Some Topics in Asymptotic Group Theory

WILLIAM M. KANTOR*

University of Oregon

1. ENUMERATION

There is nothing unusual about asymptotics in finite group theory: there are a number of known (or even well-known) asymptotic results. While these are not really the subject of this paper, it seems appropriate to begin with some especially intriguing examples (the first and last of which will be need later).

1.1. If p is prime then the number of isomorphism classes of groups of order p^k is at least $p^{\frac{2}{27}k^3 - 6k}$ (Higman [Hi]), and asymptotically $= p^{\frac{2}{27}k^3 + O(k^{8/3})}$ (Sims [Si]).

1.2 (Neumann [Ne, MN]).
The number of isomorphism classes of groups of order n is less than $n^{\frac{1}{2}(\log n)^2}$.

Here, and throughout this paper, logarithms will be to the base 2. The preceding result, as well as the next three, depend on the classification of finite simple groups.

1.3 (Holt [Ho]).

$$\frac{\#\text{ isomorphism classes of perfect groups of order} \le n}{\#\text{ isomorphism classes of groups of order} \le n} \longrightarrow 0 \text{ as } n \longrightarrow \infty.$$

1.4 (Cameron-Neumann-Teague [CaNT]).
For almost all n (in the sense of density), the only primitive permutation groups of degree n are S_n and A_n.

1.5 (Cameron [Ca]). If G is a primitive subgroup of S_n, then either

(i) $n = \binom{m}{k}^{\ell}$ and G is a subgroup of S_m *wreath* S_ℓ with S_m acting on the k-sets of an m-set and the wreathed product having the product action, or

(ii) $|G| \le n^{C \log n}$ for some constant C.

Here **1.1-1.4** are, in some sense, in a classical enumerative vein: estimate the number of groups of a certain sort. The remainder of this paper is concerned with rather different types of asymptotic questions, involving lengths of presentations or of words given in terms of generators, or the proportion of pairs of elements of a simple group that generate the group. These questions (or at least those in §§2,4) are motivated, to some extent, by questions that arose in Theoretical Computer Science. Sketches of some proofs will be given, especially in the cases of some results not in print. For a survey of related results see [BHKLS].

2. SHORT PRESENTATIONS

The *length* of a presentation $\langle X \mid R \rangle$ is the sum of $|X|$ together with the sum of the lengths of all of the members of R as words in $X \cup X^{-1}$. This is motivated, in part, by thinking of inputting $\langle X \mid R \rangle$ into a computer.

Stupid-looking Example: The usual presentation for a cyclic group of order n has length $1 + n$. A shorter presentation is $\langle X \mid R \rangle$ with $X = \{x_0,...,x_m\}$ and $R = \{ x_{i+1}x_i^{-2}, x_0^{a_0}x_1^{a_1}...x_m^{a_m} \mid i = 0,...,m \}$, where $m = [1 + \log_2 n]$ and $n = \Sigma a_i 2^i$ in base 2. Its length is at most $(m + 1) + 3m + (m + 1) \leq 5\log n + 7$. In fact, however stupid-looking this may seem to be as a way to represent a cyclic group, this method has, indeed, been used in practise.

2.1 Conjecture: *Every finite group* G *has a presentation of length* $O((\log|G|)^3)$.

Note that the constant 3 is best possible here. For, using Higman's bound **1.1**, for all $\varepsilon > 0$ and all $C > 0$ it is easy to check that

$$\frac{\#\text{ p-groups of order } p^k}{\#\text{ presentations of length} \leq C(\log p^k)^{3-\varepsilon}} \longrightarrow \infty \text{ as } k \longrightarrow \infty$$

since the denominator is straightforward to calculate.

2.2 Theorem (Babai-Kantor-Luks-Pálfy [BKLP]). *The Conjecture is true except, perhaps, if some composition factor of* G *is isomorphic to* $^2A_2(q)$, $^2B_2(q)$ *or* $^2G_2(q)$.

The remainder of this section is devoted to an indication of the ideas involved in the proof of this theorem, along with comments on the difficulties encountered with the groups $^2A_2(q)$, $^2B_2(q)$ and $^2G_2(q)$. The proof falls into two main steps: I. Simple Groups, and II. Glueing.

STEP I. Simple Groups.

2.3 Proposition. *Every simple group has a presentation of length* $O((\log|G|)^2)$, *except perhaps for the groups* $^2A_2(q)$, $^2B_2(q)$ *and* $^2G_2(q)$.

Sporadic groups can be ignored here. The usual presentation for A_n has length $< n^2$. Therefore, it remains only to consider a group G of Lie type over $\mathbb{F}_q$, of characteristic p and rank ℓ, say, in which case the order of magnitude of $\log|G|$ is $\ell^2\log q$. We presuppose various parts of [Car] here and in later portions of this paper.

Groups of Rank $\ell \geq 2$. Here the obvious approach is to try to use the Curtis-Steinberg-Tits presentation [Cu], but this is much too long (its length involves q instead of log q). Nevertheless, it is not at all surprising that this presentation can be modified so as to behave as desired. The details are as follows.

Assume, for the moment, that G is untwisted. Then the Curtis-Steinberg-Tits presentation for some perfect central extension $\tilde{G}$ of G uses generators $x_\alpha(t)$ for certain roots α, where $t \in \mathbb{F}_q$. (Specifically, α belongs to the union of the rank 2 subsystems determined by pairs of fundamental roots, so the number of these roots α has order of magnitude ℓ^2.) The relations are

$$x_\alpha(t)x_\alpha(u) = x_\alpha(t + u)$$

$$[x_\alpha(t), x_\beta(u)] = \prod_{i,j>0} x_{i\alpha+j\beta}(C_{ij\alpha\beta}t^iu^j)$$

for all relevant α and β with $\beta \neq \pm\alpha$, and all $t, u \in \mathbb{F}_q$, where i, j and $C_{ij\alpha\beta}$ are integers (and $|C_{\alpha\beta ij}| \leq 3$). In order to shorten this presentation, let $\theta_1,...,\theta_e$ be a basis of $\mathbb{F}_q$ over $\mathbb{F}_p$. Then use only the generators $x_\alpha(\theta_k)$, together with the relations of the form

$$x_\alpha(\theta_k)^p = 1, \qquad [x_\alpha(\theta_k), x_\alpha(\theta_m)] = 1$$

$$[x_\alpha(\theta_k), x_\beta(\theta_m)] = \prod_{i,j>0} x_{i\alpha+j\beta}(C_{ij\alpha\beta}\theta_k^i\theta_m^j),$$

which are interpreted as follows. For any $s_k \in \mathbb{F}_p$ and any root γ, expand $x_\gamma(\Sigma_k s_k\theta_k)$ as $\Pi_k x_\gamma(\theta_k)^{s_k}$, where expressions such as $x_\alpha(\theta_k)^p$ and $x_\gamma(\theta_k)^{s_k}$ are themselves expanded as in the Stupid-looking Example by adjoining up to 7log p additional generators and relations for each such term. The length of this presentation is dominated by that of the commutator relations, which is $O(\ell^2\ell^2 \cdot ee \cdot e\log p)$. Thus, this is a presentation for $\tilde{G}$ of length $O((\log|G|)^3)$.

