

Autorreferencia más allá de las paradojas: Enlazando el teorema de punto fijo de Lawvere con la autopoiesis

Diego Becerra^{a, b}

^a *Valparaíso Neural Dynamics Laboratory (VaNDaL),
Universidad de Valparaíso, Chile*

^b *Universidad de Santiago de Chile, Chile*

*La poesía es, en esencia, una forma de abstracción
y la matemática es, en gran medida, una forma de imaginación poética*
— Aurora Apolito (2020, traducción propia)

Introducción

La obra de Francisco Varela atraviesa diversas áreas y tópicos. Entre ellos, uno particularmente poco conocido es su incursión en la matemática. Quizá la escasa atención que ha tenido la obra matemática de Varela pueda explicarse por (i) la barrera de acceso que es consecuencia de la menor familiaridad del público (incluida la comunidad científica y filosófica) con la notación y los desarrollos matemáticos de los siglos XIX y XX; y (ii) la mayor difusión de sus tesis biológicas y filosóficas, que de cierto modo eclipsan otras contribuciones, que, como veremos, no por ello son menos fascinantes.

Así, el objetivo de este ensayo es doble: primero, presentar a un público amplio un desarrollo emocionante de la matemática contemporánea, sin presuponer familiaridad con la terminología y notación de la teoría de categorías ni sacrificar rigor; y segundo, conectar dicho desarrollo con la familia de modelos que buscan capturar la dinámica característica de lo vivo —entre ellos, la autopoiesis, pero también los sistemas de metabolismo-reparación (en adelante, sistemas (M,R)) (Rosen, 1958), los conjuntos colectivamente autocatalíticos (Eigen y Schuster, 1977; Kauffman, 1971), y las organizaciones descritas en el formalismo de la teoría de la organización química (Dittrich y di Fenizio, 2007)—.

Allende los ejercicios que resolvimos en clases durante el colegio, la matemática posee una dimensión expresiva que a menudo pasa inadvertida. Quisiera ejemplificar este punto resaltando algunas similitudes que tiene con la poesía: (i) ambas condensan muchas relaciones en pocos signos; (ii) ambas permiten ir más allá de lo que consideramos expresable en el lenguaje natural; y (iii) ninguna de las dos está restringida por cómo (afirmamos que) es el mundo, sino que ambas permiten explorar mundos distintos engendrados a partir de reglas que podemos crear y recrear con bastante (pero no infinita) libertad. Tanto en matemática como en poesía, las restricciones (por ejemplo, de métrica: hexámetro dactílico; de rima: ABAB CDCD; de espacio: 14 líneas) pueden enriquecer y embellecer lo obtenido. En matemática, comúnmente se aceptan restricciones como la no contradicción ($2 + 3 = a$ y $2 + 3 \neq a$ no pueden ser verdaderos a la vez) o el quinto axioma de Euclides (“dada una recta y un punto externo a esta, existe una única recta que pasa por ese punto y que nunca se cruza con la primera recta”). Sin embargo, es posible obtener nuevos mundos, con propiedades inesperadas, modificando o reemplazando dichas restricciones (por ejemplo, los sistemas paraconsistentes en lógica o las geometrías no euclídeas, respectivamente).¹ Por ello, la matemática aplicada —contexto en el que se enmarca

¹ Quien ya esté familiarizado con los ejemplos anteriores quizá aún pueda sorprenderse al descubrir el análisis no estándar (que acepta la existencia de infinitésimos; véase Abraham Robinson) o las probabilidades negativas (véase Espen Haug o Richard Feynman).

el segundo objetivo de este ensayo— es meramente un ápice del poder expresivo de la matemática, el cual exploraremos modestamente a continuación como parte del primer objetivo.

A continuación se presentan las paradojas de autorreferencia y, luego, el argumento de diagonalización de Cantor. Después se conecta lo anterior con el teorema de punto fijo de Lawvere y, por último, se explora la relevancia de los puntos fijos en los tratamientos formales del metabolismo o la dinámica de lo viviente. De estas secciones, la dedicada al teorema de Lawvere es la única matemáticamente densa; por ello, su último párrafo contiene un resumen del argumento que busca hacer más amable su relectura para quienes presenten dificultades con el formalismo.

Este enunciado es falso

Dícese que siglos atrás, en el poblado de Curanilahue vivía un único barbero, el cual afeitaba exclusivamente a todas aquellas personas que no se afeitaban a sí mismas. Entonces, ¿quién afeitaba al único barbero de Curanilahue? Si decimos “pues, él mismo, ¿quién más?”, estaríamos contradiciendo la premisa (ya que, si el barbero se afeitara a sí mismo, sería falso que sólo afeitaba a quienes no se afeitan), pero si el barbero no se afeitaba a sí mismo, entonces sí se afeitaba a sí mismo (porque, recordemos, él afeitaba a todos quienes no se afeitaban a sí mismos).² Hemos dado con la paradoja de Russell; o, más bien, con una popularización de esta, ligeramente modificada (Russell, 1919). Definamos provisionalmente paradoja como una conclusión aparentemente inaceptable derivada de un razonamiento aparentemente aceptable desde premisas aparentemente aceptables (Sainsbury, 2009).

² Quien lea con atención podría contraargumentar “pues ese barbero tenía una barba larguísima, ya que nunca se afeitaba” u otra salida similar, pero notemos que, dada la premisa, toda persona que no se afeitaba a sí misma era afeitada por ese barbero (y, siendo el barbero una persona, la formulación lo incluye). ¿Una invasión descarada de la autonomía corporal? Sin duda; pero la paradoja se mantiene.

