Elementary equations over free groups

Contents

1 Elementary equations over free groups

1.1 General remarks

1.1.1 Equations and some classical algorithmic problems

Let G be a group and $X = \{x_1, \cdots, x_n\}$ be a set of variables.

An *equation* in variables $x_1, \cdots, x_n$ with coefficients in G is a formal expression of the form

$$g_1 {x_{i_1}}^{\varepsilon_1} g_2 {x_{i_2}}^{\varepsilon_2} \cdots {x_{i_k}}^{\varepsilon_k} g_n = 1 \tag{1}$$

where $g_j \in G,\ x_{i_j} \in X,\ \varepsilon_j \in \{1, -1\}$.

Notice, that we allow equations without coefficients (in which case $g_1 = \ldots = g_n = 1$). Such equations are also called *coefficient-free* equations. Usually, we write equations in the functional notation:

$$f(x_1, \cdots, x_n,\ g_1, \cdots, g_n) = 1$$

or simply as

$$f(x_1, \cdots,\ x_n) = 1,$$

omitting coefficients.

Let

$$f(x_1, \cdots, x_n) = 1$$

be an equation in variables $x_1, \cdots, x_n$ with coefficients in G and let H be a group containing G as a fixed subgroup (in this case we say that H is a *G-group*). A tuple $(h_1, \ldots, h_n)$ of elements from H is a *solution* of $f(x_1, \ldots, x_n) = 1$ in H if the substitution $x_i \to h_i$ turns the formal expression 1 into equality in H, i.e., $f(h_1, \ldots, h_n) = 1$ in H.

Similarly, one can consider arbitrary systems of equations in variables $x_1, \ldots, x_n$ with coefficients in G:

$$f_i(x_1, \cdots, x_n) = 1, \quad i \in I. \tag{2}$$

A tuple $(h_1, \ldots, h_n) \in H^n$ is a solution of the system 2 if it is a solution of every equation $f_i = 1, i \in I$.

Let $S(X) = 1$ be a system of equations with coefficients from G. By $V_H(S)$ we denote the solution set of $S(X) = 1$ in a G-group H:

$$V_H(S) = \{(h_1, \ldots, h_n) \in H^n \mid f_i(h_1, \ldots, h_n) = 1, i \in I\}.$$

If $V_H(S) \neq \emptyset$ then the system $S = 1$ is called *consistent* over H. Two systems $S(X) = 1$ and $T(X) = 1$ are termed *equivalent* over H if $V_H(S) = V_H(T)$.

There are several typical questions about a given system of equation $S(X) = 1$ with coefficients in G and a G-group H:

1) Is $S(X) = 1$ consistent over H?

2) What are all possible solutions of $S(X) = 1$ in H?

3) How one can find a solution (all solutions) of $S(X) = 1$ in H?

Now we state the questions above in precise algorithmic formulations. The first question is one of the most intriguing problems in group theory, it attracted a lot of attention over the years:

Diophantine Problem (DP) Let G be a group. Does there exist an algorithm which for any equation $f = 1$ (finite system of equations $S(X) = 1$) with coefficients in G determines whether $f = 1$ ($S(X) = 1$) has a solution in G or not.

If such an algorithm exists we call it a *decision algorithm* and say that the Diophantine problem is *decidable* over G.

Remark. *A decision algorithm for DP takes equations* $f(x_1, \dots, x_n,\ g_1, \dots, g_n) = 1$ *as its inputs. Hence this equations have to be described in a constructive way. By definition these equations are group words in* X *and* G*. So it would suffices to describe effectively elements from* G*. Usually, we assume that the group* G *comes equipped with a given finite (or countable) generating set* A*. In this event, every element from* G *can be represented by a word in* $A^{\pm 1}$*. From now on we will always assume (if not said otherwise) that elements of* G *are already given as words in* $A^{\pm 1}$*.*

The this question can be formulated also as

Search Diophantine Problem (SDP) Let G be a group. Does there exist an algorithm which for any equation $f = 1$ (finite system of equations $S(X) = 1$) with coefficients in G determines whether $f = 1$ ($S(X) = 1$) has a solution in G or not, and if it does finds a solution.

G. Makanin proved that the DP is decidable over free groups [**?**]. Namely, he showed that there exists a recursive function $g : \mathbb{N} \to \mathbb{N}$ such that if a system of equations $S(X) = 1$ with coefficients in a free group F is consistent then it has a solution in F of the length at most $g(|S|)$.

Observe that if G is finitely presented group and an equation over G has a solution in G then one can find a solution of this equation effectively (in theory).

Theorem 1 *Let* G *be a finitely presented group or a finitely generated group with decidable word problem. If an equation* $S = 1$ *over* G *has a solution in* G *then one can find a solution of* $S = 1$ *effectively.*

In general DP is very difficult. In fact, decidability of the DP in a group G implies the decidability of quite a few classical algorithmic problems in G. To explain this we need to generalize slightly the Diophantine problem over G. Let $\mathcal{C}$ be an arbitrary collection of equations (or finite systems of equations) over a group G. We say that DP is *decidable for* $\mathcal{C}$ *over* G, if there exists an algorithm which for any equation (system) $f = 1$ from $\mathcal{C}$ decides whether $f = 1$ has a solution in G or not.

The Word Problem in G (WP) Let $A = \{a_1, \dots, a_n\}$ be a finite generating set for G. The word problem is decidable in G if the DP is decidable for the set of equations

$$S = \{{a_{i_1}}^{\varepsilon_1} \cdots {a_{i_n}}^{\varepsilon_n} = 1 \mid a_{i_j} \in A, \varepsilon_j \in \{1, -1\}, n \in N\}.$$

The Conjugacy Problem in G (CP) The conjugacy problem is decidable in G if the DP is decidable for the collection of equations

$$S = \{x^{-1}gx = f \ \mid \ g, f \in G\}.$$

We mention here few more algorithmic problems in groups.

The Power Problem in G (PP) The power problem is decidable in G if the DP is decidable for the collection of equations

$$S = \{x^n = f \ \mid \ n \in \mathbf{N}, \ f \in G\}.$$

The Commutator Problem in G (ComP) The commutator problem is decidable in G if the DP is decidable for the collection of equations

$$S = \{[x, y] = f \ \mid \ f \in G\}.$$

1.1.2 Cancellation schemes

Let $F(A)$ be a free group on $A = \{a_1, \dots, a_n\}$.

