Complejidad de factores de la sucesión de valores de Grundy del juego de Collatz con atajo

Una fórmula exacta para todas las longitudes

Miguel Cerdá Bennassar

Octubre de 2026

Resumen

Se estudia la complejidad de factores de la sucesión γ de valores de Grundy del juego de Collatz con atajo, recogida en la OEIS como A400826. La sucesión coincide, salvo en $n = 1$ y en las potencias de dos, con una recodificación por bloques de longitud dos de la sucesión de duplicación de periodo, cuya complejidad es conocida. Los factores de γ de frecuencia positiva son exactamente los de esa sucesión auxiliar, y los demás proceden de las ventanas que contienen una potencia de dos. Se clasifican estos últimos y se obtiene una fórmula exacta para la complejidad en todas las longitudes: si $m \geq 2$ y $k = 2^m + q$ con $0 \leq q < 2^m$, el número de factores distintos de longitud k es $4 \cdot 2^m$ si $q = 0$, $4 \cdot 2^m + 5q + 1$ si $1 \leq q < 2^{m-1}$ y $5 \cdot 2^m + 3q - 1$ si $2^{m-1} \leq q < 2^m$. Se determina también, en cada longitud, cuántos de esos factores tienen frecuencia positiva y cuántos frecuencia cero.

1. Introducción

En [2] se considera el juego imparcial con atajo asociado a la iteración de Collatz. Las posiciones son los enteros positivos; desde n par se pasa a $n/2$, desde $n > 1$ impar se puede pasar a $3n+1$ o directamente a $(3n+1)/2$, y la posición 1 no tiene opciones. Cuando la órbita de n llega a 1, el juego es finito y el Teorema 3.3 de [2] da en forma cerrada su valor de Grundy. Escrito con la valoración 2-ádica ν_2, ese valor es $\gamma(n)$, donde

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

La sección 3.1 de [2] define mediante la misma fórmula una función aritmética sobre todos los enteros positivos, sin hipótesis de terminación, y el Teorema 3.8 de [2] demuestra que es la única solución global del sistema local de ecuaciones mex con la condición terminal $\gamma(1) = 0$. Esa función incondicional es el objeto del presente escrito; no se afirma que una posición cuya órbita no terminara tuviera un valor de Grundy ordinario. Los términos de la sucesión A400826 de la OEIS [4] son estos valores. Los primeros son

$$0, 1, 1, 0, 2, 0, 2, 1, 2, 0, 1, 1, 2, 0, 1, 0, 2, 0, 1, 1, \ldots$$

En lo que sigue γ designa la función aritmética definida por (1). Los resultados se refieren a ella y no dependen de la conjetura de Collatz.

Para una sucesión $x = (x(n))_{n\geq 1}$ sobre un alfabeto finito escribimos

$$W_k^x(n) = x(n)\,x(n+1)\cdots x(n+k-1), \qquad \mathcal{L}_k(x) = \{W_k^x(n) : n \geq 1\}, \qquad P_x(k) = |\mathcal{L}_k(x)|.$$

La función P_x es la complejidad de factores de x. Las posiciones relativas de una ventana $W_k^x(n)$ se numeran de 0 a $k-1$, de modo que el índice $n+r$ ocupa la posición r. Un factor w de longitud k tiene *frecuencia positiva* si el conjunto $\{n \geq 1 : W_k^x(n) = w\}$ tiene densidad natural y esta es positiva, y *frecuencia cero* si ese conjunto tiene densidad nula.

Introducimos la sucesión auxiliar h, que sigue las mismas reglas que γ sin excepciones:

$$h(n) = \begin{cases} 0, & n \text{ par y } \nu_2(n) \text{ impar}, \\ 1, & n \text{ par y } \nu_2(n) \text{ par}, \\ 2, & n \text{ impar y } \nu_2(n+1) \text{ impar}, \\ 1, & n \text{ impar y } \nu_2(n+1) \text{ par}. \end{cases} \tag{2}$$

Sea $D = \{1\} \cup \{2^e : e \geq 1\}$. Las sucesiones h y γ coinciden fuera de D y difieren en todo punto de D: $h(1) = 2$, mientras que $\gamma(1) = 0$, y para $e \geq 1$ se tiene $h(2^e) = 1 - (e \bmod 2)$ y $\gamma(2^e) = e \bmod 2$.

