On the diameter of finite groups

L. Babai
University of Chicago
and
Eötvös University, Budapest, Hungary

G. Hetyei
Mathematical Institute of the
Hungarian Academy of Sciences
and
M. I. T.

W.M. Kantor
University of Oregon, Eugene, OR

A. Lubotzky
Hebrew University, Jerusalem, Israel

Á. Seress
Ohio State University, Columbus, OH

Abstract

The diameter of a group G with respect to a set S of generators is the maximum over $g \in G$ of the length of the shortest word in $S \cup S^{-1}$ representing g. This concept arises in the contexts of efficient communication networks and Rubik's cube type puzzles. "Best" generators (giving minimum diameter while keeping the number of generators limited) are pertinent to networks, "worst" and "average" generators seem a more adequate model for puzzles. We survey a substantial body of recent work by the authors on these subjects. Regarding the "best" case, we show that while the structure of the group is essentially irrelevant if $|S|$ is allowed to exceed $(\log |G|)^{1+c}$ $(c > 0)$, it plays a heavy role when $|S| = O(1)$. In particular, every nonabelian finite simple group has a set of ≤ 7 generators giving logarithmic diameter. This cannot happen for groups with an abelian subgroup of bounded index. – Regarding the worst case, we are concerned primarily with permutation groups of degree n and obtain a tight $\exp((n \ln n)^{1/2}(1 + o(1)))$ upper bound. In the average case, the upper bound improves to $\exp((\ln n)^2(1 + o(1)))$. As a first step toward extending this result to simple groups other than A_n, we establish that almost every pair of elements of a classical simple group G generates G, a result previously proved by J. Dixon for A_n. In the limited space of this article, we try to illuminate some of the basic underlying techniques.

1 Introduction

The diameter of finite groups has been investigated in connection with efficient communication networks (cube-connected cycles, etc. [Sto], [PV] and others) and generalizations of Rubik's puzzles (cube, rings) ([DF], [McK]). Determining the exact diameter with respect to a given set of generators is NP-hard even in the case of elementary abelian 2-groups (every element has order 2) (Even, Goldreich [EG]). The length of the shortest positive word (no inverses) representing an element in terms of given generators of a permutation group is known to be $PSPACE$-complete (Jerrum, [Je]). The difficulty of the problem of determining the diameter is indicated by the gap between the best known upper and lower bounds for Rubik's cube (see [FMSST]).

In this paper we report some progress on estimating the best, worst, and average case diameters of various classes of finite groups. In connection with the average case problem, we consider the probability that a small set of random elements generates a given group.

The results were obtained by various subsets of the authors. The full proofs will appear in a number of separate papers [Ba2], [BH], [BKL2], [BS2], [Ka], [KL].

Let G be a finite group and S a set of generators of G. The *diameter* $\mathrm{diam}(G, S)$ is the maximum over $g \in G$ of the minimum word length expressing g as a product of elements of $S \cup S^{-1}$. Taking the position that we wish to minimize the diameter, we define the *worst case* diameter of G to be $\mathrm{diam}_{\max} G = \max_S \mathrm{diam}(G, S)$ (the generators are brought in by an adversary).

CH2925-6/90/0000/0857$01.00 © 1990 IEEE

We have to be a little careful when defining the best case: when referring to friendly generators (selected by ourselves), there should be a limit on how many generators we are entitled to use. We use $\mathrm{diam}_{\min}(G,k)$ to denote the minimum of $\mathrm{diam}(G,S)$ over all sets of k generators. (If k elements do not suffice to generate G then this quantity is infinite.)

Similarly in the average case: we consider bounds on $\mathrm{diam}(G,S)$ which hold for *almost every choice* of the k-set $S \subset G$. (This is not to be confused with the *average diameter* over the same collection of Cayley graphs; we do not invesigate this quantity here.)

In this connection it is of interest to find out what the minimum number of generators is, and to find the tradeoff between the size of a random subset S of a group and the probability that S generates G.

We note that in the network context one is most interested in the best case; in the puzzle context, the worst and average cases seem more relevant. Families of expanders are relevant for the best case; we shall comment on this connection after the statement of Theorem 2.1.

2 Statement of the main results

The *Cayley graph* $\Gamma(G,S)$ of the group G with respect to the set S of generators has G for its vertex set; the edges are the pairs $\{g,gs\}$ ($g \in G$, $s \in S$). The diameter of this graph is $\mathrm{diam}(G,S)$. Note that $\mathrm{diam}(G,S) = \mathrm{diam}(G, S \cup S^{-1})$.

