Nacimiento, crecimiento exponencial y coincidencias de las palabras de Grundy en las órbitas de Collatz

Miguel Cerdá Bennassar

Octubre de 2026

Resumen

Sea γ la función de los valores de Grundy del juego de Collatz con atajo (A400826) y sea $B(p)$ el número de palabras distintas $\gamma(n)\gamma(C(n))\cdots\gamma(C^{p-1}(n))$ formadas por los números n cuya órbita llega a 1 en exactamente p pasos de la función de Collatz C. En un escrito anterior se mostró que estas palabras forman la parte de la complejidad de las palabras de Grundy que depende del árbol inverso de 1. Aquí se demuestra que las palabras nuevas de cada nivel son exactamente las de los impares de etiqueta 2, y que $B(p+1) - B(p) \geq 2^{\lfloor (p-3)/14 \rfloor}$ para $p \geq 17$; junto con la cota superior que da el lenguaje de tipo finito, resulta $2^{1/14} \leq \liminf B(p)^{1/p} \leq \limsup B(p)^{1/p} \leq 1{,}4656$. Se caracterizan los puntos del árbol en que dos números con la misma palabra se separan, y se demuestra que una misma palabra puede corresponder a tantos números como se quiera: si $m(p)$ es el máximo de esas multiplicidades en el nivel p, entonces $m(p) \geq \frac{1}{2}\, p^{\log_9 2}$ para $p \geq 81$. Las construcciones combinan la libertad de los exponentes entre impares con un control 3-ádico de los restos. Las palabras de Grundy se obtienen de las palabras de subidas y bajadas de las órbitas, de modo que $B(p)$ cuenta las clases en que quedan agrupadas, al leer solo sus etiquetas, las $M(p)$ trayectorias que llegan a 1 en p pasos (A005186); se recoge la relación con la constante $(3+\sqrt{21})/6$, propuesta mediante argumentos heurísticos para el crecimiento de $M(p)$, a la que se aproximan los datos de $B(p)$. Ningún resultado depende de la conjetura de Collatz.

1. Introducción

En [1] se estudia el juego imparcial con atajo asociado a la iteración de Collatz, y se define una función aritmética γ sobre todos los enteros positivos que da el valor de Grundy de cada posición cuya órbita llega a 1; sus valores forman la sucesión A400826 de la OEIS [8]. En [2] se leen esos valores a lo largo de las órbitas de la función de Collatz $C(n) = n/2$ si n es par, $C(n) = 3n+1$ si n es impar, mediante las palabras

$$W_p(n) = \gamma(n)\,\gamma(C(n))\cdots\gamma(C^{p-1}(n)).$$

Allí se demuestra que las palabras que no contienen el factor 00 forman un lenguaje de tipo finito, con $a(p+2)$ palabras de longitud p, donde a es la sucesión A097333, y que la complejidad total se obtiene sumando a esa cantidad la contribución $B(p)$ de los números cuya órbita llega a 1 en exactamente p pasos [2, Teorema 7.1]. La cantidad $B(p)$ se calculó hasta $p = 40$, y quedaron abiertas tres cuestiones [2, sección 9]: el comportamiento de la razón entre $B(p)$ y el número $M(p)$ de esos números, el crecimiento de $B(p)$ y la clasificación de las coincidencias, es decir, de los números distintos que comparten palabra.

Este escrito trata las dos últimas. En la sección 3 se demuestra que las palabras nuevas de cada nivel son exactamente las de los impares de etiqueta 2 (proposición 3.2). En la sección 4, una construcción aritmética con exponentes elegidos de forma adaptativa da $B(p+1) - B(p) \geq 2^{\lfloor (p-3)/14 \rfloor}$, de modo que $B(p)$ crece exponencialmente (teorema 4.3 y corolario 4.4). En la sección 5 se caracterizan los puntos de ambigüedad, en los que dos números con la misma palabra se separan (lema 5.1 y proposición 5.3), y se demuestra que las multiplicidades no están acotadas, con la cota $m(p) \geq \frac{1}{2}\, p^{\log_9 2}$ (teorema 5.5 y corolario 5.6). La sección 6 sitúa estos resultados respecto de las palabras de subidas y bajadas de las órbitas y de lo que se sabe sobre el árbol inverso de Collatz, y la sección 7 recoge datos numéricos y distingue lo demostrado de lo observado.