En la paradoja original, no hay barberos, poblados, ni personas, sino, conjuntos³ y relaciones entre conjuntos:

(A) Sea R el conjunto de todos los conjuntos que no son miembros de sí mismos. Si R no es miembro de sí mismo, entonces R es miembro de sí mismo. Y si R es miembro de sí mismo, entonces R no es miembro de sí mismo.

La notación matemática formal del siglo XX nos permite condensar esas tres líneas en la siguiente fórmula:

$$(B) R = \{x \mid x \notin x\} \therefore R \in R \Leftrightarrow R \notin R$$

Donde $|$ quiere decir “tal que”, $\notin$ “no pertenece”, $\therefore$ “por lo tanto”, $\in$ “pertenece”, y $\Leftrightarrow$ denota el conector bicondicional (que se lee como “si y solo si”). Pero ¿qué tiene que ver la paradoja de conjuntos con la autopoiesis o con la obra de Varela?

En 1984 Jorge Soto-Andrade y Francisco Varela publicaron un artículo titulado *Self-reference and fixed points: A discussion and an extension of Lawvere's theorem*, sobre el teorema de punto fijo de Lawvere (1969). Dicho teorema conecta una serie de paradojas, argumentos y otros teoremas que capturaron la imaginación filosófica del siglo XX (por ejemplo, Cantor, Russell, Gödel, Turing y Tarski). En el artículo, Soto Andrade y Varela sugieren la importancia de sus resultados para las ciencias computacionales y la biología. No obstante, ya en la conclusión, la única mención a lo viviente es la siguiente: “en el material biológico, la interdependencia interna de los procesos es primordial” (Soto-Andrade y Varela, 1984, p. 17, traducción propia). Para ponderar su impacto en la definición de lo viviente, primero debemos comprender el alcance formal del teorema; por eso, antes de embarcarnos en ese pequeño ejercicio de conexión entre ambos dominios, exploremos los puentes que el teorema original tiende dentro de las matemáticas.

³ Un conjunto es una colección de objetos matemáticos (números, puntos, variables, otros conjuntos, etcétera). Por ejemplo, $A = \{a, b, c, d\}$; $B = \{1, 2, 3\}$; $N = \{1, 2, \dots\}$ son conjuntos.

En 1891 Cantor demostró que hay infinitos más grandes que otros (incidentalmente, estuvo recluido en un hospital psiquiátrico 7 años antes y 8 después de esto). En 1901 Russell encontró la paradoja anteriormente expuesta que amenazaba la consistencia de la entonces naciente teoría de conjuntos. En 1931 Gödel demostró que si un sistema aritmético es consistente, este estará necesariamente incompleto. En otras palabras: admitirá enunciados no falsos que no pueden ser demostrados (5 años después fue recluido en un hospital psiquiátrico, pero esto puede deberse en mayor medida al impacto del asesinato de su antiguo profesor y amigo Moritz Schlick a manos de un exestudiante que posteriormente se afiliaría al partido nazi). En 1933 Tarski postuló su teorema de indefinibilidad de la verdad en sistemas formales. En 1936 Turing y Church demostraron, cada uno por su lado, que el problema de decisión (o *Entscheidungsproblem* en su formulación original) no tiene solución.⁴ Varios filósofos y filósofas pasaron la siguiente mitad del siglo XX tratando de lidiar con (y, en más de un caso, de exagerar) las consecuencias que dichos trabajos tuvieron en la discusión sobre la certidumbre y la posibilidad del conocimiento humano.

¿Qué es lo que tienen entonces en común todos estos resultados? En cada uno de ellos, tal como en el título de esta sección, aparece la autorreferencia.

Infinitos más grandes que otros infinitos

La paradoja de Russell es autorreferente porque aplica la propiedad que define el conjunto R (R contiene cualquier conjunto que no se autocontenga) sobre el mismo conjunto R . Cuando decimos *este enunciado es falso* estamos evaluando el valor veritativo del mismo enunciado mediante el cual realizamos la evaluación. ¿Pero en qué sentido la demostración de Cantor de que existen infinitos más grandes que otros infinitos es autorreferente?

⁴ Para conocer más a fondo las biografías de Cantor, Russell, Gödel, von Neumann y Turing, además de sus principales desarrollos matemáticos, véase Mosterín (2000).

Para entenderlo, primero esbozaremos una versión coloquial de dicha demostración, conocida como el argumento de diagonalización de Cantor.

A modo de motivación, observemos que los números naturales son infinitos (es decir, no es posible pensar en un número concreto que sea el más grande de la sucesión) y que los números reales también lo son. Pero si de cada conjunto tomamos cualquier intervalo acotado por ambos lados — por ejemplo, todos los números entre 5 y 7 (inclusive)— notamos que dicho intervalo extraído de los naturales ya no es infinito (tiene tres elementos), pero cuando es extraído de los reales sigue siendo infinito (entre cualesquiera dos números del intervalo hay otro número real). Esto, sin embargo, no basta para demostrar que los naturales y los reales tienen distinta cantidad de elementos (en adelante, diremos *cardinalidad*), ya que, si consideramos ese mismo intervalo, pero ahora de los números racionales, obtenemos infinitos números sin que esto signifique que haya tantos números racionales como reales en ese o cualquier intervalo.

De hecho, todo toma un giro contraintuitivo cuando nos enteramos de que previamente el mismísimo Cantor demostró que hay tantos números racionales como naturales.⁵ Cabe preguntarse: ¿cómo podemos comparar cardinalidades de conjuntos infinitos? La respuesta es similar a cuando queremos saber cuál de dos bolsas tiene más dulces sin contarlos: abrimos ambas bolsas y extraemos un dulce de cada una en simultáneo, repetidas veces.

