

A SHORT PROOF OF BUGEAUD'S BOUND FOR STURMIAN DIGIT STRINGS

TacoYaci



tacoyaci.github.io

ABSTRACT. I give a short proof of Bugeaud's S -part bound for integers with Sturmian digit strings. The argument uses the p -adic Subspace Theorem in dimensions two and three and also applies to finite words of linear factor complexity.

1. INTRODUCTION

Bugeaud proved that only finitely many powers of 3 have Sturmian binary representations, as part of a more general S -part bound [Bu]. I give a shorter proof and a finite-word consequence.

Fix an integer $b \geq 2$ and a finite nonempty set S of primes. For a positive integer x , write $[x]_S$ for its largest divisor whose prime factors belong to S , and write its base- b expansion as

$$x = a_0 b^{N-1} + \cdots + a_{N-1}, \quad a_0 \neq 0, \quad a_i \in \{0, \dots, b-1\}.$$

A finite word is *Sturmian* if it is a factor of an infinite word with exactly $u+1$ distinct factors of length u for every $u \geq 1$. Consider integers x whose digit words are Sturmian and have at least two nonzero digits. The following result is due to Bugeaud [Bu, Theorem 1.4].

Theorem 🍌. *For every $\varepsilon > 0$ and all sufficiently large such x , we have*

$$[x]_S < x^{3/4+\varepsilon}.$$

The threshold depends only on b, S, ε . The exponent is unchanged and the bound remains ineffective. Bugeaud's repeated-block method is applied in dimension two for short shifts and dimension three for long shifts.¹ The rest is counting, with a little balance.

2. A REPEATED BLOCK

Let x be any positive integer with N base- b digits. Suppose equal length- u blocks start at positions r and $r+s$, counted from 0, with $u, s \geq 1$ and $r+s+u \leq N$.

Proposition 2. *For every $\eta > 0$ and all sufficiently large x , we have*

$$[x]_S \geq x b^{-u/2+\eta N} \implies b^{N-r} \mid x.$$

The threshold depends only on b, S, η , uniformly over all pairs of repeated blocks.

Proof of Proposition 2. Let f be the difference between the integers represented by the first $r+s$ and first r digits, with an empty prefix representing 0. Cancellation gives

$$(1) \quad E = (b^s - 1)x - b^{N-r}f, \quad |E| < b^{N-r-u}, \quad 0 < f < b^{r+s}.$$

Put $z = x/[x]_S$ and let T consist of S and the prime divisors of b . Suppose there are arbitrarily large counterexamples. They satisfy

$$(2) \quad z \leq b^{u/2-\eta N}, \quad u \geq 2\eta N.$$

I use the p -adic Subspace Theorem [Sc] as stated in [Bu, Theorem 2.1]. For fixed independent forms at each place, an absolute-value product bounded by a fixed negative power of the height puts the integer vectors in finitely many proper rational subspaces. The height is the largest coordinate in absolute value. At primes in T , use coordinate forms and $|p|_p = p^{-1}$. One of the following cases contains infinitely many counterexamples.

Date: Version 1, 11 October 2026.

2020 Mathematics Subject Classification. Primary 11A63, 11J87.

Key words and phrases. Sturmian words, S -parts, Subspace Theorem, digit expansions.

¹Changing a shift can shorten the guaranteed matching length. For the word 1111111, the choices in [Bu, Section 2.2] are $x = 127$, $n = u = 6$, $r = 0$, $s' = 1$, $s = 2$ and $f = 3$. The asserted estimate (2.3) then gives $|127 \cdot 3 - 128 \cdot 3| \leq 2$. Keeping the original shift avoids this issue.

Case $s \leq u/2$. Use the integer vector

$$((b^s - 1)x, b^{N-r}f)$$

and the forms $X_1, X_1 - X_2$ at infinity. The absolute-value product is at most

$$(b^s - 1)x b^{N-r-u} [x]_S^{-1} b^{-(N-r)} < z b^{s-u} \leq b^{-\eta N}.$$

The height is less than $b^{N+s} \leq b^{2N}$, so the Subspace Theorem applies. One rational line contains infinitely many of these vectors. Their coordinate ratio tends to 1, since

$$\frac{|E|}{(b^s - 1)x} < \frac{b^{1-r-u}}{b^s - 1} \rightarrow 0.$$

On that line the ratio is fixed, hence it is 1, i.e. $E = 0$.