El plan es el siguiente. En la sección 2 se relaciona h con la sucesión de duplicación de periodo, lo que da su complejidad en todas las longitudes, y se identifican los factores de γ de frecuencia positiva. En la sección 3 se describen las parejas de símbolos que las potencias de dos introducen en γ. Las secciones 4 y 5 estudian las dos familias de factores de frecuencia cero: las ventanas que contienen una potencia grande y las ventanas iniciales. La sección 6 demuestra que no hay otras y obtiene la fórmula de la complejidad (Teorema 6.2). La sección 7 recoge valores numéricos y algunas cuestiones abiertas.

2. La sucesión de duplicación de periodo

Sea $a(n) = \nu_2(n) \bmod 2$ para $n \geq 1$. Es la sucesión de duplicación de periodo, punto fijo de la sustitución $0 \mapsto 01$, $1 \mapsto 00$; figura en la OEIS como A096268, con un desplazamiento de una unidad en el índice. Su complejidad de factores es conocida. La sucesión A275202 de la OEIS la recoge, junto con las recurrencias $P_a(2t) = 2P_a(t)$ y $P_a(2t+1) = P_a(t) + P_a(t+1)$ para $t \geq 2$, y Poláková [5] da una demostración directa de la fórmula cerrada siguiente.

Teorema 2.1 (Poláková [5, Teorema 1.1])**.** *Sea $p = P_a$. Se tiene $p(1) = 2$ y, si $k \geq 1$ y $0 \leq q < 2^k$,*

$$p(2^k + q) = \begin{cases} 3 \cdot 2^{k-1} + 2q, & q \leq 2^{k-1}, \\ 4 \cdot 2^{k-1} + q, & q > 2^{k-1}. \end{cases}$$

La sucesión h se obtiene de a mediante una recodificación por bloques de longitud dos.

Lema 2.2. *Sea φ la aplicación definida en $\{00, 01, 10\}$ por $\varphi(00) = 1$, $\varphi(01) = 2$ y $\varphi(10) = 0$. Para todo $n \geq 1$ el bloque $a(n)a(n+1)$ es distinto de 11 y*

$$h(n) = \varphi(a(n)a(n+1)).$$

Demostración. Uno de los enteros n, $n+1$ es impar, y en él a vale 0; por tanto $a(n)a(n+1) \neq 11$. Si n es par, $a(n+1) = 0$, y $h(n) = 0$ si y solo si $\nu_2(n)$ es impar, es decir, si y solo si $a(n) = 1$; en otro caso $h(n) = 1 = \varphi(00)$. Si n es impar, $a(n) = 0$, y $h(n) = 2$ si y solo si $\nu_2(n+1)$ es impar, es decir, si y solo si $a(n+1) = 1$; en otro caso $h(n) = 1 = \varphi(00)$. □

Corolario 2.3. *Para todo $k \geq 1$ se tiene $P_h(k) = P_a(k+1)$. En consecuencia, si $m \geq 1$ y $0 \leq q < 2^m$,*

$$P_h(2^m + q) = \begin{cases} 3 \cdot 2^{m-1} + 2q + 2, & 0 \leq q < 2^{m-1}, \\ 2^{m+1} + q + 1, & 2^{m-1} \leq q < 2^m. \end{cases}$$

En particular, $P_h(2^m) = 3 \cdot 2^{m-1} + 2$.

Demostración. Por el Lema 2.2, $W_k^h(n)$ se obtiene de $W_{k+1}^a(n)$ aplicando φ a cada uno de sus k bloques consecutivos de longitud dos. Esta correspondencia entre palabras es inyectiva, porque φ lo es y una palabra queda determinada por sus bloques consecutivos de longitud dos. Así $P_h(k) = P_a(k+1)$.

Sea $k = 2^m + q$. Si $q + 1 < 2^m$, el Teorema 2.1 aplicado a $2^m + (q+1)$ da $3 \cdot 2^{m-1} + 2(q+1)$ cuando $q + 1 \leq 2^{m-1}$, es decir, cuando $q < 2^{m-1}$, y $4 \cdot 2^{m-1} + q + 1$ en otro caso. Si $q + 1 = 2^m$, el mismo teorema da $p(2^{m+1}) = 3 \cdot 2^m = 2^{m+1} + q + 1$. □

