ONE-VARIABLE EQUATIONS IN FREE GROUPS¹

K. I. APPEL

1. Introduction and notation. Let G be the free group on free generators $a_1, a_2, \dots, a_r$. Let $w(x)$ be any freely reduced word in $G * \{x\}$ other than the empty word. Define $U(w)$ to be the set of elements $u = u(a_1, \dots, a_r)$ of G which satisfy $w(u) = 1$. Let $\gamma_1, \gamma_2, \dots$ be a set of parameters (variables for integers). An expression $W = w_1^{\alpha_1} w_2^{\alpha_2} \dots w_m^{\alpha_m}$ is called a parametric expression over G if the w_i are words on $a_1 \dots a_r$ and the α_j are polynomials in the parameters γ_j . A value of W is the word on $a_1, \dots, a_r$ resulting from substituting integer values for the γ_i appearing in W . Lyndon [3] effectively associates with $w(x)$ a finite set of parametric expressions over G such that $U(w)$ is precisely the set of all values of these parametric expressions. Lyndon's argument does not bound the number of parameters required. Lorenc [1] showed that at most two parameters were required in each parametric expression. We prove the following theorem.

THEOREM. *$U(w)$ is the set of values obtained by substitution of positive integers into a finite set of parametric expressions which are effectively obtainable from $w(x)$ and contain at most one parameter in each expression.*

Much of the notation and many of the arguments in this paper may be found either explicitly or implicitly in [3].

If w is a cyclically reduced word on $a_1, \dots, a_r$ and w is not a proper power, (i.e., w is not a formal m -fold repetition of one of its subwords for $m > 1$), we call w *primitive*. We use $|w|$ to denote the length of w as a freely reduced word on the a_i . The following lemma is proved in [4].

LEMMA LS. *If v and w are primitive and have powers v^m and w^n , $m, n > 0$ with common initial segment of length $|v| + |w|$ then $v = w$.*

A consequence of Nielsen's proof of the subgroup theorem is the fact that if n elements generate a free group of rank n then they generate it freely. We use it in the following form:

LEMMA N. *If $W(x, y)$ is a word on generators x and y in $G * \{x\} * \{y\}$*

Received by the editors March 13, 1967.

¹ I should like to thank Professor Wilson Zaring for an extremely careful reading of a first draft of this paper and many suggestions. I should also like to thank the referee for several suggestions simplifying the exposition.

and $u(a_1, \dots, a_r); v(a_1, \dots, a_r)$ satisfy $W(u, v) = 1$ in G then u and v are powers of a common element.

The following notational conventions are observed. Variables on words range over freely reduced words in $a_1, \dots, a_r$. We write $uv = u \cdot v$ if uv is freely reduced, that is, if the final letter of u is not the inverse of the initial letter of v . Writing $uvw = u \cdot v \cdot w$ includes the requirement that $v = 1$ implies $uw = u \cdot w$.

Since $w(x)$ is conjugate to a word ending in a power of x except in the trivial case in which it is x -free and $w(x) = 1$ has no solutions, and since solutions are preserved under conjugation, we may assume $w(x) = c_1 x^{e_1} c_2 \cdot \dots \cdot c_t x^{e_t}$ where $e_i = \pm 1$ for $1 \leq i \leq t$ and $e_i = -e_{i-1}$ implies $c_i \neq 1$. The c_i are called the *coefficients* of $w(x)$, t is called the *degree* of $w(x)$ and M , twice the maximum of the lengths of the c_i , is called the *size* of $w(x)$.

The ordered triple of words u, v, w is called a *singular triple* if there exist u', w', p and q such that $u = u' \cdot p$, $w = q \cdot w'$ and $v = p^{-1} \cdot q^{-1}$. The following proposition, due to Nielsen, is used in a form found in [3].

PROPOSITION 1. *If $w_1 w_2 \cdot \dots \cdot w_n = 1$, $n \geq 1$, then at least one of the triples $(1, w_1, w_2)$, (w_1, w_2, w_3) , $\dots$, $(w_{n-1}, w_n, 1)$ is singular.*

The following notational convention is used. When a word s is being considered, s_{2i+1} and s_{2i+2} will be used for initial and terminal segments of s respectively, with $s = s_{2i+1} \cdot s_{2i+2}$ (e.g. $s = s_1 \cdot s_2 = s_3 \cdot s_4$ etc.).

