Complexity of Unification in Free Groups and Free Semi-groups

Antoni Kościelski[*] Leszek Pacholski[†]

Abstract The exponent of periodicity is an important factor in estimates of complexity of word-unification algorithms. We prove that the exponent of periodicity of a minimal solution of a word equation is at most $2^{2.54n}$, where n is the length of the equation. Since the best known lower bound is $2^{0.31n}$ our upper bound is almost optimal and exponentially better than the original bound $(6n)^{2^{2n^4}} + 2$. Thus our result implies exponential improvement of known upper bounds on complexity of word-unification algorithms. Moreover we give some evidence that, contrary to the common belief, the algorithm deciding satisfiability of equations in free groups, given by Makanin in not primitive recursive.

The proofs are only sketched here. More details will be given in the full version.

0 Introduction.

In this note we improve the known upper bound on the exponent of periodicity thus obtaining exponential speed-up of several word unification algorithms. We also comment on the complexity of the Makanin's algorithm deciding satisfiability of equations in free groups.

By $\mathbf{N}$ we denote the set of non-negative integers, $\mathbf{N}^+$ is the set of positive integers. Given any non empty set Σ by Σ^* we denote the set of all words in Σ. Σ^+ is the set of non-empty words in Σ. If W is a word, then $|W|$ denotes the length of W. ε is the empty word. Let Σ, Ξ be two disjoint, non-empty, finite alphabets. $\Sigma = \{a_1, \ldots, a_n\}$ is the set of (constant) letters and $\Xi = \{x_1, \ldots, x_n\}$ is the set of variable letters. A word equation in (Σ, Ξ) is a pair $\mathcal{E} = (W_1, W_2)$ of words in $(\Sigma \cup \Xi)^*$, also denoted by $W_1 = W_2$. A solution of $\mathcal{E}$ is a function $v : \Sigma \to \Xi$ such that $W_1(v(x_1)/x_1, \ldots, v(x_m)/x_m) = W_2(v(x_1)/x_1, \ldots, v(x_m)/x_m)$, where $W(v(x_i)/x_i)$ denotes the word obtained from W by replacing each occurrence of x_i by $v(x_i)$. The length of a solution v is $\sum_{i=1}^{m} v(x_i)$. A solution is minimal if it has minimal length.

It was shown by Makanin [MA1], that the problem if a word equation has a solution is decidable. Later related variants of word-unification problem, namely the problems of finding a solution, finding a minimal solution and finding all minimal solutions, were studied by various researchers (see e.g. [APE], [PEC], [JAF]). Moreover, some variants of Makanin's algorithm have been implemented (see [ABD]).

In [JAF] Jaffar gave a procedure generating for a word equation $\mathcal{E}$ the minimal and complete collection of unifiers. This procedure stops with a positive answer when $\mathcal{E}$ is satisfiable. To stop the procedure in the case when $\mathcal{E}$ is not satisfiable a bound B, depending on the size of $\mathcal{E}$, is placed on the length of each path of the reduction tree, B being an increasing function of the exponent of periodicity of a minimal solution of $\mathcal{E}$. Thus the number of steps of the generation procedure of Jaffar will, in the case of unsatisfiable equation depend on the known bounds on the periodicity exponent.

In spite of the fact that the algorithm of Makanin and its variants seem to have important applications and have been intensively studied, no serious investigations of their complexity have been undertaken. It seems that the understanding of the nature of the algorithm of Makanin is still very low.

This paper contains a report on an attempt to understand the complexity of the Makanin's algorithm for semi-groups and the complexity of the problem of solvability of word equations. An important factor in estimates of the complexity of the Makanin's algorithm is the periodicity exponent of a minimal solution of a word equation. A periodicity exponent of a word W is the maximal integer p such that $W = U_1 U^p U_2$ for some non-empty word U. An important fact used in the Makanin's algorithm and its variants is that the periodicity exponent can be bounded by a recursive function of the length of an equation. In fact V.K.Bulitko [BUL] proved that if n is the length of an equation, then the index of periodicity of its minimal solution does not exceed $(6n)^{2^{2n^4}} + 2$.

In [KPA] we forced this bound down to n^{2n^4}. The method, we have used, was based on Makanin's reduction lemma and consisted of obtaining better bounds on the size of minimal positive integer solutions of sets of linear diophantine equations. The bounds we have obtained are close to the ones obtained recently by E.Bombieri and J.Vaaler [BVA] for minimal absolute values of integer solutions (not necessarily positive) of such equations, and seem to be close to optimal. On

[*] Institute of Computes Science, Wrocław University, Przesmyckiego 20, 51-151 Wrocław, Poland

[†] Institute of Mathematics, Polish Academy of Sciences, Kopernika 18, 51-617 Wrocław, Poland

the other hand we gave a lower bound $2^{0.31n}$, which we believe is the best, so it was evident that some further work was necessary, and that the problem could not be solved by an analysis of general diophantine linear equations alone.