Lema 2.4. *Sea $k \geq 1$.*

1. *Todo $w \in \mathcal{L}_k(h)$ tiene en h una frecuencia, y esta es positiva.*
2. *$\mathcal{L}_k(h) \subseteq \mathcal{L}_k(\gamma)$. Los factores de γ de longitud k con frecuencia positiva son exactamente los elementos de $\mathcal{L}_k(h)$, y su frecuencia en γ es la misma que en h. Los demás factores de γ de longitud k tienen frecuencia cero.*

Demostración. (1) Sea $S = \{n \geq 1 : W_k^h(n) = w\}$. Los símbolos $h(n), \ldots, h(n+k-1)$ dependen solo de las paridades de $\nu_2(j)$ para j en $J(n) = \{n, n+1, \ldots, n+k\}$. Fijemos $M \geq 1$ y sea E_M el conjunto de los n tales que algún elemento de $J(n)$ es múltiplo de 2^M. Si $n \notin E_M$, cada $j \in J(n)$ tiene $\nu_2(j) = \nu_2(r_j)$, donde r_j es el resto no nulo de j módulo 2^M, de modo que la pertenencia de n a S depende solo de n mód 2^M. Existe entonces un conjunto R_M de restos tal que $S \setminus E_M = \{n \notin E_M : n \bmod 2^M \in R_M\}$. Como E_M es unión de $k+1$ clases de restos módulo 2^M, las densidades inferior y superior de S están ambas en el intervalo $[(|R_M| - k - 1)/2^M, (|R_M| + k + 1)/2^M]$. Al hacer $M \to \infty$ se concluye que coinciden, es decir, que S tiene densidad.

Para la positividad, sea $n_0 \in S$ y elijamos M con $2^M > n_0 + k$. Si $n \equiv n_0 \pmod{2^M}$, entonces $\nu_2(n+t) = \nu_2(n_0+t)$ para $0 \leq t \leq k$, ya que $0 < n_0 + t < 2^M$. Por tanto S contiene la clase de n_0 módulo 2^M y su densidad es al menos 2^{-M}.

(2) Sea Z_k el conjunto de los n tales que el intervalo $[n, n+k-1]$ corta a D. Cada elemento de D pertenece a lo sumo a k de esos intervalos y D tiene $O(\log x)$ elementos menores que x, así que Z_k tiene densidad nula. Si $n \notin Z_k$, entonces $W_k^\gamma(n) = W_k^h(n)$. En consecuencia, para cada palabra w los conjuntos de posiciones de w en γ y en h difieren en un subconjunto de Z_k. Si $w \in \mathcal{L}_k(h)$, su conjunto de posiciones en γ tiene la misma densidad positiva que en h, y en particular no es vacío. Si $w \notin \mathcal{L}_k(h)$, sus posiciones en γ están en Z_k y su frecuencia es cero. □

3. Las marcas de las potencias de dos

Los factores de longitud tres de a son 010, 100, 000, 001 y 101. Por el Lema 2.2, sus imágenes dan los factores de longitud dos de h:

$$\mathcal{L}_2(h) = \{20,\ 01,\ 11,\ 12,\ 02\}.$$

En particular, h no contiene las parejas 10 ni 21. Las llamaremos *marcas*. La posición de una marca en una ventana es la posición relativa de su segundo símbolo.

Lema 3.1. *Sea $p = 2^e \geq 4$. La pareja $(\gamma(p-1), \gamma(p))$ es la marca* 10 *si e es par y la marca* 21 *si e es impar; diremos que es la marca de p y que su* tipo *es e* mód 2*. Cualquier otra pareja de símbolos consecutivos de γ pertenece a $\mathcal{L}_2(h)$.*

Demostración. Como $p - 1 > 1$ es impar y $\nu_2(p) = e$, se tiene $\gamma(p-1) = 2$ si e es impar y $\gamma(p-1) = 1$ si e es par, mientras que $\gamma(p) = e$ mód 2. Esto da las dos marcas. Una pareja $(\gamma(n), \gamma(n+1))$ en la que ni n ni $n+1$ están en D es una pareja de h. Las restantes son $(\gamma(1), \gamma(2)) = 01$, $(\gamma(2), \gamma(3)) = 11$, las ya descritas $(\gamma(p-1), \gamma(p))$ con $p \geq 4$, y las $(\gamma(p), \gamma(p+1))$ con $p = 2^e \geq 4$, que valen (e mód 2, 2) porque $\nu_2(p+2) = 1$; son 02 o 12. Todas, salvo las marcas, están en $\mathcal{L}_2(h)$. □

Así, una ventana de γ contiene la marca de $p = 2^e \geq 4$ si y solo si contiene los índices $p-1$ y p, y no contiene ninguna otra marca.