In order to shorten this presentation somewhat, and at the same time kill the center of $\tilde{G}$, choose each θ_k as θ^k for a generator θ of $\mathbb{F}_q^*$. In addition to the generators $x_\alpha(\theta^k)$, introduce generators $w_\alpha(1)$, $w_\alpha(\theta)$ and h_α, together with the following relations for all α and $\beta \neq \pm\alpha$ restricted as above:

$$w_\alpha(1) = x_\alpha(1)x_{-\alpha}(1)^{-1}x_\alpha(1) \qquad w_\alpha(\theta) = x_\alpha(\theta)x_{-\alpha}(-\theta^{-1})x_\alpha(\theta)$$

$$h_\alpha = w_\alpha(\theta)w_\alpha(1)^{-1} \qquad [h_\alpha, h_\beta] = 1$$

$$x_\alpha(\theta^k)^{h_\alpha} = x_\alpha(\theta^2\theta^k) \qquad x_\alpha(\theta^k)^{h_\beta} = x_\alpha(\theta^{2(\alpha,\beta)/(\beta,\beta)}\theta^k)$$

(where $x_{-\alpha}(-\theta^{-1})$, $x_\alpha(\theta^2\theta^k)$ and $x_\alpha(\theta^{2(\alpha,\beta)/(\beta,\beta)}\theta^k)$ are computed by expansion as above, and (α, β) denotes the usual inner product of roots). Then our earlier relations can be shortened to

$$x_\alpha(\theta^k)^p = 1, \qquad [x_\alpha(\theta^k), x_\alpha(\theta^m)] = 1$$

$$[x_\alpha(\theta^k), x_\beta(\theta^m)] = \prod_{i,j>0} x_{i\alpha+j\beta}(C_{ij\alpha\beta}\theta^{ik+jm}),$$

with $0 \le k, m \le 1$. For, these and conjugation by the various elements h_α imply the remaining ones given earlier. Now there are $O((\ell e + \ell + \ell)\log p)$ generators, and the relations have total length $O(\ell + \ell\cdot e\log p + \ell + e\ell + \ell\log p + \ell e + \ell^2\cdot e\log p) = O(\log|G|)$. The center of $\tilde{G}$ is killed using products of the elements h_α, where the required relations are explicitly written down on a case by case basis.

All of this involves an unfortunate loss of the beauty of the original Curtis-Steinberg-Tits presentation in order to achieve efficiency.

Twisted Groups of Rank $\ell \ge 2$. While there are straightforward modifications of the above presentations valid for twisted groups, an annoying snag does occur. Namely, we used the fact that $x_\alpha(t)^{h_\alpha} = x_\alpha(\theta^2 t)$ in order to see that $\langle h_\alpha \rangle$ had very few orbits on X_α. Such a situation does not occur for odd-dimensional unitary groups, in which one type of root group is nonabelian. Nevertheless, in this case a presentation of length $O((\log|G|)^2)$ can still be obtained.

Very briefly, $G = PSU(2\ell + 1, q)$, $\ell \ge 2$, is best viewed as having a "root system" of type BC_ℓ, namely, the union of a B_ℓ and a C_ℓ system. However, [Gr] provides a presentation suitable for our purposes using a C_ℓ system Φ. His generators are $x_\alpha(t)$ with α short, $t\in\mathbb{F}_{q^2}$, as well as $x_\alpha(t,u)$ with α long, $t, u\in\mathbb{F}_{q^2}$ and $u + \bar{u} = \varepsilon_\alpha t\bar{t}$, where $\varepsilon_\alpha = \pm 1$ and the involutory field automorphism of $\mathbb{F}_{q^2}$ is $t \longrightarrow \bar{t}$. The obvious sorts of commutator relations then suffice for a presentation (cf. [Gr]). These can be shortened by using generators $x_\alpha(\theta^k)$ for short roots α, as well as suitable generators $x_\alpha(0, \theta^m)$ and $x_\alpha(\theta^k, \theta^m)$ for long roots α, where $q^2 = p^e$ and θ is a generator of $\mathbb{F}^*_{q^2}$; and then introducing further generators h_α for all α as well as additional generators required in order to take various powers of generators. Relatively little care is needed to obtain a presentation of length $O((\log|G|)^2)$. It is presently not known how to obtain a presentation for G of length $O(\log|G|)$. This is the only "bad" twisted case of rank $\ell > 2$: in all other cases all root groups are abelian, and the usual group H has at most 3 orbits on each root group. Similar considerations reappear as we examine rank 1 groups, but there they produce a much more serious obstacle.

Groups of Rank $\ell = 1$. The standard presentations in this case involve all of the elements of a field $\mathbb{F}_q$, but this time *it is not at all clear how to cut the presentations down to the desired size* -- assuming, of course, that these groups do, indeed, have presentations of the desired lengths! The easiest way to explain the problem is to give a suitably short presentation for the group PSL(2, q) (similar to [To]).

A Presentation for PSL(2, q).
Let $q = p^e$ with p odd, and $\kappa := (q - 1)/2$.
Generators:
$h, r, x_i, \quad i = 0,\dots,e - 1.$
Relations:
$h^\kappa = 1, r^2 = 1, h^r = h^{-1},$
$[x_0, x_i] = [x_1, x_i] = 1, x_0^p = x_1^p = 1, r = x_0 x_0^r x_0, h^{-1}r = x_1 x_1^{h^{-1}r} x_1,$
$x_{i+2} = x_i^h$ for $0 \le i \le e - 2$, $x_{e-2}^h = \Pi_k x_k^{a_k}, x_{e-1}^h = \Pi_k x_k^{b_k},$
with the $a_k, b_k\in\mathbb{F}_p$ dictated by a single irreducible polynomial over $\mathbb{F}_p$ used to define $\mathbb{F}_q$ (i. e., h corresponds to multiplication by the square of a generator θ of $\mathbb{F}_q^*$ and $\theta^e = \Sigma_k a_k\theta^k$, $\theta^{e+1} = \Sigma_k b_k\theta^k$), where powers $x_k^{a_k}$, $x_k^{b_k}$ are viewed as being expanded as in the Stupid-looking Example (i.e., by adjoining $O(\log p)$ additional generators and relators for each such term). The length of the resulting presentation for PSL(2,q) is $O(\log q)$.

In order to understand more clearly the preceding presentation, consider any rank 1 group $G = PSL(2, q)$, ${}^2A_2(q)$, ${}^2B_2(q)$ and ${}^2G_2(q)$. There is a standard presentation for G. Namely, let $B = UH$ be a Borel subgroup, where $H = \langle h \rangle$ is (isomorphic to) a subgroup of index 1, 2 or 3 in $\mathbb{F}_q^*$ ($\mathbb{F}_{q^2}^*$ in the unitary case) and hence $\kappa = |H|$ is explicitly known. Moreover, $N = H\langle r \rangle$ is dihedral except in the unitary case, where it has the presentation $\langle h, r \mid h^\kappa = 1, r^2 = 1, h^r = h^q \rangle$. Then a presentation for G is obtained by starting with ones for U and N, by giving the action of h on U, and finally by giving all the relations of the form $w = uvu'$ with $w\in\langle h \rangle r$, $u,u'\in U$, and $v\in U^r$ [St2].