2. Proof of the Theorem. First we prove a combinatorial lemma.

LEMMA 1. *Let $v = s^n p t^m$ with s, t primitive with both $|s^n|$ and $|t^m|$ greater than $|p| + |s| + |t|$. Let j and k be the largest integers such that $|s^j| < |s^n| - |p| - |s| - |t|$, $|t^k| < |t^m| - |p| - |s| - |t|$. Then either there exists some nonempty word w such that $v = s^j \cdot w \cdot t^k$ or for some g, h , we have $s = g \cdot h$, $t = g^{-1} \cdot h^{-1}$ and $v = (gh)^\alpha g$ for some integer α . If in particular $t = s^{-1}$ and $m = n$ then either $v = s^j \cdot w \cdot s^{-j}$ or $v = p = s^\beta$ for some integer β .*

PROOF. Let $s^n p = S \cdot q$, S an initial segment of s^n and q a terminal segment of p . Let $q t^m = r \cdot T$, r an initial segment of q , T a terminal segment of t^m . If $r \neq 1$ then $v = S \cdot r \cdot T$. If $r = 1$, first consider the case that the cancellation between S and T is less than $|s| + |t|$. In both these cases v contains the required initial s^j and terminal t^k . If the cancellation between S and T exceeds $|s| + |t|$, by Lemma LS we have $s = s_1 \cdot s_2$, $t = t_1 \cdot t_2$, $S = s_1 (s_2 s_1)^\alpha$, $T = (t_2 \cdot t_1)^\beta t_2$ with $s_2 s_1 = (t_2 t_1)^{-1}$. The shorter of S and T will cancel entirely into the longer, say, by symmetry, T into S where $S = s^\alpha g T^{-1}$ with $s = g \cdot h$. It follows that t

$= g^{-1} \cdot h^{-1}$ and $v = (gh)^{\alpha g}$. If $s^{-1} = t$ then the condition $s_2 s_1 = (t_2 t_1)^{-1}$ becomes $s_2 s_1 = (s_3^{-1} s_4^{-1})^{-1}$ and $s_2 s_1 = s_4 s_3$. But $s_1 = s_3$, for if not $s_2 s_1$, a cyclic permutation of a primitive word and hence a primitive word, is equal to a cyclic permutation of itself contradicting primitivity. But now $v = s^r s_1 s_1^{-1} s^{-q} = s^{r-q}$. Thus $p = s^{r-q}$.

LEMMA 2. *If $y, y, s \cdot y$ or $y, y, s \cdot y^{-1}$ is a singular triple with $|y| > |s|$ then $y = s^n s_3 \cdot s_1^{-1} s^{-m}$, $m, n \geq 0$ and s is cyclically reduced. The triples $y, s \cdot y, y$ and $y \cdot s, y \cdot s, y^{-1}$ cannot be singular.*

PROOF. If $y, y, s \cdot y$ is singular, write $y = s^n s_3 \cdot z \cdot s_1^{-1} s^{-m}$ where $n, m, |s_1|$ and $|s_3|$ are maximal. Then we have $s^n s_3 z s_1^{-1} s^{-m}$, $s^n s_3 z s_1^{-1} s^{-m}$, $s^{n+1} s_3 z s_1^{-1} s^{-m}$ singular. If $m \leq n$ we have $s^n s_3 z s_1^{-1}$, $s^{n-m} s_3 z$, $s_2 s^{n-m} s_3 z s_1^{-1} s^{-m}$ singular. But now we note that if $z \neq 1$ the maximality hypotheses yield a contradiction. Furthermore, if $s^n s_3 \neq 1$ we have either $ss_3 = s \cdot s_3$ with s_3 nontrivial or $s^2 = s \cdot s$, hence s is cyclically reduced. On the other hand, if $y = s_1^{-1} s^{-m}$, we have $s_1^{-1} s^{-m}$, $s_1^{-1} s^{-m}$, $s \cdot s_1^{-1} s^{-m}$ singular, with $m \geq 1$ by our length hypothesis. This leaves $(s_1^{-1} s_2^{-1})^m s_1^{-1}$, $(s_1^{-1} s_2^{-1})^{m-1} s_1^{-1}$, $s_1^{-1} s^{-m}$ singular. But this is a contradiction. The argument for the case $m > n$ and that for $y, y, s \cdot y^{-1}$ are very similar.

