Anti-powers in words: a new notion of regularity based on diversity

Gabriele Fici

Workshop on Words and Complexity

Lyon, France, 19-23 February 2018

Regularities in Combinatorics

In combinatorics, the notion of regularity is often associated with repetitions of some equal objects.

Regularities in Combinatorics

In combinatorics, the notion of regularity is often associated with repetitions of some equal objects.

For example, a regular polygon is one in which all the sides have *the same* length.

Regularities in Combinatorics

In combinatorics, the notion of regularity is often associated with repetitions of some equal objects.

For example, a regular polygon is one in which all the sides have *the same* length.

We consider a different point of view, in which we look for diversity. That is, definitions of regularity based on all-distinct objects.

Regularities in Combinatorics

In combinatorics, the notion of regularity is often associated with repetitions of some equal objects.

For example, a regular polygon is one in which all the sides have *the same* length.

We consider a different point of view, in which we look for diversity. That is, definitions of regularity based on all-distinct objects.

Of course, being all distinct is *a priori* more common than being all equal. Still, only a few works have been devoted to enumerating all-distinct configurations of some combinatorial structure.

Ramsey Theory

Ramsey theory looks for unavoidable regularity inside large combinatorial structures.

Ramsey Theory

Ramsey theory looks for unavoidable regularity inside large combinatorial structures.

A basic example could be the pigeonhole principle: if n items are put into m containers, with $n > m$, then at least one container must contain more than one item.

Ramsey Theory

Ramsey theory looks for unavoidable regularity inside large combinatorial structures.

A basic example could be the pigeonhole principle: if n items are put into m containers, with $n > m$, then at least one container must contain more than one item.

The two most influential results of this type are probably:

1. Ramsey's theorem for graphs;
2. Van der Waerden's theorem for coloring of the positive integers.

Ramsey's Theorem for Graphs

Let us consider undirected graphs in which edges are colored either red or blue.

Ramsey's Theorem for Graphs

Let us consider undirected graphs in which edges are colored either red or blue.

Ramsey's theorem (1930) states that for every pair of integers $r, b \geq 2$, there exists an integer $N = N(r, b)$ such that every possible coloring of the complete graph K_N creates a red K_r or a blue K_b.

Ramsey's Theorem for Graphs

Let us consider undirected graphs in which edges are colored either red or blue.

Ramsey's theorem (1930) states that for every pair of integers $r, b \geq 2$, there exists an integer $N = N(r, b)$ such that every possible coloring of the complete graph K_N creates a red K_r or a blue K_b.

For example, $N(3, 3) = 6$. Hence, in every group of six people, one can always find three of them that are pairwise friends or three of them that are pairwise strangers.

Ramsey's Theorem for Graphs

Let us consider undirected graphs in which edges are colored either red or blue.

Ramsey's theorem (1930) states that for every pair of integers $r, b \geq 2$, there exists an integer $N = N(r, b)$ such that every possible coloring of the complete graph K_N creates a red K_r or a blue K_b.

For example, $N(3, 3) = 6$. Hence, in every group of six people, one can always find three of them that are pairwise friends or three of them that are pairwise strangers.

Note that the smallest $N(r, b)$, called Ramsey numbers, are hard to compute. For example, nobody knows the exact value of $N(5, 5)$.

Anti-Ramsey Theory

There also exists the notion of a rainbow (an edge-colored graph in which no color is repeated) and the corresponding anti-Ramsey theory, see:

P. Erdős, M. Simonovits, V. T. Sós. Anti-Ramsey theorems, infinite and finite sets (Colloq. Keszthely, 1973; dedicated to P. Erdős on his 60th birthday). Colloq. Math. Soc. János Bolyai, pages 633–643, 1975.

Van der Waerden's Theorem

Van der Waerden's theorem (1927) states that every finite coloring of the positive integers contains arbitrarily long monochromatic arithmetic subsequences.