When $G = PSL(2,q)$ we greatly decreased the number of generators by building in conjugation by h: there are at most 2 orbits of $\langle h \rangle$ on $U-\{1\}$, and we gave the action of h on a representative of each such orbit. Then all q of the relations $w = uvu'$ could be deduced from at most two of them, simply by conjugating by h. For each of the remaining rank 1 groups $\langle h \rangle$ has at least q orbits on $U-\{1\}$. The problem in those cases is to find some way to *deduce* most of these relations from a *bounded* number of them. Until some way is found to deal with this problem (or somehow circumvent it by using a different type of presentation for G), **2.1** will remain open: *the existence of short presentations of these rank* 1 *groups is the* ***only*** *obstacle to* **2.1**.

STEP II. Glueing.

Now consider any finite group G, and let N be a maximal normal subgroup of G. We may assume that $N \neq 1$. Then, by induction, there are presentations

$$G/N = \langle X \mid R \rangle \text{ and } N = \langle Y \mid S \rangle$$

each of which is suitably short (i.e., of respective lengths $O((\log|G/N|)^2)$ and $O((\log|N|)^3)$. The problem is to glue these together to form a new presentation that is itself sufficiently short. Glueing together the two presentations is standard, so once again it is necessary to find a way to proceed efficiently. This is less obvious and more interesting than Step I.

By abuse of language, view X as a subset of G and Y as a subset of N. Then R consists of elements of N, so each $r \in R$ is a word in $Y \cup Y^{-1}$. However, the presentations $\langle X \mid R \rangle$ and $\langle Y \mid S \rangle$ have nothing to do with one another, so there is no reason to expect that r will be a "nice" word in $Y \cup Y^{-1}$. In particular, it may have very large length as a word in $Y \cup Y^{-1}$, which would be unacceptable for our purposes. Fortunately, there is a way around this difficulty using the following surprising result:

2.4 Proposition [BS]. *Let* N *be a finite group and* Y *a set of generators of* N. *Let* $r \in N$. *Then there is a sequence* $w_1, \ldots, w_k = r$ *of elements of* N *such that*

each w_i *is either in* Y

or is the product of two previous w_j*'s*

or is the inverse of a previous w_j,

and $k < 2(\log|N| + 1)^2$.

Proof (based on remarks by E. M. Luks). We may assume that $N \neq 1$. We will construct a sequence A of elements of N and a subsequence $B \subseteq A$ such that the following all hold: each term in A is either in Y or is the product or inverse of terms occurring earlier in the sequence, $|A| < 2(\log|N|)^2$, $|B| \leq \log|N|$, and $N = \Pi(B)^{-1}\Pi(B)$. Here, for any sequence $B = (b_1, \ldots, b_k)$ of elements of N we write

$$\Pi(B) := \{ b_1^{\varepsilon(1)} \cdots b_k^{\varepsilon(k)} \mid \text{each } \varepsilon(i) = 0 \text{ or } 1 \}.$$

The construction of the sequences A and B will be accomplished by successive increasing approximations.

Start with A = B consisting of one element $\neq 1$ of Y (so initially $|\Pi(B)| = 2$). If, after several increases, we still have $N \neq \Pi(B)^{-1}\Pi(B)$, then $\Pi(B)^{-1}\Pi(B)Y \neq \Pi(B)^{-1}\Pi(B)$, so there exist $u, v \in \Pi(B)$ and $y \in Y$ such that $z := u^{-1}vy \notin \Pi(B)^{-1}\Pi(B)$. Then extend A and B to the following sequences by appending the indicated terms or sequences:

A': A, $\boxed{\text{viaB}}$, u, $\boxed{\text{viaB}}$, v, u^{-1}, $u^{-1}v$, z

B': B, z.

Here $\boxed{\text{viaB}}$ refers to the fact that a product such as $b_1 \cdots b_k$ (with $b_1, \ldots, b_k$ in B, in order) can be embedded in a sequence $b_1b_2, b_1b_2 \cdot b_3, \ldots, b_1 \cdots b_{k-1} \cdot b_k$ of $k - 1$ terms, each of which is a product of terms either in B or occurring earlier in this appended sequence.

Now observe that $|\Pi(B')| = 2|\Pi(B)|$ since $\Pi(B) \cap \Pi(B)z = \emptyset$, so that at most $\log|N| - 1$ increases can take place: $|B| \leq \log|N|$. Also, $|A'| \leq |A| + 2(|B| - 1) + 3$, where A is increased at most $\log|N| - 1$ times, so at the end of all the increases we have $|A| \leq 1 + (\log|N| - 1)(2\log|N| + 1)$. This completes the construction of the desired sequences A and B.

Finally, each element of B (in fact, of A) occurs in a sequence of the sort required in **2.4**, of length $\leq |A|$; and we saw that each element of $\Pi(B)$ occurs in such a sequence of length $\leq |A| + (|B| - 1)$. For the same reason, each element of $N = \Pi(B)^{-1}\Pi(B)$ occurs in such a sequence of length $\leq |A| + |B| + (|B| - 1) \leq 2(\log|N|)^2 + \log|N| - 1$. □

Note that this proof is short and ingenious while not looking at all like standard group theory. For somewhat sharper bounds and an effective version of **2.4**, see [BCFS].

Returning to the situation preceding the Proposition, adjoin a sequence using **2.4** in order to obtain additional generators for *each* $r \in R$, together with the relations implicit in the sequence. Similarly, adjoin further generators and relations in order to express the fact that $N \triangleleft G$. This readily produces a presentation of length $O((\log|G|)^5)$. Much more careful bookkeeping turns the exponent 5 into a 3 [BKLP].

3. THE PROBABILITY OF GENERATING

If G is a finite group generated by 2 elements, what proportion of the pairs of elements of G generate G? In other words, what is the probability that two randomly chosen elements of G generate G? This section will consider this question in the case of nearly simple groups. The most lovely result along these lines is due to Dixon:

3.1 Theorem [Di]. $\Pr(x, y \in S_n \text{ generate } S_n) \longrightarrow \frac{3}{4}$ *as* $n \longrightarrow \infty$,

$\Pr(x, y \in S_n \text{ generate } A_n) \longrightarrow \frac{1}{4}$ *as* $n \longrightarrow \infty$.

In other words, x and y "almost always" generate A_n or S_n, depending upon the parity of x and y. In order to show that

$$\Pr(\, x, y\in S_n \text{ do } \textit{not} \text{ generate } A_n \text{ or } S_n\,) \to 0 \text{ as } n\to\infty,$$

Dixon used two ingredients:

1. *Number Theory* ([Di], based on [ET]):

$$\Pr\left(\begin{matrix} x\in S_n \text{ has a cycle of length a prime} \le n-3 \text{ while}\\ \text{all other cycles have length relatively prime to this one}\end{matrix}\right)\to 1 \text{ as } n\to\infty.$$

2. 1873 *Group Theory* [Jo]:
If $G\le S_n$ is a primitive permutation group containing a p-cycle for some prime $p\le n-3$, then G is A_n or S_n.

