

Parity Collapse of Run-Refined Even-Zero-Run Words to a Single Binomial Coefficient

Augusto Rehfeldt

Text generated by language models in the mathforge pipeline; not reviewed by a person.

Abstract

We study binary words in which every run of zeros has even length, refined by length and number of runs. Let $R(n, r)$ denote the number of words of length n with exactly r runs and all zero-runs of even length. We prove that for every $n \geq 1$ and $m \geq 0$, $R(n, 2m)$ is even and $R(n, 2m + 1) \equiv \binom{n}{3m+1} \pmod{2}$. The proof uses run-length encoding, stars and bars, and reduction of the bivariate run generating functions modulo 2 via the identity $1 - t^2 = (1 - t)^2$ over $\mathbf{F}_2$. The result was verified by exhaustive enumeration for $1 \leq n \leq 20$ and by an independent check of every lemma. A bounded automated search over arXiv, Crossref, and OpenAlex returned no matching prior statement. No Lean formalization was attempted.

1. Introduction

1.1 Context

Let a binary word be an element of $\{0, 1\}^n$. A run is a maximal interval of consecutive equal letters. We consider the restricted language consisting of words in which every zero-run has even length. No restriction is placed on one-runs.

The unrefined enumeration of this language is classical: the total number of such words of length n satisfies a Fibonacci recurrence. Run-refined enumeration by the Goulden-Jackson cluster method and by transfer-matrix methods is standard, as are divisibility properties of Fibonacci polynomials. Congruences modulo 2 in combinatorics are commonly approached through generating functions over $\mathbf{F}_2$ and through Lucas' theorem on the parity of binomial coefficients.

1.2 What is new

The present result is a precise parity statement for the run-refined counts $R(n, r)$. It asserts vanishing modulo 2 for all even run-counts, and collapse to a single binomial coefficient $\binom{n}{3m+1}$ modulo 2 for odd run-counts $r = 2m + 1$, via a characteristic-two identity applied to the even-zero-run composition constraint. To the extent established by the bounded automated search described in Section 5, this exact formulation has not appeared as a named theorem, exercise, or paper lemma, and is best regarded as a small apparently new observation rather than a famous result.

2. Notation

A binary word of length n is an element $w = (w_1, \dots, w_n)$ in $\{0, 1\}^n$. A run of w is a maximal interval of consecutive equal letters; $r(w)$ is the number of runs, i.e. 1 plus the number of indices i in $\{2, \dots, n\}$ with $w_i \neq w_{i-1}$ (for $n \geq 1$), and $r = 0$ only for the empty word. A zero-run is a run with letter 0. For integers $n \geq 0$ and $r \geq 0$, $R(n, r)$ denotes the number of words w in $\{0, 1\}^n$ with $r(w) = r$ such that every zero-run of w has even length. For integers $a \geq 0$ and $b \geq 0$, $C(a, b)$ denotes the ordinary binomial coefficient, with the convention $C(a, b) = 0$ if $b > a$.

3. Results

Theorem. *For every integer $n \geq 1$ and every integer $m \geq 0$, (i) $R(n, 2m)$ is even, i.e. $R(n, 2m)$ is congruent to 0 modulo 2; and (ii) $R(n, 2m + 1)$ is congruent to $C(n, 3m + 1)$ modulo 2, i.e. the integer $R(n, 2m + 1) - C(n, 3m + 1)$ is divisible by 2.*

Proof. The following proof is reproduced verbatim from the verified proof. □

Parity of words with even zero-runs

Notation and theorem

Let $n \geq 1$. A binary word of length n is $w = (w_1, \dots, w_n) \in \{0, 1\}^n$. For $n \geq 1$, let $r(w)$ be 1 plus the number of indices $i \in \{2, \dots, n\}$ with $w_i \neq w_{i-1}$. A run is a maximal interval of consecutive equal letters, so $r(w)$ is the number of runs. For $n \geq 0$ and $r \geq 0$, let $R(n, r)$ be the number of $w \in \{0, 1\}^n$ with $r(w) = r$ such that every zero-run has even length. For $r = 0$ and $n \geq 1$, there is no such word, so $R(n, 0) = 0$.