For obvious counting reasons, the diameter is at least logarithmic as a function of the order of the group:

$$\mathrm{diam}_{\min}(G,k) > \log|G|/\log(2k). \qquad (1)$$

This bound is tight, up to a constant factor, for every group G, assuming $k > (\log|G|)^{1+c}$ for some positive constant c (Prop. 3.2). On the other hand, the ratio of the two sides of inequality (1) may diverge while $|G| \to \infty$ if $k \le (\log|G|)^{1+o(1)}$ (Prop. 3.3).

The greatest interest is in the case of groups generated by a bounded number of elements. It is natural to begin the "best case" study of these groups with the finite simple groups. These groups turn out to behave quite favorably.

By a result essentially due to R. Steinberg, every finite simple group is generated by two elements ([Ste]). We are unable to find two generators which would produce optimal ($O(\log|G|)$) diameter, but 7 generators suffice for this purpose:

Theorem 2.1 [BKL2] *Every nonabelian finite simple group G has a set S of at most 7 generators such that the resulting diameter is $O(\log|G|)$.*

In other words, $\mathrm{diam}_{\min}(G,7) = O(\log|G|)$. This result is *algorithmically efficient;* given an element of G, the corresponding word can be found using $O(\log|G|)$ group operations. (Here we assume that the elements of G are given in their "natural" matrix representation.)

A comment on *expanders* is in order here. For simplicity, we use the term for the not necessarily bipartite version: an ϵ-expander is a graph in which the boundary of every set A of at most half of the vertices has size $\ge \epsilon|A|$. A family of expanders is a family of ϵ-expanders for a fixed $\epsilon > 0$.

Expanders on n vertices have diameter $O(\log n)$. The current champion of bounded degree explicit expanders arises as a family of Cayley graphs of the simple groups $PSL(2,q)$ (Margulis [Ma], Lubotzky - Phillips - Sarnak [LPS]). Some other families of finite simple groups have also been known to give rise to families of expanders (notably, $PSL(d,q)$ for fixed dimension $d \ge 3$, see Alon - Milman [AM]). It is not known, however, whether or not all finite simple groups admit bounded degree expander Cayley graphs. Even the seemingly most accessible cases are open.

Problem 2.2. Do the alternating groups and the linear groups $PSL(d,q)$ for fixed q admit families of bounded degree expander Cayley graphs?

We should also point out that the known constructions of explicit expanders do not yield the algorithmic consequence stated after Theorem 2.1. (Observe that in comparison to the size of the Cayley graphs constructed, we find short paths in *logarithmic* number of group operations.)

The number 7 in Theorem 2.1 can probably be improved. The following result exists in this direction.

Theorem 2.3 [Ka] *If $n \ge 10$ then there exist two generators of $G = PSL(n,q)$ such that the diameter is $O(\log|G|)$. One of the two generators may be required to be an involution (i.e. an element of order two).*

(This reduces the degree of the Cayley graph, i.e. the size of the set $S \cup S^{-1}$ to 3 which is minimum.) The same result holds for the other classes of classical simple (matrix) groups of sufficiently large dimension, as well as for the alternating groups [Ka].

We suspect that finite simple groups behave quite nicely even in the worst case. More precisely, the following conjecture has been made.

Conjecture 2.4 [BS1] For every finite simple group G, $\mathrm{diam}_{\max}(G) = (\log |G|)^{O(1)}$.

Unfortunately we are unable to verify this conjecture even in the case of the alternating groups A_n of order $n!/2$. The only result in this direction gives a *moderately exponential* upper bound:

Theorem 2.5 [BS1] $\mathrm{diam}_{\max}(A_n) < \exp(\sqrt{(n \ln n)}(1 + o(1))$.

Conjecture 2.4 would require an upper bound, polynomial in n. Such upper bounds have been verified for special classes of generators only (Driscoll-Furst [DF], McKenzie [McK]).

The upper bound stated for the alternating groups in Theorem 2.5 actually holds for all permutation groups of degree n:

Theorem 2.6 [BS2] *If $G \leq S_n$ is a permutation group of degree n then* $\mathrm{diam}_{\max}(G) < \exp(\sqrt{(n \ln n)}(1 + o(1))$.

This result is tight, as shown by the cyclic group generated by the product of as many cycles of different prime lengths as will fit in a set of n elements. A potentially much stronger bound holds for transitive permutation groups (cf. the Conjecture above):

Theorem 2.7 [BS2] *If $G \leq S_n$ is a transitive permutation group of degree n then* $\mathrm{diam}_{\max}(G) < \exp(c(\log n)^3)\mathrm{diam}_{\max}(A_n)$.

