PACIFIC JOURNAL OF MATHEMATICS
Vol. 43, No. 3, 1972

RESIDUAL PROPERTIES OF FREE GROUPS

STEPHEN J. PRIDE

In this paper the following theorem is proved: if π is an infinite set of primes and n is an odd integer greater than one, then free groups are residually $\{PSL(n, p); p \in \pi\}$. As a by-product of the proof new generators of $SL(n, p)$ are obtained for nearly all primes p.

1. The main result. For unexplained notation the reader is referred to [8].

Let $\mathscr{A}_1$ and $\mathscr{A}_2$ be sets of groups. $\mathscr{A}_1$ is said to be *residually* $\mathscr{A}_2$ iff, for each group G belonging to $\mathscr{A}_1$ and each non-identity element g of G there is a homomorphism φ (depending on G and g) which maps G onto some element H of $\mathscr{A}_2$, and is such that $\varphi(g)$ is not the identity of H. An equivalent formulation is: for each G in $\mathscr{A}_1$ there is a set of normal subgroups $\{N_i\}_{i\in I}$ of G such that $\bigcap_{i\in I} N_i = 1$ and for each i in I, G/N_i is isomorphic to an element of $\mathscr{A}_2$. It is obvious that if $\mathscr{A}_1$ and $\mathscr{A}_2$ are sets of groups and some or all of the members of $\mathscr{A}_1$ and $\mathscr{A}_2$ are replaced by isomorphic copies, yielding new sets $\mathscr{A}_1'$ and $\mathscr{A}_2'$ respectively, then $\mathscr{A}_1$ is residually $\mathscr{A}_2$ iff $\mathscr{A}_1'$ is residually $\mathscr{A}_2'$. It is also easy to see that if $\mathscr{A}_1$ is residually $\mathscr{A}_2$, and $\mathscr{A}_2$ is residually $\mathscr{A}_3$, then $\mathscr{A}_1$ is residually $\mathscr{A}_3$.

Let $\{x_1, x_2, x_3, \cdots\}$ be a fixed but arbitrary countably infinite set, and let F_n be the free group freely generated by $\{x_1, x_2, \cdots, x_n\}$. Denote by $\mathscr{F}$ the set $\{F_n: n \geqq 2\}$. In recent years there has been some investigation into which sets, $\mathscr{A}$, of groups are such that $\mathscr{F}$ is residually $\mathscr{A}$. The two-generator groups in $\mathscr{A}$ must of necessity generate the variety, $\mathscr{O}$, of all groups. It has been conjectured by S. Meskin that this condition is also sufficient. A rich source of sets of groups which generate $\mathscr{O}$ is a result of Heineken and Neumann [3] which states that every infinite set of pairwise non-isomorphic known (1967) finite non-abelian simple groups generates the variety of all groups. This theorem has presumably inspired several of the results obtained so far. Thus Katz and Magnus [5] have proved that $\mathscr{F}$ is residually $\{A_n: n \in J\}$, where A_n is the alternating group on $\{1, 2, \cdots, n\}$ and J is an infinite set of positive odd integers; and Gorčakov and Levčuk [2] have proved that $\mathscr{F}$ is residually any infinite subset of the set of simple groups $PSL(2, p^r)$. This latter result generalizes theorems obtained in [6], [5] and [7], which consider the cases $r = 1$ and p variable, $r > 1$ and fixed and p variable, $p > 11$ and fixed and r variable, respectively.

In this paper the following main result is obtained.

THEOREM 1. *Let n be an odd integer greater than one, and let π be an infinite set of primes. Then $\mathscr{F}$ is residually $\{PSL(n, p)$: $p \in \pi\}$.*

