de la transformación natural.

Ejemplos 1.2.9.

- a) Los funtores contravariantes $\mathcal{O}, \mathcal{C}: \text{Top}^{op} \longrightarrow \text{Set}$, donde \mathcal{O} es el funtor que lleva espacios topológicos X a la familia de conjuntos abiertos de X y que a funciones continuas $f: X \rightarrow Y$ las lleva a $\mathcal{O}f: \mathcal{O}Y \rightarrow \mathcal{O}X$, con $\mathcal{O}f$ la función que lleva a un subconjunto $U \subset Y$ a su preimagen bajo f , $f^{-1}(U)$ y \mathcal{C} es el funtor análogo pero llevando los espacios a la familia de conjuntos cerrados de la topología son naturalmente isomórficos, con la transformación natural $\alpha: \mathcal{O} \rightarrow \mathcal{C}$, donde los morfismos α_X para cada $X \in \text{Top}$ llevan a los conjuntos abiertos a su complemento, que es un conjunto cerrado. Para cada uno de estos, su inverso α_X^{-1} está dado como el morfismo que lleva los conjuntos cerrados de X a su complemento, que es un conjunto abierto.
- b) Recordemos que dados órdenes parciales (P, \leq_P) , (Q, \leq_Q) vistos como categorías, un funtor entre ellos es una función que preserva órdenes. Consideremos funtores \mathcal{F}, \mathcal{G} con dominio P y codominio Q . Si para todo $p \in P$ tenemos que $\mathcal{F}p \leq_Q \mathcal{G}p$ podemos definir morfismos $\alpha_p: \mathcal{F}p \rightarrow \mathcal{G}p$ tales que el diagrama

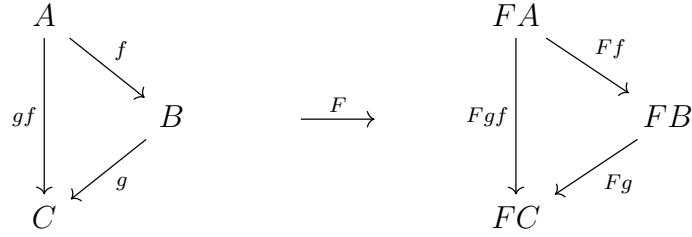
$$\begin{array}{ccc} \mathcal{F}p & \longrightarrow & \mathcal{F}q \\ \alpha_p \downarrow & & \downarrow \alpha_q \\ \mathcal{G}p & \longrightarrow & \mathcal{G}q \end{array}$$

conmuta, gracias a la transitividad del orden parcial (Q, \leq_Q) . Por otro lado, la existencia de morfismos $\alpha_p: \mathcal{F}p \rightarrow \mathcal{G}p$ implica que $\mathcal{F}p \leq_Q \mathcal{G}p$, por lo que vemos que una transformación natural entre los funtores \mathcal{F}, \mathcal{G} existe si y solo si $\mathcal{F}p \leq_Q \mathcal{G}p$ para todo $p \in P$.

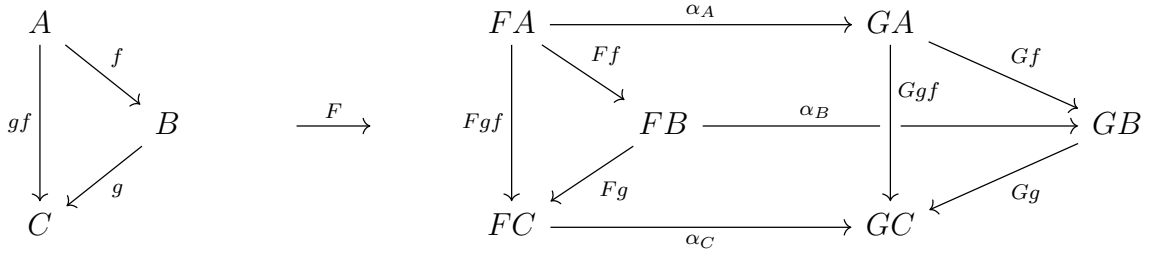
Las transformaciones naturales empiezan a resaltar la importancia de los diagramas para representar la información dada. Nos permiten entender de una manera más directa lo que está ocurriendo a través de los objetos y los morfismos. Por ejemplo, dado el diagrama

$$\begin{array}{ccc} A & & \\ \downarrow gf & \searrow f & \\ & B & \\ & \swarrow g & \\ & C & \end{array}$$

decir que es conmutativo representa la idea de la composición en una categoría. Ahora, un funtor lo podemos entender como el traslado o mapeo de la información de la categoría a otra dada y la propiedad $\mathcal{F}(gf) = \mathcal{F}g\mathcal{F}f$ implica que el mapeo del diagrama conserva la conmutatividad, es decir, el triángulo derecho del siguiente diagrama también es conmutativo.



