Probabilistic Group Theory

John D. Dixon
Carleton University

September 27, 2004

Abstract

This survey discusses three aspects of the ways in which probability has been applied to the theory of finite groups: probabilistic statements about groups; construction of randomized algorithms in computational group theory; and application of probabilistic methods to prove deterministic theorems in group theory. It concludes with a brief summary of related results for infinite groups.

Cet article donne un apercu sur trois aspects des façons dont la probabilité est appliquée à la théorie des groupes finis: les faits probabilistiques des groupes; la construction d'algorithmes aléatoires dans la computation; et l'application des moyens probabilistiques pour obtenir les theorems déterministiques dans la théorie des groupes. On termine avec un bref sommaire de resultats se rapportant aux groupes infinis.

Mathematics Subject Classification 2000: 20-02, 20D60, 20F69, 20-04

In the past 20 years, and particularly during the last decade, there has been a growing interest in the use of probability in finite groups. It has been my experience that many pure mathematicians still look on probability theory as an "applied" subject (perhaps because of the way it is taught in our universities), and are dubious about the validity of using probabilistic reasoning in their own discipline. Kolmogorov's axiomatization [51] of probability theory still seems to be a well kept secret. However, no-one should be uncomfortable in a discussion of the applications of probability theory to finite groups, since in these cases the probabilistic statements can be always be simply understood in terms of proportions.

In the current article I shall consider three aspects of the ways in which probability has been applied to problems in group theory. These are: probabilistic statements about groups which give some alternative description of the structure of the group and its elements (Sects. 1 and 2); applications of probability to construct algorithms in computational group theory (Sect 3); and applications of probabilistic methods to prove deterministic theorems in group theory (see Sect. 4). The last section (Sect. 5) deals with some related results in infinite groups.

Since my focus is on group theory, I shall ignore several very important areas where the primary interest is in probability theory such as random walks

on groups and amenable groups (but see Sect. 3.2). There is also interesting recent work on probability and conjugacy classes of the classical groups (see [36] and [37]).

There are two surveys by Shalev (see [74] and [76]) which partially overlap with this paper.

1 Probabilistic questions about elementary properties of groups

1.1 "Statistics" of the symmetric group

During the 1960's Erdös and Turán published a series of papers [31], [32], [33] and [34] on the "statistics" of the symmetric group S_n. A typical result describes the probability distribution of the logarithm of the order of a random element x from S_n; they prove that the distribution of $\ln(ord(x))$ is asymptotically normal with mean $\frac{1}{2}\ln^2 n$ and variance $\frac{1}{3}\ln^3 n$. This contrasts sharply with the classical result of E. Landau that the maximum of $\ln(ord(x))$ is $(1+o(1))\sqrt{n\ln(n)}$ (see [68]) The results of Erdös and Turán have been refined and extended in a number of other papers such as [17], [7], [29], [45], [30], [11] (see also [73]). All of these results are essentially combinatorial and do not use significant group properties of S_n. A little more group theory is used to prove the results in [23].

Similar results for the finite classical linear groups are found in [35].

1.2 Group laws

An earlier example of a probabilistic statement about groups describes how commutative a nonabelian group can be (I am not sure who first made this observation):

- If G is a nonabelian finite group with $k(G)$ conjugacy classes, then

$$\frac{|\{(x,y)\in G\times G \,|\, xy=yx\}|}{|G|^2} = \frac{k(G)}{|G|} \leq \frac{5}{8}$$

This can be interpreted as saying that, for any finite nonabelian group, the probability that two elements chosen at random from G commute is at most $5/8$ (the bound is achieved when G is nonabelian group of order 8). Elementary extensions of this result can be found in [43], [77] and [78].

Note: We are assuming (as we shall generally assume for finite groups throughout this paper) that random elements are chosen independently with the uniform distribution on G. Thus every pair (x,y) has the same probability $1/|G|^2$of being chosen.

This suggests the following general question:

- Let $w := w(X_1, X_2, ..., X_m)$ be a nontrivial word. Does there exist a constant $\eta < 1$ (depending on w) such that, if $w = 1$ is not a law for a finite group G, then a random m-tuple $(x_1, x_2, ..., x_m)$ of elements from G satisfies $w(x_1, x_2, ..., x_m) = 1$ with probability $\leq \eta$?

Of course the theorem quoted above is just the case when $w(X, Y) = X^{-1}Y^{-1}XY$. Although the question has been answered positively in some special cases, for example, for some words representing nilpotent varieties and metabelian varieties (unpublished work), the problem appears to be open in the general case. For results related to nilpotent varieties see the recent paper [38].

Some related results are known. For example, [41] shows that if G is a finite group which is not solvable, then the probability that two random elements generate a solvable subgroup is at most 11/30. A further result appears in [21] (see Sect. 4.2). Both of these results require the classification of finite simple groups for their proofs.

2 Generators

2.1 Generating the symmetric group

In 1969 I was reading the classical book of E. Netto [66] and came across the following claim (p. 90 of the English translation):

> If we arbitrarily select two or more substitutions of n elements, it is to be regarded as extremely probable that the group of lowest order which contains these is the symmetric group, or at least the alternating group. In the case of two substitutions the probability in favor of the symmetric group may be taken as about $\frac{3}{4}$, and in favor of the alternating, but not symmetric, group as about $\frac{1}{4}$. In order that any given substitutions may generate a group which is only a part of the $n!$ possible substitutions, very special relations are necessary, and it is highly improbable that arbitrarily chosen substitutions [..] should satisfy these conditions. The exception most likely to occur would be that all the given substitutions were severally equivalent to an even number of transpositions and would consequently generate the alternating group.

