A NOTE ON SOMOS SEQUENCES

HERNANDEZ JUAN

Abstract. We prove that the standard Somos-k sequence contains a noninteger for every $k \geq 8$, and give an explicit bound for its first occurrence. We classify the three-term Somos recurrences with independent coefficients and unit initial values whose terms are polynomials in the coefficients, and determine the possible first denominators. We also give exact prime-adic valuation formulas for the standard Somos-4 sequence.

hexagon:2611.00597v1

1. Introduction

For an integer $k \geq 2$, the standard Somos-k sequence is defined by

$$a_0 = \cdots = a_{k-1} = 1, \qquad a_n a_{n-k} = \sum_{j=1}^{\lfloor k/2 \rfloor} a_{n-j} a_{n-k+j} \quad (n \geq k). \tag{1.1}$$

Somos discovered the order-six example in 1982 during his study of addition formulas and presented its integrality problem in 1989 [18, 17]. Gale's accounts [5, 6] described the early computations and the related recurrences proposed by Raphael Robinson. Several independent proofs of integrality for orders four and five were found in 1990; Malouf published one in 1992 [11]. Propp's chronology [14] records Hickerson's proof for order six and Lotto's proof for order seven, both obtained in April 1990.

Fomin and Zelevinsky placed these examples in their theory of the Laurent phenomenon [4]. With independent initial variables, the terms in the relevant recurrences are Laurent polynomials over the coefficient ring. Their Theorems 1.4 and 1.8 apply respectively to the three-term and two-term right-hand sides of the Gale–Robinson recurrences. Setting the initial variables equal to one gives the integrality of Somos-4 through Somos-7. Subsequent combinatorial formulas express terms of the octahedron and cube recurrences through perfect matchings and groves [20, 2]. Explicit perfect-matching models for the two-term right-hand-side Gale–Robinson recurrences were developed by Bousquet-Mélou, Propp, and West [1].

1.1. Higher orders.

The first noninteger in Somos-8 is

$$a_{17} = \frac{420514}{7}.$$

The all-orders question is whether every standard Somos-k sequence with $k \geq 8$ contains a noninteger. Stanley recorded this question in his list of bijective proof problems [21, p. 68]. Somos's 2014 comment in OEIS A030127 also records the question of a published proof for all orders [12].

Let $N(k)$ denote that position, with the indexing in (1.1). Theorem 2.1 proves that $N(k)$ exists for every $k \geq 8$, with

$$2k \leq N(k) \leq 2k + \max\left\{1, \left\lfloor \frac{\lfloor k/2 \rfloor - 1}{4} \right\rfloor\right\}.$$

For $k = 2, 3$, every term is 1 by induction. Together with the known results for orders four through seven, this gives the exact classification: the standard sequence (1.1) is integral precisely for $2 \leq k \leq 7$. Hanna's 2004 conjecture predicts the sharper location $N(k) \in \{2k, 2k+1\}$ [12, 13].

Date: November 18, 2026.

The proof for $k \geq 8192$ begins with an exact elimination of the first two blocks of the recurrence. If a specified interval of terms were integral, explicit quotients with cubic numerators would be integers. An integer difference operator annihilates sixteen terms of a geometric expansion. The remaining values have absolute value less than one and consequently vanish. Clearing denominators gives an exponential polynomial with more real zeros than Rolle's theorem permits. The finite interval $8 \leq k \leq 8191$ is checked by exact integer remainders in the original recurrence; the complete verifier is included in Section 2.

1.2. **Prime powers in Somos-4.** The divisibility properties of Somos-4 belong to the arithmetic of elliptic divisibility sequences. Ward studied the repetition of primes in these sequences [26, 25]. Robinson established periodicity properties for Somos sequences and recorded observations about the spacing of prime-power divisors [15]. The elliptic interpretation of Somos-4 was developed by Swart and Hone [22, 7]; Hone and Swart subsequently studied stronger Laurent rings and arithmetic conditions for integrality [9]. Uchida described valuations of general Somos-4 sequences using canonical local heights [23].

For the standard Somos-4 sequence, the curve and point are

$$E:\ y^2 + y = x^3 - x, \qquad P = (0, 0).$$

The denominator correspondence appears in Everest–McLaren–Ward [3, Examples 2.7–2.8]; Jones and Rouse give explicit coordinates for $[2n-3]P$ in terms of neighboring Somos terms [10, proof of Corollary 1.2]. Together with the classical local denominator-growth formula [3, Lemma 3.1], these identities yield the following consequence. Write v_p for the p-adic valuation normalized by $v_p(p) = 1$. If an odd prime p divides a_n, then

$$v_p\left(a_{\frac{(2n-3)p^j+3}{2}}\right) = v_p(a_n) + j \qquad (j \geq 0).$$

Theorem 3.1 supplies a direct proof, including the coordinate calculation and the convergence range of the formal logarithm. It also gives the complete 2-adic exception and the absence of the bad prime 37. Corollary 3.2 describes all prime-power divisibility indices by the reduction order of P.

Van der Kamp proved arithmetic-progression results for prime-power divisibility indices and reviewed Robinson's observations [24, Section 8]. The explicit formula above answers the arbitrary-height question there for every occurring odd prime in the standard Somos-4 sequence. Its proof uses the established elliptic correspondence and local multiplication law.

1.3. **Three independent coefficients.** Gale's 1991 account extended Robinson's unit-coefficient observations to recurrences with three coefficients [5, 19]. On 11 April 1999, Somos formulated a refined polynomial version of the Gale–Robinson conjecture [19]. For integers k, p, q, r with $0 < p < q < r \leq k/2$, set $a_0 = \cdots = a_{k-1} = 1$ and

$$a_n a_{n-k} = x a_{n-p} a_{n-k+p} + y a_{n-q} a_{n-k+q} + z a_{n-r} a_{n-k+r} \qquad (n \geq k), \tag{1.2}$$

where x, y, z are independent indeterminates. He conjectured that every term belongs to $\mathbb{Z}[x, y, z]$ if and only if

$$r = p + q \qquad \text{or} \qquad k = p + q + r.$$

He also specified an interval for the first failure and two possible divisors. The positive direction follows from Fomin–Zelevinsky's Laurent theorem [4, Theorem 1.4]: the second equality is their Gale–Robinson condition, and the first becomes that condition after replacing the offset r by $k - r$. Hone, Kouloukas, and Ward identify these as the three-summand families in the Laurent classification recalled in their treatment of Hirota–Miwa reductions [8, p. 6]. The criterion here concerns polynomiality in the coefficients when all initial terms equal one.

Theorem 4.1 proves the full three-coefficient polynomial classification and its first-failure refinement. Outside the two displayed families, the first nonpolynomial term has index between $2k$ and $2k + p$, inclusive, and its reduced denominator is

$$x + y + z \qquad \text{or} \qquad x(x + y + z) + y + z.$$

These indices use the zero-based convention of this paper; Somos's original initial block is indexed from one.

The necessary conditions are obtained directly in the coefficient ring. The terms through index $2k-1$ are polynomial, and the next $p+1$ terms have the two displayed candidate divisors. Restriction to the first divisor reduces the numerator to a six-point polynomial identity. Coefficient formulas for a two-delay recurrence settle most parameters. Restriction to the second divisor, followed by first variations and pole calculations, treats the remaining parameters. All boundary cases and coefficient computations are included in Section 4.

Throughout the paper, an empty sum is zero.

2. Nonintegrality of the standard Somos sequences

For an integer $k \geq 2$, let $(a_n)_{n\geq 0}$ be defined by

$$a_0 = \cdots = a_{k-1} = 1, \qquad a_n a_{n-k} = \sum_{j=1}^{\lfloor k/2 \rfloor} a_{n-j} a_{n-k+j} \quad (n \geq k). \tag{2.1}$$

Every term is a positive rational number, by induction. In particular, the recurrence is defined for all $n \geq 0$. Write $N(k)$ for the least index of a noninteger term, when such an index exists.

Theorem 2.1. *For every integer $k \geq 8$, the index $N(k)$ exists and satisfies*

$$2k \leq N(k) \leq 2k + \max\left\{1, \left\lfloor \frac{\lfloor k/2 \rfloor - 1}{4} \right\rfloor\right\}.$$

We first prove the assertion for $k \geq 8192$. Write

$$k = 2m + \varepsilon, \qquad \varepsilon \in \{0,1\}, \qquad A = m-1, \qquad \ell = m+\varepsilon, \qquad H = 2^m, \qquad R = \left\lfloor \frac{m-1}{4} \right\rfloor.$$

Thus $m \geq 4096$ and $4R < m$.

2.1. **The first block.** Put $b_t = a_{k+t}$ and $c_t = b_t - 1$ for $0 \leq t < k$. The divisors in (2.1) are 1 throughout this block, so b_t, c_t are integers. Expanding the products in (2.1) gives

$$c_t = A + \sum_{u=0}^{t-1} c_u + h_t, \tag{2.2}$$

where $h_t = 0$ for $t < \ell$ and

$$h_{m+s} = c_s + \sum_{i=0}^{s} c_i c_{2s-i} \qquad (\varepsilon = 0,\ 0 \leq s < m), \tag{2.3}$$

$$h_{m+1+s} = \sum_{i=0}^{s} c_i c_{2s+1-i} \qquad (\varepsilon = 1,\ 0 \leq s < m). \tag{2.4}$$

Indeed, the two linear ranges cover $0, \ldots, t-1$; for even k the central index s occurs twice. The products have precisely the indices in (2.3) and (2.4). Equation (2.2) therefore gives

$$c_t = A2^t \quad (0 \leq t < \ell). \tag{2.5}$$

Every product in (2.3) or (2.4) has its first index $i < \ell$, so its first factor is given by (2.5).

Set

$$Z = c_{k-1}, \qquad W = \sum_{u=\ell}^{k-2} 2^{k-2-u} c_u, \qquad F = \begin{cases} AH/2 + A^2H^2/4, & \varepsilon = 0, \\ A^2H^2/2, & \varepsilon = 1. \end{cases}$$

At the last index of (2.3), the terms c_{m-1} and c_{m-1}^2 give F. At the last index of (2.4), the product $c_{m-1}c_m$ gives F. Reindexing the other products by $u = k-2-i$ gives AW in both cases. Thus

$$h_{k-1} = F + AW, \qquad \sum_{u=0}^{k-1} c_u = 2Z - A - F - AW. \tag{2.6}$$

We shall also need an upper bound for these two integers. Put

$$\mathcal{B} = 2A(k+1), \qquad d_t = 1 + \left\lfloor \log_2 \frac{k}{k-t} \right\rfloor \quad (0 \le t < k), \qquad d = \lfloor \log_2 k \rfloor + 1.$$

Then

$$0 < c_t \le 2^t \mathcal{B}^{d_t}, \qquad Z \le 2^{k-1}\mathcal{B}^d, \qquad W \le k2^{k-2}\mathcal{B}^d. \tag{2.7}$$

Here is the induction proving the first inequality. Let $f_t = c_t/2^t$ and $M_v = \max_{0\le u\le v} f_u$. Solving the prefix-sum recurrence (2.2) gives

$$c_t = A2^t + h_t + \sum_{j=\ell}^{t-1} 2^{t-1-j} h_j.$$

For $j \ge \ell$ put $v = 2j - k$. Each product in h_j has the form $A2^v f_{v-i}$, and there are at most k such products. The extra linear term in the even case is $A2^{j-m}$. It follows that

$$\begin{aligned} f_t &\le A + Ak2^{-m} + Ak \sum_{j=\ell}^{t} 2^{j-k} M_{2j-k} \\ &\le A(k+1) + AkM_{2t-k}. \end{aligned}$$

The second inequality uses $\sum_{j=\ell}^{t} 2^{j-k} < 2^{t-k+1} \le 1$. For $t < \ell$, $d_t = 1$ and the desired bound follows from (2.5). For $t \ge \ell$, $d_{2t-k} = d_t - 1$, since $k - (2t-k) = 2(k-t)$. The numbers d_t are nondecreasing. Induction therefore gives $M_{2t-k} \le \mathcal{B}^{d_t - 1}$, and

$$A(k+1) + Ak\mathcal{B}^{d_t-1} \le \mathcal{B}^{d_t}.$$

This proves the first inequality in (2.7). The other two follow by substituting the bound in the definitions of Z and W and using $d_t \le d$.