⁵ En otras palabras, el tamaño del conjunto de los racionales, $\mathbb{Q}$, es infinito numerable tal como el de $\mathbb{N}$, mientras que el tamaño de $\mathbb{R}$ es un infinito no numerable. ¿Cómo es esto posible? En 1874 Cantor demostró la equinumerosidad de $\mathbb{N}$ y $\mathbb{Q}$ (más bien, de $\mathbb{N}$ y los números algebraicos, que incluyen a $\mathbb{Q}$) ordenando los números racionales como fracciones de la forma a/b en una cuadrícula donde el denominador crece hacia la derecha (todas las columnas tienen el mismo denominador) y el numerador crece hacia abajo (todas las filas tienen el mismo numerador). La secuencia resultante de tomar el zigzag de dicha cuadrícula, empezando desde arriba a la izquierda, es: $\frac{1}{1}, \frac{1}{2}, \frac{2}{1}, \frac{1}{3}, \frac{2}{2}, \frac{3}{1}, \frac{1}{4}, \dots$ Luego, hay una función $f\left(\frac{a}{b}\right) = a + \frac{(a+b-1)(a+b-2)}{2}$ que permite enumerar cada fracción, demostrando así su equinumerosidad con los naturales. El dibujo del método zigzag y una propuesta más sencilla que la original de Cantor pueden consultarse en Almada (2010). Un adelanto: en vez de pensar la cuadrícula resultante como una sobreyección de $\mathbb{N}$ sobre las fracciones, podemos pensarla como una función de emparejamiento $f: \mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$ donde a cada natural se le asigna un par (a, b) de naturales, demostrando suequinumerosidad.

Si se acaban los dulces de una de las bolsas y la otra aún tiene dulces, obtenemos nuestra respuesta. Al caso donde cada dulce extraído de la segunda bolsa tiene al menos una pareja proveniente de la primera, lo denominamos *sobreyección*.⁶ Teniéndolo presente, podemos al fin pasar al argumento de diagonalización de Cantor.

$$\begin{array}{ll}
 s_1 = & 00000000 \dots \\
 s_2 = & 11111111 \dots \\
 s_3 = & 01010101 \dots \\
 s_4 = & 10101010 \dots \\
 s_5 = & 11010110 \dots \\
 s_6 = & 00110110 \dots \\
 s_7 = & 10001000 \dots \\
 s_8 = & 00110011 \dots \\
 \vdots & \vdots \vdots \vdots \vdots \vdots \vdots \vdots \vdots \\
 u = & 10111010 \dots
 \end{array}$$

Figura 1. El argumento de diagonalización de Cantor. Cada fila $s_1, s_2, s_3, \dots$ es una sucesión binaria infinita de la enumeración T . Los dígitos de la diagonal (resaltados) se invierten mediante la función *bit-flip* para construir la sucesión u , que difiere de cada s_n en la posición n y, por lo tanto, no figura en la enumeración.

Sea S el conjunto de todas las sucesiones binarias de largo infinito, y sea $T: \mathbb{N} \rightarrow S$ una enumeración de todos los elementos del conjunto S , es decir, la lista $s_1, s_2, s_3, \dots$. Podemos entonces, construir un elemento $u \in S$ (la sucesión u de la Figura 1) que no exista en la enumeración T de S . Esto se logra transformando el valor de s_n en la diagonal mediante una función α , tal que $\alpha(0) = 1$ y $\alpha(1) = 0$. Podemos llamar a dicha función *bit-flip*, ya que α invierte el valor de cada bit. Como, por construcción, u pertenece a S pero no aparece en la numeración, concluimos que no es posible sobreyectar $\mathbb{N}$ en S a través de T ; ya que por más que busquemos en la lista infinita de sucesiones no hallaremos a u por ninguna parte. En este momento, la

⁶ La definición formal de sobreyección es la siguiente: una función f que asocia cada elemento x del dominio A con al menos un elemento y del codominio B , tal que $f(x) = y$. Hay una propiedad más fuerte, llamada *biyección*, pero, dado que no la utilizaremos en lo que sigue del ensayo, preferí omitirla.

lectora suspicaz dirá: “oye, pero si la lista es infinita, entonces en algún lado tiene que estar 10111010..., no me engañas, don Cantor”. A lo que suponemos que don Cantor respondería: “digamos que la sucesión s_{13125} coincide con u en los primeros diez mil o trece mil bits, y te parece que entonces u podría estar en nuestra lista infinita. Pero recuerda que utilizamos cada una de las sucesiones enumeradas para construir a u , y eso incluye a s_{13125} , por lo que necesariamente s_{13125} debe ser distinta de u en al menos un bit. Quizá pienses que entonces basta encontrar la siguiente sucesión idéntica a u hasta ese bit, digamos, la sucesión $s_{9100507033154}$; pues te adelantaré que también utilizamos esa sucesión para construir a u , por lo que también son diferentes en un bit, el cual está en otra posición de la sucesión.”⁷

Notemos que la cardinalidad que T induce en S , al ser una enumeración, es la misma que la de $\mathbb{N}$: infinita numerable (porque emparejamos uno a uno sus elementos con los números naturales). Del mismo modo, la cardinalidad de S es la misma que la de $\mathbb{R}$: otro conjunto infinito, pero con más elementos que el infinito numerable.

