Periodicity and order

A. Belov, M. Kharitonov

August 10, 2012

1 Introduction

We consider words in some alphabet. We fix some word W. If we are very lucky, W is periodic, i.e. it is some short word w repeated many times. This word w is called *period* of W. We write $c = ba$, for words a, b, c, if c is obtained from b and a by gluing together (we note that any word is a product of several letters an that this product is noncommutative, i.e. $ab \neq ba$, even if a, b are letters). From the point of view of this product, any periodic word is a power of some (short) word.

An arbitrary word does not tend to be periodic. More often a word is a product of several periodic words. We call such words *piecewise-periodic*. Any word can be expressed as a product of several periodic words and some small pieces inbetween them.

Let now fix an alphabet $\mathcal{A} = \{a_1, \dots, a_s\}$. Order $a_1 \prec a_2 \prec \cdots \prec a_s$ on the set of letters induces a lexicographical order on the set of words. We write $U \prec V$, whenever the first letter of U is smaller than the first letter of V, if they coincide, if the second letter of U is smaller than the second letter of V e.t.c.

If a word U starts from a word V, U and V are *incomparable*, i.e. neither $U \prec V$, nor $V \prec U$. Similarly words are ordered in a dictionary (in this case the shortest from two incomparable words usually goes earlier). A given word A may have a property to be *lexicographically ordered*: this means that a letter with a smaller index always goes earlier than a letter with a bigger index. If a word W is not ordered, it has several *disorders*. There are more powerful and less powerful disorders. We say that a word W has a *k-disorder* if W contains k subwords $W_1, ..., W_k$ such that
1) W_i does not intersect W_j for $i \neq j$;
2) if $i < j$, then W_i goes earlier than W_j;
3) if $i < j$, then $W_i \succ W_j$.
By definition, a word W is *k-divisible*, if W has a k-disorder.

It appear so that the "degree of divisibility" of a word W and the number of periodic subwords of W are closely related. In this project we convert this metamathematical statement to a theorem and try to amuse participants by related facts. We present this theorem right now.

Shirshov's Theorem of height. The set of non k-divisible words is piecewise-periodic with period $(k-1)$ or less, i.e. there exists a function $H(k,s)$ such that any word is either k-divisible, or can be splitted on $H(k,s)$ pieces such that any piece is a subword of a periodic word with a length of period $(k-1)$ or less.

The main goal of this project is to produce estimates for the function $H(k,s)$ of type $H(k,s) \leq sk^{C\ln k}$, where C is some constant, which does not depend on k and s. To achieve this goal we will need some deep combinatorial results, in particular Dilworth's theorem. Another goal is to count polylinear [1] words, which have no k-disorders [2]. We also estimates the number of subwords with a given period, and here the graphs point of view is useful, in particular we explain a connection with a Ramsey's theory.

The key problems in the first problem set are 2.10, 3.7, 4.11, 4.12.

We also attach to a project a problem set called "Application to ring theory and some history", which is dedicated to the mentioned topics. This problem set is completely independent from all other problem sets, but it provides an area of mathematics, for which the results of our project are also very interesting. It turn's out that the Shirshov's theorem allows to solve some problems of this area, which seems to be completely nonrelated to it and stay unsolved for more than 20 years.

2 De arte kombinatoria

Problem 2.1. Karlsson know how to write only words which does not contain subwords with two or more different letters. How many words of length n Karlsson know how to write, if his alphabet contains l letters?

Problem 2.2. The dictionary of Winnie-the-Pooh tribe has 20 letters. The language of this tribe consider as a phrase any combination of this words. There exists two verbal spells "earth stay on Great Crocodile" and "every evening Crocodile eat Sun", which evoke an earthquake. How many phrases from 10 words does not provide the tribe with an earthquake?

Problem 2.3. The alphabet of a small-wide-tribe "Smeshariki" consists of l letters. May the language of this tribe contains a word of length l which contains precisely

a) $l+1$

b) $\frac{l(l-1)}{2}-1$

c*) $2l$

different subwords?