(*Historical comment*: The preceding result appears to be the first published application of Sylow's Theorem, which had been published only a year earlier. It is, of course, only the conjugacy part required here -- in the case of Sylow subgroups of prime order.)

In Dixon's situation,

$$\begin{aligned}&\Pr(\, x, y \text{ do not generate } A_n \text{ or } S_n\,)\\ &\qquad\le \Pr(\, x, y \text{ generate a primitive group} \ne A_n, S_n\,) + \Sigma|L|^2/|G|^2\end{aligned}$$

summed over all subgroups L of S_n maximal with respect to being intransitive or imprimitive. (This is a very crude estimate: equality would require that the various subgroups be pairwise disjoint sets!) By **1**, if x and y are randomly chosen in S_n then each of them probably has a power that is a p-cycle for some prime $p\le n-3$, and then **2** implies that $\Pr(\, x, y$ generate a primitive group $\ne A_n, S_n\,)$ is negligible. The terms in $\Sigma|L|^2/|G|^2$ involve the orders of obvious subgroups $S_k\times S_{n-k}$ and S_k *wreath* S_ℓ of S_n. Estimating this sum is made slightly simpler by noting that an upper bound is $\Sigma(|L|^2/|G|^2)\cdot(|G|/|N_G(L)|)\le\Sigma(|L|^2/|G|^2)\cdot(|G|/|L|)=\Sigma|L|/|G|$ where L ranges over one representative $S_k\times S_{n-k}$ or S_k *wreath* S_ℓ from each conjugacy class of such subgroups. Thus, it was only necessary for Dixon to check that this latter sum $\to 0$ as $n\to\infty$.

Almost 20 years after Dixon's paper, Babai [Ba] showed that

$$\Pr(\, x, y \text{ do not generate } A_n \text{ or } S_n\,) = 1/n + O(1/n^2),$$

where the leading term 1/n corresponds to the fact that 2 elements not generating A_n or S_n "probably" have a common fixed point! However, in this case the proof no longer used **1** and **2** above: Babai used the classification of finite simple

groups. Dixon had shown that $\Sigma|L|/|G|$, summed over one representative $S_k\times S_{n-k}$ or S_k *wreath* S_ℓ from each conjugacy class, is $1/n + O(1/n^2)$. Consequently, it was only necessary to obtain a bound on $\Pr(\, x, y$ generate a primitive group $\ne A_n, S_n\,)$ significantly better than one obtained in [Di]. (Better bounds had been known [Bo; BoW], obtained using number theory and generating functions; but they are not quite good enough to produce the desired result.)

Babai's argument runs as follows. It is only necessary to estimate $\Sigma|L|/|G|$ summed over one representative of each subgroup $L\ne A_n, S_n$ of S_n maximal with respect to being primitive. The possibility **1.5**(i) occurs with miniscule probability and hence can be ignored. Then $|L|\le m:=n^{C\log n}$ by **1.5**. Let K be a minimal normal subgroup of L, so that $L=N_G(K)$. By the classification, there are at most m characteristically simple groups of order m. Any such group K has at most $m^{\log m}$ subgroups (since any group of order at most m is generated by at most log m elements, by Lagrange's Theorem), and hence has at most $m^{\log m}$ transitive permutation representations. Consequently, there are at most $m\cdot m^{\log m}$ transitive characteristically simple subgroups K of S_n, and hence the desired sum $\Sigma|L|/|G|$ is at most $m\cdot m\cdot m^{\log m}/n! = O(1/n^2)$ (the upper bound $n^{\sqrt{n}}/n!$ is obtained in [Ba]).

At the same time that Babai was making Dixon's theorem more precise, a result for classical groups corresponding to Dixon's was being proved:

3.2 Theorem [KaLu]. *Let* G_0 *denote a finite simple classical group, and let* $G_0\le G\le \mathrm{Aut}(G_0)$. *If* P(G) *is the probability that two randomly chosen elements of* G ***do not*** *generate a group containing* G_0, *then* $P(G)\to 0$ *as* $|G|\to\infty$.

The methods used in the proof were similar to those of Dixon and Babai. A theorem of Aschbacher [As] asserts that each maximal subgroup L of G falls into one of nine families of subgroups of G. Eight of the families are defined very explicitly in terms of the vector space V over $\mathbb{F}_q$ used to define G_0 (the stabilizer of a subspace; the stabilizer of a direct sum or a tensor decomposition; the stabilizer of a field extension or the centralizer of a field automorphism; a classical group embedded as usual; the normalizer of a symplectic-type r-group for a prime r other than the characteristic p of $\mathbb{F}_q$). In the ninth family, $L=N_G(S)$ with S a nonabelian simple subgroup of PSL(V) such that $S\le L\le \mathrm{Aut}(S)$, and the projective representation of S on V is absolutely irreducible and is defined over no proper subfield of $\mathbb{F}_q$.

The number of conjugacy classes within each of the eight explicit families is discussed in [As] (and in greater detail in [KlLi]), which makes it easy to obtain a suitable upper bound on $\Sigma|L|/|G|$ restricted to each such family. By [Li], $|L|\le q^{3n}$

for L in the ninth family, so that it is only necessary to show that there are not too many summands $|L|/|G|$ arising from this family -- exactly the same sort of question we saw Babai had to deal with. For this purpose, once again we will see that ridiculously crude estimates suffice.

Namely, as above, there are $\le q^{3n}$ simple groups of order $\le q^{3n}$. Fix such a simple group S. The number of (equivalence classes of) absolutely irreducible projective representations of S in characteristic p is at most the order of the universal cover of S. For each such representation, maximality forces L to be the normalizer of (the image of) S; and L is isomorphic to a subgroup of Aut(S) containing S, so that $|L| \le |S| \log |S|$. These crude estimates are enough to yield a proof of **3.2** when $n \ge 21$. Slightly more care is needed for the remaining small values of n.

An examination of the argument in [KaLu] gives slightly more information than in **3.2**. For purposes of the next result we assume, temporarily, that $PSp(2\ell, 2^e)$ with $\ell \ge 2$, $P\Omega(3, q)$, $P\Omega^+(6, q)$ and $P\Omega^-(6, q)$ are replaced by the respective isomorphic groups $P\Omega(2\ell + 1, 2^e)$, $PSL(2, q)$, $PSL(4, q)$ and $PSU(4, q)$. As above let V be the underlying vector space.

3.3 Theorem. *In the situation of* **3.2**, $P(G) = \Sigma |G_0{:}M|^{-1} + O(q^{-7(n-1)/6})$, *where* M *ranges over a representative of each* G_0*-conjugacy class of maximal subgroups of* G_0 *of each of the following types*:

(i) *The stabilizer of a point or hyperplane of* V;

(ii) *The image of a group* (i) *under a triality automorphism of* $G_0 = P\Omega^+(8, q)$;

(iii) *The stabilizer of a totally isotropic* 2*-space when* $G_0 = PSp(4, q)$ *or* $PSU(5, q)$, *or of a totally singular* 3*-space when* $G_0 = P\Omega(7, q)$.