2. Notación

Para un entero positivo n escribimos $\nu_2(n)$ para su valoración 2-ádica. Según [1, sección 3.1],

$$\gamma(n) = \begin{cases} \nu_2(n) \bmod 2, & n = 2^v,\ v \geq 0, \\ 1 - (\nu_2(n) \bmod 2), & n \text{ par, no potencia de 2}, \\ 2, & n > 1 \text{ impar},\ \nu_2(n+1) \text{ impar}, \\ 1, & n > 1 \text{ impar},\ \nu_2(n+1) \text{ par}. \end{cases}$$

Para $n \geq 1$ sea $d(n)$ el menor $p \geq 0$ con $C^p(n) = 1$, si existe. Para $p \geq 0$ ponemos

$$\mathcal{W}_p = \{W_p(n) : d(n) = p\}, \qquad \mathcal{N}_p(w) = \{n : d(n) = p,\ W_p(n) = w\},$$

$$M(p) = \big|\{n : d(n) = p\}\big|, \qquad B(p) = |\mathcal{W}_p|, \qquad S(p) = B(p{+}1) - B(p), \qquad m(p) = \max_{w \in \mathcal{W}_p} |\mathcal{N}_p(w)|.$$

$W_0(n)$ es la palabra vacía, $\mathcal{W}_0$ consta solo de ella y $\mathcal{N}_0$ de ella es $\{1\}$. Cada m tiene la preimagen $2m$ por C, y la preimagen impar $(m-1)/3$ cuando $m \equiv 4 \pmod 6$ y $m > 4$; para $m = 4$ esa preimagen sería 1, que tiene $d(1) = 0$. Por tanto $n \leq 2^{d(n)}$, y todos los conjuntos anteriores son finitos. La sucesión M es A005186 [8].

Usaremos dos resultados de [2]. Si $\gamma(x) = \gamma(C(x)) = 0$, entonces $x = 1$ [2, Lema 4.1]. Las palabras que no contienen 00 y aparecen en las órbitas son exactamente las que evitan 00, 11, 12, 22 y 202, y hay $a(p+2)$ de longitud p [2, Teoremas 4.4 y 5.1]; la sucesión a cumple $a(n) = a(n-1) + a(n-3)$, y $a(n+1)/a(n)$ tiende a la raíz real $\lambda \approx 1{,}4656$ de $\lambda^3 = \lambda^2 + 1$.

3. Nacimiento de las palabras

Lema 3.1 (Duplicación)**.** *Para todo $x \geq 1$, $\gamma(2x) = \operatorname{mex}\{\gamma(x)\} \in \{0, 1\}$. Para $p \geq 0$ y $w \in \mathcal{W}_p$, todos los $2n$ con $n \in \mathcal{N}_p(w)$ tienen la misma palabra ℓw, con $\ell = \operatorname{mex}\{\gamma(n)\}$, y $d(2n) = p + 1$. La aplicación $w \mapsto \ell w$ es inyectiva, su imagen son las palabras de los números pares N con $d(N) = p + 1$, y ninguna de ellas empieza por 2. En particular, $|\mathcal{N}_{p+1}(\ell w)| \geq |\mathcal{N}_p(w)|$, y $m(p+1) \geq m(p)$.*

Demostración. La igualdad $\gamma(2x) = \operatorname{mex}\{\gamma(x)\}$ es una de las ecuaciones locales que cumple γ [1, sección 3.1]. Todos los $n \in \mathcal{N}_p(w)$ tienen la misma primera letra $\gamma(n)$ si $p \geq 1$, y para $p = 0$ el único es $n = 1$; por tanto $\gamma(2n)$ es la misma para todos. Como $C(2n) = n$, se tiene $d(2n) = p + 1$ y $W_{p+1}(2n) = \gamma(2n)\, W_p(n)$. La inyectividad se debe a que w es el sufijo de ℓw, y todo par N con $d(N) = p + 1$ es de la forma $2n$ con $d(n) = p$. □

Proposición 3.2. *Para todo $p \geq 0$,*

$$S(p) = B(p+1) - B(p) = \big|\{W_{p+1}(m) : \ m \text{ impar}, \ d(m) = p+1, \ \gamma(m) = 2\}\big|.$$

Equivalentemente, $S(p)$ es el número de palabras distintas del nivel $p+1$ que empiezan por 2; en particular, $S(p) \geq 0$.

Demostración. Por el lema 3.1, el nivel $p+1$ contiene las $B(p)$ palabras de los números pares, ninguna de las cuales empieza por 2. Sea m impar con $d(m) = p+1$; entonces $m > 1$. Si $\gamma(m) = 2$, su palabra empieza por 2 y es distinta de aquellas. Si $\gamma(m) = 1$, entonces $\nu_2(m+1)$ es par y positiva, luego $m \equiv 3 \pmod 4$. El número $q = 3m+1$ cumple $\nu_2(q) = 1$ y no es potencia de 2, porque $q/2$ es impar y mayor que 1; por tanto $\gamma(q) = 0$ y $\gamma(2q) = 1$. Como $C(m) = q = C(2q)$, se tiene $d(2q) = p+1$ y $W_{p+1}(m) = 1\,W_p(q) = W_{p+1}(2q)$, que es la palabra de un número par. □

Lema 3.3 (Paso hacia atrás)**.** *Sea x un impar no divisible por 3. Los enteros $a \geq 1$ con $2^a x \equiv 1 \pmod 3$ son los pares si $x \equiv 1$ y los impares si $x \equiv 2 \pmod 3$. Para ellos, $y(a) = (2^a x - 1)/3$ es un entero impar y $y(a+2) = 4\,y(a) + 1$. Además:*