Problem 2.4. An alphabet of Endeans has N letters, and any word of Endeans consists of letters of their alphabet. It is known that a word W repeated twice means the same with W, and that the meaning of a word $W_1W_2W_3$ is the same with $W_1W_2W_2W_3$. For example "BC" means the same with "BBC". Prove that the number of words with different meanings is finite if

a) $N=2$;

b) $N=3$.

By u^t we denote t copies of a word u putted in one line.

Problem 2.5. Fix an alphabet $\mathcal{A}=\{a,b\}$. What is the minimal number of words $\{W_1,...,W_k\}$ such that the set of all words of length 100, which does not contain $\{W_1,...,W_k\}$ as subwords, consists of $(ab)^{50}$ and $(ba)^{50}$?

Problem 2.6. Let $k,t\in\mathbb{Z}_{\geq 1}$. Prove that, if a word V of length $k\cdot t$ has not more than k different subwords of length k, then for some word v the word V contains a subword v^t.

[1] i.e. such words W that any letter is used in W not more than once

[2] For example, the number of words which are not 3-divisible equals to a Catalan's number

Problem 2.7. Provide a bijection between the following sets:

- sequences of natural numbers $1 \leqslant a_1 \leqslant a_2 \leqslant \ldots \leqslant a_n$, where $a_i \leqslant i$;
- transpositions of numbers $1, 2, \ldots, n$, such that the length of any decreasing sequence is 2 or less.

Problem 2.8. Hundred man-eaters come to a feast. During a feast man-eaters eat themselves. Therefore appear a sequences of man-eaters such that a man-eater eats a man-eater which eats a man-eater which eats a man eater... What is the smallest possible the longest such sequence of man-eaters with additional condition that from any 10 man-eaters any one eats the other one?

Similar problems to Problem 2.8 appear in subsection "Dilworth's theorem".

Definition 2.1. *We call a word u* non-cyclic*, if u is not equal to v^k, for any word v and any $k > 1$.*

Problem 2.9. Let u, v be different non-cyclic words of length m and n respectively. Assume that a word W contains subwords $u' = u^{m \cdot n}$ and $v' = v^{m \cdot n}$. Prove that the length of the common part of u' and v' does no exceed $m + n - 2$.

Problem 2.10. An infinite band is filled by numbers $\{1,..., 9\}$. Prove that either one can cut of from it 10 non-intersecting numbers with 1000 digits each, which form an increasing sequence on a band, or there exists a number with 10 or less digits which repeats 50 times in succession.

3 Dilworth's theorem

Problem 3.1. Is it true that, for any sequence of numbers of length 5, there exists a subsequence of length 3 which is ordered (i.e. it is or increasing, or decreasing)?

Problem 3.2. Is it true that, for any sequence of numbers of length 9, there exists a subsequence of length 4 which is ordered (i.e. it is or increasing, or decreasing)?

Problem 3.3. Prove that that for any sequence of numbers of length 10 there exists a subsequence of length 4 which is ordered (i.e. it is or increasing, or decreasing).

Problem 3.4. Prove that that for any sequence of numbers of length $mn+1$ either there exists a decreasing subsequence of length $m+1$ or there is an increasing subsequence of length $n+1$.

By definition a *partially ordered set* (POS) is a set M with a relation $\prec$ on it such that, for any two elements a and b of M, or $a \prec b$ is true or is false. This relation should satisfies the following axioms:
1. if $a \prec b$ and $b \prec c$, then $a \prec c$ (transitivity);
2. if $a \prec b$, then a is not b.

Problem 3.5. May $a \prec b$ and $b \prec a$ be true simultaneously?

Problem 3.6. Prove that the set of words with a lexicographical order is a POS.