En el resto del escrito fijamos $m \geq 2$, ponemos $L = 2^m$ y consideramos longitudes $k = L + q$ con $0 \leq q < L$, de modo que $L \leq k < 2L$. Cada longitud $k \geq 4$ se escribe de esta forma de manera única.

4. Ventanas que contienen una potencia grande

Para $e \geq m + 2$ y $0 \leq i \leq k - 1$ definimos

$$B_{e,i} = W_k^\gamma(2^e - i),$$

la ventana de longitud k en la que la potencia 2^e ocupa la posición relativa i.

Lema 4.1. *Sean $e \geq m + 2$ y $0 \leq i \leq k - 1$.*

1. *El único elemento de D en el intervalo $[2^e - i,\ 2^e - i + k - 1]$ es 2^e.*
2. *Los símbolos de $B_{e,i}$ son los siguientes. En el índice $2^e + j$, con $-i \leq j \leq k - 1 - i$: si $j = 0$, el símbolo es e* mód 2*; si $j \neq 0$ es par, depende de la paridad de $\nu_2(j)$ como en (2); si j es impar, depende de la paridad de $\nu_2(j+1)$ cuando $j \neq -1$, y de la paridad de e cuando $j = -1$. En particular, $B_{e,i}$ depende solo de i y de la paridad de e.*

Demostración. (1) Como $k < 2L \leq 2^{e-1}$, el intervalo está contenido en $(2^{e-1}, 2^{e+1})$.

(2) Se tiene $|j| \leq k - 1$ y $|j + 1| \leq k < 2^e$. Si $j \neq 0$ es par, $\nu_2(2^e + j) = \nu_2(j)$, porque $0 < |j| < 2^e$. Si j es impar y $j \neq -1$, $\nu_2(2^e + j + 1) = \nu_2(j + 1)$ por la misma razón, y si $j = -1$ esa valoración es e. Por (1), en todos los índices distintos de 2^e el símbolo de γ es el de h. □

Por el lema, escribimos B_i^ε, con $\varepsilon \in \{0,1\}$, para el valor común de las palabras $B_{e,i}$ con $e \geq m+2$ y $e \equiv \varepsilon \pmod 2$, y ponemos

$$\mathcal{B}_k = \{B_i^\varepsilon : \ 0 \leq i \leq k-1, \ \varepsilon \in \{0,1\}\}.$$

Lema 4.2. *Se tiene $|\mathcal{B}_k| = 2k$. Las dos palabras B_0^0 y B_0^1 pertenecen a $\mathcal{L}_k(h)$, y las $2k-2$ restantes contienen exactamente una marca y no pertenecen a $\mathcal{L}_k(h)$.*

Demostración. Sea $i \geq 1$. La ventana $B_{e,i}$ contiene los índices $2^e - 1$ y 2^e, y por el Lema 4.1(1) ningún otro elemento de D; contiene, pues, una sola marca, la de 2^e, en la posición i y de tipo ε. La posición determina i y el tipo determina ε. Resultan $2k-2$ palabras distintas, ninguna de las cuales está en $\mathcal{L}_k(h)$, porque h no tiene marcas.

Para $i = 0$ la ventana no contiene marcas y su primera letra es ε; las palabras B_0^0 y B_0^1 son distintas entre sí y de las anteriores. Para ver que $B_0^\varepsilon \in \mathcal{L}_k(h)$, tomemos $e' \geq m+2$ con $e' \not\equiv \varepsilon \pmod 2$. Entonces $h(2^{e'}) = 1 - (e' \bmod 2) = \varepsilon$, y para $1 \leq t \leq k-1$ el símbolo $h(2^{e'}+t)$ depende de $\nu_2(t)$ o de $\nu_2(t+1)$ exactamente como el símbolo de índice $2^e + t$ en el Lema 4.1(2). Así $W_k^h(2^{e'}) = B_0^\varepsilon$. □

5. Las ventanas iniciales

Definimos el conjunto de comienzos

$$S_k = \begin{cases} \{1, \ldots, L/2\}, & q = 0, \\ \{1, \ldots, L/2\} \cup \{L-q, \ldots, L\}, & 1 \leq q < L, \end{cases} \qquad \mathcal{C}_k = \{W_k^\gamma(n) : n \in S_k\}.$$