Los puntos fijos delatan la autorreferencia

Lo que la paradoja de Russell y el argumento de Cantor (y el primer teorema de incompletud de Gödel, y el teorema de Tarski, y el problema de decisión de Turing, y las paradojas de viajes en el tiempo, y las máquinas autorreplantes de von Neumann, y...⁸) tienen en común es un esquema que involucra *puntos fijos*. Una función $f(\)$ posee un punto fijo si para alguna evaluación de esta, obtenemos $f(x) = x$ (Figura 2). Entonces, un punto fijo es un elemento que no cambia bajo cierta transformación (en este caso, la función f).

⁷ Algo elegante de este argumento es que podemos saber exactamente dónde cada s_i es distinta de u : en el caso de $s_{9100507033154}$, será en la posición 9 100 507 033 154.

⁸ Más sobre la aplicación del teorema de punto fijo de Lawvere a esos teoremas o paradojas autorreferenciales y varias otras en Yanofsky (2003).

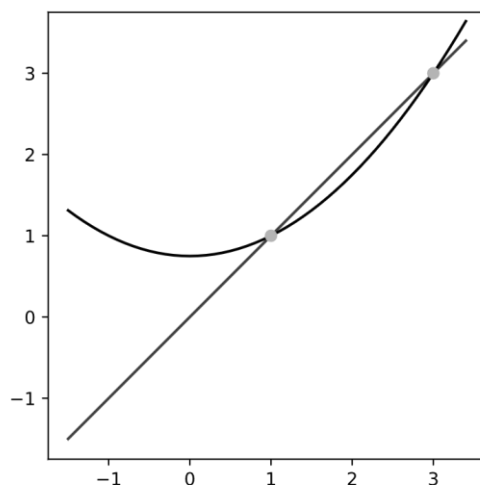


Figura 2. Ejemplo de una función con dos puntos fijos, $f(x) = \frac{1}{4}x^2 + \frac{3}{4}$ (curva) tiene un punto fijo en $x = 1$ y otro en $x = 3$. Vemos que en esos puntos nuestra función se interseca con la función identidad $f(x) = x$ (línea recta).

¿Dónde está el punto fijo en el argumento de diagonalización de Cantor? Cantor demostró que no existe una función sobreyectiva tal que $\mathbb{N} \rightarrow 2^{\mathbb{N}} = \wp(\mathbb{N})$. Denominamos $\wp(\mathbb{N})$ al conjunto potencia⁹ de los números naturales, el cual contiene todos los subconjuntos posibles de $\mathbb{N}$, mientras que $2^{\mathbb{N}}$ es el conjunto de funciones características¹⁰ g que van de los naturales al conjunto que tiene dos elementos, representado como $2 = \{0,1\}$. Dicho de otro modo, describimos cada subconjunto de $\mathbb{N}$ mediante su función característica g . Supongamos que existe una función $\alpha: 2 \rightarrow 2$ que no tiene puntos fijos ($\alpha(0) = 1$ y $\alpha(1) = 0$), podemos llamarla *negación*, siguiendo a Yano-

⁹ Entendamos el conjunto potencia con un ejemplo: la cardinalidad del conjunto $B = \{1, 2, 3\}$ es 3. Entonces, la cardinalidad de $\wp(B)$ es 2^3 : hay ocho elementos en total. En efecto, $\wp(B) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$, donde $\emptyset$ es el conjunto vacío. Sorprendentemente, $\wp(\mathbb{N})$ tiene la misma cardinalidad que $\mathbb{R}$. Para una demostración, véase Rayo (2019).

¹⁰ La función característica g de un conjunto A es aquella que asigna el valor 1 a todo objeto miembro de A y 0 a todo objeto que no sea miembro de A . Podemos pensarla como el conjunto de los valores veritativos o propiedades de los elementos de A . Formalmente, $g(a) = \begin{cases} 1 & \text{si } a \in A \\ 0 & \text{si } a \notin A \end{cases}$.

fsky (2003). Alternativamente, Soto-Andrade y Varela (1984) la llaman *involución no trivial*, lo que a oídos no matemáticos suena como si la estuviesen insultando.¹¹ Luego, sea $T: \{s_1, s_2, s_3, \dots, s_n\}$ una enumeración propuesta de $\wp(\mathbb{N})$; notemos que ahora $s_1, s_2, s_3, \dots$ son subconjuntos de $\mathbb{N}$. La función característica de T es $f: \mathbb{N} \times \mathbb{N} \rightarrow 2$, tal que $f(n, m) = \begin{cases} 1 & \text{si } n \in T_m \\ 0 & \text{si } n \notin T_m \end{cases}$ en otras palabras, f toma un número n y una posición m y permite evaluar si n se encuentra en el m -ésimo lugar de la enumeración T . Finalmente, construimos g como $g(n) = \alpha(f(n, n))$, que es lo mismo que decir que recorrer el camino inferior desde $\mathbb{N}$ hacia el conjunto de dos elementos (Figura 3, izquierda) equivale a recorrer el camino superior.

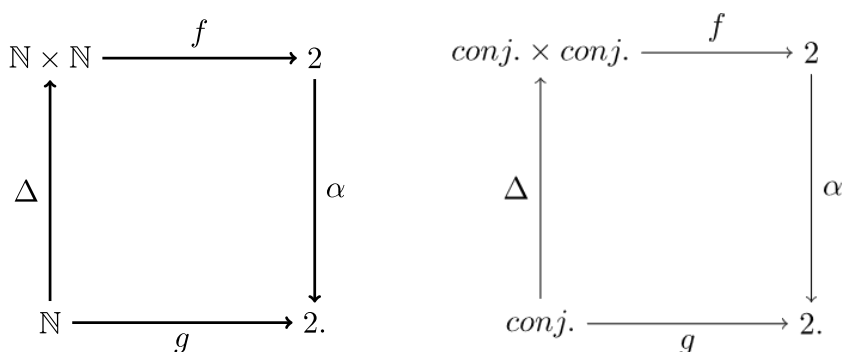


Figura 3. Esquema del argumento de diagonalización de Cantor (izquierda) y de la paradoja de Russell (derecha). Se denomina *diagonal* o *delta* a la función de emparejamiento $\Delta: \mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$ (véase la nota 5).