Perhaps Netto was expressing his frustration after trying to generate interesting subgroups of S_n from random permutations. He gives no supporting evidence for his claim.

Let A_n denote the alternating group. Then Netto's conjecture can be written:

$$p_n := \frac{|\{(x, y) \in S_n \times S_n \mid \langle x, y\rangle \geq A_n\}|}{|S_n|^n} \to 1 \text{ as } n \to \infty$$

where $\langle x, y\rangle$ denotes the subgroup generated by x and y. Since $|S_n : A_n| = 2$ for $n \geq 2$, we have $\langle x, y\rangle \leq A_n$ for exactly $\frac{1}{4}$ of the pairs, so the rest of his

claim follows easily. Netto's conjecture can be rephrased as "almost all pairs of elements of S_n generate either A_n or S_n as $n \to \infty$" or

- the probability p_n that two elements chosen at random from S_n generate either A_n or S_n tends to 1 as $n \to \infty$.

Netto's conjecture was proved in [22] where it is shown that $p_n > 1 - 2(\ln \ln n)^{-2}$ for all sufficiently large n. The proof consists of two main steps. Let x, y be random elements of S_n. Then it is shown that: (i) the probability that $\langle x, y \rangle$ is a primitive subgroup of S_n is $1 - 1/n + O(1/n^2)$; and (ii) the probability that neither $\langle x \rangle$ nor $\langle y \rangle$ contains a p-cycle for some prime $p < n - 2$ is $< 1.8(\ln \ln n)^{-2}$ for all n large enough. The result now follows by applying a classical theorem of Jordan: a primitive subgroup of S_n which contains a p-cycle for some prime $p < n - 2$ must contain A_n.

The result in [22] was progressively refined in [13] and in [12]. Finally, assuming the classification of finite simple groups, Babai [5] proved that $p_n = 1 - 1/n + O(1/n^2)$ as conjectured in [22].

2.2 Generating finite simple groups

It follows from Netto's conjecture that almost all pairs of elements from A_n generate all of A_n as $n \to \infty$. At the end of [22] the author made the conjecture that a similar result might be true for the other finite simple groups. More precisely, as S runs through the finite nonabelian simple groups:

- if x, y are random elements from S, then the probability that $\langle x, y \rangle = S$ tends to 1 as $|S| \to \infty$

This was a rash conjecture in 1969 since the proof that every finite simple group is 2-generator is based on the classification of finite simple groups (announced in 1980). However it turned out to be very fruitful. Naturally the complete proof of this conjecture was much more difficult than the special case where $G = A_n$. The general proof follows a different approach (closer to the one used in [5] for Netto's conjecture). We shall describe this approach now.

In 1936 Philip Hall [44] introduced the *Eulerian function* $\varphi(G, d)$ which is defined to be equal to the number of d-tuples from G which generate the finite group G. (The ordinary Euler function $\varphi(n)$ counts the number of 1-tuples which generate the cyclic group of order n.) He proved that his function has the form

$$\varphi(G, d) = \sum_{H \leq G} \mu(G, H) \, |H|^d$$

where the sum is over all subgroups H of G and μ denotes the Möbius function on the lattice of subgroups of G. (Specifically μ is defined recursively by $\mu(G, G) = 1$ and $\sum_{H \leq K \leq G} \mu(G, K) = 0$ for all subgroups $H < G$.) The *zeta function* $\zeta(G, s)$ for G is then defined to be the reciprocal of the finite Dirichlet

series

$$P(G,s) := \frac{\varphi(G,s)}{|G|^s} = \sum_{n=1}^{\infty} \frac{a_n(G)}{n^s} \text{ where } a_h(G) := \sum_{H \leq G, |G:H|=n} \mu(G,H)$$

Clearly, $P(G,d)$ is the probability that a random d-tuple of elements from G generates G, and for specific groups Hall's formula may be used to compute this probability exactly (see [1] and [46]). The general zeta function has a number of interesting properties, many of which are not well understood. See, for example, [14].

A lot of information about G is usually required if we want to calculate $P(G,s)$ exactly. However, in many cases a useful estimate can be obtained by using just the terms involving the maximal subgroups of G. If M is maximal in G, then $\mu(G,M) = -1$, and for all $s \geq 0$ we have the inequality

$$P(G,s) \geq 1 - \sum_{m=2}^{\infty} \frac{b_m(G)}{m^s}$$

where $b_m(G)$ is the number of maximal subgroup of index m in G. If we are lucky, then sufficient knowledge of the maximal subgroups of G may lead to a nontrivial estimate on $P(G,d)$ for sufficiently large d.

For example, in the alternating group A_n the maximal subgroups which are not primitive are easily described, and the maximal subgroups of A_n which are primitive are known to have small orders (see [26] Sect.8.5 and Theorem 5.6B) It is therefore possible to show that $P(A_n, 2) \to 1$ as $n \to \infty$. This is the idea behind the proof in [5] although it is not stated in exactly this way. Indeed it is now known [55] that there are exactly $n/2 + o(n)$ conjugacy classes of maximal subgroups in A_n, and Babai's theorem follows easily from this because every proper subgroup of A_n has index at least n.