For integers $a \geq 0$ and $b \geq 0$, let $C(a, b)$ be the ordinary binomial coefficient, with $C(a, b) = 0$ if $b > a$.

Theorem. *For every integer $n \geq 1$ and every integer $m \geq 0$, (i) $R(n, 2m)$ is even, and (ii) $R(n, 2m + 1) - C(n, 3m + 1)$ is divisible by 2.*

The proof proceeds by run-length encoding, counting compositions with restricted parities, and reduction of generating functions modulo 2.

Lemmas

Lemma 1 (Run-length encoding). *Fix $n \geq 1$, $r \geq 1$, and $s \in \{0, 1\}$. Let $W(n, r, s)$ be the set of $w \in \{0, 1\}^n$ with $r(w) = r$ and $w_1 = s$. Then $W(n, r, s)$ is in bijection with the set of compositions $c_1 + \dots + c_r = n$ with $c_i \geq 1$ for all i .*

Proof. Given $w \in W(n, r, s)$, let c_i be the length of its i -th run. Then each $c_i \geq 1$ and $\sum c_i = n$, because the runs partition the n positions. This defines a map from words to compositions. Conversely, given a composition $c_1 + \dots + c_r = n$, define w by filling the first c_1 positions with letter s , the next c_2 positions with letter $1 - s$, and so on, alternating the letter at each new block. Since $r \geq 1$ and each $c_i \geq 1$, the resulting word has length n , starts with s , and changes letter exactly at the $r - 1$ block boundaries, so it has exactly r runs. The two constructions are mutually inverse: run lengths determine the word with given starting letter, and the word determines its run lengths. Hence the map is a bijection. □

Lemma 2 (Parametrisation under the even zero-run condition). *Fix $n \geq 1$, $r \geq 1$, $s \in \{0, 1\}$. A word $w \in W(n, r, s)$ has all zero-runs of even length if and only if, in its run-length composition $c_1 + \dots + c_r = n$, every c_i whose block has letter 0 is even.*

Proof. The i -th block of w is a zero-run exactly when its letter is 0, which is s for odd i if $s = 0$ and for even i if $s = 1$, and $1 - s$ otherwise. Its length is c_i . The condition that every zero-run has even length is therefore exactly the condition that c_i is even for every i whose block letter is 0. No condition is imposed on blocks with letter 1. □

For integers $a, b \geq 0$ and $n \geq 0$, define $N_{a,b}(n)$ as the number of tuples $(e_1, \dots, e_a, d_1, \dots, d_b)$ with $e_i \geq 1$ and $d_j \geq 1$ satisfying

$$\sum_{i=1}^a e_i + 2 \sum_{j=1}^b d_j = n,$$

with the convention that an empty sum equals 0, so $N_{0,0}(0) = 1$ and $N_{0,0}(n) = 0$ for $n > 0$.

Lemma 3 (Stars and bars). *Let $k \geq 1$ and $K \geq 0$. The number of k -tuples of integers ≥ 1 with sum K equals $C(K-1, k-1)$ (which is 0 if $K < k$).*

Proof. If $K < k$, no such tuple exists since the sum of k positive integers is at least k , and $C(K-1, k-1) = 0$ under the convention in force when $K = 0$, while for $1 \leq K < k$ the upper entry $K-1$ is smaller than $k-1$, so the binomial coefficient is 0. Suppose $K \geq k \geq 1$. Given $(e_1, \dots, e_k)$ with $e_i \geq 1$ and sum K , define partial sums $s_j = e_1 + \dots + e_j$ for $j = 1, \dots, k-1$. Then $1 \leq s_1 < \dots < s_{k-1} \leq K-1$, so $\{s_1, \dots, s_{k-1}\}$ is a $(k-1)$ -element subset of $\{1, \dots, K-1\}$. Conversely, given such a subset written in increasing order $s_1 < \dots < s_{k-1}$, define $e_1 = s_1$, $e_j = s_j - s_{j-1}$ for $2 \leq j \leq k-1$, and $e_k = K - s_{k-1}$. Then each $e_j \geq 1$ and the sum is K . The two maps are inverse to each other, so the number of tuples equals the number of subsets, namely $C(K-1, k-1)$. □

