BRACKET WORDS: A GENERALISATION OF STURMIAN WORDS ARISING FROM GENERALISED POLYNOMIALS

BORIS ADAMCZEWSKI AND JAKUB KONIECZNY

ABSTRACT. Generalised polynomials are maps constructed by applying the floor function, addition, and multiplication to polynomials. Despite superficial similarity, generalised polynomials exhibit many phenomena which are impossible for polynomials. In particular, there exist generalised polynomial sequences which take only finitely many values without being periodic; examples of such sequences include the Sturmian words, as well as more complicated sequences like $\left\lfloor 2\left\{\pi n^2 + \sqrt{2}n\left\lfloor\sqrt{3}n\right\rfloor\right\}\right\rfloor$.

The purpose of this paper is to investigate letter-to-letter codings of finitely-valued generalised polynomial sequences, which we call *bracket words*, from the point of view of combinatorics on words. We survey existing results on generalised polynomials and their corollaries in terms of bracket words, and also prove several new results. Our main contribution is a polynomial bound on the subword complexity of bracket words.

CONTENTS

1. INTRODUCTION

Generalised polynomial sequences are expressions built up from the usual polynomials with the use of addition, multiplication, and the floor function, such as

$$g(n) = 2n \left\lfloor \sqrt{2}n \lfloor \sqrt{3}n \rfloor^2 + \sqrt{5}n^3 \right\rfloor - 7n^2.$$

They have been investigated by many authors, with particular emphasis placed on problems involving uniform distribution and taking their source in Weyl's equidistribution theorem for classical polynomials. See, for instance, [Hål93, Hål94, BH96, BL07, Lei12, BHKS20] and the references therein.

An important feature that distinguishes generalised polynomials from ordinary polynomials is that they can be bounded without being constant. As a basic example, for any polynomial $p(x) \in \mathbb{R}[x]$, the sequence of fractional parts $(\{p(n)\})_{n=0}^{\infty}$ is a generalised polynomial sequence with values in $[0, 1)$, which is non-constant for most choices of p. In fact, it is also possible for non-constant generalised polynomial sequences to take only finitely many values. A notable example of a class of such finitely-valued generalised polynomial sequences is provided by Sturmian words. Indeed, every Sturmian word $\mathbf{a} = (a_n)_{n=0}^{\infty}$ over $\{0, 1\}$ can be defined by a generalised polynomial formula

$$(1) \qquad a_n = \lfloor \alpha n + \beta \rfloor - \lfloor \alpha(n-1) + \beta \rfloor$$

for some $\alpha \in [0, 1) \setminus \mathbb{Q}$ and $\beta \in [0, 1)$ (possibly with the floor function $\lfloor \cdot \rfloor$ replaced by the ceiling function $\lceil \cdot \rceil$).

Motivated by this fundamental example, we endeavour to investigate finitely-valued generalised polynomials from the perspective of combinatorics on words. More precisely, we study letter-to-letter codings of finitely-valued generalised polynomial sequences, which we dub *bracket words* (cf. Definition 2.2). Thus, a bracket word $\mathbf{a} = (a_n)_{n=0}^{\infty}$ over an alphabet Σ takes the form $a_n = c(g(n))$, where $g \colon \mathbb{N}_0 \to \mathbb{R}$ is a finitely-valued generalised polynomial sequence and c is a map from the finite set $g(\mathbb{N}_0)$ to Σ. Throughout the paper, we let $\mathbb{N} = \{1, 2, \dots\}$ denote the set of positive integers and put $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. Bracket words are thus first thought of as a broad arithmetical generalisation of Sturmian words.

Among the several equivalent definitions of Sturmian words, one involves codings of circle rotations. Indeed, the word $\mathbf{a}$ defined by (1) can be described with the help of the rotation $R_\alpha \colon \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$, $x \mapsto x + \alpha$, as

$$(2) \qquad a_n = \begin{cases} 1 & \text{if } R_\alpha^n(\beta) \in [0, \alpha), \\ 0 & \text{otherwise.} \end{cases}$$

A seminal paper of Bergelson and Leibman [BL07] provides a dynamical representation of any bounded generalised polynomial sequence in terms of nilrotations, which are translations on nilmanifolds, hence linking the theory of generalised polynomials with nilpotent dynamics. As a consequence, each bracket word can be obtained as a coding of a nilrotation with respect to a semialgebraic partition. Conversely, each coding of a nilrotation which involves a semialgebraic partition gives rise to a bracket word.

Nilsystems have received a considerable amount of attention in the past two decades. A major source of interest stems from their significance in the study of multiple ergodic averages, as demonstrated by Host and Kra [HK05] and Ziegler [Zie07]. Nilsystems also play a key role in additive combinatorics, specifically in the

relatively modern branch of higher order Fourier analysis, initiated by Gowers in his work on an alternative proof of Szemerédi's theorem [Gow01]. The importance of nilsystems was revealed by Green, Tao, and Ziegler, who characterised Gowers uniformity in terms of correlations with nilsequences [GTZ12], see also [GT10a]. For an introduction to higher order Fourier analysis, we refer to [Tao12]. Thus, the characterisation of bracket words mentioned above, which we discuss in more detail in Section 4, allows us to take advantage of deep pre-existing results.

A great number of combinatorial results concerning Sturmian words can be derived from the continued fraction expansion of the irrational parameter α and some related expansion, the so-called Ostrowski expansion (see e.g. [Ber01]), of the real parameter β (these parameters being defined as in (1) or (2)). In contrast, the study of bracket words involves much more general Diophantine problems related to simultaneous rational approximations of real numbers. For this reason, one cannot expect to always obtain results as accurate as for Sturmian words. For instance, Sturmian words can be characterised in terms of their subword complexity. Recall that for an infinite word $\mathbf{a}$, the subword complexity $p_{\mathbf{a}}(N)$ is the count of distinct length-N subwords which appear in $\mathbf{a}$, and that $\mathbf{a}$ is Sturmian if and only if $p_{\mathbf{a}}(N) = N + 1$ for each $N \geq 1$, which is the slowest rate of growth possible for a word which is not eventually periodic. Contrary to Sturmian words, bracket words cannot be characterised in terms of their subword complexity, but we still prove the following polynomial upper bound, which is our main new result. It is also shown in Section 12 that this result is essentially the best possible (c.f. Propositions 12.2 and 12.3).

Theorem A. *Let $\mathbf{a}$ be a bracket word. Then there exists a constant $C > 0$ such that $p_{\mathbf{a}}(N) = O(N^C)$ for all $N \in \mathbb{N}$.*

In another direction, Sturmian words can also be characterised in terms of balance: a word $\mathbf{a}$ over $\{0,1\}$ is Sturmian if and only if it is not eventually periodic and for each $N \geq 1$ and each pair u, v of length-N subwords of $\mathbf{a}$, the number of occurrences of the symbol 1 in u and in v is either the same or differs by 1. For bracket words, we have weaker estimates on frequencies of symbols, which we explore in Sections 7 and 8.

Since Sturmian words are bracket words, all of our results apply in particular to Sturmian words. In all instances, the special case involving Sturmian words was shown earlier than the general case, or follows quickly from earlier results. However, there are several facts which we suspect might not be very widely known. These facts include existence of frequencies along the primes (Theorem 7.6) and the IP* recurrence (Theorem 7.11).

A different source from which we draw inspiration is the theory of automatic sequences, and computation theory in general. Recall that a word $\mathbf{a} = (a_n)_{n=0}^{\infty}$ is k-automatic if there exists a deterministic finite automaton which, given the base-k expansion of n as input, produces a_n as output (for extensive introduction, see [AS03]). Although k-automatic words are generally not bracket words (cf. Remark 1.1 below), we find this comparison helpful because the classes of k-automatic sequences and the bracket words enjoy similar closure properties with respect to Cartesian products, codings, finite modifications and certain types of rearrangements, as discussed at length in Section 6. For instance, bracket words taking values in a ring, equipped with coordinatewise addition and multiplication, form

a ring. In contrast, the class of Sturmian sequences is considerably too small for analogous closure properties to be true.

Access to the closure properties mentioned above allows us to look at bracket words from a computational perspective. For example, given three bracket words $\mathbf{a}^{(0)}, \mathbf{a}^{(1)} \in \Sigma^\infty$, and $\mathbf{b} \in \{0, 1\}^\infty$, we can construct a new bracket word $\mathbf{a}$ given by

$$a_n = \begin{cases} a_n^{(0)} & \text{if } b_n = 0, \\ a_n^{(1)} & \text{if } b_n = 1. \end{cases}$$

As a consequence, bracket words can encode any finite computation including real constants, the basic arithmetic operations $+, \times$, the operation of taking the integer part $\lfloor \cdot \rfloor$ and conditional statements involving $=$ and $<$ (c.f. Proposition 5.4).

Remark 1.1. We stress that, even though we use automatic sequences as a motivating example, automatic words in general are not bracket words. In fact, there are no automatic words that are also bracket words, except for eventually periodic ones (see Theorem 11.23). Rather, our interest in automatic sequences stems from the fact that they share with bracket words certain properties, as discussed above. Other classes of sequences which we discuss in this context are morphic sequences and regular sequences (in the sense of Allouche and Shallit [AS92]).

In light of the discussion above, it is natural to ask which properties of bracket words can be tested algorithmically. We will say that a problem is *decidable* if there exists an algorithm which solves it on all inputs, and *undecidable* otherwise. Many properties are known to be decidable for k-automatic sequences, see e.g., [CRS12] and references therein; a practical implementation is discussed in [GHS13] and [Mou16]. Similarly, there are many results on decidability for k-regular sequences (see e.g. [AS92], [KS22]) and for morphic sequences (see e.g. [Dur13a, Dur13b, BPR21]).

For problems involving generalised polynomial sequences, we assume that the algorithm is provided with a formula involving only polynomials, addition, multiplication and the floor function which represents the sequence. In [Lei12], Leibman constructed a "canonical" representation of a bounded generalised polynomial, which is essentially unique. As a consequence, the problem of determining if a given generalised polynomial is zero almost everywhere is decidable. Likewise, it is decidable whether two given bracket words are equal almost everywhere. Here, a statement $\varphi(n)$, involving a parameter $n \in \mathbb{N}_0$, is said to hold almost everywhere if the set $\{n \in \mathbb{N}_0 \mid \neg\varphi(n)\}$ of positions where it is false has asymptotic density zero. Somewhat surprisingly, the problem of verifying equality everywhere turns out to be undecidable, as shown in Section 9. (For terminology used, see Section 2.2 and Remark 2.5.)

Theorem B. *It is undecidable if two given bracket words* $\mathbf{a}$ *and* $\mathbf{b}$ *with algebraic coefficients defined over a finite alphabet* Σ *are equal.*

Finally, we discuss examples of "naturally occurring" words for which we can show that they are, or that they are not, bracket words. In the positive direction, we note that the characteristic word $\mathbf{1}_F = 1111010010010001\cdots$ of the Fibonacci numbers is a bracket word, a property which can be traced back to the observation that the golden ratio $(1 + \sqrt{5})/2$ is a Pisot unit and the group of units of the field it generates has rank 1. In Section 10 we discuss generalisations of this example,

corresponding to other Pisot and Salem numbers. Similar but slightly weaker results were obtained in [BK18].

In the negative direction, several criteria for proving that an infinite word is not a bracket word follow from results discussed in the reminder of the paper. Several other techniques were developed in a series of papers by Byszewski and the second-named author [BK18, BK20, Kon21], leading to a proof that automatic sequences which are not eventually periodic are not bracket words. As explicit applications of the aforementioned methods, we mention that the characteristic words of primes and of squares are not bracket words. The same applies to many words coming from number theory, such as $(\varphi(n) \bmod q)_{n=0}^{\infty}$, where φ denotes the Euler totient function and $q \geq 3$.

We point out that a significant part of the paper is devoted to a survey of known results concerning generalised polynomial sequences and their interpretation in terms of bracket words. Beyond that, we prove several new results, including Theorems A and B.

Organisation of the paper. Sections 2 and 3 are concerned with setting up the terminology and providing examples of bracket words. In Section 4, we discuss the connection between dynamics on nilmanifolds and generalised polynomials, which is one of the key tools used in subsequent sections. For the sake of readability, we delegate some related material to Appendix A. In Sections 5 and 6, we discuss closure properties of bracket words and other ways in which bracket words can be constructed. These results allow us to perform many basic operations on bracket words later in the paper. Sections 7–9 each concern a different facet of bracket words and can mostly be read independently from one another. In Sections 7 and 8, we discuss frequencies of symbols and subwords in bracket words; qualitative results are included in Section 7 and quantitative — in Section 8. In Section 9, we discuss the canonical representation of bracket words, based on [Lei12], and its consequences in terms of decidability; next, we prove several undecidability results, including Theorem B. In Section 10, we consider a class of bracket words, consisting of characteristic words of certain sets of integers described by a linear recurrence. In Section 11, we consider the problem of proving that a given word is not a bracket word. We collect several criteria from previous sections and from [BK18, Kon21], and give a number of new examples and applications. In Section 12, we discuss subword complexity of bracket words and lay the groundwork for Theorem A, which we prove in Sections 13–15.

2. Definitions and notation

2.1. **Combinatorics on words.** An alphabet Σ is a finite set of symbols, also called letters. A finite word over Σ is a finite sequence of letters in Σ or, equivalently, an element of $\Sigma^* = \bigcup_{\ell=0}^{\infty} \Sigma^{\ell}$, the free monoid generated by Σ with respect to the concatenation of finite words. The length of a finite word w, that is, the number of symbols in w, is denoted by $|w|$. We let ϵ denote the empty word, the neutral element of Σ^*. An infinite word $\mathbf{a} = (a_n)_{n=0}^{\infty}$ over Σ is an element of Σ^{∞}, or, equivalently, an infinite sequence with values in Σ (i.e., a map from $\mathbb{N}_0$ to Σ). It is sometimes represented as $\mathbf{a} = a_0 a_1 \cdots$. Throughout, we use bold symbols $\mathbf{a}, \mathbf{b}, \ldots$ to denote infinite words.

Let Σ and Π be two alphabets. A *morphism* is a map $\sigma \colon \Sigma^* \to \Pi^*$ that obeys the identity $\sigma(uv) = \sigma(u)\sigma(v)$ for all words $u, v \in \Sigma^*$. Note that a morphism σ is

uniquely determined by the knowledge of $\sigma(x)$ for all $x \in \Sigma$. A map from Σ to Π^* naturally (and uniquely) extends as a morphism from Σ^* to Π^*. A morphism is said to be *non-erasing* if $\sigma(x) \neq \epsilon$ for all $x \in \Sigma$. A morphism σ over Σ^* is said to be *k-uniform* if $|\sigma(a)| = k$ for every letter a in Σ, and just *uniform* if it is k-uniform for some k. A 1-uniform morphism is called a *coding*. Furthermore, there is a natural way to extend the action of a non-erasing morphism to infinite words, that is, as a map from Σ^∞ to Π^∞ defined by $\sigma(\mathbf{a}) = \sigma(a_0)\sigma(a_1)\cdots$ for $\mathbf{a} \in \Sigma^\infty$. Such a map is still called a morphism and denoted by σ.

2.2. Generalised polynomials and bracket words.
For $x \in \mathbb{R}$, we let $\lfloor x \rfloor \in \mathbb{Z}$ denote the integer part of x (also known as the floor), which is the unique integer with $\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1$. Similarly, we let $\{x\} = x - \lfloor x \rfloor \in [0, 1)$ denote the fractional part, $\lceil x \rceil = -\lfloor -x \rfloor \in \mathbb{Z}$ — the ceiling, $\lfloor x \rceil = \lfloor x + 1/2 \rfloor \in \mathbb{Z}$ — the nearest integer, and $\|x\| = |x - \lfloor x \rceil| \in [0, 1/2]$ — the distance to the nearest integer. We extend the notions introduced above to $x = (x_i)_{i=1}^d \in \mathbb{R}^d$ for $d \geq 2$ coordinate-wise, meaning that $\lfloor x \rfloor = (\lfloor x_i \rfloor)_{i=1}^d \in \mathbb{Z}^d$, etc. Similarly, for a map $f \colon X \to \mathbb{R}$ (where X is any set) we define $\lfloor f \rfloor \colon X \to \mathbb{Z}$ by $\lfloor f \rfloor (x) = \lfloor f(x) \rfloor$.

Let $d \in \mathbb{N}$. We define *GP maps* (or generalised polynomial maps) $\mathbb{R}^d \to \mathbb{R}$ as the smallest family such that

   (i) each polynomial map $\mathbb{R}^d \to \mathbb{R}$ is a GP map;
   (ii) if $g, h \colon \mathbb{R}^d \to \mathbb{R}$ are GP maps then $g + h$ and $g \cdot h$ are GP maps;
   (iii) if $g \colon \mathbb{R}^d \to \mathbb{R}$ is a GP map then $\lfloor g \rfloor$ is a GP map.

Note that $x \mapsto \{x\}, \lceil x \rceil, \lfloor x \rceil$ are GP maps, and the definition of a GP map does not change if in (iii) we replace $\lfloor \cdot \rfloor$ with $\{\cdot\}$, $\lceil \cdot \rceil$ or $\lfloor \cdot \rceil$. The distance for the nearest integer can also be expressed by a generalised polynomial formula, such as $\|x\| = (x - \lfloor x \rfloor) \cdot (2 \lceil x - \lfloor x \rfloor \rceil - 1)$. (Since all maps $\mathbb{R}^d \to \mathbb{R}$ obtained from polynomials using addition, multiplication and $\|\cdot\|$ are continuous, replacing $\lfloor \cdot \rfloor$ with $\|\cdot\|$ in (iii) yields a strictly smaller family of maps.)

Remark 2.1. If $g \colon \mathbb{R} \to \mathbb{R}$ and $h \colon \mathbb{R}^d \to \mathbb{R}$ are GP maps, then $g \circ h$ is also a GP map.

A GP map on a domain $\Omega \subseteq \mathbb{R}^d$ (e.g. $\Omega = \mathbb{Z}^d$ or $\mathbb{N}_0^d$) is simply the restriction of a GP map on $\mathbb{R}^d$. In particular, each GP map on Ω can be extended to $\mathbb{R}^d$, but the extensions are usually not unique. In this paper, we are particularly interested in finitely-valued GP sequences (i.e., GP maps $\mathbb{N}_0 \to \mathbb{R}$), as seen from the perspective of combinatorics on words. This motivates us to pose Definition 2.2.[1]

Definition 2.2. A *bracket word* over a (finite) alphabet Σ is an infinite word $\mathbf{a} = (a_n)_{n=0}^\infty \in \Sigma^\infty$ of the form $a_n = c(g(n))$ for all $n \in \mathbb{N}_0$, where $g \colon \mathbb{N}_0 \to \mathbb{R}$ is a finitely-valued GP map and $c \colon g(\mathbb{N}_0) \to \Sigma$ is an arbitrary map.

Remark 2.3. The inclusion of the coding in Definition 2.2 does not significantly increase the level of generality. Indeed, if $\mathbf{a} = (a_n)_{n=0}^\infty$ is a bracket word over an alphabet Σ and $\phi \colon \Sigma \to \mathbb{R}$ is an arbitrary map then the map $g \colon \mathbb{N}_0 \to \mathbb{R}$ given by $g(n) = \phi(a_n)$ is a GP map as a consequence of Corollary 5.3. However, from the perspective of combinatorics on words, it would be unnatural to restrict our

---

[1]Several authors have expressed the sentiment that a better name for "generalised polynomials" would have been "bracket polynomials", and the main reason to not adopt the latter name is that it is already used in knot theory (see e.g. [Lei12]). Fortunately, similar considerations do not apply to the term "bracket word".

attention to words over alphabets contained in $\mathbb{R}$. We also point out that in Section 6 it is frequently more natural to consider words over more general alphabets, especially when it comes to closure under Cartesian products in Proposition 6.2.

Remark 2.4. In analogy with Definition 2.2, one can define *d-dimensional bracket words* over a finite alphabet Σ to be d-dimensional infinite words

$$\mathbf{a} = (a_{n_1, n_2, \ldots, n_d})_{n_1, n_2, \ldots, n_d = 0}^{\infty} \in \Sigma^{\infty} \times \Sigma^{\infty} \times \cdots \times \Sigma^{\infty}$$

of the form $a_{n_1, n_2, \ldots, n_d} = c(g(n_1, n_2, \ldots, n_d))$ where $g \colon \mathbb{N}_0^d \to \mathbb{R}$ is a finitely-valued GP map and $c \colon g(\mathbb{N}_0^d) \to \Sigma$ is an arbitrary map. We limit the discussion to the 1-dimensional case for the sake of clarity, but many of the results have their multi-dimensional analogues. In particular, we point out that multidimensional Sturmian sequences (as defined in [Fer06, Def. 4]), and more generally rotation sequences (as defined in [Fer06, Def. 5]) are multidimensional bracket words (cf. [Fer06, Prop. 2]).

Remark 2.5. For a ring $\mathbb{Z} \subseteq A \subseteq \mathbb{R}$ we define *GP maps with coefficients in A* in a fully analogous way, except that in (i) we only include polynomial maps with coefficients in A. We will be especially interested in GP maps with algebraic coefficients, i.e., $A = \overline{\mathbb{Q}}$. One can also define the set of coefficients of a generalised polynomial (see e.g. [Hål94]), but we avoid using this notion since it depends on the choice of a representation, which is usually not unique (see Section 9 for further discussion). Slightly informally, we will say that an infinite word $\mathbf{a}$ over Σ is a *bracket word arising from a GP map with algebraic coefficients*, or simply a *bracket word with algebraic coefficients*, if $a_n = c(g(n))$ for every $n \in \mathbb{N}_0$, where $g \colon \mathbb{N}_0 \to \mathbb{R}$ is a finitely-valued GP map with coefficients in $\overline{\mathbb{Q}}$ and $c \colon g(\mathbb{N}_0) \to \Sigma$ is a map.

For $k \geq 2$, a map $g = (g_i)_{i=1}^k \colon \mathbb{R}^d \to \mathbb{R}^k$ is GP if for each i, $1 \leq i \leq k$, the coordinate map g_i is GP. In most cases, we find it simpler to speak of k-tuples of GP maps instead.

A *GP subset of $\Omega \subseteq \mathbb{R}^d$* (or simply a *GP set*, if Ω is clear from the context) is the zero locus of a GP map, that is, a set $E \subseteq \Omega$ which takes the form $E = \{x \in \Omega \mid g(x) = 0\}$ for some GP map $g \colon \Omega \to \mathbb{R}$. As we will see in Section 5, $E \subseteq \mathbb{N}_0$ is a GP set if and only if $\mathbf{1}_E$ is a bracket word, cf. Proposition 5.4. (Here and elsewhere, $\mathbf{1}_E \in \{0, 1\}^{\infty}$ is given by $(\mathbf{1}_E)_n = 1$ if $n \in E$ and $(\mathbf{1}_E)_n = 0$ otherwise). We stress that this notion depends on Ω; in particular, a GP subset of $\mathbb{N}_0$ will usually not be a GP subset of $\mathbb{Z}$.

2.3. Other notation.
We briefly summarise some other pieces of notation we use.

For $N \in \mathbb{N}_0$, we let $[N] = \{0, 1, \ldots, N-1\}$. For a quantity X, we let $O(X)$ denote any quantity bounded in absolute value by CX, where C is a constant. When the C is additionally allowed to depend on a parameter Y, we write $O_Y(X)$ instead. Similarly, assuming that $X > 0$, we let $\Omega(X)$ denote any quantity bounded from below by cX, where $c > 0$ is a constant. If $X = O(Y)$ and $Y = \Omega(X)$, we write $Y = \Theta(X)$. We occasionally also use the notation $Y \ll X$ when $Y = O(X)$. Lastly, we write $o_{n \to \infty}(X)$ for any quantity Y with $\lim_{n \to \infty} Y/X = 0$; if the parameter n is clear from the context, we write $o(X)$ instead.

In what follows, we will use the Iverson bracket notation. For a sentence φ we put $[\![\varphi]\!] = 1$ if φ is a true and $[\![\varphi]\!] = 0$ otherwise. By a slight abuse of notation, we also use the Iverson bracket to define infinite words over the alphabet $\{0, 1\}$. For

instance, if X, Y are sets with $Y \subseteq X$ and $f \colon \mathbb{N}_0 \to X$ is a map then $[\![f \in Y]\!] \in \{0,1\}^\infty$ is given by $[\![f \in Y]\!]_n = [\![f(n) \in Y]\!]$ for all $n \in \mathbb{N}_0$.

3. Examples

Let us now present several examples of bracket words. In some cases, the fact that the sequence under consideration indeed is a bracket word will follow directly from the definition, while in other cases it may be more surprising.

Example 3.1. Let $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ and let **a** be the bracket word defined by $a_n = \lfloor 1 - \{ n\alpha \} \rfloor$ for $n \in \mathbb{N}_0$. Then $\mathbf{a} = \mathbf{1}_{\{0\}} = 100\cdots$. More generally, all eventually constant sequences are bracket words. Explicitly, if **a** is a word over Σ with $a_n = b$ for all $n \geq N$ then $a_n = c(g(n))$, where $g \colon \mathbb{Z} \to \{-1, 0, 1, \dots, N-1\}$ is the GP map given by

$$g(n) = -1 + \sum_{m=0}^{N-1} (m+1) \lfloor 1 - \{ (n-m)\alpha \} \rfloor,$$

and $c \colon \{-1, 0, 1, \dots, N-1\} \to \Sigma$ is given by $c(n) = a_n$ if $n \neq -1$ and $c(-1) = b$.

Example 3.2. Let $Q \in \mathbb{N}$. Then $\mathbf{a} \in [Q]^\infty$ defined by $a_n = n \bmod Q = Q \{ n/Q \}$ is a bracket word. More generally, all eventually periodic words are bracket words. For details, see Section 6.

Example 3.3. One way to define Sturmian words is by an explicit formula. Namely, an infinite word **a** over $\{0,1\}$ is Sturmian if it is the discrete derivative of a Beatty sequence, meaning that it takes one of the following forms:

$$(3) \qquad\qquad a_n = \lfloor n\alpha + \beta \rfloor - \lfloor \alpha(n-1) + \beta \rfloor, \quad \text{or}$$

$$(4) \qquad\qquad a_n = \lceil n\alpha + \beta \rceil - \lceil \alpha(n-1) + \beta \rceil,$$

for some $\alpha, \beta \in [0,1)$ with α irrational. (Note that (3) and (4) differ for at most one value of n.) Thus, Sturmian words are bracket words (while Beatty sequences are unbounded GP sequences). We also point out that Sturmian words arise from codings of rotations, which gives another way to see that they are bracket words; we explore this point of view further in Section 4.

More concretely, setting $\alpha = \frac{\sqrt{5}-1}{2}$ and $\beta = 0$ we obtain the Fibonacci word, whose initial values are:

$$1010110101101101011010110110101101101010110101010\cdots.$$

Example 3.4. As a generalisation of Example 3.3, let $p \colon \mathbb{R} \to \mathbb{R}$ be a polynomial and let $I \subseteq [0,1)$ be an interval (or a finite union thereof). Let **a** be the infinite word over $\{0,1\}$ defined as

$$a_n = \begin{cases} 1 & \text{if } \{p(n)\} \in I, \\ 0 & \text{otherwise.} \end{cases}$$

Then **a** is a bracket word. As a concrete illustration, let $p(x) = \varphi x^2$, where $\varphi = \frac{\sqrt{5}+1}{2}$ is the golden ratio, and $I = [0, 1/4) \cup (3/4, 1)$. Then

$$a_n = \begin{cases} 1 & \text{if } \left\| \varphi n^2 \right\| < 1/4, \\ 0 & \text{if } \left\| \varphi n^2 \right\| \geq 1/4. \end{cases}$$

The initial values of **a** are:

$$1000101001111011101110111010011111011101111000111110111\cdots.$$

Example 3.5. Let $F = \{0, 1, 2, 3, 5, 8, 13, \dots\}$ be the set of all Fibonacci numbers. Then $\mathbf{1}_F$ is a bracket word. Similarly, let $(t_i)_{i=0}^{\infty}$ be the sequence given by $t_0 = 0$, $t_1 = t_2 = 1$ and $t_{i+3} = t_{i+2} + t_{i+1} + t_i$ for all $i \in \mathbb{N}_0$, sometimes called the Tribonacci numbers, and let $T = \{t_i \mid i \in \mathbb{N}_0\}$. Then $\mathbf{1}_T$ is a bracket word. These are special cases of Theorem 10.1. In an upcoming preprint by the second-named author and Byszewski [BK22], it is shown that, more generally, for each $E \subseteq F$, $\mathbf{1}_E$ is a bracket word.

Example 3.6. Let $\alpha, \beta \in \mathbb{R}$ and $\varepsilon > 0$. Then the infinite word **a** over $\{0, 1\}$ defined by

$$a_n = \begin{cases} 1 & \text{if } n \neq 0 \text{ and } \|\alpha n\| \cdot \|\beta n\| < \varepsilon/n, \\ 0 & \text{otherwise} \end{cases}$$

is a bracket word, as follows from Proposition 5.4. We point out that a famous conjecture in Diophantine approximation, the Littlewood conjecture, is equivalent to the statement that, for each choice of $\alpha, \beta, \varepsilon$, the bracket word **a** defined above is not identically zero. Indeed, in its usual formulation, Littlewood's conjecture asserts that

$$(5) \qquad \liminf_{n \to \infty} n \cdot \|\alpha n\| \cdot \|\beta n\| = 0$$

for all $\alpha, \beta \in [0, 1)$. A landmark result towards its resolution is due to Einsiedler, Katok and Lindenstrauss [EKL06]: the set of possible exceptions (i.e., the set of pairs (α, β) for which (5) is false) has Hausdorff dimension zero.

Example 3.7. Let $(n_i)_{i=0}^{\infty}$ be a sequence of positive integers with $n_{i+1} \geq n_i^2$ for all i, such as $n_i = 2^{2^i}$. Put $E = \{n_i \mid i \in \mathbb{N}_0\}$. Then $\mathbf{1}_E$ is a bracket word. This is a special case of Proposition 8.4.

4. DYNAMICAL REPRESENTATION

In this section, we discuss a dynamical description of bounded GP sequences and bracket words. Specifically, we briefly introduce basic facts about nilmanifolds and nilsystems and explain their relation to generalised polynomials, as established by Bergelson and Leibman [BL07].

4.1. **Nilsystems and generalised polynomials.** Classical theory of nilpotent dynamics can be found in [AGH63]. In order to maintain the introductory nature of this section, we delegate some of the more technical results to Appendix A. We also refer to cited references, such as [BL07], for precise definitions and a more detailed discussion.

4.1.1. *Nilpotent Lie groups.* Let G be a group. The *lower central series* $(G_i)_{i \geq 0}$ is the chain of subgroups of G inductively defined by $G_0 = G_1 = G$ and $G_{i+1} = [G, G_i]$ for $i \geq 1$. Here, we let $[G, H]$ denote the group generated by the commutators $[g, h] = ghg^{-1}h^{-1}$ for $g \in G, h \in H$. The group G is nilpotent if there exists s such that $G_{s+1} = \{\mathrm{id}_G\}$. The smallest such s is called the nilpotency class of G and G is said to be nilpotent of class s or an s-step nilpotent group. We recall that a Lie group is a smooth manifold obeying the group properties and that satisfies the

additional condition that the group operations are differentiable. A nilpotent Lie group is a Lie group that is nilpotent.

4.1.2. *Nilrotations.* A *nilmanifold* is a quotient space G/Γ where G is a nilpotent Lie group and Γ is a discrete cocompact subgroup. A *nilsystem* is a dynamical system of the form $(G/\Gamma, T_g)$, where G/Γ is a nilmanifold and T_g is a *nilrotation*. That is, there exists some $g \in G$ such that $T_g(h\Gamma) = gh\Gamma$ for all $h \in G$. In general, there is no guarantee that G is connected, and we let G° denote the connected component of id_G. We may assume without loss of generality that G° is simply connected. The simplest example of a nilsystem is the aforementioned rotation on the torus, where we take $G = \mathbb{R}$, $\Gamma = \mathbb{Z}$ (see Section 4.2).

4.1.3. *Mal'cev basis.* Let G be a connected and simply connected s-step nilpotent Lie group and let $\Gamma < G$ be a discrete cocompact subgroup. In this case, for $g \in G$ and $t \in \mathbb{R}$, one can use the Lie algebra of G to define $g^t \in G$. We let $\dim G$ denote the dimension of G as a Lie group. A *Mal'cev basis* of G is a sequence $h_1, h_2, \ldots, h_d \in \Gamma$ satisfying the following conditions.

(i) Every $g \in G$ has a unique representation $h_1^{t_1} h_2^{t_2} \cdots h_d^{t_d}$ with $t_i \in \mathbb{R}$.
(ii) There exists an increasing sequence of natural numbers

$$1 = k_1 < k_2 < \cdots < k_s = d + 1$$

such that for each $1 \le j \le s$, the quotient G_j/G_{j+1} is spanned by $h_{k_j}, \ldots, h_{k_{j+1}-1}$.

Existence of such bases was established by Mal'cev [Mal49, Mal51]. Given a Mal'cev basis, we let $\widetilde{\tau} \colon G \to \mathbb{R}^d$ denote the coordinate map, characterised by the property that

$$(6) \qquad \widetilde{\tau}\left(h_1^{t_1} h_2^{t_2} \cdots h_d^{t_d}\right) = (t_1, t_2, \ldots, t_d), \qquad t_i \in \mathbb{R}.$$

This induces also the coordinate map $\tau \colon G/\Gamma \to [0,1)^d$, similarly characterised by

$$(7) \qquad \tau\left(h_1^{t_1} h_2^{t_2} \cdots h_d^{t_d}\Gamma\right) = (t_1, t_2, \ldots, t_d), \qquad t_i \in [0,1).$$

Thus, the nilmanifold G/Γ can be identified with a cube $[0,1)^{\dim G}$ via *Mal'cev coordinates* $\tau \colon G/\Gamma \to [0,1)^{\dim G}$. The coordinate map τ is a bijection; τ^{-1} is continuous, its restriction to $(0,1)^{\dim G}$ is a diffeomorphism. The nilmanifold G/Γ carries a natural probability measure, the Haar measure, which we denote by $\mu_{G/\Gamma}$.

4.1.4. *Semialgebraic sets.* An algebraic variety in $\mathbb{R}^d$ is a set defined by a finite number of polynomial equations. More generally, a semialgebraic set is a set defined by a finite number of polynomial equations and inequalities, or a finite union of sets of this form. A map $f \colon \mathbb{R}^d \to \mathbb{R}$ is *piecewise polynomial* if there exists a partition $\mathbb{R}^d = S_1 \cup S_2 \cup \cdots \cup S_r$ into semialgebraic pieces such that, for every i, $1 \le i \le r$, the restriction $f|_{S_i}$ is a polynomial map. A map $f \colon G/\Gamma \to \mathbb{R}$ is piecewise polynomial if it takes the form $f = \widetilde{f} \circ \tau$ for a piecewise polynomial map $\widetilde{f} \colon \mathbb{R}^{\dim G} \to \mathbb{R}$. The notion of a piecewise polynomial map is independent of the choice of Mal'cev basis (see [BL07, Sec. 0.18]).