Before discussing the proof of Theorem 1 some notation and definitions will be introduced. Let R be a commutative ring with identity 1. The ring of polynomials in the indeterminant x with coefficients from R will be denoted by $R[x]$. The degree of an element $f(x)$ of $R[x]$ will be written as $\deg(f(x))$. As is well-known (see [4], page 56) the $n \times n$ matrices with entries from R form a ring with identity. The identity will be denoted by E. The $n \times n$ matrix with 1 in its ith row and jth column and zeros elsewhere will be denoted by E_{ij} $(i, j = 1, 2, \cdots, n)$, and $E_{(n+i)j}$, $E_{(n+i)(n+j)}$, $E_{i(n+j)}$ $(i, j = 1, 2, \cdots, n)$ will all be defined to be equal to E_{ij}. The multiplicative semigroup of the ring of $n \times n$ matrices with entries from R has a sub-semigroup consisting of all matrices which have a single nonzero entry, namely 1, in each row and each column. This sub-semigroup is in fact a group, isomorphic to the symmetric group on $\{1, 2, \cdots, n\}$. An isomorphism is given by:

$$\sigma \longrightarrow \sum_{i=1}^{n} E_{i\sigma(i)} ,$$

where σ is a permutation of $\{1, 2, \cdots, n\}$. The matrix $\sum_{i=1}^{n} E_{i\sigma(i)}$ will be called the *permutation matrix corresponding to σ*. When no confusion can arise, and if it is convenient to do so, the matrix $\sum_{i=1}^{n} E_{i\sigma(i)}$ will be denoted by the permutation σ.

For the rest of this section n will denote a fixed but arbitrary odd integer greater than one, and p (possibly subscripted) will stand for a prime number. To simplify the proof of Theorem 1, use is made of the following two results:

(i) *$\mathscr{F}$ is residually $\{F_2\}$,*

(ii) *For each $k \geqq 2$, $\{F_2\}$ is residually $\{T_k\}$, where $T_k = (a, b \mid a^k)$.*

The former result is proved in [6], whilst Lemma 1 of [5] proves (ii) for the case $k = 2$, and the proof for $k > 2$ is entirely analogous. Using (i) and (ii) reduces the proof of Theorem 1 to showing that $\{T_n\}$ is residually $\{PSL(n, p)$: $p \in \pi\}$.

The first step in proving that $\{T_n\}$ is residually $\{PSL(n, p)$: $p \in \pi\}$ is to find a group of $n \times n$ matrices which is isomorphic to T_n. Consider the ring of $n \times n$ matrices with entries from $Z[x]$. The multiplicative semigroup of this ring has a sub-semigroup consisting of all matrices with determinant ± 1. This sub-semigroup is a group, called the *group of units*. The permutation matrix X corresponding to the permutation $(1, 2, 3, \cdots, n)$, and the matrix $Y = E + x \sum_{j=2}^{n} E_{j1}$ are in the group of units. They therefore generate a subgroup, $\mathscr{U}_n$, of this group. Notice that in this group X has order n and Y has infinite order. In § 2 the following result is proved.

LEMMA 1. *When a product of the form*

$$(*) \qquad Y^{\nu} X^{\delta_1} Y^{m_1} \cdots X^{\delta_r} Y^{m_r} X^{\mu}$$

—where $r \geqq 0$, the δ_i can have the values $1, 2, \cdots, n-1$, the m_i can have any integer values except zero, ν can have any integer value, μ can be $0, 1, 2, \cdots, n-1$, ν and μ cannot be zero simultaneously unless $r \geqq 1$—is multiplied out, it has an entry of degree at least one, provided ν and r are not both zero.

From this lemma follows immediately

THEOREM 2. *$\mathscr{U}_n$ and T_n are isomorphic.*

The problem is now reduced to showing that $\{\mathscr{U}_n\}$ is residually $\{PSL(n, p): p \in \pi\}$. There are plenty of homomorphisms from $\mathscr{U}_n$ into $SL(n, p)$. In fact, let α be a nonzero element of $GF(p)$. Then, by Theorem 4 of Chapter III [4], there is a ring homomorphism of $Z[x]$ onto $GF(p)$ which maps x to α. This homomorphism induces a homomorphism φ_α from the multiplicative semigroup of all $n \times n$ matrices with entries from $Z[x]$ to the multiplicative semigroup of all $n \times n$ matrices with entries from $GF(p)$. The value of φ_α at the matrix M is obtained by replacing all appearances of x in M by α, and replacing all integers appearing as coefficients in the polynomials in M by their congruence classes modulo the prime p. When restricted to $\mathscr{U}_n$, φ_α is a group homomorphism with range contained in $SL(n, p)$. Let $\varphi_\alpha(X) = C$ and $\varphi_\alpha(Y) = D(\alpha)$. It is easy to see that the subgroup of $SL(n, p)$ generated by C and $D(\alpha)$ is the same as that generated by C and $D = D(1)$. For there are integers t and u such that $t\alpha = 1$ and $u1 = \alpha$, and so $D(\alpha)^t = D$ and $D^u = D(\alpha)$. In § 3 the following result is proved.