Here we give a further improvement of the upper bound to $2^{2.54n}$. The paper is divided into four parts. In the first we study presentations of words in a special form and prove the uniqueness of such presentations. In the second we construct a set $\mathcal{N}$ of linear diophantine equations whose minimal solutions describe the periodicity exponent of a minimal solution of a word equation $\mathcal{E}$. The third part gives an upper bound on the size of minimal solutions of this set of linear diophantine equations.

The problem of solvability of word equations can, using a different terminology, be rephrased as a problem of solvability of equations in a free (finitely generated) semi-group. In 1983 Makanin ([MA2]) proved that a similar, but much more difficult, problem of solvability of an equation in a free group is also solvable. The problem of generation of all minimal solutions of equations in a free group was studied by A.A. Razborov ([RAZ]).

In the fourth part we comment on the complexity of Makanin's algorithm for free groups. We argue that, contrary to the common belief, the algorithm given in [MA2] is not primitive recursive. In fact one can prove (see [KOS]) that the functions defined in [MA2] describing the number of iterations of elementary steps of Makanin's algorithm are not primitive recursive. However, it is impossible to give a concise and comprehensive proof of this statement without copying a large part of Makanin's paper, since definitions of the actual functions used in [MA2] are quite complicated and are mixed with the proof of correctness of algorithm and with algebraic arguments. Instead, we introduce a notion of an abstract Makanin's algorithm which, in our opinion, describes the main algorithmic properties used to prove the decidability of the satisfiability problem for equations in free groups. Then we prove the halting property of abstract Makanin's algorithms. Moreover, we prove that there are abstract Makanin's algorithms which are not primitive recursive. Since, as we believe, the notion of an abstract Makanin's algorithm describes the properties of the Makanin's algorithm, this gives some evidence that the algorithm given in [MA2] is not primitive recursive.

1 Presentation of words.

Here we prove some facts necessary to obtain a reduction of a problem concerning word equations to some problem concerning linear diophantine equations.

Lemma 1.1. For any words W_1, W_2, if $W_1W_2 = W_2W_1$, then $W_1 = U^m$, and $W_2 = U^n$ for some word U and integers m, n.

Definition 1.2. A non-empty word U is simple if $U \neq V^n$, for every word V and every integer $n \geq 2$.

Lemma 1.3. Suppose that U is simple. Then,

(i) if $U^2 = U_1UU_2$, then either $U_1 = \varepsilon$ and $U_2 = U$, or $U_1 = U$ and $U_2 = \varepsilon$

(ii) if $U^3 = U_1UU_2$, then either $U_1 = U^2$ and $U_2 = \varepsilon$, or $U_1 = U_2 = U$, or $U_1 = \varepsilon$ and $U_2 = U^2$.

Definition 1.4. Let n be a positive integer and let P be a non-empty word. A sequence $(U_0, ..., U_n)$ is P-stable if

(i) for $i \leq n$, P^3 is not a subword of U_i,

(ii) for $0 < i < n$, $U_i \neq P$,

(iii) for $i < n$, P is a suffix of U_i and P^2 is not a suffix of U_i,

(iv) for $0 < i \leq n$, P is a prefix of U_i and P^2 is not a prefix of U_i.

Clearly if a sequence $(U_0, ..., U_n)$ is stable, then any subsequence of it is stable. Moreover $|U_i| > |P|$ for $0 < i < n$ and if $n > 0$, then $|U_0| \geq |P|$, and $|U_n| \geq |P|$.

Definition 1.5. Let $n \in \mathbf{N}$, $W_0, \ldots, W_n, P \in \Sigma^*$. Then $[W_0, \ldots, W_n]_P : (\mathbf{N}^+)^n \to \Sigma^*$ is the function such that
$[W_0, ..., W_n]_P(k_1, ..., k_n) =$
$= W_0P^{k_1}W_1P^{k_2}...P^{k_{n-1}}W_{n-1}P^{k_n}W_n$.
Since P is often fixed, we sometimes omitt the subcript P and write $[W_0, ..., W_n]$ instead of $[W_0, ..., W_n]_P$

Lemma 1.6. Let $P \in \Sigma^*$ be a simple word, and $u, v, k_1, ..., k_u, l_1, ..., l_v \in \mathbf{N}$. Assume that sequences $\vec{U} = (U_0, ..., U_u)$ and $\vec{V} = (V_0, ..., V_v)$ are P-stable. If

(1) $[U_0, ..., U_u]_P(k_1, ..., k_u) = [V_0, ...V_v]_P(l_1, ..., l_v)$,

then $u = v$, $k_i = l_i$ for each $i = 1, 2, ..., u$ and $U_i = V_i$ for each $i = 0, 1, ..., u$.