Van der Waerden's Theorem

Van der Waerden's theorem (1927) states that every finite coloring of the positive integers contains arbitrarily long monochromatic arithmetic subsequences.

An r-coloring of the positive integers can be viewed as an infinite word over r letters (usually denoted by $0, 1, \ldots, r-1$).

Van der Waerden's Theorem

Van der Waerden's theorem (1927) states that every finite coloring of the positive integers contains arbitrarily long monochromatic arithmetic subsequences.

An r-coloring of the positive integers can be viewed as an infinite word over r letters (usually denoted by $0, 1, \ldots, r-1$).

Note that one can avoid *infinite* monochromatic arithmetic subsequences, as shown by the infinite binary word

$$w = 011000111100000 \cdots$$

Van der Waerden's Theorem

A finite version of Van der Waerden's theorem is the following: For every pair of positive integers r, k there exists a positive integer $W = W(r, k)$ such that if the integers $1, 2, \ldots, W$ are colored, each with one of r different colors, then there are at least k integers that form a monochromatic arithmetic progression. The smallest such W is the van der Waerden number $W(r, k)$.

Van der Waerden's Theorem

A finite version of Van der Waerden's theorem is the following: For every pair of positive integers r, k there exists a positive integer $W = W(r, k)$ such that if the integers $1, 2, \ldots, W$ are colored, each with one of r different colors, then there are at least k integers that form a monochromatic arithmetic progression. The smallest such W is the van der Waerden number $W(r, k)$.

As well as Ramsey numbers, van der Waerden numbers are hard to compute — for example, nobody knows the exact value of $W(4, 4)$.

Van der Waerden's Theorem

A finite version of Van der Waerden's theorem is the following: For every pair of positive integers r, k there exists a positive integer $W = W(r, k)$ such that if the integers $1, 2, \ldots, W$ are colored, each with one of r different colors, then there are at least k integers that form a monochromatic arithmetic progression. The smallest such W is the van der Waerden number $W(r, k)$.

As well as Ramsey numbers, van der Waerden numbers are hard to compute — for example, nobody knows the exact value of $W(4, 4)$.

However, T. Gowers [Geom. Funct. Anal., 2001] proved that

$$W(r,k) \leq 2^{2^{r^{2^{2^{k+9}}}}}$$

Unavoidable Regularities

We are interested in combinatorial properties of finite factors of infinite words.

Unavoidable Regularities

We are interested in combinatorial properties of finite factors of infinite words.

For example, we want to know if some kind of pattern does/does not appear inside a given infinite word.

Unavoidable Regularities

We are interested in combinatorial properties of finite factors of infinite words.

For example, we want to know if some kind of pattern does/does not appear inside a given infinite word.

Example

The fixed point starting with 2 of the substitution $0 \mapsto 1, 1 \mapsto 20, 2 \mapsto 210$:

$$h = 2102012101202102012021012 1 \cdots$$

does not contain any square, that is, a pattern of the form XX, where X is any nonempty block of consecutive letters.

Unavoidable Regularities

Some patterns are avoidable; others are not. Often, this depends on the size of the alphabet.

Unavoidable Regularities

Some patterns are avoidable; others are not. Often, this depends on the size of the alphabet.

Remark

Every infinite word over a 2-letter alphabet must contain a square.

Unavoidable Regularities

Some patterns are avoidable; others are not. Often, this depends on the size of the alphabet.

Remark

Every infinite word over a 2-letter alphabet must contain a square.

So for example squares are avoidable over a 3-letter alphabet, but they are unavoidable over a 2-letter alphabet.

Unavoidable Regularities

Some patterns are avoidable; others are not. Often, this depends on the size of the alphabet.

Remark

Every infinite word over a 2-letter alphabet must contain a square.

So for example squares are avoidable over a 3-letter alphabet, but they are unavoidable over a 2-letter alphabet.

As another example, cubes (powers of order 3) are avoidable over a 2-letter alphabet (for example, the Thue-Morse word $t = 011010011001011010010110011 0\cdots$ does not contain any cube).