4.1.5. *Connectedness.* Let us now return to the general case, where G may be disconnected. If G/Γ is connected, then it remains true that $G/\Gamma = G^\circ/\Gamma \cap G^\circ$, and the previous discussion applies verbatim (note, however, that not every translation T_g, with $g \in G$, can be represented as T_h, with $h \in G^\circ$). If G/Γ is disconnected, then it can be decomposed as a finite union of connected components which again are nilmanifolds, and we can apply the discussion above to each component separately. A map $f \colon G/\Gamma \to \mathbb{R}$ is piecewise polynomial if its restriction to each connected component of G/Γ is piecewise polynomial.

Finally, we recall that a topological dynamical system (X, T) is *minimal* if there is no closed subset $Y \subseteq X$ with $T(Y) \subseteq Y$. We have now introduced all the terminology which is needed to state the 1-dimensional case of the main result of [BL07].

Theorem 4.1 ([BL07, Thm. A])**.** *Any bounded GP map $g \colon \mathbb{Z} \to \mathbb{R}$ admits a representation $g(n) = f(T^n(x))$, where (X, T) is a minimal nilsystem, $f \colon X \to \mathbb{R}$ is piecewise polynomial, and $x \in X$. Conversely, for any nilsystem (X, T), any piecewise polynomial map $f \colon X \to \mathbb{R}$, and any $x \in X$, the map $n \mapsto f(T^n(x))$ from $\mathbb{Z}$ to $\mathbb{R}$ is a bounded GP map.*

For future reference, we record the following special case of Theorem 4.1 applicable to bracket words.

Theorem 4.2. *For any bracket word $\mathbf{a}$ over an alphabet Σ, there exist a minimal nilsystem (X, T), a point $x \in X$, and a partition $X = \bigcup_{i \in \Sigma} S_i$ into pairwise disjoint semialgebraic pieces, such that for each $n \in \mathbb{N}$ and $i \in \Sigma$,*

$$(8) \qquad a_n = i \qquad\qquad \text{if and only if} \qquad\qquad T^n(x) \in S_i.$$

Conversely, for any nilsystem (X, T), any point $x \in X$ and any partition $X = \bigcup_{i \in \Sigma} S_i$ into pairwise disjoint semialgebraic pieces, (8) defines a bracket word.

Proof. Pick a representation $a_n = c(g(n))$ $(n \in \mathbb{N}_0)$, where g is a finitely-valued GP map and c is a coding. Let $g(n) = f(T^n(x))$ be the representation of g as in Theorem 4.1. Then, for each $i \in \Sigma$, $S_i = f^{-1}(c^{-1}(i))$ is a semialgebraic set such that (8) holds. The converse implication follows along similar lines. $\qquad\square$

Remark 4.3. The spectral properties of nilsystems are well-understood: For a nilsystem (X, T) with Haar measure μ_X, the Hilbert space $L^2(\mu_X)$ can be decomposed as the direct sum $\mathcal{H} \oplus \mathcal{H}'$, with $\mathcal{H}, \mathcal{H}'$ both invariant under the action of T, such that the restriction of T to $\mathcal{H}$ has discrete spectrum, and either $\mathcal{H}'$ is trivial (this is the case when X is 1-step nilpotent) or the restriction of T to $\mathcal{H}'$ has Lebesgue spectrum of infinite multiplicity [AGH63, Sta00, HKM14]. This can be used to make inferences about the spectra of bracket words. In particular, one can construct bracket words such that the corresponding subshifts have either pure discrete spectrum or Lebesgue spectrum of infinite multiplicity.

Now, we give two emblematic examples of nilsystems and the corresponding GP maps. While we do not use them elsewhere in the paper, we believe they help to illustrate Theorem 4.1.

4.2. **One dimensional torus and Sturmian words.** As already mentioned in Section 1, Sturmian words can be dynamically represented as codings of irrational translations on the one-dimensional torus. In this case, we simply take $G = \mathbb{R}$,

$\Gamma = \mathbb{Z}$, and the nilrotation $T = T_\alpha$ is just an irrational translation on $\mathbb{R}/\mathbb{Z}$. More precisely, let **a** be the Sturmian word defined by (3) (or, respectively, by (4)). Let $T_\alpha \colon \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ denote the translation by α on $\mathbb{R}/\mathbb{Z}$, meaning that $T_\alpha(x) = x + \alpha$ ($x \in \mathbb{R}/\mathbb{Z}$). Let $I \subseteq \mathbb{R}/\mathbb{Z}$ be the interval $[0, \alpha)$ (resp. $(0, \alpha]$). Then, (3) (resp. (4)) is equivalent to:

$$a_n = 1_I(T_\alpha^n(\beta)) = \begin{cases} 1 & \text{if } T_\alpha^n(\beta) \in I, \\ 0 & \text{otherwise.} \end{cases}$$

4.3. **Heisenberg group.** Another helpful example to keep in mind is the Heisenberg nilsystem. It is a standard example, and appears e.g. in [BL07, Sec. 0.14], [GT12b, Sec. 5], [GT10a, Sec. 1], [GT12a, p. 555], [GTZ12] (as a running example). Pick any $\alpha, \beta, \gamma \in \mathbb{R}$, and set

$$G = \begin{bmatrix} 1 & \mathbb{R} & \mathbb{R} \\ 0 & 1 & \mathbb{R} \\ 0 & 0 & 1 \end{bmatrix}, \quad \Gamma = \begin{bmatrix} 1 & \mathbb{Z} & \mathbb{Z} \\ 0 & 1 & \mathbb{Z} \\ 0 & 0 & 1 \end{bmatrix}, \text{ and } h = \begin{bmatrix} 1 & \beta & \gamma + \alpha\beta/2 \\ 0 & 1 & \alpha \\ 0 & 0 & 1 \end{bmatrix}.$$

One possible choice for Mal'cev coordinates is given by

$$\tau\left(\begin{bmatrix} 1 & y & z \\ 0 & 1 & x \\ 0 & 0 & 1 \end{bmatrix} \right) = (x, y, z).$$

Then we can compute that

$$h^n \Gamma = \begin{bmatrix} 1 & n\beta & n\gamma + n^2\alpha\beta/2 \\ 0 & 1 & n\alpha \\ 0 & 0 & 1 \end{bmatrix} \Gamma$$

$$= \begin{bmatrix} 1 & \{n\beta\} & \{n\beta\{n\alpha\} - n^2\alpha\beta/2 + \gamma n\} \\ 0 & 1 & \{n\alpha\} \\ 0 & 0 & 1 \end{bmatrix} \Gamma.$$

Hence, using the nilsystem $(G/\Gamma, T_h)$ and taking $f = \tau_3$ (i.e., the third entry of the coordinate map τ), we obtain a representation of the bounded GP map

$$g(n) = \{n\beta\{n\alpha\} - n^2\alpha\beta/2 + \gamma n\}.$$

A slightly more complicated but similar construction involving matrices in dimension 4 discussed in [BL07, Sec. 0.14] provides a dynamical representation of the sequence $\{n\beta\{n\alpha\}\}$.

5. REPRESENTATIONS AND CONSTRUCTIONS

In this section we discuss methods by which bracket words can be represented and constructed.

5.1. **Representations.** It is clear from the definition that GP maps : $\mathbb{N}_0 \to \mathbb{R}$ are precisely those maps which can be expressed using (classical) polynomials, addition, multiplication and the floor function. For finitely-valued GP maps and bracket words, the situation becomes more complicated. As we will see in Section 9, it is not always possible to decide if a given GP map $g \colon \mathbb{N}_0 \to \mathbb{R}$ is finitely-valued (and hence relevant to the study of bracket words) or not. Thus, it is not always possible to see if a given formula represents a bracket word. On the other hand, for suitably constructed GP maps, it is easy to see that they must be finitely-valued. For instance, if $g(n) = \lfloor 2\{h(n)\} \rfloor$ for some GP map $h \colon \mathbb{N}_0 \to \mathbb{R}$, then evidently

$g(\mathbb{N}_0) \subseteq \{0,1\}$. In general, one can always represent a finitely-valued GP map in a form which makes it easy to estimate the cardinality of the image. We stress that Proposition 5.1 provides a concrete way to generate all finitely valued GP maps from $\mathbb{N}_0$ to $\mathbb{R}$, and hence all bracket words.

Proposition 5.1. *Let $g\colon \mathbb{N}_0 \to \mathbb{R}$ be a finitely-valued GP map. Then g can be written in the form $g(n) = f\left(\lfloor N\{h(n)\}\rfloor\right)$, where f and h are GP maps from $\mathbb{N}_0$ to $\mathbb{R}$ and $N = |g(\mathbb{N}_0)|$. Conversely, any map $g\colon \mathbb{N}_0 \to \mathbb{R}$ of the form $g(n) = f\left(\lfloor N\{h(n)\}\rfloor\right)$, where f and h are GP maps from $\mathbb{N}_0$ to $\mathbb{R}$ and $N \in \mathbb{N}$, is a GP map which takes at most N distinct values.*

In the course of the proof of Proposition 5.1, we will need the following simple fact, which clarifies the relation between bracket words and finitely-valued GP maps.

Lemma 5.2. *Let $f\colon \mathbb{N}_0 \to \mathbb{R}$ be a map taking finitely many values. Then the two following properties are equivalent.*

  (i) *The word $\mathbf{f} = (f(n))_{n=0}^{\infty}$ is a bracket word.*
  (ii) *The map $f\colon \mathbb{N}_0 \to \mathbb{R}$ is a GP map.*

Proof. It is clear that (ii) implies (i): in Definition 2.2, we can take $g = f$ and $c = \mathrm{id}$. For the converse implication, suppose that $f(n) = c(g(n))$, where $g\colon \mathbb{N}_0 \to \mathbb{R}$ has finite image and $c\colon g(\mathbb{N}_0) \to \mathbb{R}$. There exists a polynomial map $p\colon \mathbb{R} \to \mathbb{R}$ such that $p(x) = c(x)$ for all $x \in g(\mathbb{N}_0)$. Hence, $f = p \circ g$ is a GP map. $\square$

Proof of Proposition 5.1. Let $c\colon g(\mathbb{N}_0) \to [N]$ be any bijective map. Then the word $\mathbf{a} = (c(g(n))_{n=0}^{\infty}$ is a bracket word taking values in $[N]$. Hence, by Lemma 5.2, $c\circ g$ is a GP map. Let h be the GP map defined by $h(n) = c(g(n))/N$, and let $f\colon \mathbb{N}_0 \to \mathbb{R}$ be a polynomial map such that $f(c(x)) = x$ for each $x \in g(\mathbb{N}_0)$ (such f exists by polynomial interpolation). It is straightforward to check that

$$f(\lfloor N\{h(n)\}\rfloor) = f(c(g(n))) = g(n)$$

for all $n \in \mathbb{N}_0$, as needed. The converse direction holds trivially. $\square$

As a consequence of Proposition 5.1, for each bracket word we can construct a particularly convenient representation.

Corollary 5.3. *Let $\mathbf{a}$ be a bracket word defined over a finite alphabet Σ and let $N = |\Sigma|$. Then there exist a GP map $g\colon \mathbb{N}_0 \to [N]$ and a map $c\colon [N] \to \Sigma$ such that $a_n = c(g(n))$ for all $n \in \mathbb{N}_0$.*

Proof. Let $a_n = c'(g'(n))$, where $g'\colon \mathbb{N}_0 \to \mathbb{R}$ is a finitely-valued GP map and $c'\colon g'(\mathbb{N}_0) \to \Sigma$. Let $g'(n) = f(\lfloor N\{h(n)\}\rfloor)$ be the representation of g' from Proposition 5.1. It remains to define g and c by $g(n) = \lfloor N\{h(n)\}\rfloor$ and $c(x) = c'(f(x))$. $\square$

5.2. **Constructions.** Next, we discuss a basic tool, which can be used to construct potentially interesting examples of bracket words. In what follows, we will use the Iverson bracket notation, where $[\![g \in I]\!]$ denotes the word over the alphabet $\{0,1\}$ which records the positions n such that $g(n) \in I$ (see Section 2.3 for exact definition).

Proposition 5.4. *Let $g\colon \mathbb{N}_0 \to \mathbb{R}$ be a GP map and $I \subseteq \mathbb{R}$ be an interval (possibly infinite or degenerate). Then $[\![g \in I]\!]$ is a bracket word.*

Proof. The case where I is bounded is covered by [BK18, Lemma 1.2], while the case where I is unbounded follows from [Kon21, Lemma B.3]. $\square$

Remark 5.5. We point out that the analogous result with $\mathbb{N}_0$ replaced by $\mathbb{Z}$ is true for bounded I and false for unbounded I. For instance, $1_{\mathbb{N}}(n) = [\![n \in (0, \infty)]\!]$ is not a GP map on $\mathbb{Z}$ (c.f. [Kon21, Ex. B.2]). This is one of the reasons why we focus on one-sided bracket words $(a_n)_{n=0}^{\infty}$, rather than on their two-sided analogues $(a_n)_{n \in \mathbb{Z}}$.

We will use later the following slight refinement of Proposition 5.4 in the case $I = \{0\}$.

Lemma 5.6. *Let $g \colon \mathbb{N}_0 \to \mathbb{R}$ be a GP map with coefficients in some field K with $\mathbb{Q} \subsetneq K \subseteq \mathbb{R}$. Then $[\![g = 0]\!]$ is a bracket word with coefficients in K.*

Proof. Pick any $\alpha \in K \setminus \mathbb{Q}$. Note that the only solution to $\{x\} = \{\alpha x\} = 0$ in $\mathbb{R}$ is $x = 0$. Hence,

$$\left\lfloor 1 - \frac{1}{2}\{x\} - \frac{1}{2}\{\alpha x\} \right\rfloor = [\![x = 0]\!], \qquad x \in \mathbb{R}.$$

It follows that

$$[\![g(n) = 0]\!] = \left\lfloor 1 - \frac{1}{2}\{g(n)\} - \frac{1}{2}\{\alpha g(n)\} \right\rfloor, \qquad n \in \mathbb{N}_0. \qquad \square$$

One of the main reasons for interest in Lemma 5.6 is that it gives an alternative definition of bracket words, phrased in terms of fibres.

Corollary 5.7. *Let $\mathbf{a}$ be an infinite word defined over a finite alphabet Σ. Then the two following properties are equivalent.*

(i) *The word $\mathbf{a}$ is a bracket word.*
(ii) *For every $x \in \Sigma$, the fibre $\{n \in \mathbb{N}_0 \mid a_n = x\}$ is a GP subset of $\mathbb{N}_0$.*

In particular, for a set $E \subseteq \mathbb{N}_0$, E is a GP set if and only if $\mathbf{1}_E$ is a bracket word.

Proof. Let us assume that $\mathbf{a}$ is a bracket word defined over Σ and let $x \in \Sigma$. Let $a_n = c(g(n))$ be a representation of $\mathbf{a}$, where $g : \mathbb{N}_0 \to \mathbb{R}$ is a finitely valued GP map and c is a map from $g(\mathbb{N}_0)$ to Σ. Then $h = \prod_{a \in c^{-1}(x)}(g - a)$ is a GP map and it follows from Lemma 5.6 that $[\![h(n) = 0]\!] = \{n \in \mathbb{N}_0 \mid a_n = x\}$ is a GP subset of $\mathbb{N}_0$.

Conversely, let us assume that, for every $x \in \Sigma$, the fibre $F_x = \{n \in \mathbb{N}_0 \mid a_n = x\}$ is a GP subset of $\mathbb{N}_0$. Hence $\mathbf{1}_{F_x}$ is a GP map. Let us assume that $|\Sigma| = N$ and let $x_1, \ldots, x_N$ denote an enumeration of the elements of Σ. Set $h = \sum_{i=1}^{N} i \mathbf{1}_{F_{x_i}}$ and let c be the map defined by $c(i) = x_i$, $1 \leq i \leq N$. Then h is a GP map and thus $a_n = c(h(n))$ is a bracket word. $\square$

With some basic algebraic manipulations, one can extend Proposition 5.4 to apparently more complicated conditions, as shown by Example 5.8.

Example 5.8. Let g and h be GP maps from $\mathbb{N}_0$ to $\mathbb{R}$ and assume that $h(n) > 0$ for all $n \in \mathbb{N}_0$. Then $[\![g < 1/h]\!] = [\![gh < 1]\!]$ is a bracket word. More generally, for each rational exponent $\lambda = p/q \in \mathbb{Q}_{>0}$, also $[\![g < 1/h^{\lambda}]\!] = [\![g^q h^p < 1]\!]$ is a bracket word. As an explicit application, for every pair $(\alpha, c) \in \mathbb{R}^2$ and $\lambda \in \mathbb{Q}$, with c and λ positive, the formula $a_n = [\![\|\alpha n\| < c/n^{\lambda}]\!]$ defines a bracket word $\mathbf{a}$ which detects denominators of good rational approximations to α.

6. Closure properties

We will now discuss ways in which known instances of bracket words can be used to construct new ones. Compared to the earlier section, the results discussed here have a more computational flavour. For instance, we point out that, with the sole exception of Proposition 6.14, all results in this section are analogues of standard results about automatic sequences, in the sense that they remain true if the term "bracket word" is replaced with "k-automatic sequence" and the term "GP set" is replaced with "k-automatic set" for fixed $k \geq 2$ (see [AS03, Sec. 5]).

6.1. **Codings and products.** It is an almost immediate consequence of Definition 2.2 that bracket words are preserved under coding.

Lemma 6.1. *Let $\mathbf{a}$ be a bracket word defined over a finite alphabet Σ and let $\varphi \colon \Sigma \to \Pi$ be a map to some other finite alphabet Π. Then $(\varphi(a_n))_{n=0}^{\infty}$ is a bracket word over Π.*

Proof. If $a_n = c(g(n))$ is the representation of $\mathbf{a}$ as in Definition 2.2, then $\varphi(a_n) = c'(g(n))$ where $c' = \varphi \circ c$. $\square$

Next, we note that the direct product of two bracket words is again a bracket word.

Proposition 6.2. *Let $\mathbf{a}$ and $\mathbf{a}'$ be bracket words respectively defined over some finite alphabets Σ and Σ'. Then $\mathbf{a} \times \mathbf{a}' = ((a_n, a_n'))_{n=0}^{\infty}$ is a bracket word over $\Sigma \times \Sigma'$.*

Proof. Recall that, by definition of a bracket word, $\mathbf{a}$ has a representation $a_n = c(g(n))$, where $g \colon \mathbb{N}_0 \to A$ is a GP map taking values in some finite set $A \subseteq \mathbb{R}$, and $c \colon A \to \Sigma$ is an arbitrary map. Let $a_n' = c'(g'(n))$ be an analogous representation of $\mathbf{a}'$.

Replacing g with Cg for a sufficiently large positive real number $C > 0$ (and modifying A and c accordingly), we may assume that the only solutions to $x + x' = y + y'$ with $(x, y) \in A^2$ and $(x', y)' \in A'^2$ are the trivial ones: $x = x'$ and $y = y'$. Let B denote the sumset $A + A' = \{x + x' \mid x \in A, \ x' \in A'\}$, let h denote the GP map $g + g' \colon \mathbb{N}_0 \to B$, and let $d \colon B \to \Sigma \times \Sigma'$ denote the unique map such that $d(x + x') = (c(x), c'(x'))$ for all $x \in A$ and $x' \in A'$. Then $(a_n, a_n') = d(h(n))$ for all $n \in \mathbb{N}_0$. $\square$

Remark 6.3. It follows that replacing, in Definition 2.2, GP maps $\mathbb{N}_0 \to \mathbb{R}$ with GP maps $\mathbb{N}_0 \to \mathbb{R}^d$ for arbitrary $d \in \mathbb{N}$ leaves unchanged the set of words so defined.

In practice, Proposition 6.2 is mostly used via Corollary 6.4.

Corollary 6.4. *Let $\mathbf{a}$ and $\mathbf{a}'$ be bracket words respectively defined over some finite alphabets Σ and Σ', and let $f \colon \Sigma \times \Sigma' \to \Pi$ be a map to some other finite alphabet Π. Then $(f(a_n, a_n'))_{n=0}^{\infty}$ is a bracket word over Π.*

Proof. This follows directly from Proposition 6.2 and Lemma 6.1. $\square$

Remark 6.5. In particular, for any ring R, bracket words taking values in R, equipped with coordinatewise addition and multiplication, form a ring.

As a consequence of the two previous results, we see that bracket words can be defined in a "case-by-case" manner.

Proposition 6.6. *Let Σ be a finite alphabet, $\mathbb{N}_0 = \bigcup_{i=1}^{r} S_i$ be a partition of $\mathbb{N}_0$ into pairwise disjoint GP subsets, and $\mathbf{a}^{(i)}$, $1 \le i \le r$, be bracket words over Σ. Let $\mathbf{a}$ be defined by $a_n = a_n^{(i)}$ if $n \in S_i$. Then $\mathbf{a}$ is a bracket word.*

Proof. We first infer from a recursive use of Proposition 6.2 that the word

$$(\mathbf{1}_{S_1}, \ldots, \mathbf{1}_{S_r}, \mathbf{a}^{(1)}, \ldots, \mathbf{a}^{(r)})$$

defined over the alphabet $\{0,1\}^r \times \Sigma^r$ is a bracket word. Now, letting f be any function from $\{0,1\}^r \times \Sigma^r$ to Σ such that

$$f(\varepsilon_1, \ldots, \varepsilon_r, x_1, \ldots, x_r) = x_i \text{ if } \varepsilon_i = 1 \text{ and } \varepsilon_j = 0 \text{ when } j \ne i,$$

we infer from Corollary 6.4 that $f(\mathbf{1}_{S_1}, \ldots, \mathbf{1}_{S_r}, \mathbf{a}^{(1)}, \ldots, \mathbf{a}^{(r)})$ is a bracket word. It remains to see that $\mathbf{a} = f(\mathbf{1}_{S_1}, \ldots, \mathbf{1}_{S_r}, \mathbf{a}^{(1)}, \ldots, \mathbf{a}^{(r)})$. $\square$

Example 6.7. Let $\mathbf{a}$ be the infinite word over the alphabet $\{-2, -1, 0, \ldots, 10\}$ given by

$$a_n = \begin{cases} \lfloor n\{\sqrt{2}n\} \rfloor & \text{if } n\{\sqrt{2}n\} \le 10, \\ -1 & \text{if } n\{\sqrt{2}n\} > 10 \text{ and } n^2\{\sqrt{2}n\lfloor\sqrt{3}n\rfloor\} - n\{\sqrt{5}n\} + 7 > 0, \\ -2 & \text{if } n\{\sqrt{2}n\} > 10 \text{ and } n^2\{\sqrt{2}n\lfloor\sqrt{3}n\rfloor\} - n\{\sqrt{5}n\} + 7 \le 0. \end{cases}$$

Then $\mathbf{a}$ is a bracket word.

Since each eventually constant sequence is a bracket word, it follows that bracket words are also closed under finite modifications.

Corollary 6.8. *Let $\mathbf{a}, \mathbf{a}'$ be infinite words over a finite alphabet. If $\mathbf{a}$ is a bracket word and $a'_n = a_n$ for all but finitely many $n \in \mathbb{N}_0$ then $\mathbf{a}'$ is a bracket word.*

Proof. The result follows directly from Proposition 6.6 and the fact that all finite subsets of $\mathbb{N}_0$ are GP. $\square$

6.2. **Rearrangements and morphisms.** Many natural operations on infinite words can be described in terms of extracting or inserting entries in a regular manner. We record a simple observation, which can be used to find examples of operations of the aforementioned type which preserve bracket words.

Lemma 6.9. *Let $\mathbf{a}$ be a bracket word and let $h \colon \mathbb{N}_0 \to \mathbb{N}_0$ be a GP map. Then $\left(a_{h(n)}\right)_{n=0}^{\infty}$ is a bracket word.*

Proof. If $a_n = c(g(n))$ is the representation of $\mathbf{a}$ as in Definition 2.2, then $a_{h(n)} = c(g'(n))$, where $g' = g \circ h$ is a GP map (see Remark 2.1). $\square$

Below, we list some applications of this result. Recall that, for $A \in \mathbb{N}$, the map $\mathbb{N}_0 \to \mathbb{N}_0$ given by $n \mapsto n \bmod A = A\{n/A\}$ is a GP map.

Corollary 6.10. *Let $\mathbf{a}$ be a bracket word defined over a finite alphabet Σ, let $A \in \mathbb{N}$, $B \in \mathbb{N}_0$, let $\Diamond$ be a symbol not belonging to Σ, and let $\pi \colon [A] \to [A]$ be a map. Then the following infinite words are also bracket words.*

   (i) $(a_{An+B})_{n=0}^{\infty}$.
   (ii) $\mathbf{a}'$ *defined over* $\Sigma \cup \{\Diamond\}$ *by* $a'_n = a_{n/A}$ *if* $A \mid n$ *and* $a'_n = \Diamond$ *otherwise.*
   (iii) $\left(a_{\lfloor n/A\rfloor + \pi(n \bmod A)}\right)_{n=0}^{\infty}$.

Proof. The first item follows directly from Lemma 6.9. The second and third items follow from Lemma 6.9 and Proposition 6.6. $\square$

Lemma 6.11. *Let Σ and Π be two alphabets. Let $\mathbf{a}$ be a bracket word over Σ and σ be a morphism of constant length from Σ^* to Π^*. Then $\sigma(\mathbf{a})$ is a bracket word.*

Proof. Let us assume that σ has constant length k, and set $\mathbf{a}' = \sigma(\mathbf{a})$. For each $i \in [k]$ and $n \in \mathbb{N}_0$, we have $a'_{kn+i} = \sigma(a_n)_i$, where we let $\sigma(a_n)_i$ denote the ith letter occurring in $\sigma(a_n)$. It remains to apply Proposition 6.6 and Lemmas 6.9 and 6.1. $\square$

By similar techniques, we can show a result in the reverse direction: bracket words are preserved under grouping blocks of constant length, or, in other words, if the image of a word $\mathbf{a}$ by an injective morphism of constant length is a bracket word, then $\mathbf{a}$ itself is also a bracket word.

Lemma 6.12. *Let $\mathbf{a}$ be an infinite word defined over a finite alphabet Σ and let $k \in \mathbb{N}$. Consider the infinite word $\mathbf{a}'$ over Σ^k given by $a'_n = a_{kn} a_{kn+1} \cdots a_{kn+k-1}$. Then $\mathbf{a}'$ is a bracket word if and only if $\mathbf{a}$ is a bracket word.*

Proof. The proof is similar to the one of Lemma 6.11. $\square$

6.3. **Orbit closure.** Given an infinite word $\mathbf{a}$ defined over a finite alphabet Σ, we let $\mathcal{O}(\mathbf{a})$ denote the orbit of $\mathbf{a}$ under the shift, that is, the set of all infinite words $\mathbf{a}'$ given by $a'_n = a_{n+m}$ for some $m \in \mathbb{N}_0$.

Proposition 6.13. *Let $\mathbf{a}$ be a bracket word and let $\mathbf{a}' \in \mathcal{O}(\mathbf{a})$. Then $\mathbf{a}'$ is a bracket word.*

Proof. This follows from (i) of Corollary 6.10 in the special case where $A = 1$. $\square$

More generally, we can consider the orbit closure $\overline{\mathcal{O}(\mathbf{a})}$, that is, the closure of $\mathcal{O}(\mathbf{a})$ with respect to the product topology on Σ^∞. Explicitly, $\mathbf{a}'$ belongs to $\overline{\mathcal{O}(\mathbf{a})}$ if, for every $N \in \mathbb{N}$, there exists m such that $a'_n = a_{n+m}$ for all $n \in [N]$.

Proposition 6.14. *Let $\mathbf{a}$ be a bracket word and let $\mathbf{a}' \in \overline{\mathcal{O}(\mathbf{a})}$. Then $\mathbf{a}'$ is a bracket word.*

Proof. This is a direct consequence of Lemmas A.5 and A.6 given in Appendix A. $\square$

Example 6.15. Let $\mathbf{a}$ be the bracket word given by $a_n = [\![\{ \sqrt{2}n \{ \sqrt{3}n \} \} < 1/4]\!]$. Then each $\mathbf{a}' \in \mathcal{O}(\mathbf{a})$ can be written in the form

$$a'_n = [\![\{ (\sqrt{2}n + \alpha) \{ \sqrt{3}n + \beta \} + \gamma n + \delta \} < 1/4]\!],$$

where $\alpha, \beta, \gamma, \delta \in [0, 1)$. Conversely, $\overline{\mathcal{O}(\mathbf{a})}$ contains all sequences $\mathbf{a}'$ of the aforementioned form, as well as all sequences obtained from them by replacing any instances of $\{x\}$ with $\{x\}' = 1 - \{-x\}$, or by replacing the strict inequality $<$ with $\leq$.

6.4. **GP sets.** The closure properties of bracket words directly translate into closure properties of GP subsets of $\mathbb{N}_0$ thanks to Corollary 5.7. In fact, we have already used this connection at several places. For ease of reference, we gather these properties here.

Proposition 6.16. *The family of GP subsets of $\mathbb{N}_0$ is a field of sets. In other words, it contains the empty set and if $E, F \subseteq \mathbb{N}_0$ are GP sets then so are $\mathbb{N}_0 \setminus E$, $E \cup F$, and $E \cap F$.*

Proof. The result follows from Proposition 6.6. □

Proposition 6.17. *Let* $E \subseteq \mathbb{N}_0$ *be a GP set. Then the following sets are also GP sets.*

- $m \cdot E = \{nm \mid n \in E\}$, *where* $m \in \mathbb{N}_0$.
- $E/m = \{n \in \mathbb{N}_0 \mid nm \in E\}$, *where* $m \in \mathbb{N}$.
- $E - m = \{n \in \mathbb{N}_0 \mid n + m \in E\}$, *where* $m \in \mathbb{Z}$.

Proof. This is a consequence of Corollary 6.10. □

7. Frequencies and recurrence

In this section, we collect some facts concerning frequencies with which symbols and subwords appear in bracket words, as well as additive structure of the set of positions at which a given symbol appears in a bracket word. Most of these facts are consequences of results concerning uniform distribution and recurrence of orbits of point in nilsystems, and hence are related to the dynamical characterisation of bracket words discussed in Section 4. We point out that, similarly, many standard properties of Sturmian words can be inferred from dynamical properties of circle rotations.

7.1. Uniform frequency. Let Σ be a finite alphabet. Given an infinite word $\mathbf{a} \in \Sigma^\infty$ and a letter $x \in \Sigma$, the (asymptotic) *frequency* of x in $\mathbf{a}$ is defined as

$$\mathrm{freq}(\mathbf{a}, x) = \lim_{N \to \infty} |\{n \in [N] \mid a_n = x\}| / N,$$

provided that the limit exists. More generally, given a finite word $w \in \Sigma^\ell$, the frequency of w in $\mathbf{a}$ is defined as

$$\mathrm{freq}(\mathbf{a}, w) = \lim_{N \to \infty} \left|\left\{n \in [N] \mid \mathbf{a}_{[n,n+\ell)} = w\right\}\right| / N,$$

provided that the limit exists. Here and elsewhere, $\mathbf{a}_{[n,n+\ell)}$ denotes the finite subword $a_n a_{n+1} \cdots a_{n+\ell-1}$. We define also the *recurrence function* of w in $\mathbf{a}$ as the least length of a segment of $\mathbf{a}$ that is guaranteed to contain an occurrence of w:

$$\mathrm{rec}(\mathbf{a}, w) = \inf \left\{ r \in \mathbb{N} \ \middle| \ \begin{array}{l} \text{for each } m \text{ there exists } n \in [m, m+r) \\ \text{such that } \mathbf{a}_{[n,n+\ell)} = w \end{array} \right\} \in \mathbb{N} \cup \{\infty\}.$$

For bracket words, we have a strong result about the existence of frequencies, which is a rephrasing of [BL07, Thm. B].

Theorem 7.1. *Let* $\mathbf{a}$ *be a bracket word defined over a finite alphabet* Σ *and* ℓ *be a positive integer. Then for each* $w \in \Sigma^\ell$, *the frequency* $\mathrm{freq}(\mathbf{a}, w)$ *exists and, moreover, one has*

$$\left|\left\{n \in [M, M+N) \mid \mathbf{a}_{[n,n+\ell)} = w\right\}\right| / N \to \mathrm{freq}(\mathbf{a}, w)$$

uniformly in M *as* $N \to \infty$.

Proof. By definition, there exist a finitely valued GP map $g : \mathbb{N}_0 \to \mathbb{R}$ and a coding $c : g(\mathbb{N}_0) \to \Sigma$ such that $a_n = c(g(n))$. We deduce from [BL07, Thm. B] that the desired conclusion holds for the sequence $g(n)$. The same conclusion remains true after applying a coding. □

Corollary 7.2. *For* $\mathbf{a}$ *and* w *as above, either* $\mathrm{freq}(\mathbf{a}, w) = 0$ *or* $\mathrm{rec}(\mathbf{a}, w) < \infty$.

Proof. If freq($\mathbf{a}, w$) = 0, there is nothing to prove. Otherwise, freq($\mathbf{a}, w$) > 0 and uniformity of convergence in Theorem 7.1 implies that rec($\mathbf{a}, w$) < ∞. □

Remark 7.3. It can happen that freq($\mathbf{a}, w$) = 0 but still w appears in $\mathbf{a}$ infinitely often, see e.g. Example 3.5. Hence, bracket words are not guaranteed to be uniformly recurrent.[2]

The same result can be stated in terms of densities of GP sets. Recall that the *asymptotic density* of a set $E \subseteq \mathbb{N}_0$ is defined by

$$d(E) = \lim_{N \to \infty} |E \cap [N]| / N,$$

provided that the limit exists. In general, we let $\overline{d}(E)$ and $\underline{d}(E)$ denote the upper and lower asymptotic densities, obtained by replacing lim with lim sup and lim inf respectively. Additionally, we define the *upper and lower uniform (Banach) densities* by

$$d^*(E) = \limsup_{N \to \infty} \sup_M |E \cap [M, M+N)| / N$$

and

$$d_*(E) = \liminf_{N \to \infty} \inf_M |E \cap [M, M+N)| / N.$$

In general, we have the chain of inequalities: $d_*(E) \leq \underline{d}(E) \leq \overline{d}(E) \leq d^*(E)$.

Theorem 7.4. *Let E be a GP subset of $\mathbb{N}_0$. Then $d(E)$ exists. Moreover, $d^*(E) = d_*(E) = d(E)$.*

7.2. **Distribution along subsequences.** Many of the desirable properties of bracket words are preserved under passing to subsequences. As a first instance of this principle, we consider the simple case of polynomial subsequences.

Proposition 7.5. *Let $\mathbf{a}$ be a bracket word defined over a finite alphabet Σ, $p : \mathbb{N}_0 \to \mathbb{N}_0$ be a polynomial, and ℓ be a positive integer. Then for each $w \in \Sigma^\ell$, the frequency* freq$\left(\left(a_{p(n)}\right)_{n=0}^\infty, w\right)$ *exists.*

Proof. Let $a_n = c(g(n))$ be a representation of $\mathbf{a}$, where $g : \mathbb{N}_0 \to \mathbb{R}$ is a GP map. Then $g \cdot p : \mathbb{N}_0 \to \mathbb{N}_0$ is also a GP map for GP maps are closed under composition, and $a_{p(n)} = c(g \circ p(n))$. The result follows now from Theorem 7.1. □

We point out that even in the simplest cases, passing to a subsequence can alter frequencies with which symbols occur. For instance, if $\mathbf{a} = 101010 \cdots$ then freq($\mathbf{a}, 1$) = 1/2 but freq($(a_{2n})_{n=0}^\infty, 1$) = 1.