For the average case analysis we first quote a remarkable result of J. Dixon [Di]:

Theorem 2.8 (Dixon) *A randomly selected pair of permutations from S_n almost always generates A_n or S_n.*

("Almost always" refers to having probability approaching 1 as $n \to \infty$. All the $(n!)^2$ pairs of permutations have equal probability of selection.) The speed of convergence is also known:

Theorem 2.9 [Bal] *The probability that a randomly selected pair of permutations from S_n generates A_n or S_n is $1 - 1/n + O(n^{-2})$.*

While Dixon's proof is completely elementary, unfortunately the proof of this result, like the proofs of nearly all results announced here, depends on the classification of finite simple groups. (Theorems 2.5 and 2.10 are the only exceptions.) The best error-term obtained previously using generating function techniques only was $n^{-1+o(1)}$ (Bovey [Bo]).

We are able to prove that a random pair of permutations not only generates A_n or S_n, but that it generates them fairly efficiently:

Theorem 2.10 [BH] *Almost every pair of permutations generates A_n or S_n with diameter $< n^{\ln n(1/2+o(1))}$.*

As a first step toward extending this result to simple groups other than A_n, we present an analogue of Dixon's theorem for classical simple (matrix) groups. These groups are the projective linear, symplectic, orthogonal, and unitary groups over finite fields.

Theorem 2.11 [KL] *If G_0 is a classical simple group and $G_0 \leq G \leq \mathrm{Aut}(G_0)$ then almost every pair of elements of G generates a subgroup containing G_0.*

There is a hope that this result can be extended to every class of finite simple groups. We mention that Theorem 2.11 is employed in [BKL1] to construct a bounded-round interactive protocol for nonisomorphism of permutation groups.

3 General remarks

First we comment on the best case when the number of generators is allowed to grow with the size of the group. If we allow a logarithmic number of generators, then favorable sets of generators always exist; allowing a little more earns us optimal sets of generators (matching the lower bound (1)).

These statements follow from a result of Erdős and Rényi [ER] of which we quote a special case here.

Theorem 3.1 [ER] *In any finite group G there exists a sequence $g_1, \ldots, g_t$ of elements such that*

(i) $t = \lfloor \log |G| + \log\log |G| \rfloor + 2$;

(ii) every element of G can be represented as a product of the form $g_1^{\epsilon_1} \cdots g_t^{\epsilon_t}$, where $\epsilon_i \in \{0,1\}$ for every i.

(For a half-page proof, see [BE].) "log" stands for base 2 logarithms; obviously, for (ii) to hold, $t \geq \log |G|$ is a necessary condition. Erdős and Rényi also prove that if the term $+2$ in (i) is replaced by $\omega_G \to \infty$ arbitrarily slowly, then almost every choice of $g_1, \ldots, g_t$ will satisfy (ii).

It follows that G has a set of t generators, t as under (i), such that the resulting Cayley graph has diameter $\leq t$. This upper bound misses the trivial lower bound (1) by a $\log \log |G|$ factor only, and with slightly increased t, holds for almost every set of size t.

Next we show that some further growth of the set of generators makes inequality (1) tight.

Proposition 3.2. *For any constant $c > 0$, if $k \geq (\log |G|)^{1+c}$, then*

$$\operatorname{diam}_{\min}(G,k) = \Theta(\log |G| / \log k). \qquad (2)$$

Proof. Divide the sequence $g_1, \ldots, g_t$ (defined in Theorem 3.1) into segments of length s (the last segment may be shorter). Consider all the 2^s subproducts of each segment. Let S be the set of these $\leq 2^s t/s$ subproducts. Clearly $\operatorname{diam}(G,S) \leq t/s + 1$. Choosing the greatest s such that $2^s t/s < k$, the upper bound follows. The lower bound is from (1). □

In contrast, $k = (\log |G|)^{1+o(1)}$ does not suffice for inequality (1) to give always the right order of magnitude.

Proposition 3.3. *For $G = \mathbf{Z}_2^m$ and $k \geq m$ let $d = \operatorname{diam}_{\min}(G,k)$. Then $d \log(ek/d) > m$. In particular, if $\epsilon > 1/\log m$ and $k = (\log m)^{1+\epsilon}$ then*

$$\operatorname{diam}_{\min}(G,k) = \Omega(\log |G| / \epsilon \log k). \qquad (3)$$

Proof. $2^m \leq \sum_{i=0}^{d} \binom{k}{i} < (ek/d)^d$. □

Theorem 3.1 and Proposition 3.2 show that for superlogarithmic degrees, the behavior of the $\operatorname{diam}_{\min}$ function hardly depends on the group stucture. This is in sharp contrast with the case of sublogarithmic degrees and in particular with the case of constant degree.

One of our main results (Section 4) is that nonabelian simple groups behave favorably: they have bounded generating sets with respect to appropriate bounded size sets of generators.

On the other hand, abelian groups, and more generally nilpotent groups of bounded class cannot have logarithmic diameter with respect to bounded generating sets. (For the definition of nilpotent groups we refer to texts in group theory, such as [Gor]. We note that the abelian groups are exactly the nilpotent groups of class 1; and every nilpotent group is solvable but not vice versa. Among nilpotent group, the "class" parameter may serve as a measure of "nonabelianness".)