2.2. **Elimination in the next block.** Let $D_r = a_{2k+r}$ for $0 \le r \le R$. Splitting the sum in (2.1) at $j = r$ gives

$$b_r D_r = \sum_{i=0}^{r-1} D_i b_{2r-i} + Q_r, \qquad Q_r = \sum_{u=2r+1}^{m+r} b_u b_{k+2r-u}. \tag{2.8}$$

All indices in the last sum lie in the first block. In the first sum $b_{2r-i} = 1 + A2^{2r-i}$, because $2r < \ell$. Define $e_t = c_t - A2^t$. Expanding Q_r gives

$$Q_r = 2(1 + A4^r)Z + A(4^{r+1} - 1)W + B_r, \tag{2.9}$$

where

$$\begin{aligned} B_r = {} & m - r - A2^{2r+1} - F + \mathbf{1}_{\varepsilon=0} c_{m+r} \\ & - A2^{2r+1} \sum_{v=\ell}^{\ell+r-1} 2^{k-1-v} c_v + \sum_{u=\ell}^{m+r} e_u c_{k+2r-u}. \end{aligned} \tag{2.10}$$

Here $\mathbf{1}_{\varepsilon=0}$ equals 1 for even k and 0 for odd k. To verify this formula, the constant terms contribute $m - r$. The linear terms contribute

$$\sum_{u=2r+1}^{k-1} c_u + \mathbf{1}_{\varepsilon=0} c_{m+r} = 2Z - A2^{2r+1} - F - AW + \mathbf{1}_{\varepsilon=0} c_{m+r},$$

by (2.6). In the quadratic terms use $c_u = A2^u + e_u$ and then $v = k + 2r - u$. Their geometric part is

$$A2^{2r+1} \left(2W + Z - \sum_{v=\ell}^{\ell+r-1} 2^{k-1-v} c_v \right),$$

and their remaining part is the last sum in (2.10). Adding these three expressions proves (2.9).

The dependence on Z, W can now be removed. Let E denote the shift $r \mapsto r+1$ and set $L = (E-1)(E-4)$. If $U_r = b_r D_r - Q_r$, then (2.8) implies

$$U_r = \sum_{0 \le i < r} D_i + A4^r \sum_{0 \le i < r} 2^{-i} D_i.$$

Taking the two differences in L gives

$$LU_r = b_{r+3} D_{r+1} - (b_{r+2} + 3) D_r.$$

Both coefficients of Z, W in (2.9) are combinations of 1 and 4^r, and hence $LQ_r = LB_r$. Substitution gives

$$b_{r+2} D_{r+2} - 6(1 + 3A2^r) D_{r+1} + 8 b_r D_r = LB_r.$$

Equivalently, for $1 \le r \le R-1$,

$$b_{r+1}(D_{r+1} - 4D_r) - 2b_{r-1}(D_r - 4D_{r-1}) = B_{r+1} - 5B_r + 4B_{r-1}. \tag{2.11}$$

Only $B_0, \dots, B_R$ occur in this equation.

2.3. **The forcing polynomials.** The formulas for B_r needed below involve only initial segments of the deviations e_t. For integers $s \ge 0$ with $2s < m$,

$$e_{m+s} = \frac{A^2}{2}(3s+1)4^s + \left(\frac{A}{2}s + \frac{A(A+2)}{2}\right)2^s \qquad (\varepsilon = 0), \tag{2.12}$$

$$e_{m+1+s} = A^2\big((3s+1)4^s + 2^s\big) \qquad (\varepsilon = 1). \tag{2.13}$$

Indeed, the factors in $h_{\ell+i}$ are geometric in this range, so

$$h_{m+i} = A2^i + A^2(i+1)4^i, \qquad h_{m+1+i} = 2A^2(i+1)4^i.$$

Also

$$e_{\ell+s} = h_{\ell+s} + \sum_{i=0}^{s-1} 2^{s-1-i} h_{\ell+i}.$$

The sums $\sum_{i=0}^{s-1} 2^i = 2^s - 1$ and $\sum_{i=0}^{s-1}(i+1)2^i = (s-1)2^s + 1$ give (2.12) and (2.13). Since $4r < m$, these identities apply to every deviation in (2.10).

In the odd case, substitution in (2.10) yields

$$\begin{aligned} B_r = m - r - \frac{A^2H^2}{2} &\\ + \left(-2A - 2A^2 r H^2 + \frac{A^4(r+5)}{2}\right)4^r &\\ + \frac{A^4(6r^3 - 3r^2 + 13r - 10)}{4}\,16^r. & \end{aligned} \tag{2.14}$$

For the even case put $U = A^2/2$, $V = A/2$, $C = A(A+2)/2$, and define

$$H_{16}(r) = \frac{U^2(r+1)(12r^2+9r+2)}{2} + 2U\{V(6r-8) + C(6r-2)\},$$

$$H_4(r) = U\{V(10r+16) + 5C\} + \frac{V^2 r(r+1)(4r-1)}{6} + (r+1)(2rVC + C^2).$$

Then

$$\begin{aligned} B_r = m - r - F + (AH + Vr + C)2^r &\\ + \{-2A - A^2 r H^2 + U(3r+1) + AH(Vr+C) + H_4(r)\}4^r &\\ + U(3r+1)(AH + Vr + C)8^r + H_{16}(r)16^r. & \end{aligned} \tag{2.15}$$

Here are the convolution identities establishing these substitutions. For the even case let $E_s = U(3s+1)4^s + (Vs + C)2^s$. Then

$$\sum_{s=0}^{r} E_s E_{2r-s} = H_{16}(r)16^r + U(3r+1)(Vr+C)8^r + H_4(r)4^r. \tag{2.16}$$

For the odd case let $E_s = A^2((3s+1)4^s + 2^s)$. Then

$$\sum_{s=0}^{r-1} E_s E_{2r-1-s} = A^4 \left\{ \frac{6r^3 - 3r^2 + 13r - 10}{4} 16^r + \frac{r+5}{2} 4^r \right\}. \tag{2.17}$$

To prove (2.16), use

$$\sum_{s\ge 0} E_s z^s = \frac{U(1+8z)}{(1-4z)^2} + \frac{C + 2(V-C)z}{(1-2z)^2}.$$

The coefficient at z^n in its square is $L_4(n)4^n + L_2(n)2^n$, where

$$L_4(n) = U^2\left(\frac{3}{2}n^3 + 3n^2 + \frac{5}{2}n + 1\right) + 4U\{3n(C+V) - 2C - 8V\},$$
$$L_2(n) = \frac{V^2(n^3-n)}{6} + CVn^2 + (C^2 + CV + 10UV)n + C^2 + 10CU + 32UV.$$

To check this coefficient formula, substitute the four identities

$$\sum_{n\ge 0} w^n = \frac{1}{1-w}, \qquad \sum_{n\ge 0} nw^n = \frac{w}{(1-w)^2},$$
$$\sum_{n\ge 0} n^2 w^n = \frac{w(1+w)}{(1-w)^3}, \qquad \sum_{n\ge 0} n^3 w^n = \frac{w(1+4w+w^2)}{(1-w)^4}$$

in $\sum_{n\ge 0}\{L_4(n)(4z)^n + L_2(n)(2z)^n\}$. The resulting rational function is the square just displayed. The last three identities follow by repeated application of $w\,d/dw$ to the first one. The full-convolution coefficient at $2r$, plus E_r^2, is twice the left side of (2.16); substitution gives H_{16} and H_4. For the odd case use the same coefficient formula with $U = C = A^2$ and $V = 0$. Its coefficient at $2r-1$ is twice the left side of (2.17), which gives the stated expression. For $r = 0$, the right side of (2.17) is zero.

The two boundary sums in (2.10) are

$$A2^{2r+1} \sum_{v=\ell}^{\ell+r-1} 2^{k-1-v} c_v = 2^\varepsilon A^2 H^2 r 4^r + AH4^r \sum_{s=0}^{r-1} 2^{-s} E_s,$$
$$\sum_{u=\ell}^{m+r} e_u c_{k+2r-u} = AH4^r \sum_{s=0}^{r-\varepsilon} 2^{-s} E_s + \sum_{s=0}^{r-\varepsilon} E_s E_{2r-\varepsilon-s}.$$

Here $E_s = e_{\ell+s}$ is the corresponding sequence above. The first identity follows from $v = \ell + s$ and $c_{\ell+s} = 2^\varepsilon AH2^s + E_s$; the second follows from $u = \ell + s$ and $c_{k+2r-u} = AH4^r 2^{-s} + E_{2r-\varepsilon-s}$. The sums involving $2^{-s}E_s$ cancel, except for the last term when $\varepsilon = 0$. Thus the even expression before expanding its convolution is

$$m - r - F - 2A4^r - A^2H^2r4^r + AH2^r + E_r + AH2^r E_r + \sum_{s=0}^{r} E_s E_{2r-s}.$$

The odd expression is

$$m - r - 2A4^r - A^2H^2/2 - 2A^2H^2r4^r + \sum_{s=0}^{r-1} E_s E_{2r-1-s}.$$

Equations (2.16) and (2.17) now give (2.15) and (2.14) term by term.

2.4. **An integer quotient with a cubic numerator.** For $1 \le r \le R$ define

$$K_r = \frac{b_{r-1} b_r (D_r - 4D_{r-1})}{2^{r+1}}.$$

Multiplying (2.11) by $b_r/2^{r+2}$ gives

$$K_{r+1} - K_r = \frac{b_r(B_{r+1} - 5B_r + 4B_{r-1})}{2^{r+2}}. \tag{2.18}$$

Let $B(r,t)$ denote (2.14) or (2.15) with 2^{jr} replaced by t^j. We shall use the Laurent polynomial

$$\mathcal{P}(r,t) = \sum_{j=-1}^{4} p_j(r)t^j.$$

In the odd case its coefficients are

$$\begin{aligned}
p_{-1} &= -\frac{3}{2},\\
p_0 &= \frac{3Ar}{4},\\
p_1 &= \frac{3A^2(A^2-4H^2)}{8},\\
p_2 &= \frac{A^3(A^2-4H^2)}{8},\\
p_3 &= \frac{3A^4(30870r^3+8379r^2+32445r-18824)}{153664},\\
p_4 &= \frac{A^5(18r^3+9r^2+15r-8)}{64}.
\end{aligned} \tag{2.19}$$

In the even case they are

$$\begin{aligned}
p_{-1} &= -\frac{3}{2},\\
p_0 &= -\frac{Ar(2A+4H+r-9)}{16},\\
p_1 &= \frac{A^2}{16}(3A^2+12Ar+32A-12H^2+2H+6r^2+11r+28),\\
p_2 &= \frac{A^3}{288}\big(18A^2+135Ar+324A-72H^2+126Hr\\
&\qquad +180H+99r^2+375r+274\big),\\
p_3 &= \frac{A^3}{307328}\big(28812A^2r+38416A^2+57624AHr+76832AH\\
&\qquad +185220Ar^3+495831Ar^2+672763Ar\\
&\qquad +77232A+1111320r-455112\big),\\
p_4 &= \frac{A^4}{128}(36Ar^3+99Ar^2+111Ar+8A+216r-72).
\end{aligned} \tag{2.20}$$

These coefficients satisfy

$$\mathcal{P}(r+1,2t)-\mathcal{P}(r,t) = \frac{1+At}{4t}\big(B(r+1,2t)-5B(r,t)+4B(r-1,t/2)\big). \tag{2.21}$$

To verify this identity, write $B(r,t)=\sum_{j=0}^{4}\beta_j(r)t^j$, with all other $\beta_j=0$, and set

$$g_j(r) = 2^j\beta_j(r+1)-5\beta_j(r)+2^{2-j}\beta_j(r-1), \qquad q_j(r)=\frac{g_{j+1}(r)+Ag_j(r)}{4}.$$

The coefficient at t^j in (2.21) is the equality

$$2^j p_j(r+1)-p_j(r) = q_j(r). \tag{2.22}$$

For $j\neq 0$, let $\Delta q(r)=q(r+1)-q(r)$. Each q_j has degree at most 3, and the finite formula

$$p_j(r) = \sum_{h=0}^{3}\frac{(-2^j)^h}{(2^j-1)^{h+1}}\Delta^h q_j(r)$$

solves (2.22): multiplication by $(2^j-1)+2^j\Delta$ cancels consecutive terms, and $\Delta^4 q_j=0$. For $j=0$, the listed p_0 has zero constant term and $p_0(r+1)-p_0(r)=q_0(r)$, where

$$q_0(r)=\begin{cases}3A/4, & \varepsilon=1,\\ -A(2A+4H+2r-8)/16, & \varepsilon=0.\end{cases}$$

Substituting the coefficients in (2.14) and (2.15) in this finite formula gives (2.19) and (2.20). This establishes (2.21) as an identity in $\mathbb{Q}[A,H,r,t,t^{-1}]$.