Finalmente podemos entender a una transformación natural como un mapeo o traslado de la información de un funtor \mathcal{F} a otro dado \mathcal{G} , de tal modo que también el diagrama



es conmutativo. Expresar toda esta información a través de igualdad de composiciones no sólo sería tedioso y tardado, sino también confuso en la entendimiento de las ideas expuestas. Hay una clara ventaja pedagógica en el apoyo de los diagramas para expresar nuestras ideas y, como veremos, nuestras demostraciones. La sección que sigue es un ejemplo perfecto de esto, y es también uno de los resultados centrales en la teoría de categorías.

1.3. Lema de Yoneda

El Lema de Yoneda es uno de los resultados más importantes de la teoría de categorías. La primera aparición del lema en la literatura es con Grothendieck en [Gro60], aunque el resultado se la atribuye a Nobuo Yoneda. Para entender adecuadamente el resultado, necesitamos introducir algunos resultados. Primeramente, recordemos la definición 1.1.6 de un elemento global. Esta noción es más tradicional a como entendemos la pertenencia y lo denotamos global porque no depende de ningún punto de vista particular. Pero podemos extender la noción conjuntista de pertenencia y pensar en puntos de vista o de “referencia” desde los cuales podemos observar a un objeto. En la definición de elemento global usamos de punto de partida el objeto $\mathbf{1}$, pero podemos observar al elemento desde otro punto de partida. Esto nos lleva a la siguiente definición

Definición 1.3.1 (Elemento local). Sea A un objeto no isomórfico a $\mathbf{1}$. Decimos que un morfismo $A \rightarrow B$ es un **elemento local** de B en la **etapa** A .

Esta noción de observar un elemento en cierta “etapa” será útil para entender lo que el lema de Yoneda dice.

Recordemos ahora los funtores vistos en el ejemplo 1.2.4. Nos referimos a estos como **Hom-funtores**, en particular $\text{Hom}(A, -) : \mathcal{C} \rightarrow \text{Set}$ es el Hom-functor covariante y $\text{Hom}(-, A) : \mathcal{C}^{\text{op}} \rightarrow \text{Set}$ es el Hom-functor contravariante. Estos dos funtores son particularmente importantes porque nos permiten hablar de la noción de “representabilidad” que definimos a continuación.

Definición 1.3.2 (Representable). Sea \mathcal{C} una categoría. Un funtor $\mathcal{F} : \mathcal{C} \rightarrow \text{Set}$ es (covariablemente) **representable** si es naturalmente isomórfico a un Hom-functor $\text{Hom}(A, -) : \mathcal{C} \rightarrow \text{Set}$ para algún objeto A en $\text{Ob}(\mathcal{C})$. Del mismo modo, un funtor $\mathcal{G} : \mathcal{C}^{\text{op}} \rightarrow \text{Set}$ es (contravariablemente) **representable** si es naturalmente isomórfico a un Hom-functor contravariante $\text{Hom}(-, A) : \mathcal{C}^{\text{op}} \rightarrow \text{Set}$ para un objeto A en $\text{Ob}(\mathcal{C})$. En ambos casos, tal objeto A es llamado un **objeto representante** del funtor \mathcal{F} o \mathcal{G} , según corresponda. El objeto junto con la transformación natural que induce el isomorfismo son llamados una **representación del funtor** \mathcal{F} o \mathcal{G} , según corresponda.

Ponemos entre paréntesis si es covariante o contravariante porque el dominio del Hom-functor aclara inmediatamente eso. Así, si no existe confusión decimos simplemente que un funtor es representable. Antes de discutir su relación con la definición 1.3.1, veamos algunos ejemplos.

Ejemplos 1.3.3.