Suppose $u, v \in F(A)$ and the word uv is reduced as written, i.e., there is no cancellation in uv. In this event we denote the product uv by $u \circ v$. Plainly,

$$uv = u \circ v \Longleftrightarrow |uv| = |u| + |v|.$$

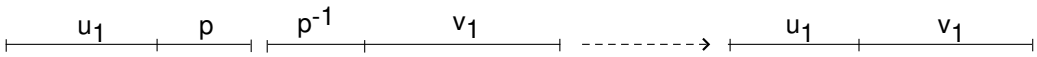
Let u, v be words in the alphabet $A^{\pm 1}$. Sometimes we write $u \equiv v$ to emphasize that the words u and v are equal as words (not only as elements of $F(A)$).

Suppose now that

$$u = u_1 \circ p, v = p^{-1} \circ v_1.$$

Then

$$uv = u_1(pp^{-1})v_1 = u_1v_1.$$



In this case we say that the subwords p and p^{-1} *cancel out* in forming the product uv. Graphically this can be expressed as

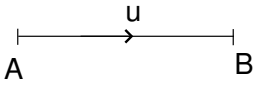
Notice, that we are not able to recover the subwords p, p^{-1} from the resulting word u_1v_1. To keep the history of cancellation we will sometimes write

$$u_1(pp^{-1})v_1$$

putting parentheses to mark the cancellation.

The graphical equivalent of this will be to "pinch" p and p^{-1} together and write the result in the following form:

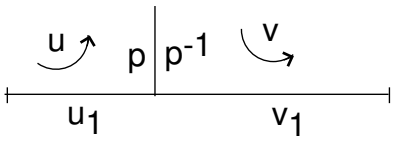
To simplify the notations we will use the following agreement. The diagram



means that the edge from A to B has the word u as a label, and the inverse edge from B to A has a label u^{-1} (i.e., we read u going from A to B). Hence the cancellation above can be expressed as

Now, let $w_1, \ldots, w_n$ be reduced words in $A^{\pm 1}$. Then, in general, we have several different ways to reduce the word

$$w \equiv w_1 w_2 \cdots w_n$$



into its reduced form. Each such reduction can be completely described by consequently putting parentheses to mark each step of the reduction (cancellation). For example, let

$$w_1 \equiv y_1y_2y_3y_1, \quad w_2 \equiv y_1^{-1}y_3^{-1}y_2^{-1}y_4, \quad w_3 \equiv y_4^{-1}y_2y_1.$$

Then

$$w = \underbrace{y_1y_2y_3y_1}_{w_1}\underbrace{y_1^{-1}y_3^{-1}y_2^{-1}y_4}_{w_2}\underbrace{y_4^{-1}y_2y_1}_{w_3}$$

and we have the following two different reduction processes of w:

$$w = \underbrace{y_1y_2(y_3y_1}_{w_1}\underbrace{y_1^{-1}y_3^{-1})(y_2^{-1}y_4}_{w_2}\underbrace{y_4^{-1}y_2)y_1}_{w_3} = y_1y_2y_1;$$

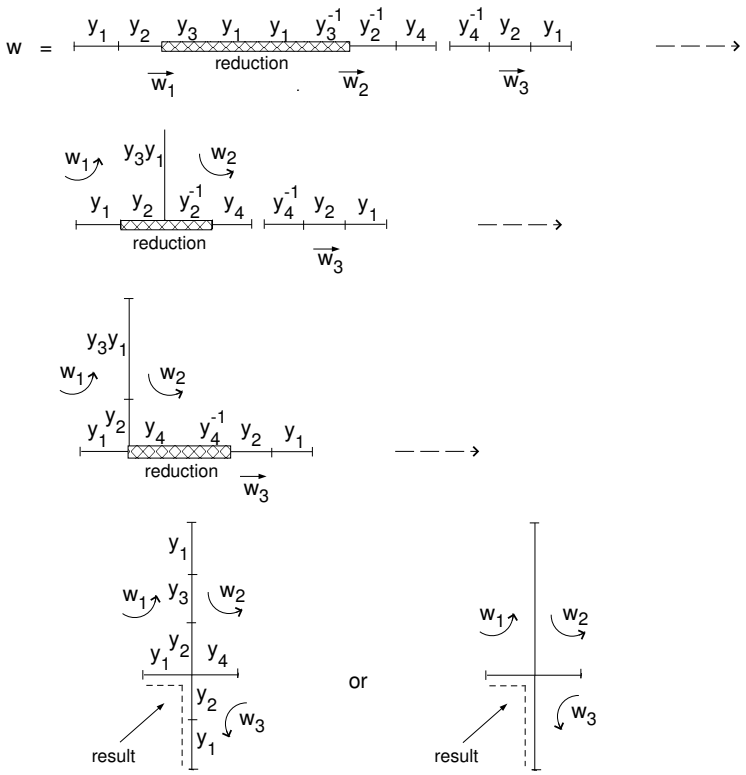
$$w = \underbrace{y_1(y_2(y_3y_1}_{w_1}\underbrace{y_1^{-1}y_3^{-1})y_2^{-1})(y_4}_{w_2}\underbrace{y_4^{-1})y_2y_1}_{w_3} = y_1y_2y_1.$$

Graphically, we can express these reductions as the following labeled trees:

Such words with parentheses or such labelled trees are called *cancellation schemes* for the word $w \equiv w_1 \cdots w_n$.

Lemma 1 *For any positive integer n there are finitely many trees $T_1, \ldots, T_{\mu(u)}$ such that for any n-tuple of words $w_1, \ldots, w_n \in F(A)$ each cancellation scheme for $w \equiv w_1 \ldots w_n$ can be obtained by labelling edges of one of the trees $T_1, \ldots, T_{\mu(u)}$ by elements from $F(A)$.*

Proof. It suffices to notice that each cancellation tree for $w \equiv w_1 \ldots w_n$ has no more than $2n$ vertices.



1.2 Equation $x^{-1}fx = g$

1.2.1 Cyclically reduced words and conjugacy problem

Let $F(A)$ be a free group with basis $A = \{a_1, \ldots, a_n\}$.

Definition 1 *A reduced word $w = y_1 \cdots y_m,\ (y_i \in A^{\pm 1})$ is called cyclically reduced if $y_1 \neq y_m^{-1}$.*

For example, the word ${a_1}^{-1}a_2a_2$ is cyclically reduced, but the word ${a_1}^{-1}a_2a_1$ is not.

Lemma 2 *Let w be a reduced word from $F(A)$. Then there exists a cyclically reduced word $\overline{w}$ and a reduced word w_1 such that*

$$w = {w_1}^{-1} \circ \overline{w} \circ w_1. \tag{3}$$

Moreover, such words $\overline{w}$ and w_1 are unique.