Lemma 4 (Generating functions over $\mathbb{Z}$). *Fix $a, b \geq 0$. In the ring $\mathbb{Z}[[t]]$ of formal power series with integer coefficients,*

$$\sum_{n \geq 0} N_{a,b}(n) t^n = \left(\frac{t}{1-t} \right)^a \left(\frac{t^2}{1-t^2} \right)^b,$$

where for $a = 0$ or $b = 0$ the corresponding factor is 1, and $t/(1-t) = \sum_{e \geq 1} t^e$, $t^2/(1-t^2) = \sum_{d \geq 1} t^{2d}$.

Proof. In $\mathbb{Z}[[t]]$, the series $1-t$ has constant term 1 and hence has a multiplicative inverse, namely $\sum_{j \geq 0} t^j$. Thus $t/(1-t) = \sum_{e \geq 1} t^e$, and similarly $t^2/(1-t^2) = \sum_{d \geq 1} t^{2d}$, with convergence in the (t) -adic topology. The product of finitely many such series has as its coefficient of t^n the number of ways to write n as a sum of the exponents of one chosen monomial from each factor, because for each fixed n only finitely many exponent combinations contribute. With a factors of the first type and b factors of the second type, a choice consists of $e_1, \dots, e_a \geq 1$ and $d_1, \dots, d_b \geq 1$ with $\sum e_i + 2 \sum d_j = n$, and the number of such choices is precisely $N_{a,b}(n)$. When $a = 0$ or $b = 0$, the empty product is 1, which matches $N_{0,0}$. □

Lemma 5 (Collapse modulo 2). *Let $\bar{N}_{a,b}(n) \in \{0, 1\}$ denote $N_{a,b}(n)$ modulo 2. In $\mathbf{F}_2[[t]]$,*

$$\sum_{n \geq 0} \bar{N}_{a,b}(n) t^n = \left(\frac{t}{1-t} \right)^{a+2b}.$$

Proof. Coefficient-wise reduction modulo 2 defines a ring homomorphism $\phi : \mathbb{Z}[[t]] \rightarrow \mathbf{F}_2[[t]]$, since addition and multiplication of formal power series involve only finite sums of coefficients in each degree. By Lemma 4, $\sum N_{a,b}(n) t^n = (t/(1-t))^a (t^2/(1-t^2))^b$ in $\mathbb{Z}[[t]]$. Applying ϕ gives $\sum \bar{N}_{a,b}(n) t^n = (t/(1-t))^a (t^2/(1-t^2))^b$ in $\mathbf{F}_2[[t]]$, where inverses are preserved because ϕ maps a series with constant term 1 to a series with constant term 1, and the image of an inverse satisfies the defining relation of the inverse, which is unique.

In $\mathbf{F}_2[[t]]$, subtraction coincides with addition, so $1-t = 1+t$ and $1-t^2 = 1+t^2$. Moreover $(1-t)^2 = 1-2t+t^2 = 1+t^2$ since $2=0$ in $\mathbf{F}_2$. Hence $1-t^2 = (1-t)^2$ in $\mathbf{F}_2[[t]]$. Since $1-t$ has constant term 1, it is invertible, and taking inverses gives $(1-t^2)^{-1} = (1-t)^{-2}$. Therefore

$$(t^2/(1-t^2))^b = t^{2b} (1-t)^{-2b}.$$

Multiplying by $(t/(1-t))^a = t^a (1-t)^{-a}$ yields $t^{a+2b} (1-t)^{-(a+2b)}$, which is $(t/(1-t))^{a+2b}$.

□

Lemma 6 (Coefficients of $(t/(1-t))^k$ modulo 2). *Fix $k \geq 1$. For every $n \geq 0$, the coefficient of t^n in $(t/(1-t))^k$ viewed in $\mathbf{F}_2[[t]]$ equals $C(n-1, k-1)$ modulo 2 (with value 0 when $n = 0$).*

Proof. Over $\mathbb{Z}$, $(t/(1-t))^k$ is the k -fold product of $\sum_{e \geq 1} t^e$, so by the same convolution argument as in Lemma 4, its coefficient of t^n is the number of k -tuples of positive integers with sum n , which by Lemma 3 equals $C(n-1, k-1)$ (and 0 for $n < k$, matching the binomial convention). Reducing coefficients modulo 2 gives the power series $(t/(1-t))^k$ in $\mathbf{F}_2[[t]]$ with coefficients $C(n-1, k-1)$ modulo 2.