If we assume $y, s \cdot y, y$ is singular let $y = y_1 \cdot s_1^{-1} s^{-m}$ with $s_1^{-1} s^{-m}$ maximal. Now if $m \geq 1$, $y_1 \cdot s_1^{-1} s^{-(m-1)}$, $y_1 \cdot s_1^{-1} s^{-m}$, $y_1 \cdot s_1^{-1} s^{-m}$ is singular and, if $m = 0$, $y_1, s_2 \cdot y_1 \cdot s_1^{-1}$, $y_1 \cdot s_1^{-1}$ is singular. In the former case maximality of $s_1^{-1} s^{-m}$ requires that y_1 be an initial segment of $s^{m-1} s_1$ or $y_1 = s^{m-1} s_1 \cdot z$. In either situation we are left with an impossible singular triple. In the case $m = 0$ the situation is similar.

If we assume $y \cdot s, y \cdot s, y^{-1}$ is singular, let $y = y_1 \cdot s_2 s^n$ with $s_2 s^n$ maximal. Now $y_1 \cdot s_2 s^{n+1}$, $y_1, s_2 s_1, y^{-1}$ is singular. Let $s^{-m} s_4^{-1}$ be the largest initial segment of $s^{-(n+1)} s_2^{-1}$ with $y_1 = s^{-m} s_4^{-1} \cdot z$. Then $s^{-m} s_4^{-1} \cdot z \cdot (s_2 s^{n+1} s^{-m} s_4^{-1})$, $z \cdot s_2 s_1, z^{-1}, s_4 s^m$ is singular. But this is contradictory since $s_2 s_1 = s_2 \cdot s_1 \neq 1$ and $z \neq 1$ implies no further cancellation except in zz .

Now we consider a set of parametric words containing the set $U(w)$ in its set of values. Since the number of words on a finite alphabet of length bounded by any constant is finite and since it is clearly effective to determine if a particular word u satisfies $w(u) = 1$ we may restrict attention to solutions of length greater than λ for any fixed λ . For simplicity of exposition we restrict attention to solutions of length $L > 4M$ where M is the size of $w(x)$.

Let u be such a solution of $w(x) = 1$. Then $(c_1 u^{e_1})(c_2 u^{e_2}) \cdots (c_r u^{e_r}) = 1$. In the case that $t \geq 3$ the part of Proposition 1 concerning end triples $(1, w_1, w_2)$ and $(w_{n-1}, w_n, 1)$ may be ignored. We will henceforth

assume, using Proposition 1, that $t \geq 3$ and for some i , $(c_{i-1}u^{e_{i-1}})$, $(c_i u^{e_i})$, $(c_{i+1}u^{e_{i+1}})$ is singular. It will easily be seen that if $t=2$ then we have a special case of the parametric word we develop while $t=1$ is trivial.

Now, since $|c| \leq M$ for every coefficient c and $|u| > 4M$ by hypothesis we must have either $u^{e_{i-1}}$, c_i , u^{e_i} or u^{e_i} , c_{i+1} , $u^{e_{i+1}}$ singular. But we may assume the former since the problem which develops from the antiautomorphism reversing left and right and replacing all elements with their inverses is identical to the original problem.

We wish to show that every such solution u is of the form $d \cdot z^a \cdot y^b \cdot f$ where d and f have length less than $2M$, z and y are primitive with length less than or equal to M and a and b are nonnegative integers. Moreover we wish to show that d , z , y and f are restricted to a finite number of possibilities determinable from the c_i . We first note that if we can obtain the form $d \cdot \bar{z}^a \cdot \bar{y}^b \cdot f$ with $\bar{z}$, $\bar{y}$ cyclically reduced then we need only let z be the unique primitive word with $\bar{z} = z^m$ and y the unique primitive word with $\bar{y} = y^n$. Hence we will just show that y and z are cyclically reduced.