Proof. We will read off $\overline{w}$ and w_1 from w using the following marking process. Let $w = y_1 \cdots y_m,\ (y_i \in A^{\pm 1})$ be a reduced word. Compare the initial literal

y_1 with the final literal y_m. If $y_1 = {y_m}^{-1}$ then mark them both, and repeat the process for the unmarked subword $y_2 \cdots y_{m-1}$. In the case $y_1 \neq {y_m}^{-1}$ we stop. Clearly, we will finish in $\leq |w|$ steps, and the unmarked subword is cyclically reduced, we denote it by $\overline{w}$. Meanwhile, the longest marked subword at the end of w is w_1. From the discussion above it follows that

$$w = {w_1}^{-1} \circ \overline{w} \circ w_1$$

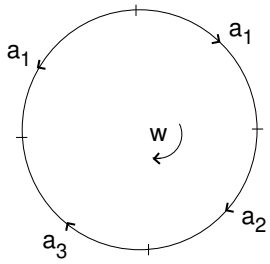
and the words $\overline{w}, w_1$ are uniquely determined by w. This proves the lemma. □

Definition 2 *Let w be an element of $F(A)$. Then the above decomposition*

$$w = {w_1}^{-1} \circ \overline{w} \circ w_1$$

is called the cyclic decomposition of w, and the word $\overline{w}$ is called the cyclically reduced form of w.

The cyclically reduced form of w has a simple visual interpretation. Every word $w \in F(A)$ defines a *cyclic word* w, which is the word w written on a circle in the clockwise orientation. For example, $w = a_1 a_2 a_3 {a_1}^{-1}$ defines the cyclic word

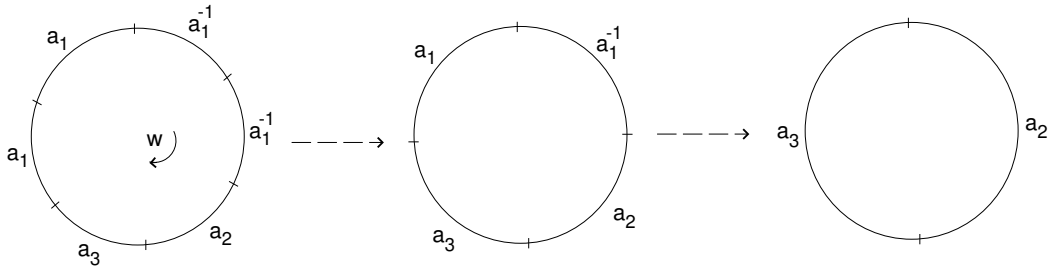


A *reduced cyclic word* is a cyclic word without neighbors of the type yy^{-1} ($y \in A^{\pm 1}$).

Every cyclic word can be transformed into a unique reduced cyclic word by means of elementary reductions exactly in the same way as for the ordinary words (see Section 1.2). Notice, that if w is a reduced word, then the reduction process for the cyclic word w is uniquely determined at each step, i.e., at each step there is only one possible elementary reduction to perform. For example, if

$$w = {a_1}^{-2} a_2 a_3 {a_1}^2$$

then the reduction process for w can be described as follows.



$$w = a_1^{-2}a_2a_3a_1^2 \;\rightarrow\; a_1^{-1}a_2a_3a_1 \;\rightarrow\; a_2a_3$$

or graphically:

This reduction procedure reflects the marking process from Lemma 2 (only instead of marking literals we delete them from the cyclic word). Therefore, the reduction of a cyclic word w results in the cyclic word $\overline{w}$. Hence, the cyclically reduced form of a word w is equal (as a cyclic word) to the reduction of a cyclic word w.

Starting with a word w we can form the corresponding cyclic word defined by w and then cut it back into a (linear) word. The resulting word depends on the cutting point and might be different from the initial word w. To describe all words that can occur here we need the following definition.

Definition 3 *Let $w = y_1 \cdots y_m$ $(y_i \in A^{\pm 1})$ be a word in $A^{\pm 1}$. A cyclic permutation of w is a word of the form $y_{i+1} \cdots y_m y_1 \cdots y_i$ $(1 \leq i \leq n)$.*

Clearly, two cyclic words defined by $u, v \in F(A)$ are equal if and only if u is a cyclic permutation of v (and hence, v is a cyclic permutation of u). Observe also, if $v = y_{i+1} \ldots y_m y_1 \ldots y_i$ is a cyclic permutation of $w = y_1 \cdots y_m$ then

$$(y_1 \cdots y_i)^{-1}\ w\ (y_1 \ldots y_i) = v.$$

So any cyclic permutation of w is a conjugate of w. Summarizing the discussion above we have the following

Proposition 1 *Let u, v be elements of $F(A)$. The following conditions are equivalent:*

1) u and v are conjugate in $F(A)$;

2) The cyclic words defined by u and v are equal;

3) $\overline{u}$ *is a cyclic permutation of* $\overline{v}$*.*

Corollary 1 *The conjugacy problem in* $F(A)$ *is decidable.*

Proof. To check whether given two elements $u, v \in F(A)$ are conjugate in $F(A)$ or not it suffices to find their cyclically reduced forms $\overline{u}$ and $\overline{v}$, and check whether the cyclic words defined by $\overline{u}$ and $\overline{v}$ are equal or not. This can be done effectively.

□

Exercise 1 *What is the time complexity of the decision algorithm for the conjugacy problem in* $F(A)$ *indicated above?*

1.2.2 Solutions of the conjugacy equation $x^{-1}ux = v$

We have already seen in the previous section how one can effectively verify whether or not the equation

$$x^{-1}ux = v \tag{4}$$

has a solution in $F(A)$.

A similar method allows one to find a particular solution (if it exists) of this equation.

Indeed, suppose $x^{-1}ux = v$ for some $x \in F(A)$. Let

$$u = {u_1}^{-1} \circ \overline{u} \circ u_1, \quad v = {v_1}^{-1} \circ \overline{v} \circ v_1$$

be the cyclic decompositions of u and v. Then

$$x^{-1}{u_1}^{-1}\overline{u}u_1x = v^{-1}\overline{v}v_1$$

and

$$v_1x^{-1}{u_1}^{-1}\overline{u}u_1x{v_1}^{-1} = \overline{v}.$$

Hence $\overline{v}$ must be a cyclic permutation of $\overline{u}$ (otherwise, (4) has no solutions in $F(A)$). It follows that some initial segment of $\overline{u}$ (viewed as a word in generators $A^{\pm 1}$), say u_2, conjugates $\overline{u}$ into $\overline{v}$. Hence

$$v = v_1^{-1}u_2^{-1}\overline{u}u_2v_1$$

and the element $x = u_1^{-1}u_2v_1$ is a solution of the equation (4) (straightforward verification).