By (2.18) and (2.21),

$$K_r=J+\mathcal{P}(r,2^r),\qquad J=K_1-\mathcal{P}(1,2)\quad(1\le r\le R). \tag{2.23}$$

Use the common integer

$$M=2765952=2^7\cdot 3^2\cdot 7^4.$$

It is divisible by 4 and clears all denominators in both coefficient lists. In addition, for every $j>0$ the polynomial Mp_j/A^j has integer coefficients, as the displayed lists show. Define

$$\Pi(r)=M\{J+\mathcal{P}(r,-1/A)\}=\widetilde{J}+f(r),\qquad f(0)=0. \tag{2.24}$$

Explicitly,

$$f(r)=-144A(6174r^3+441r^2-6741r)\qquad(\varepsilon=1), \tag{2.25}$$

$$f(r)=-16\{55566Ar^3+161406Ar^2+(86436A^2+25032A+333396)r\}\qquad(\varepsilon=0). \tag{2.26}$$

Thus the cubic coefficient of Π is $-889056A$ in both cases. In particular, Π is nonzero.

Suppose now that $D_0,\dots,D_R$ are integers. Then MK_1 and $M\mathcal{P}(1,2)$ are integers, so $MJ\in\mathbb{Z}$. For $r\ge 1$, $b_r=1+A2^r$ is odd, and $K_r\in b_r\mathbb{Z}[1/2]$, where $\mathbb{Z}[1/2]=\{z/2^v : z\in\mathbb{Z},\ v\in\mathbb{Z}_{\ge 0}\}$. The coefficient conditions above give

$$M\{\mathcal{P}(r,2^r)-\mathcal{P}(r,-1/A)\}\in b_r\mathbb{Z}[1/2].$$

Indeed, the difference for a positive power t^j is an integer polynomial multiple of $(At)^j-(-1)^j$, which is divisible by $1+At$; the difference for t^{-1} is $t^{-1}+A=(1+At)/t$. The constant term makes no contribution. The same coefficient conditions show that $M\mathcal{P}(r,-1/A)$ is an integer. Equation (2.23) therefore implies $\Pi(r)\in b_r\mathbb{Z}[1/2]\cap\mathbb{Z}=b_r\mathbb{Z}$, since b_r is odd. Consequently

$$\frac{\Pi(r)}{1+A2^r}\in\mathbb{Z}\quad(1\le r\le R). \tag{2.27}$$

We record a bound on Π which is independent of the integrality assumption. At $r=0,1$, (2.8) gives $b_0D_0=Q_0$ and $b_1D_1=D_0b_2+Q_1$. Hence

$$\begin{aligned}K_1&=\frac{(A+1)Q_1-(4A+3)Q_0}{4}\\&=-(A+1)Z+\frac{3A(A+2)}{4}W+\frac{(A+1)B_1-(4A+3)B_0}{4}.\end{aligned}$$

Inserting this expression in (2.24) gives

$$\widetilde{J}=-M(A+1)Z+\frac{3MA(A+2)}{4}W+S(A,H). \tag{2.28}$$

For odd k the polynomial S is

$$\begin{aligned}S=18(&-1152A^4+57624A^3H^2-38416A^3+192080A^2H^2\\&-115248A^2+153664AH^2-154816A).\end{aligned} \tag{2.29}$$

For even k it is

$$\begin{aligned}S=\ &-3600A^4+518616A^3H^2-264208A^3+1728720A^2H^2\\&+1037232A^2H-6914880A^2+1382976AH^2+1382976AH-5497088A+2540160.\end{aligned} \tag{2.30}$$

Both identities follow by substituting the displayed polynomials in

$$S = M\left\{\frac{(A+1)B(1,2)-(4A+3)B(0,1)}{4} - \mathcal{P}(1,2)+\mathcal{P}(0,-1/A)\right\},$$

and using $m = A+1$. Equations (2.25)–(2.30) also give $\Pi \in \mathbb{Z}[r]$ directly.

The sums of the absolute coefficients of S are 12834000 and 21270456, both less than 2^{25}. Those of f are 1923264 and 10589376, both less than 2^{24}. Every monomial of S has absolute value at most $(k+1)^4H^2$; for $0 \le r \le R$, every monomial of f has absolute value at most $(k+1)^4$. Using $M < 2^{22}$, $H^2 \le 2^k$ and (2.7) in (2.28) now gives

$$\begin{aligned} M(A+1)Z + \frac{3MA(A+2)}{4}W &< 2^{k+23}(k+1)^3\mathcal{B}^d, \\ |S| &< 2^{k+25}(k+1)^4\mathcal{B}^d, \\ \max_{0\le r\le R}|\Pi(r)| &< 2^{k+27}(k+1)^4\mathcal{B}^d. \end{aligned}$$

Here and below the maxima range over integers r. Let $s = \log_2(k+1)$. For $m \ge 4096$,

$$s \le \log_2(2m+2) \le \frac{\sqrt{m}}{4}.$$

The last inequality holds at $m = 4096$, and its difference has positive derivative thereafter, since $(m+1)\log 2 > 8\sqrt{m}$. Since $\mathcal{B} \le (k+1)^2$ and $d \le s+1$,

$$\begin{aligned} \log_2 \max_{0\le r\le R}|\Pi(r)| &< k + 27 + 4s + 2s(s+1) \\ &\le 2m + 28 + \frac{m}{8} + \frac{3}{2}\sqrt{m} < \frac{5m}{2}. \end{aligned}$$

For the last step, $28 \le m/128$ and $(3/2)\sqrt{m} \le 3m/128$ for $m \ge 4096$. We have proved

$$\max_{0\le r\le R}|\Pi(r)| < 2^{5m/2}. \tag{2.31}$$

2.5. **An obstruction to integral exponential quotients.**

Lemma 2.2. *Let $m \ge 4096$ be an integer, let $R = \lfloor (m-1)/4 \rfloor$, and let $A \ge 1$ be real. Let P be a nonzero real polynomial of degree at most 3 such that $|P(r)| \le 2^{5m/2}$ for every integer $0 \le r < R$. Then the numbers $P(r)/(1+A2^r)$, $1 \le r < R$, are not all integers.*

Proof. Let E be the shift operator and write

$$T = \prod_{h=1}^{16}(2^hE-1)^4 = \sum_{j=0}^{64} t_jE^j \in \mathbb{Z}[E].$$

Its constant coefficient is $t_0 = 1$, and

$$\sum_{j=0}^{64}|t_j| \le \prod_{h=1}^{16}(2^h+1)^4 < 2^{550}. \tag{2.32}$$

Indeed, the logarithm to base 2 of this product is

$$4\sum_{h=1}^{16} h + 4\sum_{h=1}^{16}\log_2(1+2^{-h}) < 544 + \frac{4}{\log 2} < 550.$$

If Q has degree at most 3, then $(2^hE-1)^4(Q(r)2^{-hr}) = 0$: each application replaces Q by its forward difference. All factors of T commute.

Put $u_r = P(r)/(1+A2^r)$. The finite geometric identity

$$\frac{1}{1+X} = \sum_{h=1}^{16}(-1)^{h-1}X^{-h} + \frac{X^{-16}}{1+X} \quad (X>0)$$

gives

$$u_r = \sum_{h=1}^{16} (-1)^{h-1} A^{-h} P(r) 2^{-hr} + \rho_r, \qquad |\rho_r| \le |P(r)| 2^{-17r}.$$

The operator T kills the finite sum. Take $r_0 = R - 324$. For each of $r_0, r_0 + 1, \dots, r_0 + 259$, every index used by T lies between 1 and $R - 1$, because $r_0 \ge 699$ and $r_0 + 259 + 64 = R - 1$. By (2.32),

$$|Tu_r| < 2^{550+5m/2-17r_0} \le 2^{6075-7m/4} < 1.$$

The middle inequality uses $R \ge m/4 - 1$, and the last uses $m \ge 4096$. If all u_r in the stated interval were integers, Tu_r would therefore vanish at these 260 distinct points.

Define the ordinary polynomial in independent variables r, x

$$G(r, x) = \sum_{j=0}^{64} t_j P(r + j) \prod_{\substack{0 \le h \le 64 \\ h \ne j}} (1 + A2^h x).$$

It has degree at most 64 in x and at most 3 in r, and

$$G(r, 2^r) = \left(\prod_{h=0}^{64} (1 + A2^{r+h}) \right) Tu_r.$$

It is nonzero, since

$$G(r, -1/A) = P(r) \prod_{h=1}^{64} (1 - 2^h) \ne 0$$

as a polynomial. Writing $G(r, x) = \sum_{j=0}^{64} g_j(r) x^j$, its substitution $G(r, 2^r)$ is also nonzero: after division by the largest occurring exponential, the corresponding nonzero polynomial dominates all terms with smaller exponents as $r \to +\infty$.

Consider a function

$$f(x) = \sum_{j=1}^{s} q_j(x) e^{\lambda_j x}, \qquad \lambda_1 < \cdots < \lambda_s, \qquad d_j = \deg q_j,$$

where each q_j is a nonzero real polynomial. It has at most $\sum_j (d_j + 1) - 1$ distinct real zeros. This follows by induction on s. For $s = 1$, divide by the nonvanishing exponential and use the polynomial root bound. For $s > 1$, divide by $e^{\lambda_1 x}$ and differentiate $d_1 + 1$ times. The first summand vanishes. Each other summand retains a nonzero polynomial of the same degree, since the map $q \mapsto q' + (\lambda_j - \lambda_1) q$ is invertible on polynomials of bounded degree. Repeated Rolle's theorem shows that any finite set of N zeros produces at least $N - d_1 - 1$ zeros of the resulting function. The induction hypothesis bounds that number and gives the assertion.

Apply this bound to $G(r, 2^r) = \sum_{j=0}^{64} g_j(r) e^{j(\log 2) r}$, omitting zero coefficients. There are at most 65 exponents and each polynomial has degree at most 3. Thus it has at most 259 real zeros, contrary to the 260 zeros already obtained. □

If $D_0, \dots, D_R$ were all integers, equations (2.27) and (2.31) would contradict Lemma 2.2, because Π is a nonzero cubic. Hence $N(k) \le 2k + R$ for $k \ge 8192$.

2.6. **The finite interval.** For $8 \le k \le 8191$, the following exact finite computation gives $N(k) \in \{2k, 2k + 1\}$. In this subsection put $m = \lfloor k/2 \rfloor$. The first block can be computed modulo any positive integer q, since all its divisors are 1. The divisor at index $2k$ is $a_k = m$; compute its numerator modulo m. A nonzero remainder proves $N(k) = 2k$. If the remainder is zero, a_{2k} is an integer. The next divisor is $a_{k+1} = 2m - 1$. Modulo $2m - 1$, the inverse of m is 2, so the residue of a_{2k} is twice the residue of its numerator. Substituting this residue in the original sum computes the next remainder exactly. Its nonvanishing proves $N(k) = 2k + 1$.

Here is a complete C++17 verifier. The integers in its products are residues less than 8191, and each sum has at most 4095 terms. Every sum is therefore less than $4095 \cdot 8190^2 < 2^{39}$, within the range of `uint64_t`. The output is `8184 49 0`: 8184 checked orders, 49 with first failure at $2k + 1$, and zero orders passing both tests.

```
#include <cstdint>
#include <iostream>
#include <vector>
using U = std::uint64_t;
U numerator(const std::vector<U>& a, int n, int k) {
    U v = 0;
    for (int j = 1; j <= k/2; ++j)
        v += a[n-j] * a[n-k+j];
    return v;
}
std::vector<U> block(int k, U q) {
    std::vector<U> a(2*k+2, 1);
    for (int n = k; n < 2*k; ++n)
        a[n] = numerator(a,n,k) % q;
    return a;
}
int main() {
    int checked = 0, delayed = 0, escaped = 0;
    for (int k = 8; k <= 8191; ++k) {
        U m = k/2;
        auto a = block(k,m);
        U r = numerator(a,2*k,k) % m;
        if (r == 0) {
            U q = 2*m-1;
            a = block(k,q);
            a[2*k] = 2*(numerator(a,2*k,k) % q) % q;
            r = numerator(a,2*k+1,k) % q;
            if (r != 0) ++delayed;
            else ++escaped;
        }
        ++checked;
    }
    std::cout << checked << ' ' << delayed << ' '
              << escaped << '\n';
    return escaped != 0;
}
```

For every k, the terms preceding a_{2k} are integers, since the only divisors used there are initial entries equal to 1. The finite verification and the estimate for $k \geq 8192$ prove both bounds in Theorem 2.1.

3. Prime powers in Somos-4

In this section,

$$a_0 = a_1 = a_2 = a_3 = 1, \qquad a_{n+4}a_n = a_{n+3}a_{n+1} + a_{n+2}^2 \qquad (n \geq 0).$$

For a prime p, v_p denotes the valuation normalized by $v_p(p) = 1$.