Proof. It is clear for $u = 0$, so assume that $0 < u \leq v$ and the conclusion of the Lemma holds for all stable sequences of length $\leq v$. Since (1) holds, there are words Y, Z such that

(2) $\quad U_0P^{k_1}Y = V_0P^{l_1}Z.$

First we are going to prove that $|U_0| = |V_0|$, which clearly implies that $U_0 = V_0$. Suppose that $|U_0| < |V_0|$. By the stability of $\vec{U}$ and $\vec{V}$, $U_0 = UP$ and $V_0 = VP$ for some words U, V. Clearly VP^3 has a prefix UP^3 and $|U| < |V|$. We also have, $|V| < |UP^2|$, since otherwise

$V_0 = VP$ would contain P^3 . C onsequently $V = UV'$, where $|V'| < |P^2|$. But $V'P^3$ has a prefix P^3, so P^3 has a prefix $V'P$. Now, by Lemma 1.3.(ii), either $V' = \varepsilon$, or $V' = P$ or $V' = P^2$. But it is easy to check that each of these cases gives a contradiction, so $|U_0| \geq |V_0|$. A similar argument shows that $|U_0| \leq |V_0|$, so $|U_0| = |V_0|$.

Now, we shall prove that $k_1 = l_1$. Suppose that $k_1 < l_1$. Since $U_0 = V_0$, (1) implies the existence of words X', Z' such that

(3) $\quad U_1Y' = P^rV_1Z'$ for some positive integer r.

Now, $|U_1Y'| \geq |P^2|$ and P^2 is not a prefix of U_1. Consequently $|U_1| < |P_2|$ and Y' is not the empty word. Therefore U_1 is not the last element of the stable sequence $\vec{U}$, so U_1 has a suffix P. But by (3) U_1 is a prefix of P^2 so by Lemma 1.3.(i) either $U_1 = \varepsilon$ or $U_1 = P$. It is clear that none of these cases hold, so we get a contradiction. In a similar way we prove that the assumption $k_1 > l_1$ also leads to a contradiction, which finishes the proof that $k_1 = l_1$. Now, using the inductive hypothesis and the fact that subsequences of stable sequences are stable, the conclusion of the lemma can easily be obtained.

Definition 1.7.

(i) A P-presentation of a word is a P-stable sequence $(U_0, ..., U_u)$ such that
$X = [U_0, ...U_u]_P(l_1, ..., l_u)$ for some $l_1, ..., l_u \in \mathbf{N}$.

(ii) The length of the P-presentation $(U_1, ...U_u)$ is u. A word X is of P-order u if it has a presentation of length u.

Lemma 1.8. For each simple word P each word W has a unique P-presentation.

Proof. Given any word a presentation is easily constructed. The uniquennes follows from Lemma 1.6.

Definition 1.9. If $(U_0, ..., U_u)$ is a P-presentation of X then we write $(U_0, ..., U_u) = [X]_P^{-1}$.

From now on we fix a simple word $P \in \Sigma^+$. We write order and presentation for P-order and P-presentation respectively. By $ord(X)$ we denote the P-order of X.

Lemma 1.10. If X, Y are of order 0 and XY has order > 0, then either

(i) XY has order 1 and $XY = [[XY]^{-1}](c)$ for some c such that $1 \leq c \leq 3$ or

(ii) XY has order 2 and $XY = [[XY]^{-1}](1, 1)$.

An easy proof is omitted.

Recall that Σ is a set of constant letters and $\Xi = \{x_0, ..., x_n\}$ is a set of variables. Assume that a function $v : \Xi \to \Sigma^+$ is given. For $i \leq n$ let $X_i = v(x_i)$. Let $j_{-1} = 0$ and for $k = 0, 1, ..., n$ let $j_k = \sum_{l=0}^{k} ord(X_l)$.

Definition 1.11. Let $t = j_n$. For every word $W \in (\Sigma \cup \Xi)^+$ we define a function $\{W\} : (\mathbf{N}^+)^t \to \Sigma^*$ as follows:

(i) for $\quad x_i \in \Xi$,
$\{x_i\}(l_1, ..., l_t) = [[X_i]^{-1}](l_{j_{i-1}+1}, ..., l_{j_i})$

(ii) for $\quad a \in \Sigma, \quad \{a\}(l_1, ..., l_t) = a$

(iii) for $\quad W \in (\Sigma \cup \Xi)^+ \quad$ and $\quad b \in (\Sigma \cup \Xi)$,
$\{Wb\}(l_1, ..., l_t) = \{W\}(l_1, ..., l_t)\{b\}(l_1, ..., l_t)$.