- a) El funtor potencia $\mathcal{P} : \text{Set} \rightarrow \text{Set}$ visto en 1.2.2 no es representable, sin embargo, el funtor potencia contravariante $\mathcal{P}^{\text{op}} : \text{Set}^{\text{op}} \rightarrow \text{Set}$ sí es representable. Consideremos el objeto $\mathbf{2} = \{\emptyset, \{\emptyset\}\} = \{\mathbf{0}, \mathbf{1}\}$, y a $\text{Hom}(-, \mathbf{2}) : \text{Set}^{\text{op}} \rightarrow \text{Set}$. La transformación natural $\alpha : \mathcal{P}^{\text{op}} \rightarrow \text{Hom}(-, \mathbf{2})$ está dada por la familia de morfismos $\alpha_A : \mathcal{P}(A) \rightarrow \text{Hom}(A, \mathbf{2})$ para cada conjunto A , que asigna $B \mapsto \chi_B$, donde $B \subseteq A$ y χ_B es la función característica de B ,

$$\chi_B(x) = \begin{cases} \mathbf{0} & \text{si } x \notin B \\ \mathbf{1} & \text{si } x \in B. \end{cases}$$

Es claro ver por qué estos morfismos inducen una biyección. Dado un subconjunto $B \subseteq A$, los elementos que pertenecen a éste son mandados a $\mathbf{1}$, es decir, B no es otra cosa más que la preimagen bajo χ_B de $\mathbf{1}$. Por otro lado, para ver la naturalidad necesitamos que dada una función $f : A \rightarrow B$, el siguiente diagrama conmute

$$\begin{array}{ccc} P(B) & \xrightarrow{f^{-1}} & P(A) \\ \alpha_B \downarrow & & \downarrow \alpha_A \\ \text{Hom}(B, \mathbf{2}) & \xrightarrow{f^*} & \text{Hom}(A, \mathbf{2}), \end{array}$$

donde f^* denota la precomposición. Notemos que el camino derecho nos dice que dado un conjunto $C \subseteq B$, nos da una función clasificadora de la preimagen de C , es decir, nos da $\chi_{f^{-1}(C)}$. Por otro lado, el camino izquierdo nos dice que dado ese conjunto $C \subseteq B$, nos da a través de la precomposición con f^* una función clasificadora dada por $\chi_C f$. Ver la igualdad entonces equivale a que $\chi_C f$ sea una función clasificadora de $f^{-1}(C) \subseteq A$, lo cual es claramente el caso. Por lo tanto, el diagrama conmuta.

De los argumentos anteriores vemos que α es un isomorfismo natural y por lo tanto \mathcal{P}^{op} es representado por $\mathbf{2}$. En realidad, es claro ver que cualquier conjunto con dos elementos representa a este funtor.

Pensemos ahora

- b) Consideremos el funtor identidad de Set , $1_{\text{Set}}: \text{Set} \rightarrow \text{Set}$. Este funtor está representado por el elemento $\mathbf{1}$. Es decir, $\text{Hom}(\mathbf{1}, -) \cong 1_{\text{Set}}$. La transformación natural es bastante directa, para cada X en $\text{Ob}(\text{Set})$, $\text{Hom}(\mathbf{1}, X)$ es el conjunto de todos los elementos de X , es decir, de todos los $\mathbf{1} \xrightarrow{x} X$ que recordando la discusión en 1.1.6 no es más que los $x \in X$, por lo que $\text{Hom}(\mathbf{1}, X)$ es isomórfico a X . Es inmediato ver por qué esta biyección es natural en Set también. Este funtor nos muestra de una manera clara que en Set , dado un conjunto X , $x \in X \iff \mathbf{1} \xrightarrow{x} X$.

Antes de enunciar y demostrar el lema de Yoneda, veamos el siguiente ejemplo.

Ejemplo 1.3.4. Pensemos en un orden parcial (P, \leq) como categoría y en un funtor $\mathcal{F} \rightarrow \text{Set}$. Para facilidad de consideración, supongamos que la cardinalidad del conjunto P es igual o menor a la de los naturales. Consideremos un objeto p . Tenemos que

$$\text{Hom}(p, -) = \{q \in P \mid p \leq q\}$$

De existir una transformación natural $\alpha: \text{Hom}(p, -) \Rightarrow \mathcal{F}$, debe ser tal que para cualquier $q \rightarrow r$, el diagrama