Definition 3.1. *A POS M, for any elements $a, b \in A$ of which $a = b$, $a \prec b$, $b \prec a$, is called linearly ordered. Such POS are also known as chains.*

Problem 3.7. Let m, n be natural numbers. Prove that in any POS with $mn+1$ elements there exists either a subset with $m+1$ elements which is a chain or there is a subset with $n+1$ elements which is an antichain (i.e. such that any two elements of which are incomparable).

Problem 3.8. Let M be a POS and $c(M)$ be the length of the longest chain of M. Then M can be splitted on $c(M)$ antichains.

The following theorem is in some sense dual to Problem 3.8.

Dilworth's theorem. Let M be a POS and $ad(M)$ be the length of the longest antichain of M. Then M can be splitted on $ad(M)$ chains.

4 Exponential estimates

We fix an alphabet $\mathcal{A} = \{a_1, a_2, \dots, a_l\}$ and we fix a linear order on $\mathcal{A} : a_1 \prec a_2 \prec \cdots \prec a_l$. This order introduces a lexicographical order on the set of words of $\mathcal{A}$. We consider two words u and v. If u begins from v or v begins from u, we call u and v *incomparable* (with respect to each other). Otherwise there exist words w, u', v' such that $u = wu'$, $v = wv'$ and first letters of u' and v' are different (w could be an empty word). If the first letter of u' is greater than the first letter of v', we say that u is greater than v and write $v \prec u$, otherwise we say that v is greater than u and write $v \prec u$. The set of words of $\mathcal{A}$ with respect to $\prec$ is a POS. The order $\prec$ is called *lexicographical* (see also Introduction). It would be significant later that some words are incomparable with respect to the lexicographical order $\prec$.

Problem 4.1. Let alphabet $\mathcal{A}$ consists of letters a, b, c. We introduce an order on them: $a \prec b \prec c$. Find the longest increasing sequence from the following list of words. Which pairs of these words are incomparable?

$$cb, abc, bac, abb, b, ccc, abc$$

The following definitions will be useful later.

Definition 4.1. *A word W is called n-divisible, if there exist words $u_1, ..., u_n$ such that $W = v \cdot u_1 \cdots u_n$ and $u_1 \succ \ldots \succ u_n$.*

Definition 4.2. *A word W is called k-ordered, if W is k-divisible but not $(k+1)$-divisible.*

Problem 4.2. Find a number of a) 1-ordered b) 2-ordered words of length l.

Problem 4.3. Let n be a natural number, u — noncyclic word of length n or less. Prove that the word u^{2n} is not n-divisible.

Definition 4.3. *a) A word v is called a tale of a word u if there exists a word w such that $u = vw$.*

b) If a word v has a subword u^t we say that v has a period of cyclicity t

Problem 4.4. Let l, d be some natural numbers. Prove that, for a word W of length l, or first $[l/d]$ tales are pairwise incomparable, or W has a period of length d.

Further we assume that $n \leqslant d$.

Definition 4.4. *A word W is called* (n, d)-cancellable, *either if W is n-divisible, or if there exists a word u such that u^d is a subword of W.*

Problem 4.5. Prove that if a word W has n pairwise equal non-intersecting subwords of length n, then W is (n, n)-cancellable.

Definition 4.5. *A word W is called n-divisible from tale, is there exists tales $u_1, \ldots, u_n$ such that $u_1 \succ u_2 \succ \ldots \succ u_n$, and, for any $i = 1, 2, \ldots, n-1$, u_i begins earlier than u_{i+1}.*

Problem 4.6. Prove that if a word W is
a) n^3d-divisible from tale,
b) $3n^2d$-divisible from tale,
c) $4nd$-divisible from tale,
then W is either n-divisible, or W has a subword u^d for some nontrivial word u.

Problem 4.7. For any pair of natural numbers (n, d) (except pair (1, 1)), provide an example of a word W of length $(nd-1)$ such that any set of tales of W is not increasing and W is not $(n+1, d)$-cancellable.

Problem 4.8. Try to enhance an estimate from Problem 4.6.