The following result is essentially known.

Lemma 3.4 [Wo], [Bass], [AB] *If G is a group with a subgroup H of index r which is nilpotent of class ℓ and S is a set of k generators of G then the number of elements of G representable as words of length $\leq d$ in S is at most*

$$rd^{2(kr\ell)^{\ell}}.$$

(It follows from the work of Wolf [Wo] and Bass [Bass] that an upper bound of the form $d^{c(k,r,\ell)}$ exists. The calculation for $r = 1$, using "commutator collection", was carried out by Annexstein and Baumslag [AB]. The case of general r follows by using Schreier generators.)

It is now immediate that the class of groups described in the Lemma has best diameter $\Omega(|G|^c)$ for some fixed $c > 0$. More specifically:

Corollary 3.5 [AB] *If G is a group with a subgroup H of index r which is nilpotent of class ℓ then*

$$\operatorname{diam}_{\min}(G,k) \geq (|G|/r)^{(kr\ell)^{-\ell}/2}. \qquad (4)$$

The exact relationship between group structure and best diameter is far from clear. While it seems that being "far from abelian" may help reduce the diameter, there are examples which by any standard are quite close to abelian and still behave nicely. The prime example is the "cube-connected cycles" group of order $k2^k$ which is solvable and for k itself a power of 2 it is even nilpotent. It has an abelian subgroup of very small index (the index is $k = O(\log |G|)$). Yet this group has two generators yielding logarithmic diameter.

4 Outline of the proof of the "best case" results

In this section we illustrate some of the flavor of the proofs of Theorems 2.1 and 2.3.

Apart from the 26 sporadic simple groups, the nonabelian finite simple groups fall in the following two categories: the alternating groups, and the finite simple groups of Lie type. The linear groups $PSL(n,q)$ form a subclass of the Lie type simple groups, and they are typical in many ways for the entire class. This resemblance makes it possible to illustrate the main ideas involved in the proof of Theorem 2.1 on the examples of alternating and linear groups.

Instead of A_n, we consider its close relative S_n. We remark that the familiar bubblesort generators (a transposition and an n-cycle) give rise to diameter $\Theta(n^2)$, as opposed to the optimal $O(n!) = O(n \log n)$ we require.

Next, *we construct 3 generators of S_n, producing diameter $O(n \log n)$.*

Proof. Assume n is even. (A slight modification of the argument solves the odd case.) Let S_n act on the n-set $X = \mathbf{Z}_{n-1} \cup \{\infty\}$. Let $\alpha_0 : x \mapsto 2x$ and $\alpha_1 : x \mapsto 2x+1$ be two permutations of X (both of them fix ∞). Starting from the point $0 \in X$, one can reach any positive integer $t \leq n-1$ by applying a word w_t of length $< \log n$ in α_0 and α_1. This follows by applying "Horner's rule" to the binary expansion of t:

$$t = \sum_{i=0}^{m} a_i 2^i = (\cdots(a_m \cdot 2 + a_{m-1})2 + \cdots)2 + a_0. \quad (5)$$

Then $w_t = \alpha_{a_m} \cdots \alpha_{a_0}$ is such a word. Let γ_t denote the transposition (t, ∞) $(t \in \mathbf{Z}_{n-1})$. Now $\gamma_t = w_t^{-1} \gamma_0 w_t$. Since every permutation can be written as a word of length $< 2n$ in the transpositions γ_t, it follows that the diameter resulting from the three generators $\alpha_0, \alpha_1, \gamma_0$ is $< 2n(2\log n + 1)$. □

J. J. Quisquater informed us that earlier he had found another set of equally efficient generators which are close relatives of the ones just constructed. We describe Quisquater's generators for even n; again, the modification for odd n is easy. Let now the set X be $X = \{0, 1, \ldots, n-1\}$, and the 3 permutations: β_0 is the product of all 3-cycles $(x, 2x, 2x+1)$ where $2^j \leq x < 2^{j+1}$, for all even values of j $(1 \leq x \leq (n/2)-1)$. The permutation β_1 is the product of all three-cycles of the same form for all odd values of j. Let δ_t be the transposition $(0, t)$. Like before, a word of length $< \log n$ in β_0 and β_1 takes 1 to any t $(1 \leq t \leq n-1)$, and conjugating δ_1 by such a word yields the transposition δ_t. The conclusion is the same as above. One additional feature of Quisquater's generators is that their number can be reduced to 2 in a very simple way, just observing that δ_1 and β_1 commute. It follows that $\delta_1 = (\delta_1\beta_1)^3$ and $\beta_1 = (\delta_1\beta_1)^{-2}$, hence $S_n =< \beta_0, \delta_1\beta_1 >$ and S_n has diameter $< 6n(2\log n + 1)$ with respect to this pair of generators [Qu].