Lema 5.1. *Toda palabra de $\mathcal{C}_k$ contiene al menos una marca. Las palabras $W_k^\gamma(n)$, $n \in S_k$, son distintas dos a dos y ninguna pertenece a $\mathcal{L}_k(h)$. En consecuencia,*

$$|\mathcal{C}_k| = \begin{cases} L/2, & q = 0, \\ L/2 + q + 1, & 1 \leq q < L/2, \\ L, & L/2 \leq q < L. \end{cases}$$

Demostración. Sea $n \in S_k$, de modo que $n \leq L$, y sea $I = [n, n+k-1]$. Si $n \leq L-1$, el intervalo I contiene $L-1$ y L, ya que $n + k - 1 \geq L$, y la ventana contiene la marca de L. Si $n = L$, lo que solo ocurre cuando $q \geq 1$, el intervalo $I = [L, 2L+q-1]$ contiene $2L-1$ y $2L$, y la ventana contiene la marca de $2L$. Como $n + k - 1 \leq 3L - 2 < 4L - 1$, las marcas de la ventana son de potencias a lo sumo iguales a $2L$, y su última marca es la de $2L$ si I contiene $2L$, y la de L en otro caso.

Las marcas de L y de $2L$ son de tipos distintos. Dada una palabra $W_k^\gamma(n)$ con $n \in S_k$, el tipo de su última marca indica si esta es la marca de L o la de $2L$, y su posición r da $n = L - r$ o $n = 2L - r$. Las palabras son, por tanto, distintas, y ninguna está en $\mathcal{L}_k(h)$ porque h no tiene marcas.

El cardinal es el de S_k. Si $1 \leq q < L/2$, los dos intervalos que forman S_k son disjuntos, porque $L - q > L/2$; si $q \geq L/2$, se tiene $L - q \leq L/2$ y $S_k = \{1, \ldots, L\}$. □

Lema 5.2. *Se tiene* $\mathcal{C}_k \cap \mathcal{B}_k = \varnothing$.

Demostración. Por el Lema 5.1, las palabras de $\mathcal{C}_k$ contienen alguna marca, y por el Lema 4.2 las palabras de $\mathcal{B}_k$ con marcas tienen exactamente una. Basta, pues, considerar las ventanas $W_k^\gamma(n)$, $n \in S_k$, con una sola marca; para cada una, la única palabra de $\mathcal{B}_k$ que puede coincidir con ella es B_r^ε, donde r y ε son la posición y el tipo de esa marca.

Las marcas que puede contener la ventana son las de $L/2$ (si $L/2 \geq 4$), L y $2L$. La de $L/2$ aparece si y solo si $n \leq L/2 - 1$; la de L, si y solo si $n \leq L - 1$; la de $2L$, si y solo si $n + k - 1 \geq 2L$, es decir, $n \geq L - q + 1$. De aquí se obtiene la lista de las ventanas con una sola marca.

Si $m \geq 3$, son: $n = L/2$ cuando $q \leq L/2$ (marca de L); $n = L - q$ cuando $1 \leq q < L/2$ (marca de L); y $n = L$ cuando $q \geq 1$ (marca de $2L$). Si $m = 2$, la pareja de índices $(1, 2)$ no es una marca y a esa lista se añade $n = 1$ (marca de $L = 4$). En cada caso comparamos la ventana con su candidata.

Caso $n = L/2$. La marca de L está en la posición $r = L/2$ y es de tipo $\varepsilon = m$ mód 2. La primera letra de $B_{L/2}^\varepsilon$ es $\gamma(2^e - L/2)$ con $e \equiv m \pmod 2$; como $2^e - L/2$ es par, no es potencia de dos y tiene valoración $m-1$, esa letra es $1-((m-1) \bmod 2)$. La primera letra de la ventana es $\gamma(L/2) = (m-1) \bmod 2$.

Caso $n = L - q$, $1 \leq q < L/2$. La marca de L está en la posición $r = q$ y es de tipo $\varepsilon = m$ mód 2. La última letra de la ventana es $\gamma(2L-1)$, que depende de $\nu_2(2L) = m+1$ y vale $1 + ((m+1) \bmod 2)$. La última letra de B_q^ε es $\gamma(2^e + L - 1)$, que depende de $\nu_2(2^e + L) = m$ y vale $1 + (m \bmod 2)$.