Lemma 1.12.

(i) If $\quad W = W_1W_2$, then
$\{W\}(l_1, ..., l_t) = \{W_1\}(l_1, ..., l_t)\{W_2\}(l_1, ..., l_t)$.

(ii) If $\quad x_i \in \Xi \quad$ and $\quad a \in \Sigma \quad$ then$\{x_ia\}(l_1, ..., l_t)$ can be expressed as
$[[X_ia]^{-1}](l_{j_{i-1}+1}, ..., l_{j_i})$ or
$[[X_ia]^{-1}](l_{j_{i-1}+1}, ..., l_{j_i} + 1)$ or
$[[X_ia]^{-1}](l_{j_{i-1}+1}, ..., l_{j_i}, 1)$.

Proof. Part (i) follows by a straightforward induction, part (ii) is easy.

Lemma 1.13. Let $\quad x_e, x_f \in \Xi, \quad (g_1, ..., g_u) = (l_{j_{e-1}+1}, ..., l_{j_e})$,
$(h_1, ..., h_v) = (l_{j_{f-1}+1}, ..., l_{j_f})$. Then
$\{x_ex_f\}(l_1, ..., l_t)$ can be expressed in one of the following forms:

$[[X_eX_f]^{-1}](g_1, ..., g_u, h_1, ..., h_v)$,
$[[X_eX_f]^{-1}](g_1, ..., g_u + c, h_1, ..., h_v)$,
$[[X_eX_f]^{-1}](g_1, ..., g_u, h_1 + c, ..., h_v)$,
$[[X_eX_f]^{-1}](g_1, ..., g_u + h_1 + c, ..., h_v)$,
$[[X_eX_f]^{-1}](g_1, ..., g_u, c, h_1, ..., h_v)$,
$[[X_eX_f]^{-1}](g_1, ..., g_u + 1, h_1 + 1, ..., h_v)$,
$[[X_eX_f]^{-1}](g_1, ..., g_u + 1, 1, h_1, ..., h_v)$,
$[[X_eX_f]^{-1}](g_1, ..., g_u, 1, h_1 + 1, ..., h_v)$,
$[[X_eX_f]^{-1}](g_1, ..., g_u, 1, 1, h_1, ..., h_v)$.

where $c, c' \in \mathbf{N}^+, \quad c \leq 3 \quad$ and $\quad 0 < c' < 3$.

Proof. The proof consists of a routine consideration of several possible forms of the last term of the presentation of X_e and of the first of the presentation of X_f.

Definition 1.14.

(i) If $ord(v(x)) = m$, then x is called a variable of order m.

(ii) For a fixed word $W \in (\Sigma \cup \Xi)^*$, a function $v : \Xi \to \Sigma^*$ and $i = 0, 1$, d_i is the number of variable letters of order i in W, d_2 is the number of variable letters of order > 1 in W and d_c is the number of constant letters in W.

Clearly $|W| = d_0 + d_1 + d_2 + d_c$;

In the lemma below variables w_j correspond to word variables of order 1, and variables x_j, y_j to word variables of order > 1. If $U = U_0 P^{k_1} U_1 P^{k_2} ... P^{k_u} U_u$, with $(U_0, ...U_u)$ P-stable, then we say that P^{k_1} and P^{k_u} are in boundary nesting and $P^{k_2}, ..., P^{k_{u-1}}$ are in internal nesting. The variables x_j correspond to boundary nesting of P and the variables y_j correspond to internal nesting of P.

Lemma 1.15. For every $W \in (\Sigma \cup \Xi)^*$, every simple $P \in \Sigma^*$, and every function $v : \Xi \to \Sigma^*$, there exists a sequence $\mathcal{L} = (L_1, ..., L_l)$ of linear functions $L_i = \sum c_{i,j} w_j + \sum c'_{i,j} x_j + \sum c''_{i,j} y_j + c_i$ with non-negative integer coefficients such that

(i) $\{W\}(\vec{w}, \vec{x}, \vec{y}) =$
$= [[v(W)]^{-1}](L_1(\vec{w}, \vec{x}, \vec{y}), ..., L_l(\vec{w}, \vec{x}, \vec{y}))$

(ii) $\sum_{i,j} c_{i,j} = d_1$, $\sum_{i,j} c'_{i,j} = 2d_2$, $c'_{i,j} \leq 1$, $c''_{i,j} \leq 1$.

(iii) $\forall i \quad card\{j : c'_{i,j} > 0\} \leq 2$

(iv) if for some i, j, $\quad c''_{i,j} > 0$, then $L_i = c''_{i,j} y_j$.

(v) $\sum_i c_i < 3(d_0 + d_1 + d_2) + d_c$.