In [49] and [53] Kantor, Lubotzky, Liebeck and Shalev completed the proof of the conjecture in [22] by showing that as S runs over the finite simple groups, $P(S,2) \to 1$ as $|S| \to \infty$. Their proof uses detailed knowledge of the maximal subgroups of the various classes of simple groups and is, of course, dependent on the classification. As we shall see later (Sect. 4.2), the theorem which they proved has applications to problems which seem to have nothing to do with probabilistic questions.

In the past few years much more has been proved about this problem and related questions. For example, Liebeck and Shalev [56] proved a conjecture of Kantor and Lubotzky for finite nonabelian simple groups S:

- If x is random element and y is a random involution from S then $S = \langle x, y \rangle$ with probability approaching 1 as $|S| \to \infty$.

Guralnick and Kantor [42] have also proved:

- In each finite nonabelian simple group S there is a conjugacy class C such that for each fixed element $x \neq 1$ from S and a random element y from C, the probability that $\langle x, y \rangle = S$ is at least $1/10$.

3 Algorithms

A probabilistic algorithm is an algorithm which, at some stages, does not prescribe a determined step but "tosses a coin" to decide what the next step should be. The effect of introducing randomization into the execution of an algorithm can often speed up the running time of the algorithm as well as simplify its programming. Randomized algorithms of this type have been used over the past 40 years, and have become increasingly important in solution of computational problems in combinatorics and algebra. A good general reference is [65]. The paper [15] gives a good overview of probabilistic algorithms applied to groups (see, also [16]). The papers in the proceedings [50] give a good idea of the computational problems in group theory of current interest, and the importance that probabilistic methods play.

Of particular interest are *Monte Carlo algorithms.* These are randomized algorithms whose reliability (probability of returning the correct answer) can be increased arbitrarily at the expense of extra time. A Monte Carlo algorithm which never returns an incorrect answer (but may sometimes return "fail" to indicate that it cannot find the solution) is called a *Las Vegas algorithm.*

3.1 An example: the structure of U_n

An example of a LasVegas algorithm which may be familiar is a *pseudo-prime test.* These are fast tests used to determine when a large integer n is composite and to give convincing evidence for primality when the integer is prime. The tests are based on recognizing distinguishing properties of the group $\mathbb{U}_n$ of units of the ring $\mathbb{Z}/n\mathbb{Z}$. We take a few moments to describe one of these tests here (see [72]).

Let $n > 1$ be an odd integer. Then we can represent the group $\mathbb{U}_n$ by the set of integers k with $1 \leq k < n$ with greatest common divisor $GCD(k, n) = 1$ with the operation $\cdot$ of multiplication modulo n. If n is prime, then $\mathbb{U}_n$ is a cyclic group of order $n - 1$ whose unique element of order 2 is $n - 1$. If n is not prime then: either n is a prime power and so $|\mathbb{U}_n|$ does not divide $n - 1$; or else n has at least two odd prime divisors and so $\mathbb{U}_n$ at least two elements of order 2. Suppose that n is not prime, and write $n - 1 = 2^t m$ where $t \geq 1$ and m is odd. We say that an integer k with $1 \leq k < n$ is a *witness* to the compositeness of n if any of the following hold: (i) $GCD(k, n) \neq 1$; (ii) k has order not dividing $n - 1$; or (iii) k has even order $2h$ in $\mathbb{U}_n$ but k^h is not equal to $n - 1$. We can check these three conditions as follows: a particular value of k is a witness unless $k^m = 1$ or one of the elements $k^m, k^{2m}, ..., k^{2^{t-1}m}$ is equal to $n - 1$. (The usefulness of this criterion depends on the fact that there is a fast way to compute powers of k modulo n; see, for example, [65].)

We now have a Las Vegas algorithm for checking compositeness of an odd integer. Choose a random integer k from the interval $1 \leq k < n$ and test to see whether k is a witness to the compositeness of n. If n is composite, then it can be proved that a randomly chosen k will be a witness with probability at least $1/2$. If we find a witness, then we know that n is composite and so

we are finished. The test can never give us a proof that n is prime. However, if we perform d independent repetitions of the test on n and do not find a witness, then we should become increasingly convinced that n is prime since the probability that this event happens for a composite n is $\leq (1/2)^d$. This pseudo-prime test (or a similar test) is widely used in programs such as Maple as an inexpensive partial substitute for primality testing.

A similar search for witnesses can be used to determine whether or not a given finite set of matrices from $GL(d,q)$ generates a subgroup containing $SL(d,q)$ or one of the other classical groups (see [67], [69], [70] and [75]).

3.2 Finding random elements

A problem which arises in many probabilistic algorithms in group theory is:

- If we are given a set of generators for a group G, how can we efficiently generate random elements of G?

In some cases this can be done easily; an important case is when the group is given as a permutation group and a stabilizer chain and strong generating set are known. However, in other cases, when we have less information about G or a less structured generating set, the problem may be much more difficult.

In practice we do not require that the probability distribution be exactly uniform, but it should be close to uniform. Consideration of this problem leads to the analysis of random walks on the group (more precisely, on the Cayley graph associated with the set of generators) which can be described in terms of Markov chains. Measuring the efficiency of the algorithms to generate near random elements then reduces to determining how fast the Markov chain converges. This in turn uses some interesting linear representation theory (see [20] for an excellent introduction). Along similar lines we note that [2] discusses the problem of random walks on S_n and explains why 6 random riffle shuffles of an ordinary deck of cards are not sufficient to randomize the deck, but 7 shuffles suffice.