□

Lemma 7 (Pascal's identity). *For all $N \geq 0$ and $K \geq 0$,*

$$C(N, K) + C(N, K+1) = C(N+1, K+1).$$

Proof. If $K > N$, then $C(N, K) = C(N, K+1) = 0$ and $C(N+1, K+1) = 0$ since $K+1 > N+1$, so both sides are 0. If $K = N$, the left side is $1 + 0 = 1$ and the right side is $C(N+1, N+1) = 1$. Suppose $0 \leq K < N$. If $K = 0$, the identity reads $1 + N = N+1$. If $K \geq 1$, the factorial formula gives

$$\begin{aligned} \frac{N!}{K!(N-K)!} + \frac{N!}{(K+1)!(N-K-1)!} &= \frac{N!}{K!(N-K-1)!} \left(\frac{1}{N-K} + \frac{1}{K+1} \right) \\ &= \frac{N!}{K!(N-K-1)!} \cdot \frac{N+1}{(N-K)(K+1)} = \frac{(N+1)!}{(K+1)!(N-K)!}, \end{aligned}$$

which is $C(N+1, K+1)$.

□

Proof of the theorem. Fix $n \geq 1$. For $r \geq 1$ and $s \in \{0, 1\}$, Lemmas 1 and 2 show that words in $W(n, r, s)$ with all zero-runs even correspond to compositions $c_1 + \dots + c_r = n$ in which each block with letter 0 has even length. Writing each odd-constrained block length as $e_i \geq 1$ and each even-constrained block length as $2d_j$ with $d_j \geq 1$, the count of such words equals $N_{a,b}(n)$, where a is the number of one-blocks and b the number of zero-blocks among the r runs with starting letter s .

If $r = 2m$ with $m \geq 1$, then for either starting letter there are exactly m one-blocks and m zero-blocks. Hence each starting letter contributes $N_{m,m}(n)$ words, and

$$R(n, 2m) = 2N_{m,m}(n).$$

In particular $R(n, 2m)$ is even. For $m = 0$, $r = 0$, and $R(n, 0) = 0$ since every word of positive length has at least one run. This proves (i).

Suppose $r = 2m+1$ with $m \geq 0$. If $s = 1$, the odd-positioned blocks are ones, so $a = m+1$ and $b = m$. If $s = 0$, the odd-positioned blocks are zeros, so $a = m$ and $b = m+1$. Therefore

$$R(n, 2m+1) = N_{m+1,m}(n) + N_{m,m+1}(n).$$

By Lemma 5, the generating functions of the reductions modulo 2 are $(t/(1-t))^{a+2b}$ in $\mathbf{F}_2[[t]]$. Here $(m+1) + 2m = 3m+1 \geq 1$ and $m + 2(m+1) = 3m+2 \geq 1$. By Lemma 6, equating coefficients gives, modulo 2,

$$N_{m+1,m}(n) \equiv C(n-1, 3m) \quad \text{and} \quad N_{m,m+1}(n) \equiv C(n-1, 3m+1).$$

Indeed, for $m = 0$ this remains valid: $N_{1,0}(n) = 1$ for $n \geq 1$ and $C(n-1, 0) = 1$; $N_{0,1}(n)$ is 1 if n is even and positive and 0 otherwise, which agrees with $C(n-1, 1) = n-1$ modulo 2.