THEOREM 3. *Let p be a prime which does not divide $3(n-1)$. Then C and D generate $SL(n, p)$.*

(If p divides $3(n-1)$, the validity of the theorem remains undecided.)

It follows immediately from Theorem 3 that φ_α is a homomorphism of $\mathscr{U}_n$ onto $SL(n, p)$ for all but a finite number of primes p.

Using Lemma 1 and Theorems 2 and 3, it is now possible to prove that $\{\mathscr{U}_n\}$ is residually $\{PSL(n, p): p \in \pi\}$. It is well-known (see [8], page 158) that the centre of $SL(n, p)$ consists of all scalar matrices λE, where $\lambda^n = 1$. Given a non-identity element W of $\mathscr{U}_n$, it will be shown that there is a prime p in π, and a homomorphism φ of $\mathscr{U}_n$ onto $SL(n, p)$ such that $\varphi(W)$ does not belong to the centre of $SL(n, p)$. Then the composition of φ with the natural homomorphism of $SL(n, p)$ onto $PSL(n, p)$ gives a homomorphism of $\mathscr{U}_n$ onto $PSL(n, p)$ which does not map W to the identity.

Thus, let W be a non-identity element of $\mathscr{U}_n$. Then W can be expressed uniquely as a product of the form (*) (see Lemma 1). First suppose that in the product (*) $\nu = 0$ and $r = 0$, so that $W = X^\mu$, where μ is an integer and $0 < \mu < n$. Let p_0 be a prime in π which does not divide $3(n-1)$. Then the homomorphism of $\mathscr{U}_n$ onto $SL(n, p_0)$ determined by

$$\begin{aligned} X &\longrightarrow C \\ Y &\longrightarrow D \end{aligned}$$

does not map W to the centre of $SL(n, p_0)$.

Suppose now that the product (*) is such that not both of ν and r are zero. Then by Lemma 1, W has an entry

$$a_0 + a_1 x + \cdots + a_s x^s \text{ with } a_s \neq 0,\ s \geqq 1.$$

Let p_0 be a prime in π with the property

$$p_0 - 1 > \max\{|a_s|, s(n+1)\}\ .$$

The congruence class of an integer $k \bmod p_0$ will be denoted by $\bar{k}$. Consider the polynomials

$$\begin{aligned} f(x) &= \bar{a}_0 + \bar{a}_1 x + \cdots + \bar{a}_s x^s\ , \\ g(x) &= f(x)[(f(x))^n - \bar{1}]\ , \end{aligned}$$

which are elements of $GF(p_0)[x]$. Since $\bar{a}_s \neq \bar{0}$, $\deg(f(x)) = s$, and so $\deg(g(x)) = s(n+1)$. By the choice of p_0 there is a nonzero element α of $GF(p_0)$ which is not a root of $g(x)$.

Let φ be the homomorphism of $\mathscr{U}_n$ onto $SL(n, p_0)$ determined by

$$\begin{aligned} X &\longrightarrow C \\ Y &\longrightarrow D(\alpha)\ . \end{aligned}$$

(Note that p_0 does not divide $3(n-1)$, so Theorem 3 applies.) The entries of $\varphi(W)$ are obtained from those of W by replacing x by α and working mod p_0. Hence $\varphi(W)$ has

$$f(\alpha) = \bar{a}_0 + \bar{a}_1 \alpha + \cdots + \bar{a}_s \alpha^s$$

as one of its entries. By the choice of α, $f(\alpha) \neq \bar{0}$ and $f(\alpha)^n \neq \bar{1}$, so clearly $\varphi(W)$ does not lie in the centre of $SL(n, p_0)$.

2. Proof of Lemma 1. In this and the next section it will be useful to keep in mind the following rule for calculating with permutation matrices. If M is a $u \times u$ matrix and P is the permutation matrix corresponding to a permutation σ of $\{1, 2, \cdots, u\}$, then PM is obtained from M by replacing row i by row $\sigma(i)$, and MP is obtained from M by replacing column i by column $\sigma^{-1}(i)$ $(1 \leqq i \leqq u)$.