1. *Si $x > 1$ y $a \geq 2$, entonces $y > x$, $\nu_2(y+1) = 1$ y $\gamma(y) = 2$.*
2. *Para todo $k \geq 1$, la aplicación $y \mapsto 4y+1$ recorre todos los restos módulo 3^k a partir de cualquier valor inicial. En consecuencia, cuando a recorre 3^k valores consecutivos de su paridad, $y(a)$ recorre todos los restos módulo 3^k.*

Demostración. La primera afirmación se sigue de $2^a \equiv (-1)^a \pmod 3$; el numerador $2^a x - 1$ es impar, y $3y(a+2) + 1 = 4(3y(a) + 1)$. Si $x > 1$ y $a \geq 2$, entonces $y \geq (4x-1)/3 > x > 1$; además $y + 1 = (2^a x + 2)/3$ y $2^a x + 2 \equiv 2 \pmod 4$, luego $\nu_2(y+1) = 1$ y, como $y > 1$ es impar, $\gamma(y) = 2$. La hipótesis $x > 1$ es necesaria: para $x = 1$ y $a = 2$ se obtiene $y = 1$.

Para el segundo apartado, la igualdad $3(4y+1) + 1 = 4(3y+1)$ muestra que, módulo 3^{k+1}, el número $3y+1$ se multiplica por 4. Los restos $u \equiv 1 \pmod 3$ módulo 3^{k+1} forman un grupo multiplicativo de orden 3^k, y 4 tiene en él orden exactamente 3^k, porque por inducción $4^{3^i} \equiv 1 + 3^{i+1} \pmod{3^{i+2}}$. Por tanto 4 genera ese grupo, $3y+1$ recorre todos sus elementos e y recorre todos los restos módulo 3^k. □

Para $r \geq 4$ par, el número $x_0 = (2^r - 1)/3$ es impar, $3x_0 + 1 = 2^r$, $d(x_0) = r+1$ y $\gamma(x_0) = 2$, porque $\nu_2(x_0 + 1) = 1$. También $x_0(r+2) = 4x_0(r) + 1$, de modo que el lema 3.3(2) se aplica igualmente a x_0 como función de r; en particular, $x_0 \equiv 1$, 2 o 0 (mód 3) según que $r \equiv 2$, 4 o 0 (mód 6).

4. Crecimiento exponencial

Construcción. Sean $p \geq 17$ y $h = \lfloor (p-3)/14 \rfloor \geq 1$. A cada $t = (t_0, \ldots, t_{h-1}) \in \{0,1\}^h$ le asociamos exponentes $r, a_1, \ldots, a_{h-1}, j$ del modo siguiente.

1. Si $h = 1$, sea $r = c_0 + 6t_0$, con $c_0 = 8$ si $p-1$ es par y $c_0 = 4$ si $p-1$ es impar. Si $h \geq 2$, sea $r = 4 + 6t_0$. En ambos casos $x_0 = (2^r - 1)/3$.
2. Para $1 \leq i \leq h-2$, sea c_i el menor $a \geq 2$ con $2^a x_{i-1} \equiv 1 \pmod 3$ y $(2^a x_{i-1} - 1)/3 \equiv 1 \pmod 3$; se pone $a_i = c_i + 6t_i$ y $x_i = (2^{a_i} x_{i-1} - 1)/3$.

3. Si $h \geq 2$, sea π la paridad que el lema 3.3 impone a a_{h-1} a partir de x_{h-2}, y sea ε la paridad de $p - h - r - (a_1 + \cdots + a_{h-2}) - \pi$. Sea c_{h-1} el menor $a \geq 2$ con $2^a x_{h-2} \equiv 1 \pmod 3$ y $(2^a x_{h-2} - 1)/3 \equiv (-1)^\varepsilon \pmod 3$; se pone $a_{h-1} = c_{h-1} + 6t_{h-1}$ y $x_{h-1} = (2^{a_{h-1}} x_{h-2} - 1)/3$.

4. Por último, $j = p - h - r - (a_1 + \cdots + a_{h-1})$, $n = 2^j x_{h-1}$ y $m = (n-1)/3$.

Cada c_i depende de $t_0, \ldots, t_{i-1}$, y la regla lo determina sin ambigüedad.

Proposición 4.1. *En la construcción anterior todos los pasos están definidos, $j \geq 2$, $d(n) = p$, $n \equiv 4 \pmod{12}$, y m es un impar con $d(m) = p+1$, $\gamma(m) = 2$ y $W_{p+1}(m) = 2\,W_p(n)$.*

Demostración. Si $h \geq 2$, $r \equiv 4 \pmod 6$ y $x_0 \equiv 2 \pmod 3$. Por el lema 3.3, cada c_i existe y vale a lo sumo 7, porque dentro de la paridad admisible cada resto módulo 3 corresponde a una clase módulo 6; cada x_i es un impar mayor que x_{i-1}, no divisible por 3 y de etiqueta 2, y $3x_i + 1 = 2^{a_i} x_{i-1}$. Como $a_{h-1} \equiv c_{h-1} \pmod 6$ tiene la paridad π, la paridad de j es ε, y $x_{h-1} \equiv (-1)^j \pmod 3$. Si $h = 1$, la elección de c_0 da $x_0 \equiv 1 \pmod 3$ cuando $j = p - 1 - r$ es par y $x_0 \equiv 2 \pmod 3$ cuando es impar.