Next, let us consider distribution along the primes. In [GT12a], Green and Tao obtained quantitative estimates on correlations between the Möbius function and nilsequences, which have important consequences for the question at hand (cf. [GT12a, Sec. 5]). Extending the techniques developed by Green and Tao, and the connection between GP maps and nilsystems, Bergelson, Håland Knutson and Son [BHKS20] showed that bounded GP maps have asymptotic distribution along the primes. The following result is obtained by specialising [BHKS20, Theorem 5.1] to finitely-valued sequences. Below, we let p_n denote the n-th prime. We also mention related work of Eisner [Eis20] concerning convergence of ergodic averages along primes in nilsystems. In the case of Sturmian words, this result follows from work of Davenport [Dav37].

---

[2]A sequence $\mathbf{a}$ is uniformly recurrent if, for every finite word w which appears in $\mathbf{a}$, there exists $\ell \in \mathbb{N}$ such that w appears in the length-ℓ segment $\mathbf{a}_{[n,n+\ell)}$ for all n.

Theorem 7.6. *Let* **a** *be a bracket word defined over a finite alphabet* Σ. *Then for each* $x \in \Sigma$, *the frequency* $\mathrm{freq}\left((a_{p_n})_{n=0}^{\infty}, x\right)$ *exists.*

Proof. This follows from [BHKS20, Thm. 5.1] using the same argument as in the proof of Theorem 7.1. $\qquad\square$

Considering $\mathbf{a} = 101010\cdots$ again, we see that the frequencies $\mathrm{freq}\left((a_{p_n})_{n=0}^{\infty}, x\right)$ need not be equal to $\mathrm{freq}\,(\mathbf{a}, x)$. One can also inquire if it is possible to generalise Theorem 7.6 from frequencies of symbols $x \in \Sigma$ to frequencies of words $w \in \Sigma^{\ell}$, $\ell \geq 2$. While it seems plausible that this generalisation is true, it is out of reach of the current techniques. Indeed, already in the case where $\mathbf{a} = (n \bmod q)_{n=0}^{\infty}$ is a periodic word with period $q \geq 3$, such a generalisation would require us to understand the asymptotic behaviour of

$$(9) \qquad \frac{1}{N}\left|\{n < N \mid p_n \equiv w_1, p_{n+1} \equiv w_2, \ldots, p_{n+\ell-1} = w_{n+\ell-1} \bmod q\}\right|.$$

It is conjectured that the expression in (9) converges to $1/\varphi(q)^{\ell}$, where φ denotes the totient function; in fact, the Main Conjecture in [LOS16] gives a more precise asymptotic expression. However, such estimate remains unknown for any $\ell \geq 2$, as discussed in [LOS16].

Lastly, we consider a class of subsequences which preserve the frequencies of symbols. As already alluded to earlier, Theorem 4.1 leads to a close connection between equidistribution results for nilsystems and statements about frequencies of symbols in bracket words. Recall that each nilmanifold X comes equipped with the Haar measure μ_X. A sequence $(x_n)_{n=0}^{\infty}$ is *equidistributed* in X if for each continuous map $f \colon X \to \mathbb{R}$ we have

$$(10) \qquad \lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} f(x_n) = \int_X f\,d\mu_X.$$

Definition 7.7. *Let* $t \colon \mathbb{N}_0 \to \mathbb{N}_0$. *We say that* t *is good for equidistribution in nilsystems if, for all minimal nilsystems* (X, T) *and all* $x \in X$, *the sequence* $(T^{t(n)}(x))_{n=0}^{\infty}$ *is equidistributed in* X.

Proposition 7.8. *Let* **a** *be a bracket word defined over a finite alphabet* Σ, *and let* $t \colon \mathbb{N}_0 \to \mathbb{N}_0$ *be good for equidistribution in nilsystems. Then the frequencies of letters in* **a** *along the subsequence* $t(n)$ *exist and remain the same as in* **a***:*

$$\mathrm{freq}\left((a_{t(n)})_{n=0}^{\infty}, x\right) = \mathrm{freq}(\mathbf{a}, x) \quad \forall x \in \Sigma.$$

Proof. This is a consequence of the representation of **a** coming from Theorem 4.2. Indeed, though the indicator functions $\mathbf{1}_{S_i}$ of the semialgebraic sets S_i from Theorem 4.2 are not continuous, they can be efficiently approximated by continuous functions to which Equality (10) can be applied. $\qquad\square$

Hopefully, the following result of Frantzikinakis [Fra09] provides a plentiful source of sequences that satisfy Definition 7.7 and make Proposition 7.8 relevant.

Theorem 7.9 ([Fra09, Thm. 1.1])**.** *Let us assume that* $f \colon \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ *satisfies the following properties.*

  (i) *The function* f *belongs to some Hardy field.*
  (ii) *There exists* $C > 0$ *such that* $f(x) = O(x^C)$ *as* $x \to \infty$.
  (iii) *For every* $p \in \mathbb{Z}[x]$ *and* $c \in \mathbb{R}$, *we have* $\lim_{x\to\infty} \frac{|f(x)-cp(x)|}{\log x} = \infty$.

Then $\lfloor f \rfloor$ is good for equidistribution in nilsystems.

For a definition of a Hardy field, we refer the reader to [Fra09]. Here, we just point out that one example of a Hardy field is the logarithmic-exponential functions, that is, real-valued functions defined on some interval $[c, \infty)$, $c \in \mathbb{R}$, that can be constructed from polynomials using addition, multiplication, and the functions exp and log. Hence, the following formulae define sequences that are good for equidistribution in nilsystems: $\lfloor n^{3/2} \rfloor$, $\lfloor n^2 \log^2 n \rfloor$, $\lfloor n^2 + \log^2 n \rfloor$, and $\lfloor n^2 + \sqrt{2}n \rfloor$. Conversely, because of item (iii), there is no guarantee that the same should apply to $\lfloor n^2 + \log n \rfloor$ or $\lfloor \sqrt{2}n \rfloor$. In fact, one can compute that

$$\frac{1}{\sqrt{2}} \left\lfloor \sqrt{2}n \right\rfloor = n + \frac{\{\sqrt{2}n\}}{\sqrt{2}} \equiv \frac{\{\sqrt{2}n\}}{\sqrt{2}} \bmod 1,$$

so $\lfloor \sqrt{2}n \rfloor$ is not good for equidistribution for the circle rotation by $1/\sqrt{2}$. Extensions of some results in [Fra09] were recently obtained by Richter [Ric20].

It is not possible to directly generalise Proposition 7.8 to frequencies of words instead of symbols. For instance, the sequence $t(n) = \lfloor \sqrt{n} \rfloor$ is good for equidistribution in nilsystems and constant on each interval $[N^2, (N+1)^2)$, $N \in \mathbb{N}$. Hence, already for the periodic word $\mathbf{a} = 101010 \cdots$ and $w \in \{0,1\}^2$ we have

$$\text{freq}\left((a_{t(n)})_{n=0}^{\infty}, w\right) = \begin{cases} 1/2 & \text{if } w \in \{00, 11\}, \\ 0 & \text{if } w \in \{01, 10\}. \end{cases}$$

Thus, the frequencies of words in $(a_{t(n)})_{n=0}^{\infty}$ bear little resemblance to the frequencies of symbols and words in $\mathbf{a}$.

On the other hand, under additional growth conditions, one can obtain positive results. The key technical component is Theorem 7.10 of Bergelson, Moreira and Richter. Below, we let $f^{(j)}$ denote the j-th derivative of f.

Theorem 7.10 ([BMR20, Thm. 5.6, special case]). *Let (X, T) be a minimal nilsystem, let $z \in X$, and let $f_1, f_2, \ldots, f_k$ be functions belonging to the same Hardy field. Suppose further that the following two conditions hold.*

(i) *For each $h \in \text{span}\left\{f_i^{(j)} \;\middle|\; 1 \le i \le k, j \ge 0\right\}$ and each $p(x) \in \mathbb{R}[x]$, we have either $|h(x) - p(x)| = O(1)$ or $|h(x) - p(x)| / \log x \to \infty$ as $x \to \infty$.*

(ii) *For each non-zero $h \in \text{span}\{f_i \mid 1 \le i \le k\}$ and each $p(x) \in \mathbb{R}[x]$, we have $|h(x) - p(x)| \to \infty$ as $x \to \infty$.*

Then $\left(\left(T^{\lfloor f_1(n) \rfloor}(z), T^{\lfloor f_2(n) \rfloor}(z), \ldots, T^{\lfloor f_k(n) \rfloor}(z)\right)\right)_{n=0}^{\infty}$ is equidistributed in X^k.

Let now f be a function belonging to a Hardy field, and put $t = \lfloor f \rfloor$ and $f_i(x) = f(x + i)$ for every i such that $1 \le i \le k$. Using standard techniques, such as the Taylor expansion and estimates on derivatives of functions in Hardy fields [Fra09, Lem. 2.1], one can verify that conditions 7.10.(i) and 7.10.(ii) are satisfied if

$$\frac{f(x)}{x^k \log x} \to \infty \qquad \text{and} \qquad \frac{f(x)}{x^{k+1}} \to 0 \qquad \text{as} \qquad x \to \infty.$$

Combining the conclusion of Theorem 7.10 with Theorem 4.2, one could show that for each bracket word $\mathbf{a}$ and each word $w \in \Sigma^k$ we have

$$(11) \qquad \text{freq}\left((a_{t(n)})_{n=0}^{\infty}, w\right) = \prod_{i=0}^{k-1} \text{freq}(\mathbf{a}, w_i).$$

We emphasise that the frequencies in (11) are, in general, not equal to $\mathrm{freq}(\mathbf{a}, w)$; rather, they are the values which one would expect if t exhibited random-like behaviour. We leave the details to the interested reader.

7.3. **IP recurrence.** Apart from asymptotic density, one can also inquire into combinatorial richness of GP subsets of $\mathbb{N}$. In order to state the relevant results, we introduce some terminology. A set $E \subseteq \mathbb{N}$ is an IP-*set* if it contains all finite sums of a sequence of positive integers, that is, if there exists a sequence of positive integers $(n_i)_{i=1}^{\infty}$ such that $\sum_{i \in I} n_i \in E$ for all finite set $I \subseteq \mathbb{N}$. A set $F \subseteq \mathbb{N}$ is IP* if $E \cap F \neq \emptyset$ for every IP-set E, or, equivalently, if for every sequence of positive integers $(n_i)_{i=1}^{\infty}$ there exists a finite set $I \subseteq \mathbb{N}$ with $\sum_{i \in I} n_i \in F$. We also define shifted variants of these notions. A set $E \subseteq \mathbb{N}$ is IP$_+$ if $(E + n) \cap \mathbb{N}$ is an IP-set for some $n \in \mathbb{Z}$. Likewise, a set $F \subseteq \mathbb{N}$ is IP$_+^*$ if $(F + n) \cap \mathbb{N}$ is an IP*-set for some $n \in \mathbb{Z}$. Each IP*-set is *syndetic*, meaning that for every IP*-set F there exists an integer N such that $F \cap [n, n + N) \neq \emptyset$ for all $n \in \mathbb{N}_0$. In particular, all IP*-sets have positive lower uniform density. Moreover, since for all $m \in \mathbb{N}$, the multiples of m form an IP-set, all IP*-sets contain a multiple of m. The class of IP*-sets is closed under intersection, meaning that for all pairs $F, F' \subseteq \mathbb{N}$ of IP*-sets, $F \cap F'$ is also an IP*-set. In what follows, we say that a statement $\varphi(n)$ holds for *almost all* $n \in \mathbb{N}_0$ if $d\left(\{n \in \mathbb{N}_0 \mid \neg\varphi(n)\}\right) = 0$.

Theorem 7.11 ([BL07, Thm.D]). *Let $\mathbf{a}$ be a bracket word over Σ. Then for almost all $n_0 \in \mathbb{N}_0$ and all sequences of positive integers $(n_i)_{i=1}^{\infty}$, there exists a finite set $I \subseteq \mathbb{N}$ such that $a_{n_0 + \sum_{i \in I} n_i} = a_{n_0}$. In other words, if $E \subseteq \mathbb{N}$ is a GP set with $d(E) > 0$, then E is IP$_+^*$. Furthermore, $(E - n) \cap \mathbb{N}$ is IP* for almost all $n \in E$.*

Remark 7.12. For Sturmian words, a much simpler proof is possible, see e.g. a special case of [Ber10, Cor. 7.3].

Remark 7.13. One can obtain more precise versions of this result. Specifically, it follows from [BL18] that one can additionally require that $I \subseteq [r]$ for some integer r that only depends on $\mathbf{a}$ and n_0. In a different direction, it follows from [Kon17] that one can additionally require that the gaps between consecutive elements of I are bounded from above by some integer d that only depends on $\mathbf{a}$.

8. Counting function and discrepancy

Let $\mathbf{a}$ be a bracket word defined over a finite alphabet Σ. In addition to the asymptotic frequency $\mathrm{freq}(\mathbf{a}, x)$ with which a symbol $x \in \Sigma$ appears in $\mathbf{a}$, one can inquire into more quantitative estimates on the count of occurrences of x in $\mathbf{a}$. This leads us to consider the counting function

$$(12) \qquad \pi(\mathbf{a}, x; N) = |\{n \in [N] \mid a_n = x\}|,$$

as well as the closely related notion of *symbolic discrepancy* considered in [Ada04],

$$(13) \qquad \Delta(\mathbf{a}; N) = \max_{x \in \Sigma} \left|\pi(\mathbf{a}, x; N) - N \, \mathrm{freq}(\mathbf{a}, x)\right|.$$

For example, if $\mathbf{a}$ is the Sturmian word given by $a_n = \lfloor \alpha n + \beta \rfloor - \lfloor \alpha(n-1) + \beta \rfloor$, then we have the estimates

$$(14) \qquad \pi(\mathbf{a}, 1; N) = \alpha N + O(1) \qquad \text{and} \qquad \Delta(\mathbf{a}, N) = O(1),$$

which are significantly stronger than $\mathrm{freq}(\mathbf{a}, 1) = \alpha$.

In this section, our aim is to study the possible rates of growth of $\pi(\mathbf{a}, x; N)$ and $\Delta(\mathbf{a}; N)$ when $\mathbf{a}$ is a bracket word. We note that, as a direct consequence of the definitions, for each $x \in \Sigma$ and $N \in \mathbb{N}$ we have

$$0 \leq \pi(\mathbf{a}, x; N), \ \Delta(\mathbf{a}; N) \leq N.$$

Since frequencies of all symbols exist, we also have

$$\pi(\mathbf{a}, x; N)/N \to \mathrm{freq}(\mathbf{a}, x) \text{ and } \Delta(\mathbf{a}; N)/N \to 0 \text{ as } N \to \infty.$$

This quantitative approach is especially relevant in the case where $\mathrm{freq}(\mathbf{a}, x) = 0$. We point out that if there exists $z \in \Sigma$ with $\mathrm{freq}(\mathbf{a}, z) = 1$ then

$$\Delta(\mathbf{a}; N) = N - \pi(\mathbf{a}, z; N) = \sum_{x \in \Sigma \setminus \{z\}} \pi(\mathbf{a}, x; N).$$

For instance, if we let $F = \{0, 1, 2, 3, 5, 8, \ldots\}$ denote the set of Fibonacci numbers, $\varphi = (1 + \sqrt{5})/2$ be the golden ratio, and $E = \{n \in \mathbb{N}_0 \mid \|\varphi n\| < 1/\sqrt{n}\}$, then the frequencies $\mathrm{freq}(\mathbf{1}_E, 1)$ and $\mathrm{freq}(\mathbf{1}_F, 1)$ are both zero, but

$$\pi(\mathbf{1}_F, 1; N) = \Delta(\mathbf{1}_F; N) = \Theta(\log N)$$

grows much more slowly (cf. Proposition 8.10) than

$$\pi(\mathbf{1}_E, 1; N) = \Delta(\mathbf{1}_E; N) = \Theta(\sqrt{N}).$$

For context, we mention that similar questions have been studied for other classes of sequences of combinatorial interest. For instance, it is well-known (see e.g. [Rig14, Sec. 1.2]) that for each k-automatic word $\mathbf{a}$ over a finite alphabet Σ and each $x \in \Sigma$, the count of integers n with $a_n = x$ whose base-k expansion has length $L \geq 1$, $\pi(\mathbf{a}, x; k^L) - \pi(\mathbf{a}, x; k^{L-1})$, satisfies a linear recurrence in L. Thus, either $\pi(\mathbf{a}, x; N) = \Theta(\log^d N)$ for some $d \in \mathbb{N}_0$ or $\pi(\mathbf{a}, x; N) = \Omega(N^c)$ for some $c \in (0, 1]$ [Rig14, Thm. 1.48]. In particular, if $\pi(\mathbf{a}, x; N) = N^{o(1)}$ then $\pi(\mathbf{a}, x; N) = \log^{O(1)} N$. Another convenient consequence is that if $\mathrm{freq}(\mathbf{a}, x) = 0$ then $\pi(\mathbf{a}, x; N) = O(N^c)$ for some $c \in (0, 1)$.

8.1. **Linear growth and low discrepancy.** As we have already observed, the rate of growth of $\pi(\mathbf{a}, x; N)$ can be linear and $\Delta(\mathbf{a}; N)$ can be bounded. In this context, it is helpful to recall the notion of a C-balanced word. An infinite word $\mathbf{a}$ over an alphabet Σ is C-*balanced* if the number of occurrences of a given symbol in any two subwords of $\mathbf{a}$ of equal length differs by at most C. For instance, as we have already pointed out in Section 1, Sturmian words are 1-balanced. It was shown in [Ada03, Prop. 7] that an infinite word $\mathbf{a}$ is C-balanced for a finite value $C > 0$ if and only the discrepancy $\Delta(\mathbf{a}; N)$ is bounded as $N \to \infty$.

Lemma 8.1. *For every* $\alpha \in (0, 1]$, *there exists a bracket word over* $\{0, 1\}$ *with* $\pi(\mathbf{a}, 1; N) = \alpha N + O(1)$, *and thus* $\mathrm{freq}(\mathbf{a}, 1) = \alpha$ *and* $\Delta(\mathbf{a}; N) = O(1)$.

Proof. Follows immediately from (14). $\qquad\qquad\square$

The example above stems from the fact that $[0, \alpha)$ is a bounded remainder set for the rotation by α on the torus $\mathbb{R}/\mathbb{Z}$. In general, given $d \in \mathbb{N}$ and $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d)$ a measurable set $S \subseteq \mathbb{R}^d/\mathbb{Z}^d$ is a *bounded remainder set* for the rotation by α on $\mathbb{R}^d/\mathbb{Z}^d$ if for almost all $x \in \mathbb{R}^d/\mathbb{Z}^d$, the *discrepancy*

$$\Delta(S, x; N) = \left| \sum_{n=0}^{N-1} 1_S(x + n\alpha) - \lambda(S)N \right|$$

is bounded by a constant $C(S, \alpha)$ independent of x and N. (Above, λ denotes the Lebesgue measure.) A complete classification of Riemann measurable bounded remainder sets for rotations on tori was obtained in [GL15]. In particular, by [GL15, Thm. 1], if $v_1, v_2, \ldots, v_d \in \mathbb{Z}\alpha + \mathbb{Z}^d$ are linearly independent vectors such that the parallelepiped

$$P = \left\{ \sum_{i=1}^{d} t_i v_i \;\middle|\; t_1, t_2, \ldots, t_d \in [0, 1) \right\}$$

is contained in $[0, 1)^d$ then P (or, strictly speaking, its projection to $\mathbb{R}^d/\mathbb{Z}^d$) is a bounded remainder set.

Example 8.2. Let $d \in \mathbb{N}$, let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d) \in \mathbb{R}^d$ be a d-tuple of real numbers such that $1, \alpha_1, \alpha_2, \ldots, \alpha_d$ are linearly independent over $\mathbb{Q}$, and let $S \subseteq \mathbb{R}^d/\mathbb{Z}^d$ be a bounded remainder set for the rotation by α on $\mathbb{R}^d/\mathbb{Z}^d$ which is also a semialgebraic set. Then $\mathbf{a} = (1_S(n\alpha))_{n=0}^{\infty}$ is a bracket word with $\operatorname{freq}(\mathbf{a}, 1) = \lambda(S)$ and $\Delta(\mathbf{a}; N) = O(1)$ as $N \to \infty$.

A different way to generalise Lemma 8.1 is to consider iterated discrete derivatives of polynomials of higher degrees.

Example 8.3. Let $\alpha \in (0, 1)$ and let $\mathbf{a}$ be the bracket word given by

$$a_n = \lfloor \alpha(n+2)^2 \rfloor - 2 \lfloor \alpha(n+1)^2 \rfloor + \lfloor \alpha n^2 \rfloor + 1$$
$$= -\{\alpha(n+2)^2\} + 2\{\alpha(n+1)^2\} - \{\alpha n^2\} + 1 + 2\alpha.$$

It follows directly from the formula above that $a_n \in \mathbb{Z}$ and $-1 < a_n < 5$ for all $n \in \mathbb{N}_0$, so $\mathbf{a}$ is a word over the alphabet $\{0, 1, 2, 3, 4\}$. Consider the word $\mathbf{a}'$ over the alphabet $\{0, 1\}$ given by

$$a'_{4n+i} = \begin{cases} 1 & \text{if } a_n > i, \\ 0 & \text{otherwise,} \end{cases} \qquad \text{for } n \in \mathbb{N}_0, \ i \in \{0, 1, 2, 3\}.$$

Then we can compute that

$$\pi(\mathbf{a}', 1; N) = \sum_{n=0}^{N-1} a'_n = \sum_{n=0}^{N/4-1} a_n + O(1)$$
$$= N/4 + \lfloor \alpha(N/4 + 1)^2 \rfloor - \lfloor \alpha(N/4)^2 \rfloor + O(1) = (1/4 + \alpha/8)N + O(1).$$

Hence, $\pi(\mathbf{a}', 1; N) = (1/4 + \alpha/8)N + O(1)$ and $\Delta(\mathbf{a}'; N) = O(1)$.

We recall the words in Examples 8.2 and 8.3 are C-balanced for some finite constant C. It follows from Theorem 4.2 that there exists a nilsystem (X, T), a point $x \in X$ and a semialgebraic set $S \subseteq X$ such that for the word $\mathbf{a}'$ from Example 8.3 we have $a'_n = 1$ if and only if $T^n(x) \in S$. Thus, the discrepancy $\left| \sum_{n=0}^{N-1} 1_S(T^n(x)) - \mu_X(S)N \right|$ is bounded as $N \to \infty$. In fact, we can take X to be the 2-step nilmanifold which is the product of the Heisenberg nilmanifold discussed in Section 4.3 and a circle $\mathbb{R}/\mathbb{Z}$.

8.2. **Slow growth and small discrepancy.** At the other extreme in terms of the possible rates of growth of $\pi(\mathbf{a}, x; N)$, it is possible for $\pi(\mathbf{a}, x; N)$ to tend to ∞ arbitrarily slowly. In the case where $|\Sigma| = 2$, $\Delta(\mathbf{a}; N) = \pi(\mathbf{a}, x; N)$ tends to ∞ at the same rate. As a first example, we mention a result [BK18, Theorem C] asserting that any "sufficiently sparse" subset of $\mathbb{N}_0$ is a GP set.

Proposition 8.4 ([BK18, Thm. C]). *Let $E = \{n_i \mid i \in \mathbb{N}_0\} \subseteq \mathbb{N}_{\geq 2}$ be a set with*

$$\liminf_{i \to \infty} \frac{\log n_{i+1}}{\log n_i} > 1.$$

Then $\mathbf{1}_E$ is a bracket word.

The proof of Proposition 8.4 relies on an application of Proposition 5.4 for a GP map of the form $g(n) = [\![\|\alpha n\| < 1/n^d]\!]$ with large $d \in \mathbb{N}$ and a carefully constructed $\alpha \in \mathbb{R}$. Specifically, the authors of [BK18] take α to be an irrational element of the intersection of the sets $\bigcap_{i=0}^{\infty} \{\beta \in \mathbb{R}/\mathbb{Z} \mid \|\beta n_i\| < n_i^d\}$, which is uncountable under mild additional assumptions on the rate of growth of $(n_i)_{i=0}^{\infty}$; we refer to [BK18] for details. Of course, this result provides examples of bracket words over $\{0, 1\}$ for which the growth of $\pi(\mathbf{a}, 1; N)$ tends to infinity as slowly as wanted.

Now, we prove the following general result that also covers growth order of type $\log N$.

Proposition 8.5. *Let $f \colon \mathbb{N}_0 \to \mathbb{R}_{>0}$ be a non-decreasing map and assume that there exists a positive real number C such that $f(2n) \leq f(n) + C$ for all $n \in \mathbb{N}$. Then there exists a bracket word $\mathbf{a}$ over $\{0, 1\}$ such that $\pi(\mathbf{a}, 1; N) = f(N) + O(1)$.*

Proof. Let F be the set of all Fibonacci numbers, and let $F' = F + [C]$. Note that for all sufficiently large $n \in \mathbb{N}$, we have $|F' \cap [n, 2n]| \geq C$. In [BK22], it is shown that any subset E of F is a GP set, and hence the same also applies to each $E \subseteq F'$. One can inductively construct a set $E \subseteq F'$ such that $|E \cap [N]| = f(N) + O(1)$ for all N: for each $n \in F'$, assuming that $E \cap [n]$ has already been constructed, we let $n \in E$ if and only if $f(n) > |E \cap [n]|$. It remains to set $\mathbf{a} = \mathbf{1}_E$. $\square$

Example 8.6. For every $\alpha \in \mathbb{R}_{>0}$, there is a bracket word $\mathbf{a}$ over $\{0, 1\}$ such that $\pi(\mathbf{a}, 1; N) = \alpha \log N + O(1)$.

For the sake of completeness, we mention that for even slower rates of growth, we can reformulate [BK18, Thm. C] to obtain a slightly more precise result which does not involve an error term.

Proposition 8.7. *Let $f \colon \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ be a continuously differentiable increasing function with $\sup_x |f'(x)| < \infty$. Then there exists a bracket word $\mathbf{a}$ over $\{0, 1\}$ with $\pi(\mathbf{a}, 1; N) = \lfloor f(\log \log N) \rfloor$ for all sufficiently large N.*

Proof. Since $f(\log \log x)$ is increasing in x and its derivative tends to 0 as $x \to \infty$, we can construct a set $E = \{m_1 < m_2 < \cdots\}$ with $|E \cap [N]| = \lfloor f(\log \log N) \rfloor$ for all sufficiently large N. Thus, for sufficiently large n we have

$$f(\log \log m_n) < n \leq f(\log \log(m_n + 1)).$$

As a consequence, we find that

$$f^{-1}(n - 1/2) \leq \log \log m_n \leq f^{-1}(n).$$

Applying this bound to n and $n + 1$, we obtain

$$\frac{\log m_{n+1}}{\log m_n} \geq \exp\left(f^{-1}(n + 1/2) - f^{-1}(n)\right).$$

Since $f'(x)$ is bounded as $x \to \infty$, we have

$$\liminf_{n \to \infty} \left(f^{-1}(n + 1/2) - f^{-1}(n)\right) > 0,$$

and consequently

$$\liminf_{n\to\infty} \frac{\log m_{n+1}}{\log m_n} > 1.$$

Thus, $\mathbf{1}_E$ is a bracket word by Proposition 8.4. □

Example 8.8. There are bracket words $\mathbf{a}$ over $\{0,1\}$ such that $\pi(\mathbf{a}, 1; N)$ is any of the following: $\lfloor \alpha \log\log N \rfloor$ for any $\alpha \in \mathbb{R}_{>0}$, $\lfloor (\log\log N)^c \rfloor$ for any $c \in [0,1)$, $\lfloor \log\log\log N \rfloor$, and so on.

8.3. **Moderate behaviours.** Bearing in mind the examples mentioned above, we are left with the question of determining which rates of growth between linear and logarithmic are possible for $\pi(\mathbf{a}, x; N)$. Towards this end, we will make use of results concerning the asymptotic behaviour of Diophantine expressions of the form $\prod_{i=1}^{d} \|\alpha_i n\|$, which have received considerable attention in connection with the Littlewood conjecture (see Example 3.6). While the conjecture remains unsolved, one can obtain considerably more precise estimates for a generic choice of $(\alpha_1, \alpha_2, \ldots, \alpha_d)$. For future reference, let us define the set

$$E(\lambda, \alpha) = \left\{ n \in \mathbb{N} \,\middle|\, \prod_{i=1}^{d} \|n\alpha_i\| < n^{-1+\lambda} \right\},$$

where $d \in \mathbb{N}$, $\lambda \in [0,1]$, and $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d) \in \mathbb{R}^d$. Note that $E(\lambda, \alpha)$ is a GP set if $\lambda \in \mathbb{Q}$.

Proposition 8.9. *For every $\lambda \in [0,1) \cap \mathbb{Q}$ and $c \in \mathbb{N}_0$, there exists a bracket word $\mathbf{a}$ over $\{0,1\}$ with $\pi(\mathbf{a}, 1; N) = \Theta(N^\lambda \log^c N)$.*

Proof. It follows from a variant of the main result in [WY81], as cited in [CT22, Thm. 1.6] and specialised to $\psi(n) = n^{-1+\lambda}$, that for each $d \in \mathbb{N}$ and almost all $\alpha \in \mathbb{R}^d$, we have the asymptotic formula

$$(15) \qquad |E(\lambda, \alpha) \cap [N]| = \begin{cases} \Theta(x^\lambda \log^{d-1} x) & \text{if } \lambda \neq 0, \\ \Theta(\log^d x) & \text{if } \lambda = 0. \end{cases}$$

If $\lambda = c = 0$ then one can simply take $\mathbf{a} = \mathbf{1}_{\{0\}}$, so suppose this is not the case. Let $d = c + 1$ if $\lambda \neq 0$ and $d = c$ if $\lambda = 0$. It follows from (15) that for almost all $\alpha \in \mathbb{R}^d$ we may take $\mathbf{a} = \mathbf{1}_{E(\lambda, \alpha)}$. □

Unfortunately, the main result in [WY81] does not provide any explicit example of $\alpha \in \mathbb{R}^d$ for which (15) holds, and as a consequence our proof of Proposition 8.9 does not provide any explicit example of bracket words $\mathbf{a}$ such that $\pi(\mathbf{a}, 1; N)$ has prescribed asymptotic behaviour. However, such examples can be constructed in the special case where (using notation from Proposition 8.9) $d = 0$. Below, we let φ denote the golden ratio.

Proposition 8.10. *Let $\lambda \in (0,1) \cap \mathbb{Q}$, and let $\mathbf{a} = \mathbf{1}_E$, where*

$$E = \left\{ n \in \mathbb{N} \,\middle|\, \|n\varphi\| \le n^{-1+\lambda} \right\}.$$

Then $\pi(\mathbf{a}, 1; N) = \Theta(N^\lambda)$.

Proof. Recall that every positive integer n has a unique expansion

$$n = \sum_{i=2}^{\infty} \epsilon_i(n) f_i,$$

where f_i is the i-th Fibonacci number, $\epsilon_i(n) \in \{0, 1\}$ and $(\epsilon_i(n), \epsilon_{i+1}(n)) \neq (1, 1)$ for all $i \geq 2$. Let $\nu(n)$ denote the least index $i \geq 1$ such that $\epsilon_{i+1}(n) = 1$, and let $\mu(n) = -\log_\varphi \|n\varphi\|$. It follows from [DMS18, Lemma 1] that the difference between $\nu(n)$ and $\mu(n)$ is bounded; in fact, $|\mu(n) - \nu(n)| \leq 3$ for all $n \in \mathbb{N}$.

One can show that for each $\ell \in \mathbb{N}$, the number of sequences in $\{0, 1\}^\ell$ with no pair of consecutive 1s is $f_{\ell+2} = \Theta(\varphi^\ell)$. Indeed, each sequence of this type starts with 0 or 10, yielding the same recurrence relation as the one satisfied by the Fibonacci numbers. As a consequence, for each sufficiently large $\ell \in \mathbb{N}$ and each $c \in (-10, 10)$, we have

$$(16) \qquad |\{n \in [f_{\ell-1}, f_\ell) \mid \nu(n) \geq (1 - \lambda)\ell + c\}| = \Theta\left(\varphi^{\lambda\ell}\right).$$

Since $|\mu(n) - \nu(n)| \leq 3$ and $\left|\log_\varphi n - \ell\right| < 2$ for all $n \in [f_{\ell-1}, f_\ell)$, it follows that

$$(17) \qquad \left|\left\{n \in [f_{\ell-1}, f_\ell) \mid \mu(n) \geq (1 - \lambda)\log_\varphi n\right\}\right| = \Theta\left(\varphi^{\lambda\ell}\right).$$

Let N be a large integer, and let i be the index with $f_{i-1} \leq N < f_i$. Summing (17) over all $\ell \leq i$ (resp. $\ell \leq i - 1$), we conclude that

$$|E \cap [N]| = \left|\left\{n \in [N] \mid \mu(n) \geq (1 - \lambda)\log_\varphi n\right\}\right| = \Theta\left(\varphi^{\lambda i}\right) = \Theta\left(N^\lambda\right). \qquad \square$$

In general, it is not clear how to construct $\alpha \in \mathbb{R}^d$ which satisfy (15) for $d \geq 2$. However, some progress in that direction follows from recent work of Chow and Technau. For instance, it follows as a special case of [CT22, Theorem 1.9] that for each $\lambda \in (0, 1)$ and each $\alpha_1 \in \mathbb{R}$ which is not a Liouville number (i.e., $\|n\alpha_1\| \gg 1/n^C$ for some constant C), for almost all α_2 we have the one-sided variant of (15),

$$|E(\lambda, (\alpha_1, \alpha_2)) \cap [N]| \gg N^\lambda \log N.$$

8.4. **Almost linear growth.** It might come as a surprise that there exist bracket words $\mathbf{a}$ over $\{0, 1\}$ such that the growth rate $\pi(\mathbf{a}, 1; N)$ is slower than, but arbitrarily close to, linear. This is closely related to a classical result of Khinchine asserting that there exist pairs $\alpha_1, \alpha_2 \in \mathbb{R}$, with $1, \alpha_1, \alpha_2$ linearly independent over $\mathbb{Q}$, such that $\min_{|n| < N} \|n_1\alpha_1 + n_2\alpha_2\|$ tends to 0 arbitrarily fast as $N \to \infty$, see e.g. [Cas72, Ch. V, Thm. XIV]. We will need the following elementary lemma.

Lemma 8.11. *Let $(N_i)_{i=1}^\infty$ be a sequence of positive integers and $(\varepsilon_i)_{i=1}^\infty$ be a sequence with values in $(0, 1)$ such that*

$$N_{i+1} \geq 2N_i/\varepsilon_i.$$

Then there exists $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ with $\|N_i\alpha\| \leq \varepsilon_i$ for all $i \in \mathbb{N}$.

Proof. For every $j \in \mathbb{N}$, we set

$$A_j = \{\alpha \in \mathbb{R}/\mathbb{Z} \mid \|N_i\alpha\| \leq \varepsilon_i \text{ for all } 1 \leq i \leq j\}.$$

It will suffice to show that the set $A = \bigcap_{j=1}^\infty A_j$ is uncountable.