Before proving Lemma 1, it should be pointed out that the result is also valid when n is even (the proof given below does not depend upon n being odd), but in this case the permutation matrix corresponding to $(1, 2, 3, \cdots, n)$ has determinant -1, so that the result is not of any use here.

A product of the form (*) (as in the statement of Lemma 1) in which $\nu = \mu = 0$ will be called a *product of type-*(XY). When such a product is multiplied out, a matrix with entries $\xi_{ij}^{(r)}$ $(i, j = 1, 2, \cdots, n)$ from $Z[x]$ is obtained. The following assertion will be proved by induction on r.

$$(++) \qquad \begin{aligned} &\deg(\xi_{11}^{(r)}) = r \\ &\deg(\xi_{1j}^{(r)}) < r \text{ for } j = 2, 3, \cdots, n\,. \end{aligned}$$

For $r = 1$ the product is just $X^{\delta_1}Y^{m_1}$, which is equal to $X^{\delta_1} + m_1 x \sum_{j=2}^{n} E_{(n+j-\delta_1)1}$. Thus

$$\xi_{i1}^{(1)} = \begin{cases} m_1 x & i \neq n + 1 - \delta_1 \\ 1 & i = n + 1 - \delta_1\,. \end{cases}$$

All other entries of $X^{\delta_1}Y^{m_1}$ are either zero or one. Since $0 < \delta_1 < n$, it follows that $1 < n + 1 - \delta_1 < n + 1$, so that $\xi_{11}^{(1)}$ is $m_1 x$. Thus $(++)$ holds when $r = 1$.

Now assume $(++)$ holds for all $s < r$, where $r > 1$. The first row of $X^{\delta_1}Y^{m_1} \cdots X^{\delta_{r-1}}Y^{m_{r-1}}X^{\delta_r}Y^{m_r}$ is obtained from that of $X^{\delta_1}Y^{m_1} \cdots X^{\delta_{r-1}}Y^{m_{r-1}}$ by right multiplication by $X^{\delta_r}Y^{m_r}$. Thus

$$\xi_{11}^{(r)} = \sum_{\substack{1 \leqq j \leqq n \\ j \neq n+1-\delta_r}} m_r x \xi_{1j}^{(r-1)} + \xi_{1(n+1-\delta_r)}^{(r-1)}\,.$$

Since $1 < n + 1 - \delta_r < n + 1$, it follows that

$$\begin{aligned} \deg(\xi_{11}^{(r)}) &= \deg(\xi_{11}^{(r-1)}) + 1 \\ &= r\,. \end{aligned}$$

Now except for column one, every column of $X^{\delta_r}Y^{m_r}$ contains only zeros and ones. Hence for $2 \leqq j \leqq n$,

$$\begin{aligned}\deg(\xi_{1j}^{(r)}) &\leqq \max\{\deg(\xi_{1t}^{(r-1)})\colon t = 1, 2, \cdots, n\}\\ &\leqq r-1\\ &< r\,.\end{aligned}$$

This shows that $(++)$ holds for r, and completes the induction proof.

Now take a product of the general form (*) in which not both of ν and r are zero, and let W be the matrix obtained when this product is multiplied out. It is required to show that W has an entry of degree at least one.

Case (i). $\nu = \mu = 0$. The product is of type-(XY), so W has an entry of degree r, by $(++)$.

Case (ii). $\nu \neq 0$, $\mu \neq 0$. Since

$$W^{-1} = X^{n-\mu} Y^{-m_r} X^{n-\delta_r} \cdots Y^{-m_1} X^{n-\delta_1} Y^{-\nu}$$

and the product on the right is of type-(XY), W^{-1} has an entry of degree at least one by $(++)$; consequently W has also.

Case (iii). $\nu \neq 0$, $\mu = 0$. If $r = 0$, W is just Y^{ν}, which has νx as one of its entries. Suppose then that $r \geqq 1$. $X^{\delta_1} Y^{m_1} \cdots X^{\delta_r} Y^{m_r}$ is a product of type-(XY), so the entries $\xi_{1j}^{(r)}$ $(j = 1, 2, \cdots, n)$ in the first row of the matrix U obtained when this product is multiplied out satisfy $(++)$. The first row of W is the same as that of U, so W has an entry of degree r.