Unavoidable Regularities

Some patterns are avoidable; others are not. Often, this depends on the size of the alphabet.

Remark

Every infinite word over a 2-letter alphabet must contain a square.

So for example squares are avoidable over a 3-letter alphabet, but they are unavoidable over a 2-letter alphabet.

As another example, cubes (powers of order 3) are avoidable over a 2-letter alphabet (for example, the Thue-Morse word $t = 0110100110010110100101100110\cdots$ does not contain any cube).

There is a vaste literature on the avoidability of exact and approximate repetitions...

Unavoidable Regularities

Definition

An unavoidable regularity is a property P such that it is not possible to construct arbitrarily long words not satisfying P.

Unavoidable Regularities

Definition

An unavoidable regularity is a property P such that it is not possible to construct arbitrarily long words not satisfying P.

Only a few unavoidable regularities are known without restrictions on the size of the alphabet.

Unavoidable Regularities

Definition

An unavoidable regularity is a property P such that it is not possible to construct arbitrarily long words not satisfying P.

Only a few unavoidable regularities are known without restrictions on the size of the alphabet.

Unavoidable regularities are important because finding them in a given sequence does not allow one to derive that that sequence has special properties.

Unavoidable Regularities

For example, if someone observes that in the word of length 76:

$$CCGCTACGATGTCCTATAACCTCGCAAGGTGCCACGCA\cdot$$

$$CCGTCAGCGACAGGTCGA{\color{red}T}GGCC{\color{red}T}TCGC{\color{red}T}ATGGACTAA$$

there are 3 T's at the same distance, then we are not surprised since we know that $W(4,3) = 76$.

Unavoidable Regularities

For example, if someone observes that in the word of length 76:

$$CCGCTACGATGTCCTATAACCTCGCAAGGTGCCACGCA\cdot$$
$$CCGTCAGCGACAGGTCGA{\color{red}T}GGCC{\color{red}T}TCGC{\color{red}T}ATGGACTAA$$

there are 3 T's at the same distance, then we are not surprised since we know that $W(4,3) = 76$.

(Therefore, not only the existence of an unavoidable regularity is important, but also the computation of the "avoidability thresholds" is.)

Unavoidable Regularities

An example of unavoidable regularity for words is given by the Zimin patterns: $X_1, X_1X_2X_1, X_1X_2X_1X_3X_1X_2X_1, \ldots$

Unavoidable Regularities

An example of unavoidable regularity for words is given by the Zimin patterns: $X_1, X_1X_2X_1, X_1X_2X_1X_3X_1X_2X_1, \ldots$

Theorem (Zimin, 1984)

All Zimin patterns are unavoidable.

Unavoidable Regularities

An example of unavoidable regularity for words is given by the Zimin patterns: $X_1, X_1X_2X_1, X_1X_2X_1X_3X_1X_2X_1, \ldots$

Theorem (Zimin, 1984)

All Zimin patterns are unavoidable.

For example, over a k-ary alphabet, every word of length at least $2k+1$ must contain a factor of the form XYX.

Unavoidable Regularities

An example of unavoidable regularity for words is given by the Zimin patterns: $X_1, X_1X_2X_1, X_1X_2X_1X_3X_1X_2X_1, \ldots$

Theorem (Zimin, 1984)

All Zimin patterns are unavoidable.

For example, over a k-ary alphabet, every word of length at least $2k+1$ must contain a factor of the form XYX.

E.g., over 3 letters, the word $w = 001122$ does not contain any factor of the form XYX, but it cannot be extended by one letter keeping this property.

Unavoidable Regularities

Another example of unavoidable regularity is given by the Shirshov's theorem.

Unavoidable Regularities

Another example of unavoidable regularity is given by the Shirshov's theorem.