For example, if $G_0 = PSL(n, q)$ then, in the unlikely event that $\langle g, h\rangle$ does not contain G_0, $\langle g, h\rangle$ "probably" fixes a point or hyperplane. It should be noted that the constant 7/6 best possible in **3.3**, as is seen when $G_0 = P\Omega(7, q)$ and M is either the stabilizer of a totally singular line or $G_2(q)$ (but for no infinite collection of pairs G_0, M disjoint from this one).

There is no doubt that, for any simple group G_0 and any group G such that $G_0 \le G \le Aut(G_0)$, the probability that two randomly chosen elements of G do not generate a group containing G_0 approaches 0 as $|G| \to \infty$. There is sufficient published information to prove this conjecture for various choices of G ($^2B_2(q)$, $^2G_2(q)$, $G_2(q)$, $^3D_4(q)$ or $E_6(q)$). Recent work [LS] reported in Seitz's lectures at this Symposium probably handles all the exceptional groups G_0 for characteristic p not too small (namely, $p > 113$).

4. WORD LENGTH.

While the preceding results say something about how *often* two elements generate a given group, they say nothing about *how* this generation takes place. In order to explain the difference, consider the following standard

Example. $S_n = \langle (1,...,n), (1,2) \rangle$. Write $S = \{ (1,...,n), (1,2) \}$. Then every element S_n has length $\le n^2$ in $S \cup S^{-1}$; but some elements have length $\ge n^2/6$ (e.g., the involution $z \to n + 1 - z$). (N.B. -- The length of each element of S_n does not seem to be known -- which is rather surprising in view of the standard nature of this pair S of generators.)

This leads to the consideration of the *diameter* of a group G with respect to a set S of generators of G. Temporarily write $T = S \cup S^{-1}$, and enumerate the elements of G as follows:

T	$|T|$ elements
TT	$\le |T|^2$ elements (actually, $\le |T|(|T| - 1)$ elements $\ne 1$)
$\vdots$	
$TT \cdots T$	$\le |T|^d$ elements.

The diameter is the smallest d such that these sets cover G, and in that case $|G| \le 1 + \Sigma_1^d |T|^i$ (or, more precisely, $|G| \le 1 + \Sigma_1^d |T|(|T| - 1)^i))$, so that $d \ge \frac{\log|G| - 1}{\log 2|S|}$. Note that d is the same as the diameter of the (undirected) Cayley graph determined by the pair G, T; and the preceding inequality is essentially the "Moore bound" for this graph.

For example, when $G = S_n$ and $|S| = 2$ we have $d \ge \frac{\log n! - 1}{2}$, which suggests that one might be able to do better than in the above Example. That this is, indeed, the case, is seen both in the next result and in **4.4**.

4.1 Theorem [BKL]. *If* G *is a nonabelian finite simple group then there is a set* S *of at most* 7 *generators of* G *such that the corresponding diameter is* $O(\log|G|)$ (better: the diameter is $\le 10^{10}\log|G|$).

Note that this result is false for cyclic groups. Namely, if G is cyclic and $G = \langle S\rangle$ then the diameter is easily seen to be greater than $\frac{1}{2}(n^{1/|S|} - 1)$ -- in fact, this holds for any abelian group. Thus, this is a particularly useless way to distinguish between nonabelian and abelian simple groups.

The theorem is constructive: a set S is more or less explicitly constructed; and, implicit in the proof, there is an *algorithm* which, given $g \in G$, will compute an expression for g as a word in S using $O(\log|G|)$ group operations. However, this is not the same question as determining an expression for g of shortest length in the generators (the S-*length* of g) -- nor even the exact length of g as a word in the generators.

Two generators. Of course, in **4.1** one naturally expects that there is a set S of 2 generators producing diameter $O(\log|G|)$. This has been verified in "most" cases (a few of these are discussed are in [Ka2]):

4.2 Theorem. (i) *If* G *is an alternating group, or a group of Lie type and rank* > 1, *then there is a set* S *of* 2 *generators of* G *such that the corresponding diameter is* $O(\log|G|)$.

(ii) *If* G *is an alternating group, or a group of Lie type and rank* ≥ 20, *then there is a set* S *of* 2 *generators of* G, *one having order* 2, *such that the corresponding diameter is* $O(\log|G|)$.

In (ii), the corresponding undirected graph is *trivalent*. The rank assumption is unfortunate, and in many instances the arguments sketched below can be modified so as to work in somewhat lower ranks (much lower when the characteristic is 2); see [Ka2] for examples of this. However, despite the more tractable appearance of the smaller rank cases, the general version of (ii) remains open and seems to require a less naive approach than will be presented below. In both (i) and (ii) there is an associated algorithm in the sense indicated previously.

Question 1: Clearly (i) is aimed at extending Steinberg's result [St1] that groups of Lie type have 2-element generating sets. Do Steinberg's 2 generators produce diameter $O(\log|G|)$? The proof in [St1] uses roughly ℓ = rank G commutations, therefore producing words of length $> 2^{\ell}$, which is too large. Note, however, that if ℓ is *bounded* and > 1, then Steinberg's proof shows that his generators do, indeed, produce diameter $O(\log|G|)$.

Question 2: Is **4.2**(ii) true for all $\ell \geq 2$? Presumably it is in all such cases, and also when $\ell = 1$. However, the latter is open even for **4.2**(i) even in the most familiar rank 1 instance:

Question 3: If G = PSL(2, q) with q *not* a prime, find a set S of 2 generators producing diameter O(log q). (For a set S of 3 generators producing this diameter see **4.3**.)

Question 4: Give a *constructive* proof that, when p is prime, PSL(2, p) has diameter O(log p) with respect to $\left\{ \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right\}$. The fact that the diameter is O(log p) -- in fact, $\leq 500\log p$ -- is due to Lubotzky and Sarnak (see [BKL, 8.1]). However, their proof is nonconstructive, using [We]. In order to see the difficulty inherent in this question, consider the much more restricted -- but also open -- question:

Write $\begin{pmatrix} 1 & \frac{1}{2}(p-1) \\ 0 & 1 \end{pmatrix}$ as a word of length O(log p) in the above generators.

Question 5: Prove that "most" S produce small diameter. For example, prove that $\Pr\left(\begin{matrix} x,y,z \in S_n, \langle x,y \rangle = S_n \\ z \text{ has length } O(\log n!) \text{ in } \{x,y,x^{-1},y^{-1}\} \end{matrix} \right) \longrightarrow 1$ as $n \longrightarrow \infty$.
On the other hand, all S ought to come close to working. For example, there is the following conjecture: if $S_n = \langle x,y \rangle$, then the corresponding diameter is $O(n^2)$.

Sketch of the parts of the proof of 4.2.
See **4.4** for the case of alternating groups. When G is classical we will replace it by the corresponding linear group, which will then also be called G. We will generally assume that q is odd, the even case being similar but simpler.

Example I. G = SL(2, q), q *odd*.
Write $x(t) := \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$ for $t \in \mathbb{F}_q$, $h(b) := \begin{pmatrix} b^{-1} & 0 \\ 0 & b \end{pmatrix}$ for $b \in \mathbb{F}_q^*$, and $r := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Then
$$x(t+u) = x(t)x(u) \text{ and } x(t)^{h(b)} = x(tb^2) \text{ for all } b \neq 0,\ t,\ u \in \mathbb{F}_q.$$

4.3 Proposition: (i) *If* q *is an odd prime then* G *has diameter* $O(\log|G|)$ *with respect to* $S := \{x(1), r'\}$, *where* $r' := h(\frac{1}{2})r$.

