LANG'S CONJECTURE AND COMPLEXITY OF ALGEBRAIC NUMBERS

ANUP B DIXIT, PURUSOTTAM RATH AND ANANTH SHANKAR

Abstract. A folklore conjecture of Borel asserts that algebraic irrationals have exponential complexity. A remarkable recent work of Bugeaud and Adamczwski shows that algebraic irrationals do not have linear complexity. We show that a conjecture of Lang in Diophantine approximation implies there exists a constant $c > 1$ such that the complexity of any algebraic irrational exceeds c^n for almost all n.

1. Introduction

For any $b \in \mathbb{N}$, consider the base b expansion of a real number α

$$\alpha = [\alpha] + \frac{a_1}{b} + \frac{a_2}{b^2} + \cdots .$$

This associates to α, an infinite word $\mathbf{w} = a_1 a_2 \cdots$ with letters in the alphabet set $\{0, 1, \cdots b-1\}$. We define the complexity function $P_\alpha(n)$ of α in base b to be the number of distinct words of length n that occur in $\mathbf{w}$. Thus $1 \leq P_\alpha(n) \leq b^n$ and it is easy to see that $P_\alpha(n)$ is either a strictly increasing function or is bounded. It is bounded if and only if α is rational. Thus for any irrational α, $P_\alpha(n) \geq n+1$.

Almost all real numbers α have optimal complexity b^n. A weak version of a folklore conjecture due to Borel [3] suggests that an algebraic irrational number α has maximal complexity $P_\alpha(n) = b^n$ for any base b. For instance, it suggests that the elements in the Cantor set are either rational or transcendental. Though this conjecture is far from being solved, a remarkable breakthrough has been made recently by Adamczewski and Bugeaud [1]. Using the subspace theorem, they prove the following result, which is perhaps the first significant progress towards the elusive Borel conjecture:

The complexity function $P_\alpha(n)$ of any algebraic irrational number α satisfies

$$\liminf_{n\to\infty} \frac{P_\alpha(n)}{n} = \infty.$$

In particular, algebraic irrationals cannot have sub-linear complexity.

The proof of the above theorem centres around a transcendence criterion derived from the subspace theorem.

1991 *Mathematics Subject Classification.* Primary 11F20; Secondary 11F11.

Lang has formulated a conjecture which extends Roth's theorem, and in the next section, we formulate a generalised subspace version of this conjecture. Using the techniques developed by Bugeaud and Adamczewski, we prove the following:

Theorem 1.1. *Let $b \geq 2$ be an integer, and α be an algebraic irrational. Then the Subspace-Lang conjecture implies that there exists a constant $c > 1$ (depending on α) such that $P_\alpha(n) \geq c^n$ for almost all n.*

An immediate consequence is the following:

Corollary 1.2. *Subspace-Lang conjecture implies that algebraic irrationals cannot have polynomial complexity .*

The program proposed by E. Borel suggests that algebraic irrationals "behave" like almost all real numbers. Since almost every real number has optimal (exponential) complexity, Borel's hypothesis predicts that algebraic irrationals should also have exponential complexity. Interestingly, the conjecture of Lang is also a vindication of Borel's philosophy. That is, almost all real numbers satisfy the assertion in Lang's conjecture. It is indeed pleasing to see the interrelation between the consequences of Borel's hypothesis in two seemingly different setups.

2. Preliminaries

We begin with some preliminaries on the complexity of real numbers. Let α be any real number. It is a classical result that α is normal in base b if and only if the sequence of fractional parts $\{\alpha b^n\}$ is uniformly distributed in $[0, 1]$. Recall that α is normal in base b, if in the b-ary expansion of α, every word of length n occurs with the same frequency $1/b^n$. Further α has maximal complexity b^n if and only if the sequence of fractional parts $\{\alpha b^n\}$ is dense in $[0, 1]$

We now detail the tools from Diophantine approximation relevant for our investigation. Let us first fix the notation we shall be using. Let K be a number field which is Galois over $\mathbb{Q}$. We denote by M_K (resp. M_∞) the set of all valuations (resp. archimedean valuations) of K; for each valuation ν, we denote by $|\cdot|_\nu$ the absolute value corresponding to ν normalized with respect to K; by this we mean there is an automorphism $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$ of K such that for any $x \in K$,

$$|x|_\nu = |\sigma(x)|^{d(\sigma)/[K:\mathbb{Q}]},$$

where $d(\sigma)$ is the local degree which is 1 if the embedding is real, and 2 if the embedding is complex. We normalise the non-archimedean valuation

accordingly so that the product formula holds and the absolute Weil height reads

$$H(x) = \prod_{\nu \in M_K} \max\{1, |x|_\nu\}.$$

For a vector $\mathbf{x} = (x_1, \cdots, x_n) \in K^n$ and a valuation ν in M_K, we denote by $||\mathbf{x}||_\nu$ the ν-norm of $\mathbf{x}$, namely,

$$||\mathbf{x}||_\nu := \max\{|x_1|_\nu, \cdots, |x_n|_\nu\},$$

We define the projective height $H(\mathbf{x})$ of $\mathbf{x}$ by

$$H(\mathbf{x}) := \prod_{\nu \in M_K} \max\{|x_1|_\nu, \cdots, |x_n|_\nu\}.$$

Finally, we denote the group of S-units in K by $\mathcal{O}_S^\star$:

$$\mathcal{O}_S^\star = \{u \in K : \ |u|_\nu = 1 \text{ for all } \nu \notin S\}.$$

We will be requiring the following p-adic version of the Subspace Theorem (see [2] or [4], for instance) over number fields. We choose for every valuation ν in K an extension of the absolute value $|\cdot|_\nu$ to $\overline{\mathbb{Q}}$, the algebraic closure of $\mathbb{Q}$, which we also denote by the same symbol $|\cdot|_\nu$.