Recall that a sequence $(u_1, u_2, \ldots, u_n)$ of nonempty words is an n-division of the word $u = u_1u_2 \cdots u_n$ if for any permutation $\sigma \in S_n$ one has $u > u_{\sigma(1)}u_{\sigma(2)} \cdots u_{\sigma(n)}$ (in the lexicographic order).

Unavoidable Regularities

Another example of unavoidable regularity is given by the Shirshov's theorem.

Recall that a sequence $(u_1, u_2, \ldots, u_n)$ of nonempty words is an n-division of the word $u = u_1 u_2 \cdots u_n$ if for any permutation $\sigma \in S_n$ one has $u > u_{\sigma(1)} u_{\sigma(2)} \cdots u_{\sigma(n)}$ (in the lexicographic order).

Theorem (Shirshov, 1957)

Given a finite ordered alphabet A, for any integers $n, k > 1$ there exists $S = S(|A|, n, k)$ such that every word of length S contains either a factor that is a k-power or a factor that is n-divided.

Our Result

We introduce the notion of an anti-power and show that it gives rise to a new unavoidable regularity.

Our Result

We introduce the notion of an anti-power and show that it gives rise to a new unavoidable regularity.

Definition

An anti-power of order k (or simply a k-anti-power) is a word of the form $x_1 x_2 \cdots x_k$ where the x_i have the same length and are all distinct.

Our Result

We introduce the notion of an anti-power and show that it gives rise to a new unavoidable regularity.

Definition

An anti-power of order k (or simply a k-anti-power) is a word of the form $x_1 x_2 \cdots x_k$ where the x_i have the same length and are all distinct.

Example

The prefix of length 12 of

$$h = 2102012101202102012021012 1 \cdots$$

is a 3-anti-power: $2102 \cdot 0121 \cdot 0120$, while the prefix of length 16 is not a 4-anti-power.

Our Result

We proved the following

Theorem (G. F., A. Restivo, M. Silva, L. Zamboni)

Every infinite word contains powers of any order or anti-powers of any order.

Our Result

We proved the following

Theorem (G. F., A. Restivo, M. Silva, L. Zamboni)

Every infinite word contains powers of any order or anti-powers of any order.

That is, the presence of consecutive blocks of the same length that are all equal or all different within any infinite word is an unavoidable regularity.

Our Result

The proof of the theorem is purely combinatorial, and actually allows us to state a stronger result:

Theorem

Let w be an infinite word.

$$AP(w,k) = \{m \in \mathbb{N} \mid \text{ the prefix of } w \text{ of length } km \text{ is a } k\text{-anti-power}\}.$$

Suppose that the lower density $\underline{d}(AP(w,k))$ verifies

$$\underline{d}(AP(w,k)) < \left(1 + \binom{k}{2}\right)^{-1} = \frac{2}{2+k(k-1)}$$

for some $k \in \mathbb{N}$. Then there exists u with $0 < |u| \leq (k-1)\binom{k}{2}$ such that u^l is a factor of w for every $l \geq 1$.

Some Consequence

Let us give a corollary of our main result.

Definition

An infinite word w is said to be uniformly recurrent if every finite factor of w occurs syndetically in it (that is, it occurs infinitely often and with bounded gaps).

Some Consequence

Let us give a corollary of our main result.

Definition

An infinite word w is said to be uniformly recurrent if every finite factor of w occurs syndetically in it (that is, it occurs infinitely often and with bounded gaps).

Corollary

Let w be a uniformly recurrent aperiodic word. Then anti-powers of any order occur at every position of w.

Some Consequence

So for example the Thue-Morse word

$$t = 01101001100101101001011001 10\cdots$$

contains anti-powers of any order starting at every position.

Some Consequence

So for example the Thue-Morse word

$$t = 0110100110010110100101100110\cdots$$

contains anti-powers of any order starting at every position.

We conjectured that the length $n_t(k)$ of the shortest prefix of t that is a k-anti-power grows linearly with k.