(ii) *If* q *is odd, and if* θ *generates* $\mathbb{F}_q^*$, *then* G *has diameter* $O(\log|G|)$ *with respect to* $S := \{x(1), r', h(\theta)\}$.

Proof. If ad - bc = 1 then a straightforward calculation yields that, for $c \neq 0$,
$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = x(-c^{-1} + ac^{-1})x(-c)^r x(-c^{-1} + dc^{-1}).$$
In case c = 0 use rg instead of g. This reduces the proof to showing that the S-length of each x(a), $a \in \mathbb{F}_q$, is O(log q) with respect to the given set S.

If q = p is an odd prime write $\theta = 2$; if q > p let θ be as in (ii). In either case, $\mathbb{F}_q = \mathbb{F}_p(\theta^2)$. Every element $t \in \mathbb{F}_q$ can be written in the form
$$t = \Sigma_0^m a_i\theta^{2i} = (\cdots(a_m\theta^2 + a_{m-1})\theta^2 + \cdots)\theta^2) + a_0$$
where either

$q = p$, $m + 1 \leq \frac{1}{2}\log q$, and $a_i \in \{0,1,2,3\}$ (base 4 representation of t), or
$q > p$, $m + 1 \leq \log_p q$, and $a_i \in \mathbb{F}_p$.

Suppose that $q = p$. Each $x(t)$ is a word $x(t) = (\cdots(x(a_m)^{h(2)}x(a_{m-1}))^{h(2)}\cdots)^{h(2)}x(a_0)$ in $m + 1$ elements $x(a_i)$ and $2m$ elements $h(\theta)^{\pm 1}$. Here, each $x(a_i) = x(1)^{a_i}$ has length ≤ 3, while (by matrix multiplication) $h(2)^{-1} = x(1)^{-2}(x(1)^2)^{r'}x(1)(x(1)^{-4})^{r'}$ has length ≤ 13. Thus, $x(t)$ has length $O(\log p)$, as required.

Suppose that $q > p$. As above, $x(t) = (\cdots(x(a_m)^{h(\theta)}x(a_{m-1}))^{h(\theta)}\cdots)^{h(\theta)}x(a_0)$, where we just saw that each $x(a)$, $a \in \mathbb{F}_p$, has S-length $O(\log p)$. Thus, each $x(t)$ has length $m \cdot O(\log p) = O(\log q)$. □

Remark. By crudely counting the lengths in the above arguments, it is easy to check that the diameters are $\leq 45\log|G|$ in (i) and $\leq 135\log|G|$ in (ii). Namita Sarawagi has observed that $h(2) = x(1)r'x(1)^4r'x(1)r'^{-1}$ has length ≤ 9, thereby improving these estimates.

Example **IIa**. $SL(n, q)$, q *odd.*
Let $s := r_{n-1}\cdots r_1$, so sH is an n-cycle within W. Let d_1 denote an involutory diagonal automorphism of G centralizing a hyperplane, normalizing H and L_{α_1}, and inverting X_{α_1}; write $d_{i+1} := d_1^{s^i}$.

If $g := r_1d_1 \cdot h_{\alpha_3}(2)r_3d_3 \cdot h_{\alpha_5}(2\theta)r_5d_5 \cdot d_7 \cdot x_{\alpha_9}(1)d_9$ *then* $S := \{s, g\}$ *behaves as required* (cf. [Ka2]). The point here, and in the other examples of **4.2**(ii) sketched below, is that g is chosen so that its eigenspaces and those of suitable shifts (conjugates by powers of s) will have very well-behaved overlaps: if $g' := gg^{s^2}$ then $[g'^{4s^{-1}}, g'^4]^{s^{-7}g'} = x_{\alpha_2}(a)$ with $a \in \mathbb{F}_p^*$. Then $x_{\alpha_3}(a)$ has length $O(1)$, while $x_{\alpha_3}(a)^{g'} = x_{\alpha_3}(4a)$. As in **4.3** we can use conjugation by g' in order to see first that all elements of $x_{\alpha_3}(\mathbb{F}_p)$ have length $O(\log p)$ and then that all elements of X_{α_5} have length $O(\log q)$; then so do all elements of X_{α_1} and $X_{-\alpha_1} = (X_{\alpha_1})^g$. As in **4.3** it follows that all elements of L_{α_1} have length $O(\log q)$, and then so do $z := sr_1$ and all elements of H_{α_1}. Note that $U \subset YY^s\cdots Y^{s^{n-1}}$ where $Y := X_{\alpha_1}X_{\alpha_1}^z\cdots X_{\alpha_1}^{z^{n-2}}$, and there are cancellations occurring in these products since $s^k(s^{k+1})^{-1} = s^{-1}$ and $z^k(z^{k+1})^{-1} = z^{-1}$. It follows that each element of Y has length $O(n \cdot \log q)$, so that each element of U has length $O(n \cdot n\log q)$. Each element of $H = H_{\alpha_1}H_{\alpha_1}^s\cdots H_{\alpha_1}^{s^{n-2}}$ also has length $O(n\log q)$. On the other hand, each element of $W = N/H$ has $\{r_1, s\}$-length $O(n^2)$. Then each element of N has S-length $O(n^2\log q) = O(\log|G|)$, and hence so does each element of $G = UNU$.

The proof is always easier when q is even because root elements have order $p = 2$ and hence are more readily accessible:

Example **IIb**. $SL(n, q)$, q *even*:
This time let $g := r_1 \cdot h_{\alpha_4}(\theta)r_4 \cdot x_{\alpha_7}(1)$ and $S := \{s, g\}$; again write $g' := gg^s$. Then $(g'^6)^{s^{-6}g} = x_{\alpha_2}(1)$, so that $gx_{\alpha_7}(1) = r_1 \cdot h_{\alpha_4}(\theta)r_4$ has length $O(1)$, as does $u := gx_{\alpha_7}(1)(gx_{\alpha_7}(1))^{s^3} = r_1 \cdot h_{\alpha_4}(\theta) \cdot h_{\alpha_7}(\theta)r_7$. Since $x_{\alpha_4}(b)^u = x_{\alpha_4}(b\theta^2)$ for all b, as above we find that all elements of X_{α_4} and $X_{-\alpha_4} = (X_{\alpha_4})^g$ have length $O(\log q)$. Now proceed as before.

Example **IIc**. $Sp(2\ell, q)$, q *odd.*
This time $s := r_\ell\cdots r_1$ induces a 2ℓ-cycle. The support V_{α_1} of L_{α_1} is a nonsingular 4-space of V. Let d_{α_1} denote an involution in G that normalizes L_{α_1}, induces the identity on $V_{\alpha_1}^\perp$ and inverts X_{α_1}; write $d_{i+1} := d_1^{s^i}$. If $\sigma := \alpha_{\ell-1}^{r_\ell}$ then $\alpha_\ell = (\sigma + \alpha_{\ell-1})^s$. Recall that $\ell \geq 16$ and write $\alpha := \sigma^{s^{14}}$ and $\beta := \alpha^{s^2}$; define d_α and $d_\beta := d_\alpha^{s^2}$ in the obvious manner. Note that $V_\alpha = V_{\alpha_{13}}$.