Note that, for each $j \in \mathbb{N}$, A_{j+1} is the intersection of A_j with the periodic set $\{\alpha \in \mathbb{R}/\mathbb{Z} \mid \|N_{j+1}\alpha\| \leq \varepsilon_{j+1}\}$, which is the union of the intervals

$$I_{j+1}^{(m)} = \left[\frac{m - \varepsilon_{j+1}}{N_{j+1}}, \frac{m + \varepsilon_{j+1}}{N_{j+1}}\right], \qquad m \in [N_{j+1}].$$

This motivates us to recursively define a family of sets A_j' by $A_1' = A_1$ and

$$A_{j+1}' = \bigcup\left\{I_{j+1}^{(m)} \mid m \in [N_{j+1}] \text{ and } I_{j+1}^{(m)} \subseteq A_j'\right\}.$$

A routine inductive argument shows that $A'_j \subseteq A_j$ for each $j \in \mathbb{N}$, and it follows directly from the definition that A'_j is a union of intervals with lengths $2\varepsilon_j/N_j$. Set $A' = \bigcap_{j=1}^\infty A'_j \subseteq A$.

Since $2\varepsilon_j/N_i \geq 4/N_{j+1}$, for each $m \in [N_j]$ there exists $m' \in [N_{j+1}]$ such that

$$(18) \qquad I_{j+1}^{(m')}, I_{j+1}^{(m'+1)} \subseteq \left[\frac{m'-1}{N_{j+1}}, \frac{m'+2}{N_{j+1}}\right] \subseteq I_j^{(m)}.$$

Applying this observation inductively and using Cantor's intersection theorem, we conclude that $I_j^{(m)} \cap A' \neq \emptyset$ for each $j \in \mathbb{N}$ and $m \in [N_j]$ such that $I_j^{(m)} \subseteq A_j$. In fact, since the left hand side of (18) includes two disjoint intervals, we can use a similar argument to produce an injective map from $\{0,1\}^{\mathbb{N}}$ to A'. Thus, A' is uncountable, and so is A. $\qquad\square$

Proposition 8.12. *Let $f\colon \mathbb{N} \to (0,1]$ be any function with $f(N) \to 0$ as $N \to \infty$. There exists a bracket word $\mathbf{a}$ over $\{0,1\}$ such that $\mathrm{freq}(\mathbf{a}, 1) = 0$ but $\pi(\mathbf{a}, 1; N) = \Delta(\mathbf{a}; N) \geq f(N)N$ for all $N \in \mathbb{N}$.*

Proof. We may assume without loss of generality that $f(n) \geq 1/n$ for all $n \in \mathbb{N}$ and that f is non-increasing; if this is not the case, we can freely replace $f(n)$ with $\max(1/n, f(n), f(n+1), f(n+2), \ldots)$. For each N, let $h(N)$ denote the least positive integer with $f(h(N)) < 1/N$, and let $(N_i)_{i=1}^\infty)$ be the sequence defined by $N_1 = 1$ and

$$N_{i+1} = 2h(N_i)^2.$$

Put also $\varepsilon_i = N_i/h(N_{i+1})^2 = 2N_i/N_{i+2}$. Note that $h(N) > N$ so $N_{i+1} > 2N_i^2$, and in particular the sequence N_i is increasing.

By Lemma 8.11, there exist $\alpha_0, \alpha_1 \in \mathbb{R} \backslash \mathbb{Q}$ such that $\|\alpha_j N_i\| \leq \varepsilon_i$ for all $j \in \{0,1\}$ and $i \in \mathbb{N}$ with $i \equiv j \bmod 2$. For $j \in \{0,1\}$, set

$$E_j = \{n \in \mathbb{N} \mid \|n\alpha_j\| \leq 1/n\},$$

$E = E_0 \cup E_1 \cup [h(N_1)]$, and $\mathbf{a} = \mathbf{1}_E$. We claim that $\mathbf{a}$ satisfies the required conditions. Since α_0 and α_1 are irrational, we have

$$\mathrm{freq}(\mathbf{a}, 1) = d(E) \leq d(E_0) + d(E_1) = 0.$$

It remains to show that $\pi(\mathbf{a}, 1; N) = |E \cap [N]| \geq f(N)N$ for all N. If $N \leq h(N_1)$ then $[N] \subseteq E$, so we may assume that $N > h(N_1)$. Let i denote the unique index such that $h(N_i) \leq N < h(N_{i+1})$, and put $\alpha = \alpha_{i \bmod 2}$. Since $\|N_i\alpha\| \leq \varepsilon_i$, we have

$$\|mN_i\alpha\| \leq m\|N_i\alpha\| \leq m\varepsilon_i \leq \frac{1}{mN_i}$$

and thus $mN_i \in E$ for all integers m with $1 \leq m \leq h(N_{i+1})/N_i$. It follows that

$$|E \cap [N]| \geq N/N_i \geq f(h(N_i))N \geq f(N)N. \qquad\square$$

Example 8.13. There exists a bracket word $\mathbf{a}$ over $\{0,1\}$ with $\mathrm{freq}(\mathbf{a}, 1) = 0$ and $\pi(\mathbf{a}, 1; N) \geq N/\log\log\log N$ for all sufficiently large N. Note, however, that Proposition 8.12 does not ensure the existence of a bracket word $\mathbf{a}$ with $\pi(\mathbf{a}, 1; N) = \Theta(N/\log\log\log N)$.

8.5. **Algebraic coefficients.** The construction in Proposition 8.12 relies on finding a pair of real numbers (α_0, α_1) with some rather unusual Diophantine properties. If we restrict our attention to bracket words with algebraic coefficients (c.f. Remark 2.5), such constructions are no longer possible and some gap appears in the possible growth rates.

Proposition 8.14. *Let $\mathbf{a}$ be a bracket word defined over a finite alphabet Σ which arises from a GP map with algebraic coefficients and let $x \in \Sigma$. If $\mathrm{freq}(\mathbf{a}, x) = 0$, then there exists $c > 0$ such that $\pi(\mathbf{a}, x; N) = O(N^{1-c})$.*

Proof. In this proof, we will make extensive use of the material discussed in Appendix A. We also recall that definitions of algebraic varieties and semialgebraic sets are given in Section 4.1.4.

By Theorem A.1, there exists a nilmanifold $X = G/\Gamma$ with G connected and simply connected, as well as a polynomial sequence $g\colon \mathbb{Z} \to G$ and a semialgebraic set $S \subseteq X$ such that for each $n \in \mathbb{N}_0$ we have the equivalence: $a_n = x \iff g(n)\Gamma \in S$. The nilmanifold X can be equipped with a Mal'cev coordinate $\tau\colon X \to [0,1)^d$, where we put $d = \dim G$. Recall also from (6) that we have a closely related coordinate map $\widetilde{\tau}\colon G \to \mathbb{R}^d$. Furthermore, inspecting the construction in [BL07], we see that X can be equipped with a Mal'cev coordinate τ such that $\widetilde{\tau} \circ g$ is a polynomial with algebraic coefficients and $\tau(S)$ is defined by equations and inequalities with algebraic coefficients. Since $\mathrm{freq}(\mathbf{a}, x) = 0$, the measure of S is zero. As a consequence (see e.g. [BCR98, Sec. 2.8]), S is contained in a proper algebraic variety $V \subsetneq \mathbb{R}^d$, which is defined over $\mathbb{Q}$.

Now, we pick a non-zero polynomial $p\colon \mathbb{R}^d \to \mathbb{R}$ that vanishes on $V \cup \partial[0,1]^d$, which exists because $V \cup \partial[0,1]^d$ is contained in a proper algebraic sub-variety of $\mathbb{R}^d$ (specifically, in the union of the algebraic variety V and $2d$ hyperplanes). Throughout the argument, we allow all implicit constants to depend on X, g, and S. There are now two cases to consider, depending on the distribution of $g(n)\Gamma$.

Suppose first that the sequence $(g(n)\Gamma)_{n=0}^{\infty}$ is equidistributed in X. Then by Lemma A.4, there exists $c_1 > 0$ such that for each N the sequence $(g(n)\Gamma)_{n=0}^{N-1}$ is $O(N^{-c_1})$-equidistributed for all $N \in \mathbb{N}$. For $\delta > 0$, consider the function $H_\delta\colon X \to [0,1]$ defined by

$$H_\delta(x) = \max\left(1 - \frac{|p(\tau(x))|}{\delta}, 0\right).$$

Then $\|H_\delta\|_{\mathrm{Lip}} = O(1/\delta)$ and $\int H_\delta d\mu_X = O(\delta^{c_2})$ for some $c_2 > 0$. We can now estimate

$$\pi(\mathbf{a}, x; N) \leq \sum_{n=0}^{N-1} H_\delta(g(n)\Gamma) \ll \frac{N^{1-c_1}}{\delta} + \frac{N}{\delta^{c_2}}.$$

Letting $\delta = N^{-c/c_2}$ for a sufficiently small $c > 0$, we obtain $\pi(\mathbf{a}, x; N) = O(N^{1-c})$.

Let us assume now that the sequence $(g(n)\Gamma)_{n=0}^{\infty}$ is not equidistributed in X. In that case, there exists a horizontal character $\eta\colon X \to \mathbb{R}/\mathbb{Z}$ such that $\eta \circ g$ is constant. Let G' be the connected component of $\ker \eta$ and let $\Gamma' = \Gamma \cap G'$. Then, $g(n)$ can be decomposed as $g(n) = g'(n)\gamma(n)$, where g' takes values in G' and γ is periodic. This can be shown by adapting the proof of [GT12b, Prop. 9.2], or using

this result as a black-box and passing to the limit $\delta \to 0$; periodicity of γ follows from Lemma A.12 therein. Let q denote a period of γ. Each of the sequences $\mathbf{a}^{(i)}$, $0 \le i < q$, defined by $a_n^{(i)} = a_{qn+i}$ can be represented using the nilmanifold G'/Γ' rather than X. Reasoning by induction with respect to the dimension d, we may assume that for each i, $0 \le i < q$, there exists c_i such that $\pi(\mathbf{a}^{(i)}, x; N) = O(N^{1-c_i})$. Letting $c = \min_{0 \le i < q} c_i$, we see that $\pi(\mathbf{a}, x; N) = O(N^{1-c})$. $\qquad\square$

8.6. Concluding remarks. We close this section by mentioning several rates of growth about which we do not know if they can be realised. It is also not clear if restriction to bracket words with algebraic coefficients influences the answer.

Question 8.15. Does there exist a bracket word $\mathbf{a}$ over $\{0,1\}$ such that one of the following rates of growth holds?

(i) $\pi(\mathbf{a}, 1; N) = \Theta(N^\lambda)$, where $\lambda \in (0,1) \setminus \mathbb{Q}$.
(ii) $\pi(\mathbf{a}, 1; N) = \Theta(\log^c N)$, where $c \in (1, \infty) \setminus \mathbb{N}$.
(iii) $\pi(\mathbf{a}, 1; N) = \Theta\left(e^{\sqrt{\log N}}\right)$.

9. COMPUTABILITY

As we have already seen, a single bracket word can be represented by many different formulae. This happens, for instance, in Lemma 5.6. An even simpler example is provided by the bracket word $\mathbf{a} = \mathbf{1}_0$, which can be represented as $a_n = \lfloor 1 - \{\alpha n\}\rfloor$ for any $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. A more surprising representation of the same word is attributed to Håland Knutson in [GO10, p. 2]:

$$a_n = \left\lfloor \lfloor \sqrt{2}n \rfloor 2\sqrt{2}n \right\rfloor - \left\lfloor \sqrt{2}n \right\rfloor^2 - 2n^2 + 1.$$

Existence of multiple representations of a single bracket word is not, in and of itself, surprising or worrying. After all, the same phenomenon occurs for usual polynomials, say with real coefficients. The reason why this ambiguity does not lead to problems in that case is that each polynomial $p(x) \in \mathbb{R}[x]$ has a canonical representation $p(x) = \sum_{i=0}^d \alpha_i x^d$, where $\alpha_i \in \mathbb{R}$ and $\alpha_d \ne 0$.

Leibman [Lei12] constructed a similar canonical representation for GP sequences. The full statement of the main result [Lei12, Thm. 0.1 & 0.2] is rather technical, so we state only a simplified version. Specifically, we restrict our attention to bracket words (Leibman's result concerns arbitrary bounded GP maps from $\mathbb{Z}^d$ to $\mathbb{R}$) and do not discuss the details of the construction (Leibman explicitly describes the families $\mathcal{F}_M$ appearing below).

Theorem 9.1. Let $\mathbf{a}$ be a bracket word defined over a finite alphabet Σ. Then there exist families $\mathcal{F}_M$, $M \in \mathbb{N}$, of GP maps from $\mathbb{N}_0$ to $\mathbb{R}$, depending only on polynomials which appear in a given representation of $\mathbf{a}$, with the following properties.

(i) For each $M, d \in \mathbb{N}$, and each d-tuple of different maps $v_1, v_2, \ldots, v_d \in \mathcal{F}_M$, the sequence $(\{v_1(n)\}, \{v_2(n)\}, \ldots, \{v_d(n)\})_{n=0}^\infty$ is equidistributed in $[0,1)^d$.

(ii) There exist integers $M, Q, d \in \mathbb{N}$, maps $v_1, v_2, \ldots, v_d \in \mathcal{F}_M$, and a partition of

$$\{0, 1/Q, \ldots, (Q-1)/Q\} \times [0,1)^d$$

into pairwise disjoint semialgebraic pieces S_i, $i \in \Sigma$, such that for each $n \in \mathbb{N}$ and $i \in \Sigma$,

$$(19) \quad a_n = i \quad \text{if and only if} \quad (\{n/Q\}, \{v_1(n)\}, \{v_2(n)\}, \ldots, \{v_d(n)\}) \in S_i.$$

Remark 9.2.

(i) The families $\mathcal{F}_M$ are explicitly constructed in [Lei12]. We start with polynomial maps $p_1(x), p_2(x), \ldots \in \mathbb{R}[x]$ that span the $\mathbb{Q}$-algebra generated by the polynomials which appear in some fixed representation of **a**. As the first step, $p_1/M, p_2/M, \ldots \in \mathcal{F}_M$. In subsequent steps, we add to $\mathcal{F}_M$ some elements of the form $u\{v\}$ where $u, v \in \mathcal{F}_M$ have already been constructed. For details on exactly which of such elements should be added, we refer to the original paper.

(ii) The representation in 9.1(ii) can be explicitly computed for a given representation of the bracket word **a**. For fixed M, Q, and d, the sets S_i are determined uniquely up to a set of measure zero.

As a consequence of Theorem 9.1, for a given bracket word **a** over an alphabet Σ and $i \in \Sigma$, one can check if $a_n = i$ for almost all $n \in \mathbb{N}_0$ (cf. [Lei12, p. 133]). Indeed, it is enough to verify if S_i has full measure. Since, for any two bracket words **a** and **b** over the same alphabet, we can construct the bracket word $(\llbracket a_n = b_n \rrbracket)_{n=0}^\infty$, one can determine whether $a_n = b_n$ for almost all $n \in \mathbb{N}_0$. The situation is totally different when we insist on exact equality, even when restricting our attention to bracket words with algebraic coefficients.

Theorem 9.3. *Given a GP map $g\colon \mathbb{N}_0 \to \{0,1\}$ with algebraic coefficients, the following problems are undecidable.*

(i) *Determine if $g(n) = 0$ for at least one $n \in \mathbb{N}_0$.*
(ii) *Determine if $g(n) = 0$ for infinitely many $n \in \mathbb{N}_0$.*
(iii) *Determine if $g(n) = 0$ for all but finitely many $n \in \mathbb{N}_0$.*
(iv) *Determine if $g(n) = 0$ for all $n \in \mathbb{N}_0$.*

Corollary 9.4. *It is undecidable if a given GP map $g\colon \mathbb{N}_0 \to \mathbb{R}$ with algebraic coefficients takes only finitely many values.*

Proof. Take an arbitrary GP map $g\colon \mathbb{N}_0 \to \{0,1\}$ and consider the GP map $h\colon \mathbb{N}_0 \to \mathbb{R}$ defined by $h(n) = g(n)n$. Then h takes finitely many values if and only if $g(n) = 0$ for all but finitely many $n \in \mathbb{N}_0$. $\square$

Remark 9.5.

(i) In Theorem 9.3, item (i) is reminiscent of the undecidability of Hilbert's tenth problem, concerning the existence of integer solutions to polynomial equations. In fact, our argument proceeds by a reduction to this problem.

(ii) In the case of k-automatic sequences, the analogues of properties in 9.3(i)–(iv) are easily seen to be decidable.

(iii) Another related result, due to Allouche and Shallit [AS92, Thm. 5.2], asserts that it is undecidable if a given k-regular sequence with integer values has at least one vanishing term. We point out that, conversely to Theorem 9.3(iv), it is decidable if a k-regular sequence is identically zero [KS22, Thm. A].

A key component in the proof of Theorem 9.3 is the existence of a surjective GP map from $\mathbb{N}_0$ to $\mathbb{Z}^d$ for each $d \in \mathbb{N}$.

Proposition 9.6. *For each $d \in \mathbb{N}$, there exists a GP map $h_d\colon \mathbb{N}_0 \to \mathbb{Z}^d$ with algebraic coefficients such that for each $x \in \mathbb{Z}^d$ there exist infinitely many $n \in \mathbb{N}_0$ with $h_d(n) = x$. In particular, h_d is surjective.*

Before we proceed to construct the maps mentioned in Proposition 9.6, let us show how their existence can be used to deduce Theorem 9.3 from classical undecidability results.

Proof of Theorem 9.3 *assuming Proposition* 9.6. Let p be an arbitrary polynomial in $\mathbb{Z}[x_1, \ldots, x_d]$ and let $h \colon \mathbb{N}_0 \to \mathbb{Z}^d$ be the map constructed in Proposition 9.6. Define g_p by $g_p(n) = [\![p(h(n)) = 0]\!]$ for $n \in \mathbb{N}_0$. Then g_p is a GP map from $\mathbb{N}_0$ to $\{0, 1\}$ with algebraic coefficients, and the following conditions are equivalent.

  (a) There exists $n \in \mathbb{N}_0$ with $g_p(n) = 0$.
  (b) There exist infinitely many $n \in \mathbb{N}_0$ with $g_p(n) = 0$.
  (c) There exist $x_1, x_2, \ldots, x_d \in \mathbb{Z}$ with $p(x_1, x_2, \ldots, x_d) = 0$.

Hilbert's tenth problem is known to be undecidable (see, for instance, [Mat93]). Hence, comparing (a) and (c), we conclude that there is no algorithm to determine whether an element of the set $\{g_p \mid d \geq 1,\, p \in \mathbb{Z}[x_1, \ldots, x_d]\}$ has a zero. Similarly, comparing (b) and (c) we conclude that there is no algorithm to determine whether an element of the aforementioned set has infinitely many zeroes. This finishes the proof in cases (i) and (ii).

To derive case (iii) from (ii), it is enough to notice that, for any map $g \colon \mathbb{N}_0 \to \{0, 1\}$, the condition that $g(n) = 0$ for all but finitely many $n \in \mathbb{N}_0$ is equivalent to the condition that there are finitely many $n \in \mathbb{N}_0$ such that $1 - g(n) = 0$. Likewise, (iv) follows from (i) and the observation that $g(n) = 0$ for all $n \in \mathbb{N}_0$ is equivalent to the condition that there does not exist $n \in \mathbb{N}_0$ such that $1 - g(n) = 0$. $\qquad\square$

We now turn to the proof of Proposition 9.6. The main difficulty lies in the construction of a surjective GP map $\mathbb{N}_0 \to \mathbb{N}_0^2$; once such a map is constructed, it will not be difficult to use it to construct surjective maps $\mathbb{N}_0 \to \mathbb{Z}^d$ for all $d \in \mathbb{N}$. As a source of motivation, let us consider a (random) map $\mathbf{f} \colon \mathbb{N}_0 \to \mathbb{N}_0^2$ given by $\mathbf{f}(n) = \left(\lfloor \mathbf{X}_n^2 n \rfloor, \lfloor \mathbf{Y}_n^2 n \rfloor \right)$, where $\mathbf{X}_n, \mathbf{Y}_n$ for $n \in \mathbb{N}_0$ are jointly independent random variables, uniformly distributed in $[0, 1)$. One can explicitly compute that, for fixed $k, l \in \mathbb{N}$ and sufficiently large N, we have

$$\mathbb{P}(\mathbf{f}(n) = (k, l)) = \frac{\left(\sqrt{k+1} - \sqrt{k}\right)\left(\sqrt{l+1} - \sqrt{l}\right)}{n}.$$

Hence, by the second Borel–Cantelli lemma, almost surely there exist infinitely many $n \in \mathbb{N}_0$ with $\mathbf{f}(n) = (k, l)$. This prompts us to consider generalised polynomials of the form $\left(\lfloor \{\sqrt{2}n\}^2 n \rfloor, \lfloor \{\sqrt{3}n\}^2 n \rfloor \right)$, or, more generally

$$\left(\left\lfloor \{\sqrt{2}n\}^A n \right\rfloor, \left\lfloor \{\sqrt{3}n\}^A n \right\rfloor \right)$$

for $A \geq 2$, and exploit equidistribution properties of the sequence $\left(\{\sqrt{2}n\}, \{\sqrt{3}n\}\right)$ in $[0, 1)^2$. As a first step in that direction, we mention the following quantitative equidistribution estimate.

Lemma 9.7. *There exist $N_0 > 0$ and $c > 0$ such that for each $M \in \mathbb{Z}$, $N \geq N_0$, and $(x, y) \in [0, 1)^2$, there exists $n \in [M, M + N)$ such that*

$$\max\left\{ \left\| \sqrt{2}n - x \right\|, \left\| \sqrt{3}n - y \right\| \right\} \leq 1/N^c.$$

Proof. A more precise version follows from [Che00, Thm 1.(ii)] combined with standard estimates on the quality of approximate rational relations between algebraic

numbers. Specifically, we have the well-known estimate

$$\min_{(n,m)\in[N]^2\setminus\{(0,0)\}} \left\| n\sqrt{2} + m\sqrt{3} \right\| \gg N^{-3},$$

which is easily obtained by noticing that the norm $\prod_{\sigma,\rho\in\{\pm 1\}} \left(k + \sigma n\sqrt{2} + \rho m\sqrt{3} \right)$ is an integer for all $k, n, m \in \mathbb{Z}$. (Stronger estimates follow from Schmidt's subspace theorem [Sch72], but are not needed for our purposes.)

The theorem is applied with $n = 2$, $m = 1$, $\alpha_1 = x$, $\alpha_2 = y$, $T_1 = M$, $T_1' = M+N$, $\delta_1 = \delta_2 = 1$, and $M = N^{1/7}$; then $\Delta = 2$, $\Lambda_2 \gg N^{-3}$, and $\|S(T)\| = N$. The conclusion, after elementary manipulations, asserts that there exists $n \in [M, M+N)$ such that

$$\max\left\{ \left\| \sqrt{2}n - x \right\|, \left\| \sqrt{3}n - y \right\| \right\} \ll N^{-2/7}.$$

Alternatively, one can also derive this estimate as a special case of Theorem A.3. $\square$

Lemma 9.8. *For $A \in \mathbb{N}$, define the map $g_A \colon \mathbb{N}_0 \to \mathbb{N}_0^2$ by*

$$g_A(n) = \left(\left\lfloor \{\sqrt{2}n\}^A n \right\rfloor, \left\lfloor \{\sqrt{3}n\}^A n \right\rfloor \right).$$

There exists a constant A_0 such that for all $A > A_0$, for all $(k, l) \in \mathbb{N}_0$ the set

$$\{ n \in \mathbb{N}_0 \mid g_A(n) = (k, l) \}$$

is infinite. In particular, g_A is surjective.

Proof. Let c be the constant from Lemma 9.7 and let $A > 3/c$. Fix $(k, l) \in \mathbb{N}_0^2$, let N be a sufficiently large integer (to be determined in the course of the argument), and put $M = N^2$. It will suffice to show that there exists $n \in [M, M + N)$ with $g_A(n) = (k, l)$.

Pick $(x, y) \in [0, 1)^2$ such that $x^A M = k$ and $y^A M = l$ (that is, $x = (k/M)^{1/A}$ and $y = (l/M)^{1/A}$). By Lemma 9.7, we can find $n \in [M, M + N)$ such that $x \leq \{\sqrt{2}n\} < x + 1/N^c$ and $y \leq \{\sqrt{3}n\} < y + 1/N^c$. It remains to show that

$$(20) \qquad k \leq \{\sqrt{2}n\}^A n < k + 1 \text{ and } l \leq \{\sqrt{3}n\}^A n < l + 1.$$

We only consider the first of these two conditions, the second one is fully analogous. The lower bound is immediate:

$$\{\sqrt{2}n\}^A n \geq x^A n = kn/M \geq k.$$

For the upper bound, we first obtain that

$$\{\sqrt{2}n\}^A n < (x + N^{-c})^A (M + N).$$

If $k = 0$, then $x = 0$ and thus $(x + N^{-c})^A (M + N) < 2N^{2-Ac} < 1$ (recall that $A > 3/c$). If $k \geq 1$, then $x > N^{-2c/3}$ and hence

$$(x + N^{-c})^A (N_0 + N) = x^A N_0 \left(1 + N^{-c}/x\right)^A (1 + N/M)$$
$$\leq k\left(1 + N^{-c/3}\right)^A (1 + 1/N) \leq k \exp\left((A + 1)N^{-c/3}\right).$$

Thus, we can find $N_0 = N_0(k) = O_A(k^{3/c})$ such that, for all $N \geq N_0$, we have

$$\{\sqrt{2}n\}^A n \leq k \exp\left((A + 1)N^{-c/3}\right) < k + 1,$$

as needed. $\square$

Proof of Proposition 9.6. Let $g_2 \colon \mathbb{N}_0 \to \mathbb{N}_0^2$ denote the map constructed in Lemma 9.8. For $d \geq 2$, define $g_d \colon \mathbb{N}_0 \to \mathbb{N}_0^d$ inductively by

$$g_{i+1}(n) = (g_2(n)_1, g_i(g_2(n)_2)) .$$

An inductive argument shows that for each $x \in \mathbb{N}_0^d$, there are infinitely many $n \in \mathbb{N}_0$ with $g_d(n) = x$. Next, define $h_d \colon \mathbb{N}_0 \to \mathbb{Z}^d$ by

$$h_d(n) = (g_{2d}(n)_1 - g_{2d}(n)_{d+1}, g_{2d}(n)_2 - g_{2d}(n)_{d+2}, \ldots, g_{2d}(n)_d - g_{2d}(n)_{2d}) .$$

Since the map $\mathbb{N}_0^2 \to \mathbb{Z}$, $(n,m) \mapsto n - m$ is surjective, for each $x \in \mathbb{Z}^d$, there are infinitely many $n \in \mathbb{N}_0$ with $h_d(n) = x$. □

9.1. Consequences for bracket words.

As a consequence of Theorem 9.3, we deduce Theorem B, which we restate below for the reader's convenience.

Theorem B. *It is undecidable if two given bracket words* **a** *and* **b** *with algebraic coefficients defined over a finite alphabet* Σ *are equal.*

Proof. Given an arbitrary GP map $g \colon \mathbb{N}_0 \to \{0,1\}$, define **a** by $a_n = g(n)$ for all $n \in \mathbb{N}_0$, and let **b** be defined by $b_n = 0$ for all $n \in \mathbb{N}_0$. Then $\mathbf{a} = \mathbf{b}$ if and only if $g(n) = 0$ for all $n \in \mathbb{N}_0$. Thus, the result follows from Theorem B. □

10. LINEAR RECURRENCES

In earlier sections, we have encountered examples of bracket words related to linear recurrence sequences, such as the Fibonacci numbers. Here, we discuss these results in more detail and provide some new arguments.

10.1. Results.

We recall that a *Pisot number* is a real algebraic integer $\beta > 1$ such that all Galois conjugates of β have absolute value strictly less than 1. Similarly, a *Salem number* is a real algebraic integer $\beta > 1$ whose Galois conjugates all have absolute value no greater than 1, and at least one of which has absolute value exactly 1. The minimal polynomial for a Salem number must be reciprocal, which implies that $1/\beta$ is a Galois conjugate of β, and that all other roots have absolute value exactly one. As a consequence, a Salem number is a unit in the ring of algebraic integers.

Theorem 10.1. *Let* $\beta > 1$ *be an algebraic unit with minimal polynomial* $p(x) = x^d - \sum_{i=1}^d x^{d-i} a_i$ *for some* $d \in \mathbb{N}$ *and* $a_1, a_2, \ldots, a_d \in \mathbb{Z}$. *Let* $(n_i)_{i=0}^\infty$ *be a sequence of non-negative integers satisfying the linear recurrence*

$$n_{i+d} = \sum_{i=1}^d a_i n_{i+d-1}, \qquad i \in \mathbb{N}_0,$$

and put $E = \{n_i \mid i \in \mathbb{N}_0\}$. *Then* $\mathbf{1}_E$ *is a bracket word if one of the following holds.*

(i) $d = 2$ *(in this case,* β *must be a Pisot number).*

(ii) $d = 3$ *and* β *is a Pisot number with no real Galois conjugate (i.e., the discriminant of* p *is negative).*

(iii) β *is a Salem number.*

Proof. Case (i) is covered by [BK18, Thm. B], using an argument which relies on best rational approximations to quadratic irrationals. Case (ii) likewise follows from [BK18, Thm. B] under mild additional assumptions. In Theorem 10.2 we give a complete proof of Case (ii) which has a distinctly algebraic flavour, in contrast to

the argument in [BK18] relying on Diophantine approximation. Finally, case (iii) is covered in the upcoming preprint [BK22] using methods analogous to those used in Theorem 10.2. $\square$

In the remainder of this section, we will mostly speak of GP subsets of $\mathbb{N}_0$ rather than bracket words. We recall that, as defined in Section 2, these terms are closely related, and the connection between them if further elucidated in Corollary 5.7.

10.2. **Cubic Pisot units.** The main ingredient in the proof of Theorem 10.1(ii) is the following result. In fact, the two results are equivalent due to a reduction obtained in [BK18, Prop. 5.1].

Theorem 10.2. *Let $\beta > 1$ be a cubic Pisot unit with a pair of complex Galois conjugates $\alpha, \overline{\alpha}$ and let $E = \left\{ \lfloor \beta^i \rceil \mid i \in \mathbb{N}_0 \right\}$. Then $\mathbf{1}_E$ is a bracket word.*

Proof. We will devise a procedure that verifies if, for a given integer $n \in \mathbb{Z}$, we have

$$(21) \qquad n = \lfloor \beta^i \rceil \text{ for some } i \in \mathbb{N}_0.$$

Later, we will explain how this procedure can be encoded using a generalised polynomial formula. Throughout, we assume that $|n|$ is sufficiently large, which we may do because GP sets are closed under finite modifications. Set $L = \mathbb{Q}(\beta, \alpha, \overline{\alpha})$ and $K = \mathbb{Q}(\beta)$, and let $p(x) = x^3 - ax^2 - bx - 1$ denote the minimal polynomial of β.

Suppose for a moment that (21) holds, and hence in particular $n > 0$. Our first goal is to compute β^i as a GP function of n. Since $\beta^i + \alpha^i + \overline{\alpha}^i$ is an integer and $\left| \alpha^i \right| = \beta^{-i/2} \ll 1/\sqrt{n}$, we see that

$$(22) \qquad n = \beta^i + \alpha^i + \overline{\alpha}^i.$$

Similarly, for every $j \in \mathbb{N}$, we have

$$\beta^j n = \beta^{i+j} + \alpha^{i+j} + \overline{\alpha}^{i+j} + O(\beta^j/\sqrt{n}),$$

and, as a consequence, bearing in mind that n is large enough, we obtain

$$(23) \qquad \lfloor \beta n \rceil = \beta^{i+1} + \alpha^{i+1} + \overline{\alpha}^{i+1},$$
$$(24) \qquad \lfloor \beta^2 n \rceil = \beta^{i+2} + \alpha^{i+2} + \overline{\alpha}^{i+2}.$$

Thus, we have expressed $(n, \lfloor \beta n \rceil, \lfloor \beta^2 n \rceil)$ as a linear function of $(\beta^i, \alpha^i, \overline{\alpha}^i)$. As a consequence, we can compute $(\beta^i, \alpha^i, \overline{\alpha}^i)$ by solving a system of linear equations involving $(n, \lfloor \beta n \rceil, \lfloor \beta^2 n \rceil)$ and algebraic coefficients belonging to L.

For each $n \in \mathbb{Z}$, let $g(n)$, $h(n)$, and $h^*(n)$ be the solution to the system of equations

$$(25) \qquad g(n) + h(n) + h^*(n) = n,$$
$$(26) \qquad g(n)\beta + h(n)\alpha + h^*(n)\overline{\alpha} = \lfloor \beta n \rceil,$$
$$(27) \qquad g(n)\beta^2 + h(n)\alpha^2 + h^*(n)\overline{\alpha}^2 = \lfloor \beta^2 n \rceil.$$

Note that (25)–(27) are non-singular, so $g(n)$, $h(n)$, and $h^*(n)$ are well-defined and unique. In fact, one can explicitly compute that

$$(28) \qquad g(n) = \frac{\lfloor \beta n \rceil (\alpha + \overline{\alpha}) - \alpha \overline{\alpha} n - \lfloor \beta^2 n \rceil}{(\beta - \alpha)(\beta - \overline{\alpha})} = \frac{\lfloor \beta n \rceil (a - \beta) - n/\beta - \lfloor \beta^2 n \rceil}{2\beta^2 - a\beta + 1/\beta},$$

$$(29) \qquad h(n) = \frac{\lfloor \beta n \rceil (\beta + \overline{\alpha}) - \beta \overline{\alpha} n - \lfloor \beta^2 n \rceil}{(\alpha - \overline{\alpha})(\alpha - \beta)} = \frac{\lfloor \beta n \rceil (a - \alpha) - n/\alpha - \lfloor \beta^2 n \rceil}{2\alpha^2 - a\alpha + 1/\alpha},$$

$$(30) \qquad h^*(n) = \frac{\lfloor \beta n \rceil (\alpha + \beta) - \alpha \beta n - \lfloor \beta^2 n \rceil}{(\overline{\alpha} - \alpha)(\overline{\alpha} - \beta)} = \frac{\lfloor \beta n \rceil (a - \overline{\alpha}) - n/\overline{\alpha} - \lfloor \beta^2 n \rceil}{2\overline{\alpha}^2 - a\overline{\alpha} + 1/\overline{\alpha}}.$$

The key reason for our interest in g, h, h^* is that, if (21) holds, then it follows from the discussion above that $g(n) = \beta^i$, $h(n) = \alpha^i$ and $h^*(n) = \overline{\alpha}^i$. Indeed, $(\beta^i, \alpha^i, \overline{\alpha}^i)$ is a solution to (25)–(27), and the solution to (25)–(27) is unique.