Proof. The sequence $\mathcal{L}$ such that (i) holds is constructed by induction on the length of W using Lemmas 1.12 and 1.13. Properties (ii) - (v) easily follow from the construction.

2 Main reduction.

Assume we are given an equation $\mathcal{E} = (W, W')$ in (Σ, Ξ) and a fixed simple word $P \in \Sigma^*$. Let d_0, d_1, d_2, d_c denote the parameters fixed for W at the end of the last section, let d'_0, d'_1, d'_2, d'_c denote the corresponding parameters for W' and finally let $d_0^+ = d_0 + d'_0$, $\quad d_1^+ = d_1 + d'_1$, and so on. Assume that v is a solution of $\mathcal{E}$ and $V = v(W)$, $\quad V' = v(W')$. Then clearly $V = \{W\}(\vec{z}) = [[V]_P^{-1}]_P(L_1(\vec{z}), ..., L_l(\vec{z}))$ and $V' = \{W'\}(\vec{z}) = [[V]_P^{-1}]_P(L'{}_1(\vec{z}), ..., L'_l(\vec{z}))$. Since $V = V'$ and the function $[[\quad]^{-1}]$ is one to one, we get $l = l'$ and $L_i = L'_i$, for $0 < i \leq l$.

All linear expressions except the ones of the form y_i and c_i for $c_i \leq 3$ are called proper. An equation $L_i = L'_i$ is proper if either L_i or L'_i is proper.

Lemma 2.1. There is at most $d_0 + d_1 + 2d_2 + \frac{1}{2} d_c$ proper expressions in the set $\{L_1, ..., L_l\}$ and at most $d'_0 + d'_1 + 2d'_2 + \frac{1}{2} d'_c$ in $\{L'_1, ..., L'_l\}$.

Proof. It follows by calculation from the fact that variables in internal nesting can not appear in proper expressions.

Lemma 2.2. There is at most $d_0^+ + d_1^+ + 2d_2^+ + \frac{1}{2} d_c^+$ proper equations in the set
$\mathcal{L} = \{L_1 = L'_1, ..., L_l = L'_l\}$ of equations.

Proof. Immediate from Lemma 2.1.

Lemma 2.3. The system $\mathcal{L}$ can be transformed into a system
$\mathcal{M} = \{M_1(\vec{u}) = M'_1(\vec{u}), ..., M_m(\vec{u}) = M'_m(\vec{u})\}$ of linear diophantine equations with
$M'_i = \sum_j m_{i,j} u_j + m_i, \quad M_i = \sum_j m'_{i,j} u_j + m'_i$ such that:

(i) if p is a coordinate of a solution of $\mathcal{L}$ and $p > 3$, then p is a coordinate of a solution of $\mathcal{M}$,

(ii) $m \leq d_0^+ + d_1^+ + 2d_2^+ + \frac{1}{2} d_c^+$,

(iii) $\sum_{i,j} m_{i,j} + \sum_{i,j} m'_{i,j} \leq d_0^+ + 2d_1^+ + 4d_2^+ + \frac{1}{2} d_c^+$,
$\sum_i m_i + \sum_i m' - i \leq 3d_0^+ + 6d_1^+ + 9d_2^+ + d_c - 1$.

Proof. Let $\equiv_{\mathcal{L}}$ be the equivalence relation in the set of variables of order > 1 in internal nesting such that $y_i \equiv_{\mathcal{L}} y_j$ iff $(y_i = y_j) \in \mathcal{L}$. If $[y]$ is an $\equiv_{\mathcal{L}}$-equivalence class and for some $y_i \in [y]$, $\quad (y_i = c) \in \mathcal{L}$ then all non-proper equations are deleted from $\mathcal{L}$ and all occurences of variables in $[y]$ are replaced by c. Otherwise, if for every $y_i \in [y]$ and each constant c, $\quad (y_i = c) \notin \mathcal{L}$, then an element $y_i \in [y]$ is choosen, non-proper equations are deleted from $\mathcal{L}$ and all occurences in $\mathcal{L}$ of variales in $[y]$ are replaced by y_i.

Theorem 2.4. (Reduction Lemma). If $p > 6$ is an exponent of periodicity of a minimal solution of a word equation $\mathcal{E}$ of length d and such that $d_c^+ \geq 2$, then $p - 3$ is a coordinate of a minimal solution of a set $\mathcal{N} = \{N_1(\vec{q}) = 0, ..., N_s(\vec{q}) = 0\}$ of linear diophantine equations $N_i = \sum_j n_{i,j} q_j + n_i$, with $\vec{q} = (q_1, ..., q_r)$, such that

(i) $\quad s \leq 2d - 3$ and $r \leq 4d - 7$

(ii) $\quad \sum_{i,j} |n_{i,j}| \leq 4d - 7, \quad \sum_i |n_i| \leq 13d - 24$.