If $g := r_1d_1 \cdot h_{\alpha_3}(2)r_3d_3 \cdot h_{\alpha_5}(2\theta)r_5d_5 \cdot d_7 \cdot x_{\alpha_9}(1)d_9 \cdot r_\alpha d_\alpha$ *then* $S := \{s, g\}$ *behaves as required.* Once again $g' := gg^{s^2}$ satisfies $[g'^{4s^{-1}}, g'^4]^{s^{-7}g'} = x_{\alpha_2}(a)$ with $a \in \mathbb{F}_p^*$. As before we can use conjugation by s and g' in order to see that all elements of X_{α_5} and $X_{-\alpha_5}$ have length $O(\log q)$; and then (as in in **4.3**) so do $z := sr_1$ and all elements of H_{α_5}. Moreover, so do all elements of $(X_{\alpha_{12}})^{gr_{12}s^{-12}} = X_{-\alpha_\ell}$ and $(X_{\alpha_{14}})^{gr_{14}s^{-14}} = X_{\alpha_\ell}$; and then so do r_ℓ and all elements of H_{α_ℓ}. After suitably ordering the positive roots we find that $U \subset YY^s\cdots Y^{s^{2\ell-1}}X_{\alpha_\ell}X_{\alpha_\ell}^s\cdots X_{\alpha_\ell}^{s^{2\ell-1}}$ with $Y := X_{\alpha_1}X_{\alpha_1}^z\cdots X_{\alpha_1}^{z^{2\ell-3}}$. Also, H and N/H are easily handled exactly as in the previous Examples.

Example **IId**. $\Omega^-(2\ell + 2, q)$, q *odd.*
Define s as in Example **IIc**, as well as the support V_γ for every root γ of the B_ℓ root system for G. Let V_0 be the anisotropic 2-space $\langle V_\gamma \mid \gamma \text{ is long}\rangle^\perp$, and let j denote an involution in G that interchanges V_0 with a subspace of $V_{\alpha_{\ell-3}}$ while inducing the identity on $\langle V_0, V_0^j\rangle^\perp$; note that $|jj^{s^2}| = 3$. Let d_{α_1} be an involution in G that normalizes L_{α_1}, induces the identity on $V_{\alpha_1}^\perp$ and inverts X_{α_1}. Define σ, α, β, d_i, d_α and d_β as in Example **IIc**; once again $V_\alpha = V_{\alpha_{13}}$. Since $\ell \geq 20$, α and β are perpendicular to $\alpha_{\ell-3}$ and $\alpha_{\ell-1}$.

If $g := r_1d_1 \cdot h_{\alpha_3}(2)r_3d_3 \cdot h_{\alpha_5}(\theta)r_5d_5 \cdot d_7 \cdot x_{\alpha_9}(1)d_9 \cdot r_\alpha d_\alpha \cdot j$ *then* $S := \{s, g\}$ *behaves as required.* This time $g' := gg^{s^2}$ satisfies $[g'^{6s^{-1}}, g'^6] = x_{\alpha_8+\alpha_9}(36)$, so if $p \neq 3$ then we can proceed as in **4.3** in order to see that all elements of X_{α_5} and $X_{-\alpha_5}$ have length $O(\log q)$. Then so do all elements of $(X_{\alpha_{12}})^g$ and $(X_{\alpha_{14}})^g$. Conjugating by s^{-1} we find that, if $Y \cong \Omega^+(8, q)$ denotes the orthogonal group on $V_{\alpha_1} \perp V_{\alpha_3}$, then all the long root groups X_γ lying in Y have length $O(\log q)$. Using the usual method we see that all elements of Y have length $O(\log q)$; the same is then true for the orthogonal group $Y^{s^{-4}}$ on $V_{\alpha_{\ell-3}} \perp V_{\alpha_{\ell-1}}$. However, $Y^{s^{-4}g} = Y^{s^{-4}j}$ contains L_{α_ℓ}! Then r_ℓ and all elements of H_{α_ℓ} and X_{α_ℓ} have length $O(\log q)$, and hence we can proceed exactly as before (cf. [BKL]).

If $p = 3 < q$ write $u := g'^{2s^4}$ and $v := [g'^2, u] = x_{\alpha_9}(-\theta^2 + 1)$, note that $v^u = x_{\alpha_9}((-\theta^2 + 1)\theta^2)$, and obtain all of X_{α_9} as in **4.3**. If $q = 3$ then $[g'^4, g'^{3s^{-2}}] = x_{\alpha_9}(1)$. Now proceed as before. □

We conclude with a purely combinatorial argument.

4.4 Proposition. *There are trivalent Cayley graphs for* A_n *and* S_n *having diameter* $O(n\log n)$.

The following proof is motivated by an idea due to Quisquater ([Qu]; cf. [BHKLS]). My original approach was slightly more complicated, very similar in spirit to the partitioning method of [BKL] but using [Ka2]. The two generators constructed below have the added property that their orders are bounded -- 2 and 15 -- whereas one of those obtained as in [BKL] has order roughly $C\log n$.

Proof. Let $m \geq 4$, and consider the m-set $X = \{0,1,2,\ldots,m-1\}$; expressions such as x, $2x+1$, etc., are always assumed to refer to elements of X. Write

$$b_0 := \prod_{\substack{2^j \leq x < 2^{j+1} \\ j \text{ even}}} (x,2x,2x+1) \quad \text{and} \quad b_1 := \prod_{\substack{2^j \leq x < 2^{j+1} \\ j \text{ odd}}} (x,2x,2x+1) \quad \text{if } m \text{ is even,}$$

$$b_1 := \prod_{\substack{2^j \leq x < 2^{j+1} \\ j > 0 \text{ even}}} (x-1,2x-1,2x) \quad \text{and} \quad b_0 := \prod_{\substack{2^j \leq x < 2^{j+1} \\ j \text{ odd}}} (x-1,2x-1,2x) \quad \text{if } m \text{ is odd.}$$

Note that each product consists of pairwise commuting 3-cycles. In each case, $\langle b_0, b_1\rangle$ fixes 0 and b_1 fixes 1. If $x \in X$ and $x > 1$ then $b_i^{\pm 1}$ moves x to a smaller member of X (in fact, to a member $\leq \frac{1}{2}x$) for some i. Thus, $1 = x^w$ for a word w in $\{b_0, b_1\}$ of length $\leq \log m$. It follows that $\{(0,1), b_0, b_1\}$ generates S_m with diameter $O(m\log m)$ [Qu]. Namely, each transposition $(0,x) = (0,1)^{w^{-1}}$ has length $\leq 2\log m + 1$; and it is easy to see that each element of S_m has length $\leq m$ in these $m - 1$ transpositions.