Clearly, there exists an effective algorithm to find a solution of the type $x = u_1^{-1}u_2v_1$ from above (check one by one all initial segments u_2 of the word $\overline{u}$).

Corollary 2 *The conjugacy search problem is decidable in $F(A)$.*

Exercise 2 *What is the time complexity of the algorithm for solving the search conjugacy problem in $F(A)$ above?*

To describe all solutions of the equation (4) we rewrite it in the following form

$$[u, x] = g, \quad (here\ g = u^{-1}v). \tag{5}$$

This is a particular type of a *commutator equation* which we discuss in Section 1.4. Notice that if $x = b$ and $x = c$ are solutions of the equation 5 then $[u, b] = [u, c]$ which implies that $[u, cb^{-1}] = 1$. Hence, if b is a particular solution of (5), then all other solutions of (5) can be described as

$$x = yb,$$

where y is a solution of the homogeneous equation

$$[u, y] = 1. \tag{6}$$

We will see in Section 1.4.1 how to describe all solutions of this homogeneous equation.

Note. *Observe, that the description above reminds remotely the description of solution sets of liner systems of equations over an arbitrary ring or over an abelian group. Unfortunately, for equations over non-abelian groups, as we will see later, this situation is rather an exceptional one.*

1.3 Equation $x^n = g$

For an element $g \in F(A)$ consider the *power equation*

$$x^n = g \quad (n \in \mathbf{N}) \tag{7}$$

Lemma 3 *The power equation $x^n = g$ has at most one solution in $F(A)$.*

Proof. Suppose $u^n = g$ for some $u \in F(A)$. Let

$$u = {u_1}^{-1} \circ \bar{u} \circ u$$

be the cyclic decomposition of u. Then

$$u^n = {u_1}^{-1} \circ \bar{u}^n \circ u_1$$

is the cyclic decomposition of

$$g = u^n.$$

Hence

$$u_1 = g_1, \bar{u}^n = \bar{g}.$$

It follows that (7) has a solution in $F(A)$ if and only if, firstly, n divides $|\bar{g}|$ and , secondly, if

$$\bar{g} = f_1 \circ f_2 \circ \cdots \circ f_n$$

is the partition of $\bar{g}$ into n subwords of equal length, then

$$f_1 = f_2 \cdots = f_n$$

In this event

$$x = g^{-1}{}_1 f_1 g_1$$

is the only solution of (7) in $F(A)$.

Notice, that the argument above shows that for $n \geq |g|$ the equation $x^n = g$ has no solution in $F(A)$. Another remark is that if the power equation has a solution in $F(A)$, one can effectively find this solution in $O(|g|)$ steps.

Definition 4 *The Power Problem is decidable in a group G if there exists an algorithm which determines whether a power equation $x^n = g (g \in G, n \in N)$ has a solution in G or not.*

Corollary 3 *The Power Problem is decidable in a free group.*

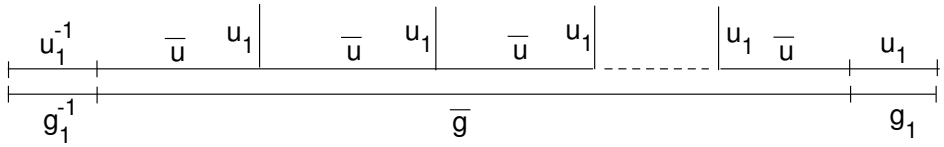
Notice, that there are finitely presented groups in which the power problem is undecidable.

The uniqueness of solutions of power equations in $F(A)$ allows one to define n-th root of elements of $F(A)$.

Definition 5 *The solution of the power equation $x^n = g$ is called the n-th root of g in $F(A)$.*

The *maximal root* of g is the n-th root of g with n maximal possible. Observe that the maximal root of a non-trivial element on $g \in F(A)$ exists and it is unique.

The following picture (a cancellation scheme) explains geometrically how to find a solution of a power equation $x^n = g$. If $x = u$ is a solution of this equation then we have the following cancellation scheme:



1.4 Equation $[x, y] = g$

1.4.1 Homogenous equation $[x, y] = 1$

We start with the classical approach using induction on the total length of solutions. Then we demonstrate how to solve the equation using cancellation schemes.

Lemma 4 *Let $u, v \in F(A)$. If $uv = vu$ then there exists $w \in F(A)$ and integers m, n such that*

$$u = w^m, v = w^n$$

Proof. Induction on $|u| + |v|$.

If $|u| + |v| = 1$, then $u = 1$ or $v = 1$ and in this case the conclusion of the lemma is obvious. Suppose now that $|u|, |v| \geq 1$ and $|u| \leq |v|$.

Case 1. Assume that there is no cancellation in uv, i.e., $|uv| = |u| + |v|$. Then $|vu| = |uv| = |u| + |v|$ and there is no cancellation in vu. Hence vu and uv are reduced. Now from $uv = vu$ and $|u| \leq |v|$ we deduce that $v = u \circ v_1$. Therefore $uv_1 = v_1 u$ and $|v_1| < |v|$. By induction, u and v_1 (hence u and v) are powers of some $w \in F(A)$, as desired.

Case 2. Suppose u cancels completely in v. Then $v = u^{-1} \circ v_1$ and the equality $uv = vu$ takes the form $uu^{-1}v_1 = u^{-1}v_1 u$. The latter can be written as

$$uv_1 = v_1 u.$$

Since $|v_1| < |v|$ we can proceed by induction.

Case 3. There is a cancellation in uv, but u does not cancel completely in uv. In this event

$$u = u_1 \circ y, v = y^{-1} \circ v_1, \quad y \in A^{\pm 1}.$$

Then $uv = vu$ can be rewritten as

$$u_1 v_1 = y^{-1} v_1 u_1 y \tag{8}$$

Since $|v| \geq |u|$ and v does not cancel completely in uv, it follows that vu begins with y^{-1}. Similarly, u_1 does not cancel completely in $u_1 v_1$ (otherwise

u cancels out in uv), therefore u_1 begins with y^{-1}. Hence $u_1 = y^{-1} \circ u_2 \circ y, v_1 = y^{-1} \circ v_1$ and it follows from (8) that

$$u_2 v_1 = v_1 y^{-1} u_2 y \tag{9}$$

Notice (comparing lengths) that cancellation must occur in the right-hand side of (9). There is only one possible place where the cancellation may occur, namely in $v_1 y^{-1}$, so $v_1 = v_2 \circ y$. In this event, $u = y^{-1} \circ u_2 \circ y, v = y^{-1} \circ v_2 \circ y$, and equality $uv = vu$ takes the form

$$u_2 v_2 = v_2 u_2.$$

Now the conclusion of the lemma follows by induction.