Case $s > u/2$. Use the integer vector

$$(xb^s, x, b^{N-r}f)$$

and the forms $X_1, X_2, X_1 - X_2 - X_3$ at infinity. The product is at most

$$x^2 b^s b^{N-r-u} [x]_S^{-2} b^{-s} b^{-(N-r)} = z^2 b^{-u} \leq b^{-2\eta N}.$$

Again the height is less than b^{2N} . There are therefore fixed integers A, B, D , not all zero, such that

$$Axb^s + Bx + Db^{N-r}f = 0$$

along an infinite subsequence. Substituting (1) gives

$$(A + D)b^s + B - D = DE/x \rightarrow 0,$$

since $|E|/x < b^{1-u}$. Now $s > u/2 \rightarrow \infty$, so $A + D = B - D = 0$. This forces $D \neq 0$ and $E = 0$.

In either case, $(b^s - 1)x = b^{N-r}f$. Since $b^s - 1$ is coprime to b , it follows that $b^{N-r} \mid x$, contrary to the choice of counterexamples. $\square$

3. TWO CONSEQUENCES

Finite Sturmian words have at most $u + 1$ distinct length- u factors and are balanced, i.e. equal-length factors differ in the number of occurrences of either letter by at most one (see [Lo, Chapter 2]).

Proof of Theorem 1.1. Take $u = \lfloor (N - 1)/2 \rfloor$. The $N - u + 1 > u + 1$ starting positions give a repeated factor. The word also has at most u trailing zeros. To see this, suppose there are $Z > 0$ trailing zeros. The alphabet is then $\{0, a\}$ with $a \neq 0$. If the second a has index j , balance gives $j \geq Z$, since otherwise the prefix of length $j + 1$ contains two copies of a and the suffix of that length contains none. As $j \leq N - Z - 1$, this gives $2Z \leq N - 1$.

The divisibility in Proposition 2 would give at least $N - r \geq u + s > u$ trailing zeros. Therefore, for any fixed $\eta > 0$ and all sufficiently large x ,

$$[x]_S < xb^{-u/2+\eta N} \leq b^{1/2} x b^{(-1/4+\eta)N}.$$

Choose $0 < \eta < \min\{\varepsilon, 1/4\}$. Since $x < b^N$, the last expression is less than $b^{1/2} x^{3/4+\eta}$, and hence less than $x^{3/4+\varepsilon}$ for sufficiently large x . $\square$

Taking $b = 2$ and $S = \{3\}$ gives the result for powers of 3. For a finite word w , let $p(u, w)$ count its distinct length- u factors.

Corollary 3. Fix $C \geq 1$ and $\varepsilon > 0$. For sufficiently large x with $b \nmid x$, we have

$$p(u, w) \leq Cu \text{ for } 1 \leq u \leq N \implies [x]_S < x^{1-1/(2(C+1))+\varepsilon},$$

where w is the base- b word of x .

Proof. Choose $u = \lfloor N/(C+1) \rfloor$. The $N - u + 1 > Cu$ starting positions give a repeated factor. The divisibility alternative in Proposition 2 is impossible because $b \nmid x$. Apply that proposition with η sufficiently small and use $u \geq N/(C+1) - 1$. $\square$

No infinite-word extension is needed. Integers supported on S with this complexity bound therefore lie in finitely many sets $\{yb^v : v \geq 0\}$ with y fixed. To see this, remove all trailing zeros and apply Corollary 3 to y , since removing a suffix cannot increase factor complexity. This gives a finite-word form of [Bu, Theorem 1.2].

REFERENCES

- [Bu] Y. Bugeaud, *On the binary representation of powers of 3*, [arXiv:2608.23017v1](https://arxiv.org/abs/2608.23017), 2026.
- [Lo] M. Lothaire, *Algebraic Combinatorics on Words*, Cambridge University Press, 2002, Chapter 2, pp. 45–110, [doi:10.1017/CBO9781107326019.003](https://doi.org/10.1017/CBO9781107326019.003).
- [Sc] H. P. Schlickewei, *The p -adic Thue–Siegel–Roth–Schmidt theorem*, Arch. Math. **29** (1977), 267–270, [doi:10.1007/BF01220404](https://doi.org/10.1007/BF01220404).