Rather than handling the *projective* linear groups $PSL(d,q)$, we shall consider the groups $SL(d,q)$ consisting of those $n \times n$ matrices over the field $\mathbf{F}_q$ of order q (q a prime power) with determinant 1. "Good" generators for $SL(d,q)$ are good for its simple factor group $PSL(d,q)$ as well.

First we treat the case $d = 2$. For this case, [Ma] and [LPS] independently arrived at the same family of rapidly expanding Cayley graphs (Ramanujan graphs). Although these graphs clearly have logarithmic diameter, no algorithm is known to actually find logarithmic length paths in reasonable (polylog) amount of time.

We present a simple algorithmic proof that another set of *3 generators of $SL(2,q)$ gives diameter $O(\log q)$*. (The order of $SL(2,q)$ is $q(q^2-1)$.)

Proof. Write

$$x(t) = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}; \quad h(b) = \begin{pmatrix} b^{-1} & 0 \\ 0 & b \end{pmatrix};$$

and

$$r = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

for $b \neq 0, t \in \mathbf{F}_q$. Then

$$x(t+u) = x(t)x(u) \quad \text{and} \quad h(b)^{-1}x(t)h(b) = x(tb^2) \quad (6)$$

for all $b \neq 0, t, u \in \mathbf{F}_q$. If $ad - bc = 1$ then a straightforward calculation shows that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = x(-c^{-1}+ac^{-1})r^{-1}x(-c)rx(-c^{-1}+dc^{-1}).$$

(In the case $c = 0$ one can find a simpler expression, or multiply our matrix by r first.) This reduces the proof to showing that the length of each $x(a)$, $a \in \mathbf{F}_q$, is small with respect to a fixed set S of ≤ 3 generators.

In the case $q = p$, odd prime, we take $S = \{x(1), h(1/2)r\}$.

In the case $q = p^k$ is odd, $k \geq 2$, we take $S = \{x(1), h(1/2)r, h(\theta)\}$, where θ is a generator of $\mathbf{F}_q$ over its prime field $\mathbf{F}_p$.

In the case q is even, we take $S = \{x(1), r, h(\theta)\}$.

We indicate the proof in the case $q = p$. In this case we write t in base 4 and use Horner's rule (5) for this case. Using identities (6), one can turn the Horner expansion into a word of length $O(\log p)$ in the generating set S, expressing $x(t)$. $\square$

We omit the general case from this survey.

5 Outline of the proof of the "worst case" results

A *normal series* of a group G is a chain of subgroups

$$G = G_0 \geq G_1 \geq \ldots \geq G_m = \{1\}, \qquad (7)$$

where each G_i is a normal subgroup of G.

For large n, the permutation groups $G \leq S_n$ have normal series with remarkable properties. The construction of such normal series is a central part of the proof of Theorem 2.6. The overall structure is inspired by the "augmented structure forests", introduced in E. M. Luks's seminal paper on parallel algorithms for permutation groups [Lu], cf. [BLS]. Along the way, we make use of several consequences of the classification of finite simple groups and in particular of the classification of those primitive permutation groups of degree n whose order is divisible by some prime $p > \sqrt{n}$ [LS].

For primes $q < p$, where $q|p-1$, let $H(p,q)$ denote the (unique) nonabelian group of order pq.

Direct products of simple groups are called *semisimple*. A group is *characteristically simple* if it is the direct product of isomorphic simple groups. A subgroup $G \leq H_1 \times \cdots \times H_k$ is a *subdirect product* of the H_i if G projects *onto* each factor H_i.

Let $G = G_0 \geq G_1 \cdots \geq G_m = 1$, $G_i \triangleleft G$ be a normal series of the group G. We refer to the factors G_{i-1}/G_i as the *levels* of the series. We call the level G_{i-1}/G_i an *alternating level* if all composition factors of G_{i-1}/G_i are alternating groups. On a *small level*, all composition factors are nonabelian, nonalternating. (The reason for this terminology is that nonalternating simple permutation groups of degree n have fairly small order: $\exp(c\log_2^2 n)$ [Cam].) G_{i-1}/G_i is an *abelian level* if G_{i-1}/G_i is abelian. Finally, a *metacyclic level* is the subdirect product of cyclic groups of prime order and metacyclic groups of the form $H(p,q)$ (see above). A normal series is *organized* if (i) it has at most one metacyclic level; (ii) all the other levels are semisimple; (iii) each level is either metacyclic or abelian or alternating or small.

Let $G \leq S_n$. A ***giant level*** is an alternating level involving an alternating group of degree $> \sqrt{n}$. A normal series is ***well organized*** if it is organized and has at most one giant level.