Proof. It follows by an easy calculation from Lemma 2.3 and the fact that $d = d_0^+ + d_1^+ + d_2^+ + d_c^+$.

3 The bounds.

Lemma 3.1. Assume $(a_{i,j})$ is a square matrix and $\sum_{i,j} |a_{i,j}| = A$. Then
$|det(a_{i,j})| \leq (e^{\frac{1}{e}})^A$.

Proof. Let r be the dimension of $(a_{i,j})$, an easy induction gives that $det(a_{i,j}) \leq \prod_i a_i$, where $a_i = \sum_j |a_{i,j}|$. Clearly $\prod_{i<r+1} a_i \leq \left(\frac{A}{r}\right)^r \leq (e^{\frac{1}{e}})^A$.

Definition 3.2. We say that a sequence $\vec{q} = (q_1, ..., q_r)$ is positive if for each
$0 < i \leq r, \quad q_i \geq 0$ and for some $0 < i \leq r \quad q_i > 0$. We write $\vec{q} > 0$ to denote that $\vec{q}$ is positive.

Theorem 3.3. If $\vec{q}_0 = (q_{0,1}, ..., q_{0,r}$ is a minimal positive integer solution of a system

$$\mathcal{L} = \left\{ \sum_{i=1}^{r} n_{i,j} q_i + n_j = 0 : \quad j = 1, ..., s \right\}$$

of linear diophantine equations, then for each $0 < i \leq r$ we have $q_{0,i} \leq (w + r)(e^{\frac{1}{e}})^c$, where $w = \sum_{j=1}^{s} |n_j|$ and $c = \sum_{j=1}^{s} \sum_{i=1}^{r} |n_{i,j}|$.

The proof of this theorem will be divided into lemmas. First we decompose a minimal solution into a sum of two vectors whose size will then be estimated.

Definition 3.4. Let $\mathcal{L}$ be the set of linear diophantine equations as in Theorem 3.3. We put,
$\mathbf{W} = \{\vec{q} \in \mathbf{R}^r : \vec{n}_i \vec{q} + n_i = 0$ for $i = 1, 2, ..., s, \quad \vec{q} \geq 0\}$,
$\mathbf{V} = \{\vec{q} \in \mathbf{R}^r : \vec{n}_i \vec{q} = 0 \quad$ for $\quad i = 1, 2, ..., s, \quad \vec{q} \geq 0\}$,
$\mathbf{A} = \{\vec{q} \in \mathbf{R}^r : q$ satisfies an independent set of r equations of the form $\vec{n}_i \vec{q} + n_i = 0, \quad q_i = 0\}$.

Theorem 3.5. (see [VZGS]) $\mathbf{W} = \mathbf{H}(\mathbf{A}) + \mathbf{V}$ where $\mathbf{H}(\mathbf{A})$ is the convex hull of $\mathbf{A}$ and $+$ denotes the complex sum.

Lemma 3.6. If $\vec{h} = (h_1, ..., h_r) \in \mathbf{H}(\mathbf{A})$, then for each $0 < i \leq r$ we have $h_i \leq w(e^{\frac{1}{e}})^c$.

Proof. We can assume that $\vec{h} \in \mathbf{A}$. For each $0 < i \leq r$, we have $|h_i| = |\frac{det(A_i)}{det(A)}|$ for some matrices A_i, A, so $|h_i| \leq |det(A_i)|$. To avoid double indexing assume that $\vec{h}$ is a solution of the set $\{\vec{n}_i \vec{q} + n_i = 0 : i = 1, ..., k\} \cup \{q_i = 0 : i = k + 1, ..., r\}$ and estimate h_1. Then clearly

$$A_1 = \begin{pmatrix} n_1 & n_{12} & \dots & \dots & \dots & n_{1r} \\ n_2 & n_{22} & \dots & \dots & \dots & n_{2r} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ n_k & n_{k2} & \dots & \dots & \dots & n_{kr} \\ 0 & 0 & \dots & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & 1 \end{pmatrix}.$$

Consequently $det(A_1) = \sum_{l=1}^{k} \pm n_l det(C_l)$ where for $0 < l \leq k$, C_l is a submatrix of $(n_{i,j})$. Since $\sum_{j=1}^{s} \sum_{i=1}^{r} |n_{i,j}| = c$, by Lemma 3.1 we get $|h_1| \leq |det(A_1)| \leq \sum_{l=1}^{k} |n_l| |det(C_l|) \leq \sum_{l=1}^{k} |n_l| (e^{\frac{1}{e}})^c \leq w(e^{\frac{1}{e}})^c$.

