

Bounds for binary words in divisibility patterns

Sulayman Bowles

19 October 2026

Abstract

For a binary word w , let $N(w)$ be the least positive integer whose divisor indicator contains w . In response to a question of Ilkka Törmä, we prove $\log N(01^{m-1}) \sim 2m$, and obtain the general upper bound $\log \max_{|w|=m} N(w) \leq (1 + o(1))m(\log m)^2 / \log \log m$. These give improved bounds relative to those displayed in the question; the growth rates of the maximum and arithmetic mean remain undetermined. Historical priority is not asserted.

For $n \geq 1$, write $d_n(j) = 1$ if $j \mid n$, and $d_n(j) = 0$ otherwise. A word $w = w(0) \cdots w(m-1)$ occurs in d_n at $k \geq 1$ if $d_n(k+i) = w(i)$ for every $0 \leq i < m$. Put

$$N(w) = \min\{n \geq 1 : w \text{ occurs in } d_n\}, \quad M(m) = \max_{w \in \{0,1\}^m} N(w),$$

and let $A(m) = 2^{-m} \sum_{w \in \{0,1\}^m} N(w)$ be the arithmetic mean. The construction below shows that these quantities exist. All logarithms are natural.

Törmä's question [1] gives $\log M(m) \leq (1 + o(1))m^2 \log m$ and the lower bound $N(01^{m-1}) \geq \frac{1}{2} \binom{2m}{m}$. We sharpen both displayed bounds, using standard prime-number estimates and Rankin's method. The least-common-multiple estimate appearing in the original question is credited there to Farhi [2]; we give the elementary special case needed here.

1 The word 01^{m-1}

Write

$$\vartheta(x) = \sum_{p \leq x} \log p, \quad \psi(x) = \log \operatorname{lcm}(1, \dots, \lfloor x \rfloor).$$

We use the prime number theorem in the forms $\vartheta(x) = x + o(x)$ and $\psi(x) = x + o(x)$, together with its consequence that there is a prime $p \geq m$ satisfying $p = (1 + o(1))m$.

Lemma 1. *For positive integers k, r , let $D = \operatorname{lcm}(k+1, \dots, k+r)$. Then*

$$\operatorname{lcm}(1, \dots, r) \mid D, \quad \binom{k+r}{r} \mid D.$$

In particular, $D \geq k^r / r!$.

Proof. Every block of r consecutive integers contains a multiple of every integer at most r , which proves the first assertion. For the second, fix a prime p . For each $j \geq 1$, the number of multiples of p^j in the block differs from $\lfloor r/p^j \rfloor$ by either zero or one. Thus the p -adic valuation of $(k+1) \cdots (k+r)/r!$ is a sum of zeros and ones. A summand can be nonzero only if p^j divides some member of the block. The sum is therefore at most the largest p -adic valuation of a block member, which is the valuation of D . $\square$

Theorem 2. As $m \rightarrow \infty$,

$$\log N(01^{m-1}) = (2 + o(1))m.$$

Proof. Suppose 01^{m-1} occurs in d_n at k . For $m \geq 2$ we must have $k \geq m$: otherwise $2k$ is one of the following $m-1$ positions and divides n , contradicting $k \nmid n$. Set $r = m-1$ and $D = \text{lcm}(k+1, \dots, k+r)$, so that $D \mid n$.

First suppose $m \leq k \leq 3m$. By Lemma 1, D is divisible by $\text{lcm}(1, \dots, r)$. It is also divisible by every prime in $(k, k+r]$. These primes exceed r , so their product is coprime to that lcm. Consequently

$$\log D \geq \psi(r) + \vartheta(k+r) - \vartheta(k) = 2m + o(m).$$

The error is uniform over these k : the prime number theorem gives a uniform $o(m)$ error for arguments in $[m, 4m]$. More explicitly, if $\vartheta(x) = x + x\eta(x)$ with $\eta(x) \rightarrow 0$, then the two endpoint errors are bounded by $7m \sup_{x \geq m} |\eta(x)| = o(m)$.

If $k > 3m$, Lemma 1 and Stirling's formula instead give

$$\log D \geq r \log k - \log r! \geq (1 + \log 3)m - O(\log m) > 2m - o(m).$$

Both cases establish $\log n \geq (2 - o(1))m$ for every occurrence. No monotonicity of consecutive-interval lcms is used.

For the reverse inequality, choose a prime $p \geq m$ with $p = (1 + o(1))m$, and put

$$n = \text{lcm}(p+1, \dots, p+m-1).$$

All entries in this block lie strictly between p and $2p$, so none is divisible by p . Since p is prime, $p \nmid n$, and 01^{m-1} occurs at p . Finally,

$$\log n \leq \psi(p+m-1) = (2 + o(1))m.$$

□

Corollary 3. The original maximum and arithmetic mean satisfy

$$\log M(m) \geq (2 - o(1))m, \quad \log A(m) \geq (2 - \log 2 - o(1))m.$$

Proof. Use $M(m) \geq N(01^{m-1})$ and $A(m) \geq 2^{-m}N(01^{m-1})$. □

2 A universal upper bound

An integer is y -smooth if all its prime factors are at most y ; 1 is included. Let $\Psi(X, y)$ count these integers up to X .