We define the ***multiplicity free part*** of the integer $p_1^{\alpha_1}\cdots p_k^{\alpha_k}$ as $\nu(p_1^{\alpha_1}\cdots p_k^{\alpha_k}) = p_1 \cdots p_k$ where the p_i are all distinct primes.

Theorem 5.1 [BS2] *Let $G \leq S_n$. Then G has a well organized normal series $G = G_0 \geq G_1 \geq \cdots \geq G_m = 1$, $G_i \triangleleft G$ with the following properties.*

(i) $m = O(\log^3 n)$.

(ii) Let s_i be the multiplicity free part of the order of G_{i-1}/G_i. If G_{i-1}/G_i is abelian then $s_i \leq \exp(n^{0.4}(1+o(1)))$.

(iii) If G_{i-1}/G_i is metacyclic then all primes r dividing its order satisfy $n^{0.4} < r < n^{0.7}$.

Remark 5.2. Instead of the exponents 0.4 and 0.7 we could have chosen any numbers $a < \frac{1}{2} < b$ satisfying $a > \frac{1}{2}b$ and $a + b > 1$.

The case of transitive groups is much simpler.

Theorem 5.3 [BS2] *Let $G \leq S_n$ be transitive. Then G has a normal series $G = G_0 \geq G_1 \geq \cdots \geq G_m = 1$, $G_i \triangleleft G$ with the following properties.*

(i) $m = O(\log^3 n)$.

(ii) Each factor G_{i-1}/G_i is characteristically simple.

(iii) At most one of the levels G_{i-1}/G_i is giant, i.e., the product of alternating groups of degree $> \sqrt{n}$.

Note that the normal series described in Theorem 5.3 is well organized. The estimate in part (ii) of Theorem 5.1 is replaced by the trivial bound $s_i \leq n$.

6 Metacyclic groups

Recall from the previous section that when p, r are primes and $r|p-1$ then there exists a unique nonabelian group $H(p,r)$ of order pr. It has a cyclic normal subgroup of order p; the corresponding factor group is cyclic of order r. This seemingly simplest of nonabelian groups offers a number of interesting cases in the study of the diameter, with a surprising array of relevant mathematical tools and large regions of unsolved cases. Here is a summary of some of the results.

We observe that these groups have 2 generators and every non-redundant set of generators has two elements. The notation $\mathrm{diam}_{\min}$ will in this section always refer to *pairs* of generators.

Theorem 6.1 [Ba2]. *If r is bounded while $p \to \infty$ then both $\mathrm{diam}_{\max}(H(p,r))$ and $\mathrm{diam}_{\min}(H(p,r))$ have the order of $\Theta(p^{1/(r-1)})$.*

This unison disappears when r becomes large.

Theorem 6.2 [Ba2]. *If $r > p^{1/2+c}$ for some constant $c > 0$ then*

(i) $\mathrm{diam}_{\min}(H(p,r)) = \Theta(r^{1/2})$.

(ii) $\mathrm{diam}_{\max}(H(p,r)) = \Theta(r)$.

The proof of Theorem 6.1 is elementary, based on a resultant argument involving the irreducibility of the r^{th} cyclotomic polynomial. Theorem 6.2 on the other hand rests on Weil's estimate on the number of solutions of the diagonal equation over finite fields which we quote below ([We]), cf. [Jo, p. 57]).

Theorem 6.3 (A. Weil) *The number of solutions $N(b)$ of the diagonal equation*

$$x_1^k + \cdots + x_t^k = b \tag{8}$$

over the finite field $\mathbf{F}_q$, where $b \neq 0$, satisfies the inequality

$$|N(b) - q^{t-1}| \leq (k-1)^t q^{(t-1)/2}. \tag{9}$$

The connection is made through an elementary reduction of the diameter of $H(p,r)$ to the diameter of the r^{th} cyclotomic graph mod p, defined on the vertex set $\mathbf{F}_p$ by the rule that i and j are adjacent if $(i-j)^r = \pm 1$ in $\mathbf{F}_p$. This is a Cayley graph of $\mathbf{Z}_p$ and it is easy to see that its diameter is in close connection with the solvability of the diagonal equation (8) with $q = p$ and $k = (p-1)/r$. Finding the diameter of the directed version of this graph (when ± 1 is replaced by 1 in the definition) is also referred to as the Waring problem mod p and has extensive literature.

7 Average case

We describe some of the prerequisites of the proof of Theorem 2.10. The required information on the cycle structure of most permutations is provided in a series of papers by Erdős and Turán.