Lemma 3.7. There exists a set $\mathbf{G} \subset \mathbf{N}^r \cap \mathbf{V}$ such that

(i) $\mathbf{V}$ is a positive closure of $\mathbf{G}$ i.e. for every $\vec{v} \in \mathbf{V}$ there exists $\vec{q}_1, ..., \vec{q}_r \in \mathbf{G}$ and $\alpha_1, ..., \alpha_r \geq 0$ such that $\vec{v} = \sum_{i=1}^{r} \alpha_i g_i$.

(ii) If $\vec{g} \in \mathbf{G}, \quad \vec{g} = (g_1, ..., g_r)$ then $g_i \leq (e^{\frac{1}{e}})^c$.

Proof. Let

$$Q_k = (-1)^{k+1} det((a_{i,j})_{i=1,...,k-1,k+1,...,r; j=1,...,r-1}).$$

Clearly we have

$$\sum_{i=1}^{r} a_{i,j} Q_i = det \begin{pmatrix} a_{1,j} & a_{1,1} & a_{1,2} & \dots & a_{1,r-1} \\ a_{2,j} & a_{2,1} & a_{2,2} & \dots & a_{2,r-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{r,j} & a_{r,1} & a_{r,2} & \dots & a_{r,r-1} \end{pmatrix} = 0$$

The vectors $(Q_1, ..., Q_r)$ and $(-Q_1, ..., -Q_r)$ are called the standard solutions of the set
$\mathcal{U} = \{\sum_{i=1}^{r} a_{i,j} q_i = 0 : j = 1, ..., r - 1\}$.
We put
$\mathbf{G} = \{\vec{g} \in \mathbf{N}^r \cap \mathbf{V} : \vec{g}$ is the standard positive solution of a set of $r - 1$ independent equations of the form $\vec{n}_j \vec{q} = 0$, or $q_i = 0\}$.

Proof of Theorem 3.3. Let $\vec{h} \in \mathbf{H}(\mathbf{A}), \vec{g}_1, ..., \vec{g}_r \in \mathbf{G}$, and $\alpha_i \geq 0$ for $0 < i \leq r$. Clearly if $\vec{q} = \vec{h} + \sum_{i=1}^{r} \alpha_i \vec{g}_i$ is a minimal element of $\mathbf{W} \cap \mathbf{N} - \{(0, 0, ..., 0)\}$, then $\alpha_i < 1$ for each $0 < i \leq r$. But then for each $0 < j \leq r$ we have $q_j \leq h_j + \sum_{i=1}^{r} \alpha_i g_{i,j} \leq w(e^{\frac{1}{e}})^c + r(e^{\frac{1}{e}})^c = (w+r)(e^{\frac{1}{e}})^c$.

Theorem 3.8. If $\mathcal{E}$ is a word equation of length d and p is the exponent of periodicity of a minimal solution of $\mathcal{E}$, then $p \leq e^{1.81d} \approx 2^{2.54d}$. Moreover, if d is large, then $p \leq e^{1.49d} \approx 2^{2.14d}$.

Proof. By a combination of Theorem 3.3 and Theorem 2.4 we get that, if $p \geq 6$ and $d_c^+ \geq 2$, then $(17d - 31)(e^{\frac{1}{e}})^{4d-7}$, which immediately gives the second part of the thesis. To obtain the first part notice that for each d, $(17d - 30) \leq (e^{\frac{1}{e}})^{cd+7}$ where $c \approx 0.78$ is the solution of the equation $elog\frac{17}{c} - 7 = \frac{30c}{17}$.q

Example 3.9. (The lower bound) Consider the equation

$$x_1 b x_2 b ... b x_n = x_2 x_2 x_2 b x_3 x_3 x_3 b ... b x_n x_n x_n baaa$$

Clearly the length of this equation is $d = 6n - 2$ and it has the unique solution with $x_1 = a^{3^n}$, so, as it can easily be computed, the exponent of periodicity is $> 2^{0.316n}$.

4 Equations in free groups.

In this chapter we comment on the complexity of Makanins's algorithm for groups.

Definition 4.1.

(i) Let O be an infinite set of "objects", P a set of "parameters", $r : O \to O$ a "reduction function", $p : O \to P$ a parameter function, $s : O \to \mathbf{N}$ a "size function" and $Q \subset O$ a halting set.

(ii) Given $X \in O$ we put $X_0 = X$, $X_{n+1} = X_i$ if $X_i \in Q$ else $r(X_i)$. For $k \in \mathbf{N} \cup \{\infty\}$ we put $A_k(X) = \{p(X_i) : i < k\}$ and $A(X) = A_\infty(X)$.