Some Consequence

So for example the Thue-Morse word

$$t = 0110100110010110100101100110\cdots$$

contains anti-powers of any order starting at every position.

We conjectured that the length $n_t(k)$ of the shortest prefix of t that is a k-anti-power grows linearly with k.

Recently, C. Defant [Elect. J. Combin., 2017] proved this conjecture, together with other interesting results on the combinatorics of anti-powers in the Thue-Morse word.

Some Consequence

So for example the Thue-Morse word

$$t = 0110100110010110100101100110\cdots$$

contains anti-powers of any order starting at every position.

We conjectured that the length $n_t(k)$ of the shortest prefix of t that is a k-anti-power grows linearly with k.

Recently, C. Defant [Elect. J. Combin., 2017] proved this conjecture, together with other interesting results on the combinatorics of anti-powers in the Thue-Morse word.

What other words w are such that $n_w(k)$ grows linearly with k?

Avoiding Anti-Powers

Periodic words avoid anti-powers, the period length being an upper bound for the maximal number of distinct consecutive blocks of the same length.

Avoiding Anti-Powers

Periodic words avoid anti-powers, the period length being an upper bound for the maximal number of distinct consecutive blocks of the same length.

Lemma

Let w be an infinite word avoiding 3-anti-powers. Then w is ultimately periodic. i.e., $w = uvvvv\cdots$ for some blocks u and v.

Avoiding Anti-Powers

Periodic words avoid anti-powers, the period length being an upper bound for the maximal number of distinct consecutive blocks of the same length.

Lemma

Let w be an infinite word avoiding 3-anti-powers. Then w is ultimately periodic. i.e., $w = uvvvv\cdots$ for some blocks u and v.

Proposition

There exist aperiodic words avoiding 4-anti-powers.

Avoiding Anti-Powers

Periodic words avoid anti-powers, the period length being an upper bound for the maximal number of distinct consecutive blocks of the same length.

Lemma

Let w be an infinite word avoiding 3-anti-powers. Then w is ultimately periodic. i.e., $w = uvvvv\cdots$ for some blocks u and v.

Proposition

There exist aperiodic words avoiding 4-anti-powers.

An example is the word $w = 1 \cdot 0^3 \cdot 1 \cdot 0^{19} \cdot 1 \cdot 0^{99} \cdot 1 \cdots$ where there is a 1 in every position that is a power of 5, and 0 elsewhere.

Avoiding Anti-Powers

$$w = 1 \cdot 0^3 \cdot 1 \cdot 0^{19} \cdot 1 \cdot 0^{99} \cdot 1 \cdots$$

This word avoids 4-anti-powers but is not recurrent (a word is recurrent if every factor occurs infinitely often).

Avoiding Anti-Powers

$$w = 1 \cdot 0^3 \cdot 1 \cdot 0^{19} \cdot 1 \cdot 0^{99} \cdot 1 \cdots$$

This word avoids 4-anti-powers but is not recurrent (a word is recurrent if every factor occurs infinitely often).

Adding the hypothesis of recurrence we proved the following

Proposition

There exist recurrent aperiodic words avoiding 6-anti-powers.

Avoiding Anti-Powers

$$w = 1 \cdot 0^3 \cdot 1 \cdot 0^{19} \cdot 1 \cdot 0^{99} \cdot 1 \cdots$$

This word avoids 4-anti-powers but is not recurrent (a word is recurrent if every factor occurs infinitely often).

Adding the hypothesis of recurrence we proved the following

Proposition

There exist recurrent aperiodic words avoiding 6-anti-powers.

Let $w_0 = 0$ and $w_n = w_{n-1}1^{3|w_{n-1}|}w_{n-1}$ for every $n > 0$. The infinite word w obtained as the limit of the sequence of words $(w_n)_{n\geq 1}$ is recurrent and avoids 6-anti-powers.