Como $t_i \leq 1$, se tiene $r \leq 14$ y $a_i \leq 13$, de modo que $r + \sum a_i \leq 13h + 1$ y $j \geq p - 14h - 1 \geq 2$. La órbita de n consta de j divisiones hasta x_{h-1}, de los bloques que llevan de cada x_i a x_{i-1} en $a_i + 1$ pasos, y de los $r + 1$ pasos de x_0 a 1, sin pasar antes por 1; por tanto $d(n) = j + \sum(a_i + 1) + r + 1 = p$. Como $4 \mid n$ y $n \equiv (-1)^j x_{h-1} \equiv 1 \pmod 3$, se tiene $n \equiv 4 \pmod{12}$; m es un impar mayor que 1 con $C(m) = n$, luego $d(m) = p + 1$, y de $3m + 1 \equiv 0 \pmod 4$ se sigue $m \equiv 1 \pmod 4$ y $\gamma(m) = 2$. □

Lema 4.2 (Recuperación)**.** *La palabra $W_p(n)$ de la construcción contiene exactamente h símbolos 2, en las posiciones de $x_{h-1}, \ldots, x_0$, y de esas posiciones se recupera t. Por tanto, tuplas distintas dan palabras $W_{p+1}(m)$ distintas.*

Demostración. Los únicos impares de la órbita de n antes de 1 son $x_{h-1}, \ldots, x_0$, todos de etiqueta 2; los demás números son pares y tienen etiqueta 0 o 1. Numerando las posiciones desde 0, x_{h-1} está en la posición j, cada x_{i-1} está $a_i + 1$ posiciones después de x_i, y x_0 está en la posición $p - 1 - r$. Las posiciones dan, pues, j, $a_{h-1}, \ldots, a_1$ y r. De r se obtienen t_0, porque c_0 depende solo de h y de la paridad de p, y x_0; conocido x_{i-1}, la regla determina c_i, y de a_i se obtienen $t_i = (a_i - c_i)/6$ y x_i. □

Teorema 4.3. *Para todo $p \geq 17$,*

$$B(p+1) - B(p) \;\geq\; 2^{\lfloor (p-3)/14 \rfloor}.$$

Demostración. Por la proposición 4.1 y el lema 4.2, las 2^h tuplas dan 2^h palabras distintas del nivel $p + 1$ que empiezan por 2, y la proposición 3.2, en su forma equivalente, concluye. □

Corolario 4.4. *Para todo $p \geq 18$, $2^{\lfloor (p-4)/14 \rfloor} \leq B(p) \leq \min\{M(p),\, a(p+2)\}$. En consecuencia,*

$$2^{1/14} \leq \liminf_{p\to\infty} B(p)^{1/p} \leq \limsup_{p\to\infty} B(p)^{1/p} \leq \lambda \approx 1{,}4656.$$

Demostración. La cota inferior se sigue de $B(p) \geq S(p-1)$ y del teorema 4.3. Para la superior, si $d(n) = p$ la órbita de n no pasa por 1 antes del paso p, de modo que $W_p(n)$ no contiene 00 por [2, Lema 4.1]; es, por tanto, una de las $a(p+2)$ palabras del lenguaje de tipo finito. □

5. Coincidencias

Llamamos *punto de ambigüedad* a un entero $z > 4$ con $z \equiv 4 \pmod 6$ cuyas dos preimágenes $2z$ y $(z-1)/3$ tienen la misma etiqueta.

Lema 5.1. *Un entero $z > 4$ es un punto de ambigüedad si y solo si $z = 2u$, con $u > 1$ impar, $u \equiv 2 \pmod 3$ y $\gamma(u) = 2$. En ese caso, $\gamma(2z) = \gamma((z-1)/3) = 1$.*

Demostración. Como $\gamma(2z) \in \{0, 1\}$ y la etiqueta de un impar mayor que 1 es 1 o 2, la etiqueta común solo puede ser 1. Sea $m = (z-1)/3$. Se tiene $\gamma(m) = 1$ si y solo si $\nu_2(m+1)$ es par y positiva; en particular $m \equiv 3 \pmod 4$, es decir, $z = 3m + 1 \equiv 2 \pmod 4$ y $z = 2u$ con u impar, $u > 1$. Entonces $m + 1 = (z+2)/3$ y $\nu_2(z+2) = 1 + \nu_2(u+1)$, que es par si y solo si $\nu_2(u+1)$ es impar, es decir, si y solo si $\gamma(u) = 2$. Por otra parte, $z = 2u$ no es potencia de 2 y tiene valoración 1, de modo que $\gamma(z) = 0$ y $\gamma(2z) = 1$. Por último, $z \equiv 4 \pmod 6$ equivale a $u \equiv 2 \pmod 3$. □