It is apparent from formulae (28)–(30) (or from the symmetries of (25)–(27)) that $h^*(n) = \overline{h(n)}$ and $g(n) \in K = \mathbb{Q}(\beta)$, $h(n) \in \mathbb{Q}(\alpha)$ for all $n \in \mathbb{Z}$. We also see that $g(n)$ is a linear combination of n, $\lfloor \beta n \rceil$, and $\lfloor \beta^2 n \rceil$. Hence g is a GP map. For the same reason, $\mathrm{Re}(h)$ and $\mathrm{Im}(h)$ are GP maps.[3]

Our next goal is to express condition (22) in terms of the maps g, h, h^*. By the Dirichlet unit theorem, we know that the group of units of $\mathcal{O}_K$ has rank 1. Since β is a unit, it generates a group that has finite index in the group of all units of $\mathcal{O}_K$. Assume for now that β is a fundamental unit, meaning that all units in $\mathcal{O}_K$ take the form $\pm \beta^i$ for $i \in \mathbb{Z}$. We address the general case at the end of the proof.

Recall that if (21) holds for some $n \in \mathbb{Z}$, then $g(n) = \beta^i$ and $h(n) = \alpha^i$. Thus, $g(n) \in \mathcal{O}_K$ and $g(n)h(n)h^*(n) = 1$. If, additionally, $|n|$ is sufficiently large, then $g(n) \in [n-1, n+1)$. Suppose, conversely, that for some $n \in \mathbb{Z}$ we have $g(n) \in \mathcal{O}_K$ and $g(n)h(n)h^*(n) = 1$. Then $g(n)$ is a unit, since its norm $N_{L/\mathbb{Q}}(g(n)) = g(n)h(n)h^*(n)$ is equal to 1. As a consequence, we have $g(n) = \pm\beta^i$ for some $i \in \mathbb{Z}$. If additionally $g(n) \geq n-1$ then $g(n) = \beta^i$ for some $i \in \mathbb{N}$. Thus, for all but finitely many $n \in \mathbb{Z}$, (21) is equivalent to

$$(31) \qquad g(n) \in \mathcal{O}_K, \quad g(n)h(n)h^*(n) = 1, \qquad \text{and} \qquad g(n) \in [n-1, n+1).$$

Our final goal is to express the conditions in (31) in terms of generalised polynomials. As a first step in this direction, we will need a more precise description of g, h, h^*. Since $g(n) \in \mathbb{Q}(\beta)$, for each $n \in \mathbb{Z}$, there exists a decomposition

$$(32) \qquad\qquad g(n) = u(n) + v(n)\beta + w(n)\beta^2,$$

where $u(n)$, $v(n)$, and $w(n) \in \mathbb{Q}$. If $\sigma \in \mathrm{Gal}(L/\mathbb{Q})$ is an automorphism with $\sigma(\beta) = \alpha$, then $\sigma(g(n)) = h(n)$, and as a consequence we also have

$$(33) \qquad\qquad h(n) = u(n) + v(n)\alpha + w(n)\alpha^2,$$

$$(34) \qquad\qquad h^*(n) = u(n) + v(n)\overline{\alpha} + w(n)\overline{\alpha}^2.$$

Arguing along similar lines as above, we can express $u(n)$, $v(n)$, and $w(n)$ as linear combinations of $g(n)$, $h(n)$, and $h^*(n)$, and hence also as a linear combination

---

[3]Note that we have not introduced the notion of a GP map $\mathbb{N}_0 \to \mathbb{C}$. However, if one identifies $\mathbb{C}$ with $\mathbb{R}^2$ in the standard way, then h is a GP map under this identification.

of n, $\lfloor \beta n \rceil$, and $\lfloor \beta^2 n \rceil$. For instance,

$$u(n) = \frac{1}{\Delta} \Big(\left(-2a^3 + a^2 b^2 - 10ab + 4b^3 - 9 \right) n$$
$$+ \left(a^3 b - a^2 + 4ab^2 + 6b \right) \lfloor \beta n \rceil + \left(-a^2 b + 3a - 4b^2 \right) \lfloor \beta^2 n \rceil \Big),$$

where $\Delta = -4a^3 + a^2 b^2 - 18ab + 4b^3 - 27$ is the discriminant of p. In particular, u, v, and w are GP maps.

Since g, $\mathrm{Re}(h)$ and $\mathrm{Im}(h)$ are generalised polynomials, it follows that

$$E_1 = \{ n \in \mathbb{Z} \mid g(n)h(n)h^*(n) = 1 \} \text{ and } E_2 = \{ n \in \mathbb{Z} \mid g(n) - n \in [-1,1) \}$$

are GP sets. Let us consider the set

$$\Lambda = \left\{ (x,y,z) \in \mathbb{Q}^3 \mid x + y\beta + z\beta^2 \in \mathcal{O}_K \right\}.$$

Clearly, Λ is a lattice. It follows that Λ is a GP subset of $\mathbb{R}^3$. Indeed, if $\Lambda = A\mathbb{Z}^3$ for a matrix $A \in \mathrm{GL}(3; \mathbb{R})$, then

$$\Lambda = \left\{ x \in \mathbb{R}^3 \mid \left\{ \left(A^{-1} x \right)_1 \right\} + \left\{ \left(A^{-1} x \right)_2 \right\} + \left\{ \left(A^{-1} x \right)_3 \right\} = 0 \right\}.$$

Recalling that u, v, and w are GP maps, we conclude that

$$E_3 = \{ n \in \mathbb{Z} \mid g(n) \in \mathcal{O}_K \} = \{ n \in \mathbb{N}_0 \mid (u(n), v(n), w(n)) \in \Lambda \}$$

is a GP set. Hence, $E_1 \cap E_2 \cap E_3$ is a GP set by Proposition 6.16, and it is precisely the set of all $n \in \mathbb{Z}$ which satisfy (31), as needed.

Let us now consider the case where the fundamental unit of K is some $\widetilde{\beta} \neq \pm\beta$. We may assume that $\widetilde{\beta} > 0$. Since β is a unit, we have $\beta = \widetilde{\beta}^k$ for some $k \in \mathbb{N}$. It is straightforward to verify that $\widetilde{\beta}$ is again a cubic Pisot unit with a pair of complex conjugates, say $(\gamma, \overline{\gamma})$, so the above discussion applies with $\widetilde{\beta}$ in place of β. Set $\widetilde{n}_i = \widetilde{\beta}^i + \gamma^i + \overline{\gamma}^i$, for every $i \in \mathbb{N}_0$. Then $(\widetilde{n}_i)_{i=0}^{\infty}$ satisfies a linear recurrence relation, and, for all sufficiently large i, we have $\widetilde{n}_i = \lfloor \widetilde{\beta}^i \rceil$ and $\widetilde{n}_{i+1} = \lfloor \widetilde{\beta} \widetilde{n}_i \rceil$. Furthermore, $\widetilde{E} = \{ \widetilde{n}_i \mid i \in \mathbb{N}_0 \}$ is a GP subset of $\mathbb{N}_0$. One can find an integer m such that the sequence $(\widetilde{n}_i \bmod m)_{i=0}^{\infty}$ is periodic with minimal period ℓ that is a multiple of k [War33].[4] Note that, for each sufficiently large i and for each j, $0 \leq j < \ell$, we have

$$\left\lfloor \widetilde{\beta}^j \widetilde{n}_i \right\rceil \equiv \widetilde{n}_{i+j \bmod \ell} \bmod m.$$

This allows us to identify, for each residue class r, $0 \leq r < \ell$, the elements $n \in \widetilde{E}$ which take the form $n = \widetilde{n}_i$ with $i \equiv r \bmod \ell$. With finitely many exceptions, these are precisely those $n \in \widetilde{E}$ for which $\left\lfloor \widetilde{\beta}^j \widetilde{n} \right\rceil \equiv \widetilde{n}_{r+j \bmod \ell} \bmod m$ for all j, $0 \leq j < \ell$. It follows that the set $\left\{ \lfloor \beta^i \rceil \mid i \in \mathbb{N}_0 \right\}$ differs by finitely many elements from

$$(35) \qquad \left\{ n \in \widetilde{E} \mid (\exists\, 0 \leq a < \ell/k)\, (\forall\, 0 \leq j < \ell)\, \lfloor \widetilde{\beta}^j n \rceil \equiv \widetilde{n}_{ak+j} \bmod m \right\}.$$

Since the set in (35) is a GP set, and since the property of being a GP set is preserved under finite modifications, we conclude that the set $\left\{ \lfloor \beta^i \rceil \mid i \in \mathbb{N}_0 \right\}$ is also a GP set. $\qquad\square$

---

[4]More precisely, by [War33, Thm. 1], it is enough to consider the case where k is a power of a prime p. By [War33, Thm. 10.2 & Cor. 3], the period of $(\widetilde{n}_i \bmod p^N)_i$ is a multiple of the order of $\widetilde{\beta}$ in $\mathbb{F}_q[\widetilde{\beta}]^{\times}$, where $q = p^{N-C}$ for a constant C. In particular, the period of $(\widetilde{n}_i \bmod p^N)_i$ is a multiple of k for sufficiently large N.

Proof of Theorem 10.1(ii). Follows immediately from Theorem 10.2 and [BK18, Prop. 5.1]. □

10.3. **Concluding remarks.** Let us now briefly discuss potential generalisations of Theorem 10.1. Let $p(x) = x^d - \sum_{i=1}^d a_i x^{d-i} \in \mathbb{Z}[x]$ be a monic irreducible polynomial of degree $d \in \mathbb{N}$ with a root $\beta > 1$ which is either a Pisot unit or a Salem number, and let $\alpha_1, \alpha_2, \ldots, \alpha_{d-1}$ denote the remaining roots.

Much of the reasoning in Theorem 10.2 carries through to this more general context, except that the group of units of $\mathcal{O}_K$ now no longer has rank 1 (here, $K = \mathbb{Q}(\beta)$). Adapting the argument, we can hope to show that

$$F = \left\{ \mathrm{Tr}_{L/\mathbb{Q}}(\mu) \mid \mu \text{ is a unit of } \mathcal{O}_K \right\}$$

is a GP subset of $\mathbb{Z}$, where $L = \mathbb{Q}(\beta, \alpha_1, \ldots, \alpha_{d-1})$ is the splitting field of p.

Consider also the sequence given by $n_i = \mathrm{Tr}(\beta^i) = \beta^i + \sum_{k=1}^{d-1} \alpha_k^i \in \mathbb{Z}$, and note that n_i obeys the linear recurrence

$$n_{i+d} = \sum_{j=1}^d a_j n_{i+d-j}, \qquad i \in \mathbb{N}_0.$$

Set $E = \{ n_i \mid i \in \mathbb{N}_0 \} \cap \mathbb{N}$. Note that the assumption on β ensures that $n_i = \lfloor \beta^i \rceil + O(1)$ for all $i \in \mathbb{N}_0$. In analogy with Theorem 10.1, one can ask if E is a GP subset of $\mathbb{N}$.

In the case where the group of units of $\mathcal{O}_K$ has rank 1, the set F is essentially equal to E, which is one of the key observations behind the proof of Theorem 10.2. In general, we have the inclusion $E \supset F$, but E will usually be a proper subset of F and there is no clear way of describing E inside of F with a GP formula.

In the case where β is a Pisot number but not a unit, we have a negative result for $d = 1$: if $k \geq 2$ is and integer, then $E = \left\{ k^i \mid i \in \mathbb{N}_0 \right\}$ is not a GP set. (We revisit this example in Section 11.) When $d \geq 2$, nothing is known, but, based on the previously mentioned result, it seems reasonable to expect a negative result. In the case where β is neither a Pisot nor a Salem number (i.e., either β is transcendental or β has a Galois conjugate with absolute value greater than 1) nothing is known, and the techniques discussed in this section appear not to be applicable. We end this section with the following two general problems.

Problem 10.3.

(i) Classify all sequences of positive integers $(n_i)_{i=1}^\infty$ satisfying a linear recurrence such that the characteristic word $\mathbf{1}_E$ of the set $E = \{ n_i \mid i \in \mathbb{N}_0 \}$ is a bracket word.

(ii) Classify all real numbers $\beta > 1$ such that the characteristic word $\mathbf{1}_E$ of the set $E = \left\{ \lfloor \beta^i \rceil \mid i \in \mathbb{N}_0 \right\}$ is a bracket word.

11. Negative results

In this section we discuss the problem of proving that a given infinite word cannot be described by a generalised polynomial formula. We remark that, as a general principle, it is often harder to verify that a sequence does not admit a representation of a specified form than it is to find such a representation when it exists. This phenomenon is particularly prevalent from a computational point of view. For context, we also note that Allouche, Shallit, and Yassawi [ASY22] recently survey ways in which one can show that a sequence is not automatic.

11.1. **General conditions.** Let us now review criteria that can be used to show that a given infinite word **a** is not a bracket word. In principle, each "positive" result about bracket words gives rise to such a criterion, and *vice versa*. For each property P which can apply to finitely-valued sequences, the implication "*If **a** is a bracket word then $P(\mathbf{a})$*" is tantamount to "*If $\neg P(\mathbf{a})$ then **a** is not a bracket word*". In practice, depending on the aesthetic appeal of the property $P(\mathbf{a})$ and on the ease of verifying $\neg P(\mathbf{a})$, one of these implications is more interesting than the other. Proposition 11.1 gathers several such properties that arise from positive results discussed earlier. We recall that $\mathrm{freq}(\mathbf{a}, w)$ and $\mathrm{rec}(\mathbf{a}, w)$ are defined in Section 7.1.

Proposition 11.1. *Let **a** be an infinite word defined over a finite alphabet Σ. Then **a** is not a bracket word if one of the following conditions hold.*

(i) *There is $w \in \Sigma^*$ such that $\mathrm{freq}(\mathbf{a}, w)$ does not exist.*

(ii) *There is $w \in \Sigma^*$ such that $\mathrm{freq}(\mathbf{a}, w) > 0$ but $\mathrm{rec}(\mathbf{a}, w) = \infty$.*

(iii) *The subword complexity of **a** satisfies $\limsup_{n \to \infty} \log(p_{\mathbf{a}}(N))/\log N = \infty$.*

(iv) *There exists $x \in \Sigma$ with $\mathrm{freq}(\mathbf{a}, x) < 1$ and $t \colon \mathbb{N}_0 \to \mathbb{N}_0$ that is good for equidistribution on nilsystems (cf. Section 7.2) such that $a_{t(n)} = x$ for all $n \in \mathbb{N}_0$.*

Proof. These are immediate consequences of Theorem 7.1, Corollary 7.2, Theorem A, and Proposition 7.8. $\qquad\square$

Another way to discriminate bracket words that we want to mention is related to periodicity. Following [BK20], we say that an infinite word **a** is *weakly periodic* if for every infinite arithmetic progression $(kn + r)_{n=0}^{\infty}$ ($k \in \mathbb{N}$, $r \in \mathbb{N}_0$) there exist two distinct sub-progressions with equal steps $(k'n + r')_{n=0}^{\infty}$ and $(k'n + r'')_{n=0}^{\infty}$ ($k' \in \mathbb{N}$ and $k \mid k'$, r' and r'' belong to $\mathbb{N}_0$, $r' \equiv r'' \equiv r \bmod k$, and $r' \neq r''$) such that the corresponding restrictions of **a** are the same: $a_{k'n+r'} = a_{k'n+r''}$ for all $n \in \mathbb{N}_0$. For instance, all automatic sequences are weakly periodic, all Toeplitz sequences are weakly periodic, and the characteristic sequence of the square-free integers is weakly periodic.[5] We will also say that **a** is *almost everywhere periodic* if there exists an infinite periodic word $\mathbf{a}'$ such that $a_n = a'_n$ for almost all n, that is, $d\left(\{n \in \mathbb{N}_0 \mid a_n \neq a'_n\}\right) = 0$. The following result is a rephrasing of [BK20, Thm. 2.6].

Proposition 11.2 ([BK20, Thm. 2.6])**.** *Let **a** be an infinite word that is weakly periodic but not almost everywhere periodic. Then **a** is not a bracket word.*

The proof relies on Theorem 4.1 and the fact that each nilsystem (X, T) can be partitioned into a finite number of components, $X = X_1 \cup X_2 \cup \cdots \cup X_d$, such that (X_i, T^d) is a totally minimal[6] dynamical system for all i, $1 \le i \le d$. Next, we observe that, for a totally minimal dynamical system (X, T), $x \in X$, open $S \subseteq X$, $k \in \mathbb{N}$, $r, r' \in \mathbb{N}_0$ with $r \neq r'$, if for all $n \in \mathbb{N}_0$ we have the implication

$$T^{kn+r}(x) \in S \Rightarrow T^{kn+r'}(x) \in S,$$

then S is either empty or dense in X (see [BK20, Lem. 2.4]).

---

[5]A sequence **a** is a Toeplitz sequence if for each n there exists a period $q \in \mathbb{N}$ such that $a_{n+qm} = a_n$ for all $m \in \mathbb{N}$; see, for instance, [Dow05] for background. An integer n is square-free if, for every prime p, one has $p^2 \nmid n$.

[6]A topological dynamical system (X, T) is said to be totally minimal if (X, T^d) is minimal for every $d \ge 1$.

Example 11.3. The indicator function of the square-free integers is not a bracket word (which can also be derived from item (iii) of Proposition 11.1). A non-periodic Toeplitz sequence is not a bracket word (note that a Toeplitz sequence is periodic if and only if it is almost everywhere periodic). It was shown in [CK97] that a large class of Toeplitz sequences has polynomial subword complexity, so item (iii) of Proposition 11.1 cannot be applied in this case, nor can item (ii) of Proposition 11.1.

As a counterpart to Theorem 7.11, it is shown in [BK18] that any GP subset $E \subseteq \mathbb{N}_0$ with $d(E) = 0$ is very poor in terms of additive structure. Here, we cite a slightly weaker (but more succinct) variant of [BK18, Thm. A].

Proposition 11.4 ([BK18, Thm. A]). *Let $E \subseteq \mathbb{N}_0$ be a set. Suppose that $d(E) = 0$ and E contains an IP set. Then E is not a GP set.*

Corollary 11.5. *Let $\mathbf{a} = (a_n)_{n=0}^{\infty}$ be an infinite word over a finite alphabet Σ. Suppose that there exist $x \in \Sigma$ and an IP set $E \subseteq \mathbb{N}$ such that $\mathrm{freq}(\mathbf{a}, x) = 0$ and $a_n = 0$ for all $n \in E$. Then $\mathbf{a}$ is not a bracket word.*

Example 11.6. The characteristic word of the set $\left\{\sum_{n \in I} n! \mid I \subseteq \mathbb{N}, |I| < \infty\right\}$ is not a bracket word.

Finally, we cite a result recently obtained by the second-named author [Kon21, Thm. A], as a final ingredient needed to finish the classification of automatic bracket words, cf. Theorem 11.23. Recall that a set $E \subseteq \mathbb{N}_0$ is *thick* if its complement is not syndetic, or equivalently if for each ℓ there exists n with $n, n+1, \ldots, n+\ell \in E$.

Theorem 11.7 ([Kon21, Thm. A]). *Let $k \geq 2$ be an integer, let $E \subseteq \mathbb{N}_0$, and put*

$$F = \{m \in \mathbb{N} \mid mk^n \in E \text{ for infinitely many } n \in \mathbb{N}_0\}.$$

If $\{k^n \mid n \in \mathbb{N}_0\} \subseteq E$ and $\mathbb{N} \setminus F$ is thick then E is not a GP set.

Example 11.8. Let $k \geq 2$ be an integer. Then the characteristic word of the set $\{k^n \mid n \in \mathbb{N}_0\}$ is not a bracket word.

For a set $E \subseteq \mathbb{N}_0$ and $k \in \mathbb{N}$, we set $E/k = \{n \in \mathbb{N}_0 \mid kn \in E\}$. Restricting our attention to sets $E \subseteq \mathbb{N}_0$ with $E/k = E$, we obtain a cleaner statement, which can be phrased either in terms of GP sets or bracket words.

Corollary 11.9. *Let $k \geq 2$ and $E \subseteq \mathbb{N}$. If $E \neq \emptyset$, $d(E) = 0$, and $E/k = E$, then E is not a GP set.*

Proof. Pick $n_0 \in E$ and let $E' = E/n_0$. Then $d(E') = 0$, $k^n \in E'$ for all n, and

$$F' = \{m \in \mathbb{N} \mid mk^n \in E' \text{ for infinitely many } n \in \mathbb{N}_0\} = E'.$$

Since $d(F') = 0$, the set $\mathbb{N} \setminus F'$ is thick and we can apply Theorem 11.7. $\qquad\square$

Corollary 11.10. *Let $\mathbf{a} = (a_n)_{n=0}^{\infty}$ be an infinite word over a finite alphabet Σ, let $k \geq 2$ be an integer and let $x \in \Sigma$. Suppose that (i) $a_{kn} = a_n$ for all $n \in \mathbb{N}$, (ii) $\mathrm{freq}(\mathbf{a}; x) = 0$, and (iii) $a_n = x$ for at least one $n \in \mathbb{N}$. Then $\mathbf{a}$ is not a bracket word.*

11.2. **Primes and squares.** Let us consider two standard examples: the primes $\mathcal{P} = \{2, 3, 5, 7, \dots\}$ and the squares $\mathcal{S} = \{0, 1, 4, 9, \dots\}$. In both cases, we can show that the corresponding characteristic word is not a bracket word. The arguments are short and rely heavily on theorems from nilpotent dynamics.

Proposition 11.11. *The characteristic word $\mathbf{1}_\mathcal{P}$ of the set of primes is not a bracket word.*

Proof. As a special case of [GT12a, Thm. 5.2], for each GP map $g\colon \mathbb{N}_0 \to [0, 1]$ and for all sufficiently large integers N, we have

$$\text{(36)} \qquad \left| \mathbb{E}_{n < N} \mu(n) g(n) \right| = O(\log^{-2} N),$$

where we let $\mu(n)$ denote the Möbius function (i.e., $\mu(n) = 0$ if n is divisible by a square and $\mu(n) = (-1)^k$ if n is the product of k different primes), and the implicit constant is allowed to depend on g. If $\mathbf{1}_\mathcal{P}$ was a bracket word, then (36) would imply that

$$\text{(37)} \qquad |\mathcal{P} \cap [N]| = O(N / \log^2 N),$$

which would contradict the Prime Number Theorem. $\qquad\qquad\square$

Remark 11.12. Note, in particular, that the cited theorem implies that μ is not a bracket word. The same applies, with virtually the same proof, to the Liouville function, which is defined by $\lambda(n) = (-1)^k$ if n is the product of k primes (counting multiplicities). Results from [FH17] allow one to extend this observation to a wider class of multiplicative functions. In a different direction, using the results in [GT12a], one should be able to strengthen the result of Proposition 11.11 by showing that for each bounded GP map $g\colon \mathbb{N}_0 \to \mathbb{R}$, the averages $\mathbb{E}_{p \in \mathcal{P} \cap [N]} g(p)$ converge to the same limit as the averages $\mathbb{E}_{n \in [N]} g(n)$ as $N \to \infty$.

We next consider the set of squares.

Proposition 11.13. *The characteristic word $\mathbf{1}_\mathcal{S}$ of the set of squares is not a bracket word.*

This follows immediately from the following more general result.

Proposition 11.14. *Let $p\colon \mathbb{Z} \to \mathbb{Z}$ be a polynomial with $p(\mathbb{N}_0) \subseteq \mathbb{N}_0$ and $\deg p \geq 2$. Then the set $\{p(n) \mid n \in \mathbb{N}_0\}$ is not a GP subset of $\mathbb{N}_0$.*

Proof. We will prove marginally more, namely that if $p_1, p_2, \dots, p_s\colon \mathbb{Z} \to \mathbb{Z}$ are polynomials with $p_j(\mathbb{N}_0) \subseteq \mathbb{N}_0$ and $\deg p_j \geq 2$ for all j, $1 \leq j \leq r$, then the set

$$\text{(38)} \qquad E = \bigcup_{j=1}^{r} p_j(\mathbb{N}_0)$$

is not a GP subset of $\mathbb{N}_0$. For the sake of contradiction, suppose that the converse is true. Throughout the proof, we use terminology introduced in Appendix A.

Applying Theorem A.1, we can find a nilmanifold $X = G/\Gamma$ with G connected and simply connected, a semialgebraic subset $S \subseteq X$, and a polynomial sequence $g\colon \mathbb{Z} \to G$ such that $E = \{n \in \mathbb{N}_0 \mid g(n)\Gamma \in S\}$.

Let us first consider the case where $(g(n)\Gamma)_{n=0}^{\infty}$ is equidistributed in X. Since $d(E) = 0$, we have $\mu_X(S) = 0$. The polynomial sequence $(\pi_{\mathrm{ab}}(g(n)\Gamma))_{n=0}^{\infty}$ is equidistributed in X_{ab}. Hence, we can conclude from Weyl's equidistribution theorem that

$(\pi_{\mathrm{ab}}(g(p(n)\Gamma)))_{n=0}^{\infty}$ is equidistributed in X_{ab} for each non-constant polynomial map $p\colon \mathbb{N}_0 \to \mathbb{N}_0$. It now follows from Theorem A.2 (that is, [Lei05, Thm. C]) that $(g(p(n)\Gamma))_{n=0}^{\infty}$ is equidistributed in X. In particular $g(p(n))\Gamma \notin S$ for almost all $n \in \mathbb{N}_0$. Taking $p = p_1$ we reach a contradiction.

Suppose next that $(g(n)\Gamma)_{n=0}^{\infty}$ is not equidistributed in X. Then, it follows from [Lei05, Thm. B] that there exist $q \in \mathbb{N}$ and sub-nilmanifolds $Y_0, Y_1, \ldots, Y_{q-1} \subseteq X$ such that $(g(qn+i)\Gamma)_{n=0}^{\infty}$ is equidistributed in Y_i for each $i \in [q]$. It remains to apply the previously considered special case to the sets $(E - i)/q$ for each $i \in [q]$ (note that these sets again take the form (38), possibly for some larger r). $\qquad\square$

11.3. **Number-theoretical functions.** Number theory provides a plentiful source of examples of finitely-valued sequences, for which one can inquire into the existence of a generalised polynomial representation. Multiplicative functions constitute a particularly interesting and well-studied class of sequences, with many applications to other problems. Recall that a sequence $f\colon \mathbb{N} \to \mathbb{C}$ is *multiplicative* if $f(nm) = f(n)f(m)$ for all $n, m \in \mathbb{N}$ with $\gcd(n, m) = 1$. For instance, the Möbius function, mentioned in the previous sub-section, is often used in the study of the prime numbers. The Prime Number Theorem is equivalent to $\frac{1}{N}\sum_{n=1}^{N}\mu(n) \to 0$ as $N \to \infty$, while the more quantitative bound $\left|\sum_{n=1}^{N}\mu(n)\right| = O(N^{1/2+\varepsilon})$ for each $\varepsilon > 0$ is equivalent to the Riemann hypothesis. The results from [GT12a] on correlations of the Möbius function were a crucial ingredient in the work of Green and Tao on linear patterns in the primes [GT10b]. As we pointed out earlier, the Möbius function is not a bracket word, and the same applies to the closely related Liouville function.

While the Möbius function and the Liouville function are bounded, there are many interesting examples of unbounded integer-valued multiplicative functions. In order to obtain finitely-valued sequences, we use reduction modulo a fixed integer. Thus, in this section we investigate infinite words of the form $((f(n) \bmod q)_{n=1}^{\infty}$, where $q \geq 2$ is an integer and $f\colon \mathbb{N} \to \mathbb{Z}$ is multiplicative. Analogous questions, with automatic sequences in place of bracket words, were investigated in [Yaz01]. As an illustrative example, we begin with the Euler totient function ϕ, given by

$$\phi(n) = |(\mathbb{Z}/n\mathbb{Z})^*| = n \prod_{\mathcal{P} \ni p \mid n} \left(1 - \frac{1}{p}\right), \qquad n \in \mathbb{N}.$$

Proposition 11.15. *Let $q \geq 3$ be an integer. Then $(\phi(n) \bmod q)_{n=1}^{\infty}$ is not a bracket word.*

Proof. Let us assume by contradiction that $(\phi(n) \bmod q)_{n=1}^{\infty}$ is a bracket word. Replacing q with a divisor, we may assume that $q = 4$ or q is a prime.

We first assume that $q = 4$. By direct inspection we have:

$$\phi(n) \bmod 4 = \begin{cases} 1 & \text{if } n \in \{1, 2\}; \\ 2 & \text{if } n = 4 \text{ or } n \in \{p, 2p\} \text{ for a prime } p \equiv 3 \bmod 4; \\ 0 & \text{otherwise.} \end{cases}$$

Since $(\phi(n) \bmod 4)_{n=1}^{\infty}$ is a bracket word, we deduce that $\mathcal{P} \cap (4\mathbb{Z} + 3)$ is a GP set. This is impossible by direct repetition of the argument in Proposition 11.11.

Suppose now that q is a prime and set

$$E = \{n \in \mathbb{N} \mid \phi(n) \not\equiv 0 \bmod q\}.$$

Alternatively, E is described by the condition: $n \in E$ if and only if $q^2 \nmid n$ and n has no prime divisors congruent to 1 modulo q. Hence, applying the prime number theorem in arithmetic progressions, we conclude that $d(E) = 0$. By assumption, E is a GP set. Pick a prime $p > q$ with $q \nmid p - 1$. Then $E/p = \{n \in \mathbb{N} \mid pn \in E\} = E$, and hence we infer from Corollary 11.9 that E is not a GP set. This provides a contradiction. $\qquad\square$

Applying the same ideas, we can obtain the following result, which can be used to deal with most other "naturally occurring" multiplicative sequences.

Proposition 11.16. *Let $f \colon \mathbb{N} \to \mathbb{Z}$ be a multiplicative sequence and let $q \in \mathbb{N}$. Suppose that the two following properties hold.*

(i) *There exists $p \in \mathcal{P}$ such that the sequence $(f(p^n) \bmod q)_{n=1}^\infty$ is eventually periodic but not eventually zero.*

(ii) *There exist infinitely many $p \in \mathcal{P}$ such that $f(p^n) \equiv 0 \bmod q$ for some $n \in \mathbb{N}$.*

Then $(f(n) \bmod q)_{n=1}^\infty$ is not a bracket word.

Proof. We argue by contradiction, assuming that $(f(n) \bmod q)_{n=1}^\infty$ is a bracket word. Using Lemma 6.1, we deduce that $(f(n) \bmod p)_{n=1}^\infty$ is also a bracket word for every $p \mid q$. Replacing q with a prime divisor if necessary, we may thus assume that q is prime. Consider the GP set

$$E = \{n \in \mathbb{N} \mid f(n) \not\equiv 0 \bmod q\}.$$

Let p be a prime satisfying the conditions in (i). Let $c \in \mathbb{N}_0$ and $d \in \mathbb{N}$ be such that $f(p^{n+d}) \equiv f(p^n) \bmod q$ for all $n \geq c$ and $f(p^c) \not\equiv 0 \bmod q$. Let E' be the GP set defined by $E' = E/p^c = \{n \in \mathbb{N} \mid p^c n \in E\}$. Then $E'/p^d = E'$.

Next, we show that $d(E) = 0$. Suppose, conversely, that $d(E) > 0$. Then, by Theorem 7.4, E is syndetic, meaning that we can find N such that E intersects any interval $[n, n+N)$ ($n \in \mathbb{N}$). By (ii), we can pick pairwise coprime integers $r_0, r_1, \ldots, r_{N-1}$ with $f(r_i) \equiv 0 \bmod q$ for all $i \in [N]$. By the Chinese remainder theorem, we can find $n \in \mathbb{N}$ such that $n + i \equiv r_i \bmod r_i^2$ for all $i \in [N]$. Then we have $f(n+i) \equiv 0 \bmod q$ and $n + i \notin E$ for all $i \in [N]$, contradicting the defining condition of N. Hence $d(E) = 0$.

Since $d(E) = 0$, we also have $d(E') = 0$. Thus, we infer from Corollary 11.9 that E' is not a GP set, which contradicts previous observations. $\qquad\square$

Example 11.17. Given an integer $k \geq 0$, we let $\sigma_k \colon \mathbb{N} \to \mathbb{N}$ denote the k-th power divisor-sum function defined by

$$\sigma_k(n) = \sum_{d \mid n} d^k = \prod_{p \mid n} \frac{p^{k(\nu_p(n)+1)} - 1}{p^k - 1}, \qquad n \in \mathbb{N},$$

where the sum runs over all positive integers d that divide n. Note that $\sigma_0(n) = d(n)$ is the number of divisors of n and $\sigma_1(n) = \sigma(n)$ is the sum of divisors. For each integer $q \geq 2$, the assumptions of Proposition 11.16 are satisfied: if $k \geq 1$ then in (i) we can take $p = q$, and in (ii) we can take any $p \equiv 1 \bmod q$ and $n = 1$ (infinitude of such primes follows from Dirichlet's theorem). If $k = 0$, both (i) and (ii) hold for all primes p. Hence, $(\sigma_k(n) \bmod q)_{n=1}^\infty$ is not a bracket word.

Example 11.18. Given an integer $k \geq 1$, we let $\tau_k \colon \mathbb{N} \to \mathbb{N}$ denote the number of representations as the product of k positive integers, that is

$$\tau_k(n) = \left| \left\{ (d_1, d_2, \ldots, d_k) \in \mathbb{N}^k \ \middle| \ \prod_{i=1}^{k} d_i = n \right\} \right| = \prod_{p \mid n} \binom{k + \nu_p(n) - 1}{k - 1}, \qquad n \in \mathbb{N}.$$

For each integer $q \geq 2$, the assumptions of Proposition 11.16 are satisfied: both (i) and (ii) hold for all primes p. Hence, $(\tau_k(n) \bmod q)_{n=1}^{\infty}$ is not a bracket word.

Example 11.19. Let $\tau \colon \mathbb{N} \to \mathbb{Z}$ denote the Ramanujan function, and let q be an integer with $\gcd(q, 6) = 1$. For each prime p, the sequence $(\tau(p^n))_{n=1}^{\infty}$ satisfies the recurrence

$$\tau(p^{n+2}) = \tau(p)\tau(p^{n+1}) - p^{11}\tau(p^n),$$

and hence the sequence $(\tau(p^n) \bmod q)_{n=1}^{\infty}$ is eventually periodic. In particular, since $\tau(2) = -24 \not\equiv 0 \bmod q$, in Proposition 11.16 assumption (i) is satisfied for $p = 2$. It is known that the set of $n \in \mathbb{N}$ with $\tau(n) \not\equiv 0 \bmod q$ has asymptotic density 0 [AAB$^+$88]. As a consequence, assumption 11.16(ii) is also satisfied (otherwise, there would exist $P \in \mathbb{N}$ such that $\tau(n) \not\equiv 0 \bmod q$ for all $n \in P\mathbb{N} + 1$). Hence, $(\tau(n) \bmod q)_{n=1}^{\infty}$ is not a bracket word.

We close this section on multiplicative functions with another result, Corollary 11.21, which applies to multiplicative sequences that exhibit non-trivial behaviour only for a finite set of primes. In particular, these sequences do not satisfy condition (ii) of Proposition 11.16. In fact, we can state a somewhat more general result. For a prime p and an integer n, we let $\nu_p(n)$ denote the p-adic valuation of n.

Proposition 11.20. *Let $\ell \in \mathbb{N}$, let $p_1, p_2, \ldots, p_\ell$ be primes, let Σ be a finite alphabet, and let $F \colon \mathbb{N}_0^\ell \to \Sigma$. Let $\mathbf{a}$ be the infinite word defined over Σ by*

$$a_n = F\left(\nu_{p_1}(n), \nu_{p_2}(n), \ldots, \nu_{p_\ell}(n)\right), \qquad n \in \mathbb{N}.$$

Suppose that $\mathbf{a}$ is not almost everywhere periodic. Then $\mathbf{a}$ is not a bracket word.