Lemma 4. For every $\epsilon > 0$ and all sufficiently large m , there is a length- m interval contained in $[1, X]$ with no m -smooth integer, where

$$X = \left\lfloor \exp \left((1 + \epsilon) \frac{(\log m)^2}{\log \log m} \right) \right\rfloor.$$

Proof. Put $L = \log m$, $B = \log L$, $\delta = B/L$, and $s = 1 - \delta$. We claim

$$\log \prod_{p \leq m} (1 - p^{-s})^{-1} = o(L).$$

For sufficiently large m , $s \geq 3/4$. In the logarithmic expansion, the terms of degree at least two have bounded total, uniformly for these s . It remains to bound $\sum_{p \leq m} p^{-s}$. Chebyshev's estimate $\pi(t) = O(t/\log t)$ and partial summation give

$$\sum_{p \leq m} p^{-s} = O\left(1 + \int_{\log 2}^L \frac{e^{\delta u}}{u} du\right) = O(B + L/B) = o(L).$$

For the last estimate, split the integral at $1/\delta$. The first piece is $O(\log(1/\delta)) = O(B)$; in the second use $v = \delta u$ and $\int_1^B e^v dv/v = O(e^B/B) = O(L/B)$.

Rankin's inequality [3, §3.10] is

$$\Psi(X, m) \leq X^s \prod_{p \leq m} (1 - p^{-s})^{-1}.$$

Indeed, for every m -smooth $a \leq X$, the indicator of $a \leq X$ is at most $(X/a)^s$, and summing over all smooth a produces the Euler product. Hence

$$\frac{\Psi(X, m)}{X} \leq \exp(-\delta \log X + o(L)) = m^{-1-\epsilon+o(1)}.$$

There are $X - m + 1$ length- m intervals in $[1, X]$. A smooth integer lies in at most m of them, so at most $m\Psi(X, m) = o(X)$ intervals meet a smooth integer. Since $m = o(X)$, an interval remains. $\square$

Theorem 5. As $m \rightarrow \infty$,

$$\log M(m) \leq (1 + o(1)) \frac{m(\log m)^2}{\log \log m}.$$

The same upper bound holds with $A(m)$ in place of $M(m)$.

Proof. Choose $k, \dots, k + m - 1$ as in Lemma 4. For every i there is a prime $p_i > m$ dividing $k + i$. This prime divides no other block member, because their nonzero differences have magnitude less than m . For any word w , put

$$n_w = \prod_{i:w(i)=1} (k + i),$$

where the empty product is 1. The one-positions divide n_w . At a zero-position i , the prime p_i divides no selected factor, and therefore $k + i \nmid n_w$. Thus w occurs at k and $N(w) \leq n_w \leq X^m$. Lemma 4, with arbitrarily small fixed ϵ , gives the result. Since $A(m) \leq M(m)$, the mean obeys it as well. $\square$

3 Exact finite checks and remaining questions

The accompanying program computes $H(m) = N(01^{m-1})$ exactly for $1 \leq m \leq 250$. Given a starting point k , the smallest possible witness is $\text{lcm}(k + 1, \dots, k + m - 1)$, provided that k does not divide this lcm. Every other witness at that start is its multiple. A prime start supplies an initial finite upper bound. Lemma 1 then bounds the search: once $k^{m-1} > H(m)(m - 1)!$, no later start can improve the best witness. The stopping test uses exact integers.

m	1	2	3	4	5	6
$H(m)$	1	3	20	168	504	3960

These are family values, not the general maximum; for example, the original question reports $M(5) = 660$, whereas $H(5) = 504$. The program also checks 18,000 instances of both lcm divisibilities

in Lemma 1. A second program independently matches all family values through $m = 40$ using prime valuations and a binomial cutoff; a separate literal divisor-indicator search checks $m \leq 5$. The proofs do not depend on the finite checks.

The gap between Corollary 3 and Theorem 5 remains substantial. In particular, no asymptotic formula for $M(m)$ or $A(m)$ is claimed. A focused search of the question, its terminology, hard-family initial values, and related consecutive-integer lcm literature did not locate a matching result; that search does not establish novelty. Törmä’s question had no answer according to its page and the Stack Exchange API checked on 19 October 2026.

Authorship and computational disclosure. This note was prepared for Sulayman Bowles using OpenAI Codex. Codex selected and developed the arguments, wrote the exposition and verification programs, and carried out separately prompted mathematical reviews and exact computations. These checks are not independent human peer review or formal verification. The note is a research draft; the author must understand and take responsibility for it before a scholarly submission.

References

- [1] I. Törmä, *Occurrence of binary words in divisibility patterns*, MathOverflow question 372973, 30 September 2020.
<https://mathoverflow.net/q/372973>.
- [2] B. Farhi, *Minorations non triviales du plus petit commun multiple de certaines suites finies d’entiers*, C. R. Math. Acad. Sci. Paris **341** (2005), no. 8, 469–474. doi:10.1016/j.crma.2005.09.019.
- [3] A. Granville, *Smooth numbers: computational number theory and beyond*, in *Algorithmic Number Theory*, MSRI Publications **44**, Cambridge University Press, 2008, 267–323; especially §3.10.
<https://library.slmath.org/books/Book44/files/09andrew.pdf>.