Proposición 5.2 (Reconstrucción)**.** *Sean $w \in \mathcal{W}_p$ y $\ell \in \{0, 1, 2\}$. Entonces $\mathcal{N}_{p+1}(\ell w)$ está formado por los números $2z$ y $(z-1)/3$, con $z \in \mathcal{N}_p(w)$, que existen y tienen etiqueta ℓ. Cada z aporta a lo sumo uno, salvo cuando $\ell = 1$ y z es un punto de ambigüedad, en cuyo caso aporta dos.*

Demostración. Es la descripción de las preimágenes por C, junto con el lema 5.1. □

Proposición 5.3 (Separación)**.** *Si $n \neq n'$ pertenecen a $\mathcal{N}_p(w)$, sea s el menor índice con $C^s(n) = C^s(n')$, es decir, el primer paso en que coinciden los iterados del mismo orden, y sea $z = C^s(n)$. Entonces z es un punto de ambigüedad y $\{C^{s-1}(n), C^{s-1}(n')\} = \{(z-1)/3,\ 2z\}$.*

Demostración. Como $C^p(n) = C^p(n') = 1$, el índice s existe, y $1 \leq s \leq p$. Los números $C^{s-1}(n)$ y $C^{s-1}(n')$ son preimágenes distintas de z, luego son $2z$ y $(z-1)/3$, con $z \equiv 4 \pmod 6$; además $z > 4$, porque la preimagen impar de 4 es 1, que no aparece antes del paso p. Ocupan la misma posición en dos órbitas con la misma palabra, de modo que tienen la misma etiqueta, y z es un punto de ambigüedad. □

Para la construcción de multiplicidades grandes, sea c_K el único resto módulo 3^K con $10c_K \equiv -1 \pmod{3^K}$. Se tiene $c_K \equiv 2 \pmod 3$, y $c_K \equiv 8 \pmod 9$ si $K \geq 2$.

Lema 5.4. *Sean $K \geq 2$ y $u_1, \ldots, u_s$ impares mayores que 1, todos congruentes con c_K módulo 3^K. Entonces los $2s$ números $4u_i$ y $(2u_i - 1)/3$ son congruentes entre sí módulo 3^{K-1} y congruentes con 2 módulo 3. Si j es impar, $j \geq 3$, los números $2^j \cdot 4u_i$ y $2^j(2u_i - 1)/3$ son pares, congruentes con 1 módulo 3, y tienen valoraciones impares y al menos 3; sus preimágenes impares son impares de etiqueta 2, congruentes entre sí módulo 3^{K-2}.*

Demostración. Se tiene $4u_i - (2u_i - 1)/3 = (10u_i + 1)/3$, y $10u_i + 1 \equiv 0 \pmod{3^K}$, de modo que la diferencia es divisible por 3^{K-1}. Si $u_i \equiv u_{i'} \pmod{3^K}$, también $4u_i \equiv 4u_{i'} \pmod{3^K}$ y $(2u_i - 1)/3 \equiv (2u_{i'} - 1)/3 \pmod{3^{K-1}}$. Como $u_i \equiv 2 \pmod 3$, $4u_i \equiv 2 \pmod 3$, y por la congruencia anterior, con $K - 1 \geq 1$, también $(2u_i - 1)/3 \equiv 2 \pmod 3$. Si j es impar, $2^j \equiv 2 \pmod 3$ y los números duplicados son $\equiv 1 \pmod 3$; sus valoraciones son $j + 2$ y j, ya que $(2u_i - 1)/3$ es impar. Sus preimágenes impares tienen etiqueta 2 por el lema 3.3(1), aplicado a las partes impares u_i y $(2u_i - 1)/3$, que son mayores que 1 porque $u_i \geq 5$, y dividir por 3 números congruentes módulo 3^{K-1} que son $\equiv 1 \pmod 3$ da números congruentes módulo 3^{K-2}. □

Teorema 5.5. *Para todo $R \geq 1$ existen p_R y $w \in \mathcal{W}_{p_R}$ con $|\mathcal{N}_{p_R}(w)| \geq 2^R$, y puede tomarse $p_R \leq 9^R$ si $R \geq 2$.*

Demostración. Consideramos conjuntos de *ramas*: números distintos que ocupan la misma posición en una construcción hacia atrás desde 1, con la misma palabra hasta 1 y sin pasar por 1 antes. Para $k \geq 0$ y $K \geq 1$, sea $I(k, K)$ la afirmación de que hay 2^k ramas, todas impares, mayores que 1, de etiqueta 2 y congruentes con c_K módulo 3^K.

Arranque. Por el lema 3.3(2) aplicado a $x_0(r)$, existe un par $r \geq 4$, con $r \leq 2 \cdot 3^{2R-1} + 2$, tal que $x_0 = (2^r - 1)/3 \equiv c_{2R-1} \pmod{3^{2R-1}}$. La rama x_0, que ocupa la posición $r + 1$ contada desde 1, da $I(0, 2R - 1)$.