Theorem 2.1. *Let K be an algebraic number field. let $n > 1$ be an integer. Let S be a finite set of valuations of K containing all the infinite valuations. For each ν in S, let $L_{1,\nu}, \cdots L_{n,\nu}$ be linear forms with algebraic coefficients which are linearly independent. Then for any $\epsilon > 0$, the set of solutions $\mathbf{x} \in K^n$ to the inequality*

$$\prod_{\nu \in S} \prod_{i=1}^{n} \frac{|L_{i,\nu}(\mathbf{x})|_\nu}{||\mathbf{x}||_\nu} \leq H(\mathbf{x})^{-n-\epsilon}$$

lie in finitely many proper subspaces of K^n.

In this context, we have the following conjecture of Lang on the order of approximation of algebraic numbers by rational numbers. Lang suggests that the polynomial exponent in Roth's theorem can be replaced by a logarithm factor.

Lang's Conjecture. *Let α be an algebraic number. Then for any $\epsilon > 0$, there exists only finitely many rationals p/q satisfying*

$$0 < |\alpha - p/q| < \frac{1}{q^2 \log^{(1+\epsilon)} q}.$$

We note that by Khintchine's theorem, almost all real numbers satisfy the above criterion. Hence Lang's conjecture is in conformity with the philosophy of E. Borel which proposes that algebraic irrationals behave like almost all real numbers. We formulate a subspace version of this conjecture as follows.

Subspace-Lang conjecture. *Let K be an algebraic number field. Let $n > 1$ be an integer. Let S be a finite set of valuations of K containing all the infinite valuations. For each ν in S, let $L_{1,\nu}, \cdots L_{n,\nu}$ be linear forms with algebraic coefficients which are linearly independent. Then for any $\epsilon > 0$, the set of solutions $\mathbf{x} \in K^n$ to the inequality*

$$\prod_{\nu \in S} \prod_{i=1}^{n} \frac{|L_{i,\nu}(\mathbf{x})|_\nu}{|\mathbf{x}|_\nu} \leq \frac{1}{H(\mathbf{x})^n \log^{n-1+\epsilon} H(\mathbf{x})}$$

lie in finitely many proper subspaces of K^n.

3. Proof of the theorem

We will show that any c satisfying $1 < c < b^{\frac{1}{16}}$ works. We assume that α is algebraic and that $P_\alpha(n) < c^n$ for infinitely many n.

We will prove that α is necessarily rational. Without loss of generality, we may assume that $0 < \alpha < 1$.

There are two steps in the proof. The first step is combinatorial in which we show the following:

There are infinitely many natural numbers N such that the prefix W_N of the b-ary expansion of α has two equal disjoint substrings of length at least $\frac{n}{4}$ and that $N < 2c^n$.

Let us first prove this assertion. By our assumption, there exist infinitely many n satisfying $P_\alpha(n) < c^n$. Fix one such n and set $N = [c^n] + n + 1$. By the Pigeonhole principle, $W_N = w_1 w_2 ... w_N$ contains two equal substrings of length n. If they are disjoint, we are done. Otherwise, W_N contains a substring $W = ABC$, where the strings A, B and C are nonempty and AB and BC are equal and of length n. Suppose the length of A exceeds that of B, then we are again done, as the length of A would be greater than $\frac{n}{3}$ and there would be two disjoint strings both equal to A occurring in W_N. So we can assume that the length of A is less than the length of B. Since the strings AB and BC are equal, A must be a prefix of B. Let $B = AB_1$. Then $AAB_1 = AB_1C$. The argument can be repeated to get $B = AA...AB_i$ until the length of A exceeds that of B_i. But now, it is clear that we get two equal disjoint words (a concatenation of a number of As) whose length is at least $\frac{n}{4}$ as required. Therefore, W_N has a prefix $ABCB$ with length of $B \geq \frac{n}{4}$, where n is as above. Clearly, for large enough n, the inequality $N < 2c^n$ will be satisfied.

Now we come to the second step of the proof which is diophantine. We approximate α by rationals. Let ζ be the rational number with the b-ary expansion $.ABCBC.....$ We have

$$\zeta = \frac{M}{b^r(b^s - 1)}$$

with r and s, the lengths of A and BC respectively and $M < b^r(b^s - 1)$, a positive integer.