Theorem 7.1 [ET1] *Let $g(\pi)$ denote the number of cycles of $\pi \in S_n$. If $\omega_n \to \infty$ arbitrarily slowly, then for almost all $\pi \in S_n$ we have*

$$|g(\pi) - \ln n| \leq \omega_n \cdot \sqrt{\ln n}. \tag{10}$$

Theorem 7.2 [ET2] *Let $1 \leq a_1 < a_2 < \ldots < a_s \leq n$ be a sequence of integers. Then the number of those $\pi \in S_n$ having no cycles of length a_i for any i cannot exceed*

$$\frac{n!}{\sum_{\nu=1}^{s} 1/a_\nu}. \tag{11}$$

Just as in the worst case results (Theorems 2.5 and 2.6), the bottleneck in obtaining a better bound is the order of the permutations involved. While the maximum order of a permutation is $\exp(\sqrt{n \ln n}(1 + o(1))$, explaining the bound obtained in those theorems, the next result explains the main bottleneck in the average case.

Theorem 7.2 [ET1], [Gon] *For arbitrary $\epsilon > 0$, the order of almost all permutations is between the bounds* $\exp((1/2 \pm \epsilon)(\ln n)^2)$.

The *support* of a permutation is the set of elements actually displaced. The size of the support is the *degree* of the permutation. Our strategy is to construct permutations of small degree. Eventually we obtain 3-cycles, and rapidly build any (even) permutation.

The two operations we employ to achieve small degree is taking large powers and commutators. The following lemma helps calibrate the degree of a power of a random permutation.

Lemma 7.4. *Given $0 < r < 1$, almost every $\sigma \in S_n$ has a power of degree $n^{r+o(1)}$.*

This result is one of the tools in establishing the following lemma, from which the proof of Theorem 2.10 is essentially immediate.

Lemma 7.5. *For almost every $\sigma, \pi \in S_n$ there exist powers σ^s and π^t such that the commutator of σ^s and its conjugate $\pi^t \sigma^s \pi^{-t}$ is a 3-cycle.*

What we have to show in effect is that the support of σ^s and its image under π^t share precisely one element for appropriately chosen s and t. We choose s such that the degree of σ^s be about $n^{0.05}$ and regard its support M as a random subset of this size. Now M has two elements, a and b, on the same π-cycle (since π has a logarithmic number of cycles only). Let t be the smallest positive integer such that π^t maps a to b. We claim that with large probability, $\pi^t(M) \cap M = \{b\}$. The proof involves case distinctions and counting.

8 Randomly selected generators

The basic principle of the proofs of Theorems 2.9 and 2.11 is similar.

Proposition 8.1. *Let G be a group, and M a subgroup of G. The probability that k randomly chosen elements of G generate a subgroup contained in some conjugate of M is $\leq |G : M|^{1-k}$.*

Indeed, the number of conjugates of M is $|G : N_G(M)| \leq |G : M|$, and the probability that each of the k elements selected belongs to a particular one of these conjugates is $|G : M|^{-k}$. □

Corollary 8.2. *Let G be a group. The probability that k randomly chosen elements generate a proper subgroup of G is $< \sum_M |G : M|^{1-k}$, where the summation is extended over representatives of the conjugacy classes of the maximal subgroups of G.* □

We can apply this estimate directly when G is simple ($G = G_0$ in Theorem 2.11). In most cases one can prove that M either has a specific structure (such as an imprimitive permutation group in the case $G = A_m$), or has large index. The work then consists of showing that there are not too many conjugacy classes of the latter kind; and of a detailed analysis of the cases of the first kind. Aschbacher's work on maximal subgroups is the relevant tool for groups of Lie type [Asch]; the case of the alternating group rests on consequences of the O'Nan-Scott Theorem (cf. [Cam]).

References

[AB] Annexstein, F., Baumslag, M.: Limitations on constructing expanders with Cayley graphs, manuscript 1989.

[AM] Alon, N., Milman, V.D.: λ_1, isoperimetric inequalities for graphs, and superconcentrators, *J. Comb. Theory - B* **38** (1985), 73-88.

[Asch] Aschbacher, M.: On the maximal subgroups of the finite classical groups, *Inventiones Math.* **76** (1984), 469-514.

[Ba1] Babai, L.: The probability of generating the symmetric group, *J. Comb. Theory - A* **52** (1989), 148-153.

[Ba2] Babai, L.: On the diameter of metacyclic groups, in preparation.

[Bass] Bass, H., The degree of polynomial growth of finitely generated nilpotent groups, *Proc. London Math. Soc.* **25** (1972), 603-614.

[BE] Babai, L., Erdős, P.: Representation of group elements as short products, in: "Theory and Practice of Combinatorics" (A. Rosa, G. Sabidussi, J. Turgeon, eds.), *Ann. Discr. Math.* **12** (1982), 21-26.

[BH] Babai, L., Hetyei, G.: On the diameter of a random Cayley graph of the symmetric group, in preparation

[BKL1] Babai, L., Kannan, S., Luks, E. M.: A bounded round interactive proof of permutation group nonisomorphism, manuscript, 1990.