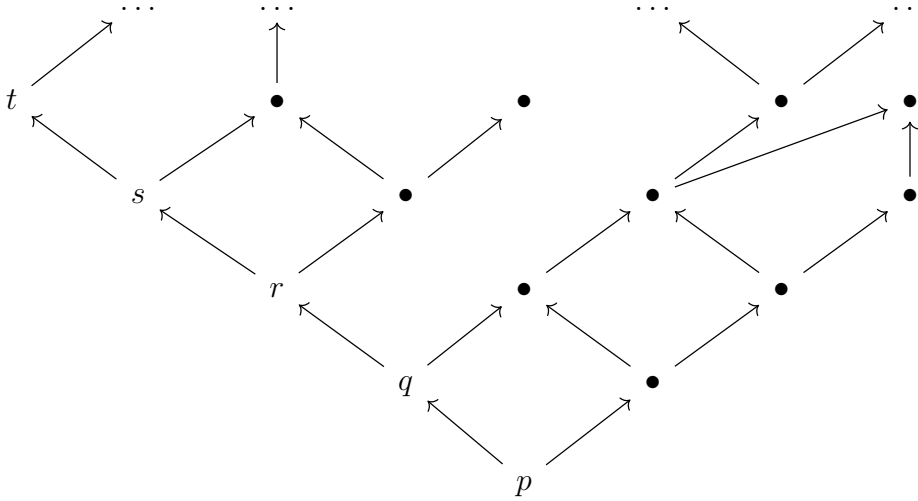
$$\begin{array}{ccc} \text{Hom}(p, q) & \longrightarrow & \text{Hom}(p, r) \\ \alpha_q \downarrow & & \downarrow \alpha_r \\ \mathcal{F}q & \longrightarrow & \mathcal{F}r \end{array}$$

conmuta. Si p y q no son comparables bajo \leq el diagrama no aporta información alguna, del mismo modo que si $q \leq p$, pues sería vacío en su contenido. Si $p \leq q$, $\text{Hom}(p, q)$ es simplemente la flecha $p \rightarrow q$ y del mismo modo $\text{Hom}(p, r)$ es la flecha $p \rightarrow r$, es decir, ambos son unitarios. $\text{Hom}(p, q) \rightarrow \text{Hom}(p, r)$ es simplemente la transitividad del orden parcial en acción: $p \leq q \leq r$ implica $p \leq r$. De esto observamos que el diagrama determina elementos $\alpha_q \in \mathcal{F}q$ y $\alpha_r \in \mathcal{F}r$. Si denotamos por $f_{q,r}$ a la $\mathcal{F}q \rightarrow \mathcal{F}r$ del diagrama, vemos que la naturalidad implica $\alpha_r = f_{q,r} \alpha_q$. Si seguimos

este proceso con una cadena que inicie en p , i.e.,

$$\begin{array}{ccccccc}
 \mathrm{Hom}(p, p) & \longrightarrow & \mathrm{Hom}(p, q) & \longrightarrow & \mathrm{Hom}(p, r) & \longrightarrow & \mathrm{Hom}(p, s) \longrightarrow \dots \\
 \alpha_p \downarrow & & \alpha_q \downarrow & & \alpha_r \downarrow & & \alpha_s \downarrow \\
 \mathcal{F}p & \xrightarrow{f_{p,q}} & \mathcal{F}q & \xrightarrow{f_{q,r}} & \mathcal{F}r & \xrightarrow{f_{r,s}} & \mathcal{F}s \longrightarrow \dots
 \end{array}$$

se vuelve un ejercicio inductivo ver que para todos estos elementos son iguales a α_p . Por supuesto, el orden parcial puede tener muchas ramificaciones en el proceso. Sin embargo, hay un componente α que “absorbe” a todos los demás valores, a saber, $\alpha_p : \mathrm{Hom}(p, p) \rightarrow \mathcal{F}p$. En principio, nuestro orden parcial visualizando a todos los elementos comparables con p podría verse de este estilo



donde los \bullet son otros elementos del orden parcial. Las flechas nos indican que los elementos están bajo la relación \leq . El conjunto pudiera ser denso también, pero como todos estos elementos son comparables con p , la naturalidad nos dará siempre que estos elementos son iguales a $\alpha_p \in \mathcal{F}p$. Por lo tanto, la transformación natural $\alpha : \mathrm{Hom}(p, -) \Rightarrow \mathcal{F}$ está completamente determinada por la elección $\alpha_p \in \mathcal{F}p$. Podemos pensar en este elemento como la imagen del morfismo 1_p bajo la transformación natural $\alpha_p : \mathrm{Hom}(p, p) \rightarrow \mathcal{F}p$. ¿Cuáles son las posibles elecciones de elementos? Vemos que es precisamente los $x \in \mathcal{F}p$.