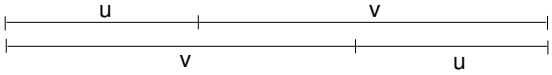
□

Corollary 4 *The set* $\{(w^n, w^m) | w \in F(A), n, m \in \mathbf{Z}\}$ *is the solution set of the equation* $[x, y] = 1$.

Now we illustrate the proof above by cancellation schemes.

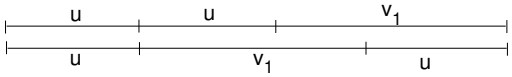
Let $u, v \in F(A)$ and $[u, v] = 1$. Consider all possible cancellation schemes in the product $uvu^{-1}v^{-1}$.

Case 1. No cancellation in uv. Then the corresponding scheme is the following:



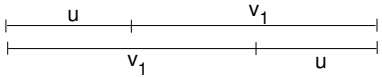
(we assume here that $|u| \leq |v|$).

Clearly, the point A divides v into two parts u and v_1 ($v = u \circ v_1$). Draw u on both occurrences of v in the scheme:



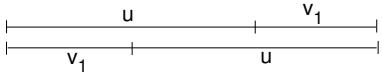
Cut out the common initial segment u from the scheme, the resulting scheme will be of the following type:

It is exactly the initial scheme provided $|u| \leq |v_1|$. We can repeat the process until this is possible and in finitely many steps (say, m_1 steps) we will have:



$$\begin{cases} v = u^{m_1} v_1, \ |u| > |v_1| \\ u = u \end{cases}$$

The resulting scheme will be like this:

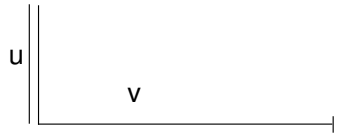


Now we see that $u = v_1 \circ u_1$ and we can again cut out the common initial segment v_1 from the scheme and proceed to do so until this is possible, i.e., until $|u_1| < |v_1|$. In finitely many steps, say K_1 steps, we will have

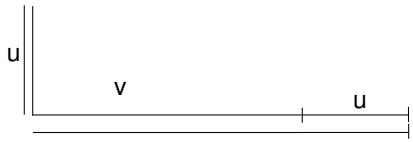
$$\begin{cases} v = u^{m_1} \circ v_1, \ |u_1| \leq |v_1| \\ u = v_1^{K_1} \circ u_1 \end{cases}$$

Now again we have the initial cancellation scheme and we can repeat the process again. Since, we cannot decrease the length of $|u|$ or $|v|$ forever, the process will stop eventually. At this point we will have $u_1 = 1$ or $v_1 = 1$. In any event, u and v will be powers of either u_1 or v_1.

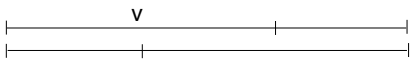
Case 2. u cancels out in uv.



Then there exists only one possible scheme:

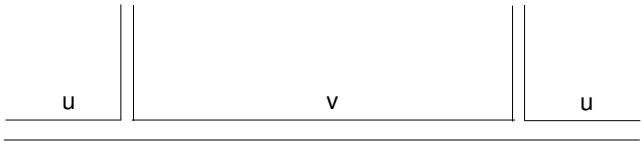


If we draw it in the following way

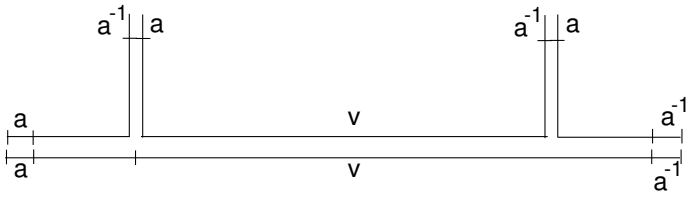


then we can see that the process from Case 1 will work in this case also.

Case 3. There is a cancellation in uv. Then there is a cancellation in vu and the corresponding scheme is as follows:



Let a be the first literal in u, i.e., $u = a \circ u_1$. Draw a in the scheme



Then we see that $u = a \circ u_2 \circ a^{-1}, v = a \circ v_2 \circ a^{-1}$.

Clearly, we can cut out the segment labelled by a everywhere in the scheme and the resulting scheme will be the cancellation scheme for $[u_2, v_2] = 1$. It is easy to see that in finitely many steps we will cut out either the whole initial segment up to the point A, or we will cut out both pinches. In either case we will get the scheme from case 1) or case 2).

□

Exercise 3 *Define a long commutator* $[x, u_1, \ldots, u_n]$ *inductively as follows:*

$$[x, u_1, \ldots, u_n] = [[x, u_1, \ldots, u_{n-1}], u_n].$$

For given $u_1, \ldots, u_n \in F(A)$ *solve the equation*

$$[x, u_1, \ldots, u_n] = 1.$$

1.4.2 Centralizers in $F(A)$. Commutative-transitive groups

Definition 6 *Let* G *be a group and let* M *be a subset of* G*. Then the set*

$$C_G(M) = \{g \in G | [g, m] = 1 \ \forall m \in M\}$$

is called the centralizer of M *in* G*.*

It is easy to see that $C_G(M)$ is always a subgroup of G.

Proposition 2 *Let* F *be a free group and* M *be a subset of* F*. If* $M \neq \{1\}$*, then the centralizer* $C_F(M)$ *is cyclic.*

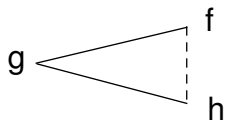
Proof. If $g \in M$, then $C_F(M) \subseteq C_F(g)$. Therefore to prove the proposition it suffices to prove that the centralizer $C_F(g)$ of any non-trivial element $g \in F$ is cyclic. Let q_0 be the maximal root of g in F, then $g = {q_0}^k$ for some positive k. We claim that for any $f \in F$ if $[{q_0}^n, f] = 1$, then $[q_0, f] = 1$. Indeed, if ${q_0}^n f = f {q_0}^n$, then

$${q_0}^k = f {q_0}^k f^{-1} = (f q_0 f^{-1})^k$$

since k-roots of elements in F are unique (see Section 1.3), then $q_0 = fq_0f^{-1}$, i.e., $[q_0, f] = 1$. It follows that $C_F(g) = C_F(q_0)$. Now if $f \in C_F(q_0)$, then by Lemma 4 q_0 and f are powers of some element $w \in F$. Since q_0 is not a proper power, then $q_0 = w^{\pm 1}$ and f is a power of q_0. Consequently, the maximal root q_0 is a generator of the centralizer $C_F(g)$.