¿Qué pasa cuando evaluamos autorreferencialmente f ? Al tomar $f(n, n)$ obtenemos que:

$$(C) f(n, n) = g(n) = \alpha(f(n, n))$$

¹¹ La función *negación* no es necesaria para el argumento, sino que es la forma más sencilla de obtener la contradicción: si, por ejemplo, utilizáramos un conjunto de 3 elementos en vez de 2 (podemos representar sus valores como *verdadero*, *falso* y *ni verdadero ni falso* o *sinsentido*), la función sin punto fijo sería un *desplazamiento* ($\alpha(0) = 1, \alpha(1) = 2, \alpha(2) = 0$), no una negación (Domanov, 2025).

Encontramos entonces que $\alpha(f(n, n)) = f(n, n)$: un punto fijo en α , que, por definición, es una función sin punto fijo. Queda así demostrado que no podemos sobreyectar $\mathbb{N}$ en $\wp(\mathbb{N})$.

¡El esquema de la paradoja de Russell (Figura 3, derecha) es idéntico! Basta con reemplazar $\mathbb{N}$ por un universo de conjuntos (*conj.*) y notar que la función característica f de su producto cartesiano, *conj.* $\times$ *conj.* es la siguiente:

$$(D)f(c, t) = \begin{cases} 1 & \text{si } c \in t \\ 0 & \text{si } c \notin t \end{cases}$$

Al evaluarla autorreferencialmente, $f(c, c) = 1$ si un conjunto es miembro de sí mismo, y 0 si no lo es. Llegamos a la misma conclusión: α tiene un punto fijo cuando evaluamos el argumento de f respecto a sí mismo, lo cual contradice la premisa.

Finalmente, tras esta explicación podemos formular el teorema de punto fijo de Lawvere en un lenguaje más abstracto: el de la teoría de categorías, rama de las matemáticas originada en la segunda mitad del siglo XX. Por razones de espacio y atendiendo al objetivo del ensayo, no nos detendremos en el detalle del teorema que enunciaremos a continuación:

Teorema 1. Sea $\mathcal{C}$ una categoría cartesiana cerrada. Decimos que un objeto Y en $\mathcal{C}$ tiene la propiedad de punto fijo si y solo si cada endomorfismo de Y tiene un punto fijo. Si existe un objeto A en $\mathcal{C}$, y una función g sobreyectiva $A \rightarrow Y^A$, entonces Y tiene la propiedad de punto fijo.

Soto-Andrade y Varela (1984) realizan dos extensiones de dicho teorema: (i) postulan que cualquier estructura con propiedad de punto fijo es una “sombra” (retracción) de un dominio reflexivo de orden superior;¹² no obstante, Björner (1985) ofreció rápidamente un contraejemplo; y (ii) plantean

¹² Un dominio reflexivo es una estructura matemática donde no hay distinción entre operadores y operandos; podemos entenderlo como completamente autorreferencial.

que los fractales, objetos autosimilares definidos recursivamente, son instancias de la propiedad de punto fijo; posteriormente, Hayashi (1985) demostró un teorema en línea con este postulado.

Las paradojas tienen cierto aire misterioso y definitivo. Si no podemos fundamentar adecuadamente las matemáticas (Russell con la teoría de conjuntos ingenua; Gödel con las demostraciones en sistemas aritméticos), definir la verdad en sistemas formales (Tarski), ni determinar mediante algoritmos si un programa se detendrá o correrá por siempre (Turing), las proyecciones del conocimiento humano parecen limitadas. Por el contrario, tal como el argumento de diagonalización nos permite imaginar y operar en nuevos mundos,¹³ es posible mostrar que las paradojas de autorreferencia, más que restricciones inexplicables e impredecibles a lo que podemos conocer, son consecuencias de ciertas estructuras matemáticas (las categorías cartesianas cerradas) cuya desmitificación permite abrir la puerta a nuevas investigaciones. El teorema nos indica que los puntos fijos son requisitos para que un sistema formal pueda hablar acerca de sí mismo, y su desarrollo formal nos ayuda a describir ciertos procesos del mundo cuya operación característica debe ser descrita autorreferencialmente: por ejemplo, el metabolismo, la autorreparación y la autoconstrucción como características de lo vivo, transportándonos así al terreno de las matemáticas aplicadas.

La autopoiesis y el teorema de punto fijo

A una definición de cierta función que invoca a la misma función como argumento la llamamos recursiva. Esta requiere uno o más casos base: argumentos donde la función no necesita invocarse a sí misma para producir un resultado y un paso recursivo que reduce todos los argumentos a aplicaciones del caso base. Por ejemplo, la ecuación de recurrencia de Fibonacci,

¹³ En defensa de los números transfinitos de Cantor (es decir, la cardinalidad de conjuntos infinitos de distinto tamaño), el también matemático David Hilbert aseguró, en una conferencia pronunciada en 1925, que “del paraíso que Cantor creó para nosotros, nadie nos echará” (Hilbert, 1926, p. 170).