De este análisis vemos que las transformaciones naturales de $\mathrm{Hom}(p, -)$ al funtor \mathcal{F} están determinadas por los elementos de $\mathcal{F}p$. Esta idea, como veremos, no depende de nuestra elección p o del funtor en particular, ni de haber trabajado en un conjunto parcialmente ordenado.

Volveremos a este ejemplo en un momento, pero ahora enunciamos el teorema más importante en teoría de categorías.

Teorema 1.3.5 (Lema de Yoneda). *Para cualquier funtor $\mathcal{F}: \mathbf{C} \rightarrow \mathbf{Set}$, con \mathbf{C} una categoría localmente pequeña, y cualquier objeto A en $\mathbf{Ob}(\mathbf{C})$, la siguiente función es una biyección:*

$$\begin{aligned}\varphi: \text{Nat}(\text{Hom}(A, -), \mathcal{F}) &\cong \mathcal{F} A, \\ \varphi(\alpha) &= \alpha_A(1_A)\end{aligned}$$

donde $\text{Nat}(\text{Hom}(A, -), \mathcal{F})$ denota la colección de transformaciones naturales entre $\text{Hom}(A, -)$ y \mathcal{F} .

Demostración. Queremos construir una $\phi: \mathcal{F} A \rightarrow \text{Nat}(\text{Hom}(A, -), \mathcal{F})$ que sea la inversa de φ . Para este fin, para cada elemento $x \in \mathcal{F} A$, $\phi(x)$ debe ser una transformación natural, por lo que tenemos que definir una colección de morfismos $\phi(x)_B: \text{Hom}(A, B) \rightarrow \mathcal{F} B$ para cada objeto B en $\mathbf{Ob}(\mathbf{C})$ que definan a $\phi(x): \text{Hom}(A, -) \Rightarrow \mathcal{F}$. Para este fin, definimos para cada $x \in \mathcal{F} A$ y cada objeto B en $\mathbf{Ob}(\mathbf{C})$,

$$\phi(x)_B(f) = \mathcal{F} f(x)$$

recordando que $f: A \rightarrow B$. Para corroborar que $\phi(x)$ es una transformación natural, debemos verificar que para cualquier $g: B \rightarrow C$ con C en $\mathbf{Ob}(\mathbf{C})$, el diagrama

$$\begin{array}{ccc}\text{Hom}(A, B) & \xrightarrow{g_*} & \text{Hom}(A, C) \\ \downarrow \phi(x)_B & & \downarrow \phi(x)_C \\ \mathcal{F} B & \xrightarrow{Fg} & \mathcal{F} C\end{array}$$

conmuta, donde g_* denota la post-composición. Evaluemos el camino izquierdo del diagrama.

$$\begin{aligned}\mathcal{F} g(\phi(x)_B(f)) &= \mathcal{F} g \mathcal{F} f(x) \\ &= \mathcal{F} g f(x) \\ &= \phi(x)_C(gf) \\ &= \phi(x)_C(g_*(f)).\end{aligned}$$

Por lo tanto el diagrama conmuta. Notemos que el paso de la primera igualdad a la segunda se cumple por los axiomas de funtorialidad vistos en 1.2.1.

Recordando que $\varphi(x)$ es una transformación natural y por definición de ϕ tenemos que $\phi\varphi(x) = \phi(x)_A(1_A)$, por definición de $\phi(x)_A$ esto es igual a $\mathcal{F} 1_A(x)$ y por axiomas de funtorialidad esto es $1_{\mathcal{F} A}(x) = x$. Por lo tanto, $\phi\varphi(x) = x$.