Avoiding Anti-Powers

$$w = 1 \cdot 0^3 \cdot 1 \cdot 0^{19} \cdot 1 \cdot 0^{99} \cdot 1 \cdots$$

This word avoids 4-anti-powers but is not recurrent (a word is recurrent if every factor occurs infinitely often).

Adding the hypothesis of recurrence we proved the following

Proposition

There exist recurrent aperiodic words avoiding 6-anti-powers.

Let $w_0 = 0$ and $w_n = w_{n-1} 1^{3|w_{n-1}|} w_{n-1}$ for every $n > 0$. The infinite word w obtained as the limit of the sequence of words $(w_n)_{n \geq 1}$ is recurrent and avoids 6-anti-powers.

Do aperiodic recurrent words exist that avoid anti-powers of order k for $k = 4, 5$?

Ramsey-like Version

Another consequence of our main result is the following finite version of our result.

Theorem

For all integers $l > 1$ and $k > 1$ there exists an integer $N = N(l, k)$ such that every word of length N contains an l-power or a k-anti-power. Furthermore, for $k > 2$, one has $k^2 - 1 \leq N(k, k) \leq k^3\binom{k}{2}$.

Ramsey-like Version

Another consequence of our main result is the following finite version of our result.

Theorem

For all integers $l > 1$ and $k > 1$ there exists an integer $N = N(l, k)$ such that every word of length N contains an l-power or a k-anti-power. Furthermore, for $k > 2$, one has $k^2 - 1 \leq N(k, k) \leq k^3\binom{k}{2}$.

Thanks to Jeff Shallit, the sequence $N(k, k) = 1, 2, 9, 24, 55, \cdots$ is now in the Online Encyclopedia of Integer Sequences (sequence A274543).
To do: compute more terms.

Ramsey-like Version

Another consequence of our main result is the following finite version of our result.

Theorem

For all integers $l > 1$ and $k > 1$ there exists an integer $N = N(l, k)$ such that every word of length N contains an l-power or a k-anti-power. Furthermore, for $k > 2$, one has $k^2 - 1 \leq N(k, k) \leq k^3\binom{k}{2}$.

Thanks to Jeff Shallit, the sequence $N(k, k) = 1, 2, 9, 24, 55, \cdots$ is now in the Online Encyclopedia of Integer Sequences (sequence A274543). To do: compute more terms.

A future direction of investigation consists in improving the bounds on these numbers.

OEIS

Other interesting sequences are present in the OEIS:

OEIS

Other interesting sequences are present in the OEIS:

A274449: Number of binary words of length n having the minimum possible number of different anti-power periods:
2, 2, 8, 4, 32, 2, 128, 16, 176, 32, 2048, 4, 8192, 128, ...

OEIS

Other interesting sequences are present in the OEIS:

A274449: Number of binary words of length n having the minimum possible number of different anti-power periods:
2, 2, 8, 4, 32, 2, 128, 16, 176, 32, 2048, 4, 8192, 128, ...

A274450: Largest number of anti-power periods possible for a binary word of length n:
1, 2, 1, 2, 1, 3, 1, 3, 2, 2, 1, 4, 1, 2, 3, 3, 1, 4, 1, ...

OEIS

Other interesting sequences are present in the OEIS:

A274449: Number of binary words of length n having the minimum possible number of different anti-power periods:
2, 2, 8, 4, 32, 2, 128, 16, 176, 32, 2048, 4, 8192, 128, ...

A274450: Largest number of anti-power periods possible for a binary word of length n:
1, 2, 1, 2, 1, 3, 1, 3, 2, 2, 1, 4, 1, 2, 3, 3, 1, 4, 1, ...

A274451: Number of possible sets of anti-power periods for binary words of length n:
1, 2, 1, 2, 1, 4, 1, 3, 2, 2, 1, 6, 1, 2, 4, 3, 1, 5, 1, ...

OEIS

Other interesting sequences are present in the OEIS:

A274449: Number of binary words of length n having the minimum possible number of different anti-power periods:
2, 2, 8, 4, 32, 2, 128, 16, 176, 32, 2048, 4, 8192, 128, ...