The general problem of generating random elements in a group is not yet satisfactorily solved, and it is clear that naive methods of computing random elements are not adequate (see [6] and [71]). The problem is particularly important when G is a group of matrices over a finite field.

We remark that Babai, Luks and Seress have introduced a simple technique called *random subproducts* which can sometimes be used to substitute for the problem of finding random elements. This is based on the following easily proved proposition (see [15] Prop. 2.1):

- If H is a proper subgroup of G and $x_1, ..., x_m$ is a set of generators of G, then with probability $\geq 1/2$ a random element from the set

$$\{x_1^{\varepsilon_1} x_2^{\varepsilon_2} ... x_m^{\varepsilon_m} \mid \varepsilon_1, \varepsilon_2, ..., \varepsilon_m \in \{0,1\}\}$$

  does *not* lie in H.

3.3 Recognizing S_n

Let $f(X)$ be a monic separable polynomial of degree n with integer coefficients, and let $G = Gal(f)$ denote the Galois group of the splitting field of f where G is considered as a permutation group on the set of n roots. For each prime p we can consider the factorization

$$f(X) = f_1(X)f_2(X)...f_r(X) \pmod p$$

where the $f_i(X)$ are monic irreducible modulo p. Suppose that p does not divide the discriminant $disc(f)$ of f (so the factors $f_i(X)$ are distinct) and let $n_1 \leq n_2 \leq ... \leq n_r$ be the degrees of the factors. Then Frobenius showed that G contains permutations with cycle type $(n_1, n_2, ..., n_r)$ (see [82] Sect. 61); and later Chebotarev showed that, in a suitable sense, the proportion of p which give rise to a particular cycle type is equal to the proportion of permutations in G with that cycle type (see [81]). This theorem is used to help identify the Galois groups of irreducible polynomials.

Van der Waerden also proved that "almost all" irreducible polynomials over the rationals have the full symmetric group as their Galois group. It is therefore worthwhile having a quick test to determine whether this is true for $Gal(f)$. This leads to the following heuristic. For a sequence of "random" primes p find the associated cycle type $(n_1, n_2, ..., n_r)$ which must appear in $Gal(f)$ according to Frobenius' theorem. Try to determine whether the existence of these cycle types in $Gal(f)$ implies that $Gal(f)$ is the full symmetric group. Recalling that two elements of S_n have the same cycle type exactly when they are conjugate in S_n, and taking into consideration Chebotarev's theorem, we have the following (slightly idealized question) about recognizing when we have the full symmetric group.

We shall say that a list $x_1, x_2, ..., x_d$ from a group G *invariably generates* G if $\langle y_1, y_2, ..., y_d \rangle = G$ whenever y_i is conjugate to x_i in G for $i = 1, 2, .., d$. We then ask:

- Given $d \geq 2$ what is the probability that d random elements of S_n invariably generate S_n?

This problem was first posed by John McKay (private communication). He conjectured from numerical experiments that the expected number of random elements required to invariably generate S_n is a constant (about 5) independent of n. It is shown in [25] that $O((\ln n)^{1/2})$ random elements are enough, and soon after that Łuczak and Pyber [59] improved this to show that for each $\varepsilon > 0$ there exists a constant C (depending on ε but independent of n) such that C random elements of S_n invariably generate S_n with probability at least $1 - \varepsilon$. A good value of C is still not known.

3.4 Other algorithmic problems

There are many other algorithmic problems in which probabilistic methods play a part. For example, suppose that we are given a permutation group G. How

difficult is it to find an element of order p in G? One answer is given by [47] where it is shown:

- If G is a permutation group of degree n, and p is a prime which divides $|G|$, then the probability that a random element of G has its order divisible by p is at least $1/n$.

Thus there is a good probability that a randomly chosen element will have its order divisible by p, and then some power of this element has order p. Surprisingly, the proof uses the classification of finite simple groups.

The paper [52] proposes a probabilistic algorithm for determining when a linear group has a tensor product decomposition, [79] describes a probabilistic algorithm to find the structure of a finite abelian group, and [64] considers the discrete logarithm problem in $GL(n, q)$.

4 Applications to deterministic theorems

A famous theorem of Georg Cantor states that, because the set of real numbers is uncountable and the set of algebraic real numbers is countable, therefore the set of transcendental real numbers is uncountable; in particular, transcendental real numbers exist. In 1947 Paul Erdös [28] popularized similar arguments (they had occasionally been used earlier by other authors) to prove existence theorems in finite structures. This extension of the pigeonhole principle is now called the *probabilistic method* (see [3]) and has been used with great success in combinatorics. Recently probabilistic methods have been successfully applied to problems in group theory.

The possibility of such applications was predicted by Paul Turán. In a letter dated (Budapest. 16.3.1970) to the author, Turán concludes with:

> *My "Einstellung" with statistical group theory will be perhaps more understandable by repeating how I came to the idea of statistical group theory. My "old age dream" (an expression, imitating "Kronecker's Jugendtraum") is to disprove Burnside's conjecture (if G is finitely generated and for all elements x we have with the same n $x^n = e$ then G is finite) by finding for such groups an appropriate representation in a space so that one could find that in this space the "points" belonging to finite groups form a "small" set. But I could not find a good representation so far.*

So far no-one has succeeded in tackling Burnside's problem in this way, but in recent years there have been a number of successful applications of probabilistic group theory somewhat along the lines which Turán describes. We discuss some of these.