Para probar la biyección, nos resta probar que $\varphi\phi(\alpha) = \alpha$. Consideremos un morfismo $f: A \rightarrow B$. Evaluando,

$$\begin{aligned}\varphi(\phi(\alpha))_B(f) &= \varphi(\alpha_A(1_A))_B(f) \\ &= F(f)(\alpha_A(1_A))\end{aligned}$$

Ahora, α es una transformación natural, por lo que el diagrama

$$\begin{array}{ccc} \text{Hom}(A, A) & \xrightarrow{f_*} & \text{Hom}(A, B) \\ \alpha_A \downarrow & & \downarrow \alpha_B \\ \mathcal{F} A & \xrightarrow{Ff} & \mathcal{F} B \end{array}$$

conmuta. La última igualdad, $F(f)(\alpha_A(1_A))$ es el camino izquierdo del diagrama, por lo que es igual a $\alpha_B(f_*(1_A))$. Tenemos,

$$\begin{aligned} \alpha_B(f_*(1_A)) &= \alpha_B(f(1_A)) \\ &= \alpha_B(f) \end{aligned}$$

Por lo tanto $\varphi(\phi(\alpha))_B(f) = \alpha_B(f)$. Como el morfismo fue elegido arbitrariamente concluimos que $\varphi(\phi(\alpha))_B = \alpha_B$, y como B fue elegido arbitrariamente, concluimos que para todo morfismo de la colección que conforman a α y a $\varphi(\phi(\alpha))$ se cumple la igualdad. Como coinciden en todos sus componentes, es el caso que son la misma transformación natural. Por lo tanto, $\varphi(\phi(\alpha)) = \alpha$. De esto concluimos que las funciones ϕ y φ son inversas y por lo tanto ϕ es una biyección. \square

Existe la versión del lema de Yoneda para pregavillas, es decir, tenemos

Corolario 1.3.6 (Contravariante del lema de Yoneda). *Para cualquier funtor $\mathcal{F}: \mathbf{C}^{op} \rightarrow \text{Set}$, con \mathbf{C}^{op} una categoría localmente pequeña, y cualquier objeto A en $\text{Ob}(\mathbf{C})$, la siguiente función es una biyección:*

$$\begin{aligned} \varphi: \text{Nat}(\text{Hom}(-, A), \mathcal{F}) &\cong \mathcal{F} A, \\ \varphi(\alpha) &= \alpha_A(1_A^{op}) \end{aligned}$$

Demostración. La demostración de este teorema es completamente análoga a la demostración del lema de Yoneda en su versión covariante. \square

Podemos extender el lema. La biyección φ de 1.3.5 es una correspondencia natural tanto en A y \mathcal{F} . Pero expresar a qué se refiere esto en particular requiere la introducción de algunos conceptos.

Definición 1.3.7 (Categoría producto). Sean \mathbf{C} y \mathbf{D} categorías. Definimos la categoría producto $\mathbf{C} \times \mathbf{D}$ dada por

- objetos que son parejas ordenadas (c, d) donde c es un objeto en $\text{Ob}(\mathbf{C})$ y d es un objeto en $\text{Ob}(\mathbf{D})$,
- morfismos que son parejas ordenadas $(f, g): (c, d) \rightarrow (e, f)$ con $f: c \rightarrow e$ morfismo en \mathbf{C} y $g: d \rightarrow f$ morfismo en \mathbf{D}

la composición y la identidad están definidas componentemente.

Definición 1.3.8. Dadas categorías \mathbf{C} y \mathbf{D} , definimos la categoría $\mathbf{D}^{\mathbf{C}}$ donde

- los objetos son funtores $\mathcal{F}: \mathbf{C} \rightarrow \mathbf{D}$,
- los morfismos son transformaciones naturales entre funtores $\mathbf{C} \xrightarrow[\mathcal{G}]{\mathcal{F}} \mathbf{D}$
- dado un funtor $\mathcal{F}: \mathbf{C} \rightarrow \mathbf{D}$ la transformación natural identidad $1_{\mathcal{F}}: \mathcal{F} \Rightarrow \mathcal{F}$ está definida por los morfismos $(1_{\mathcal{F}})_c = 1_{\mathcal{F}c}$.
- dadas transformaciones naturales $\alpha: \mathcal{F} \Rightarrow \mathcal{G}$, $\beta: \mathcal{G} \Rightarrow \mathcal{H}$ la composición $\beta\alpha: \mathcal{F} \Rightarrow \mathcal{H}$ está definida en sus componentes $(\beta\alpha)_A = \beta_A\alpha_A$.