Case (iv). $\nu = 0$, $\mu \neq 0$. If U is the matrix obtained when $X^{\delta_1} Y^{m_1} \cdots X^{\delta_r} Y^{m_r}$ is multiplied out, then U has an entry of degree r, and since W is just obtained from U by a permutation of columns, W also has an entry of degree r.

This completes the proof of Lemma 1.

3. Proof of Theorem 3. The following definitions are used. A matrix of the form $E + \lambda E_{ij}$, where $\lambda \neq 0$ and $i \neq j$, will be called a *transvection*. In a group G the *commutator* $[g_1]$ of $g_1 \in G$ will be defined to be g_1, the *commutator* $[g_1, g_2]$ of $g_1, g_2 \in G$ will be defined to be $g_1 g_2 g_1^{-1} g_2^{-1}$, and for $n \geqq 3$, $[g_1, g_2, \cdots, g_n]$ will be defined to be $[[g_1, \cdots, g_{n-1}], g_n]$. If S is a nonempty subset of G then $sgpS$ will denote the subgroup of G generated by S.

Let n denote a fixed but arbitrary odd integer greater than one, and let p be a fixed but arbitrary prime which does not divide $3n - 3$. It is required to show that the elements

$$C = \sum_{i=1}^{n} E_{i(i+1)}$$

$$D = E + \sum_{j=2}^{n} E_{j1} ,$$

of $SL(n, p)$ generate this group. It will be shown below that the transvection $E + E_{1n}$ belongs to $sgp\{C, D\}$, and from this the result follows, as is now indicated.

It is well-known (see [8], page 158) that the transvections

$$E + \lambda E_{ij} \ (i \neq j; i, j = 1, 2, \cdots, n) ,$$

where λ ranges over the nonzero elements of $GF(p)$, generate $SL(n, p)$. In fact, it is enough to choose one value of λ, say λ_{ij}, for each pair (i, j). For λ_{ij} has order p in the additive group of $GF(p)$, and so as t runs through the integers from 1 to $p - 1$, $t\lambda_{ij}$ assumes every nonzero element of $GF(p)$. Since

$$(E + \lambda_{ij}E_{ij})^t = E + (t\lambda_{ij})E_{ij} \ (i \neq j; i, j = 1, 2, \cdots, n)$$

all transvections can be obtained from the $E + \lambda_{ij}E_{ij}$. Notice that, in particular, the value 1 can be chosen for each λ_{ij}.

Let $\mathscr{H} = sgp\{E + E_{1n}, C\}$. Now for $i, j = 1, \cdots, n$

$$(^{**}) \qquad CE_{ij}C^{-1} = E_{(n+i-1)(n+j-1)} .$$

Therefore

$$\begin{aligned} C^r(E + E_{1n})C^{-r} &= E + E_{(n+1-r)(n-r)} \\ &= \tau_r, \text{ say } (0 \leqq r \leqq n - 1) . \end{aligned}$$

It is easily shown that

$$[\tau_0, \tau_1, \cdots, \tau_s] = E + E_{1(n-s)} \ (0 \leqq s \leqq n - 2) .$$

Thus $\mathscr{H}$ contains all the transvections

$$E + E_{1h} \ h = 2, 3, \cdots, n .$$

Finally, using $(^{**})$ k times $(0 \leqq k \leqq n - 1)$ gives

$$C^k(E + E_{1h})C^{-k} = E + E_{(n+1-k)(n+h-k)}, \ h = 2, 3, \cdots, n,$$

and so $\mathscr{H}$ contains all the transvections

$$E + E_{ij} \ (i \neq j; i, j = 1, 2, \cdots, n) .$$

Therefore $\mathscr{H} = SL(n, p)$.

It will now be shown that $E + E_{1n}$ belongs to $sgp\{C, D\}$. Straightforward computations show

$$\begin{aligned}[D^{-1}, C^{-1}]D &= E + E_{11} + E_{12} - E_{21} - E_{22} \\ &= P, \text{ say} \\ [D^{-1}, C^{-2}]D &= E + E_{11} + E_{13} - E_{31} - E_{33} \\ &= Q, \text{ say} \\ C^{-1}([D^{-1}, C^{-1}]D)C &= E + E_{22} + E_{23} - E_{32} - E_{33} \\ &= R, \text{ say.}\end{aligned}$$

Let t be an integer such that $6t \equiv 1 \bmod p$ (such a t exists since p is not 2 or 3). Then

$$(QP^{-1}R^{-1})^{2t} = E - E_{13} + E_{23} .$$

This element will be denoted by T. It turns out to be extremely useful.