fib, que nos permite computar la famosa sucesión de Fibonacci,¹⁴ presenta dos casos base: $fib(0) = 0$ y $fib(1) = 1$; y luego, un paso recursivo: $fib(n) = fib(n - 1) + fib(n - 2)$, para $n > 1$. Vemos entonces que $fib(6) = fib(5) + fib(4) = fib(4) + fib(3) + fib(3) + fib(2)$ y que, si desarrollamos todo eso, tenemos $fib(6) = 8$. La recursión se enlaza con la autorreferencia en el paso recursivo de la definición aquí presentada.¹⁵

Pensemos ahora en qué hace que una colección de moléculas sea un ser autopoiético. Tomando la definición de Varela (1979/2025), con una pequeña modificación que bloquea algunos de los contraejemplos y objeciones más comunes en la literatura (Razeto-Barry, 2012), llegamos a que un sistema es autopoiético si y solo si está organizado como una red fisicoquímica de procesos de producción de componentes que produce al menos algunos de los componentes que (I) regeneran y realizan continuamente la red de procesos que los produjo y (II) constituyen el sistema como una unidad concreta en el espacio (mantienen la proximidad entre componentes). Podemos llamar a (I) condición de *autoproducción* y a (II) condición de *autodistinción* (Di Paolo, 2018).

Para poder evaluar si ambas condiciones son autorreferenciales y recursivas, primero hay que formalizarlas, y esto puede tomar dos caminos distintos: el de una teoría en cierto lenguaje matemático y el de un modelo computacional. No obstante, en ambos casos los fenómenos centrales y sus relaciones son abstraídos y representados mediante objetos y relaciones (morfismos) matemáticos. Y aquí nos topamos con que la autopoiesis no tiene una formalización única ni principal —digamos, más popular— en la literatura. Tenemos que elegir entonces, de entre varias opciones, un lenguaje que capture elementos clave del fenómeno (dejando atrás otros) y, en el caso de los modelos, que nos permita obtener predicciones sobre qué

¹⁴ Más bien, en rigor histórico, sucesión de Pingala (पिङ्गल), poeta y matemático indio del siglo IV a.e.c. (Singh, 1985).

¹⁵ Nótese que hablar de funciones recursivas puede ser un abuso del lenguaje: lo que es susceptible o no de ser recursivo es la definición elegida para cierta función, y pueden existir infinitas definiciones equivalentes para una función dada. Por ejemplo, la fórmula de Binet nos permite calcular el n -ésimo número de la sucesión de Fibonacci sin utilizar los números anteriores; en otras palabras, hay una definición no recursiva de $fib(n)$. Agradezco a Fernando Lehue por mostrarme esto.

podría ocurrir en escenarios difíciles de explorar experimentalmente, como la frontera entre lo vivo y lo no vivo. Aquí se vuelve urgente un pequeño desvío para enfatizar una precisión epistemológica: la referencia es una propiedad de los sistemas de signos. Al individualizar regiones, procesos o eventos del mundo, el nombrar refiere. Si abstraemos eso, *el nombrar* refiere al nombrar. Entonces, como la autorreferencia aparece en el lenguaje (las matemáticas aquí incluidas), la autopoiesis será autorreferencial solo en la medida en que es una abstracción acerca de la dinámica de lo vivo. El caso de la recursión es aún más ilustrativo: los procesos recursivos operan sobre los resultados de esos mismos procesos en el pasado solo si concebimos el tiempo como discreto (esto es, compuesto de pequeños pasos de tiempo separados por un intervalo Δt). Un ejemplo de esto es que solemos representar cualquier fenómeno natural descriptible por ecuaciones diferenciales sin solución analítica (como la gran mayoría de nuestros modelos: péndulos, movimiento de fluidos viscosos, actividad eléctrica de nuestros cerebros, etcétera) como un sistema de ecuaciones recursivo. Pero, más que algo fundamental acerca del mundo, esto es una limitación de los métodos numéricos, utilizados extensamente por nuestros programas de cómputo, que nos entregan trayectorias aproximadas en el tiempo sobre cómo una variable cambia en relación con otra u otras.

Las formalizaciones existentes de la autopoiesis aprovechan los desarrollos formales de teorías afines, como los sistemas (M,R), que utilizan ecuaciones funcionales¹⁶ en teoría de categorías (Soto-Andrade et al., 2011; Chastain, 2023); los conjuntos colectivamente autocatalíticos, donde la autopoiesis se caracterizaría como autocatálisis más individuación espacial (McMullin, 2000); y la teoría de la organización química, donde la autopoiesis puede caracterizarse utilizando teoría de conjuntos y ecuaciones diferenciales (Heylighen y Bussieniers, 2023).

¹⁶ Las ecuaciones funcionales son aquellas donde la incógnita no es un parámetro ni una variable, sino una función. Un ejemplo común son las ecuaciones diferenciales: aquellas que relacionan funciones con sus derivadas; en otras palabras, expresan la tasa de cambio instantánea de una variable con respecto a otra.