4.1 The $(2,3)$-generator problem

The *(2,3)-generator problem* was open for nearly a century. It arose from the study of groups acting on Riemann surfaces and asks:

- Which finite simple groups S can be generated by two elements x, y of orders 2 and 3, respectively?

The modular group $PSL(2,\mathbb{Z})$ is isomorphic to a free product $\langle x\rangle * \langle y\rangle$ of a group of order 2 and a group of order 3. Therefore the $(2,3)$-generator problem is equivalent to: which simple groups S are homomorphic images of $PSL(2,\mathbb{Z})$?

It is easily verified that A_n is $(2,3)$-generated for all $n > 8$, but for the other families of simple groups the problem is more complicated. In 1996 it was shown that the simple groups $PSL(d,q)$ are $(2,3)$-generated for odd q except when $d = 2$ and $q = 9$ (see [18] and [19]). At that time it was conjectured that, with a finite number of exceptions, every finite simple group was $(2,3)$-generated. A strengthened form of this conjecture was tackled by Liebeck and Shalev [54] who proved:

- If S runs over the finite classical simple linear groups which are *not* of the form $PSp(4,q)$, then the probability that two random elements of order 2 and 3, respectively, generate S tends to 1 as $|S| \to \infty$. Moreover, the corresponding probability as S runs over the groups $PSp(4,q)$ with $q \neq 2^k$ or 3^k is $1/2$.

Unexpectedly, it turned out that the groups from the two infinite families $PSp(4,2^k)$ and $PSp(4,3^k)$ fail to be $(2,3)$-generated. However, Liebeck and Shalev's result shows that, except for these families, all finite simple classical groups — with finitely many possible exceptions — are $(2,3)$-generated (what the exceptions may be is still unknown).

Since then Lübeck and Malle [57] settled the $(2,3)$-generation problem for exceptional groups of Lie type using more direct methods. They show that, except for $G_2(2)'$ and the Suzuki groups, all of these groups are $(2,3)$-generated.

4.2 Residual properties of free groups

A group G is called *residually-$\mathcal{C}$* for a class $\mathcal{C}$ of groups if for each $x \neq 1$ in G there is a normal subgroup N of G with $x \notin N$ and G/N isomorphic to a group in $\mathcal{C}$. It is well known that any free group F of rank ≥ 2 is residually finite; indeed F is a residually finite p-group for each prime p. In 1969 Magnus [60] asked the question:

- Is it true that F residually-$\mathcal{X}$ for every infinite set $\mathcal{X}$ of finite nonabelian simple groups?

Equivalently, is it true that for each $x \neq 1$ in F there exists a normal subgroup N_x such that $x \notin N_x$ and F/N_x is isomorphic to one of the groups S in $\mathcal{X}$?

The problem is easily reduced to the case where F has rank 2 since every free group of rank > 2 is residually free of rank 2. After several partial solutions, Magnus' question was completely answered in the affirmative by Weigel in a series of three long papers ([85], [83] and [84]). More recently, a stronger probabilistic version of Weigel's theorem has proved in [21] (both theorems require the classification of finite simple groups). We can explain the latter result as follows.

Let F be the free group on two generators X, Y and let $w(X, Y)$ be a nontrivial word in F. In order to prove Magnus' conjecture it is necessary to show that there exists $S \in \mathcal{X}$ and a homomorphism of F onto S such that $w(X, Y)$ is not mapped onto the identity of S. Equivalently, there exist $x, y \in S$ such that $S = \langle x, y \rangle$ and $w(x, y) \neq 1$. In [21] the following is proved.

- Let $w(X, Y)$ be a nontrivial word in F. Then, as S runs over the set of all finite nonabelian simple groups, the probability that two random elements x, y from S generate S and satisfy $w(x, y) \neq 1$ tends to 1 as $|S| \to \infty$.

Weigel's theorem clearly follows from this. The proof is simplified because we already know (see Sect. 2.2) that x, y generate S with probability tending to 1, so it is enough to show that $w(x, y) \neq 1$ also with probability tending to 1 as $|S| \to \infty$. This is done by considering separately each family in a finite set of infinite families of nonabelian simple groups.

We illustrate the proof for the family of alternating groups A_n (the easiest case). Assume that $w(X, Y)$ is a reduced word of length $r \geq 1$, and write $w(X, Y) = w_1(X, Y)...w_r(X, Y)$ where $w_i(X, Y) \in \left\{X, X^{-1}, Y, Y^{-1}\right\}$ for each i. Suppose that s of these factors are X or X^{-1} and t of the factors are Y or $Y^{-1}$and assume that $n \geq r + 2$. Let α_0 be a fixed element from the set Ω on which A_n acts. Now for each $(r + 1)$-tuple $(\alpha_0, \alpha_1, ..., \alpha_r)$ of distinct points in Ω, there exist $(n - s)!/2$ values of $x \in A_n$ such that $\alpha_i = \alpha_{i-1}^{w_i(x,y)}$ for the indices where $w_i(X, Y) \in \{X, X^{-1}\}$, and $(n - t)!/2$ values of y such that $\alpha_i = \alpha_{i-1}^{w_i(x,y)}$ for the indices where $w_i(X, Y) \in \{Y, Y^{-1}\}$. Since $\alpha_0^{w(x,y)} = \alpha_r \neq \alpha_0$ for such choices of x and y, we must have $w(x, y) \neq 1$. Since there are $(n-1)!/(n-r-1)!$ $(r + 1)$-tuples of distinct points starting with α_0, this guarantees that there are at least $\frac{1}{4}(n - s)!(n - t)!(n - 1)!/(n - r - 1)!$ pairs (x, y) from A_n for which $w(x, y) \neq 1$. This latter number is asymptotic to $[\frac{1}{2}n!]^2$ and so the probability that two random elements from A_n satisfy $w(x, y) \neq 1$ tends to 1 as $n \to \infty$. In particular, this gives a simple solution to Magnus' question in the special case whenever $\mathcal{X}$ is an infinite set of alternating groups.