Ronda. Supongamos $I(k, K)$ con $K \geq 3$. Se duplican todas las ramas: $z_i = 2u_i$ es un punto de ambigüedad por el lema 5.1, porque $u_i \equiv 2 \pmod 3$ y $\gamma(u_i) = 2$, y todas tienen la misma etiqueta $\gamma(z_i) = 0$. Por la proposición 5.2, cada z_i da las dos ramas $4u_i$ y $(2u_i - 1)/3$, de etiqueta 1; son 2^{k+1} ramas distintas: las ramas pares $4u_i$ son distintas entre sí porque $u \mapsto 4u$ es inyectiva, las impares $(2u_i - 1)/3$ lo son porque $u \mapsto (2u - 1)/3$ también lo es, y ninguna rama par coincide con una impar. Se duplican después j veces, con j impar y $j \geq 3$, lo que conserva la igualdad de etiquetas por el lema 3.1, y se toma la preimagen impar de todas, que por el lema 5.4 existe y da 2^{k+1} impares de etiqueta 2 congruentes entre sí módulo 3^{K-2}. Por el lema 3.3(2), aplicado a una de las ramas, se puede elegir $j \leq 2 \cdot 3^{K-2} + 1$ de modo que su resto, y por tanto el de todas, sea c_{K-2}. Se obtiene $I(k + 1, K - 2)$, y la ronda ocupa $j + 3 \leq 2 \cdot 3^{K-2} + 4$ posiciones.

Bifurcación final. Partiendo de $I(0, 2R - 1)$, las rondas con $K = 2R - 1, 2R - 3, \ldots, 3$ dan $I(R - 1, 1)$. Una duplicación y una bifurcación más, que solo requieren que cada rama sea un impar de etiqueta 2 congruente con 2 módulo 3, dan 2^R números distintos con la misma palabra y el mismo tiempo de parada p_R, que ninguna rama alcanza 1 antes.

Longitud. Sumando,

$$p_R \leq (2 \cdot 3^{2R-1} + 3) + \sum_{i=1}^{R-1} (2 \cdot 3^{2R-1-2i} + 4) + 2 = 2 \cdot 3^{2R-1} + \tfrac{3}{4}(9^{R-1} - 1) + 4R + 1,$$

que es menor o igual que 9^R para $R \geq 2$. □

Corolario 5.6. *$m(p) \to \infty$, y $m(p) \geq 2^{\lfloor \log_9 p \rfloor} \geq \frac{1}{2} p^{\log_9 2}$ para $p \geq 81$.*

Demostración. Por el teorema 5.5, $m(p_R) \geq 2^R$ con $p_R \leq 9^R$, y m no decrece por el lema 3.1. Si $p \geq 81$ y $R = \lfloor \log_9 p \rfloor \geq 2$, entonces $p \geq 9^R \geq p_R$ y $m(p) \geq 2^R \geq 2^{\log_9 p - 1}$. □

6. Palabras de Collatz y árbol inverso

Para n con $d(n) = p$, llamamos *palabra de Collatz* de n a $U_p(n) = u_0 u_1 \cdots u_{p-1}$, con $u_i = \mathtt{u}$ si $C^i(n)$ es impar y $u_i = \mathtt{d}$ si es par. Es el patrón de subidas y bajadas de la trayectoria hasta su primera llegada a 1, que estudian Trümper [7] y Lang [5].

Proposición 6.1. *Sea $p \geq 0$.*

1. *La aplicación $n \mapsto U_p(n)$ es inyectiva sobre $\{n : d(n) = p\}$.*
2. *La palabra $W_p(n)$ queda determinada por $U_p(n)$, sin necesidad de conocer n.*

En consecuencia, $M(p)$ es el número de palabras de Collatz de longitud p, $B(p)$ es el número de sus imágenes por la aplicación $U_p(n) \mapsto W_p(n)$, y dos números distintos con $d(n) = d(n') = p$ comparten palabra de Grundy exactamente cuando esa aplicación no es inyectiva sobre sus palabras de Collatz.

Demostración. Sea $x_i = C^i(n)$, de modo que $x_p = 1$ y $x_i \neq 1$ para $i < p$. Leyendo la palabra desde el final, $x_i = 2x_{i+1}$ si $u_i = \texttt{d}$ y $x_i = (x_{i+1} - 1)/3$ si $u_i = \texttt{u}$; esto prueba el primer apartado.

Para el segundo, sea $i < p$. Si $u_i = \texttt{d}$, sea r la longitud de la racha de letras `d` que empieza en i. Si la racha llega hasta la posición $p - 1$, entonces $x_i = 2^r$ y $\gamma(x_i) = r$ mód 2; en otro caso x_{i+r} es un impar mayor que 1, x_i no es potencia de 2, $\nu_2(x_i) = r$ y $\gamma(x_i) = 1 - (r \text{ mód } 2)$. Si $u_i = \texttt{u}$, entonces $x_i > 1$ es impar; sea $e \geq 1$ la longitud de la racha de letras `d` que empieza en $i+1$. Si $e \geq 2$, entonces $3x_i + 1 \equiv 0 \pmod 4$, luego $x_i \equiv 1 \pmod 4$, $\nu_2(x_i + 1) = 1$ y $\gamma(x_i) = 2$. Si $e = 1$, el número $x_{i+2} = (3x_i + 1)/2$ es un impar mayor que 1, y $\gamma(x_i) = 3 - \gamma(x_{i+2})$ por [2, Lema 2.1]. Procediendo desde el final de la palabra hacia el principio, todas las etiquetas quedan determinadas. □

La proposición se refiere solo a las palabras de trayectorias que llegan a 1 por primera vez en el paso p: no toda sucesión de letras `u` y `d` es la palabra de Collatz de algún número. En este lenguaje, la proposición 5.3 dice que dos palabras de Collatz con la misma imagen difieren, en la última posición en que difieren, en una letra `u` frente a una letra `d`, correspondientes a las preimágenes $(z - 1)/3$ y $2z$ de un punto de ambigüedad z.