We may assume $e_i = 1$ since replacing x by x^{-1} in $w(v)$ and replacing u by u^{-1} makes no essential change. First we suppose $e_{i-1} = 1$. Then $c_i = p^{-1} \cdot r^{-1}$, $u = r \cdot v \cdot p$ where $|v| > 3M$.

Case 1. $e_{i+1} = 1$. Then $c_{i-1}u$, $c_i u$, $c_{i+1}u = c_{i-1}(r \cdot v \cdot p)$, $(p^{-1} \cdot r^{-1})(r \cdot v \cdot p)$, $c_{i+1}(r \cdot v \cdot p)$ is singular. Hence $c_{i-1}rv$, $v \cdot p$, $c_{i+1}rv \cdot p$ is singular, and since $|v| > 3M$ and each other letter represents a word of length at most $M/2$, we must have $c_{i+1}rv = p^{-1} \cdot w$ where either $w = s \cdot v$ or $v = s \cdot w$ for some s . Now $v = s \cdot w$ leads to w , $s \cdot w$, w singular which is impossible by Lemma 2 and hence $w = s \cdot v$ where s is a subword of $c_{i+1}r$ and hence one of a finite number of possibilities determinable from c_i and c_{i+1} of length $\leq M$. By length considerations and the fact that less than half of v may cancel in the product vw we have v , v , $s \cdot v$ singular. Hence, by Lemma 2, s is cyclically reduced and $u = r \cdot s^n s_3 \cdot s_1^{-1} s^{-m} \cdot p = r s_3 \cdot (s_4 s_3)^n \cdot (s_1^{-1} s_2^{-1})^m \cdot s_1^{-1} p$ and we have obtained the required form since $s_4 s_3$ and $s_1^{-1} s_2^{-1}$ as cyclic permutations of primitive words are primitive.

Case 2. $e_{i+1} = -1$. Then $c_{i-1}u$, $c_i u$, $c_{i+1}u^{-1} = c_{i-1}(r \cdot v \cdot p)$, $(p^{-1} r^{-1})(r \cdot v \cdot p)$, $c_{i+1}(p^{-1} \cdot v^{-1} \cdot r^{-1})$ is singular and hence $c_{i-1}rv$, $v \cdot p$, $c_{i+1}p^{-1}v^{-1}r^{-1}$ is singular. Now $c_{i+1}p^{-1} = p^{-1} \cdot w^{-1}$ and either $v^{-1} = s \cdot w^{-1}$ or $w^{-1} = s \cdot v^{-1}$. The former leads to $w \cdot s^{-1} = v$, $w \cdot s^{-1}$, $w \cdot s^{-1}$, w^{-1} singular which is impossible by Lemma 2. In the latter case v , v , $s \cdot v^{-1}$ is singular and Lemma 2 gives the desired form as before.

Next we suppose $e_{i-1} = -1$. Then $c_i = g \cdot h^{-1} \cdot g^{-1}$ with h cyclically reduced and $u = gh^m h_1 \cdot v$ for some v .

Case 3. $e_{i+1} = -1$. Let $c_{i+1} = k^{-1} \cdot l^{-1} \cdot k$, l cyclically reduced. Then v is some terminal segment of $l^n k$ for some $n \geq 0$ and $u = gh^m h_1 \cdot l_2 l^r k$ or $u = gh^m h_1 \cdot k_2$ and each is expressible in the proper form.

Case 4. $e_{i+1} = 1$. This is essentially symmetric to Case 2.