Theorem 3.1. *If p is an odd prime and $p \mid a_n$, then, for every integer $j \geq 0$,*

$$n_j = \frac{(2n-3)p^j + 3}{2}, \qquad v_p(a_{n_j}) = v_p(a_n) + j.$$

Moreover,

$$v_2(a_n) = \begin{cases} 1, & n \equiv 4 \pmod 5, \\ 0, & n \not\equiv 4 \pmod 5, \end{cases} \qquad 37 \nmid a_n.$$

We give the coordinate and local calculations explicitly. The elliptic curve correspondence is due to the earlier work discussed in the introduction; the coordinates below are those in [10, proof of Corollary 1.2].

3.1. The elliptic curve and its denominators.

Set

$$F(a,b,c,d) = a^2d^2 - 4abcd + ac^3 + b^3d + b^2c^2.$$

With $e = (bd + c^2)/a$, expansion gives

$$F(b,c,d,e) = \frac{e}{a}F(a,b,c,d). \tag{3.1}$$

The recurrence first defines positive rational numbers. Since $F(1,1,1,1) = 0$, equation (3.1) proves that F vanishes on every consecutive quadruple.

We now prove simultaneously that all terms are integers and that the terms in every consecutive quadruple are pairwise coprime. Suppose this holds for a, b, c, d. Reduction of $F = 0$ modulo a gives $b^2(bd + c^2) \equiv 0 \pmod a$. Since $\gcd(a, b) = 1$, the next term $e = (bd + c^2)/a$ is an integer. A prime dividing e and b, or e and d, would divide c. A prime dividing e and c would divide bd. Each possibility contradicts pairwise coprimality of a, b, c, d. Thus b, c, d, e are pairwise coprime, completing the induction.

Consider

$$E: \ y^2 + y = x^3 - x, \qquad P = (0,0).$$

A bar denotes reduction modulo the prime under consideration. The identity of the group law is denoted by O, and $[m]P$ denotes the integer multiple of P. The equation has discriminant 37, and $2P = (1,0)$. It is minimal at every prime: discriminant valuations of isomorphic integral Weierstrass equations differ by a multiple of 12, whereas the present valuations are at most 1. For $n \geq 1$, define

$$X_n = \frac{a_n^2 - a_{n-1}a_{n+1}}{a_n^2}, \qquad Y_n = \frac{a_{n-1}^2a_{n+2} - 2a_{n-1}a_na_{n+1}}{a_n^3}. \tag{3.2}$$

We claim that

$$(X_n, Y_n) = [2n-3]P. \tag{3.3}$$

For its verification, write

$$X = \frac{b^2 - ac}{b^2}, \qquad Y = \frac{a^2d - 2abc}{b^3}, \qquad (a,b,c,d) = (a_{n-1}, a_n, a_{n+1}, a_{n+2}).$$

Then

$$Y^2 + Y - X^3 + X = \frac{a^2}{b^6}F(a,b,c,d) = 0.$$

Positivity gives $X \neq 1$. The line through (X, Y) and $(1, 0)$ has slope $\lambda = (2bc - ad)/(bc)$. Hence their sum has coordinates

$$X_+ = \lambda^2 - X - 1, \qquad Y_+ = -Y - 1 + \lambda(X - X_+).$$

Let X', Y' be the expressions for X, Y with (a, b, c, d) replaced by (b, c, d, e). The two required identities are

$$X_+ - X' = \frac{F(a,b,c,d)}{b^2c^2}, \qquad Y_+ - Y' = \frac{a^2d - 2abc - b^3}{ab^3c^3}F(a,b,c,d).$$

They prove that $(X_{n+1}, Y_{n+1}) = (X_n, Y_n) + 2P$. Since $(X_1, Y_1) = (0, -1) = -P$, equation (3.3) follows.

If a rational point has $v_p(x) < 0$, its equation forces $v_p(y) < 0$ and $2v_p(y) = 3v_p(x)$: on each side the term of lowest valuation is unique. Thus every negative valuation of x is even. For every multiple $[m]P \neq O$, write its reduced x-coordinate as

$$x([m]P) = \frac{A_m}{B_m^2}, \qquad A_m \in \mathbb{Z}, \quad B_m > 0, \quad \gcd(A_m, B_m) = 1.$$

Negation sends (x, y) to $(x, -y - 1)$, so $B_{-m} = B_m$. Pairwise coprimality yields

$$\gcd(a_n^2 - a_{n-1}a_{n+1}, a_n) = 1.$$

Consequently,

$$B_{|2n-3|} = a_n \qquad (n \geq 0). \tag{3.4}$$

At $n = 0$ this says $B_3 = 1$, as follows from $3P = (-1, -1)$. The construction already shows that each odd multiple appearing here is finite. The next calculation also shows that P has infinite order.

3.2. **A local valuation calculation.** We recall the formal group of an integral Weierstrass equation at O; see [16, Chapter IV]. Put $t = -x/y$ and $w = -1/y$. For the present equation,

$$w = t^3 + (1 - t)w^2.$$

Successive comparison of coefficients gives $w = t^3U(t)$, with $U(t) \in 1 + t\mathbb{Z}[[t]]$. The chord-and-tangent law at O therefore defines a one-dimensional commutative formal group

$$\mathcal{F}(T_1, T_2) \in \mathbb{Z}[[T_1, T_2]], \qquad \mathcal{F}(T_1, T_2) = T_1 + T_2 + \text{terms of total degree at least 2.}$$

For any prime p, the points reducing to O form the group parametrized by $t \in p\mathbb{Z}_p$, with $t = 0$ corresponding to O. For $t \neq 0$,

$$x = t^{-2}U(t)^{-1}, \qquad y = -t^{-3}U(t)^{-1}. \tag{3.5}$$

Here $U(t)$ is a unit. Conversely, $v_p(x(R)) < 0$ gives $v_p(y(R)) = 3v_p(x(R))/2$, so the projective point $(x(R) : y(R) : 1)$ reduces to $(0 : 1 : 0) = O$. Thus, if $x(R) = A/B^2$ is reduced, then $p \mid B$ exactly when R reduces to O, and in this case $v_p(t(R)) = v_p(B)$. This construction uses the smooth point O also when $p = 37$.

Let R be a nonzero point in this formal neighborhood and put $e = v_p(t(R)) \geq 1$. For an integer u prime to p, the integral multiplication series

$$[u]_{\mathcal{F}}(T) = uT + O(T^2)$$

shows that $v_p(t([u]R)) = e$.

For odd p, the precise formula for every nonzero integer u is

$$v_p(t([u]R)) = e + v_p(u). \tag{3.6}$$

Here is the convergence and valuation argument. The invariant differential of the formal group has the form

$$\omega(T) = (1 + c_1T + c_2T^2 + \cdots)\, dT, \qquad c_i \in \mathbb{Z}_p.$$

Its integral

$$\log_{\mathcal{F}}(T) = T + \sum_{\nu \geq 2} \frac{c_{\nu-1}}{\nu} T^\nu$$

satisfies $\log_{\mathcal{F}}(\mathcal{F}(T_1, T_2)) = \log_{\mathcal{F}}(T_1) + \log_{\mathcal{F}}(T_2)$. For $e \geq 1$, odd p, and $\nu \geq 2$,

$$\nu e - v_p(\nu) \geq e + 1.$$

Indeed $v_p(2) = 0$; for $\nu \geq 3$, the inequality $v_p(\nu) \leq \nu - 2$ follows from $p \geq 3$. The estimates tend to infinity with ν, so the series converges and preserves the valuation of every nonzero element of $p\mathbb{Z}_p$. Moreover,

$$\log_{\mathcal{F}}([u]_{\mathcal{F}}(t)) = u \log_{\mathcal{F}}(t).$$

For $u \neq 0$ and $t \neq 0$, the right-hand side is nonzero. Consequently $[u]_{\mathcal{F}}(t) \neq 0$, and taking valuations proves (3.6). This also proves that the formal neighborhood has no nonzero torsion for odd p.

Equation (3.2) gives $7P = (-5/9, 8/27)$, whose parameter $15/8$ has positive 3-adic valuation. Hence $7P$, and therefore P, has infinite order. Combining (3.5) and (3.6) gives

$$v_p(B_{mu}) = v_p(B_m) + v_p(u) \quad \text{if } p \text{ is odd and } p \mid B_m. \tag{3.7}$$

This is the elliptic-denominator multiplication law [3, Lemma 3.1], in the square-root denominator convention.

Proof of Theorem 3.1. If an odd prime p divides a_n, then $n \geq 4$. Set $m = 2n - 3 > 0$. Equations (3.4) and (3.7), with $u = p^j$, give

$$v_p(B_{mp^j}) = v_p(a_n) + j.$$

The integer mp^j is positive and odd, so its corresponding Somos index is $n_j = (mp^j + 3)/2$. This proves the first assertion.

At $p = 2$, the reduced curve has exactly five points:

$$O, \quad (0,0), \quad (0,1), \quad (1,0), \quad (1,1).$$

Thus the reduction of P has order 5. Good reduction and the kernel of reduction give $2 \mid B_m$ if and only if $5 \mid m$. Since $5P = (1/4, -5/8)$, one has $v_2(B_5) = 1$. For odd $m = 5u$, multiplication by the odd integer u preserves this valuation. Therefore $v_2(B_m) = 1$ for odd multiples of 5, and is zero at all other odd indices. The condition $5 \mid 2n - 3$ is $n \equiv 4 \pmod 5$.

At 37, the unique singular affine point is $(5, 18)$. The reduction of P lies in the smooth part. The addition law there gives

m	1	2	4	8
$[m]\overline{P}$	$(0,0)$	$(1,0)$	$(2,34)$	$(26,3)$
m	16	18	19	38
$[m]\overline{P}$	$(15,31)$	$(31,33)$	$(27,18)$	O

The slopes for the first four doublings are $36, 2, 20, 20$, respectively, modulo 37. Adding $2\overline{P}$ to $16\overline{P}$ gives $18\overline{P}$ with slope 26, and adding $\overline{P}$ gives $19\overline{P}$ with slope 13. The point $(27, 18)$ is a smooth nonidentity point with $2y + 1 = 0$, so it has order 2. Since $2\overline{P} \neq O$ and $19\overline{P} \neq O$, the order of $\overline{P}$ is 38. Points with smooth reduction form a subgroup, and reduction on this subgroup is a homomorphism whose kernel is the formal neighborhood of O; see [16, Proposition VII.2.1]. Every multiple of P belongs to this subgroup. Hence $37 \mid B_m$ if and only if $38 \mid m$. An odd index $m = 2n - 3$ cannot satisfy this condition. □

The same argument describes every prime-power divisibility index.

Corollary 3.2. *Let $p \neq 2, 37$ be prime and let d_p be the order of $\overline{P}$ in $E(\mathbb{F}_p)$. If d_p is even, then $p \nmid a_n$ for every n. If d_p is odd, put $w_p = v_p(B_{d_p})$. Then*

$$v_p(a_n) = \begin{cases} 0, & d_p \nmid 2n - 3, \\ w_p + v_p((2n-3)/d_p), & d_p \mid 2n - 3. \end{cases}$$

When d_p is odd, for every integer $s \geq 1$,

$$p^s \mid a_n \quad \Longleftrightarrow \quad d_p p^{\max(0, s - w_p)} \mid 2n - 3.$$

Proof. Good reduction gives $p \mid B_m$ exactly when $d_p \mid m$. For $m = d_p u$, apply (3.7); then use (3.4). Since $d_p p^{\max(0,s-w_p)}$ is odd, the displayed divisibility condition describes one residue class of n modulo that integer. □

4. Three-term polynomial recurrences

Let x, y, z be algebraically independent. For integers

$$0 < p < q < r \le k/2,$$

define $a_0 = \cdots = a_{k-1} = 1$ and

$$a_n a_{n-k} = x a_{n-p} a_{n-k+p} + y a_{n-q} a_{n-k+q} + z a_{n-r} a_{n-k+r} \qquad (n \ge k). \tag{4.1}$$

These rational functions are nonzero: their values at positive real x, y, z are positive, by induction. Thus the recurrence is defined for every n.

Theorem 4.1. *All terms of* (4.1) *belong to* $\mathbb{Z}[x, y, z]$ *if and only if*

$$r = p + q \qquad \text{or} \qquad k = p + q + r. \tag{4.2}$$

If neither equality holds and N *is the first index for which* $a_N \notin \mathbb{Z}[x, y, z]$*, then*

$$2k \le N \le 2k + p.$$

In lowest terms, the denominator of a_N*, normalized to be one of the following primitive polynomials, is*

$$\begin{cases} x + y + z, & N < 2k + p, \\ x(x + y + z) + y + z, & N = 2k + p. \end{cases}$$

The sufficiency in (4.2) is the three-term Gale–Robinson Laurent theorem of Fomin and Zelevinsky [4, Theorem 1.4]. For $r = p + q$, replace the third offset by $k - r$; the two factors in its summand are interchanged, and $p + q + (k - r) = k$. Specializing the initial Laurent variables to 1 gives coefficients in $\mathbb{Z}[x, y, z]$. We prove necessity, including the assertion about the first denominator, in the following subsections.