Caso $n = L$, $q \geq 1$. La marca de $2L$ está en la posición $r = L$ y es de tipo $\varepsilon = (m+1)$ mód 2. La primera letra de B_L^ε es $\gamma(2^e - L)$, de valoración m, que vale $1 - (m \bmod 2)$. La primera letra de la ventana es $\gamma(L) = m \bmod 2$.

Caso $m = 2$, $n = 1$. La marca de 4 está en la posición $r = 3$ y es de tipo 0. La primera letra de B_3^0 es $\gamma(2^e - 3) = 2$, porque $\nu_2(2^e - 2) = 1$, y la de la ventana es $\gamma(1) = 0$.

En todos los casos las primeras o las últimas letras difieren, y la ventana no pertenece a $\mathcal{B}_k$. □

6. Clasificación y fórmula de la complejidad

Lema 6.1. *Se tiene* $\mathcal{L}_k(\gamma) = \mathcal{L}_k(h) \cup \mathcal{B}_k \cup \mathcal{C}_k$.

Demostración. La inclusión del segundo miembro en el primero se sigue del Lema 2.4(2) para $\mathcal{L}_k(h)$ y de las definiciones para $\mathcal{B}_k$ y $\mathcal{C}_k$. Sea $n \notin S_k$ y sea $I = [n, n+k-1]$. Distinguimos cuatro casos, que agotan los comienzos fuera de S_k.

Caso 1: $L/2 < n < L - q$. Se tiene $n < L \leq n + k - 1 \leq 2L - 2$, de modo que $I \cap D = \{L\}$. Sea $i = L - n$, que cumple $q < i < L/2$. Los índices de la ventana son $L + j$ con $-i \leq j \leq k - 1 - i \leq L - 2$. Si $j \neq 0$ es par, $\nu_2(L + j) = \nu_2(j)$, porque $0 < |j| < L$. Si j es impar, $j + 1$ está entre $-L/2 + 2$ y $L - 1$, de modo que $\nu_2(L + j + 1) = \nu_2(j + 1)$ si $j \neq -1$, y vale m si $j = -1$. Por último, $\gamma(L) = m$ mód 2. Comparando con el Lema 4.1(2), la ventana es B_i^ε con $\varepsilon = m$ mód 2.

Caso 2: $n = L$ *y* $q = 0$. Afirmamos que

$$W_L^\gamma(L) = W_L^h(3L/2). \tag{3}$$

Para $0 \le j \le L-1$ comparamos $\gamma(L+j)$ con $h(3L/2+j)$. Como $m \ge 2$, $L/2$ es par y ambos índices tienen la misma paridad. En $j = 0$, $\gamma(L) = m$ mód 2 y $h(3L/2) = 1-((m-1)$ mód $2) = m$ mód 2, pues $\nu_2(3L/2) = m-1$. Para $1 \le j \le L-1$, el índice $L+j$ no está en D y $\gamma(L+j) = h(L+j)$. Si $L+j$ es par, hay que comparar $\nu_2(L+j)$ con $\nu_2(3L/2+j)$: ambas valen $\nu_2(j)$ si $\nu_2(j) < m-1$, y el único j del intervalo con $\nu_2(j) \ge m-1$ es $j = L/2$, para el que las valoraciones son $\nu_2(3L/2) = m-1$ y $\nu_2(2L) = m+1$, de igual paridad. Si $L+j$ es impar, hay que comparar $\nu_2(L+j+1)$ con $\nu_2(3L/2+j+1)$, con $2 \le j+1 \le L$: ambas valen $\nu_2(j+1)$ si $\nu_2(j+1) < m-1$; para $j+1 = L/2$ las valoraciones son $m-1$ y $m+1$, y para $j+1 = L$ son $\nu_2(2L) = m+1$ y $\nu_2(5L/2) = m-1$, de igual paridad en ambos casos. Esto prueba (3), y la ventana pertenece a $\mathcal{L}_L(h)$.

Caso 3: $L < n \le 2L$. Se tiene $n+k-1 \ge 2L$ y $n+k-1 \le 4L-2$, de modo que $I \cap D = \{2L\}$. Sea $i = 2L-n \in [0, L-1]$. Los índices son $2L+j$ con $-(L-1) \le j \le k-1 \le 2L-2$. Si $j \ne 0$ es par, $\nu_2(j) \le m$ y $\nu_2(2L+j) = \nu_2(j)$. Si j es impar, $-(L-2) \le j+1 \le 2L-1$, de modo que, si $j \ne -1$, $\nu_2(j+1) \le m$ y $\nu_2(2L+j+1) = \nu_2(j+1)$, mientras que, si $j = -1$, esa valoración es $m+1$. Por último, $\gamma(2L) = (m+1)$ mód 2. La ventana es B_i^ε con $\varepsilon = (m+1)$ mód 2.