Proof. By Proposition 11.2, it will suffice to check that $\mathbf{a}$ is weakly periodic. Pick any $k \in \mathbb{N}, r \in \mathbb{N}_0$. We can find k' with $k \mid k'$ such that $\nu_{p_i}(k') > \nu_{p_i}(r)$ for all $1 \leq i \leq \ell$. Thus, $a_{k'n+r} = a_{k'n+k'+r}$ for all $n \in \mathbb{N}_0$, so, with notation as in the definition of weak periodicity, we may take $r' = r$ and $r'' = r + k'$. $\qquad \square$

Corollary 11.21. *Let $f \colon \mathbb{N} \to \mathbb{Z}$ be a multiplicative sequence and let $q \in \mathcal{P}$. Suppose that $(f(n) \bmod q)_{n=1}^{\infty}$ is not periodic and that there exists $p_0 \in \mathcal{P}$ such that $f(p^n) \equiv 1 \bmod q$ for all primes $p \geq p_0$ and all $n \in \mathbb{N}$. Then $(f(n) \bmod q)_{n=1}^{\infty}$ is not a bracket word.*

Proof. Note that the sequence $(f(n) \bmod q)_{n=1}^{\infty}$ is Toeplitz, since, for every positive integer n, one has $f(n + dm) = f(n)$ for all $m \in \mathbb{N}$, for any choice of $d \in \mathbb{N}$ such that $\nu_p(d) > \nu_p(n)$ for all $p < p_0$. Note also that $(f(n) \bmod q)_{n=1}^{\infty}$ is not periodic, and thus it is also not almost everywhere periodic. It remains to apply Proposition 11.20. $\qquad \square$

Example 11.22. The sequence $\left((-1)^{\nu_2(n)+\nu_3(n)}\right)_{n=1}^{\infty}$ is not a bracket word.

11.4. **Automatic sequences.** Let us now consider the problem of classifying automatic sequences which are bracket words. To begin with, we very briefly recall the definition of an automatic sequence; for extensive background, see [AS03].

Let $k \in \mathbb{N}$, let $\Sigma_k = \{0, 1, \ldots, k-1\}$ denote the base-k alphabet, and let Ω be a finite set. A sequence $\mathbf{a}$ over Ω is k-automatic if there exists a finite k-automaton which computes $\mathbf{a}$. More explicitly, this means that there exist a finite set of states Q, a distinguished state $q_0 \in Q$, a transition function $\delta \colon Q \times \Sigma_k \to Q$, and an output map $\tau \colon Q \to \Omega$ such that for each integer $n \in \mathbb{N}_0$ with base-k expansion $n = n_0 + kn_1 + k^2 n_2 + \cdots + k^\ell n_\ell$ we have

$$a_n = \tau \left(\delta \Big(\delta \big(\ldots \delta(\delta(q_0, n_0), n_1) \ldots, n_{\ell-1} \big), n_\ell \Big) \right).$$

It is relatively easy to show that if an automatic sequence $\mathbf{a} = (a_n)_{n=0}^\infty$ coincides with a bracket word $\mathbf{b} = (b_n)_{n=0}^\infty$ almost everywhere, that is, if

$$d \left(\{ n \in \mathbb{N}_0 \mid a_n \neq b_n \} \right) = 0,$$

then $\mathbf{a}$ must also be periodic almost everywhere (i.e., there exists a periodic word $\mathbf{c}$ such that $d \left(\{ n \in \mathbb{N}_0 \mid a_n \neq c_n \} \right) = 0$). Indeed, this result follows from Proposition 11.2 combined with the fact that a k-automatic sequence has a finite k-kernel; see [BK20] for details. Here, the k-kernel of a sequence $\mathbf{a} = (a_n)_{n=0}^\infty$ is defined as the set of subsequences $\left\{ (a_{k^i n + r})_{n=0}^\infty \mid r, i \in \mathbb{N}_0, \ r < k^i \right\}$.

As a consequence of this "density 1" result, in order to classify all automatic sequences which are bracket words, it is enough to consider sequences $\mathbf{a}$ over $\{0, 1\}$, with $\mathrm{freq}(\mathbf{a}, 1) = 0$. Equivalently, we let $\mathbf{a} = \mathbf{1}_E$, where $E \subseteq \mathbb{N}_0$ and $d(E) = 0$. In [BK18], we investigated such "sparse" bracket words and showed a slightly stronger variant of Proposition 11.4. Combining it with facts concerning additive richness of automatic sets, a full classification was obtained in [BK20], conditional on the conjecture that for $k \geq 2$ the set $\left\{ k^i \mid i \in \mathbb{N}_0 \right\}$ of powers of k is not a GP subset of $\mathbb{N}_0$. This conjecture was finally proved in [Kon21], leading to the following result.

Theorem 11.23 ([Kon21, Thm. B]). *Let $\mathbf{a}$ be an automatic sequence that is not eventually periodic. Then $\mathbf{a}$ is not a bracket word.*

Note that Theorem 11.23 extends the classical result claiming that a Sturmian word cannot be automatic.

12. Subword complexity

Since generalised polynomials are defined by relatively simple formulae, it is natural to inquire if they also have low complexity when viewed from different perspectives. In particular, we investigate *subword complexity*, that is, the number $p_\mathbf{a}(N)$ of different length-N factors (or subwords) of an infinite word $\mathbf{a}$. An overview of various results concerning subword complexity can be found in a number of surveys, such as [CN10] or [All94], and connections with the theory of dynamical systems are discussed in [Fer99].

12.1. **Background.** For an infinite word $\mathbf{a} = (a_n)_{n=0}^\infty$ defined over an alphabet Σ, the complexity function of $\mathbf{a}$ is the function that associates with each positive integer n the positive integer

$$p_\mathbf{a}(N) = \left| \{ a_n a_{n+1} \cdots a_{n+N-1} \mid n \in \mathbb{N}_0 \} \right|.$$

The simplest sequences, in terms of subword complexity, are the eventually periodic ones: If $\mathbf{a}$ is eventually periodic then $p_{\mathbf{a}}(N)$ is bounded. Conversely, it was shown by Morse and Hedlund that if there exists N with $p_{\mathbf{a}}(N) \leq N$ then $\mathbf{a}$ must be eventually periodic (see, for instance, [MH38] or [Fer99, Prop. 2]). Thus, for any sequence $\mathbf{a}$, which is not eventually periodic, and every N we have $p_{\mathbf{a}}(N) \geq N+1$. The Sturmian words, discussed in Section 4.2, are characterised by the property that $p_{\mathbf{a}}(N) = N+1$ for all N. Recall that Sturmian words are, in particular, bracket words. At the other extreme, if $|\Sigma| = k$, we have the upper bound $p_{\mathbf{a}}(N) \leq k^N$ for all N, and there exist sequences for which equality holds. (In fact, this is the case for almost all sequences with respect to the natural probability measure on $\Sigma^{\mathbb{N}_0}$). Because of the elementary inequality $p_{\mathbf{a}}(N + M) \leq p_{\mathbf{a}}(N)p_{\mathbf{a}}(M)$, the limit

$$h(\mathbf{a}) = \lim_{N \to \infty} \frac{\log p_{\mathbf{a}}(N)}{N}$$

exists for any finitely-valued sequence $\mathbf{a}$; its value is called the *entropy* of $\mathbf{a}$, and is closely connected to the notion of topological entropy coming from the theory of dynamical systems.

12.2. **New result.** Because bracket words can be represented using nilsystems as in Theorem 4.2 and because nilsystems have zero entropy, one could show[7] that $h(\mathbf{a}) = 0$ for each bracket word $\mathbf{a}$, meaning that $p_{\mathbf{a}}(N) = \exp(o(N))$ as $N \to \infty$. Our main new result concerning the complexity of bracket words is a polynomial bound, given in Theorem A and restated below for the reader's convenience.

Theorem A. *Let $\mathbf{a}$ be a bracket word. Then there exists a constant $C > 0$ such that $p_{\mathbf{a}}(N) = O(N^C)$ for all $N \in \mathbb{N}$.*

Remark 12.1. We note that related results concerning topological complexity of nilsystems were obtained in [HKM14], but are not directly applicable in our context (the crucial obstacle is the fact that the representation in Theorem 4.2 involves a partition into semialgebraic sets S_i which are neither open nor closed, while the results in [HKM14, Sec. 3] are applicable to open covers).

The proof of this result is carried out in Sections 13, 14, and 15.

12.3. **Complementary results.** In view of Theorem A, it is natural to inquire into more precise asymptotics for $p_{\mathbf{a}}(N)$, where $\mathbf{a}$ is a bracket word. Let

$$\lambda(\mathbf{a}) = \limsup_{N \to \infty} \frac{\log p_{\mathbf{a}}(N)}{\log N}$$

be the smallest exponent such that $p_{\mathbf{a}}(N) \leq N^{\lambda(\mathbf{a})+o(1)}$ as $N \to \infty$. Theorem A asserts that $\lambda(\mathbf{a})$ is finite for all bracket words $\mathbf{a}$. The exact value of $\lambda(\mathbf{a})$ is known only in the simplest examples, such as the Sturmian words. Already in relatively simple cases, such as $a_n = [\![\{ \sqrt{2}n \lfloor \sqrt{3}n \rfloor \} > 1/4]\!]$, it would be interesting to compute $\lambda(\mathbf{a})$ exactly.

Theorem 3.4 of [HKM14] asserts that the topological complexity $S(\varepsilon, N)$ of a minimal nilsystem (X, T) obeys the bounds $C(\varepsilon)N^c \leq S(\varepsilon, N) \leq C'(\varepsilon)N^c$, where c is the total commutator dimension of X (see the original paper for the relevant

---

[7]Since we are about to prove a more precise estimate, we do not go into the details of the argument. If in Theorem 4.2 the sets S_i were open, then the conclusion would be a consequence of standard facts about topological entropy, see e.g. [Wal82, Sec. 7.1]. In general, the sets S_i are not open, but their boundaries ∂S_i have zero $\mu_{G/\Gamma}$-measure, which is sufficient for this application.

definitions). Let $\mathbf{a}$ be a bracket word represented on the nilsystem (X, T) as in Theorem 4.2. It seems plausible that, under mild additional assumptions, similar estimates might hold for $p_{\mathbf{a}}(N)$, implying in particular that $\lambda(\mathbf{a}) = c$.

We prove now that $\lambda(\mathbf{a})$ can be arbitrarily large. In fact, for each $d \in \mathbb{N}$, we can construct a bracket word $\mathbf{a}$ with $\lambda(\mathbf{a}) = d$.

Proposition 12.2. *Fix $d \in \mathbb{N}$, and $\alpha_1, \alpha_2, \ldots, \alpha_d \in \mathbb{R}$ such that $1, \alpha_1, \alpha_2, \ldots, \alpha_d$ are $\mathbb{Q}$-linearly independent. For every i, $1 \leq i \leq d$, let $\mathbf{a}^{(i)}$ be the Sturmian word defined by $a_n^{(i)} = \lfloor n\alpha_i \rfloor - \lfloor (n-1)\alpha_i \rfloor$ for $n \in \mathbb{N}_0$, and let $\mathbf{a}$ be the Cartesian product $\prod_{i=1}^{d} \mathbf{a}^{(i)}$. Then $\mathbf{a}$ is a bracket word with $\lambda(\mathbf{a}) = d$.*

Proof. We first observe that, according to Corollary 6.4, and since Sturmian words are bracket words, the word $\mathbf{a}$ is a bracket word.

For each i, $1 \leq i \leq d$, we have $p_{\mathbf{a}^{(i)}}(N) = N + 1$. As a general fact about products, we have $p_{\mathbf{a}}(N) \leq \prod_{i=1}^{d} p_{\mathbf{a}^{(i)}}(N)$. It follows that $\lambda(\mathbf{a}) \leq \sum_{i=1}^{d} \lambda\left(\mathbf{a}^{(i)}\right) = d$.

Conversely, for each i, $1 \leq i \leq d$, and each length-N subword $w^{(i)}$ of $\mathbf{a}^{(i)}$, there exists an interval (not degenerate to a point) $I \subseteq \mathbb{R}/\mathbb{Z}$ such that, for $n \in \mathbb{N}_0$, $\mathbf{a}^{(i)}|_{[n,n+N)} = w^{(i)}$ if and only if $n\alpha_i \bmod \mathbb{Z} \in I$. Since by assumption the numbers $1, \alpha_1, \alpha_2, \ldots, \alpha_d$ are $\mathbb{Q}$-linearly independent, the sequence $(n\alpha_1, n\alpha_2, \ldots, n\alpha_d) \bmod \mathbb{Z}^d$ is equidistributed in $\mathbb{R}^d/\mathbb{Z}^d$. It follows that for any d-tuple $w^{(1)}, w^{(2)}, \ldots, w^{(d)}$ of factors of $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \ldots, \mathbf{a}^{(d)}$ respectively, $\prod_{i=1}^{d} w^{(i)}$ is a factor of $\mathbf{a}$. Hence, $p_{\mathbf{a}}(N) \geq \prod_{i=1}^{d} p_{\mathbf{a}^{(i)}}(N)$, and consequently $\lambda(\mathbf{a}) \geq d$. This ends the proof. $\square$

We end this section with an explicit example showing that $\lambda(\mathbf{a})$ takes on arbitrarily large values for bracket words $\mathbf{a}$, even when the size of the alphabet of $\mathbf{a}$ remains bounded.

Proposition 12.3. *Fix $d \in \mathbb{N}$. Let $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ and let $\mathbf{a}$ be the bracket word over $\{0, 1, \ldots, 9\}$ given by $a_n = \lfloor 10\{\alpha n^d\} \rfloor$. Then $p_{\mathbf{a}}(N) \gg_d N^{d(d-1)/2}$ for all $N \in \mathbb{N}$.*

Proof. For each pair $(m, n) \in \mathbb{Z}^2$, we can compute that $a_{n+m} = \lfloor 10\{h_m(n)\} \rfloor$, where

$$h_m(n) = \sum_{i=0}^{d} \alpha_i(m) n^{d-i} \qquad \text{and} \qquad \alpha_i(m) = \left\{ \alpha m^i \binom{d}{i} \right\}, \qquad 0 \leq i \leq d.$$

Note that $\alpha_0(m) = \{\alpha\}$ for all m. Let $\alpha(m) = (\alpha_i(m))_{i=1}^{d} \in [0,1]^d$, and let $A = \{\alpha(m) \mid m \in \mathbb{Z}\}$. By Weyl's equidistribution theorem, A is dense in $[0,1]^d$.

Let N be a sufficiently large integer. Suppose that for some $m, m' \in \mathbb{N}_0$, $\mathbf{a}|_{[m,m+N)} = \mathbf{a}|_{[m',m'+N)}$. Then

(39) $$\|h_m(n) - h_{m'}(n)\| \leq 1/5, \quad \forall n \in [N].$$

In general, if $\beta = (\beta_0, \beta_1, \ldots, \beta_{d-1}) \in [0,1)^d$ is such that $\left\| \sum_{i=0}^{d-1} \beta_i n^i \right\| \leq 1/5$ for all $n \in [N]$, then it follows from the quantitative version of the Weyl equidistribution theorem that there exists a non-zero integer $q = O_d(1)$ such that $\|q\beta_i\| = O_d(N^{-(d-i)})$ for all i, $0 \leq i \leq d$. (This can be derived using [Vau97, Lemma 2.4], or as a special case of a much more general statement in Theorem A.3.) Applying this observation with $\beta = \alpha(m) - \alpha(m')$, we see that there exists a set $R \subseteq \mathbb{R}^d/\mathbb{Z}^d$, which is a union of $O_d(1)$ rectangles with side lengths $O_d(N^{-(d-i)})$, $1 \leq i \leq d$, such

that (39) implies $\alpha(m) - \alpha(m') \bmod \mathbb{Z}^d \in R$. Note that the measure of R is

$$0 < \lambda(R) = O_d\left(N^{-(d-1)}N^{-(d-2)}\cdots N^0\right) = O_d\left(N^{-d(d-1)/2}\right).$$

Put $r = \lfloor 1/\lambda(R) \rfloor$. We can inductively construct an increasing sequence of integers $m_1, m_2, \ldots, m_r$ with the property that $\alpha(m_j) - \alpha(m_i) \bmod \mathbb{Z}^d \notin R$ for all $1 \le i < j \le r$. Indeed, if $m_1, m_2, \ldots, m_{j-1}$ have already been constructed $(1 \le j \le r)$ then it is enough to pick any m_j with

$$\alpha(m_j) \bmod \mathbb{Z}^d \notin \bigcup_{i=1}^{j-1} (R + \alpha(m_i)),$$

which is possible because the measure of the union above is strictly less than 1 and A is dense in $[0,1]^d$. By construction, all factors $\mathbf{a}|_{[m_i, m_i+N)}$ with $1 \le i \le r$ are distinct. It follows that $p_{\mathbf{a}}(N) \ge r \gg_d N^{d(d-1)/2}$. $\qquad\square$

12.4. **Outline of the proof of Theorem A.** We illustrate our strategy of the proof of Theorem A with a specific example, which already employs most of the tools used in the proof of the general case. Since the subsequent discussion serves only as illustration and motivation, we skip some of the technical details.

Let $\mathbf{a}$ be the bracket word over the alphabet $\{0,1\}$ given by

$$a_n = \left[\!\left[\left\{\sqrt{2}n\left\lfloor\sqrt{3}n\right\rfloor\right\} < \frac{1}{2}\right]\!\right].$$

(Above, we use Iverson bracket, introduced in Section 2.3.) Our goal is to obtain a polynomial bound for $p_{\mathbf{a}}(N)$. We may write $a_n = c(g(n))$, where

$$g(n) = \left\lfloor 2\left\{\sqrt{2}n\left\lfloor\sqrt{3}n\right\rfloor\right\}\right\rfloor,$$

and $c\colon \{0,1\} \to \{0,1\}$ is given by $c(0) = 1$, $c(1) = 0$. Of course, the subword complexity of $\mathbf{a}$ is the same as the subword complexity of $(g(n))_{n=0}^{\infty}$.

Our first step (cf. Example 13.6) is to note that for each $m \in \mathbb{N}_0$ there exist $\alpha, \beta, \gamma, \delta \in [0,1)$ such that for all $n \in \mathbb{N}_0$ we have

$$g(n+m) = \left\lfloor 2\left\{(\sqrt{2}n+\alpha)\left\lfloor\sqrt{3}n+\beta\right\rfloor + \gamma n + \delta\right\}\right\rfloor = \widetilde{g}_{\alpha,\beta,\gamma,\delta}(n).$$

Thus, instead of estimating the number of subwords of $(g(n))_{n=0}^{\infty}$ of a given length N, it will suffice to estimate the number of sequences $(\widetilde{g}_{\alpha,\beta,\gamma,\delta}(n))_{n=0}^{N-1}$, where $(\alpha, \beta, \gamma, \delta)$ varies over $[0,1)^4$. This step is the key reason why we introduce the notion of a parametric GP map in Section 13.2.

For the purpose of counting, it will be more convenient to work with the operation $\lfloor\cdot\rfloor$ rather than $\{\cdot\}$. We can remove $\{\cdot\}$ from the definition of $\widetilde{g}$ by writing

$$\widetilde{g}_{\alpha,\beta,\gamma,\delta}(n) = 2\left\lfloor(\sqrt{2}n+\alpha)\left\lfloor\sqrt{3}n+\beta\right\rfloor + \gamma n + \delta\right\rfloor$$
$$+ \left\lfloor 2\left((\sqrt{2}n+\alpha)\left\lfloor\sqrt{3}n+\beta\right\rfloor + \gamma n + \delta\right)\right\rfloor = 2g'_{\alpha,\beta,\gamma,\delta}(n) - g''_{\alpha,\beta,\gamma,\delta}(n).$$

Thus, to obtain a polynomial bound for $p_{\mathbf{a}}(N)$, it will suffice to obtain polynomial bounds for the number of sequences $\left(g'_{\alpha,\beta,\gamma,\delta}(n)\right)_{n=0}^{N-1}$ and $\left(g''_{\alpha,\beta,\gamma,\delta}(n)\right)_{n=0}^{N-1}$, where $(\alpha, \beta, \gamma, \delta)$ varies over $[0,1)^4$. We only consider the first of these two bounds, the second one being analogous. (Actually, the two classes of sequences are related by $g'_{\alpha,\beta,\gamma,\delta}(n) = \left\lfloor g''_{\alpha,\beta,\gamma,\delta}(n)/2\right\rfloor$, but we do not need this fact.)

Expanding the brackets inside $\lfloor \cdot \rfloor$, we may bring $g'_{\alpha,\beta,\gamma,\delta}(n)$ into a more convenient form

$$g'_{\alpha,\beta,\gamma,\delta}(n) = \left\lfloor \sqrt{2}n \left\lfloor \sqrt{3}n + \beta \right\rfloor + \alpha \left\lfloor \sqrt{3}n + \beta \right\rfloor + \gamma n + \delta \right\rfloor$$

$$= \left\lfloor \sqrt{2}h^{(1)}_{\beta}(n) + \alpha h^{(2)}_{\beta}(n) + \gamma h^{(3)}(n) + \delta h^{(4)}(n) \right\rfloor,$$

where $h^{(1)}_{\beta}(n) = n \left\lfloor \sqrt{3}n + \beta \right\rfloor$, $h^{(2)}_{\beta}(n) = \left\lfloor \sqrt{3}n + \beta \right\rfloor$, $h^{(3)}(n) = n$ and $h^{(4)}(n) = 1$ are integer-valued GP maps, which are strictly simpler than $g'_{\alpha,\beta,\gamma,\delta}(n)$ in a sense that is made precise in Section 13.1. Note also that for each of the maps $h^{(i)}$ and each $n \in [N]$ we have $0 \leq h^{(i)}(n) < 10N^2$.

Suppose that we have already proved a polynomial bound on the number of sequences $\left(h^{(1)}_{\beta}(n) \right)_{n=0}^{N-1}$ and $\left(h^{(2)}_{\beta}(n) \right)_{n=0}^{N-1}$ as β varies over $[0,1)$. (Our proof of Theorem A proceeds by induction, cf. Section 13.3. Here, we omit the discussion of relatively simple and uninteresting cases.) Since the only dependence of $g'_{\alpha,\beta,\gamma,\delta}(n)$ on β is through $h^{(1)}_{\beta}$ and $h^{(2)}_{\beta}$, it will suffice to obtain for each $\beta \in [0,1)$ a polynomial bound for the number of sequences $\left(g'_{\alpha,\beta,\gamma,\delta}(n) \right)_{n=0}^{N-1}$, where (α, γ, δ) varies over $[0,1)^3$. Thus, we have reduced the number of parameters from 4 to 3. Let us fix the choice of $\beta \in [0,1)$, and let $h^{(1)}(n) = h^{(1)}_{\beta}(n)$, $h^{(2)}(n) = h^{(2)}_{\beta}(n)$. We point out that from this point we will no longer need any information about the maps $h^{(i)}$ ($1 \leq i \leq 4$) other than that they map $[N]$ to $[10N^2]$; in particular, we will not use the fact that they are GP (cf. Proposition 15.2).

We are thus left with the task of obtaining a polynomial bound on the number of sequences

$$\left(g'_{\alpha,\beta,\gamma,\delta}(n) \right)_{n=0}^{N-1} = \left(\left\lfloor \sqrt{2}h^{(1)}(n) + \alpha h^{(2)}(n) + \gamma h^{(3)}(n) + \delta h^{(4)}(n) \right\rfloor \right)_{n=0}^{N-1}$$

as α, γ, δ varies over $[0,1)^3$. A natural approach at this point is to approximate α, γ, δ by rational numbers $\alpha^*, \gamma^*, \delta^*$ with denominators Q (to be optimised in the course of the argument) and (say) $0 \leq \alpha - \alpha^*,\ \gamma - \gamma^*, \delta - \delta^* \leq 1/Q$. Indeed, as long as we have a polynomial bound $Q = N^{O(1)}$, the choice of $\alpha^*, \gamma^*, \delta^*$ will only contribute a polynomial factor to our bound on the subword complexity of $\mathbf{a}$. At the same time, for $n \in [N]$ we expect (cf. eq. (50)), at least heuristically, that

$$g'_{\alpha,\beta,\gamma,\delta}(n) = \left\lfloor \sqrt{2}h^{(1)}(n) + \alpha h^{(2)}(n) + \gamma h^{(3)}(n) + \delta h^{(4)}(n) \right\rfloor$$

$$= \left\lfloor \sqrt{2}h^{(1)}(n) + \alpha^* h^{(2)}(n) + \gamma^* h^{(3)}(n) + \delta^* h^{(4)}(n) \right\rfloor = g'_{\alpha^*,\beta,\gamma^*,\delta^*}(n).$$

Let us make that last point somewhat more precise. We have

$$\left| (\alpha - \alpha^*)h^{(2)}(n) + (\gamma - \gamma^*)h^{(3)}(n) + (\delta - \delta^*)h^{(4)}(n) \right| \leq 30N^2/Q.$$

Thus (cf. eq. (52)), $g'_{\alpha,\beta,\gamma,\delta}(n) = g'_{\alpha^*,\beta,\gamma^*,\delta^*}(n)$ as long as we have

(40) $$\left\| \sqrt{2}h^{(1)}(n) + \alpha h^{(2)}(n) + \gamma h^{(3)}(n) + \delta h^{(4)}(n) \right\| > 30N^2/Q.$$

Next, we will need to better understand the situation where (40) does not hold. It will be convenient to define (cf. eq. (51))

$$h^{(0)}(n) = -\left\lfloor \sqrt{2}h^{(1)}(n) + \alpha^* h^{(2)}(n) + \gamma^* h^{(3)}(n) + \delta^* h^{(4)}(n) \right\rfloor.$$

Then we have the bound $0 \geq h^0(n) > -50N^2$, and we may express (40) in a slightly more convenient form

$$\left| h^{(0)}(n) + \sqrt{2}h^{(1)}(n) + \alpha h^{(2)}(n) + \gamma h^{(3)}(n) + \delta h^{(4)}(n) \right| > 30N^2/Q.$$

Let $B \subseteq \mathbb{Z}^5$ denote the set of vectors $m = (m_0, m_1, m_2, m_3, m_4)$ with $0 \geq m_0 > -50N^2$ and $0 \leq m_1, m_2, m_3, m_4 < 10N^2$ (thus, $\left(h^{(0)}(n), h^{(1)}(n), \ldots, h^{(4)}(n)\right) \in B$ for $n \in [N]$). The set of vectors $m = (m_0, m_1, m_2, m_3, m_4)$ such that

$$(41) \qquad \left| m_0 + \sqrt{2}m_1 + \alpha m_2 + \gamma m_3 + \delta m_4 \right| \leq 30N^2/Q$$

can be thought of as a discrete approximation of a hyperplane, and thus we expect it to be additively structured. Indeed, we show in Section 14 (Proposition 14.4) that there exists a lattice Λ with the following properties:

- If $m \in B$ and (41) holds then $m \in \Lambda$.
- If $m \in B \cap \Lambda$ then

$$(42) \qquad \left| m_0 + \sqrt{2}m_1 + \alpha m_2 + \gamma m_3 + \delta m_4 \right| \leq CN^{12}/Q$$

  for some absolute constant $C > 0$.

Consider $n \in [N]$ such that $\left(h^{(0)}(n), h^{(1)}(n), \ldots, h^{(4)}(n)\right) \in \Lambda$. Then (cf. eq. (53))

$$g'_{\alpha,\beta,\gamma,\delta}(n) = -h^{(0)}(n) + \left\lfloor h^{(0)}(n) + \sqrt{2}h^{(1)}(n) + \alpha h^{(2)}(n) + \gamma h^{(3)}(n) + \delta h^{(4)}(n) \right\rfloor,$$

where the expression under the $\lfloor \cdot \rfloor$ is bounded in absolute value by CN^{12}/Q. We will take $Q > CN^{12}$, meaning that

$$(43) \qquad g'_{\alpha,\beta,\gamma,\delta}(n) = -h^{(0)}(n), \qquad \text{or} \qquad g'_{\alpha,\beta,\gamma,\delta}(n) = -h^{(0)}(n) - 1.$$

Next, we need to develop a better understanding of when each of the two possibilities mentioned above occurs. Put

$$\Lambda^+ = \left\{ m \in \Lambda \ \middle| \ m_0 + \sqrt{2}m_1 + \alpha m_2 + \gamma m_3 + \delta m_4 \geq 0 \right\},$$

$$\Lambda^- = \left\{ m \in \Lambda \ \middle| \ m_0 + \sqrt{2}m_1 + \alpha m_2 + \gamma m_3 + \delta m_4 < 0 \right\}.$$

Then $\Lambda = \Lambda^+ \cup \Lambda^-$ is a partition obtained by cutting Λ with a hyperplane. We can now make (43) more precise (cf. eq. (54)):

$$(44) \qquad g'_{\alpha,\beta,\gamma,\delta}(n) = \begin{cases} -h^{(0)}(n) & \text{if } \left(h^{(0)}(n), h^{(1)}(n), \ldots, h^{(4)}(n)\right) \in \Lambda^+, \\ -h^{(0)}(n) - 1 & \text{if } \left(h^{(0)}(n), h^{(1)}(n), \ldots, h^{(4)}(n)\right) \in \Lambda^-. \end{cases}$$

Recall from (40) that if $\left(h^{(0)}(n), h^{(1)}(n), \ldots, h^{(4)}(n)\right) \notin \Lambda$ then $g'_{\alpha,\beta,\gamma,\delta}(n) = g'_{\alpha^*,\beta,\gamma^*,\delta^*}(n)$. Combining this observation with (44), we see that the sequence $\left(g'_{\alpha,\beta,\gamma,\delta}(n)\right)_{n=0}^{N-1}$ is completely determined by the following data:

- the rational approximations $\alpha^*, \gamma^*, \delta^*$;
- the restrictions of the sequences $h_\beta^{(1)}$ and $h_\beta^{(2)}$ to $[N]$;
- the intersection of the lattice Λ with the box B;
- the hyperplane partition $\Lambda \cap B = (\Lambda^+ \cap B) \cup (\Lambda^- \cap B)$.

As we have seen, we may take Q of the form $Q = C'N^{12}$ for a large constant C'. Then the number of choices of $\alpha^*, \gamma^*, \delta^*$ is $O(N^{36})$, and hence polynomial in N. The contribution from $h_\beta^{(1)}$ and $h_\beta^{(2)}$ is polynomial by the inductive assumption. Thus, it remains to estimate the number of ways in which the set B can be partitioned into the components $B = (B \setminus \Lambda) \cup (\Lambda^+ \cap B) \cup (\Lambda^- \cap B)$, as described above. A polynomial estimate on the number of such partitions is obtained in Section 14 (Proposition 14.3) using the techniques of additive geometry.

13. Proof of Theorem A: Notation and induction

In this section we set up the inductive scheme and introduce notation which will be used in the proof of Theorem A.

13.1. **Height.** One of several measures of complexity of a generalised polynomial is the *height*, that is, the number of nested instances of the floor function. Let $d \in \mathbb{N}$, and let GP_0 denote the polynomial maps from $\mathbb{R}^d$ to $\mathbb{R}$. Inductively, for each $i \in \mathbb{N}$, let GP_i be the smallest class of maps from $\mathbb{R}^d$ to $\mathbb{R}$ that is closed under sums and products, and which contains g and $\lfloor g \rfloor$ for each $g \in \mathrm{GP}_{i-1}$. By definition, for each GP map $g \colon \mathbb{R}^d \to \mathbb{R}$, there exists an integer $i \in \mathbb{N}_0$ such that $g \in \mathrm{GP}_i$. The height of a GP map $g \colon \mathbb{R}^d \to \mathbb{R}$, denoted by $\mathrm{ht}(g)$, is the least of such integers $i \in \mathbb{N}_0$.

More generally, if $\Omega \subseteq \mathbb{R}^d$ and $g \colon \Omega \to \mathbb{R}$ is a GP map then $\mathrm{ht}(g)$ is defined as the least possible value of $\mathrm{ht}(\widetilde{g})$ where $\widetilde{g} \colon \mathbb{R}^d \to \mathbb{R}$ is a GP map and $\widetilde{g}|_\Omega = g$. For instance, if g is the GP map $\mathbb{Z} \to \mathbb{Z} \subseteq \mathbb{R}$ given by $g(n) = \left\lfloor \sqrt{2}n \lfloor \sqrt{3}n \rfloor + \sqrt{5}n^2 \right\rfloor$ then $\mathrm{ht}(g) \leq 2$. For the following result, we recall that by convention we treat the empty product $\prod_{j=1}^0 (\cdots)$ as being identically equal to 1.

Lemma 13.1. *Let $d \in \mathbb{N}$ and let $g \colon \mathbb{R}^d \to \mathbb{R}$ be a GP map. Then g can be written as*

$$(45) \qquad g = \sum_{i=1}^s p_i \prod_{j=1}^{r_i} \lfloor h_{i,j} \rfloor,$$

where $s \in \mathbb{N}$, $p_i \colon \mathbb{R}^d \to \mathbb{R}$ $(1 \leq i \leq s)$ are polynomials, and for each $1 \leq i \leq s$, $r_i \in \mathbb{N}_0$ and $h_{i,j} \colon \mathbb{R}^d \to \mathbb{R}$ $(1 \leq j \leq r_i)$ are GP maps with $\mathrm{ht}(h_{i,j}) < \mathrm{ht}(g)$.

Proof. We proceed by induction on $\mathrm{ht}(g)$. If g is a polynomial, there is nothing to prove, since we can take $s = 1$ and $p_1 = g$. If $g = \lfloor h \rfloor$ for some h with $\mathrm{ht}(h) < \mathrm{ht}(g)$ then we can take $s = 1$, $r_1 = 1$, $p_1 = 1$, and $h_{1,1} = h$. If $g = h + h'$ or $g = h \cdot h'$ for some h and h' with $\max\{\mathrm{ht}(h), \mathrm{ht}(h')\} < \mathrm{ht}(g)$, then a representation of g of the form (45) can be obtained from the analogous representations of h and h'. $\square$

13.2. **Parametric generalised polynomials.** It will be convenient to state some of our results in terms of *parametric GP maps*, by which we mean families of GP maps which include real-valued parameters as coefficients. For instance, the formula

$$g_{\alpha,\beta}(n) = \left\lfloor \alpha n \lfloor \beta n \rfloor + \sqrt{2}n^2 \right\rfloor$$

defines, for each $\alpha, \beta \in \mathbb{R}$, a GP map from $\mathbb{Z}$ to $\mathbb{R}$, and we will refer to $g_\bullet$ as a parametric GP map[8] $\mathbb{Z} \to \mathbb{R}$. We make this notion precise in Definition 13.2.

---

[8]We use "$\bullet$" as a placeholder for a variable. Thus, in the discussion above, we let $g_\bullet$ denote the map $\mathbb{R}^2 \to \mathbb{R}^{\mathbb{Z}}$, $(\alpha, \beta) \mapsto g_{\alpha,\beta}$, which can be identified with a map $\mathbb{R}^2 \times \mathbb{Z} \to \mathbb{R}$ in a natural way.

Below and elsewhere, if $I \subseteq J$ are finite sets and $\beta \in \mathbb{R}^J$, we let $\beta|I \in \mathbb{R}^I$ denote the restriction of β to I, that is, $(\beta|I)_i = \beta_i$ for all $i \in I$.