Definition 7 *A group G is called* commutative-transitive *if the commutation is an equivalence relation on $G - \{1\}$, i.e., if $[g, f] = 1$ and $[g, h] = 1$ then $[f, h] = 1$ provided f, g, h are non-trivial elements of G.*

The *commutation graph* of G consists of elements from $G - \{1\}$ as vertices of this graph and two vertices q, h are connected by an edge if and only if $[g, h] = 1$. So whenever we have



then there exists an edge between f and h. Now it is clear, that a group G is commutative-transitive if and only if in the commutation graph of G each connected component is a complete subgraph (i.e., any two vertices of the component are connected by an edge).

Now the following lemma is obvious.

Lemma 5 *A group G is commutative-transitive if all proper centralizers of G are commutative.*

Corollary 5 *Every free group is commutative-transitive.*

Notice, that commutative transitive groups have very simple structure of centralizers: any two proper centralizers in such a group either are equal or intersect in identity.

Notice also, that many non-free groups are commutative-transitive. For example, proper centralizers in a torsion-free hyperbolic group are cyclic, hence each such group is commutative-transitive.

Exercise 4 *Let G be a commutative-transitive group. Prove that if $x, y \in G$ and $x^n = y^m$ then $[x, y] = 1$.*

1.4.3 CSA groups

A subgroup H of a groups G is called *malnormal* if for any $g \in G$ the following condition holds:

$$H^g \cap H \neq 1 \Longrightarrow g \in H.$$

A group G is called a *CSA group* if every maximal abelian subgroup of G is malnormal. Observe, that every CSA group is commutative-transitive.

Theorem 2 *Let G be a torsion-free group such that every proper centralizer of G is cyclic. Then G is a CSA group.*

Corollary 6 *Every free groups is CSA.*

Exercise 5 *Give a direct proof that a free group is CSA.*

1.4.4 Wicks' theorem

Let G be a group and g be an element of G. The equation

$$[x, y] = g$$

has a solution in G if and only if g is a commutator in G. It follows that the Diophantine problem is decidable for the class of commutator equations

$$\{[x, y] = g | g \in G\}$$

if and only if there exists an effective procedure for recognizing commutators in G.

The following result shows that there exists an algorithm for recognizing commutators in free groups.

Theorem 3 *(Wicks'). Let F be a free group, g an element in F, and $\bar{g}$ the cyclically reduced form of g. Then g is a commutator in F if and only if some cyclic permutation of $\bar{g}$ is of the form*

$$a \circ b \circ c \circ a^{-1} \circ b^{-1} \circ c^{-1} \tag{10}$$

for some $a, b, c \in F$.

$\triangleright$. Notice, that since g and $\bar{g}$ are conjugate then if one of them is a commutator, then the other one also does.

Obviously, every commutator $u^{-1}v^{-1}uv$ can be presented as a product $abca^{-1}b^{-1}c^{-1}$ (say $a = u^{-1}, b = v^{-1}, c = 1$), but there may be some cancellation in there. The point is to find such a presentation without cancellation.

Observe, that for any $a, b, c \in F$ one has

$$abca^{-1}b^{-1}c^{-1} = (ab)(ca^{-1})(ab)^{-1}(ca^{-1})^{-1}$$

hence $abca^{-1}b^{-1}c^{-1}$ is a commutator.

Now suppose g is a cyclically reduced commutator. As we have mentioned above any cyclic permutation of g can be presented in the form $abca^{-1}b^{-1}c^{-1}$. Among all such presentations chose one with the minimal total length $|a| + |b| + |c|$. We claim, that in this event there is no cancellation in the product $abca^{-1}b^{-1}c^{-1}$. To prove this it suffices to consider all possible cases where cancellation can occur.

Suppose $b = b_1 \circ y, c = y^{-1} \circ c_1$.

Then

$$abca^{-1}b^{-1}c^{-1} = ab_1c_1a^{-1}y^{-1}{b_1}^{-1}{c_1}^{-1}y$$

Conjugating by a, we obtain the following cyclic permutation of g:

$$b_1c_1a^{-1}y^{-1}{b_1}^{-1}{c_1}^{-1}ya = b_1c_1(a^{-1}y^{-1})b_1^{-1}c_1^{-1}(a^{-1}y^{-1})^{-1}$$

with a shorter total length of representatives $|b_1|+|c_1|+|a|+1 < |a|+|b|+|c|$-contradiction. Similar argument shows that cancellation does not occur in all other places, which proves the theorem.

□

Corollary 7 *Commutators are effectively recognizable in free groups.*

Indeed, let g be an element of a free group F. We can effectively find the cyclically form $\bar{g}$ of g. If $|\bar{g}|$ is of odd length, then $\bar{g}$ (as well as g) is not a commutator. If $|\bar{g}|$ is of even length, then we can take the left half of $\bar{g}$ and partition in into all possible products of the type $a \circ b \circ c$.

Now for every such partition abc we check whether the right half of $\bar{g}$ is equal to $a^{-1}b^{-1}c^{-1}$ or not. If yes, then g is a commutator, if not, then we take one by one all cyclic permutations of $\bar{g}$ and repeat the process for each of them.

Exercise 6 *What is the complexity of the algorithm to recognize commutators in $F(A)$ described above?*

Exercise 7 *Let $u, v \in F(A)$ and $[u, v] \neq 1$. Then equation*

$$[u, v] = x^n$$

has a solution if and only if $n = \pm 1$.

1.4.5 Elementary transformations and minimal solutions

The following lemma shows how one can obtain infinitely many solutions of the equation $[x, y] = g$ from a given particular solution $x = u, y = v$.

Lemma 6 *Let (u, v) be a solution of an equation*

$$[x, y] = g.$$

Then pairs

$$(v^m u, v) \text{ and } (u, u^m v), m \in \mathbf{Z}$$

are also solutions of this equation.