5 Infinite groups and the ubiquity of free subgroups

When we consider infinite groups, even the statement of probabilistic questions becomes a little subtle. First we need a suitable probability distribution defined on the group. For some groups this can be done very naturally. For

example, there is a natural probability distribution on a profinite group defined in terms of the uniform distribution on its finite quotients (its Haar measure). In this context Mann and Shalev ([62] and [61]) have considered the problem (for integers $k \geq 1$) :

- For which profinite groups G is there a positive probability that the (closed) subgroup generated by a random k-tuple of elements from G is equal to G?

In particular, they show that, if G satisfies this condition for some k, then G also satisfies the condition of polynomial maximal subgroup growth. Related theorems are proved in [58], [8] and [48].

If we have no natural probability distribution on our group, it may still be possible to make "almost all" statements in a sense similar to "almost all real numbers are transcendental". These are not probabilistic statements, but they have much the same flavour as "probability 1" statements.

An early example of such a theorem is due to Epstein [27] who proved that

- If G is a simple Lie group, then almost all k-tuples from G generate a free group of rank k.

In this case "almost all" means all but a set of measure 0 in the natural measure on G. If G is not compact, this measure does not define a probability distribution on G.

In 1990 the author proved in [24] a parallel result on the ubiquity of free subgroups in the infinite symmetric group of countably infinite degree:

- If $k \geq 2$, then almost all k-tuples from $Sym(\mathbb{N})$ generate a subgroup which is free of rank k. Moreover, almost all of these subgroups are m-transitive for every $m \geq 1$.

This gave a nonconstructive proof of the existence of highly transitive free subgroups of $Sym(\mathbb{N})$ (examples of such subgroups had been constructed earlier in [63]). For $G = Sym(\mathbb{N})$ there is no natural measure. However, we can define a simple metric d on G by setting $d(x, y) := 2^{-t}$ if xy^{-1} fixes $0, 1, ..., t-1$ but does not fix t. Under this metric G is a complete metric space and also a topological group. In particular, the Baire category theorem holds in G^k, and so it makes sense to consider meagre sets (= sets of the "first category") as "null". So in this context we say that a subset of G^k includes *almost all* k-tuples in G if its complement is meagre.

The result above has been extended by Glass and others (see [39], [80] and [40]). The proper setting for these theorems appears to be in the context of metrizable topological groups which are Polish spaces (see [10]) since these are precisely the spaces in which a Baire category theorem holds.

Other theorems on the ubiquity of free subgroups in other contexts are found in [9] and [4].

Acknowledgement This work was supported in part by NSERC under Grant A7171.

References

[1] Vincenzo Acciaro, *The probability of generating some common families of finite groups*, Utilitas Math. **49** (1996), 243–254.

[2] D. Aldous and P. Diaconis, *Shuffling cards and stopping times*, Amer. Math. Monthly **93** (1986), 333–348.

[3] Noga Alon and Joel H. Spencer, *The probabilistic method*, Wiley-Interscience, New York, 2000.

[4] G.N. Arzhantseva and A.Yu. Ol'shanshij, *The class of groups all of whose subgroups with lesser number of generators are free is generic*, Math. Notes **59** (1996), 350–355, transl. from Mat. Zametki 59 (1996) 489-496.

[5] László Babai, *The probability of generating the symmetric group*, J. Combin. Theory Ser. A **52** (1989), 148–153.

[6] László Babai and Igor Pak, *Strong bias of group generators: an obstacle to the "product replacement algorithm"*, Proc.of Eleventh Annual ACM-SIAM Symposium on Discrete Algorithms (San Francisco, 2000) (New York), ACM, 2000, pp. 627–635.

[7] M.R. Best, *The distribution of some variables on symmetric groups*, Nederl. Akad. Wetensch. Proc. Ser. A **73**=Indag. Math. **32** (1970), 385–402.

[8] Meenaxi Bhattacharjee, *The probability of generating certain profinite groups by two elements*, Israel J. Math. **86** (1994), 311–329.

[9] ———, *The ubiquity of free subgroups in certain inverse limits of groups*, J. Algebra **172** (1995), 134–146.

[10] N. Bourbaki, *General topology (part 2)*, Hermann, Paris, 1966.

[11] J.D. Bovey, *An approximate probability distribution for the order of the elements of the symmetric group*, Bull. London Math. Soc. **12** (1980), 41–46.

[12] ———, *The probability that some power of a permutation has small degree*, Bull. London Math. Soc. **12** (1980), 47–51.

[13] John Bovey and Alan Williamson, *The probability of generating the symmetric group*, Bull. London Math. Soc. **10** (1978), 91–96.