4.1. **The first two possible denominators.** Put

$$S = x + y + z, \qquad T = xS + y + z, \qquad b_t = a_{k+t} \quad (0 \le t < k),$$

and set $b_t = 1$ for $t < 0$. The first block consists of polynomials and satisfies

$$b_t = x b_{t-p} b_{t-k+p} + y b_{t-q} b_{t-k+q} + z b_{t-r} b_{t-k+r} \quad (0 \le t < k). \tag{4.3}$$

In particular, $b_0 = \cdots = b_{p-1} = S$ and $b_p = T$. For $0 \le t < p$, let

$$R_t = x b_{k+t-p} b_{t+p} + y b_{k+t-q} b_{t+q} + z b_{k+t-r} b_{t+r}.$$

Every index here is less than k. Directly from (4.1),

$$a_{2k+t} = \frac{R_t}{S} \quad (0 \le t < p), \qquad a_{2k+p} = \frac{U}{ST}, \tag{4.4}$$

where

$$U = x R_0 b_{2p} + S(y b_{k+p-q} b_{p+q} + z b_{k+p-r} b_{p+r}). \tag{4.5}$$

Consequently it suffices to prove that some R_t is nonzero modulo S, or that U is nonzero modulo T. The polynomials S, T are coprime and primitive irreducibles: $T = (x + 1)S - x$, and, as a polynomial in z, its coefficients $x + 1$ and $x^2 + (x + 1)y$ are coprime. If the first failure is in the first part of (4.4), its reduced denominator is therefore S. Otherwise $S \mid R_0$, whence $S \mid U$, and its reduced denominator is T. Gauss's lemma shows that polynomial quotients here have integer coefficients. This also proves that the first failure over $\mathbb{Q}[x, y, z]$ and over $\mathbb{Z}[x, y, z]$ is the same.

4.2. **A two-delay coefficient formula.** For positive integers $p < q$, define

$$C_n = 1 \ (n < 0), \qquad C_n = x(C_{n-p} - C_{n-q}) \ (n \geq 0).$$

Its generating function is

$$\sum_{n\geq 0} C_n u^n = \frac{-xu^p(1-u^{q-p})}{(1-u)(1-xu^p+xu^q)}. \tag{4.6}$$

Indeed, summing the recurrence and collecting its negative-index terms gives this identity. Expanding the geometric denominator, the coefficient of x^m is the negative of $u^{mp}(1-u^{q-p})^m/(1-u)$. For $n \geq 0$ and $m \geq 1$, the identity $\sum_{j=0}^{v}(-1)^j\binom{m}{j} = (-1)^v\binom{m-1}{v}$ for $0 \leq v < m$ gives

$$[x^m]C_n = -(-1)^v\binom{m-1}{v}, \qquad v = \left\lfloor \frac{n-mp}{q-p} \right\rfloor, \tag{4.7}$$

when $0 \leq v < m$; the coefficient is zero otherwise. In particular, if $q \geq 2p$ and $n \geq p$, then $\deg C_n = \lfloor n/p \rfloor$ and its leading coefficient is -1. If $q = p+d < 2p$ and $n \geq 0$, write $n = mp+u$, $0 \leq u < p$. Then

$$C_n \neq 0 \quad \Longleftrightarrow \quad u < md. \tag{4.8}$$

If $u < md$, then $v = \lfloor u/d \rfloor < m$ makes $[x^m]C_n$ nonzero. Conversely, $u \geq md$ implies $n \geq mq$; every coefficient of degree at most m then vanishes by (4.7), and no higher degree is possible.

4.3. **The six-point identity on $S = 0$.** Suppose $r \geq 2p$. Work in the domain $\mathbb{Q}[x,y,z]/(S)$, and write $B_t = b_t \bmod S$. Thus $B_t = 1$ for $t < 0$ and $B_t = 0$ for $0 \leq t < p$. Set

$$d = q-p, \qquad f = r-q, \qquad e = k-q-r.$$

The inequality $k \geq 2r$ gives $e \geq f > 0$. For $0 \leq t < p$, define

$$h = \mathbf{1}_{t\geq d}, \quad g = \mathbf{1}_{t<f}, \quad i = \mathbf{1}_{t<e},$$
$$A = B_{e+d+t}, \ G = B_{d+f+t}, \ D = B_{e+d+f+t},$$
$$E = B_{d+t}, \ F = B_{e+t}, \ H = B_{f+t}.$$

Lemma 4.2. *In the quotient by S,*

$$R_t = xyzK_t, \qquad K_t = h(DE-AG) + (1-h)gD + EFH + iAH + gGF - gFH. \tag{4.9}$$

Proof. The recurrence (4.3) gives

$$B_{t+p} = -x - hy, \qquad B_{k+t-p} = yDE + zAG,$$
$$B_{k+t-q} = x(1-h)D + zFH, \qquad B_{t+q} = xE + zgi,$$
$$B_{k+t-r} = xA + ygF, \qquad B_{t+r} = xG + yiH.$$

For example, the missing summand in B_{k+t-p} contains $B_t = 0$; the small factors in the other two equations are either 0 or 1, according to the displayed indicators. The bound $t+p < 2p \leq r$ justifies the first equation. Also $gi = g$, since $e \geq f$. Substitution in R_t and use of $x+y+z = 0$ gives (4.9). The points A, G, E, F, H all have indices less than $k-r$. Only D can encounter the nonlinear part of (4.3). □

Because xyz is nonzero in this domain, it is enough to prove that one K_t is nonzero. A nonzero specialization of K_t to any coefficient axis suffices.

Lemma 4.3. *If $r \geq 2p$, specialize $y = 0, z = -x$ in the first block. The resulting polynomial B_t is nonzero for every $p \leq t < k$.*

Proof. For $p \leq t < k-p$, the p summand is xB_{t-p}. When the r summand has a negative index, it has lower degree than $\lfloor t/p \rfloor$, except that the initial interval $p \leq t < 2p$ gives $B_t = -x$ directly. When both its indices are nonnegative and outside the zero block, its degree is at most $1 + \lfloor (2t-k)/p \rfloor$. This is strictly less than $\lfloor t/p \rfloor$ for $t \leq k-2p$. For $k-2p < t < k-p$, its two predecessors and the p predecessor lie below $k-2p$ and have leading coefficient -1.

Hence this summand adds another negative leading coefficient if it reaches the largest degree. Induction gives degree $\lfloor t/p \rfloor$ and leading coefficient -1 or -2 throughout $p \le t < k-p$. For $k-p \le t < k$, the p summand vanishes, while both indices of the r summand lie in $[p, k-2p-1]$. Its product, multiplied by $-x$, has leading coefficient -1 and is nonzero. □

If $k < p+q+r$, then $f \le e < p$. At $t=0$, $F = H = 0$, $h = 0$ and $g = i = 1$ in (4.9), so $K_0 = D$. Its index $k-p-q = e+r-p$ lies in $[p, k-1]$. Lemma 4.3 proves $K_0 \neq 0$. Thus, for $r \ge 2p$, only $k > p+q+r$ remains to be considered outside the two families (4.2).

4.4. **The region $2p \le r < p+q$.** First suppose $r = 2p$ and put $d = q-p$, $e = k-q-2p > p$. At $y = 0$, the linear $p, 2p$ sequence is constant on blocks of length p, with block values $P_0 = 0$, $P_1 = -x$, $P_m = x(P_{m-1} - P_{m-2})$ for $m \ge 2$. Their leading terms are $-x^m$ for $m \ge 1$. If $e \ge 2p$, equation (4.9) at $t = 0$ gives $K_0 = -xC_{e-p} \neq 0$ on this axis. If $e = p+a$ with $1 \le a < p$, the following choices cover all possibilities:

Condition	Specialization and index	Nonzero value or coefficient
$a \ge 2d$	$x = 0$, $t = 0$	$K_0 = -y^2$
$d \le a < 2d$	$z = 0$, $t = 0$	$K_0 = x^2$
$a < d$, $a > 2d-p$	$z = 0$, $t = d-a$	$K_t = x^2$
$a < d$, $a \le 2d-p$	$z = 0$, $t = p-1$	$[x^2]K_t = -1$

Here are the coefficient checks. In the first row $K_0 = B_{2p+a}$ and $2q \le 2p+a < q+2p$, giving $-y^2$ from the $q, 2p$ recurrence. In the next two rows (4.9) reduces to $-xC_{e+t-d}$, whose indicated index lies in $[p, q)$. For the last row $2d > p$, and (4.9) gives

$$K_{p-1} = -x(C_{2p+a+d-1} - C_{2p+a-d-1}).$$

The second index lies in $[p, q)$, whereas the first is at least q; (4.7) says that their coefficients of x are respectively -1 and 0. All the indices used on the $z = 0$ axis are less than $k-q = e+2p$, so that the linear recurrence applies.

Now let $2p < r < p+q$. If $q \ge 2p$, take $t = p-1$ and specialize $z = 0$. With $n = e+p-1 \ge 2p$, (4.9) becomes

$$K_{p-1} = -x(C_{n+q-p} + C_{q-1}C_n).$$

Write $q = Mp+b$ with $1 \le b \le p$ and $M = \lfloor (q-1)/p \rfloor$, and $n = mp+u$ with $m \ge 2$, $0 \le u < p$. Then $C_{q-1} = -(x + \cdots + x^M)$. If $u + b < p$, the product in parentheses has uniquely largest degree. If $u + b \ge p$, formula (4.7) gives the following first nonzero term after cancellation of the largest terms:

Condition	Coefficient or identity for $J = C_{n+q-p} + C_{q-1}C_n$
$M \ge 3$	$[x^{m+M-1}]J = 1$
$M = 1$ $(q = 2p)$	$J = -xC_{n-p}$
$M = 2$, $u < b$	$[x^{m+1}]J = 1$
$M = 2$, $u \ge b$, $m > 2$	$[x^{m+1}]J = 2-m$
$M = m = 2$, $u \ge b$	$J = x^2$

For the last identity the two polynomials are $C_n = -x^2$ and $C_{n+q-p} = -x^4 - x^3 + x^2$; all other rows follow by using $v = 0$ or 1 in (4.7) for the two largest possible degrees. Each row is nonzero.

If $q < 2p$, write $q = p+d$, $r = 2p+\delta$ with $0 < \delta < d < p$, and use the $y = 0$ axis. At $t = p-1$, the six-point values satisfy $E = H = -x$, $G = -x-x^2$, and $D = xC_{n+\delta}$, where C_n is the linear p, r sequence and $n = e+p-1$. The r summand in D vanishes because its small factor is $B_{f-1} = 0$. Thus

$$K_{p-1} = x^2(C_n + C_{n+d} - C_{n+\delta}).$$

Write $n = mp+u$, $m \ge 2$, $0 \le u < p$. If the three indices stay in one p block, the leading coefficient in parentheses is -1. If only $n+d$ crosses to the next block, its largest term is

unique. If both shifted indices cross, their leading terms cancel; (4.7) gives coefficient -1 or $m-1$ at x^m. These values are nonzero. This settles $2p \le r < p+q$.

4.5. **A linear kernel for $q \ge 2p$.** We next prove the remaining S obstruction when $r > p+q$ and $q \ge 2p$. Here $e \ge f > p$. On the $z=0$ axis, all six entries in (4.9) are values of the linear p,q sequence. For $0 \le t < p$, put

$$E = C_{q-p+t}, \quad F = C_{e+t}, \quad H = C_{f+t},$$
$$A = C_{e+q-p+t}, \quad G = C_{f+q-p+t}, \quad D = C_{e+f+q-p+t}.$$

The specialized kernel is

$$K_t = D + EFH + AH + GF - FH. \tag{4.10}$$

We show that these p polynomials cannot all be zero.