Adding the two contributions modulo 2,

$$R(n, 2m+1) \equiv C(n-1, 3m) + C(n-1, 3m+1) \pmod{2}.$$

By Lemma 7 with $N = n - 1$ and $K = 3m$, $C(n - 1, 3m) + C(n - 1, 3m + 1) = C(n, 3m + 1)$ as integers, hence also modulo 2. Therefore

$$R(n, 2m + 1) \equiv C(n, 3m + 1) \pmod{2},$$

which is (ii). This holds for every $n \geq 1$ and $m \geq 0$. □

4. Computational verification

4.1 Exhaustive search

For each length n with $1 \leq n \leq 20$ and each r with $0 \leq r \leq n$, all 2^n binary words w of length n were enumerated. For each word, every maximal block of consecutive zeros was tested for even length, the number $r(w)$ of maximal constant blocks was counted, and the exact count $R(n, r)$ was formed. It was then checked that $R(n, 2m) \bmod 2 = 0$ and $R(n, 2m + 1) \bmod 2 = \binom{n}{3m+1} \bmod 2$ for all m with $2m \leq n$ and $2m + 1 \leq n$. The total search covered $\sum_{n \leq 20} 2^n$ words (about two million words, 2,097,150 masks, 230 relations), each checked in linear time.

The run reported exit code 0 with no counterexample. Sanity checks confirmed hand values $R(1 \dots 4)$ as $\{1 : [0, 1], 2 : [0, 2, 0], 3 : [0, 1, 2, 0], 4 : [0, 2, 2, 1, 0]\}$ and agreement between a fast implementation and a naive implementation on all words for $n = 1 \dots 10$ (3,069 words checked).

4.2 Independent check

A separate independent check with exit code 0 confirmed Lemma 1 (run-length encoding bijection), Lemma 2 (even zero-run parametrisation), Lemma 3 (stars and bars), Lemma 4 (generating functions over $\mathbb{Z}$), Lemma 5 (collapse modulo 2), Lemma 6 (coefficients of $(t/(1-t))^k \bmod 2$), Lemma 7 (Pascal identity), the theorem counting decomposition, and parts (i) and (ii) of the theorem. All checks passed with no repairs.

4.3 Lean 4 status

Lean 4 formalization was not attempted for this theorem. The status is therefore: not attempted. There is no Lean statement to compare, so no back-translation faithfulness verdict applies. In particular, no sorry-free Lean proof is claimed, and the correctness claim rests on the informal proof and the computational checks above. A sorry-free Lean proof, if produced in future, would certify only its own formal statement, which would still have to be read carefully against the informal statement.

5. Limitations and open questions

No **##** GAP markers remain from the verification record; there are no hidden gaps to carry forward. The referee report returned verdict VALID with no issues.

The result is limited to parity (modulo 2) and to refinement by number of runs. It does not address refinement by weight, higher moduli, exact formulas over $\mathbb{Z}$, asymptotics, or words with additional restrictions. Natural open questions are exact expressions for $R(n, r)$ over the integers, analogues modulo odd primes, and joint refinements by length, weight, and number of runs.

5.1 Novelty

The novelty verdict recorded is APPARENTLY_NEW, with closest known results listed as Lucas' theorem on the parity of binomial coefficients, exact and asymptotic enumeration of restricted integer compositions, and the fixed-point-free involution method for proving congruences modulo 2. The matching-hits list is empty.

The exact search queries run were:

Queries: 1. “parity of binary words with all zero-runs of even length enumerated by number of runs modulo 2 binomial coefficient” 2. “enumeration of binary compositions with restricted run lengths and restricted part sizes modulo 2 parity results” 3. “fixed-point-free involution technique for congruences modulo 2 in combinatorial enumeration with Lucas theorem on binomial coefficient parity”

Search terms: 1. “binary words even zero-runs runs enumeration” 2. “restricted compositions parity modulo 2” 3. “Lucas theorem binomial parity runs”

Follow-up queries: 1. “binary words all zero-runs even length number of runs parity” 2. “compositions with even parts on alternating positions modulo 2 binomial coefficient” 3. “involution proof $R(n,r)$ congruent binomial coefficient mod 2”

These were run in 1 round against arXiv, Crossref, and OpenAlex. A total of 39 records were retrieved (15 from arXiv, 14 from Crossref, 10 from OpenAlex), with no matching prior statement. The evidence base is retrieval, not memory-only.

This is a bounded automated search over those indexes only. It is not a literature review, and it is blind to books, journals outside those indexes, and to anything phrased differently from the queries above. Novelty therefore means apparently new relative to that bounded search plus model recall, not certified novelty against all literature.