Todas las formulaciones mencionadas, al tratar con procesos metabólicos (entre otros, como la reparación de membrana y la replicación), tienen en común el capturar dinámicas cíclicas: podemos llegar al mismo punto donde “empezamos” (aunque, en un ciclo, hablar de inicios o destinos no deja de ser una abstracción¹⁷). Podemos ejemplificar la condición de autoproducción tomando un subconjunto de las reacciones metabólicas típicas de una célula. Sea $R: \{R_1, R_2, \dots\}$ el conjunto de todas las reacciones metabólicas de una célula o contexto geoquímico C , y sea $S: \{H^+, H_2, CO_2, HCO_3^-, \dots\}$ el conjunto de todas las especies químicas que participan en R . Tenemos, por ejemplo, $CO_2 + 2H^+ + 2e^- + \text{catalizador} \rightarrow CO + H_2O + \text{catalizador}$.¹⁸ Podemos construir una función con una definición recursiva, $R_i: S \rightarrow S$. Si tomamos la k -ésima reacción del conjunto, tenemos $R_k(x) = a$, donde x y a son, respectivamente, los reactivos y los productos de la reacción. Así llegamos a la concatenación de reacciones $R_i \circ R_{i+1} \circ \dots \circ R_{k-1}(a) = x$; en palabras más simples, si recorremos desde la primera (especificada arbitrariamente) a la reacción inmediatamente anterior a k (esto es, $k - 1$) e ingresamos a como reactivo, obtenemos x como producto y aparece un punto fijo. Un ejemplo más detallado biológicamente que usa el lenguaje de teoría de categorías puede ser consultado en López-Díaz y Gershenson (2025).

¿Es esto suficiente para caracterizar lo viviente? No. Hay fenómenos autocatalíticos inorgánicos que cumplen con dicha propiedad, como la reacción de Belousov–Zhabotinsky (Miyazaki, 2013), por lo que se requiere algo más. Existen modelos de ecuaciones diferenciales de sistemas colectivamente autocatalíticos simplificados y de esta reacción en particular: por ejemplo, el Brusselator (Prigogine y Lefever, 1968)¹⁹ y el Oregonator (Field

¹⁷ Sin embargo, tanto histórica como ontogenéticamente, hubo un inicio de la dinámica: estos modelos formales se utilizan también para estudiar el origen del metabolismo (un pequeño ejemplo: Becerra, 2022).

¹⁸ El catalizador puede ser tanto inorgánico (sulfuro de hierro, oro, plata, cobre, estaño, zinc) como orgánico, como la enzima monóxido de carbono deshidrogenasa (CODH).

¹⁹ La simplicidad del Brusselator merece que nos extendamos un poco más. Formulado en la Université Libre de Bruxelles (de ahí el nombre) por Prigogine y Lefever, está compuesto por apenas cuatro reacciones ($A \rightarrow X$; $2X + Y \rightarrow 3X$; $B + X \rightarrow Y + D$; $X \rightarrow E$), cuya dinámica está descrita por el siguiente sistema de dos ecuaciones diferenciales que indica la tasa de cambio en el tiempo de las concentraciones $[X]$ de las moléculas X y Y : $d[X]/dt =$

y Noyes, 1974), respectivamente. La condición de autodistinción resulta necesaria (y un tanto más difícil de formalizar) si lo que queremos es pasar de la autocatálisis colectiva a la dinámica de lo vivo.

Conclusión

En este pequeño paseo por algunos desarrollos de la matemática del siglo pasado, apenas cubrimos ciertos aspectos intuitivos de los sistemas biológicos: ocurren como ciclos en el tiempo, con dinámicas que reproducen patrones y que, literalmente, reproducen la propia organización y la multiplican. Lawvere formula su teorema mostrando que ciertos teoremas, como los de Gödel y Tarski, más que amenazar impredeciblemente la posibilidad de conocer, son instancias de una categoría cartesiana cerrada donde aparecen puntos fijos. Soto-Andrade y Varela (1984) extendieron dicho teorema a estructuras reflexivas (siguiendo con la jerga técnica: aquellas isomórficas a su espacio de endomorfismos). Cuando un sistema formal lidia con sus propias propiedades, las consecuencias son muchas veces contraintuitivas. En la paradoja del barbero esto es especialmente claro: la función característica del producto cartesiano entre universos de conjuntos que definimos, al evaluar la membresía de cada conjunto respecto a sí mismo, revela un punto fijo donde, por premisa, no debería haber ninguno. Pero esto no debe desalentarnos: la aparición de un punto fijo también nos permite distinguir unos infinitos de otros, y el teorema revela la estructura subyacente común a muchos resultados que antes parecían inconexos. Si una red de moléculas produce al menos algunas de las moléculas que son reactivos para la red, es inevitable encontrar puntos fijos ($f(x) = x$) a la hora de modelarla; cada formalización de la autoproducción y de la autodistinción traerá consigo, del mismo modo, puntos fijos. El desarrollo de las consecuencias de aquello

$A + [X]^2[Y] - B[X] - [X]$; $d[Y]/dt = B[X] - [X]^2[Y]$. El sistema tiene un atractor de punto fijo en $[X] = A$, $[Y] = B/A$, lo que en sistemas dinámicos significa que ese es un estado estacionario: al llegar ahí, no habrá más cambio (esto es, $d[X]/dt = 0$). En cambio, cuando $B > 1 + A^2$, el sistema se torna oscilatorio.

para la reflexión filosófica acerca de lo viviente queda como ejercicio para quien lee.