Caso 4: $n > 2L$. El intervalo I no contiene 1 ni dos potencias de dos, porque dos potencias $p < p'$ mayores que $2L$ distan al menos $p > 2L > k-1$. Si I no corta a D, la ventana es $W_k^h(n) \in \mathcal{L}_k(h)$. Si contiene una potencia 2^e, entonces $2^e \ge n > 2L$, luego $e \ge m+2$, y la ventana es $B_{e,i}$ con $i = 2^e - n \in [0, k-1]$.

Los comienzos $n \le L/2$ y, si $q \ge 1$, los comprendidos entre $L-q$ y L están en S_k; los demás están en alguno de los cuatro casos. □

Teorema 6.2. *Sean* $m \ge 2$, $L = 2^m$, $0 \le q < L$ *y* $k = L+q$. *Entonces*

$$P_\gamma(k) = \begin{cases} 4L, & q = 0, \\ 4L+5q+1, & 1 \le q < L/2, \\ 5L+3q-1, & L/2 \le q < L. \end{cases}$$

De esos factores, los de frecuencia positiva son los $P_h(k)$ *factores de* h, *dados por el Corolario 2.3, y el número de factores de frecuencia cero es*

$$P_\gamma(k) - P_h(k) = \begin{cases} 5L/2-2, & q = 0, \\ 5L/2+3q-1, & 1 \le q < L/2, \\ 3L+2q-2, & L/2 \le q < L. \end{cases}$$

Demostración. Por los Lemas 6.1, 4.2, 5.1 y 5.2, $\mathcal{L}_k(\gamma)$ es unión disjunta de $\mathcal{L}_k(h)$, de $\mathcal{B}_k \setminus \mathcal{L}_k(h)$, que tiene $2k-2$ elementos, y de $\mathcal{C}_k$. Por tanto

$$P_\gamma(k) = P_h(k) + 2k - 2 + |\mathcal{C}_k|.$$

Con el Corolario 2.3 y el Lema 5.1: si $q = 0$, la suma es $(3L/2+2)+(2L-2)+L/2 = 4L$; si $1 \le q < L/2$, es $(3L/2+2q+2)+(2L+2q-2)+(L/2+q+1) = 4L+5q+1$; si $L/2 \le q < L$, es $(2L+q+1)+(2L+2q-2)+L = 5L+3q-1$. Por el Lema 2.4(2), los factores de frecuencia positiva son los de $\mathcal{L}_k(h)$ y los demás tienen frecuencia cero. □

Proposición 6.3. *Se tiene* $P_\gamma(1) = 3$, $P_\gamma(2) = 7$ *y* $P_\gamma(3) = 11$.

Demostración. Los tres símbolos aparecen en γ. Por el Lema 3.1, $\mathcal{L}_2(\gamma)$ está contenido en $\mathcal{L}_2(h) \cup \{10, 21\}$; contiene $\mathcal{L}_2(h)$ por el Lema 2.4(2) y las dos marcas, que aparecen en los índices $(3, 4)$ y $(7, 8)$. Así $P_\gamma(2) = 5 + 2 = 7$.

Para $k = 3$, el Corolario 2.3 da $P_h(3) = 6$. Si $n \geq 3$, el intervalo $[n, n + 2]$ contiene a lo sumo una potencia de dos, porque dos potencias $p < p'$ con $p \geq 3$ distan al menos 4. Si no contiene ninguna, la ventana es de h. Si contiene $p = 2^e \geq 4$ en la posición $i \in \{0, 1, 2\}$, sus índices son $p + j$ con $-2 \leq j \leq 2$, todos mayores que 2. Los índices pares $p \pm 2$ tienen valoración 1, el símbolo de p es e mód 2, y los de los índices impares $p - 1$ y $p + 1$ dependen de $\nu_2(p) = e$ y de $\nu_2(p + 2) = 1$; la ventana depende solo de i y de la paridad de e. Para $i = 0$ son 020 y 120, que están en $\mathcal{L}_3(h)$, como en el Lema 4.2; para $i = 1, 2$ son las cuatro palabras 102, 212, 010 y 021, que contienen una marca y no están en $\mathcal{L}_3(h)$. Quedan las ventanas iniciales: $W_3^\gamma(1) = 011 = W_3^h(2)$, que está en $\mathcal{L}_3(h)$, y $W_3^\gamma(2) = 110$, que contiene la marca 10 en la posición 2 y difiere de 010, única palabra anterior con esa marca en esa posición. Así $P_\gamma(3) = 6 + 4 + 1 = 11$. □