Definition 13.2. Let I be a finite set. A *parametric GP map* $g_\bullet \colon \mathbb{Z} \to \mathbb{R}$ *with index set* I is a map $\mathbb{R}^I \to \mathbb{R}^{\mathbb{Z}}$, $\alpha \mapsto g_\alpha$, such that the combined map $\mathbb{R}^I \times \mathbb{Z} \to \mathbb{R}$, $(\alpha, n) \mapsto g_\alpha(n)$ is a GP map. The height of the parametric GP map $g_\bullet$, denoted by $\mathrm{ht}(g_\bullet)$, is the height of the corresponding GP map $(\alpha, n) \mapsto g_\alpha(n)$.

Definition 13.3. Let $g_\bullet, h_\bullet \colon \mathbb{Z} \to \mathbb{R}$ be two parametric GP maps with index sets I and J respectively. We define the sum $g_\bullet + h_\bullet$ to be the parametric GP map with index set $I \cup J$ given by $(g + h)_\alpha(n) = g_{\alpha|I}(n) + h_{\alpha|J}(n)$, $\alpha \in \mathbb{R}^{I \cup J}$, $n \in \mathbb{Z}$. The product $g_\bullet \cdot h_\bullet$ is defined accordingly by $(g \cdot h)_\alpha(n) = g_{\alpha|I}(n) \cdot h_{\alpha|J}(n)$, $\alpha \in \mathbb{R}^{I \cup J}$, $n \in \mathbb{Z}$.

Our interest in parametric GP maps stems largely from Lemma 13.4, which allows us to replace the study of subwords in GP sequences by the study of prefixes in parametric GP sequences.

Lemma 13.4. *Let* $g \colon \mathbb{Z} \to \mathbb{R}$ *be a bounded GP map. Then there exists a parametric GP map* $\widetilde{g}_\bullet$ *with index set* I, *such that, for each* $m \in \mathbb{Z}$, *there exists* $\alpha \in [0, 1)^I$ *such that* $g(n + m) = \widetilde{g}_\alpha(n)$ *for all* $n \in \mathbb{Z}$.

Proof. This is just a rephrasing of Lemma A.5. □

In Examples 13.5 and 13.6, we use, as in Section 5, the Iverson bracket convention: $[\![\varphi]\!] = 1$ if φ is a true sentence and $[\![\varphi]\!] = 0$ otherwise.

Example 13.5. Let $d \in \mathbb{N}$ and $g(n) = [\![\{\sqrt{2}n^d\} < \frac{1}{2}]\!]$. Pick any $m \in \mathbb{Z}$. Then $g(n + m) = \widetilde{g}_\alpha(n)$ for all $n \in \mathbb{Z}$, where $\alpha = (\alpha_i)_{i=0}^d$ is given by $\alpha_i = \left\{ \sqrt{2}m^i \binom{d}{i} \right\}$ $(0 \le i \le d)$ and $\widetilde{g}_\bullet$ is given by

$$\widetilde{g}_\alpha(n) = \left[\!\!\left[\left\{ \sum_{i=0}^d \alpha_i n^{d-i} \right\} < \frac{1}{2} \right]\!\!\right], \qquad \alpha \in \mathbb{R}^{d+1}, \ n \in \mathbb{Z}.$$

Example 13.6. Let $g(n) = [\![\{\sqrt{2}n \lfloor \sqrt{3}n \rfloor\} < \frac{1}{2}]\!]$. Pick any $m \in \mathbb{Z}$ and let $a = \lfloor \sqrt{2}m \rfloor$, $\alpha = \sqrt{2}m$, $b = \lfloor \sqrt{3}m \rfloor$, $\beta = \{\sqrt{3}m\}$, $\gamma = \{\sqrt{2}b\}$, and $\delta = \{\alpha b\}$. Then

$$g(n + m) = \left[\!\!\left[\left\{ \left(\sqrt{2}n + a + \alpha \right) \left(\lfloor \sqrt{3}n + \beta \rfloor + b \right) \right\} < \frac{1}{2} \right]\!\!\right]$$

$$= \left[\!\!\left[\left\{ (\sqrt{2}n + \alpha) \lfloor \sqrt{3}n + \beta \rfloor + \sqrt{2}bn + \alpha b \right\} < \frac{1}{2} \right]\!\!\right]$$

$$= \left[\!\!\left[\left\{ (\sqrt{2}n + \alpha) \lfloor \sqrt{3}n + \beta \rfloor + \gamma n + \delta \right\} < \frac{1}{2} \right]\!\!\right]$$

$$= \widetilde{g}_{\alpha, \beta, \gamma, \delta}(n),$$

where $\widetilde{g}_\bullet$ is the parametric GP map with index set $\{1, 2, 3, 4\}$ given by

$$\widetilde{g}_{\alpha_1, \alpha_2, \alpha_3, \alpha_4}(n) = \left[\!\!\left[\left\{ (\sqrt{2}n + \alpha_1) \lfloor \sqrt{3}n + \alpha_2 \rfloor + \alpha_3 n + \alpha_4 \right\} < \frac{1}{2} \right]\!\!\right].$$

When the dependence of a parametric GP map on the parameters becomes too complicated, it is often more convenient to instead work with a parametric GP map

which has more parameters but depends on them in a simpler way. For instance, if $g_\bullet$ is a parametric GP map of the form

$$(46) \qquad g_\alpha(n) = f_1(\alpha)h_\alpha^{(1)}(n) + f_2(\alpha)h_\alpha^{(2)}(n) + f_3(\alpha)h_\alpha^{(3)}(n),$$

where f_1, f_2, and f_3 are GP maps and $h_\bullet^{(1)}$, $h_\bullet^{(2)}$, and $h_\bullet^{(3)}$ are parametric GP maps, then it might be preferable to instead work with the parametric GP map

$$(47) \qquad g'_{\alpha,\beta}(n) = \beta_1 h_\alpha^{(1)}(n) + \beta_2 h_\alpha^{(2)}(n) + \beta_3 h_\alpha^{(3)}(n).$$

We make this idea precise in Definition 13.7.

Definition 13.7. Let $g_\bullet$ and $h_\bullet$ be two parametric GP maps with index sets I and J respectively. Then we say that $h_\bullet$ *extends* $g_\bullet$, denoted $h_\bullet \succeq g_\bullet$, if there exists a GP map $\varphi \colon \mathbb{R}^I \to \mathbb{R}^J$ such that $g_\alpha = h_{\varphi(\alpha)}$ for all $\alpha \in \mathbb{R}^I$.

It is routine to check that the relation $\succeq$ defined above is a partial order.

Example 13.8. If $g_\bullet$ and $g'_\bullet$ are respectively given by (46) and (47), then $g'_\bullet \succeq g_\bullet$. One can take $\varphi(\alpha) = (\alpha, f_1(\alpha), f_2(\alpha), f_3(\alpha))$.

13.3. **Induction scheme.** Using the terminology introduced above, we are ready to explain the induction scheme that will be used in the proof of Theorem A. It can be construed as an analogue of the inductive definition of generalised polynomials, but restricted to $\mathbb{Z}$-valued sequences. Note that GP maps from $\mathbb{Z}$ to $\mathbb{R}$ can be identified with parametric GP maps with an empty index set.

Proposition 13.9. *Let $\mathcal{G}$ be a family of parametric GP maps from $\mathbb{Z}$ to $\mathbb{Z}$ with index sets contained in $\mathbb{N}_0$. Suppose that $\mathcal{G}$ has the following closure properties.*

(i) *All GP maps $\mathbb{Z} \to \mathbb{Z}$ belong to $\mathcal{G}$.*

(ii) *For every $g_\bullet$ and $h_\bullet \in \mathcal{G}$, it holds that $g_\bullet + h_\bullet \in \mathcal{G}$ and $g_\bullet \cdot h_\bullet \in \mathcal{G}$.*

(iii) *For every $g_\bullet \in \mathcal{G}$, $\mathcal{G}$ contains all the parametric GP maps $g'_\bullet \colon \mathbb{Z} \to \mathbb{Z}$ satisfying $g_\bullet \succeq g'_\bullet$.*

(iv) *For every pair of disjoint finite sets $I \subseteq \mathbb{N}$, $J \subseteq \mathbb{N}$, and every sequence of parametric GP maps $h_\bullet^{(i)} \in \mathcal{G}$, $i \in I$, with index set J, $\mathcal{G}$ contains the parametric GP map $g_\bullet$ defined by*

$$g_{\alpha,\beta}(n) = \left\lfloor \sum_{i \in I} \alpha_i h_\beta^{(i)}(n) \right\rfloor, \qquad n \in \mathbb{Z}, \ \alpha \in \mathbb{R}^I, \ \beta \in \mathbb{R}^J.$$

Then $\mathcal{G}$ contains all parametric GP maps $\mathbb{Z} \to \mathbb{Z}$ with index sets contained in $\mathbb{N}_0$.

Proof. Since each parametric GP map from $\mathbb{Z}$ to $\mathbb{Z}$ takes the form $\lfloor g_\bullet \rfloor$ for some parametric GP map $g_\bullet \colon \mathbb{Z} \to \mathbb{R}$, it suffices to show that $\lfloor g_\bullet \rfloor \in \mathcal{G}$ for each parametric GP $g_\bullet$ with index set $I \subseteq \mathbb{N}_0$. We proceed by induction on $\mathrm{ht}(g_\bullet)$.

Suppose first that $\mathrm{ht}(g_\bullet) = 0$, that is, $g_\bullet$ is a polynomial. Expanding, we can write

$$g_\alpha(n) = \sum_{i=0}^{d} q_i(\alpha) n^i, \qquad n \in \mathbb{Z}, \ \alpha \in \mathbb{R}^I,$$

where $d \in \mathbb{N}_0$ and $q_i \colon \mathbb{R}^I \to \mathbb{R}$ are polynomials. It follows from (i) and (iv) that $\mathcal{G}$ also contains the parametric GP map with index set $\{0, 1, \ldots, d\}$ given by

$$g'_\alpha(n) = \left\lfloor \sum_{i=0}^{d} \alpha_i n^i \right\rfloor, \qquad n \in \mathbb{Z}, \ \alpha \in \mathbb{R}^{d+1}.$$

Since $g'_\bullet \succeq \lfloor g_\bullet \rfloor$, it follows from (iii) that $g_\bullet \in \mathcal{G}$.

Suppose next that $\mathrm{ht}(g_\bullet) \geq 1$. Using Lemma 13.1 and expanding out the polynomial contributions, we can represent $g_\bullet$ in the form

$$g_\beta(n) = \sum_{i=1}^{s} q_i(\beta)\widetilde{h}_\beta^{(i)}(n), \qquad n \in \mathbb{Z},\ \beta \in \mathbb{R}^J,$$

where for every i, $1 \leq i \leq s$, q_i is a polynomial, and $\widetilde{h}_\bullet^{(i)}$ is a parametric GP map of the form

$$\widetilde{h}_\beta^{(i)}(n) = n^{d_i} \prod_{j=1}^{r_i} \left\lfloor h_\beta^{(i,j)}(n) \right\rfloor \qquad n \in \mathbb{Z},\ \beta \in \mathbb{R}^J,$$

where $d_i \in \mathbb{N}_0$, $r_i \in \mathbb{N}_0$, and $\mathrm{ht}(h_\bullet^{(i,j)}) < \mathrm{ht}(g_\bullet)$ for all j, $1 \leq j \leq r_i$. By the inductive assumption, for each pair (i,j), with $1 \leq i \leq s$ and $1 \leq j \leq r_i$, we have $\left\lfloor h_\bullet^{(i,j)} \right\rfloor \in \mathcal{G}$. Consequently, applying (i) and (ii), we conclude that also $\widetilde{h}_\bullet^{(i)} \in \mathcal{G}$.

For notational convenience, assume that $\min J > s$ and put $I = \{1, \ldots, s\}$ (since the ordering of the parameters does not play any role, this does not decrease the level of generality). It follows from (iv) that $\mathcal{G}$ also contains the parametric GP map with index set $I \cup J$ that is defined by

$$g'_{\alpha,\beta}(n) = \left\lfloor \sum_{i=1}^{s} \alpha_i \widetilde{h}_\beta^{(i)}(n) \right\rfloor, \qquad n \in \mathbb{Z},\ \alpha \in \mathbb{R}^I,\ \beta \in \mathbb{R}^J.$$

Since $g'_\bullet \succeq \lfloor g_\bullet \rfloor$, it follows from (iii) that $\lfloor g_\bullet \rfloor \in \mathcal{G}$. This ends the proof. $\qquad\square$

14. Proof of Theorem A: Auxiliary results

In this section we discuss some results in additive combinatorics and Diophantine approximation which will be used in the course of the proof of Theorem A.

14.1. Additive geometry.
Let $r \in \mathbb{N}$. A *symmetric generalised arithmetic progression* of rank r in an abelian group Z with steps $x_1, x_2, \ldots, x_r \in Z$ and side lengths $\ell_1, \ell_2, \ldots, \ell_r \in \mathbb{N}_0$ is defined as the set

$$\mathsf{P}(x_1, x_2, \ldots x_r; \ell_1, \ell_2, \ldots, \ell_r) = \left\{ \sum_{i=1}^{r} n_i x_i \ \middle|\ -\ell_i < n_i < \ell_i \text{ for all } i,\ 1 \leq i \leq r \right\}.$$

Let $d \in \mathbb{N}$. By a *lattice* in $\mathbb{R}^d$ we mean a discrete subgroup of $\mathbb{R}^d$. In particular, we do not require Λ to have full rank, which is slightly non-standard but consistent (in particular with [TV06]). If $\Lambda < \mathbb{R}^d$ is a lattice of full rank, we let $\mathrm{covol}\,\Lambda = \mathrm{vol}(\mathbb{R}^d/\Lambda)$ denote its *covolume* (that is, the measure of a fundamental domain of Λ). By a *convex body* in $\mathbb{R}^d$, we mean a convex, open, non-empty, and bounded set, and by a *half-space*, we always mean a closed half-space, that is, a set of the form

$$\mathsf{H}(x,t) = \left\{ y \in \mathbb{R}^d \ \middle|\ x \cdot y \geq t \right\}$$

for some $x \in \mathbb{R}^d$ and $t \in \mathbb{R}$.

The classical John's theorem asserts that each symmetric convex body B in $\mathbb{R}^d$ can be efficiently approximated by an ellipsoid E, in the sense that $E \subseteq B \subseteq \sqrt{d}E$. We will use a discrete analogue of this result.

Theorem 14.1 (Discrete John's theorem, [TV06, Thm. 3.36])**.** *Let $d \in \mathbb{N}$, $B \subseteq \mathbb{R}^d$ be a symmetric convex body, and $\Lambda < \mathbb{R}^d$ be a lattice of rank r. Then there exist vectors $v_1, v_2, \ldots, v_r \in \mathbb{R}^d$ and integers $\ell_1, \ell_2, \ldots, \ell_r \in \mathbb{N}$ such that*

$$\mathsf{P}(v_1, v_2, \ldots, v_r; \ell_1, \ell_2, \ldots, \ell_r) \subseteq B \cap \Lambda$$
$$\subseteq \mathsf{P}(v_1, v_2, \ldots, v_r; r^{2r}\ell_1, r^{2r}\ell_2, \ldots, r^{2r}\ell_r).$$

We also mention an estimate on the number of partitions of a finite set using a hyperplane.

Theorem 14.2 ([Har67, Thm. 1])**.** *Let d and n be two positive integers, and let $S \subseteq \mathbb{R}^d$ be a set with $|S| = n$. Then the number of sets of the form $S \cap H$, where H is a half-space, is at most $2\sum_{i=0}^{d} \binom{n-1}{i}$.*

In the course of the argument, we will need estimates on the number of sets that can be obtained from a given convex body in $\mathbb{R}^d$ by intersecting it with a sublattice of $\mathbb{Z}^d$ and a half-space.

Proposition 14.3. *Let $d \in \mathbb{N}$, $\Gamma < \mathbb{R}^d$ be a lattice, and $B \subseteq \mathbb{R}^d$ be a symmetric convex body. Then the number of pairs of sets of the form $(\Lambda \cap B, \Lambda \cap B \cap H)$, where $\Lambda < \Gamma$ is a lattice and H is a half-space, belongs to $O_d(|B \cap \Gamma|^{2d})$.*

Proof. Put $M = |B \cap \Gamma|$. Let $\Lambda < \Gamma$ be a sublattice. It follows from Theorem 14.1 that there exist vectors $v_1, \ldots, v_d \in \Lambda \cap B$ such that $\Lambda \cap B \subseteq \mathrm{span}_{\mathbb{Z}}\{v_1, \ldots, v_d\}$. As a consequence,

$$\Lambda \cap B = \mathrm{span}_{\mathbb{Z}}\{v_1, \ldots, v_d\} \cap B.$$

In particular, the set $\Lambda \cap B$ is completely determined by the vectors $v_1, \ldots, v_d$ which belong to $B \cap \Gamma$, and hence can be chosen in at most M^d ways.

For a fixed choice of Λ, we infer from Theorem 14.2 that the number of sets of the form $\Lambda \cap B \cap H$, where H is a half-space, belongs to $O_d(M^d)$. Combining these two estimates yields the claim. □

14.2. **Diophantine approximation.** In this subsection, we briefly discuss approximate linear relations with integer coefficients. Let $d \in \mathbb{N}$. Given a vector $\alpha = (\alpha_i)_{i=1}^{d} \in \mathbb{R}^d$, $\varepsilon > 0$, and $N \in \mathbb{N}$, we set

$$(48) \qquad \mathcal{R}_N(\alpha, \varepsilon) = \left\{ n \in \mathbb{Z}^d \ \middle| \ \|n\|_{\infty} < N, \ \left| \sum_{i=1}^{d} n_i \alpha_i \right| < \varepsilon \right\}.$$

It is a standard fact (see e.g. [Cas72, Thm. VI]) that $\mathcal{R}_N(\alpha, \varepsilon) \neq \{0\}$ when $\varepsilon < \|\alpha\|_{\infty} N^{-(d-1)}$. In our application, we will be interested in the case where ε is a large negative power of N. In this regime, the sets $\mathcal{R}_N(\alpha, \varepsilon)$ can be approximated by lattices in a sense that is made precise by Proposition 14.4.

Proposition 14.4. *Let $d \in \mathbb{N}$. There exists a positive real number C_d such that for each $\alpha = (\alpha_i)_{i=1}^{d} \in \mathbb{R}^d$, $\varepsilon > 0$, and $N \in \mathbb{N}$, there exists a lattice $\Lambda = \Lambda(\alpha, \varepsilon, N)$ satisfying*

$$(49) \qquad \mathcal{R}_N(\alpha, \varepsilon) \subseteq \Lambda \cap (-N, N)^d \subseteq \mathcal{R}_N(\alpha, C_d N^d \varepsilon).$$

For later reference, we let $\Lambda(\alpha, \varepsilon, N)$ denote a lattice satisfying (49). The remaining part of this section is devoted to the proof of Proposition 14.4. We begin with a lemma concerning finite groups, where the situation is simpler. Recall that, for an abelian group G and $g \in G$, the order $\mathrm{ord}(g)$ is the least $q \in \mathbb{N}$ with $qg = e_G$,

the order of G is the cardinality of G, and the exponent of G is the least common multiple of $\mathrm{ord}(g)$ for $g \in G$. For a set $S \subseteq G$ an and $k \in \mathbb{N}$, we let kS denote the k-fold sumset $S + S + \cdots + S = \{a_1 + a_2 + \cdots + a_k \mid a_i \in S\}$.

Lemma 14.5. *Let G be a finite abelian group with order $M \geq 2$ and exponent q, and let $S \subseteq G$ be a generating set with $e_G \in S$. Then there exists $k \leq q \log M / \log q$ such that $kS = G$.*

Proof. We proceed by induction on M. Pick any $g \in S \setminus \{e_G\}$ and let $\ell = \mathrm{ord}(g)$. Since G is non-trivial, we may assume that $\ell \geq 2$. Put $f(x) = x/\log x$ $(x > 1)$ and note that $f(n) \leq f(m)$ for all pairs of integers (n, m) with $m \geq n \geq 2$ and $n \mid m$. Our goal is to find $k \leq f(q) \log M$ with $kS = G$.

Let $\overline{G} = G/\langle g \rangle$, and let $\pi \colon G \to \overline{G}$ be the quotient map. Then G has order $\overline{M} = M/\ell$ and exponent $\overline{q}$ which divides q. The set $\overline{S} = \pi(S)$ generates $\overline{G}$. If $\overline{G}$ is trivial then G is the cyclic group generated by g, and hence $q = M$ and $kS = G$ for $k = f(q) \log M = M$, as needed. (Note that in this step we use the assumption that $e_G \in S$.) Suppose next that $\overline{G}$ is non-trivial, meaning that $\overline{M} \geq 2$. By the inductive assumption, there exists $\overline{k} \leq f(\overline{q}) \log \overline{M}$ with $\overline{k}\,\overline{S} = \overline{G}$. It follows that $kS = G$, where $k = \overline{k} + \ell$. We can estimate

$$k \leq f(\overline{q}) \log \overline{M} + \ell = f(q) \log M - (f(q) - f(\overline{q})) \log M + (f(\ell) - f(\overline{q})) \log \ell.$$

Thus, it is enough to show that

$$(f(q) - f(\overline{q})) \log M \geq (f(\ell) - f(\overline{q})) \log \ell$$

which follows from the observations that $M \geq \ell$, $\ell \mid q$, and $\overline{q} \mid q$. $\qquad\square$

Lemma 14.6. *Let $d \in \mathbb{N}$, $B \subseteq \mathbb{R}^d$ be a symmetric convex body, and $S \subseteq B$ be a symmetric set with $0 \in S$. Put $\Lambda = \mathrm{span}_{\mathbb{Z}} S$ and assume that Λ is a lattice. Then there exists $k = O_d(|\Lambda \cap B|)$ such that $kS \cap B = \Lambda \cap B$.*

Proof. Put $M = |\Lambda \cap B|$. We proceed by induction on d, including the degenerate case $d = 0$ for which one can take $k = 1$. Assume that $d \geq 1$ and that the claim is proved for $d - 1$. We may also assume without loss of generality that Λ has full rank, since otherwise we could replace $\mathbb{R}^d$ with the subspace spanned by Λ.

Applying Theorem 14.1, we conclude that there exist vectors $w_1, w_2, \ldots, w_d \in \Lambda$ and side lengths $\ell_1, \ell_2, \ldots, \ell_d \in \mathbb{N}$ such that

$$\mathsf{P}(w_1, w_2, \ldots, w_d; \ell_1, \ell_2, \ldots, \ell_d) \subseteq B \cap \Lambda \subseteq \mathsf{P}(w_1, w_2, \ldots, w_d; D\ell_1, D\ell_2, \ldots, D\ell_d),$$

where $D = d^{2d} = O_d(1)$. Since Λ has rank d, the vectors $w_1, w_2, \ldots, w_d$ form a basis of $\mathbb{R}^d$. Note that $\prod_{i=1}^{d} \ell_i \leq M$. Applying a change of basis, we may assume that $w_i = e_i$, the i-th standard basis vector, for all i, $1 \leq i \leq d$. In particular, $\Lambda = \mathbb{Z}^d$.

Pick any d-tuple of linearly independent vectors $v_1, v_2, \ldots, v_d \in S$. Let $A \in \mathbb{R}^{d \times d}$ be the matrix satisfying $Ae_i = v_i$, and put $q = \det A = \mathrm{covol}\,(\mathrm{span}_{\mathbb{Z}}(v_1, v_2, \ldots, v_d))$. Then $q \neq 0$ and the Leibniz formula gives the estimate $|q| \leq d! D^d \prod_{i=1}^{d} \ell_i = O_d(M)$. Using Cramer's rule to find the inverse of A, we observe that the matrix qA^{-1} has integer entries and satisfies $|q(A^{-1})_{i,j}| = O_d(M/\ell_i)$ for each pair (i, j) with $1 \leq i, j \leq d$. Since $qe_i = \sum_{j=1}^{d} q(A^{-1})_{i,j} v_j$ and since $0 \in S$, we conclude that there exists $p = O_d(M)$ such that $qe_i \in \lfloor p/\ell_i \rfloor S$ for every i, $1 \leq i \leq d$.

Pick any $u \in \mathbb{Z}^d \cap B$. Since $u \in \mathbb{Z}^d$, by Lemma 14.5 applied to the group $G = \mathbb{Z}^d/q\mathbb{Z}^d$ and the set[9] $(S \bmod q) \subseteq G$, there exists $u' = (u_i')_{i=1}^d \in \mathbb{Z}^d$ such that $u \equiv qu' + nS$ for some $n \leq dq$. Then for every i, $1 \leq i \leq d$, we have $|u_i'| \leq D\ell_i$, so $qu' \in mS$, where

$$m = \sum_{i=1}^d |u_i'| \lfloor p/\ell_i \rfloor \leq dDp = O_d(M).$$

It follows that we can take $k = n + m = O_d(M)$. $\qquad\square$

Example 14.7. Pick any $d, N \in \mathbb{N}$. Let $S = \{0, \pm v_1, \pm v_2, \ldots, \pm v_d\}$, where $v_1 = e_1$ and $v_i = e_i - Ne_{i-1}$ for every i, $2 \leq i \leq d$. Thus, $\operatorname{span}_{\mathbb{Z}}(S) = \mathbb{Z}^d$. We can compute that

$$e_d = v_d + Nv_{d-1} + \cdots + N^{d-1}v_1.$$

Taking $B = (-2N, 2N)^d$, we see that $|B \cap \mathbb{Z}^d| \leq (4N)^d$ and if $kS \supset B \cap \mathbb{Z}^d$ then $k \geq N^d$. Hence, the bound in Lemma 14.6 is tight up to a constant factor.

Proof of Proposition 14.4. Apply Lemma 14.6 to $B = (-N, N)^d$ and $S = \mathcal{R}_N(\alpha, \varepsilon)$. It is immediate from the definition that $\mathcal{R}_N(\alpha, \varepsilon) \subseteq \Lambda \cap B$, where $\Lambda = \operatorname{span}_{\mathbb{Z}} S$ is a sub-lattice of $\mathbb{Z}^d$. For the other inclusion, we note that $\Lambda \cap B = kS = k\mathcal{R}_N(\alpha, \varepsilon) \subseteq \mathcal{R}_N(\alpha, k\varepsilon)$, where $k = O_d(|\Lambda \cap B|) = O_d(N^d)$. $\qquad\square$

15. Proof of Theorem A: Combining the ingredients

We are now ready to prove a proposition which serves as the inductive step in the proof of Theorem A. In the argument, we will use the following elementary fact.

Lemma 15.1. *Let x and x^* be two real numbers and assume that $|x - x^*| < \min(|x - \lfloor x^* \rfloor|, 1/2)$. Then $\lfloor x \rfloor = \lfloor x^* \rfloor$.*

Proof. If $\{x^*\} \in [0, 1/2)$, then $|x - \lfloor x^* \rfloor| > |x - x^*|$ implies that $x > \lfloor x^* \rfloor = \lfloor x^* \rfloor$, and $|x - x^*| < 1/2$ implies that $x < \lfloor x^* \rfloor + 1$. Thus, $\lfloor x \rfloor = \lfloor x^* \rfloor$. The case where $\{x^*\} \in [1/2, 1)$ is analogous. $\qquad\square$

Proposition 15.2. *Let $d \in \mathbb{N}$, let $h_1, h_2, \ldots, h_d: [N] \to \mathbb{Z}$ be sequences, with $\|h_i\|_\infty \leq H$ for every i, $1 \leq i \leq d$. For $\alpha \in \mathbb{R}^d$, we let $g_\alpha: [N] \to \mathbb{Z}$ denote the sequence defined by*

$$g_\alpha(n) = \left\lfloor \sum_{i=1}^d \alpha_i h_i(n) \right\rfloor, \quad n \in [N].$$

Then $\left|\{g_\alpha \mid \alpha \in [-R, R)^d\}\right| = O_d(R^d H^{3d^2})$.

Proof. Note that $g_\alpha(n) = g_{\{\alpha\}}(n) + g_{\lfloor \alpha \rfloor}(n)$, where $\{\alpha\} \in [0, 1)^d$ and $\lfloor \alpha \rfloor \in \{-R, -R+1, \ldots, R-2, R-1\}^d$, which is a set whose cardinality is in $O_d(R^d)$. Thus, it remains to show that the cardinality of the set $\{g_\alpha \mid \alpha \in [0, 1)^d\}$ is in $O_d(H^{3d^2})$.

Pick any $\alpha \in [0, 1)^d$. Our plan is to construct an alternative description of g_α that, instead of α, includes a finite number of parameters chosen from sets whose cardinality can be estimated in an easier way.

---

[9]Note that $\operatorname{span}_{\mathbb{Z}} S = \Lambda = \mathbb{Z}^d$. Hence it makes sense to consider the set $S \bmod q$. Furthermore, since by assumption $0 \in S$, it follows that $0 \in S \bmod q$ as needed.

Put $\varepsilon = 1/\big(100 C_d (dH)^d\big)$, where C_d is the constant from Proposition 14.4. Let $\alpha^* = (\alpha_i^*)_{i=1}^d \in [0,1)^d$ be any vector such that $|\alpha_i - \alpha_i^*| < \varepsilon/dH$ and $\alpha_i^* \in (\varepsilon/dH)\mathbb{Z}$ for every i, $1 \le i \le d$. This ensures that

$$(50) \qquad \left| \sum_{i=1}^d \alpha_i h_i(n) - \sum_{i=1}^d \alpha_i^* h_i(n) \right| < \varepsilon, \qquad \forall\, n \in [N].$$

Let $h_0 \colon [N] \to \mathbb{Z}$ be the map given by

$$(51) \qquad h_0(n) = -\left\lfloor \sum_{i=1}^d \alpha_i^* h_i(n) \right\rfloor$$

and let $h = (h_0, h_1, h_2, \ldots, h_d) \colon [N] \to \mathbb{Z}^{d+1}$. Note that $\|h_0\|_\infty \le dH$ and, since $\varepsilon < 1/2$, that $h_0(n) + g_\alpha(n) \in \{0, -1\}$ for all $n \in [N]$. Put $B = [-dH, dH] \times [-H, H]^d$. Let also $\Lambda = \Lambda(1\alpha, \varepsilon, dH)$ be the lattice constructed in Proposition 14.4 and set

$$\Lambda^+ = \{ m \in \Lambda \mid 1\alpha \cdot m \ge 0 \} .$$

Here and elsewhere, we use the shorthand $1\alpha = (1, \alpha_1, \alpha_2, \ldots, \alpha_d)$. Hence $(1\alpha)_0 = 1$ and $(1\alpha)_i = \alpha_i$ for every $i \ge 1$.

We have now introduced all the objects needed to obtain a more well-behaved formula for $g_\alpha(n)$. We consider two cases, depending on whether $h(n) \in \Lambda$ or not. If $h(n) \in B \setminus \Lambda$, then $|1\alpha \cdot h(n)| > \varepsilon$ and hence, by Lemma 15.1, we find

$$(52) \qquad g_\alpha(n) = \left\lfloor \sum_{i=1}^d \alpha_i h_i(n) \right\rfloor = \left\lfloor \sum_{i=1}^d \alpha_i^* h_i(n) \right\rfloor .$$

Next, if $h(n) \in B \cap \Lambda$, then $|1\alpha \cdot h(n)| < C_d (dH)^d \varepsilon < 1/2$. Since $h_0(n) \in \mathbb{Z}$, we have

$$(53) \qquad g_\alpha(n) = -h_0(n) + \lfloor 1\alpha \cdot h(n) \rfloor .$$

If $n \in \Lambda^+$, then $1\alpha \cdot h(n) \in [0, \frac{1}{2})$ and hence $\lfloor 1\alpha \cdot h(n) \rfloor = 0$. Similarly, if $h(n) \in \Lambda \setminus \Lambda^+$, then $1\alpha \cdot h(n) \in (-\frac{1}{2}, 0)$ and hence $\lfloor 1\alpha \cdot h(n) \rfloor = -1$. Combining (52) and (53), and expanding out the definition of h_0, we conclude that

$$(54) \qquad g_\alpha(n) = \begin{cases} \left\lfloor \displaystyle\sum_{i=1}^d \alpha_i^* h_i(n) \right\rfloor & \text{if } h(n) \in B \setminus \Lambda, \\[2.5ex] \left\lfloor \displaystyle\sum_{i=1}^d \alpha_i^* h_i(n) + \frac{1}{2} \right\rfloor & \text{if } h(n) \in B \cap \Lambda^+, \\[2.5ex] \left\lfloor \displaystyle\sum_{i=1}^d \alpha_i^* h_i(n) - \frac{1}{2} \right\rfloor & \text{if } h(n) \in B \cap \Lambda \setminus \Lambda^+. \end{cases}$$

It follows from (54) that, in order to determine $g_\alpha|_{[N]}$, it is sufficient to know the following data: α^*, $B \cap \Lambda$, and $B \cap \Lambda^+$. The number of possible choices of α^* is in $O_d(H^{d^2})$ directly by the definition. Furthermore, it follows from Proposition 14.3, applied with $\Gamma = \mathbb{Z}^d$, that the number of possible choices for $B \cap \Lambda$ and $B \cap \Lambda^+$ are both in $O_d\left(H^{2d^2}\right)$. Consequently, the number of distinct sequences of the form $g_\alpha|_{[N]}$ is in $O_d\left(H^{3d^2}\right)$. This ends the proof. $\qquad\square$

We have now collected all the components needed to prove our main result.

Proposition 15.3. *Let $g_\bullet \colon \mathbb{Z} \to \mathbb{Z}$ be a parametric GP map with index set $I \subseteq \mathbb{N}_0$. Then there exist two positive real numbers $B = B(g_\bullet)$ and $C = C(g_\bullet)$ such that*

$$\tag{55} \left| \left\{ g_\alpha|_{[N]} \mid \alpha \in \mathbb{R}^I, \ \|\alpha\|_\infty \leq R \right\} \right| = O_g\left(R^B N^C\right).$$

Proof. We proceed by structural induction with respect to $g_\bullet$, using the scheme introduced by Proposition 13.9. Let $\mathcal{G}'$ denote the set of all parametric GP maps $g_\bullet$ such that (55) holds.

(i) If $g \colon \mathbb{Z} \to \mathbb{Z}$ is a GP map (viewed as a parametric GP map with empty index set), then the set in (55) has only one element and thus $g \in \mathcal{G}$ holds trivially.

(ii) For each $g_\bullet, h_\bullet \in \mathcal{G}'$ with index sets $I, J \subseteq \mathbb{N}_0$ respectively, and for each $N \in \mathbb{N}$, we note that $(g + h)_\gamma|_{[N]}$, $\gamma \in \mathbb{R}^{I \cup J}$, is uniquely determined by $g_\alpha|_{[N]}$ and $h_\beta|_{[N]}$, where $\alpha = \gamma|I$ and $\beta = \gamma|J$. As a consequence, $g_\bullet + h_\bullet \in \mathcal{G}'$ and we can take

$$\tag{56} B(g_\bullet + h_\bullet) \leq B(g_\bullet) + B(h_\bullet) \text{ and } C(g_\bullet + h_\bullet) \leq C(g_\bullet) + C(h_\bullet).$$

The same reasoning also applies to the product $g_\bullet \cdot h_\bullet$, with the same bounds.