$\triangleright$. This follows immediately from the following identities:

$$[v^m u, v] = [v^m, v]^u [u, v] = [u, v] = g;$$

$$[u, u^m v] = [u, v][u, u^m]^v = [u, v] = g.$$

So starting from a solution (u, v) we can generate infinitely many other solutions:

$$(u, v) \to (v^{m_1} u, v) \to (v^{m_1} u, (v^{m_1} u)^{m_2} v) \to (((v^{m_1} u)^{m_2} v)^{m_3}, (v^{m_1} u)^{m_2} v) \to \dots$$

Now we introduce the following elementary transformations on pairs of elements from F:

$$U : (u, v) \to (vu, v)$$

$$V : (u, v) \to (u, uv).$$

Clearly, this transformation U, V has inverses

$$U^{-1} : (u, v) \to (v^{-1} u, v)$$

$$V^{-1} : (u, v) \to (u, u^{-1} v).$$

By U^m, V^m we denote the product of the m consequent transformation of the type U, V correspondingly.

We will say that two pairs (u_1, v_1) and (u_2, v_2) are equivalent ($(u_1, v_1) \sim (u_2, v_2)$) if one can transform (u_1, v_1) into (u_2, v_2) by a finite chain of elementary transformations. Clearly, $\sim$ is an equivalence relation on pairs

By Lemma 6 if (u, v) is a solution of $[x, y] = g$, then the equivalence class $[(u, v)]$ provides infinitely many other solutions of this equation. It follows that the solution set $V([x, y] = g)$ of the equation is a union of equivalence classes:

$$V([x, y] = g) = \cup_{i \in I} [(u_1, v_i)].$$

Our goal now is to prove that one can effectively find a finite collection of solutions $(u_1, v_1), \ldots, (u_n, v_n)$ such that

$$V([x, y] = g) = [(u_1, v_1)] \cup \cdots \cup [(u_n, v_n)].$$

To this end we need the following

Definition 8 *A solution (u, v) of the equation $[x, y] = g$ is called minimal if it has the minimal total length $|u| + |v|$ among all solutions in its equivalence class $[(u, v)]$.*

Lemma 7 *Let (u, v) be a minimal solution of the equation $[x, y] = g$. Then in each of the products*

$$uv, v^{-1}u, u^{-1}v^{-1}$$

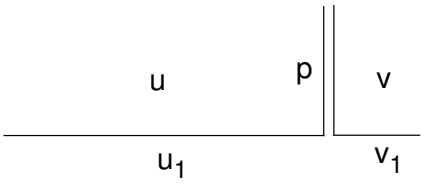
the total cancellation can not be greater than $\min\left\{\frac{|u|}{2}, \frac{|v|}{2}\right\}$

Proof. Observe, that for arbitrary $u, v \in F(A)$ the following conditions are equivalent:

1) more than half of v cancels out in uv;

2) $|uv| < |u|$

This can be seen from the corresponding cancellation scheme:



Indeed,

$$|uv| = |u_1| + |v_1| < |u| \Leftrightarrow |v_1| < |p|.$$

Now, let (u, v) be a minimal pair.

Case 1. Consider the product uv.

Suppose, that more than half of v cancels out in uv. Then $|uv| < |u|$. In this case

$$(u, v) \sim (u, uv) \sim (v^{-1}, uv)$$

and $|v^{-1}| + |uv| = |v| + |uv| < |v| + |u|$ - contradiction with the minimality of (u, v).

Suppose, that more then half of u cancels out in uv. Then $|uv| < |v|$ and the equivalence

$$(u, v) \sim (u, uv)$$

shows that (u, v) is not minimal ($|u| + |uv| < |u| + |v|$).

Case 2. Suppose more than half of u cancels out in the product $v^{-1}u$. Then $|v^{-1}| < |v^{-1}| = |v|$. In this case

$$(u, v) \sim (u, uv)$$

and $|u| + |uv| < |u| + |v|$ - contradiction.

Suppose that more than half of v cancels out in $v^{-1}u$. Then

$$|v^{-1}u| = |u|.$$

Hence,

$$(u, v) \sim (v^{-1}u, v)$$

and $|v^{-1}u| + |v| < |u| + |v|$ - contradiction.

Case 3. $u^{-1}v^{-1}$.

If more than half of u^{-1} cancels out in $u^{-1}v^{-1}$, then $|u^{-1}v^{-1}| = |vu| < |v|$. Hence

$$(u, v) \sim (vu, v) \sim (vu, u^{-1})$$

and $|vu| + |u^{-1}| < |v| + |u|$ contradiction. If more than half of v^{-1} cancels out in $u^{-1}v^{-1}$ then $|u^{-1}v^{-1}| < |u^{-1}| = |u|$. Now

$$(u, v) \sim (vu, v)$$

and $|vu| + |v| = |u^{-1}v^{-1}| + |v|$ - contradiction.

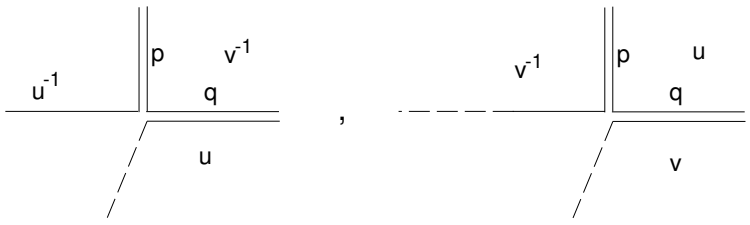
Actually, it suffices to consider just the first case, the two others are similar.

Corollary 8 *Let (u, v) be a minimal solution of the equation $[x, y] = g$ ($g \neq 1$). Then at least one literal from each of the factors u^{-1}, v^{-1}, u, v occurs in the reduced form of $u^{-1}v^{-1}uv$.*

$\triangleright$. By the lemma above at most half of each factor can cancel out in the products $u^{-1}v^{-1}, v^{-1}u, uv$. Hence if u has odd length, then the literal in the middle of u does not cancel out in the product $v^{-1}uv$. Similarly, if v^{-1} is of odd length, then it does not cancel completely in $u^{-1}v^{-1}u$.

Suppose now that u is of even length, $u = pq, |p| = |q|$. If u cancels out completely in $v^{-1}uv$, then $p = q^{-1}$ and $u = 1$. This implies $g = 1$-contradiction. Hence u does not cancel completely in $v^{-1}uv$, as well as v^{-1} in $u^{-1}v^{-1}u$.