Comentario 6.4. La demostración describe explícitamente los factores de frecuencia cero: los $2k - 2$ que contienen una sola potencia 2^e con $e \geq m + 2$ en una posición distinta de la primera, y los de las ventanas que comienzan en S_k. Los factores de frecuencia positiva son los de h, cuyas frecuencias pueden obtenerse de las de la sucesión de duplicación de periodo, descritas en [5, Teorema 1.2]. La fórmula del Teorema 6.2 no vale para $m = 1$: daría 12 en lugar de $P_\gamma(3) = 11$.

7. Valores numéricos y cuestiones abiertas

La tabla siguiente recoge los primeros valores de P_γ y su descomposición según el Teorema 6.2 y la Proposición 6.3: los $P_h(k)$ factores de frecuencia positiva y los $P_\gamma(k) - P_h(k)$ de frecuencia cero.

k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Frec. positiva	3	5	6	8	10	11	12	14	16	18	20	21	22	23	24	26
Frec. cero	0	2	5	8	12	14	16	18	22	25	28	30	32	34	36	38
$P_\gamma(k)$	3	7	11	16	22	25	28	32	38	43	48	51	54	57	60	64
k	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
Frec. positiva	28	30	32	34	36	38	40	41	42	43	44	45	46	47	48	50
Frec. cero	42	45	48	51	54	57	60	62	64	66	68	70	72	74	76	78
$P_\gamma(k)$	70	75	80	85	90	95	100	103	106	109	112	115	118	121	124	128

Los valores coinciden con los obtenidos enumerando los factores del prefijo $\gamma(1), \ldots, \gamma(2^{20})$ para $1 \leq k \leq 256$. Una comprobación sobre un prefijo finito no constituye una demostración: solo da cotas inferiores para la complejidad, y la exhaustividad de la clasificación la garantizan las demostraciones anteriores.

La sucesión γ es 2-automática: lo es la sucesión de duplicación de periodo, lo son sus recodificaciones por bloques, y la modificación de una sucesión 2-automática sobre el conjunto 2-automático de las potencias de dos conserva esa propiedad [1]. La complejidad de factores de una sucesión automática puede calcularse de forma mecánica [3], y el programa Walnut [6] permitiría una comprobación independiente del Teorema 6.2.

Quedan abiertas otras cuestiones, como la determinación de los factores especiales de γ (los que admiten más de una extensión a la derecha) o el estudio de la complejidad de otras sucesiones obtenidas al modificar una sucesión automática sobre un conjunto lacunario, como el de las potencias de dos.

Los cálculos numéricos pueden reproducirse con el programa que acompaña al escrito. Ese programa comprueba además que γ coincide con los valores de Grundy del juego con atajo para $n \leq 10^5$, el Lema 2.2, la igualdad $P_h(k) = P_a(k+1)$ con la fórmula del Teorema 2.1 para $k \leq 128$, y, para $2 \leq m \leq 6$ y todo q, las igualdades de conjuntos de los Lemas 4.2 a 6.1.

Referencias

[1] J.-P. Allouche y J. Shallit, *Automatic Sequences: Theory, Applications, Generalizations*, Cambridge University Press, 2003.

[2] 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.

[3] É. Charlier, N. Rampersad y J. Shallit, Enumeration and decidable properties of automatic sequences, *Internat. J. Found. Comput. Sci.* **23** (2012), 1035–1066.

[4] OEIS Foundation Inc., *The On-Line Encyclopedia of Integer Sequences*, `https://oeis.org`; sucesiones A096268, A275202 y A400826.

[5] M. Poláková, Complexity and invariant measure of the period-doubling subshift, preprint, 2019. arXiv:1902.08387.

[6] J. Shallit, *The Logical Approach to Automatic Sequences: Exploring Combinatorics on Words with Walnut*, London Math. Soc. Lecture Note Ser. 482, Cambridge University Press, 2022.