Una observación importante de la definición anterior es que para garantizar que ésta es efectivamente una categoría, la composición de los morfismos debe estar bien definida, y esto sólo es el caso si en efecto $\beta\alpha$ es una transformación natural. Consideremos $f: A \rightarrow B$ en \mathbf{C} , la naturalidad de α nos dice que los diagramas

$$\begin{array}{ccc} \mathcal{F}A & \xrightarrow{\alpha_A} & \mathcal{G}A \\ \mathcal{F}f \downarrow & & \downarrow \mathcal{G}f \\ \mathcal{F}B & \xrightarrow{\alpha_B} & \mathcal{G}B \end{array} \quad \begin{array}{ccc} \mathcal{G}A & \xrightarrow{\beta_A} & \mathcal{H}A \\ \mathcal{G}f \downarrow & & \downarrow \mathcal{H}f \\ \mathcal{G}B & \xrightarrow{\beta_B} & \mathcal{H}B \end{array}$$

conmutan. Podemos unir estos dos diagramas y ver que por la naturalidad de cada componente, el rectángulo exterior del diagrama

$$\begin{array}{ccccc} \mathcal{F}A & \xrightarrow{\alpha_A} & \mathcal{G}A & \xrightarrow{\beta_A} & \mathcal{H}A \\ \mathcal{F}f \downarrow & & \downarrow \mathcal{G}f & & \downarrow \mathcal{H}f \\ \mathcal{F}B & \xrightarrow{\alpha_B} & \mathcal{G}B & \xrightarrow{\beta_B} & \mathcal{H}B \end{array}$$

conmuta. De esto podemos ver que efectivamente $\beta\alpha$ es una transformación natural y la composición está bien definida.

Requerimos la construcción de estas dos categorías para hablar de una en particular. Podemos considerar dada una categoría \mathbf{C} localmente pequeña, la categoría $\mathbf{C} \times \mathbf{Set}^{\mathbf{C}}$. Aquí, podemos considerar un funtor $\text{Nat}(\text{Hom}(-, -), -): \mathbf{C} \times \mathbf{Set}^{\mathbf{C}} \rightarrow \mathbf{Set}$ que recibe un objeto (A, \mathcal{F}) y lo manda al conjunto $\text{Nat}(\text{Hom}(A, -), \mathcal{F})$. Es importante notar en este punto, que el hecho de que \mathbf{C} sea localmente pequeña, no garantiza que $\mathbf{Set}^{\mathbf{C}}$ sea localmente pequeña, por lo que no podríamos garantizar en principio que $\text{Nat}(\text{Hom}(A, -), \mathcal{F})$ sea un conjunto. Sin embargo, el lema de Yoneda precisamente garantiza esto, por lo que el codominio de este funtor está bien definido. Podemos considerar otro funtor de $\mathbf{C} \times \mathbf{Set}^{\mathbf{C}}$ a \mathbf{Set} , el funtor $\text{Ev}: \mathbf{C} \times \mathbf{Set}^{\mathbf{C}} \rightarrow \mathbf{Set}$ que manda objetos (A, \mathcal{F}) al conjunto $\mathcal{F}A$. Con estas categorías y estos funtores en vista, podemos entender a la función φ dada en el lema de Yoneda como un isomorfismo natural

$\varphi: \text{Nat}(\text{Hom}(-, -), -) \Rightarrow \text{Ev}$. Es en este sentido que decimos que tanto A como \mathcal{F} son naturales en 1.3.5. Como tal, esto es parte de la totalidad del lema de Yoneda, por lo que lo enunciamos como una “segunda parte”.

Teorema 1.3.9 (Segunda parte del Lema de Yoneda). *El isomorfismo φ visto en 1.3.5 es natural tanto en A como en \mathcal{F} cuando ambos lados son vistos como funtores de $\mathcal{C} \times \text{Set}^{\mathcal{C}}$ a Set .*

Demostración. La demostración consiste en dos partes. Primero, fijemos un funtor $\mathcal{F}: \mathcal{C} \rightarrow \text{Set}$ y consideremos un morfismo $f: A \rightarrow B$ en \mathcal{C} . Consideremos el diagrama

$$\begin{array}{ccc} \text{Nat}(\text{Hom}(A, -), \mathcal{F}) & \xrightarrow{\circ(f^*)} & \text{Nat}(\text{Hom}(B, -), \mathcal{F}) \\ \downarrow \varphi_{A, \mathcal{F}} & & \downarrow \varphi_{B, \mathcal{F}} \\ \mathcal{F} A & \xrightarrow{\mathcal{F} f} & \mathcal{F} B \end{array}$$