A274450: Largest number of anti-power periods possible for a binary word of length n:
1, 2, 1, 2, 1, 3, 1, 3, 2, 2, 1, 4, 1, 2, 3, 3, 1, 4, 1, ...

A274451: Number of possible sets of anti-power periods for binary words of length n:
1, 2, 1, 2, 1, 4, 1, 3, 2, 2, 1, 6, 1, 2, 4, 3, 1, 5, 1, ...

A274457: Shortest possible anti-power period of a binary word of length n:
1, 1, 3, 2, 5, 2, 7, 2, 3, 5, 11, 3, 13, 7, 3, 4, 17, ...

OEIS

Other interesting sequences are present in the OEIS:

A274449: Number of binary words of length n having the minimum possible number of different anti-power periods:
2, 2, 8, 4, 32, 2, 128, 16, 176, 32, 2048, 4, 8192, 128, ...

A274450: Largest number of anti-power periods possible for a binary word of length n:
1, 2, 1, 2, 1, 3, 1, 3, 2, 2, 1, 4, 1, 2, 3, 3, 1, 4, 1, ...

A274451: Number of possible sets of anti-power periods for binary words of length n:
1, 2, 1, 2, 1, 4, 1, 3, 2, 2, 1, 6, 1, 2, 4, 3, 1, 5, 1, ...

A274457: Shortest possible anti-power period of a binary word of length n:
1, 1, 3, 2, 5, 2, 7, 2, 3, 5, 11, 3, 13, 7, 3, 4, 17, ...

A275061: Number of binary words of length n avoiding 4-anti-powers:
1, 2, 4, 8, 16, 32, 64, 128, 232, 432, 808, 1512, ...

Abelian Anti-Powers

A natural generalization of the notion of an anti-power is that of an abelian anti-power.

Abelian Anti-Powers

A natural generalization of the notion of an anti-power is that of an abelian anti-power.

Recall that the Parikh (or composition) vector of a finite word w is the vector that counts the multiplicities of the letters of the alphabet in w, e.g. $P(abaac) = (3, 1, 1)$.

Abelian Anti-Powers

A natural generalization of the notion of an anti-power is that of an abelian anti-power.

Recall that the Parikh (or composition) vector of a finite word w is the vector that counts the multiplicities of the letters of the alphabet in w, e.g. $P(abaac) = (3, 1, 1)$.

Definition

An abelian anti-power of order k (or simply an abelian k-anti-power) is a word of the form $x_1 x_2 \cdots x_k$ where the x_i have the same length and different Parikh vectors.

Abelian Anti-Powers

A natural generalization of the notion of an anti-power is that of an abelian anti-power.

Recall that the Parikh (or composition) vector of a finite word w is the vector that counts the multiplicities of the letters of the alphabet in w, e.g. $P(abaac) = (3, 1, 1)$.

Definition

An abelian anti-power of order k (or simply an abelian k-anti-power) is a word of the form $x_1 x_2 \cdots x_k$ where the x_i have the same length and different Parikh vectors.

Of course an abelian k-anti-power is a k-anti-power but the converse is not always true.

Abelian Anti-Powers

Conjecture

Every infinite word contains abelian powers of any order or abelian anti-powers of any order.

Abelian Anti-Powers

Conjecture

Every infinite word contains abelian powers of any order or abelian anti-powers of any order.

An infinite word may contain both abelian powers of any order and abelian anti-powers of any order. This is the case, for example, of any word with full factor complexity.

Abelian Anti-Powers

Conjecture

Every infinite word contains abelian powers of any order or abelian anti-powers of any order.

An infinite word may contain both abelian powers of any order and abelian anti-powers of any order. This is the case, for example, of any word with full factor complexity.