La sucesión $M(p)$ es A005186 [8]. En sus términos conocidos, el cociente $M(p+1)/M(p)$ se aproxima a

$$\rho = \frac{3 + \sqrt{21}}{6} = \frac{1 + \sqrt{7/3}}{2} \approx 1{,}2637626,$$

la sucesión A176014. Merlini y Sala [6] propusieron este valor para el límite del cociente, y en la ficha de A005186 se recogen argumentos heurísticos de D. W. Wilson y de M. Sigg que conducen a la misma constante; ρ es la raíz positiva de $\rho^2 = \rho + \frac{1}{3}$, la ecuación que resulta de suponer que cada nodo m tiene la preimagen $2m$, a un paso de C, y, con frecuencia 1/3, la preimagen impar $(2m - 1)/3$ de la función acelerada, que llega a m en dos pasos de C. No se conoce una demostración de que el cociente tienda a ρ.

Los resultados rigurosos sobre el tamaño del árbol inverso se refieren a la función acelerada $T(x) = x/2$ o $(3x + 1)/2$ y miden la profundidad en pasos de T. Lagarias y Weiss [4] estudian modelos de ramificación para estos árboles; el número medio de nodos no divisibles por 3 a profundidad k, promediado sobre las clases de restos de la raíz, es exactamente $(4/3)^k$ [3, (2.8)]. Applegate y Lagarias demuestran, con ayuda del ordenador, que para toda raíz $a \not\equiv 0 \pmod 3$ el número de preimágenes de a a profundidad k está entre $1{,}302^k$ y $1{,}359^k$ para k suficientemente grande [3, Teorema 1.1]. Como un nodo a profundidad k en pasos de T que ha dado w pasos impares está a profundidad $k + w$ en pasos de C, estas cotas no se traducen directamente en cotas para $M(p)$, que dependen de la distribución de w. El corolario 4.4 da, en particular, $M(p) \geq B(p) \geq 2^{\lfloor (p-4)/14 \rfloor}$, cota que no se compara con aquellas.

7. Datos numéricos

Los valores siguientes se han obtenido recorriendo el árbol inverso de 1 por niveles. Son observaciones, no resultados demostrados, salvo en lo que coinciden con los enunciados anteriores.

p	$M(p)$	$B(p)$	$S(p)$	$m(p)$
10	6	5	2	2
15	24	17	4	2
20	72	50	13	3
25	227	161	44	3
30	732	531	139	3
35	2365	1698	451	4
40	7628	5408	1405	5
45	24561	17219	4530	6
50	79255	55204	14383	7
55	255455	177052		8

Cuadro 1: Números con tiempo total de parada p, palabras distintas, palabras nuevas y multiplicidad máxima.

En estos datos, la razón $B(p+1)/B(p)$ está entre 1,2605 y 1,2639 para $45 \le p \le 54$, muy cerca de la constante $\rho \approx 1{,}2638$ de la sección 6, y muy por encima de la cota inferior $2^{1/14} \approx 1{,}051$ del corolario 4.4; en el mismo intervalo, $M(p+1)/M(p)$ está entre 1,2632 y 1,2644. La razón $B(p)/M(p)$ está entre 0,692 y 0,726 para $20 \le p \le 55$. La multiplicidad máxima $m(p)$ vale 2 desde $p = 7$, 3 desde $p = 17$, 4 desde $p = 31$ y 8 en $p = 55$, de modo que crece mucho más deprisa que la cota $\frac{1}{2}\, p^{\log_9 2}$ del corolario 5.6. En la clase de 8 números del nivel 55 se observa la estructura de la demostración del teorema 5.5: leída desde 1, la clase se divide en tres puntos, y en cada uno se dividen a la vez todas sus ramas.

8. Conclusiones

Las palabras de Grundy de los números que llegan a 1 en exactamente p pasos nacen, de un nivel al siguiente, solo en los impares de etiqueta 2: las palabras de los números pares reproducen las del nivel anterior, y las de los impares de etiqueta 1 coinciden con las de un número par. La libertad de los exponentes entre impares, controlada mediante restos módulo 3, permite construir en cada nivel $2^{\lfloor (p-3)/14 \rfloor}$ palabras nuevas distintas, de modo que $B(p)$ crece exponencialmente, con una base comprendida entre $2^{1/14}$ y la raíz real de $\lambda^3 = \lambda^2 + 1$.