[BKL2] Babai, L., Kantor, W.M., Lubotzky, A.: Small diameter Cayley graphs for finite simple groups, *European J. Comb.* **10** (1989), 507-522.

[BLS] Babai, L., Luks, E. M., Seress, Á.: Permutation groups in *NC*, *Proc. 19th ACM STOC*, 1987, pp. 409-420.

[Bo] Bovey, J.D.: The probability that some power of a permutation has small degree, *Bull. London Math. Soc.* **12** (1980), 47-51.

[BS1] Babai, L., Seress, Á.: On the diameter of Cayley graphs of the symmetric group, *J. Comb. Theory - A* **49** (1988), 175-179.

[BS2] Babai, L., Seress, Á.: On the diameter of Cayley graphs of permutation groups, submitted

[Cam] Cameron, P. J.: Finite permutation groups and finite simple groups, *Bull. London Math. Soc.* **13** (1981), 1-22.

[Car] Carter, R.: *Simple groups of Lie type*, Wiley 1972.

[DF] Driscoll, J. R., Furst, M. L.: Computing short generator sequences, *Info. and Comput.* **72** (1987), 117-132.

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

[EG] Even, S., Goldreich, O.: The minimum length generator sequence is NP-hard, *J. Algorithms* **2** (1981), 311-313.

[ER] Erdős, P., Rényi, A.: Probabilistic methods in group theory, J. Analyse Math. **14** (1965), 127-138.

[ET1] Erdős, P., Turán, P.: On some problems of a statistical group theory I., *Z. Wahrscheinlichkeitstheorie verw. Geb.* **4** (1965), 175–186.

[ET2] Erdős, P., Turán, P.: On some problems of a statistical group theory II., *Acta Math. Acad. Sci. Hung.* **18** (1967), 151–163.

[FMSST] Fiat, A., Moses, S., Shamir, A., Shimsoni, I., Tardos, G.: Planning and learning in permutation groups, *Proc. 30th IEEE FOCS*, 1989, pp. 274-279.

[Gon] Goncharov, V.L.: On the field of combinatory analysis, *Isvestija Akad. Nauk. SSSR. Ser. Mat.* **8**(1944), 3–48 (in Russian); English translation: *Translations of the AMS* Ser. math. **2/19**(1962), 1–46.

[Gor] Gorenstein, D.: *Finite Groups*, 2nd ed., Chelsea 1980.

[Je] Jerrum, M. R.: The complexity of finding minimum length generator sequences, *Theoretical Computer Science* **36** (1985), 265-289.

[Jo] Joly, J-R.: Équations et variétés algébriques sur un corps fini, *L'Enseignement Mathématique* **19/1-2**, 1-117.

[Ka] Kantor, W.M.: Some large trivalent graphs having small diameters, submitted

[KL] Kantor, W.M., Lubotzky, A.: The probability of generating a finite classical group, in preparation

[LPS] Lubotzky, A., Phillips, R., Sarnak, P.: Ramanujan graphs, *Combinatorica* **8** (1988), 261-277.

[LS] Liebeck, M. W., Saxl, J.: Primitive permutation groups containing an element of large prime order, *J. London Math. Soc.* **31** (1985), 237-249.

[Lu] Luks, E. M.: Parallel algorithm for permutation groups and graph isomorphism, *Proc. 27th IEEE FOCS*, 1986, pp. 292-302.

[Ma] Margulis, G. A.: Explicit group theoretic constructions of combinatorial schemes and their applications for the construction of expanders and concentrators, *Problemy Peredachi Informatsii* **24/1** (Jan.-March 1988), 51-60 (in Russian); English translation: *Probl. of Info. Transm.* **24/1** (July 1988), 39-46.

[McK] McKenzie, P.: Permutations of bounded degree generate groups of polynomial diameter, *Info. Proc. Lett.* **19** (1984), 253-254.

[PV] Preparata, F., Vuillemin, J.: The cube-connected cycles: a versatile network for parallel computation, *Comm. ACM* **24** (1981), 300-309.

[Qu] Quisquater, J-J.: personal communication

[Ste] Steinberg, R.: Generators for simple groups, *Canad. J. Math.* **14** (1962), 277-283.

[Sto] Stone, H. S.: Parallel processing with the perfect shuffle, *IEEE Trans. Comput.* **C-20** (1971), 153-161.

[We] Weil, A.: Number of solutions of equations in finite fields, *Bull. Amer. Math. Soc.* **55** (1949), 497-504.

[Wi] Wielandt, H.: *Finite Permutation Groups*, Academic Press, New York 1964.

[Wo] Wolf, J.: Growth of finitely generated solvable groups and curvature of Riemannian manifolds, *J. Diff. Geom.* **2** (1968), 421-446.