[14] Kenneth S. Brown, *The coset poset and probabilistic zeta function of a finite group*, J. Algebra **225** (2000), 989–1012.

[15] Gene Cooperman and Larry Finkelstein, *Combinatorial tools for computational group theory*, Groups and computation (New Brunswick, NJ, 1991), Amer. Math. Soc., Providence, RI, 1993, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., 11, pp. 53–86.

[16] Gene Cooperman and George Havas, *Elementary algebra revisited: randomized algorithms*, Randomization methods in algorithm design (Princeton, NJ, 1997) (Providence, RI), Amer. Math. Soc., 1999, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., 43, pp. 37–44.

[17] J. Dénes, P. Erdős, and P. Turán, *On some statistical properties of the alternating group of degree n*, Enseign. Math. (2) **15** (1969), 89–99.

[18] L. Di Martino and N.A. Vavilov, $(2,3)$*-generation of* $SL(n,q)$*. I*, Comm. Algebra **22** (1994), 1321–1347.

[19] ———, $(2,3)$*-generation of* $SL(n,q)$*. II*, Comm. Algebra **24** (1996), 487–515.

[20] Persi Diaconis, *Group representations in probability and statistics*, Institute of Mathematical Statistics, Hayward, CA, 1988.

[21] J.D. Dixon, L. Pyber, A. Seress, and A. Shalev, *Residual properties of free groups and probabilistic methods*, submitted for publication.

[22] John D. Dixon, *The probability of generating the symmetric group*, Math. Z. **110** (1969), 199–205.

[23] ———, *Maximal abelian subgroups of the symmetric group*, Canad. J. Math. **23** (1971), 426–438.

[24] ———, *Most finitely generated permutation groups are free*, Bull. London Math. Soc. **22** (1990), 222–226.

[25] ———, *Random sets which invariably generate the symmetric group*, Discrete Math. **105** (1992), 25–39.

[26] John D. Dixon and Brian Mortimer, *Permutation groups*, Springer, New York, 1996.

[27] D.B.A. Epstein, *Almost all subgroups of a Lie group are free*, J. Algebra **19** (1971), 261–262.

[28] P. Erdős, *Some remarks on the theory of graphs*, Bull. Amer. Math. Soc. **53** (1947), 292–294.

[29] P. Erdős and R.R. Hall, *Probabilistic methods in group theory. II*, Houston J. Math. **2** (1976), 173–180.

[30] ———, *Some new results in probabilistic group theory*, Comment. Math. Helv. **53** (1978), 448–457.

[31] P. Erdős and P. Turán, *On some problems of a statistical group-theory. I*, Z. Wahrschein. Verw. Gebeite **4** (1965), 175–186.

[32] ———, *On some problems of a statistical group-theory. II*, Acta Math. Acad. Sci. Hungar. **18** (1967), 151–163.

[33] ———, *On some problems of a statistical group-theory. III*, Acta Math. Acad. Sci. Hungar. **18** (1967), 309–320.

[34] ———, *On some problems of a statistical group-theory. IV*, Acta Math. Acad. Sci. Hungar. **19** (1968), 413–435.

[35] Jason Fulman, *Cycle indices for the finite classical groups*, J. Group Theory **2** (1999), 251–289.

[36] ———, *A probabilistic approach toward conjugacy classes in the finite general linear and unitary groups*, J. Algebra **212** (1999), 557–590.

[37] ———, *A probabilisitic approach to conjugacy classes in the finite symplectic and orthogonal groups*, J. Algebra **234** (2000), 207–224.

[38] J.E. Fulman, M.D. Galloy, G.J. Sherman, and J.M. Vanderkam, *Counting nilpotent pairs in finite groups*, Ars Combin. **54** (2000), 161–178.

[39] A.M.W. Glass, *The ubiquity of free groups*, Math. Intelligencer **14** (1992), 54–57.

[40] A.M.W. Glass, Stephen H. McCleary, and Rubin Matatyahu, *Automorphism groups of countable highly homogeneous partially ordered sets*, Math. Z. **214** (1993), 55–66.

[41] R.M. Guralnick and J.S. Wilson, *The probability of generating a finite soluble group*, Proc. London Math. Soc (3) **81** (2000), 405–427.

[42] Robert M. Guralnick and William M. Kantor, *Probabilistic generation of finite simple groups*, J. Algebra **234** (2000), 743–792.

[43] W.H. Gustafson, *What is the probability that two group elements commute?*, Amer. Math. Monthly **80** (1973), 1031–1034.

[44] Philip Hall, *The eulerian functions of a group*, Quart. J. Math. **7** (1936), 134–151.

[45] R.R. Hall, *Extensions of a theorem of Erdős-Rényi in probabilistic group theory*, Houston J. Math. **3** (1977), 225–234.

[46] Ishai Ilani, *Zeta functions related to the group $SL_2(\mathbf{Z_p})$*, Israel J. Math. **109** (1999), 157–172.

[47] I.M Isaacs, W.M. Kantor, and N. Spaltenstein, *On the probability that a group element is p-singular*, J. Algebra **176** (1995), 139–181.

[48] Moshe Jarden and Alexander Lubotzky, *Random normal subgroups of free profinite groups*, J. Group Theory **2** (1999), 213–224.

[49] William M. Kantor and Alexander Lubotzky, *The probability of generating a finite classical group*, Geom. Dedicata **36** (1990), 67–87.