A uniformly recurrent example is given by the regular paperfolding word

$$p = 00100110001101100010011100110110\cdots$$

Proposition (G.F., M. Postic, M. Silva)

The regular paperfolding word p contains abelian powers of any order and abelian anti-powers of any order.

The fact that p contains abelian powers of any order was proved in 2013 by Štěpán Holub (J. Combin. Theory Ser. A, 120).

Algorithms

It might be useful to be able to locate efficiently anti-power factors in a finite word. For this, we proved the following

Theorem (G. Badkobeh, G.F., S. Puglisi, 2018)

Given a word w of length n and an integer $k > 1$, there is an $O(n^2/k)$ time and $O(n)$ space algorithm that locates all the factors of w that are k-anti-powers.

Algorithms

The algorithm is clearly optimal in the case of an unbounded alphabet (think of a word made by all-distinct letters). What can we say in the case of a finite alphabet?

Algorithms

The algorithm is clearly optimal in the case of an unbounded alphabet (think of a word made by all-distinct letters). What can we say in the case of a finite alphabet?

For every positive integer m, we let w_m denote the word obtained by concatenating the binary expansions of integers from 0 to m followed by a symbol $\$$. So for example

$$w_5 = 0\$1\$10\$11\$100\$101\$$$

We have that $|w_m| = \Theta(m \log m)$. Let us write $n = |w_m|$.

Algorithms

The algorithm is clearly optimal in the case of an unbounded alphabet (think of a word made by all-distinct letters). What can we say in the case of a finite alphabet?

For every positive integer m, we let w_m denote the word obtained by concatenating the binary expansions of integers from 0 to m followed by a symbol \$. So for example

$$w_5 = 0\$1\$10\$11\$100\$101\$$$

We have that $|w_m| = \Theta(m \log m)$. Let us write $n = |w_m|$.

Lemma

Every word w_m of length n contains $\Omega(\frac{n^2}{k})$ anti-powers of order k.

Algorithms

The algorithm is clearly optimal in the case of an unbounded alphabet (think of a word made by all-distinct letters). What can we say in the case of a finite alphabet?

For every positive integer m, we let w_m denote the word obtained by concatenating the binary expansions of integers from 0 to m followed by a symbol $\$$. So for example

$$w_5 = 0\$1\$10\$11\$100\$101\$$$

We have that $|w_m| = \Theta(m \log m)$. Let us write $n = |w_m|$.

Lemma

Every word w_m of length n contains $\Omega(\frac{n^2}{k})$ anti-powers of order k.

So our algorithm is optimal even for finite alphabets.

Conclusions

- We introduced the notion of an anti-power, and proved that containing powers of any order or anti-powers of any order is a new unavoidable regularity for infinite words.

Conclusions

- We introduced the notion of an anti-power, and proved that containing powers of any order or anti-powers of any order is a new unavoidable regularity for infinite words.
- We think there is space for extensions of our results to other problems in combinatorics on words.

Conclusions

- We introduced the notion of an anti-power, and proved that containing powers of any order or anti-powers of any order is a new unavoidable regularity for infinite words.
- We think there is space for extensions of our results to other problems in combinatorics on words.
- More generally, we think that our approach may be worth to be considered for other combinatorial structures.

Conclusions

- We introduced the notion of an anti-power, and proved that containing powers of any order or anti-powers of any order is a new unavoidable regularity for infinite words.
- We think there is space for extensions of our results to other problems in combinatorics on words.
- More generally, we think that our approach may be worth to be considered for other combinatorial structures.
- We provided an optimal algorithm to locate anti-powers in a finite word. We think there are other interesting algorithmic questions related to anti-powers.

References

- G. Fici, A. Restivo, M. Silva, L. Zamboni:
  Anti-Powers in Infinite Words
  Journal of Combinatorial Theory, Series A, to appear
  (Preliminary version in ICALP 2016)

- G. Badkobeh, G. Fici, S. Puglisi:
  Algorithms for Anti-Powers in Strings
  Information Processing Letters, to appear

Thank you