Write $e = \alpha p + a$, $f = \beta p + b$, $q - p = \gamma p + c$, where $0 \le a, b, c < p$. Suppose first $a, b > 0$ and, by symmetry, $a \le b$. After subtracting the common degree $\alpha + \beta + \gamma$, the degrees of D, EFH, AH, GF are respectively

$$\left\lfloor \frac{a+b+c+t}{p} \right\rfloor, \quad \left\lfloor \frac{a+t}{p} \right\rfloor + \left\lfloor \frac{b+t}{p} \right\rfloor + \left\lfloor \frac{c+t}{p} \right\rfloor,$$
$$\left\lfloor \frac{a+c+t}{p} \right\rfloor + \left\lfloor \frac{b+t}{p} \right\rfloor, \quad \left\lfloor \frac{b+c+t}{p} \right\rfloor + \left\lfloor \frac{a+t}{p} \right\rfloor.$$

Their leading coefficients are $-1, -1, 1, 1$. The term $-FH$ is of smaller degree than EFH. Let $w = (c+t) \bmod p$. Comparing the four floors, the sum of leading coefficients at their largest degree vanishes only for the following w:

Range of t	Allowed w
$[0, p-b)$	$[0, p-a-b) \cup [p-b, p-a)$
$[p-b, p-a)$	$[0, p-a-b) \cup [p-b, p-a) \cup [2p-a-b, p)$
$[p-a, p)$	$[p-b, p-a) \cup [2p-a-b, p)$

Intervals are intersected with $[0, p)$. For verification, put $A_0 = \mathbf{1}_{t \ge p-a}$ and $B_0 = \mathbf{1}_{t \ge p-b}$. The four degrees, after removing their common $\lfloor (c+t)/p \rfloor$, are

$$\left\lfloor \frac{a+b+w}{p} \right\rfloor, \quad A_0 + B_0, \quad \mathbf{1}_{w \ge p-a} + B_0, \quad \mathbf{1}_{w \ge p-b} + A_0;$$

the table is obtained by taking $(A_0, B_0) = (0,0), (0,1), (1,1)$. If all t gave cancellation, rotating the first interval by c forces it into $[p-b, p-a)$: its other allowed component has smaller length. Similarly the last interval must rotate into the same middle component. The resulting inequalities are

$$p - b \le c \le b - a, \qquad p + a - b \le c \le p - a.$$

They imply $a < p/2 < b$. The middle interval, of length $b-a$, also cannot fit in either outside component of its allowed set; their lengths are strictly smaller. Components are separated by positive gaps, including across the endpoint 0, so a rotated interval cannot straddle two of them. Hence the middle interval must rotate to itself, giving $c = 0$, in contradiction to $c \ge p - b > 0$.

It remains to treat one zero residue. By symmetry take $e = \alpha p$ with $\alpha \ge 2$. Initially put $t = p-1$, write $H = C_{hp+w}$ with $h \ge 2$, $0 \le w < p$, and set

$$J_n = C_{n+\alpha p} + FC_n.$$

Equation (4.10) reads

$$K_t = J_{f+q-p+t} + HJ_{q-p+t} - FH. \tag{4.11}$$

All coefficient calculations below use (4.7).

If $q \ge 3p$, every C_n of degree $m \ge 2$ begins $-x^m - x^{m-1} + \cdots$. Thus J_n has leading coefficient 1 and degree $\alpha + \deg C_n - 1$. Let $g = \lfloor (q-p+t)/p \rfloor$ and $j = \lfloor (f+q-p+t)/p \rfloor$. If $j \ne g + h$, the first two terms in (4.11) have different degrees, and their larger degree

exceeds that of FH. If $j = g + h$, their next coefficients give a nonzero value -1. Here are the endpoint checks of that assertion. Write $F = -x^{\alpha}(1 + x^{-1} + B_2x^{-2} + \cdots)$. For $q - p = \gamma p + c$ with $\gamma \geq 3$, the next coefficients of the two J terms, before multiplication by H, are both $B_2 + 1$. For $\gamma = 2, c > 0$, degree equality forces the relevant residues below c, and the same values occur. For $\gamma = 2, c = 0$, they are $B_2 + \alpha$ for J_{q-p+t} and $B_2 + \alpha + 1$ for $J_{f+q-p+t}$; their combination cancels, but $-FH$ contributes -1 at exactly this degree. These values can also be read directly from the following expansion. For $n = mp + v$, $m \geq 2$, $0 \leq v < p$, write $C_n = -x^m(1 + x^{-1} + c_2(n)x^{-2} + \cdots)$. Then $c_2(n) = 0$ for $m = 2$. For $m \geq 3$, it is 1 if $q - p \geq 3p$, and, when $q - p = 2p + c$, it is 1 for $v < c$ and $3 - m$ for $v \geq c$. The coefficient following the leading term of J_n is $B_2 + c_2(n) + 1 - c_2(n + \alpha p)$, which gives all three cases above.

If $q = 2p$, let $H_{-1} = 0, H_0 = 1$ and $H_m = xH_{m-1} - xH_{m-2}$ for $m \geq 1$. Then $C_n = -xH_{\lfloor n/p \rfloor - 1}$ for $n \geq p$, and

$$K_{p-1} = -x^3 H_{\alpha-2} H_{h-2} \neq 0. \tag{4.12}$$

For a direct proof of the identity, keep h fixed. Both sides satisfy the same second-order recurrence in α. At $\alpha = 1$ the kernel is zero, and at $\alpha = 2$ it is $x^2 C_{(h-1)p} = -x^3 H_{h-2}$. These initial values prove (4.12); H_m is monic for every $m \geq 0$.

Finally suppose $q = 2p + c$, $0 < c < p$. At $t = p - 1$, $E = -x^2 - x$ and $F = -x^{\alpha}(1 + (2 - \alpha)x^{-1} + B_2x^{-2} + \cdots)$. If $\alpha \geq 3$ and $w < p - c$, the term HJ_{q-p+t} in (4.11) has uniquely largest degree, with coefficient $\alpha - 2$. If $w \geq p - c$, the leading terms cancel and the next coefficient is $(\alpha - 2)c_1(H)$, where $H = -x^h(1 + c_1(H)x^{-1} + \cdots)$. Formula (4.7) shows this can vanish only when $h = 2$ and $w \geq c$. In that case write $f = p + b$, where $\max(c, p - c) < b \leq p$, and choose $t = p - b$. Then $0 \leq t < \min(c, p - c)$, $E = -x$, $H = -x^2 - x$, and

$$J_{q-p+t} = \alpha x^{\alpha} + x^{\alpha-1} + \cdots, \qquad J_{f+q-p+t} = (\alpha + 1)x^{\alpha+2} + 2x^{\alpha+1} + \cdots.$$

The coefficient of $x^{\alpha+1}$ in the kernel is $2 - (\alpha + 1) - 2 = -\alpha - 1$.

For $\alpha = 2$, two applications of the recurrence give, at $t = p - 1$,

$$D = x^2G - 2x^2H + x^2C_{f-c-1}, \qquad K_{p-1} = x^2 C_{f-c-1}.$$

This is nonzero if $f \geq p + c + 1$. Otherwise $f = p + b$, $1 \leq b \leq c$. Choose $t = c - b$ and put $\eta = \mathbf{1}_{2c-b \geq p}$, $\theta = \mathbf{1}_{2c \geq p}$. The first few blocks give

$$E = -x - \eta x^2, \quad H = -x, \quad F = -x^2 - x, \quad G = -x^2 - \theta x^3.$$

Using the recurrence in (4.11) now gives $K_t = -xG + x^2E - x^2 = -x^2 + (\theta - \eta)x^4 \neq 0$. The linear kernel is therefore nonzero for some t in every case.

4.6. **Restriction to the divisor $T = 0$.** For the remaining regions, set

$$s = \frac{x}{1 + x}, \qquad z = z_0 - y, \qquad z_0 = -xs,$$

so that $S = s$ and $T = 0$. We work in $\mathbb{Q}(x)[y]$. A nonzero coefficient of U after this substitution proves $T \nmid U$. Write $u_t = b_t|_{y=0}$ after the substitution, and define the auxiliary linear sequence

$$l_t = 1 \ (t < 0), \qquad l_t = xl_{t-p} + z_0 l_{t-r} \ (t \geq 0).$$

Summing its recurrence gives

$$\sum_{t \geq 0} l_t v^t = \frac{1}{1 - v} - \frac{1}{(1 + x)(1 - v)(1 - xv^p - z_0 v^r)}. \tag{4.13}$$

For later use put $A_m = 1 + x + \cdots + x^{m-1}$ for $m \geq 1$ and $A_0 = 0$. For $0 \leq t < r$, the values are constant on p blocks:

$$l_t = U_{\lfloor t/p \rfloor}, \qquad U_0 = s, \quad U_m = -xsA_{m-1} \quad (m \geq 1). \tag{4.14}$$

This follows from $U_m = xU_{m-1} + z_0$, starting with $U_0 = s$.

4.7. **The region $r < 2p$.** Here the allowed equality $r = p + q$ is impossible. Along $T = 0$,

$$b_t = s\ (0 \le t < p), \quad b_t = 0\ (p \le t < q), \quad b_t = y(s-1)\ (q \le t < r).$$

Put $h = k + p - q - r$ and

$$\theta = \begin{cases} s & k \le 2p + r, \\ 1 & k > 2p + r, \end{cases} \quad \eta = \begin{cases} s & k \le 2p + q, \\ 1 & k > 2p + q, \end{cases} \quad \psi = \begin{cases} s & k \le p + q + r, \\ 1 & k > p + q + r. \end{cases}$$

The first-block recurrence gives

$$\begin{aligned} u_{2p} &= z_0 s\theta, & [y]b_{2p} &= s(\eta - \theta), \\ u_{p+r} &= x\theta u_r, & [y]b_{p+r} &= x\theta [y]b_r, \\ u_{p+q} &= z_0 s\psi, & u_{k+p-q} &= xsu_{k-q}, \\ & & b_{k+p-r} &= xsb_{k-r} + ysb_h. \end{aligned}$$

For instance, the three small factors in the last equation have indices $2p - r, p + q - r, p$; their values are $s, s, 0$. In R_0 the p summand vanishes, while the q summand is divisible by y^2. Thus $R_0 = zb_{k-r}b_r + O(y^2)$. Substituting the displayed equations in (4.5), every derivative of b_r or b_{k-r} has the factor $z_0 + xs = 0$. The remaining coefficient is

$$[y]U = xz_0 s(u_r(\eta - \theta)u_{k-r} + s^2\psi u_{k-q} + s\theta u_r u_h). \tag{4.15}$$

The base sequence u_t is related to l_t by

$$u_t = \begin{cases} l_t, & t < k - r, \\ l_t + s^2 l_{t-r}, & k - r \le t < k - p, \\ sl_t, & k - p \le t < k - r + p, \\ xsu_{t-p}, & k - r + p \le t < k. \end{cases} \tag{4.16}$$

The r summand acquires its factor s at $k - r$, the p summand acquires its factor s at $k - p$, and the r summand vanishes at $k - r + p$. Before the last interval both predecessors are before the first altered boundary; use $z_0(s-1) = s^2$ to obtain the middle two lines. The early linear values needed below are

$$l_t = \begin{cases} s, & 0 \le t < p, \\ 0, & p \le t < r, \\ s^2, & r \le t < 2p, \\ z_0 s, & 2p \le t < p + r, \end{cases} \qquad l_{p+r} = xs^2.$$

If $2r \le k \le 2p + r$, let $\delta = \mathbf{1}_{h \ge r}$. Except at the boundary specified after the table, the bracket in (4.15) is as follows:

Condition	Bracket in (4.15)
$k < 2p + q$	$s^5[1 + (1 + \delta)s]$
$k = 2p + q$	$s^6(\delta - x)$
$2p + q < k < 2p + r$	$s^4[1 + (\delta - x - 1)s^2]$
$k = 2p + r$	$s^5(\delta s - x)$

To verify the table, use

$$u_h = \delta s^2, \qquad u_{k-r} = \begin{cases} s^2(1 + s) & k < 2p + r, \\ -xs^3 & k = 2p + r, \end{cases}$$

and $u_{k-q} = s^2(1 + s)$ for $k < 2p + q$, and $u_{k-q} = -xs^3$ otherwise. Usually $u_r = s^2$; at $k = 2r$ it is $s^2(1 + s)$ and $u_h = 0$. When $k = 2r > 2p + q$, the third row is accordingly replaced by

$$s^4\big((1 - s)(1 + s)^2 - xs^2\big).$$

Every displayed rational function is nonzero: the brackets with constant term 1 retain that term at $x = 0$, and the other cases are immediate for $\delta = 0, 1$.

For $2p+r<k\le p+q+r$, put $N=k-q$ and $d_0=r-p$. The bracket in (4.15) is

$$s^3(l_N+l_{N-d_0}+s^2l_{N-r}).$$

Here $2p<N\le p+r$. If $N<p+r$, the expression in parentheses is $s^2(z_0+\delta)$ for $\delta=0$ or 1, and is nonzero. At $N=p+r$ it is $xs^2+z_0s=0$; this is precisely $k=p+q+r$.