Cuando dos números tienen la misma palabra, sus órbitas se juntan por primera vez en un punto de ambigüedad, de la forma $2u$ con u impar, $u \equiv 2 \pmod 3$ y $\gamma(u) = 2$. Este mecanismo local describe cómo se produce cada coincidencia, pero no da por sí solo su número, que es una de las cuestiones abiertas. Encadenando bifurcaciones en estos puntos y conservando la congruencia de todas las ramas módulo potencias de 3, se obtienen palabras que corresponden a 2^R números distintos para cualquier R. En las dos construcciones intervienen a la vez las valoraciones 2-ádicas, que deciden las etiquetas, y los restos módulo potencias de 3, que deciden qué ramas del árbol existen; es la misma interacción que hace irregular el recuento de $B(p)$ y que no permite, por ahora, describir sus valores con una fórmula.

Visto desde las palabras de Collatz, $B(p)$ cuenta las clases en que quedan agrupadas las $M(p)$ trayectorias que llegan a 1 en p pasos cuando se leen solo sus etiquetas de Grundy. Los datos sugieren que $B(p)$ y $M(p)$ crecen al mismo ritmo, cercano a la constante $\rho = (3 + \sqrt{21})/6$ propuesta para $M(p)$, pero ninguno de los dos ritmos está demostrado.

9. Cuestiones abiertas

Problema 9.1. ¿Existe $\lim_{p\to\infty} B(p)^{1/p}$? En caso afirmativo, ¿coincide con $\rho = (3+\sqrt{21})/6$? Por el corolario 4.4, estaría entre $2^{1/14}$ y $\lambda \approx 1{,}4656$. La misma pregunta para $M(p)$ tiene la respuesta conjetural ρ [6, 8], también sin demostrar.

Problema 9.2. ¿Existe $\lim_{p\to\infty} B(p)/M(p)$? En caso afirmativo, ¿es estrictamente positivo? Esta pregunta no equivale a la anterior: si $B(p)$ y $M(p)$ tuvieran la misma base exponencial, su cociente podría aun así tender a 0, como ocurriría con ρ^p/p frente a ρ^p. La diferencia $M(p) - B(p) = \sum_{w\in\mathcal{W}_p}(|\mathcal{N}_p(w)| - 1)$ cuenta los números que sobran cuando se elige un representante de cada palabra; el número de parejas de números distintos con la misma palabra es, en cambio, $\sum_{w\in\mathcal{W}_p}\binom{|\mathcal{N}_p(w)|}{2}$. En los datos de la tabla 1, la razón $B(p)/M(p)$ está entre 0,692 y 0,726 para $20 \le p \le 55$.

Verificación

El programa `verificar_nacimiento_coincidencias.py`, que acompaña a este escrito, usa aritmética entera exacta. Comprueba la proposición 3.2 y la cota $B(p) \le a(p+2)$ para $p \le 40$; el periodo completo de $y \mapsto 4y+1$ módulo 3^k para $k \le 8$; la construcción de la sección 4 para $17 \le p \le 130$, con sus 2^h palabras distintas y las propiedades de la proposición 4.1; el lema 5.1 para $4 < z < 2\cdot 10^6$; la proposición 5.3 en las 1135 parejas de números con la misma palabra y $p \le 30$; la construcción del teorema 5.5 para $R \le 5$, comprobando en cada ronda las condiciones de $I(k,K)$ y, hacia delante, que las 2^R ramas tienen la misma palabra y el mismo tiempo de parada; los valores de la tabla 1 y las razones citadas en la sección 7; y, para los 5617 números con $d(n) \le 32$, que la palabra de Collatz determina el número y que la palabra de Grundy se obtiene de ella con las reglas de la demostración de la proposición 6.1. Ninguna demostración depende de estos cálculos.

Referencias

[1] M. Cerdá Bennassar, *El juego de Collatz en el sentido de Conway. Valores de Grundy locales y terminación global*, Zenodo, 2026. doi:10.5281/zenodo.23501310.

[2] M. Cerdá Bennassar, *Palabras de Grundy en las órbitas de Collatz. Un lenguaje de tipo finito y el árbol inverso de* 1, Zenodo, 2026. doi:10.5281/zenodo.23748643.

[3] D. Applegate y J. C. Lagarias, *Density bounds for the* $3x+1$ *problem. I. Tree-search method*, Math. Comp. 64 (1995), 411–426. doi:10.1090/S0025-5718-1995-1270612-0.

[4] J. C. Lagarias y A. Weiss, *The* $3x+1$ *problem: two stochastic models*, Ann. Appl. Probab. 2 (1992), 229–261.

[5] W. Lang, *On Collatz words, sequences, and trees*, J. Integer Seq. 17 (2014), artículo 14.11.7.

[6] D. Merlini y N. Sala, *On the Fibonacci's attractor and the long orbits in the* $3n+1$ *problem*, International Journal of Chaos Theory and Applications 4 (1999), n.º 2–3, 75–84.

[7] M. Trümper, *The Collatz problem in the light of an infinite free semigroup*, Chinese Journal of Mathematics (2014), artículo 756917.

[8] OEIS Foundation Inc., *The On-Line Encyclopedia of Integer Sequences*, https://oeis.org; sucesiones A005186, A097333, A176014 y A400826.