We claim that,

$$|\alpha - \zeta| \leq b^{-r-s-\frac{n}{4}}.$$

Indeed, the b-ary expansions of both α and ζ start with $.ABCB$ and the length of the substring is greater than $r + s + \frac{n}{4}$ and so the claim follows. We will now use the Subspace-Lang conjecture to prove that α is rational.

Let S be the set of all valuations (primes) dividing b. Let $T = S \cup \{\infty\}$. Define

$$L_{i,p}(x_1, x_2, x_3) = x_i, i \in \{1, 2, 3\}, \qquad \forall p \in S$$
$$L_{i,\infty}(x_1, x_2, x_3) = x_i, i \in \{1, 2\},$$
$$L_{3,\infty}(x_1, x_2, x_3) = \alpha x_1 + \alpha x_2 + x_3.$$

We will prove that

$$\prod_{\nu \in T} \prod_{i=1}^{3} \frac{|L_{i,\nu}(\mathbf{x})|_\nu}{\|\mathbf{x}\|_\nu} \leq H(\mathbf{x})^{-3} \log^{-3} H(\mathbf{x})$$

where $\mathbf{x} = (b^{r+s}, -b^r, -M)$. We have,

$$\prod_{\nu \in T} \prod_{i=1}^{3} \frac{|L_{i,\nu}(\mathbf{x})|_\nu}{\|\mathbf{x}\|_\nu} = \frac{\prod_{\nu \in T} |b^r|_\nu \prod_{\nu \in T} |b^{r+s}|_\nu \prod_{\nu \in S} |M|_\nu}{\prod_{\nu \in T} \|\mathbf{x}\|_\nu^3} |b^{r+s}\alpha - b^r\alpha - M|_\infty.$$

By the product formula and since T contains all valuations dividing b and the infinite valuation, we have:

$$\prod_{\nu \in T} |b^r|_\nu \, |b^{r+s}|_\nu = 1, \qquad \prod_{\nu \in S} |M|_\nu \leq 1.$$

Also,

$$\prod_{\nu \in T} \|\mathbf{x}\|_\nu^3 \;\geq\; |H(\mathbf{x})|^3$$

because T leaves out only non-archimedean valuations.

Therefore,

$$\prod_{\nu \in T} \prod_{i=1}^{3} \frac{|L_{i,\nu}(\mathbf{x})|_\nu}{\|\mathbf{x}\|_\nu} \;\leq\; |H(\mathbf{x})|^{-3} \;\; |b^{r+s}\alpha - b^r\alpha - M|_\infty.$$

But, we have

$$|b^{r+s}\alpha - b^r\alpha - M| \leq b^{-\frac{n}{4}}.$$

We note that for infinitely many n,

$$\log H(\mathbf{x}) \leq \log \|\mathbf{x}\|_\infty \ll r + s \leq N < 2c^n.$$

As $c < b^{\frac{1}{16}}$, we get that

$$\log H(\mathbf{x})^3 \ll b^{\frac{n}{4}}.$$

Thus, we have

$$\prod_{\nu\in T}\prod_{i=1}^{3}\frac{|L_{i,\nu}(\mathbf{x})|_\nu}{\|\mathbf{x}\|_\nu} \ll H(\mathbf{x})^{-3}\log^{-3}H(\mathbf{x})$$

where $\mathbf{x}=(b^{r+s},-b^r,-M)$ with $s\to\infty$.

By Subspace-Lang conjecture, infinitely many of the points $(b^{r+s},-b^r,-M)$ lie in a proper subspace of $\mathbb{Q}^3$. Thus there exists a non-zero triple $(x_0,y_0,z_0)\in\mathbb{Q}^3$ such that

$$x_0b^{r+s}-y_0b^r-z_0M=0$$

and hence

$$x_0-\frac{y_0}{b^s}-z_0\frac{M}{b^{r+s}}=0$$

As $N\to\infty$, so does s and since

$$\alpha=\lim_{N\to\infty}\frac{M}{b^{r+s}},$$

α is necessarily rational. This completes the proof.

Acknowledgements. We are thankful to Ram Murty for going through the paper. The second author is thankful to the serene working conditions provided by the Department of Mathematics, Orsay during his visit under the ARCUS program where the work got started.

References

[1] B. Adamczewski and Y. Bugeaud, *On the complexity of algebraic numbers. I. Expansion in integer bases*, Ann. of. Math, 2, 165 (2007), 547–565

[2] E. Bombieri and W. Gubler, *Heights in Diophantine geometry.* New Mathematical Monographs, 4. Cambridge University Press, Cambridge, 2006.

[3] E. Borel, *Sur les chiffres décimaux de* $\sqrt{2}$ *et divers problmes de probabilités en chaine.* C. R. Acad. Sci. Paris 230, (1950). 591-593.

[4] W.M. Schmidt, *Diophantine Approximation,* Lecture Notes in Math. 785 (1980), Springer- Verlag, New York.

Chennai Mathematical Institute, Plot No H1, SIPCOT IT Park, Padur PO, Siruseri 603103, Tamil Nadu, India.

E-mail address, Anup B Dixit: adixit@cmi.ac.in

E-mail address, Purusottam Rath: rath@cmi.ac.in

E-mail address, Ananth Shankar: ashankar@cmi.ac.in