Now we have shown that the values of a finite number of parametric expressions of the form $d \cdot z^\alpha \cdot y^\beta \cdot f$ include all solutions u of $w(x) = 1$ of length at least $4M$. We will examine such solutions. Consider a particular parametric expression $d \cdot z^\alpha \cdot y^\beta \cdot f$. If only a finite number of values of one of the parameters yields solutions, the solutions determined by this expression are also determined by a finite number of expressions involving only the other parameter. Hence in showing that we can eliminate two parameter expressions we need consider only expressions such that the values of α and β yielding solutions are unbounded. Thus, where t is the degree and M the size of $w(x)$ we examine a particular value $u = d \cdot z^\alpha \cdot y^\beta \cdot f$ where $|z^\alpha| > (t+6)M$, $|y^\beta| > (t+6)M$. We may assume $y \neq z$ and $y, z \neq 1$ for otherwise we would have an instance of $d \cdot z^\alpha \cdot f$, a one parameter expression in our list, and clearly $z \neq y^{-1}$ by definition of free reduction. We will show that under the conditions described u cannot be a solution and hence the solutions form an effective finite set of one parameter expressions. Furthermore we will note that the solution expressions $d \cdot z^\alpha \cdot f$ obtained are solutions for every a with $|z^a| > (t+6)M$ and hence, for appropriate n , $(dz^n) \cdot z^\alpha \cdot f$ gives solutions for every non-negative value of α .

Suppose that $w(u) = 1$. Then $c_1(d \cdot z^\alpha \cdot y^\beta \cdot f)^{e_1} c_2(d \cdot z^\alpha \cdot y^\beta \cdot f)^{e_2} \cdots c_i(d \cdot z^\alpha \cdot y^\beta \cdot f)^{e_i} = 1$. After cyclic permutation we may rewrite this $\gamma_1 \gamma_2 \cdots \gamma_i = 1$ where γ_i is one of the following expressions:

$$\begin{array}{ll} (+ +) & y^b f c_i d z^a \quad \quad (- +) \quad z^{-a} d^{-1} c_i d z^a \\ (+ -) & y^b f c_i f^{-1} y^{-b} \quad \quad (- -) \quad z^{-a} d^{-1} c_i f^{-1} y^{-b} \end{array}$$

We note that if $e_{i-1} = 1$ and $e_i = -1$ then γ_i has sign $+-$ etc. Hence the second sign on γ_i must correspond to the first on γ_{i+1} .

Now we will use Lemma 1 and Lemma N to show that the assumptions that $|z^a| > (t+6)M$, $|y^b| > (t+6)M$, $y \neq z^{\pm 1}$, lead to a contradiction.

First we assume that it is not the case that $y = g \cdot h$, $z = g^{-1} \cdot h^{-1}$ for any g, h . Then using the argument of Lemma 1 we note that the entire word $y f c_i d z$ cannot have cancelled in the reduction of γ_i . But $y f = y \cdot f$ and $d z = d \cdot z$ so either $\gamma_i = y^{b-1} \cdot w \cdot z^{a-1}$ or, if either z or y cancels in $y f c_i d z$, either $f c_i = v \cdot d^{-1}$ or $c_i d = f^{-1} \cdot v$, and $|f c_i d| < 2M$. Thus if b' and a' are minimal such that $|y^{b'}| \geq |y^b| - |f c_i d| - |y| - |z| \geq (t+6)M$

$-2M - M - M = (t+2)M$ and $|z^{a'}| \geq |z^a| - |fc_id| - |y| - |z| \geq (t+2)M$, we must have $(++a) \gamma_i = y^{b'} \cdot w \cdot z^{a'}$.

The situation for type $(--)$ is symmetric and yields reduced form $(--a) z^{-a'} \cdot w \cdot y^{-b'}$ with the same length restriction. By Lemma 1, γ_i of type $+-$ may have reduced form $(+-a) y^{b'} \cdot w \cdot y^{-b''}$ with $|y^{b'}|, |y^{b''}| > (t+2)M$ or $(+-b), y^n \neq 1$ where $|y^n| \leq M$.