For $k>p+q+r$, the bracket is s^2g_N, where

$$g_N=l_N+sl_{N-r+p}+s^2l_{N-r},\qquad N=k-q>p+r. \tag{4.17}$$

We verify that $g_N\neq 0$ for all these N. Let $\epsilon=x+1$ and write $N=jr+h$, $0\le h<r$. Expanding (4.13) according to the number of r steps shows that the coefficient of ϵ^{-j-1} in l_{jr+h} is

$$L_j(h)=(-1)^{j+1}c_j(h),\qquad c_j(h)=\begin{cases}1 & h<p,\\ -j & h\ge p.\end{cases} \tag{4.18}$$

Indeed the layer with j such steps is

$$\frac{(-1)^j\epsilon^{-j}(1-\epsilon)^{2j}v^{jr}}{(1+(1-\epsilon)v^p)^{j+1}},$$

multiplied by $-\epsilon^{-1}/(1-v)$. Because $h<r<2p$, its residual coefficient uses at most one p step. All lower layers have smaller poles. If $h<r-p$, then $N>p+r$ forces $j\ge 2$, and the last term of (4.17) has a unique pole of order $j+2$. If $h\ge r-p$, the coefficient at that order is $2(-1)^j$ for $h<p$, and $(2-j)(-1)^j$ for $h\ge p$. Thus only $j=2$, $p\le h<r$ remains.

The next pole coefficient of l_{jr+h} is

$$M_j(h)=(-1)^{j+1}\bigl(-2jc_j(h)+(j+1)\mathbf{1}_{h\ge p}\bigr)+(-1)^j\sum_{a=0}^{\lfloor (r+h)/p\rfloor}(-1)^a\binom{j-1+a}{a}.$$

The first summand comes from the linear term in ϵ in the highest layer; the sum comes from the next layer. For $j=2$, $h\ge p$, the three terms in (4.17) contribute $2,-7,5$ to the pole of order 3 when $r+h<3p$, and $2,-7,6$ when $r+h\ge 3p$. The latter sum is 1. In the remaining range $r+h<3p$, define

$$H_t=\sum_{\substack{a,b\ge 0\\ ap+br\le t}}\binom{a+b}{a}x^az_0^b,\qquad l_t=1-\frac{H_t}{1+x}.$$

Only $b=0,1,2$ occur, and $m=\lfloor N/p\rfloor$ is 3 or 4. The three finite sums are

$$\begin{aligned}H_N&=\sum_{a=0}^{m}x^a+z_0(1+2x+3x^2)+z_0^2(1+3x),\\ H_{N-r+p}&=1+x+x^2+x^3+z_0(1+2x+3x^2)+z_0^2,\\ H_{N-r}&=1+x+x^2+z_0(1+2x).\end{aligned}$$

Their bounds follow from $N=2r+h$, $p\le h<r$, $r+h<3p$. Substitution gives $g_N=x^4/(1+x)^2$ for $m=3$ and $g_N=-x^5/(1+x)^2$ for $m=4$. This proves the required nonvanishing throughout $r<2p$.

4.8. **The region $p<q<2p$, $r>p+q$.** This is the only remaining parameter region. Set

$$d=q-p,\qquad f=r-q>p,\qquad e=k-q-r\ge f,$$

and retain the $T=0$ notation from Subsection 4.6. Since $r>p+q$, the first-block equations give

$$\begin{gathered}b_p=0,\quad b_q=y(s-1),\quad b_{2p}=z_0+y(s-1),\quad u_{p+q}=z_0,\\ b_{k+p-r}=xb_{k-r}+yb_{e+p},\qquad b_{p+r}=xb_r+yb_{p+f},\\ u_{k+p-q}=xsu_{k-q}+z_0u_{e+p}u_{p+f}.\end{gathered}$$

As in (4.15), substitution in U cancels the derivative terms containing $z_0 + xs$. It gives

$$[y]U = xz_0Q, \quad Q = (su_r - V)u_{k-r} + sVu_{e+p} + s^2u_{k-q}, \quad V = u_r - su_{p+f}. \tag{4.19}$$

For explicit verification of the cancellation, write $R_0 = zb_{k-r}b_r + O(y^2)$ and insert the preceding two equations for b_{k+p-r}, b_{p+r} into (4.5); the coefficients multiplying $[y]b_r$ and $[y]b_{k-r}$ both vanish. The remaining terms factor as xz_0 times the displayed Q.

Write $f = \beta p + b$, $0 \le b < p$. The precise boundary formulas are

$$\begin{aligned} u_{k-r} &= l_{e+q} + s^2 l_{e-f} = x(l_{e+d} - s^2 l_{e-f}), \qquad u_{e+p} = l_{e+p}, \\ u_{k-q} &= l_{e+r} + s^2 \sum_{i=0}^{\beta} x^i A_{\beta-i+1} l_{e-ip}. \end{aligned} \tag{4.20}$$

To obtain the last equation, the interval from $k-r$ to $k-q$ has length $f < r$, so every r predecessor stays before the altered boundary. At a point with small-factor block index m, the difference from the linear recurrence is $z_0(U_m - 1)l_{t-r} = s^2 A_{m+1} l_{t-r}$. Propagation by i steps of length p gives the stated finite sum. Also $u_r = l_r$ if $e > f$, whereas $u_r = l_r + s^3$ if $e = f$.

4.8.1. *The range $e < r$.* At $e = f$, $u_{k-r} = u_r$ has a pole of order 3 at $x = -1$ with coefficient -1, and $su_r - V$ has a pole of order 4 with coefficient 1. Their product has order 7, while the other terms of Q have order at most 5. If $f < e < r - d$, then $0 < e - f < p$ and $l_{e-f} = s$. Now u_{k-r} and $su_r - V$ each have a pole of order 3 with coefficient -1, producing the unique pole of order 6 in Q. Thus only

$$e = r - v, \qquad 1 \le v \le d$$

requires another calculation.

Write

$$\begin{gathered} f = \beta p + b, \quad c = \mathbf{1}_{b+d \ge p}, \quad \rho = b + d - cp, \quad r = (\beta + 1 + c)p + \rho, \\ H = \mathbf{1}_{\rho \ge v}, \quad C = \mathbf{1}_{\rho + d - v \ge p}, \quad \gamma = \left\lfloor \frac{2\rho - v}{p} \right\rfloor, \quad h = c + H. \end{gathered}$$

Here h is 1 or 2. The possible (c, H, C, γ) are the eight rows in the next table. To check completeness, $c = 0$ implies $\rho \ge d \ge v$ and $H = 1$; then $C = 1$ implies $\gamma = 1$. If $c = H = 1$, then $\rho < d$ and $\gamma = 1$ implies $C = 1$. If $c = 1, H = 0$, then $C = 0$ and $\gamma = -1$ or 0. By (4.14) and (4.20),

$$\begin{aligned} u_r &= U_{\beta+1+c} + s^2, & u_{k-r} &= x(U_{\beta+1+c+C} + s^2), \\ u_{e+p} &= U_{\beta+1+c+H} + s^2, & u_{p+f} &= U_{\beta+1}, \\ u_{k-q} &= U_{2\beta+2+2c+\gamma} + s^2(T_{\beta+h} - xsZ_{\beta,h}), \end{aligned} \tag{4.21}$$

where

$$\begin{aligned} T_m &= \sum_{j=0}^{m} (j+1)x^j = \frac{1 - (m+2)x^{m+1} + (m+1)x^{m+2}}{(1-x)^2}, \\ Z_{\beta,h} &= \sum_{i=0}^{\beta} x^i A_{\beta-i+1} A_{\beta+h-i-1} \\ &= \frac{(1 + x^{\beta+h})A_{\beta+1} - (\beta+1)(x^{\beta+1} + x^{\beta+h-1})}{(1-x)^2}. \end{aligned}$$

The first identity is the derivative of a finite geometric sum; for the second, expand each $A_m = (1 - x^m)/(1 - x)$ and sum its four terms. For $r \le n < 2r$, the expansion of (4.13) contains at most one r step, giving $l_n = U_{\lfloor n/p \rfloor} + s^2 T_{\lfloor (n-r)/p \rfloor}$; this proves (4.21).

Here is a polynomial certificate for the calculation of all eight rows. Put $B = x^\beta$ and $D_0 = (x-1)^3(x+1)^5$. Substituting (4.21) into (4.19) gives

$$D_0 Q = x^8(2x - 1) + BP_1 + B^2 P_2. \tag{4.22}$$

The complete coefficients are specified below. Define

$$\begin{aligned}
L_1 &= x^6(x-1)(\beta x^2+\beta x-2x^2-x-1),\\
L_2 &= x^6(x-1)(\beta x^2+\beta x-x^3-x^2-2),\\
L_3 &= x^7(x-1)(\beta x^2+\beta+x^2-4x),\\
L_4 &= x^7(x-1)(\beta x^2+\beta-x^3+2x^2-3x-1),\\
M_1 &= x^5-x^4-2x^3+2x^2+2x-1,\\
M_2 &= x^4-x^3-2x^2+2x+1.
\end{aligned}$$

Row	c	H	C	γ	P_1	P_2
A	0	1	0	0	L_1	$-x^6$
B	0	1	0	1	L_1	$-x^5M_1$
C	0	1	1	1	L_2	$-x^6$
D	1	1	0	0	L_3	$-x^7$
E	1	1	1	0	L_4	$-x^7$
F	1	1	1	1	L_4	$-x^8M_2$
G	1	0	0	-1	L_2	$-x^6$
H	1	0	0	0	L_2	$-x^7M_2$

This certificate can be checked by multiplication using the two finite-sum identities above; it contains the full dependence on β. Since $x^6 \mid P_1$ and $x^5 \mid P_2$, equation (4.22) proves $Q = x^8 + O(x^9)$ for every $\beta \geq 3$. For $\beta = 1, 2$, the same certificate gives

Row	$\beta = 1$	$\beta = 2$
A	$-x^7/(1+x)^5$	0
B	$-x^7(x^2+2x+2)/(1+x)^5$	$-x^9/(1+x)^3$
C	$-x^7(x+2)/(1+x)^5$	$-x^8/(1+x)^4$
D	$2x^8/(1+x)^5$	$x^8(3x+1)/(1+x)^5$
E	$-x^8(x-1)/(1+x)^5$	$-x^8(x^2-2x-1)/(1+x)^5$
F	$-x^8(x^3+2x^2+2x-1)/(1+x)^5$	$-x^8(x^5+2x^4+x^3+x^2-2x-1)/(1+x)^5$
G	$-x^7(x+2)/(1+x)^5$	$-x^8/(1+x)^4$
H	$-x^7(x^3+2x^2+2x+2)/(1+x)^5$	$-x^8(x^3+x^2+1)/(1+x)^4$

The sole zero case is row A with $\beta = 2$. Its parameter description is

$$f = 2p+b, \qquad r = 3p+d+b, \qquad e = 3p+a, \qquad b \leq a < b+d, \quad a+b+d < p. \tag{4.23}$$

Indeed $a = b+d-v$, and the condition $\gamma = 0$ is $2b+2d-v < p$; it also forces the other row-A conditions.

We exclude (4.23) using S. Specialize $z = 0, y = -x$ in the six-point kernel and use the linear p, q sequence C_n. For $d \leq t < p-d$, one has $E = C_{d+t} = 0$, $h = 1$ and $g = i = 1$. Write $G = C_{f+d+t}$ and $A = C_{e+d+t}$. Equation (4.9) then factors as

$$K_t = -AG + AH + GF - FH = -(A-F)(G-H).$$

The recurrence gives $C_{n+q} = x(C_{n+d} - C_n)$, so

$$K_t = -\frac{C_{e+q+t}C_{f+q+t}}{x^2}. \tag{4.24}$$

If $b \geq d$, choose $t = p-b-1$. The inequalities in (4.23) imply $p > 3d$ and $d \leq t < p-d$. The two indices in (4.24) are $5p+a-b+d-1$ and $4p+d-1$, whose p residues are less than $2d$ and d. Both polynomials are nonzero by (4.8). If $b < d$ and $p > 2d$, take $t = d$. The indices are $4p+a+2d$ and $3p+b+2d$. Before reducing modulo p, their excesses are less than $4d$ and $3d$; crossing a p block only decreases the residue and increases the block number. Again both are nonzero.

Finally, if $p \le 2d$, then $a+b<d$. At $t=0$, $E=0$ and $K_0 = D + AH + GF - FH$. Formula (4.7) gives

$$H=-x^2, \quad G=x^2, \quad A=2x^3, \quad F=-x^3+\eta x^2, \qquad \eta=\mathbf{1}_{p+a<2d},$$
$$D=4x^5+\delta x^4, \qquad \delta=\begin{cases}-3 & p+a+b+d<3d,\\ 1 & p+a+b+d\ge 3d.\end{cases}$$

There are no lower coefficients: for example the potential cubic coefficient of D vanishes because $2p+a+b+d>3d$. Thus $K_0=(\delta+2\eta)x^4\neq 0$. This finishes $e<r$.