REFERENCIAS

- Almada, C. (2010). On counting the rational numbers. *International Journal of Mathematical Education in Science and Technology*, 41(8), 1096–1101. <https://doi.org/10.1080/0020739X.2010.500695>
- Apolito, A. (2020). *The problem of scale in anarchism and the case for cybernetic communism*. Center for a Stateless Society. <https://c4ss.org/content/52970>
- Becerra, D. (2022). De las redes de reacciones al surgimiento del metabolismo: Modelamiento del origen de la vida mediante la teoría de la organización química. *Revista de Modelamiento Matemático de Sistemas Biológicos*, 2(3), 9–18.
- Björner, A. (1985). Reflexive domains and fixed points. *Acta Applicandae Mathematicae*, 4, 99–100. <https://doi.org/10.1007/BF02293493>
- Chastain, E. (2023). Formal autopoiesis: Solutions of the classical and extended functional closure equations. *Biosystems*, 226, 104872. <https://doi.org/10.1016/j.biosystems.2023.104872>
- Di Paolo, E. (2018). The enactive conception of life. En A. Newen, L. De Bruin, & S. Gallagher (Eds.), *The Oxford handbook of 4E cognition: Embodied, embedded, enactive and extended* (pp. 71–94). Oxford University Press.
- Dittrich, P., & Di Fenizio, P. S. (2007). Chemical organisation theory. *Bulletin of Mathematical Biology*, 69(4), 1199–1231. <https://doi.org/10.1007/s11538-006-9130-8>
- Domanov, O. A. (2025). Самореферентность и негативность как условия парадоксов [Autorreferencia y negatividad como condiciones de las paradojas]. *Epistemology & Philosophy of Science*, 62(3), 53–58. <https://doi.org/10.5840/eps202562337>
- Eigen, M., & Schuster, P. (1977). The hypercycle. A principle of natural self-organization. *Naturwissenschaften*, 64, 541–565. <https://doi.org/10.1007/BF00450633>
- Field, R., & Noyes, R. (1974). Oscillations in chemical systems. IV. Limit cycle behavior in a model of a real chemical reaction. *The Journal of Chemical Physics*, 60(5), 1877–1884. <https://doi.org/10.1063/1.1681288>
- Hayashi, S. (1985). Self-similar sets as Tarski's fixed points. *Publications of the Research Institute for Mathematical Sciences*, 21(5), 1059–1066. <https://doi.org/10.2977/PRIMS/1195178796>

- Heylighen, F., & Busseniens, E. (2023). Modeling autopoiesis and cognition with reaction networks. *Biosystems*, 230, 104937. <https://doi.org/10.1016/j.biosystems.2023.104937>
- Hilbert, D. (1926). Über das Unendliche [Sobre el infinito]. *Mathematische Annalen*, 95(1), 161–190. <https://doi.org/10.1007/BF01206605>
- Kauffman, S. (1971). Cellular homeostasis, epigenesis, and replication in randomly aggregated macromolecular systems. *Journal of Cybernetics*, 1(1), 71–96. <https://doi.org/10.1080/01969727108545830>
- Lawvere, F. W. (1969). Diagonal arguments and cartesian closed categories. En *Category theory, homology theory and their applications II* (Lecture Notes in Mathematics, Vol. 92, pp. 134–145). Springer. <https://doi.org/10.1007/BFb0080769>
- López-Díaz, A. J., & Gershenson, C. (2025). Closing the loop: How semantic closure enables open-ended evolution? *Journal of the Royal Society Interface*, 22(223), 20250784. <https://doi.org/10.1098/rsif.2025.0784>
- McMullin, B. (2000). Remarks on autocatalysis and autopoiesis. *Annals of the New York Academy of Sciences*, 901(1), 163–174. <https://doi.org/10.1111/j.1749-6632.2000.tb06276.x>
- Miyazaki, J. (2013). Belousov–Zhabotinsky reaction. En S. Kinoshita (Ed.), *Pattern formations and oscillatory phenomena* (pp. 61–83). Elsevier. <https://doi.org/10.1016/C2011-0-07715-3>
- Mosterín, J. (2000). *Los lógicos*. Espasa-Calpe.
- Prigogine, I., & Lefever, R. (1968). Symmetry breaking instabilities in dissipative systems. *The Journal of Chemical Physics*, 48(4), 1695–1700. <https://doi.org/10.1063/1.1668896>
- Rayo, A. (2019). Infinite cardinalities. En *On the brink of paradox: Highlights from the intersection of philosophy and mathematics* (pp. 3–30). MIT Press.
- Razeto-Barry, P. (2012). Autopoiesis 40 years later. A review and a reformulation. *Origins of Life and Evolution of Biospheres*, 42, 543–567. <https://doi.org/10.1007/s11084-012-9297-y>
- Rosen, R. (1958). A relational theory of biological systems. *Bulletin of Mathematical Biophysics*, 20, 245–260. <https://doi.org/10.1007/BF02478302>
- Russell, B. (1919). *Introduction to mathematical philosophy*. Allen & Unwin.
- Sainsbury, R. (2009). *Paradoxes*. Cambridge University Press.
- Singh, P. (1985). The so-called Fibonacci numbers in ancient and medieval India. *Historia Mathematica*, 12(3), 229–244. [https://doi.org/10.1016/0315-0860\(85\)90021-7](https://doi.org/10.1016/0315-0860(85)90021-7)

- Soto-Andrade, J., Jaramillo, S., Gutiérrez, C., & Letelier, J.-C. (2011). Ouroboros avatars: A mathematical exploration of self-reference and metabolic closure. En T. Lenaerts, M. Giacobini, H. Bersini, P. Bourguine, M. Dorigo, & R. Doursat (Eds.), *Advances in artificial life, ECAL 2011: Proceedings of the 11th European conference on the synthesis and simulation of living systems* (pp. 763–770). MIT Press.
- Soto-Andrade, J., & Varela, F. (1984). Self-reference and fixed points: A discussion and an extension of Lawvere's theorem. *Acta Applicandae Mathematicae*, 2, 1–19. <https://doi.org/10.1007/BF01405490>
- Varela, F. J. (2025). *Principles of biological autonomy*. MIT Press. (Trabajo original publicado en 1979)
- Yanofsky, N. (2003). A universal approach to self-referential paradoxes, incompleteness and fixed points. *Bulletin of Symbolic Logic*, 9(3), 362–386. <https://doi.org/10.2178/bsl/1058448677>