Another useful element is

$$T^2RP = \sum_{i=4}^{n} E_{ii} + E_{12} + E_{23} + E_{31} .$$

This is just the permutation matrix corresponding to the permutation (123). Since, for $m \geqq 3$ and odd, the permutations (123) and $(123 \cdots m)$ generate the alternating group A_m ([1], page 67), it follows that $sgp\{C, D\}$ contains all even permutation matrices.

Suppose that n is greater than 3. It is easy to see that

$$(1) \qquad (34 \cdots n)T^{-1}(34 \cdots n)^{-1} = E + E_{1n} - E_{2n}$$

$$(2) \qquad (1s)(2, s+1)(E + E_{1n} - E_{2n})(1s)(2, s+1) = E + E_{sn} - E_{(s+1)n} \quad (3 \leqq s \leqq n-2)$$

and

$$(3) \qquad (123)^{-1}(E + E_{1n} - E_{2n})(123) = E + E_{2n} - E_{3n} .$$

From (1), (2) and (3) it follows that $sgp\{C, D\}$ contains all the matrices

$$\Lambda_\lambda = E + E_{\lambda n} - E_{(\lambda+1)n} \quad 1 \leqq \lambda \leqq n-2 .$$

This is also obviously true if n equals 3.

Now take the matrix

$$CDC^{-1} = E + \sum_{i=1}^{n-1} E_{in} .$$

Multiplying by Λ_{n-2} (on either side, since each Λ_λ commutes with CDC^{-1}) gives $E + \sum_{i=1}^{n-3} E_{in} + 2E_{(n-2)n}$. Then multiplying by Λ_{n-3}^2 gives $E + \sum_{i=1}^{n-4} E_{in} + 3E_{(n-3)n}$. Continuing in this manner finally gives the matrix $E + (n-1)E_{1n}$. Formally,

$$\left(\prod_{j=1}^{n-2} A^{j}_{(n-1)-j}\right)(CDC^{-1}) = E + (n-1)E_{1n} .$$

Since p does not divide $n-1$, there is an integer t such that $t(n-1) \equiv 1 \bmod p$. Then

$$(E + (n-1)E_{1n})^{t} = E + E_{1n} .$$

This shows that $sgp\{C, D\}$ contains the transvection $E + E_{1n}$, and completes the proof of Theorem 3.

Acknowledgement. I thank Dr. M. F. Newman for his help and encouragement.

References

1. H. S. M. Coxeter and W. O. J. Moser, *Generators and Relations for Discrete Groups*, 2nd edition, Ergebnisse der Mathematik, Band 14, Springer-Verlag, Berlin-Göttingen-Heidelberg-New York, 1965.
2. Yu. M. Gorčakov and V. M. Levčuk, *Concerning a residual property of free groups*, Algebra i Logika, **9** (1970), 415-421 (Russian).
3. Hermann Heineken and Peter M. Neumann, *Identical relations and decision procedures for groups*, J. Austral. Math. Soc., **7** (1967), 39-47.
4. Nathan Jacobson, *Lectures in Abstract Algebra*, Volume 1, van Nostrand, Princeton, 1951.
5. Robert A. Katz and Wilhelm Magnus, *Residual properties of free groups*, Comm. Pure Appl. Math., **22** (1969), 1-13.
6. Ada Peluso, *A residual property of free groups*, Comm. Pure Appl. Math., **19** (1966), 435-437.
7. Samuel Poss, *A residual property of free groups*, Comm. Pure Appl. Math., **23** (1970), 749-756.
8. Joseph J. Rotman, *The Theory of Groups: An Introduction*, Allyn and Bacon, Boston, 1965.

Received September 23, 1971. I would like to acknowledge the support of an Australian Commonwealth Postgraduate Award.

INSTITUTE OF ADVANCED STUDIES
AUSTRALIAN NATIONAL UNIVERSITY
CANBERRA, ACT.