□

Aquí, $\varphi_{A, \mathcal{F}}$ define la función definida en 1.3.5 en los valores A y \mathcal{F} y del mismo modo para $\varphi_{B, \mathcal{F}}$; por otro lado, $\circ f^*$ es dada a través de la composición con una transformación natural, $\circ f^*(\alpha) = \alpha f^*: \text{Hom}(A, -) \Rightarrow \text{Hom}(B, -)$. Evaluando el lado derecho del diagrama,

$$\begin{aligned} \varphi_{B, \mathcal{F}}(\circ(f^*)(\alpha)) &= \varphi_{B, \mathcal{F}}(\alpha f^*) \\ &= (\alpha f^*)_B(1_B) \\ &= \alpha_B(f) \\ &\stackrel{*}{=} \alpha_B(f_*(1_A)) \\ &\stackrel{*}{=} \mathcal{F}(f)(\alpha_A(1_A)) \\ &= \mathcal{F}(f)\varphi_{A, \mathcal{F}}(\alpha) \end{aligned}$$

La justificación a las igualdades con un asterisco encima se encuentran en la demostración de 1.3.5. La última igualdad es precisamente el lado izquierdo del diagrama, de lo que concluimos que el diagrama conmuta. Por tanto, la función es natural en A .

Ahora fijemos al objeto A en $\text{Ob}(\mathcal{C})$ y consideremos una transformación natural $\alpha: \mathcal{F} \Rightarrow \mathcal{G}$ en $\text{Set}^{\mathcal{C}}$, es decir, α es un morfismo en dicha categoría. Consideremos el diagrama

$$\begin{array}{ccc} \text{Nat}(\text{Hom}(A, -), \mathcal{F}) & \xrightarrow{\text{Ev}_A} & \text{Nat}(\text{Hom}(A, -), \mathcal{G}) \\ \downarrow & & \downarrow \\ \mathcal{F} A & \xrightarrow{\quad} & \mathcal{G} A \end{array}$$

El lema de Yoneda nos dice que podemos entender cualquier objeto de una categoría a través de cómo se comportan todos los demás objetos. Por eso mismo a los morfismos también los llamamos elementos locales, como vimos en 1.3.1: son puntos de vista de un objeto que queremos entender. Es decir, para saber todo lo que necesitamos de un objeto dado, basta con ver cómo interaccionan los demás objetos de la categoría con éste.

Regresemos al ejemplo 1.3.4. Entendemos ahora, por el lema de Yoneda, que

$$\text{Nat}(\text{Hom}(p, -), \mathcal{F}) \cong \mathcal{F}p$$

Nosotros sabíamos que las transformaciones naturales estaban determinadas por los elementos de $\mathcal{F}p$, pero con el lema de yoneda entendemos también que los elementos de $\mathcal{F}p$ están determinados por las transformaciones naturales y que esta correspondencia es biyectiva.

Bibliografía

- [Göd92] Kurt Gödel. *On Formally Undecidable Propositions of Principia Mathematica and Related Systems*. Dover Publications, 1992.
- [Gro60] Alexander Grothendieck. “Technique de descente et théorèmes d’existence en géométrie algébriques. II. Le théorème d’existence en théorie formelle des modules”. En: *Séminaire Bourbaki* (1960).
- [Law06] F. William Lawvere. “Diagonal arguments and cartesian closed categories with author commentary”. En: *Reprints in Theory and Applications of Categories* (2006).
- [Lei16] Tom Leinster. *Basic Category Theory*. 2016. arXiv: 1612.09375 [math.CT]. URL: <https://arxiv.org/abs/1612.09375>.
- [LR03] F. William Lawvere y Robert Rosebrugh. *Sets for mathematics*. Cambridge University Press, 2003.
- [Mac97] Saunders MacLane. *Categories for the Working Mathematician*. Springer, 1997.
- [Rie16] Emily Riehl. *Category Theory in Context*. Dover Publications, 2016.
- [Yan03] Noson S. Yanofsky. “A Universal Approach to Self-Referential Paradoxes, Incompleteness and Fixed Points”. En: *Bulletin of Symbolic Logic* 9.3 (sep. de 2003), págs. 362-386. ISSN: 1943-5894. DOI: 10.2178/bsl/1058448677. URL: <http://dx.doi.org/10.2178/bsl/1058448677>.