(iii) Next, let $g_\bullet$ and $g'_\bullet$ be two parametric GP maps with index sets $I, J \subseteq \mathbb{N}_0$ respectively, and suppose that $g_\bullet \in \mathcal{G}'$ and $g_\bullet \succeq g'_\bullet$. By Definition 13.7, there exists a GP map $\varphi \colon \mathbb{R}^J \to \mathbb{R}^I$ such that $g'_\beta = g_{\varphi(\beta)}$ for all $\beta \in \mathbb{R}^J$. Since φ is a GP map, there exists a constant $D = D(\varphi)$ such that $\|\varphi(\beta)\|_\infty \leq DR^D$ for all $\beta \in \mathbb{R}^J$ with $\|\beta\|_\infty \leq R$. Hence, for every pair of positive integers (N, R), we have

$$\left| \left\{ g'_\beta|_{[N]} \mid \beta \in \mathbb{R}^J, \ \|\beta\|_\infty \leq R \right\} \right| \leq \left| \left\{ g_\alpha|_{[N]} \mid \alpha \in \mathbb{R}^I, \ \|\alpha\|_\infty \leq DR^D \right\} \right|,$$

which implies that $g'_\bullet \in \mathcal{G}'$ and that we can take

$$\tag{57} B(g'_\bullet) \leq DB(g_\bullet) \text{ and } C(g'_\bullet) \leq C(g_\bullet).$$

(iv) Finally, let $I \subseteq \mathbb{N}_0$ be a finite set, and, for every $i \in I$, let $h_\bullet^{(i)} \in \mathcal{G}'$ be a parametric GP map with index sets $J^{(i)} \subseteq \mathbb{N}_0 \setminus I$. Set $J = \bigcup_{i \in I} J^{(i)}$. Our aim is to show that $\mathcal{G}'$ also contains the parametric GP map $g_\bullet$ with index set $I \cup J$ defined by

$$g_{\alpha, \beta}(n) = \left\lfloor \sum_{i \in I} \alpha_i h_\beta^{(i)}(n) \right\rfloor, \qquad n \in \mathbb{Z}, \ \alpha \in \mathbb{R}^I, \ \beta \in \mathbb{R}^J.$$

Given any $N, R \in \mathbb{N}$, we set

$$H = \max_{i \in I} \max_{n \in [N]} \sup_{\|\beta\|_\infty \leq R} \left| h_\beta^{(i)}(n) \right|.$$

Since the $h_\bullet^{(i)}$'s are parametric GP maps, we have $H = O_g(R^D N^D)$ for some constant $D = D(g_\bullet)$. For every $i \in I$, set $B_i = B(h_\bullet^{(i)})$, $C_i = C(h_\bullet^{(i)})$, and

$$\mathcal{H}_i = \left\{ h_\beta^{(i)}|_{[N]} \mid \beta \in \mathbb{R}^J, \ \|\beta\|_\infty \leq R \right\}.$$

By Proposition 15.2, for each $h_1 \in \mathcal{H}_1, h_2 \in \mathcal{H}_2, \ldots, h_d \in \mathcal{H}_d$, the number of sequences of the form $\left\lfloor \sum_{i=1}^d \alpha_i h_i \right\rfloor$, where $\alpha \in \mathbb{R}^d$ and $\|\alpha\|_\infty \leq R$, belongs to $O_d(R^d H^{3d^2})$. It follows that the number of distinct sequences of the form $g_{\alpha, \beta}|_{[N]}$, where $\alpha \in \mathbb{R}^I$, $\beta \in \mathbb{R}^J$, and $\max\{\|\alpha\|_\infty, \|\beta\|_\infty\} \leq R$, is at most

$$\prod_{i=1}^d |\mathcal{H}_i| \cdot O_d\left(R^d H^{3d^2}\right) = O_g(R^B N^C),$$

where

$$B = \sum_{i \in I} B_i + d + 3d^2 D \text{ and } C = \sum_{i \in I} C_i + 3d^2 D.$$

In particular, $g_\bullet \in \mathcal{G}'$ and we can take $B(g_\bullet) \leq B$ and $C(g_\bullet) \leq C$.

It now follows from Proposition 13.9 that $\mathcal{G}'$ contains all parametric GP maps $\mathbb{Z} \to \mathbb{Z}$ with index sets contained in $\mathbb{N}_0$. This ends the proof. $\qquad\square$

Proof of Theorem A. Let $\mathbf{a}$ be a bracket word defined over an alphabet Σ. We infer from Corollary 5.3 that there exist a finitely-valued GP map $g \colon \mathbb{N}_0 \to \mathbb{N}$ and $c \colon g(\mathbb{N}_0) \to \Sigma$ such that $\mathbf{a} = \left(c(g(n))\right)_{n=0}^\infty$. Let $\mathbf{g} = (g(n))_{n=0}^\infty$. It follows from Proposition 15.3 and Lemma 13.4 that $p_{\mathbf{g}}(N) = O_g(N^C)$ for some $C > 0$. Since $p_{\mathbf{a}}(N) \leq p_{\mathbf{g}}(N)$ for all $N \in \mathbb{N}$, this ends the proof. $\qquad\square$

APPENDIX A. NILPOTENT DYNAMICS

A.1. **Connectedness.** A technical issue which often leads to complications is that there is in general no guarantee that the nilpotent Lie group G and the nilmanifold X in Theorem 4.1 should be connected.

As for connectedness of X, in most applications this is not a major problem. In this case, it follows from [Lei05, Thm. B] that there exist $q \in \mathbb{N}$ and connected sub-nilmanifolds $Y_0, Y_1, \ldots, Y_{q-1} \subseteq X$ such that $(g(qn+i)\Gamma)_{n=0}^\infty$ is equidistributed in Y_i for each $i \in [q]$. Passing to an arithmetic progression of the form $q\mathbb{N}_0 + i$, one can then work with the connected nilsystem (Y_i, T^q).

Connectedness of G is a more fundamental issue. For instance, for $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, the GP map $g(n) = \{n^2\alpha\}$ can be represented using the nilrotation $(G/\Gamma, T_h)$, where

$$G = \begin{bmatrix} 1 & \mathbb{Z} & \mathbb{R} \\ 0 & 1 & \mathbb{R} \\ 0 & 0 & 1 \end{bmatrix}, \qquad \Gamma = \begin{bmatrix} 1 & \mathbb{Z} & \mathbb{Z} \\ 0 & 1 & \mathbb{Z} \\ 0 & 0 & 1 \end{bmatrix}, \text{ and } h = \begin{bmatrix} 1 & 1 & -\alpha \\ 0 & 1 & 2\alpha \\ 0 & 0 & 1 \end{bmatrix}.$$

Indeed, a simple computation shows that

$$g^n\Gamma = \begin{bmatrix} 1 & 0 & \{n^2\alpha\} \\ 0 & 1 & \{2n\alpha\} \\ 0 & 0 & 1 \end{bmatrix} \Gamma.$$

However, if we pass to the connected component of G, then we obtain a much simpler nilsystem $G^\circ/\Gamma \cap G^\circ \simeq \mathbb{R}^2/\mathbb{Z}^2$, where it is no more possible to represent g. Fortunately, this problem disappears if instead of sequences $g^n\Gamma$ we consider polynomial maps from $\mathbb{Z}$ to G. Given a connected and simply connected Lie group G, a *polynomial map* $p \colon \mathbb{Z} \to G$ is a sequence of the form

$$p(n) = g_1^{p_1(n)} g_2^{p_2(n)} \cdots g_r^{p_r(n)},$$

where $g_i \in G$, $r \in \mathbb{N}$, and $p_i \colon \mathbb{R} \to \mathbb{R}$ are polynomials.

Theorem A.1 ([BL07, Thm. A*]). *Any GP map $g \colon \mathbb{Z} \to [0,1)$ has a representation $g(n) = \tau_j(p(n)\Gamma)$, where $X = G/\Gamma$ is a nilmanifold with G connected and simply connected, $\tau = (\tau_1, \tau_2, \ldots, \tau_{\dim G}) \colon G/\Gamma \to [0,1)^{\dim G}$ are Mal'cev coordinates, $p \colon \mathbb{Z} \to G$ is a polynomial map, and $1 \leq j \leq \dim G$. Conversely, any map $g \colon \mathbb{Z} \to \mathbb{R}$ of the aforementioned form is a GP map.*

A.2. **Quantitative equidistribution.** Let $X = G/\Gamma$ be a nilmanifold, with G connected and simply connected. Throughout, we assume that X is equipped with a Mal'cev coordinate τ as well as a metric d_X.

Recall that X is equipped with a natural choice of measure, namely the Haar measure μ_X. A sequence $(x_n)_{n=0}^{\infty}$ is equidistributed in X if for each continuous map $f \colon X \to \mathbb{R}$,

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} f(x_n) = \int_X f \, d\mu_X.$$

We will also need a quantitative variant of this property. For $\delta > 0$ and $N \in \mathbb{N}$, a sequence $(x_n)_{n=0}^{N-1}$ is said to be δ-equidistributed in X if for every Lipschitz map $f \colon X \to \mathbb{R}$,

$$\left| \frac{1}{N} \sum_{n=0}^{N-1} f(x_n) - \int_X f \, d\mu_X \right| \le \delta \, \|f\|_{\mathrm{Lip}},$$

where $\|f\|_{\mathrm{Lip}} = \sup_{x \in X} |f| + \sup_{x, y \in X, \, x \neq y} |f(x) - f(y)| / d_X(x, y)$.

We let $X_{\mathrm{ab}} = G/[G, G]\Gamma \simeq (\mathbb{R}/\mathbb{Z})^{d_{\mathrm{ab}}}$ denote the so-called *horizontal torus*, which is the largest torus that is a factor of X, and we let $\pi_{\mathrm{ab}} \colon X \to X_{\mathrm{ab}}$ denote the corresponding projection. We cite a criterion for equidistribution obtained by Leibman.

Theorem A.2 ([Lei05, Thm. C]). *Let $g \colon \mathbb{Z} \to G$ be a polynomial map. Then $(g(n)\Gamma)_{n=0}^{\infty}$ is equidistributed in X if and only if $(\pi_{\mathrm{ab}}(g(n)\Gamma))_{n=0}^{\infty}$ is equidistributed in X_{ab}.*

A *horizontal character* $\eta \colon G \to \mathbb{R}/\mathbb{Z}$ is a continuous homomorphism with $\Gamma \subseteq \ker(\eta)$. Note that η vanishes on $[G, G]\Gamma$ and hence induces a continuous homomorphism $X_{\mathrm{ab}} \to \mathbb{R}/\mathbb{Z}$. Identifying X_{ab} with $(\mathbb{R}/\mathbb{Z})^{d_{\mathrm{ab}}}$ allows us to identify η with a vector $k \in \mathbb{Z}^{d_{\mathrm{ab}}}$, and we set $|\eta| = \|k\|_{\infty}$. The identification of X_{ab} with $(\mathbb{R}/\mathbb{Z})^{d_{\mathrm{ab}}}$ is generally not unique, but Mal'cev coordinates provide one distinguished choice. We cite a simplified variant of the quantitative version of Leibman's theorem.

Theorem A.3 ([GT12b, Thm. 1.16]). *Let $g \colon \mathbb{Z} \to G$ be a polynomial map. Then there exists a positive real number C, which depends on X, τ, and g, such that the following holds. For any $\delta \in (0, 1/2)$ and $N \in \mathbb{N}$ such that the sequence $(g(n)\Gamma)_{n=0}^{N-1}$ is not δ-equidistributed, there exists a horizontal character $\eta \colon G \to \mathbb{R}/\mathbb{Z}$ such that $0 < |\eta| \le 1/\delta^C$ and $\|\eta \circ g\|_{C^{\infty}[N]} \le 1/\delta^C N$.*

We record the following consequence of Theorem A.3, combined with the Schmidt subspace theorem. Below, we say that a polynomial map $g \colon \mathbb{Z} \to G$ has algebraic coefficients if the corresponding map $\widetilde{\tau} \circ g$ is a polynomial with algebraic coefficients (cf. eq. (6)).

Lemma A.4. *Let $g \colon \mathbb{Z} \to G$ be a polynomial map with algebraic coefficients. Suppose that the sequence $(g(n)\Gamma)_{n=0}^{\infty}$ is equidistributed. Then there exists a constant $c > 0$ (dependent on G, Γ, τ, and g) such that for each $N \in \mathbb{N}$ the sequence $(g(n)\Gamma)_{n=0}^{N-1}$ is $O(N^{-c})$-equidistributed in X.*

Proof. We argue by contradiction, assuming that the sequence $(g(n)\Gamma)_{n=0}^{N-1}$ is not $N^{-\delta}$-equidistributed in X for some small $\delta > 0$, to be determined in the course of the argument. It follows from Theorem A.3 that there exist a real number $C > 0$ and a horizontal character $\eta \colon G \to \mathbb{R}/\mathbb{Z}$ such that $|\eta| \le N^{C\delta}$ and $\|\eta \circ g\|_{C^{\infty}[N]} \le N^{C\delta}$.

Letting $\pi_{\mathrm{ab}}\colon X \to X_{\mathrm{ab}}$ denote the projection to the abelian torus and identifying X_{ab} with $(\mathbb{R}/\mathbb{Z})^{d_{\mathrm{ab}}}$, we can expand

$$(58) \qquad \pi_{\mathrm{ab}}(g(n)) = \sum_{j=0}^{r} \alpha^{(j)} n^j,$$

where for every j, $0 \le j \le r$, $\alpha^{(j)} = (\alpha_i^{(j)})_{i=1}^{d_{\mathrm{ab}}} \in X_{\mathrm{ab}}$. Identifying η with a vector $k = (k_1, k_2, \ldots, k_{d_{\mathrm{ab}}}) \in \mathbb{Z}^{d_{\mathrm{ab}}}$, we have

$$(59) \qquad \left\| k \cdot \alpha^{(j)} \right\| = \left\| \sum_{i=1}^{d_{\mathrm{ab}}} k_i \alpha_i^{(j)} \right\| \le N^{C\delta - j} \le N^{C\delta - 1},$$

for every j, $0 \le j \le r$. We have no guarantee that $1, \alpha_1^{(j)}, \alpha_2^{(j)}, \ldots, \alpha_{d_{\mathrm{ab}}}^{(j)}$ are $\mathbb{Q}$-linearly independent for some j. However, we can find a linear combination $\beta = (\beta_i)_{i=1}^{d_{\mathrm{ab}}} = \sum_j w_j \alpha^{(j)}$, with $w_j \in \mathbb{Z}$, such that $1, \beta_1, \beta_2, \ldots, \beta_{d_{\mathrm{ab}}}$ are $\mathbb{Q}$-linearly independent. Then, we infer from (59) that

$$(60) \qquad \| k \cdot \beta \| = \left\| \sum_{i=1}^{d_{\mathrm{ab}}} k_i \beta_i \right\| \ll N^{C\delta - 1}.$$

On the other hand, it follows from the subspace theorem (see [Sch72]) that

$$(61) \qquad \| k \cdot \beta \| = \left\| \sum_{i=1}^{d_{\mathrm{ab}}} k_i \beta_i \right\| \gg \left(\max_{1 \le i \le d_{\mathrm{ab}}} k_i \right)^{-d_{\mathrm{ab}} - 1} \gg N^{-C\delta(d_{\mathrm{ab}} + 1)}.$$

Provided that N is sufficiently large and δ is sufficiently small, (60) and (61) are contradictory. $\qquad\square$

A.3. **Orbit closures.** Lastly, we discuss the behaviour of bounded GP sequences under shifts. First, we record a consequence of Theorem 4.1 and other results in [BL07]. Roughly speaking, it asserts that the class of bounded GP maps representable by a formula of a given "shape" is closed under shifts. (See also Section 13 for an alternative formulation.)

Lemma A.5. *Let $g\colon \mathbb{Z} \to \mathbb{R}$ be a bounded GP map. Then there exist $d \in \mathbb{N}_0$ and a GP map $h\colon \mathbb{R}^d \times \mathbb{Z} \to \mathbb{R}$ such that, for every $m \in \mathbb{Z}$, there exists $\alpha \in [0,1)^d$ such that $g(n + m) = h(\alpha, n)$ for all $n \in \mathbb{Z}$.*

Proof. Let $g(n) = f(T^n(x))$ be the representation of g coming from Theorem 4.1 (using the notation therein). Assume first that X is connected and let $\tau\colon X \to [0,1)^d$ be some Mal'cev coordinates on X. Set $h(\alpha, n) = f(T^n(\tau^{-1}(\alpha)))$. Then h is a GP map by [BL07, Sec. 1.15]. For $m \in \mathbb{Z}$, let $\alpha(m) = \tau(T^m(x))$. Then $h(\alpha(m), n) = g(n + m)$ for all $n \in \mathbb{Z}$. If X is not connected, then we can reduce to the connected case by passing to an arithmetic progression. $\qquad\square$

Finally, we mention a closure property of bounded GP maps from $\mathbb{N}$ to $\mathbb{R}$. We stress that the same conclusion does not hold after replacing $\mathbb{N}$ with $\mathbb{Z}$.

Lemma A.6. *Let $d \in \mathbb{N}$, $h\colon \mathbb{R}^d \times \mathbb{N} \to \mathbb{R}$ be a GP map, and $(x_i)_{i=0}^{\infty}$ be a bounded sequence with values in $\mathbb{R}^d$. Let us assume that the limit*

$$g(n) = \lim_{i \to \infty} h(x_i, n)$$

exists for all $n \in \mathbb{N}$. Then g is a GP map.

Proof. This is a special case of [Kon21, Prop. 2.16]. Note that the assumption that the limit above exists allows us to avoid the use of limits along ultrafilters, which appear in [Kon21]. □

References

[AAB+88] George E. Andrews, Richard A. Askey, Bruce C. Berndt, K. G. Ramanathan, and Robert A. Rankin (eds.), *Ramanujan revisited*, Academic Press, Inc., Boston, MA, 1988. MR938955

[Ada03] Boris Adamczewski, *Balances for fixed points of primitive substitutions*, Theoret. Comput. Sci. **307** (2003), no. 1, 47–75, DOI 10.1016/S0304-3975(03)00092-6. Words. MR2014730

[Ada04] Boris Adamczewski, *Symbolic discrepancy and self-similar dynamics* (English, with English and French summaries), Ann. Inst. Fourier (Grenoble) **54** (2004), no. 7, 2201–2234 (2005). MR2139693

[AGH63] L. Auslander, L. Green, and F. Hahn, *Flows on homogeneous spaces*, Annals of Mathematics Studies, No. 53, Princeton University Press, Princeton, N.J., 1963. With the assistance of L. Markus and W. Massey, and an appendix by L. Greenberg. MR0167569

[All94] Jean-Paul Allouche, *Sur la complexité des suites infinies* (French, with English and French summaries), Bull. Belg. Math. Soc. Simon Stevin **1** (1994), no. 2, 133–143. Journées Montoises (Mons, 1992). MR1318964

[AS92] Jean-Paul Allouche and Jeffrey Shallit, *The ring of k-regular sequences*, Theoret. Comput. Sci. **98** (1992), no. 2, 163–197, DOI 10.1016/0304-3975(92)90001-V. MR1166363

[AS03] Jean-Paul Allouche and Jeffrey Shallit, *Automatic sequences*, Cambridge University Press, Cambridge, 2003. Theory, applications, generalizations, DOI 10.1017/CBO9780511546563. MR1997038

[ASY22] J.-P. Allouche, J. Shallit, and R. Yassawi, *How to prove that a sequence is not automatic*, Expo. Math. **40** (2022), no. 1, 1–22, DOI 10.1016/j.exmath.2021.08.001. MR4388977

[BCR98] Jacek Bochnak, Michel Coste, and Marie-Françoise Roy, *Real algebraic geometry*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 36, Springer-Verlag, Berlin, 1998. Translated from the 1987 French original; Revised by the authors, DOI 10.1007/978-3-662-03718-8. MR1659509

[Ber01] Valérie Berthé, *Autour du système de numération d'Ostrowski* (French, with French summary), Bull. Belg. Math. Soc. Simon Stevin **8** (2001), no. 2, 209–239. Journées Montoises d'Informatique Théorique (Marne-la-Vallée, 2000). MR1838931

[Ber10] Vitaly Bergelson, *Ultrafilters, IP sets, dynamics, and combinatorial number theory*, Ultrafilters across mathematics, Contemp. Math., vol. 530, Amer. Math. Soc., Providence, RI, 2010, pp. 23–47, DOI 10.1090/conm/530/10439. MR2757532

[BH96] Vitaly Bergelson and Inger Johanne Håland, *Sets of recurrence and generalized polynomials*, Convergence in ergodic theory and probability (Columbus, OH, 1993), Ohio State Univ. Math. Res. Inst. Publ., vol. 5, de Gruyter, Berlin, 1996, pp. 91–110. MR1412598

[BHKS20] Vitaly Bergelson, Inger J. Håland Knutson, and Younghwan Son, *An extension of Weyl's equidistribution theorem to generalized polynomials and applications*, Int. Math. Res. Not. IMRN **19** (2021), 14965–15018, DOI 10.1093/imrn/rnaa035. MR4324734

[BK18]   Jakub Byszewski and Jakub Konieczny, *Sparse generalised polynomials*, Trans. Amer.
           Math. Soc. **370** (2018), no. 11, 8081–8109, DOI 10.1090/tran/7257. MR3852458
[BK20]   Jakub Byszewski and Jakub Konieczny, *Automatic sequences and generalised polyno-
           mials*, Canad. J. Math. **72** (2020), no. 2, 392–426, DOI 10.4153/s0008414x19000038.
           MR4081697
[BK22]   J. Byszewski and J. Konieczny, *Pisot numbers, salem numbers, and generalised poly-
           nomials*, 2022, In preparation. `arXiv:2302.06242`
[BL07]   Vitaly Bergelson and Alexander Leibman, *Distribution of values of bounded generalized
           polynomials*, Acta Math. **198** (2007), no. 2, 155–230, DOI 10.1007/s11511-007-0015-y.
           MR2318563
[BL18]   V. Bergelson and A. Leibman, IP$_r^*$-*recurrence and nilsystems*, Adv. Math. **339** (2018),
           642–656, DOI 10.1016/j.aim.2018.09.032. MR3866908
[BMR20]   V. Bergelson, J. Moreira, and F. K. Richter, *Multiple ergodic averages along func-
           tions from a Hardy field: convergence, recurrence and combinatorial applications*, 2020,
           Preprint, `arXiv:2006.03558` [math.DS].
[BPR21]   M.-P. Béal, D. Perrin, and A. Restivo, *Decidable problems in substitution shifts*, 2021,
           Preprint, `arXiv:2112.14499` [math.DS].
[Cas72]   J. W. S. Cassels, *An introduction to Diophantine approximation*, Cambridge Tracts
           in Mathematics and Mathematical Physics, No. 45, Hafner Publishing Co., New York,
           1972. Facsimile reprint of the 1957 edition. MR0349591
[Che00]   Yong-Gao Chen, *The best quantitative Kronecker's theorem*, J. London Math. Soc. (2)
           61 (2000), no. 3, 691–705, DOI 10.1112/S0024610700008772. MR1766098
[CK97]   Julien Cassaigne and Juhani Karhumäki, *Toeplitz words, generalized periodicity and
           periodically iterated morphisms*, European J. Combin. **18** (1997), no. 5, 497–510, DOI
           10.1006/eujc.1996.0110. MR1455183
[CN10]   Julien Cassaigne and François Nicolas, *Factor complexity*, Combinatorics, automata
           and number theory, Encyclopedia Math. Appl., vol. 135, Cambridge Univ. Press, Cam-
           bridge, 2010, pp. 163–247. MR2759107
[CRS12]   Émilie Charlier, Narad Rampersad, and Jeffrey Shallit, *Enumeration and decidable
           properties of automatic sequences*, Internat. J. Found. Comput. Sci. **23** (2012), no. 5,
           1035–1066, DOI 10.1142/S0129054112400448. MR2983367
[CT22]   S. Chow and N. Technau, *Counting multiplicative approximations*, Ramanujan J.
           (2022).
[Dav37]   H. Davenport, *On some infinite series involving arithmetical functions (II)*, The Quar-
           terly Journal of Mathematics, os-8(1): 313–320, 01 1937.
[DMS18]   Michael Drmota, Clemens Müllner, and Lukas Spiegelhofer, *Möbius orthogonality for
           the Zeckendorf sum-of-digits function*, Proc. Amer. Math. Soc. **146** (2018), no. 9,
           3679–3691, DOI 10.1090/proc/14015. MR3825824
[Dow05]   Tomasz Downarowicz, *Survey of odometers and Toeplitz flows*, Algebraic and topo-
           logical dynamics, Contemp. Math., vol. 385, Amer. Math. Soc., Providence, RI, 2005,
           pp. 7–37, DOI 10.1090/conm/385/07188. MR2180227
[Dur13a]   Fabien Durand, *Decidability of the HD0L ultimate periodicity problem*, RAIRO Theor.
           Inform. Appl. **47** (2013), no. 2, 201–214, DOI 10.1051/ita/2013035. MR3072319
[Dur13b]   Fabien Durand, *Decidability of uniform recurrence of morphic sequences*, Internat.
           J. Found. Comput. Sci. **24** (2013), no. 1, 123–146, DOI 10.1142/S0129054113500032.
           MR3061952
[Eis20]   Tanja Eisner, *Nilsystems and ergodic averages along primes*, Ergodic Theory Dynam.
           Systems **40** (2020), no. 10, 2769–2777, DOI 10.1017/etds.2019.27. MR4138910
[EKL06]   Manfred Einsiedler, Anatole Katok, and Elon Lindenstrauss, *Invariant measures and
           the set of exceptions to Littlewood's conjecture*, Ann. of Math. (2) **164** (2006), no. 2,
           513–560, DOI 10.4007/annals.2006.164.513. MR2247967
[Fer99]   Sébastien Ferenczi, *Complexity of sequences and dynamical systems*, Discrete Math.
           206 (1999), no. 1-3, 145–154, DOI 10.1016/S0012-365X(98)00400-2. Combinatorics
           and number theory (Tiruchirappalli, 1996). MR1665394
[Fer06]   T. Fernique, *Multidimensional Sturmian sequences and generalized substitutions*, In-
           ternat. J. Found. Comput. Sci. 17(3): 575–599, 2006.

[FH17]   Nikos Frantzikinakis and Bernard Host, *Higher order Fourier analysis of multiplicative functions and applications*, J. Amer. Math. Soc. **30** (2017), no. 1, 67–157, DOI 10.1090/jams/857. MR3556289

[Fra09]   Nikos Frantzikinakis, *Equidistribution of sparse sequences on nilmanifolds*, J. Anal. Math. **109** (2009), 353–395, DOI 10.1007/s11854-009-0035-y. MR2585398

[GHS13]   Daniel Goč, Dane Henshall, and Jeffrey Shallit, *Automatic theorem-proving in combinatorics on words*, Internat. J. Found. Comput. Sci. **24** (2013), no. 6, 781–798, DOI 10.1142/S0129054113400182. MR3158968

[GL15]   Sigrid Grepstad and Nir Lev, *Sets of bounded discrepancy for multi-dimensional irrational rotation*, Geom. Funct. Anal. **25** (2015), no. 1, 87–133, DOI 10.1007/s00039-015-0313-z. MR3320890

[GO10]   Ron Graham and Kevin O'Bryant, *Can you hear the shape of a Beatty sequence?*, Additive number theory, Springer, New York, 2010, pp. 39–52, DOI 10.1007/978-0-387-68361-4_3. MR2744742

[Gow01]   W. T. Gowers, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal. **11** (2001), no. 3, 465–588, DOI 10.1007/s00039-001-0332-9. MR1844079

[GT10a]   Ben Green and Terence Tao, *An arithmetic regularity lemma, an associated counting lemma, and applications*, An irregular mind, Bolyai Soc. Math. Stud., vol. 21, János Bolyai Math. Soc., Budapest, 2010, pp. 261–334, DOI 10.1007/978-3-642-14444-8_7. MR2815606

[GT10b]   Benjamin Green and Terence Tao, *Linear equations in primes*, Ann. of Math. (2) **171** (2010), no. 3, 1753–1850, DOI 10.4007/annals.2010.171.1753. MR2680398

[GT12a]   Ben Green and Terence Tao, *The Möbius function is strongly orthogonal to nilsequences*, Ann. of Math. (2) **175** (2012), no. 2, 541–566, DOI 10.4007/annals.2012.175.2.3. MR2877066

[GT12b]   Ben Green and Terence Tao, *The quantitative behaviour of polynomial orbits on nilmanifolds*, Ann. of Math. (2) **175** (2012), no. 2, 465–540, DOI 10.4007/annals.2012.175.2.2. MR2877065

[GTZ12]   Ben Green, Terence Tao, and Tamar Ziegler, *An inverse theorem for the Gowers $U^{s+1}[N]$-norm*, Ann. of Math. (2) **176** (2012), no. 2, 1231–1372, DOI 10.4007/annals.2012.176.2.11. MR2950773

[Hål93]   Inger Johanne Håland, *Uniform distribution of generalized polynomials*, J. Number Theory **45** (1993), no. 3, 327–366, DOI 10.1006/jnth.1993.1082. MR1247389

[Hål94]   Inger Johanne Håland, *Uniform distribution of generalized polynomials of the product type*, Acta Arith. **67** (1994), no. 1, 13–27, DOI 10.4064/aa-67-1-13-27. MR1292518

[Har67]   E. F. Harding, *The number of partitions of a set of N points in k dimensions induced by hyperplanes*, Proc. Edinburgh Math. Soc. (2) **15** (1966/67), 285–289, DOI 10.1017/S0013091500011925. MR229128

[HK05]   Bernard Host and Bryna Kra, *Nonconventional ergodic averages and nilmanifolds*, Ann. of Math. (2) **161** (2005), no. 1, 397–488, DOI 10.4007/annals.2005.161.397. MR2150389

[HKM14]   Bernard Host, Bryna Kra, and Alejandro Maass, *Complexity of nilsystems and systems lacking nilfactors*, J. Anal. Math. **124** (2014), 261–295, DOI 10.1007/s11854-014-0032-7. MR3286054

[Kon17]   Jakub Konieczny, *Combinatorial properties of Nil-Bohr sets*, Israel J. Math. **220** (2017), no. 1, 333–385, DOI 10.1007/s11856-017-1520-0. MR3666828

[Kon21]   Jakub Konieczny, *Generalised polynomials and integer powers*, J. Lond. Math. Soc. (2) **105** (2022), no. 1, 154–219, DOI 10.1112/jlms.12509. MR4411322

[KS22]   Daniel Krenn and Jeffrey Shallit, *Decidability and k-regular sequences*, Theoret. Comput. Sci. **907** (2022), 34–44, DOI 10.1016/j.tcs.2022.01.018. MR4383214

[Lei05]   A. Leibman, *Pointwise convergence of ergodic averages for polynomial sequences of translations on a nilmanifold*, Ergodic Theory Dynam. Systems **25** (2005), no. 1, 201–213, DOI 10.1017/S0143385704000215. MR2122919

[Lei12]   A. Leibman, *A canonical form and the distribution of values of generalized polynomials*, Israel J. Math. **188** (2012), 131–176, DOI 10.1007/s11856-011-0097-2. MR2897727

[LOS16]   Robert J. Lemke Oliver and Kannan Soundararajan, *Unexpected biases in the distribution of consecutive primes*, Proc. Natl. Acad. Sci. USA **113** (2016), no. 31, E4446–E4454, DOI 10.1073/pnas.1605366113. MR3624386

[Mal49]   A. I. Mal′cev, *On a class of homogeneous spaces* (Russian), Izvestiya Akad. Nauk. SSSR. Ser. Mat. **13** (1949), 9–32. MR0028842

[Mal51]   A. I. Malcev, *On a class of homogeneous spaces*, Amer. Math. Soc. Translation **1951** (1951), no. 39, 33. MR0039734

[Mat93]   Yuri V. Matiyasevich, *Hilbert's tenth problem*, Foundations of Computing Series, MIT Press, Cambridge, MA, 1993. Translated from the 1993 Russian original by the author; With a foreword by Martin Davis. MR1244324

[MH38]   Marston Morse and Gustav A. Hedlund, *Symbolic Dynamics*, Amer. J. Math. **60** (1938), no. 4, 815–866, DOI 10.2307/2371264. MR1507944

[Mou16]   H. Mousavi, *Automatic theorem proving in walnut*, 2016, Preprint, `arXiv:1603.06017` [math.FL].

[Ric20]   F. K. Richter, *Uniform distribution in nilmanifolds along functions from a Hardy field*, JAMA (2022). DOI 10.1007/s11854-022-0253-0.

[Rig14]   Michel Rigo, *Formal languages, automata and numeration systems. 2*, Networks and Telecommunications Series, ISTE, London; John Wiley & Sons, Inc., Hoboken, NJ, 2014. Applications to recognizability and decidability; With a foreword by Valérie Berthé. MR3526118

[Sch72]   Wolfgang M. Schmidt, *Norm form equations*, Ann. of Math. (2) **96** (1972), 526–551, DOI 10.2307/1970824. MR314761

[Sta00]   Alexander N. Starkov, *Dynamical systems on homogeneous spaces*, Translations of Mathematical Monographs, vol. 190, American Mathematical Society, Providence, RI, 2000. Translated from the 1999 Russian original by the author, DOI 10.1090/m-mono/190. MR1746847

[Tao12]   Terence Tao, *Higher order Fourier analysis*, Graduate Studies in Mathematics, vol. 142, American Mathematical Society, Providence, RI, 2012, DOI 10.1090/gsm/142. MR2931680

[TV06]   Terence Tao and Van Vu, *Additive combinatorics*, Cambridge Studies in Advanced Mathematics, vol. 105, Cambridge University Press, Cambridge, 2006, DOI 10.1017/CBO9780511755149. MR2289012

[Vau97]   R. C. Vaughan, *The Hardy-Littlewood method*, 2nd ed., Cambridge Tracts in Mathematics, vol. 125, Cambridge University Press, Cambridge, 1997, DOI 10.1017/CBO9780511470929. MR1435742

[Wal82]   Peter Walters, *An introduction to ergodic theory*, Graduate Texts in Mathematics, vol. 79, Springer-Verlag, New York-Berlin, 1982. MR648108

[War33]   Morgan Ward, *The arithmetical theory of linear recurring series*, Trans. Amer. Math. Soc. **35** (1933), no. 3, 600–628, DOI 10.2307/1989851. MR1501705

[WY81]   Yuan Wang and Kun Rui Yu, *A note on some metrical theorems in Diophantine approximation* (English, with Chinese summary), Chinese Ann. Math. **2** (1981), no. 1, 1–12, DOI 10.1007/bf00646384. MR619457

[Yaz01]   Soroosh Yazdani, *Multiplicative functions and k-automatic sequences* (English, with English and French summaries), J. Théor. Nombres Bordeaux **13** (2001), no. 2, 651–658. MR1879677

[Zie07]   Tamar Ziegler, *Universal characteristic factors and Furstenberg averages*, J. Amer. Math. Soc. **20** (2007), no. 1, 53–97, DOI 10.1090/S0894-0347-06-00532-7. MR2257397

Université Claude Bernard Lyon 1, CNRS UMR 5208, Institut Camille Jordan, 69622 Villeurbanne Cedex, France

Email address: `boris.adamczewski@math.cnrs.fr`

Université Claude Bernard Lyon 1, CNRS UMR 5208, Institut Camille Jordan, 69622 Villeurbanne Cedex, France

Email address: `jakub.konieczny@gmail.com`