We fix an alphabet $\mathcal{A}$ of length l, a word W of this alphabet of length $r(W)$ and natural numbers $n \leqslant d$.

Further we assume that, for all words u, W does not contain a subword u^d and W is not $4nd$-divisible from tale. We consider first $[r(W)/d]$ tales of W (further we denote this set of words by Ω). Then by Dilworth's theorem we can split Ω on $(4nd-1)$ groups such that tales in one group form a chain.

In solutions of the following problems we expect that you use previous ones.

Problem 4.9. Prove that, from any $4nd^2$ tales of Ω, there exists two tales, for which the first subwords of length $4nd$ are pairwise different.

Problem 4.10. In some infinite parliament any member has not more than 3 enemies among members. Prove that there is a way to divide this parliament into two houses such that a member of any house has not more than 1 enemy among members of his own house.

In the following problems we assume that l, n, d are variables and that W is some word defined over alphabet of l-letters.

Problem 4.11 (Shirshov's lemma)**.** Prove that there exists a function $f(l, n, d)$ such that for any subword W defined over alphabet of l-letters either W is (n, d)-cancellable, or $r(W) < f(l, n, d)$.

Problem 4.12. Prove that $f(l, n, d) < l(4nd)^{4nd+2}$.

5 Advanced estimates

We fix natural numbers s, n, d such that $d \geq n$. We also fix an alphabet $\mathcal{A} := \{a_1, ..., a_s\}$ and a word W of this alphabet of length $r(W)$ such that W is not (n, d)-cancellable (i.e. either W is n-divisible or W has a subword u^d for some non-zero word u). Let $\Omega_d(W)$ be the set of tales of W which begins from the first $[r(W)/d]$-letters of W. By text after Problem 4.8 the tales of Ω are splitted on $4nd - 1$ groups which we call *colors*. We set $p_{n,d} := 4nd$. By definition *a k-start* of a word W is a subword of W which consists from the first k letters of W.

Problem 5.1. Assume that $\Omega_d(W)$ has a subset Ω_c with $p_{n,d}^2 + p_{n,d} + 1$ elements such that all elements of Ω_c have the same color and such that any two elements of Ω_c have different $2k$-starts for some number k. Prove that there exist two subwords of Ω_c which have pairwise different k-starts.

Problem 5.2. Prove that for any subset Ω_c of $\Omega_d(W)$ with $(d+1)p_{n,d}^4$ elements there exist two elements of Ω_c which have the same color and have different $\frac{p_{n,d}}{2}$-starts.

Problem 5.3. Prove that for any subset Ω_c of $\Omega_d(W)$ with $(d+1)p_{n,d}^7$ elements there exist two elements of Ω_c which have the same color and have different $\frac{p_{n,d}}{4}$-starts.

Assume that $p_{n,d} = 2^t$ for some natural number t.

Problem 5.4. Prove that for any subset Ω_c of $\Omega_d(W)$ with $(d+1)p_{n,d}^{1+3t}$ elements there exist two elements of Ω_c which have different 1-starts.

Problem 5.5. Prove that if the length of W is greater than $(d+1)p_{n,d}^{2+3t}s$, then W is either n-divisible or has a subword u^d for some non-zero word d.

Using terms of problem set "Application to ring theory and some history" and problem 5.5 one can solve the problem posed by E. Zelmanov in 1991:

Let $F_{2,m}$ be a free 2-generated associative ring with the identity $x^m = 0$. Is it true that the nilpotency class of $F_{2,m}$ depends exponentially on m?

Assume that $n = 2^q$ for some natural number q.

Problem 5.6. Let $\Gamma_{n,s}^1$ be a finite set which consists from all words of length n or less over alphabet $\mathcal{A}$ with s letters. Let $\Gamma_{n,s}$ be an infinite set which contains $\Gamma_{n,s}^1$ and all powers of all elements of $\Gamma_{n,s}^1$. Prove that if W is not n-divisible, then W is a product of not more than n^{100q} words from $\Gamma_{n,s}$.