4.8.2. *The range $e \ge r$.* Write $e=jr+h$, where $j\ge 1$ and $0\le h<r$. Expansion of (4.13) at $\epsilon = x+1=0$ gives the highest possible pole coefficient

$$[\epsilon^{-j-1}]l_{jr+h} = (-1)^{j+1}\mathcal{C}_j(a), \quad a=\lfloor h/p\rfloor, \quad \mathcal{C}_j(a)=\sum_{v=0}^{a}(-1)^v\binom{j+v}{v}. \tag{4.25}$$

Moreover,

$$(-1)^a\mathcal{C}_j(a) = [v^a]\frac{1}{(1-v^2)(1-v)^j}.$$

This identity follows by multiplying the generating series $\sum_{v\ge 0}\binom{j+v}{v}z^v = (1-z)^{-j-1}$ by $1/(1+z)$. For $j\ge 1$ the displayed coefficient is strictly positive; for $j=0$ it is 1 when a is even and 0 when a is odd. Also $u_r=l_r$ has a pole of order 2 with leading coefficient 1, while V has order at most 2. Consequently

$$su_r - V = -\epsilon^{-3}+O(\epsilon^{-2}).$$

If $h\ge f$, the term $s^2(su_r-V)l_{e-f}$ from (4.20) has a unique pole of order $j+6$. Its coefficient is nonzero by (4.25); the other terms have order at most $j+5$.

Suppose $h<f$ and write $h=ap+u$, $0\le u<p$, $\delta=\mathbf{1}_{u+d\ge p}$. Only $s^2(su_r-V)l_{e-f}$ and the terms with $i\le a$ in the convolution can reach pole order $j+5$. Their total coefficient is

$$(-1)^{j+1}\left(\mathcal{C}_{j-1}(a+1+\delta)+\sum_{\substack{0\le i\le a\\ i\equiv\beta\ (2)}}(-1)^i\mathcal{C}_j(a-i)\right). \tag{4.26}$$

The parity restriction comes from $A_{\beta-i+1}(-1)=1$ exactly when $i\equiv\beta \pmod 2$. For $\delta=1$ the terms in parentheses all have sign $(-1)^a$. The first is nonzero for $j\ge 2$; if $j=1$ and it vanishes, then a is odd and at least 1, and one of $i=0,1$ supplies a positive term after removing this sign. Hence $Q\neq 0$.

It remains to consider $h<f$ and $\delta=0$. First assume $r\ge 3p$. The following lowest terms at $x=0$ follow from a formal power-series identity:

$$l_n = \frac{1}{1+x}\sum_{\substack{a,b\ge 0\\ ap+br>n}}\binom{a+b}{a}x^a z_0^b. \tag{4.27}$$

Indeed the unrestricted sum is $(1-x-z_0)^{-1}=1+x$, which transforms (4.13) into (4.27). Since $z_0=-x^2+O(x^3)$, a pair (a,b) has smallest degree $a+2b$. For $n=jr+h$, the minimal pairs and resulting terms are

$$l_{jr+h}\sim\begin{cases}(-1)^j(j+1)x^{2j+1}, & 0\le h<p,\\ (-1)^j\dfrac{j(j+3)}{2}x^{2j+2}, & p\le h<2p,\ j\ge 1,\\ (-1)^{j+1}x^{2j+2}, & 2p\le h<r.\end{cases} \tag{4.28}$$

The pairs are respectively $(1,j)$; $(2,j),(0,j+1)$; and $(0,j+1)$. For $b=j-t$, $t\ge 1$, a path of degree at most $2j+2$ reaches at most $jr+(2-t)p$, which cannot exceed a target in the middle or final interval. The analogous bound at degree $2j+1$ is $jr+(1-t)p$. Thus there are

no omitted lowest-degree paths, even at $r = 3p$. For $j = 0$ the middle interval is identically zero and is treated as such.

The early values give

$$u_r = -2x^3 + O(x^4), \quad V = -x^3 + O(x^4), \quad su_r - V = x^3 + O(x^4), \quad sV = -x^4 + O(x^5).$$

Let $\kappa_j = j(j+3)/2$ and $I = \mathbf{1}_{\beta \geq 2}$. Applying (4.28) to (4.20) and then (4.19), the first nonzero coefficient of Q is

$a = \lfloor h/p \rfloor$	Degree	Coefficient
0	$2j+5$	$(-1)^j(j+1)$
1	$2j+6$	$(-1)^j(\kappa_j - 1 + I)$
$a \geq 2$	$2j+6$	$(-1)^{j+1}$

Here are the convolution cancellations giving the table. For $a = 0$, u_{k-r} begins with $(-1)^j(j+1)x^{2j+2}$; the coefficient of x^{2j+3} in u_{k-q} is $(-1)^j(-(j+2)+(j+1)+1) = 0$, from its linear term and $i = 0, 1$. The middle term of Q has larger degree. For $a = 1$, the leading coefficients of u_{k-r} and u_{e+p} are $(-1)^j(\kappa_j - 1)$ at degree $2j+3$ and $(-1)^{j+1}$ at degree $2j+2$. The coefficient of x^{2j+4} in u_{k-q} is

$$(-1)^j \left(-\frac{(j+1)(j+4)}{2} + \kappa_j + (j+1) + I \right) = (-1)^j(I-1).$$

Its four terms come from the linear part and $i = 0, 1, 2$, with $i = 2$ present exactly when $\beta \geq 2$. For $a \geq 2$, u_{k-r} has coefficient $2(-1)^{j+1}$ at degree $2j+3$, u_{e+p} has coefficient $(-1)^{j+1}$ at degree $2j+2$, and the linear term and $i = 0$ term in u_{k-q} cancel at degree $2j+4$. To bound all remaining indices, when $i > a$ the residual position after crossing one r block is at least

$$r + (a-i)p + u \geq (a+1)p + b + d + u > (a+1)p.$$

Including the factor x^i puts every unlisted term strictly above the orders just used. The indices $i \leq a$ follow directly from (4.28). Every coefficient in the table is nonzero for $j \geq 1$.

Finally take $r < 3p$. Write $r = 2p + \rho$, where $d < \rho < p$; then $\beta = 1$, $V = 0$, $u_r = -xs^2$, and $h < f$ forces $a = 0$ or 1. For $\delta = 0$, the parentheses in (4.26) are $-(j-1)$ when $a = 0$, and $j(j-1)/2$ when $a = 1$. This proves nonvanishing for $j \geq 2$. For $j = 1$, (4.20) simplifies (4.19) to

$$Q = xs^2 J, \qquad J = l_{e+r-p} - sl_{e+q} + s^2 l_{e-p} - s^3 l_{e-f}.$$

Put $\alpha = \mathbf{1}_{\rho+u \geq p}$. When $a = 0$, $l_{e-f} = l_{p+u+d} = 0$. The first term of J has pole order at most 2; the next two contribute -1 and $-\alpha$ to the pole of order 3. Thus $[\epsilon^{-5}]Q = 1 + \alpha \neq 0$. When $a = 1$, $h < f$ gives $u < \rho - d$. The first two terms of J contribute -1 and 2 at order 3. Also $l_{e-f} = z_0$ and $l_{e-p} = l_{r+u} = z_0(s + \alpha x)$, whence

$$s^2 l_{e-p} - s^3 l_{e-f} = \alpha x z_0 s^2 = -\alpha x^2 s^3.$$

Its order-3 coefficient is α. Consequently $[\epsilon^{-5}]Q = -(1+\alpha) \neq 0$. This completes $e \geq r$ and the last parameter region.

The alternatives treated above exhaust the complement of (4.2): first $r < 2p$; then $r \geq 2p$ with $k < p + q + r$; and, for $k > p + q + r$, the regions $r < p + q$, $r > p + q$ with $q \geq 2p$, and $r > p + q$ with $q < 2p$. In each case (4.4) has a nonpolynomial term by index $2k + p$. The discussion following (4.4) proves the denominator assertion and finishes Theorem 4.1.

References

[1] M. Bousquet-Mélou, J. Propp, and J. West, *Perfect matchings for the three-term Gale–Robinson sequences*, Electron. J. Combin. **16** (2009), no. 1, R125. doi:10.37236/214.

[2] G. D. Carroll and D. E. Speyer, *The cube recurrence*, Electron. J. Combin. **11** (2004), no. 1, R73. arXiv:math/0403417.

[3] G. Everest, G. McLaren, and T. Ward, *Primitive divisors of elliptic divisibility sequences*, J. Number Theory **118** (2006), 71–89. doi:10.1016/j.jnt.2005.08.002.

[4] S. Fomin and A. Zelevinsky, *The Laurent phenomenon*, Adv. Appl. Math. **28** (2002), 119–144. doi:10.1006/aama.2001.0770.

[5] D. Gale, *The strange and surprising saga of the Somos sequences*, Math. Intelligencer **13** (1991), no. 1, 40–42. doi:10.1007/BF03024070.

[6] D. Gale, *Somos sequence update*, Math. Intelligencer **13** (1991), no. 4, 49–50. doi:10.1007/BF03028343.

[7] A. N. W. Hone, *Elliptic curves and quadratic recurrence sequences*, Bull. London Math. Soc. **37** (2005), 161–171. doi:10.1112/S0024609304004163; corrigendum, **38** (2006), no. 5, 741–742. doi:10.1112/S0024609306018844.

[8] A. N. W. Hone, T. E. Kouloukas, and C. Ward, *On reductions of the Hirota–Miwa equation*, SIGMA **13** (2017), 057, 17 pp. doi:10.3842/SIGMA.2017.057.

[9] A. N. W. Hone and C. Swart, *Integrality and the Laurent phenomenon for Somos 4 and Somos 5 sequences*, Math. Proc. Cambridge Philos. Soc. **145** (2008), 65–85. doi:10.1017/S030500410800114X.

[10] R. Jones and J. Rouse, *Galois theory of iterated endomorphisms*, Proc. London Math. Soc. (3) **100** (2010), 763–794. doi:10.1112/plms/pdp051.

[11] J. L. Malouf, *An integer sequence from a rational recursion*, Discrete Math. **110** (1992), 257–261. doi:10.1016/0012-365X(92)90714-Q.

[12] OEIS Foundation, *A030127: Term at which n-Somos sequence first becomes nonintegral*, comments by P. D. Hanna (2004) and M. Somos (2014), `https://oeis.org/A030127`, accessed 18 November 2026.

[13] OEIS Foundation, *A101591: Indices $k \geq 8$ designating the k-Somos sequences that first become nonintegral at odd-indexed terms*, P. D. Hanna (2004), `https://oeis.org/A101591`, accessed 18 November 2026.

[14] J. Propp, *A bare-bones chronology of Somos sequences*, July 2006, `https://faculty.uml.edu/jpropp/somos/chronology.html`.

[15] R. M. Robinson, *Periodicity of Somos sequences*, Proc. Amer. Math. Soc. **116** (1992), 613–619. doi:10.1090/S0002-9939-1992-1140672-5.

[16] J. H. Silverman, *The Arithmetic of Elliptic Curves*, second ed., Graduate Texts in Mathematics, vol. 106, Springer, Dordrecht, 2009.

[17] M. Somos, *Problem 1470*, Crux Math. **15** (1989), 208.

[18] M. Somos, *Brief history of the Somos sequence problem*, 7 May 1990, updated 30 October 1994, `https://faculty.uml.edu/jpropp/somos/history.txt`.

[19] M. Somos, *The refined Gale–Robinson polynomial sequence conjecture*, 11 April 1999, updated 7 September 2022, `https://grail.eecs.csuohio.edu/~somos/robinson.html`.

[20] D. E. Speyer, *Perfect matchings and the octahedron recurrence*, J. Algebraic Combin. **25** (2007), 309–348. arXiv:math/0402452.

[21] R. P. Stanley, *Bijective proof problems*, 18 August 2009, `https://math.mit.edu/~rstan/bij.pdf`.

[22] C. S. Swart, *Elliptic curves and related sequences*, Ph.D. thesis, Royal Holloway, University of London, 2003.

[23] Y. Uchida, *Valuations of Somos 4 sequences and canonical local heights on elliptic curves*, Math. Proc. Cambridge Philos. Soc. **150** (2011), 385–397. doi:10.1017/S030500411000068X.

[24] P. H. van der Kamp, *Somos-4 and Somos-5 are arithmetic divisibility sequences*, J. Difference Equations Appl. **22** (2016), no. 4, 570–580, corrected version. doi:10.1080/10236198.2015.1113272; erratum doi:10.1080/10236198.2016.1193956.

[25] M. Ward, *The law of repetition of primes in an elliptic divisibility sequence*, Duke Math. J. **15** (1948), 941–946.

[26] M. Ward, *Memoir on elliptic divisibility sequences*, Amer. J. Math. **70** (1948), 31–74.