If all γ_i are of type b (i.e. $(+-b)$ or $(-+b)$) then $\gamma_1 \cdots \gamma_t$ is a product of factors which are alternately nontrivial powers of y and z , hence by Lemma N, y and z are powers of a common element, whence, by primitivity, $y = z^{\pm 1}$ contradicting our assumption. Now suppose that at least one of the γ_i is of type a . Then the length of its initial and terminal y^a or z^b is greater than the sum of the lengths of all the z^{n_i} and y^{n_i} which could arise from γ_i of type b . If we can show that no such initial y^a or z^b can fully cancel then we will contradict the hypothesis that $\gamma_1 \cdots \gamma_t = 1$. Consider a subword $z^{a'} y^{n_1} z^{n_2} \cdots y^{n_j} z^{-a''}$. For this to cancel $z^{a'}$ fully we must (by Lemma 1) have $y^{n_1} z^{n_2} \cdots y^{n_j} = z^n$ for some r . But this gives a nontrivial word in y and z which is 1 and by Lemma N, $y = z^{\pm 1}$. On the other hand in a subword $z^a y^{n_1} z^{n_2} \cdots z^{n_j} y^{b'}$, Lemma 1 requires $y = g \cdot h, z = g^{-1} \cdot h^{-1}$ for some g, h . Hence we are forced to the conclusion that $y = g \cdot h, z = g^{-1} \cdot h^{-1}$ for some g, h , and since $y \neq z^{-1}$ neither g nor h is 1.

Now, by Lemma 1 we may have additional forms of γ_i , $(++c)$ $(gh)^n g$ and $(--c) (g^{-1} h^{-1})^n g^{-1}$. We note $ghg^{-1} h^{-1} \neq 1$ for $gh = hg$ implies that y is not primitive. If no γ_i of type a appears then we have $\gamma_1 \cdots \gamma_n$ a word on g and h . We will show that it is nontrivial, and hence by Lemma N, g and h commute. Note that if the type of γ_i ends in $+$ (e.g. $++c, -+b$) its terminal symbol is g or h^{-1} while if it begins $+$ its initial symbol is g or h^{-1} . Since the type of γ_{i+1} must begin $+$ if that of γ_i ends $+$, trivial cancellation is here precluded. Similarly if the type of γ_i ends with $-$ its terminal symbol is g^{-1} or h as is the initial symbol of γ_{i+1} . Hence $\gamma_1 \cdots \gamma_t$ is a nontrivial word on g and h .

If any γ_i is of type a the corresponding conditions on first and last symbols hold but now Lemma 1 applies to force cancellation of gh against $g^{-1} h^{-1}$ and contradict primitivity of y . Thus the two parameter possibilities are eliminated.

Now we have a class of parametric expressions which may be effectively found in the form $d \cdot z^a \cdot f$, since our proof consisted in showing $z = y$. Now we need only show that there is an effective test to see if for particular d, z and $f, d \cdot z^a \cdot f$ is a solution for every value a of α with $|z^a| > (t+6)M$ and if not show that no such value yields a solution.

We may check solutions by defining δ_i similarly to the γ_i above, i.e., if we consider $c_1(d \cdot z^a \cdot f)^{e_1} c_2(d \cdot z^a \cdot f)^{e_2} \cdots c_t(d \cdot z^a \cdot f)^{e_t}$ we write

$\delta_i = fc_2dz^a$ or $fc_if^{-1}z^{-a}$ or $d^{-1}c_if^{-1}z^{-a}$ or $d^{-1}c_idz^a$. But now for the large values of a we consider, the solutions are clearly independent of the particular value of a as long as $|z^a| > (t+6)M$, for we obtain cancellations of $z^{a'}$ against $z^{-a''}$ where $a' - a''$ depend only on some subset of f , c_i and d . On the other hand we may choose the smallest substitution a for α with $|z^a| > (t+6)M$ so our test is effective.

REFERENCES

1. A. A. Lorenc, *The solution of systems of equations in one unknown in free groups*, Dokl. Akad. Nauk SSSR **148** (1963), 1253–1256=Soviet Math. Dokl. **4** (1963), 262–266.
2. ———, *Equations without coefficients in free groups*, Dokl. Akad. Nauk SSSR **160** (1965), 538–540=Soviet Math. Dokl. **6** (1965), 141–143.
3. R. C. Lyndon, *Equations in free groups*, Trans. Amer. Math. Soc. **96** (1960), 445–457.
4. R. C. Lyndon and M. P. Schutzenberger, *The equation $a^M = b^N c^P$ in a free group*, Michigan Math. J. **9** (1962), 289–298.

UNIVERSITY OF ILLINOIS