[50] W.M. Kantor and A. Seress (eds), *Groups and computation III (Proc. 3rd Internat. Conf. at Ohio State Univ., 1999)*, Walter de Gruyter, Berlin, 2001.

[51] A. Kolmogorov, *Foundations of the theory of probability*, Chelsea, New York, NY, 1956, transl. of "Grundbegriffe der Wahrscheinlichkeitsrechnung" (1933).

[52] C.R. Leedham-Green and E.A. O'Brien, *Recognizing tensor products of matrix groups*, Internat. J. Algebra Comput. **7** (1997), 541–559.

[53] Martin W. Liebeck and Aner Shalev, *The probability of generating a finite simple group*, Geom. Dedicata **56** (1995), 103–113.

[54] ———, *Classical groups, probabilistic methods, and the* $(2,3)$*-generation problem*, Ann. of Math. (2) **144** (1996), 77–125.

[55] ———, *Maximal subgroups of symmetric groups*, J. Combin. Theory (Ser. A) **75** (1996), 341–352.

[56] ———, *Simple groups, probabilistic methods, and a conjecture of Kantor and Lubotzky*, J. Algebra **184** (1996), 31–57.

[57] Frank Lübeck and Gunter Malle, $(2,3)$*-generation of exceptional groups*, J. London Math. Soc. (2) **59** (1999), 109–102.

[58] Alexander Lubotzky, *Random elements of a free profinite group generate a free subgroup*, Illinois J. Math. **37** (1993), 78–84.

[59] Tomasz Łuczak and László Pyber, *On random generation of the symmetric group*, Combin. Probab. Comput. **2** (1993), 505–512.

[60] W. Magnus, *Residually finite groups*, Bull. Amer. Math. Soc. **75** (1969), 305–316.

[61] Avinoam Mann, *Positively finitely generated groups*, Forum Math. **8** (1996), 429–459.

[62] Avinoam Mann and Aner Shalev, *Simple groups, maximal subgroups, and probabilistic aspects of profinite groups*, Israel J. Math. **96** (1996), 449–468.

[63] T.P. McDonough, *A permutation representation of a free group*, Quart. J. Math. Oxford (2) **28** (1977), 353–356.

[64] Alfred J. Menezes and Yi-Hong Wu, *The discrete logarithm problem in* $GL(n,q)$, Ars Combin. **47** (1997), 23–32.

[65] Rajeev Motwani and Prabhakar Raghavan, *Randomized algorithms*, Cambridge Univ. Press, Cambridge, 1995.

[66] E. Netto, *Substitutionentheorie and ihre Anwendungen auf die Algebra*, Teubner, Leipzig, 1882, English transl. 1892, second ed., Chelsea, New York, 1964.

[67] Peter M. Neumann and Cheryl E. Praeger, *A recognition algorithm for special linear groups*, Proc. London Math. Soc. (3) **65** (1992), 555–603.

[68] J.-L. Nicolas, *Sur l'ordre maximum d'un élément dans le groupe S_n des permutations*, Acta Arith. **14** (1967/68), 315–332.

[69] Alice C. Niemeyer and Cheryl E. Praeger, *Implementing a recognition algorithm for classical groups*, Groups and computation, II (New Brunswick, NJ 1995), Amer. Math. Soc., Providence, RI, 1997, pp. 273–296.

[70] ———, *A recognition algorithm for classical groups over finite fields*, Proc. London Math. Soc. (3) **77** (1998), 117–169.

[71] Igor Pak, *What do we know about the product replacement algorithm?*, Groups and computation III (Ohio State Univ., Ohio, 1999), Walter de Gruyter, Berlin, 2001, pp. 301–347.

[72] M.O. Rabin, *Probabilistic algorithm for primality testing*, Journal of Number Theory **12** (1980), 128–138.

[73] Vladamir N. Sachov, *Probabilistic methods in combinatorial analysis*, Cambridge Univ. Press, Cambridge, 1997.

[74] Aner Shalev, *Simple groups, permutation groups and probability*, Documenta Mathematica (1998), 129–137 (electronic), Proc. of International Congress of Mathematicians, vol. II (Berlin, 1998).

[75] ———, *A theorem on random matrices and some applications*, J. Algebra **199** (1998), 124–141.

[76] ———, *Asymptotic group theory*, Notices of Amer. Math. Soc. **48** (2001), 383–389.

[77] Gary Sherman, *What is the probability an automorphism fixes a group element?*, Amer. Math. Monthly **82** (1975), 261–264.

[78] ———, *A probabilistic estimate of invariance for groups*, Amer. Math. Monthly **85** (1978), 361–363.

[79] Edlyn Teske, *A space efficient algorithm for group structure computation*, Math. Comp. **67** (1998), 1637–1663.

[80] J.K. Truss, *Joint embeddings of infinite permutation groups*, Algebra Logic Appl. **9** (1997), 121–134.

[81] N. Tschebotareff, *Die Bestimmung der Dichtigkeit einer Menge von Primzahlen, welche zu einer gegebenen Substitutionsklasse gehören*, Math. Annalen **95** (1926), 191–228.

[82] B.L van der Waerden, *Modern algebra*, Ungar, New York, 1948.

[83] T. Weigel, *Residual properties of free groups II*, Comm. Algebra **20** (1992), 1395–1425.

[84] ———, *Residual properties of free groups III*, Israel J. Math. **77** (1992), 65–81.

[85] ———, *Residual properties of free groups*, J. Algebra **160** (1993), 16–41.