(iii) For O, P, Q, r, p, s as above, by $MA(O, P, Q, r, p, s)$ we denote the algorithm

```
repeat
  $r(X)$
until  $X \in Q$.
```

Definition 4.2. $MA(O, P, Q, r, p, s)$ is an abstract Makanin's algorithm (AMA) if the following hold:

.1 for each $X \in O$, $A(X)$ is finite,

.2 for each $X \in O$, $p(r(X)) \neq p(X)$,

.3 there are two increasing functions $e, f : \mathbf{N} \to \mathbf{N}$ such that

[.3.1] for each $X \in O$ and each $k \in \mathbf{N} \cup \infty$ there exists an $a \in A_k(X)$ such that $card\{i : p(X_i) = a$, and $i \leq k\} \leq e(s(X))$

[.3.2] $f(0) > 0$ and $s(r(X)) \leq f(s(X))$.

e and f are called the complexity parameters of $AM(O, P, Q, r, p, s)$.

Theorem 4.3. If $AM = MA(O, P, Q, r, p, s)$ is an AMA, then AM terminates for each input $X \in O$. In fact there exists a recursive function $g : \mathbf{N}^2 \to \mathbf{N}$ such that for every $X \in O$, $g(card(A(X)), s(X))$ is the upper bound on the number of iterations of r in AM on the input X.

Proof. Recursively we define $g : \mathbf{N}^2 \to \mathbf{N}$ and $h : \mathbf{N}^3 \to \mathbf{N}$ as follows

(i) $g(0, k) = 1$,

(ii) $h(0, t, k) = g(t, k)$,

(iii) $h(j+1, t, k) = h(j, t, k) + 1 + g(t, f^{h(j,t,k)+1}(k))$,

(iv) $g(t+1, k) = h(e(k), t, k)$,

where f^x denotes f iterated x-times. It is not difficult to show by induction on t that

(v) $h(j, t, k) \geq$ the number of executions of the reduction function r if the set $A(X)$ has cardinality $t+2$ and for some $a \in A(X)$, $card\{i : p(X_i) = a\} = j$ and $k = s(X)$.

(vi) $g(t, k) \geq$ the number of executions of the reduction function r for an input X if the set $A(X)$ has cardinality $t+1$ and $k = s(X)$.

Definition 4.4. The function g defined in Theorem 4.3 is called the time complexity of AM.

Theorem 4.5. For every increasing $e, f : \mathbf{N} \to \mathbf{N}$ with $f(0) > 0$ there exists an AMA with complexity parameters e, f whose time complexity is not primitive recursive.

Proof. Let Ack be the Ackermann function and let g, h be the functions defined in the proof of Theorem 4.5. It is easy to prove by induction on n, that if $Ack(m, n) < g(m+1, n+1)$, then $Ack(m+1, n) < h(n+1, m+1, n)$, which implies that $Ack(m, n) < g(m+1, n+1)$ for all $m, n \in \mathbf{N}$. Now, an AMA with complexity parameters e, f can be constructed which stops exactly after g steps.

References

[ABD] H. Abdulrab, Résolution d'équations sur les mots: étude et implémentation LISP de l'algorithme de Makanin, Ph.D. Thesis, Université de Rouen, 1987.

[APE] H. Abdulrab, J.P. Pecuchet, Solving word equations, to appear in the special issue of The Journal of Symbolic Computation.

[BVA] E.Bombieri, J.Vaaler, On Siegel's Lemma. Inventiones Mathematicae, 73 (1983), pp.11-32.

[BUL] V.K. Bulitko, Equations and Inequalities in a Free Group and a Free Semigroup, Tul. Gos. Ped. Inst. Ucen. Zap. Mat. Kafedr Vyp 2, Gemetr. i Algebra, pp.242-252 (Russian), 1970.

[JAF] J. Jaffar, Minimal and Complete Word Unification, Journal of ACM 37 (1990) pp.47-85.

[KOS]
A. Kościelski, Is Makanin's algorithm for groups primitive recursive?, unpublished manuscript, Fall 1989.

[KPA] A. Kościelski, L. Pacholski, On the index of periodicity of a minimal solution of a word equation, unpublished manuscript, Spring 1989.

[MA1] G.S. Makanin, The Problem of Solvability of Equations in a Free Semigroup (in Russian), Matematiceskii Sbornik, 103 (1977), pp. 147-236.

[MA2] G.S. Makanin, Equations in a Free Group (in Russian), Izvestia AN SSSR 46 (1982), pp.1199-1273.

[PEC] J.P. Pecuchet, Solutions Principales et Rang d'un Système d'équations avec Constantes dans le Monoïde Libre, Discrete Mathematics 48 (1984), pp.253-274.

[RAZ] A.A. Razborov, On Systems of Equations in a Free Group, Math. USSR Izviestiya 25 (1985), pp.115-162.

[VZGS] J. Von Zur Gathen, M. Sieveking, A bound on Solutions of Linear Integer Equations and Inequalities, Proc. AMS 72 (1978), pp.155-158.