Now consider an n-set, $n \geq 11$, which we may assume has the form $\{\infty, \infty', p\} \cup X \cup X'$ where $X' = \{x' \mid x \in X\}$, ∞ and ∞' are new symbols, and so is p if $n = 2m + 3$ is odd while $p = 0'$ if $n = 2m + 2$ is even. Let

$$t := \prod_x (x,x') \quad \text{or} \quad (\infty,\infty')\prod_x (x,x') \quad \text{depending on the parity desired, and}$$

$$g := (\infty',p,\infty,0,1')b_0b_1',$$

where, for example, $b_1' = b_1^t$ denotes the permutation of X' behaving as b_1 does on X. (In particular, b_0 fixes 0 while b_1' fixes $0'$ and $1'$.) We will show that $S := \{t, g\}$ *generates* S_n *with diameter* $O(n\log n)$.

Clearly $g^3 = (\infty,\infty',0,p,1')$, $g^{-5} = b_0b_1'$ and $(g^{-5})^t = b_0'b_1$. As seen above, for each $x \in X - \{0\}$ there is a word w of length $\leq \log m$ in $\{g^2, (g^2)^t\}$ fixing ∞, ∞', 0, p and sending $1'$ to x'. Thus, $(\infty,\infty',0,p,x')$ has S-length $O(\log n)$. Since $(\infty,\infty',0,p,1')^{-2}(\infty,\infty',0,p,x')(\infty,\infty',0,p,1') = (\infty,\infty',x')$ for $x > 1$, it follows that (∞,∞',u) has length $O(\log n)$ for each $u \in \{0,p\} \cup (X' - \{0'\})$. Now conjugate by t in order to see that (∞,∞',u) also has length $O(\log n)$ for each $u \in \{0'\} \cup (X - \{0\})$. Each element of A_n has length $\leq 2n$ in these 3-cycles, while parity can be adjusted if needed by using t. □

Postscript (January 31, 1991): There are certainly many further directions one can go in asymptotic group theory. The following very recent result, concerning the number $k(G)$ of conjugacy classes of a group G, uses the classification of finite simple groups in order to greatly improve estimates (essentially $k(G) > C\log\log |G|$) obtained by Landau and Brauer [Br] using only the class equation of G:

(Pyber [Py]) $k(G) > c\log |G|/(\log\log |G|)^8$ for some constant c.

References

[As] M. Aschbacher, On the maximal subgroups of the finite classical groups. Invent. Math. 76 (1984) 469-514.

[Ba] L. Babai, The probability of generating the symmetric group. J. Comb. Theory(A) 52 (1989) 148-153.

[BCFS] L. Babai, G. Cooperman, L. Finkelstein and Á. Seress, Nearly linear time algorithms for permutation groups with a small base, pp. 200-209 in *Proc.* 1991 *Int. Symp. Symbolic and Algebraic Computation.*

[BHKLS] L. Babai, G. Hetyei, W. M. Kantor, A. Lubotzky and Á. Seress, On the diameter of finite groups, pp. 857-865 in *Proc. 31st IEEE Symposium on Foundations of Computer Science* (1990).

[BKL] L. Babai, W. M. Kantor and A. Lubotzky, Small diameter Cayley graphs for finite simple groups. European J. Combinatorics 10 (1989) 507-522.

[BKLP] L. Babai, W. M. Kantor, E. M. Luks and P. P. Pálfy, Short presentations for simple groups (in preparation).

[BS] L. Babai and E. Szemerédi, On the complexity of matrix group problems, I, pp. 229-240 in *Proc. 25th IEEE Symposium on Foundations of Computer Science* (1984).

[Bo] J. D. Bovey, The probability that some power of a permutation has small degree. BLMS 12 (1980) 47-51.

[BoW] J. D. Bovey and A. Williamson, The probability of generating the symmetric group. BLMS 10 (1978) 91-96.

[Br] R. Brauer, Representation theory of finite groups, pp. 133-175 in *Lectures on Modern Mathematics* (ed. T. L. Saaty). Wiley, New York 1963.

[Ca] P. J. Cameron, Finite permutation groups and finite simple groups. BLMS 13 (1981) 1-22.

[CaNT] P. J. Cameron, P. M. Neumann and D. N. Teague, On the degrees of primitive permutation groups. Math. Z. 180 (1982) 141-149.

[Car] R. Carter, *Simple groups of Lie type*. Wiley, London-New York-Sydney-Toronto 1972.

[Cu] C. W. Curtis, Central extensions of groups of Lie type. J. reine angew. Math. 220 (1965) 174-185.

[Di] J. D. Dixon, The probability of generating the symmetric group. Math. Z. 110 (1969) 199-205.

[ET] P. Erdös and P. Turán, On some problems of a statistical group theory II. Acta Math. Acad. Sci. Hung. 18 (1967) 151-163.

[Gr] R. L. Griess, Schur multipliers of finite simple groups of Lie type. TAMS 183 (1973) 355-421.

[Hi] G. Higman, Enumerating p-groups, I: Inequalities. PLMS 10 (1960) 24-30.

[Ho] D. F. Holt, Enumerating perfect groups. JLMS 39 (1989) 67-78.

[Jo] C. Jordan, Sur la limite du degré des groupes primitifs non alternées. Bull. Soc. Math. France 1 (1873) 40-71.

[Ka1] W. M. Kantor, Some Cayley graphs for simple groups. *Proc. Conf. Combinatorics and Complexity*, Chicago 1987 = Discrete Applied Math. 254 (1989) 99-104.

[Ka2] W. M. Kantor, Some large trivalent graphs having small diameters (to appear).

[KaLu] W. M. Kantor and A. Lubotzky, The probability of generating a finite classical group. Geom. Ded. 36 (1990) 67-87.

[KlLi] P. B. Kleidman and M. W. Liebeck, *The subgroup structure of the finite classical groups*. LMS Lecture Note Series 129, Cambridge University Press 1990.

[Li] M. W. Liebeck, On the orders of maximal subgroups of the finite classical groups. PLMS 50 (1985) 426-446.

[LS] M. W. Liebeck and G. M. Seitz, Maximal subgroups of exceptional groups of Lie type, finite and algebraic. Geom. Ded. 35 (1990) 353-387.

[MN] A. McIver and P. M. Neumann, Enumerating finite groups. Quart. J. Math. 38 (1987) 473-488.

[Ne] P. M. Neumann, An enumeration theorem for finite groups. Quart. J. Math. 20 (1969) 395-401.

[Py] L. Pyber, Every finite group has many conjugacy classes, preprint, Math. Inst. Hung. Acad. Sci. (1990).

[Qu] J-J. Quisquater, Structures d'interconnexion: Constructions et applications, Ph. D. Thesis, Université de Paris - Sud (Orsay), July 1987.

[Si] C. C. Sims, Enumerating p-groups. PLMS 15 (1965) 151-166.

[St1] R. Steinberg, Generators for simple groups. Can. J. Math. 14 (1962) 277-283.

[St2] R. Steinberg, Generators, relations and coverings of algebraic groups, II. J. Algebra 71 (1981) 527-543.

[To] J. A. Todd, A second note on the linear fractional group. JLMS 11 (1936) 103-107.

[We] A. Weil, Sur les courbes algébriques et les variétés que s'en déduisent. Act. Sci. Ind. 1041 (1948).