It follows that reducing v^{-1} in $u^{-1}v^{-1}uv$ we do not have cancellation schemes of the types:



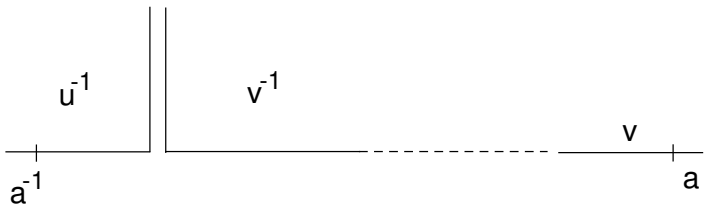
Hence, neither of the factors u^{-1}, v^{-1}, u, v cancels out completely in $u^{-1}v^{-1}uv$.

Lemma 8 *If (u, v) is a minimal solution of the equation $[x, y] = g(g \neq 1)$ then $|u| + |v| \leq |g|$.*

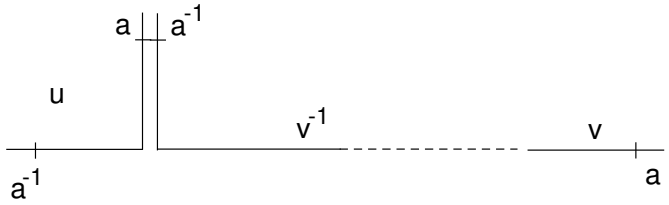
$\triangleright$. Suppose g is not cyclically reduced, say $g = a^{-1} \circ g_1 \circ a$. If we do not have pinches in a cancellation scheme for

$$u^{-1}v^{-1}uv,$$

then $|u^{-1}| + |v^{-1}| + |u| + |v| = |g|$ and the conclusion of the lemma holds. Suppose we have at least one pinch in the cancellation scheme for $u^{-1}v^{-1}uv$; say the scheme is of the type:



then we can see that we have:



hence $u = a^{-1} \circ u_1 \circ a, v = v_1 \circ a, g = {g_1}^a$.

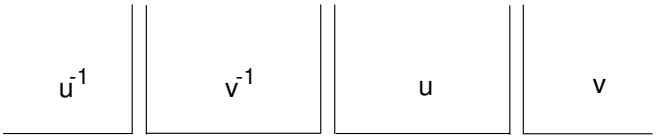
Now, $[u_1, av_1] = g_1$ and (u_1, av_1) is a minimal solution of the equation $[x, y] = g_1$. By induction on $|g|$ we have

$$|u_1| + |av_1| \leq |g_1|$$

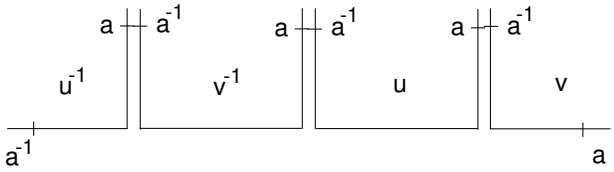
hence

$$|u| + |v| = |a^{-1} \circ u_1 \circ a| + |v_1 \circ a| \leq |a^{-1} \circ g_1 \circ a| = |g|.$$

Suppose now, that g is cyclically reduced. Then there is no cancellation scheme with all three non-trivial pinches:

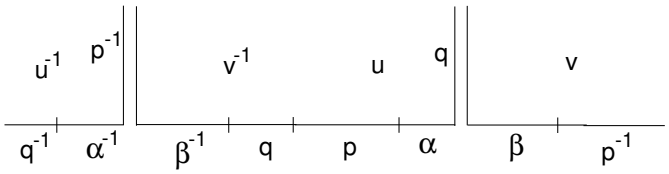


Indeed, in this event we have



So $u = a^{-1} \circ u_1 \circ a, v = a^{-1} \circ v_1 \circ a$, and hence g is not cyclically reduced-contradiction. It implies, that we have to consider the schemes of the following type (with at least one trivial pinch):

Assume, $|p| \geq |q|$. Then we have the following scheme:



where $u = p\alpha q, v = q^{-1}\beta p^{-1}$. Hence

$$q^{-1}\alpha^{-1}\beta^{-1}qp\alpha\beta p^{-1} = g \tag{11}$$

Hence

$$|u| + |v| = |p| + |\alpha| + |q| + |q| + |\beta| + |p| \leq |g|.$$

Remark 1 *Notice, that from (11) we have*

$$q^{-1}\alpha^{-1}\beta^{-1}qp\alpha\beta p^{-1} = g.$$

Hence the cyclic permutation of g:

$$\alpha^{-1}\beta^{-1}(qp)\alpha\beta(p^{-1}q^{-1})$$

is the Wicks' form of the type

$$a \circ b \circ c \circ a^{-1} \circ b^{-1} \circ c^{-1}.$$

It shows that Wicks' forms for the cyclic word g come from the minimal solutions of $[x, y] = g$.

Corollary 9 *Let $g \in F(A)$ and $g \neq 1$. Then there are only finitely many minimal solutions of the commutator equation*

$$[x, y] = g.$$

If $(u_1, v_1), \ldots, (u_n, v_n)$ are minimal solutions of the equation above, then the solution set of this equation is a finite union of the equivalence classes:

$$[(u_1, v_1)] \cup \cdots \cup [(u_n, v_n)].$$

Moreover, one can effectively find the set of minimal solutions $(u_1, v_1), \cdots, (u_n, v_n)$.

Corollary 10 *Let $F = F(a, b, \ldots)$. Then the solution set of the equation $[x, y] = [a, b]$ is equal to the equivalence class $[(a, b)]$, and (a, b) is the only minimal solution of this equation.*

$\triangleright$. Indeed, if (u, v) is a minimal solution of $[x, y] = [a, b]$, then $|u|+|v| \leq 4$. If $|u| = 1$, then $|v| = 1$ (since in this case u does not cancel in $[u, v]$). In this event $u = 1, v = b$. If $|u| = 2 = |v|$, then

$$u^{-1} = a^{-1}p^{-1}, v^{-1} = pc$$

(since $[a, b]$ begins with a^{-1}) , and also $v = c^{-1}p^{-1}$ has to end on b. It follows $p = b^{-1}$, and substituting into $[u, v]$ we see that $[u, v] \neq [a, b]$. So the case $|u| = |v| = 2$ is not possible.

$\square$

We described completely the solution set of the equation $[x, y] = g$.

Digression

It turns out that the methods we discussed here can be generalized into much more general situations. In fact, they are some of the most powerful methods in the theory of free groups. The transformations U, V are particular types of so-called Nielsen elementary transformations, the idea of minimal pairs gives rise to the notation of N-reduced sets, Wicks' forms exists for an arbitrary quadratic equation over free groups and the idea of estimation of the length of a minimal solution of an equation in terms of length of its coefficients is the principal one in Makanin's solution of the Diophantine problem over free groups