# PROBABILISTIC METHODS IN GROUP THEORY

*By*

P. ERDÖS AND A. RÉNYI
*in Budapest, Hungary*

## Introduction

The application of probabilistic methods to another chapter of mathematics (number theory, different branches of analysis, graph theory etc.) has in the last 30 years often led to interesting results, which could not be obtained by the usual methods of the chapters in question. These results are in most cases of the following character: it is shown that in some sense "most" elements of a set of mathematical objects possess a certain property. Thus these results deal with typical properties of elements of a certain set, neglecting elements which behave in a different way, provided that their number is in some sense negligibly small. If the number of elements of the set considered is infinite a certain natural measure has to be introduced and relations are studied which are valid for "almost all" elements of the set considered, i.e. with the exception of subset of measure zero according to the chosen measure. If finite systems are studied, then the most natural measure is the number of elements of the sets considered. In such cases the point of view usually adopted is as follows: if $S_n (n = 1, 2, \cdots)$ is a sequence of finite sets, the number of elements of $S_n$ being equal to $N(n)$ where $\lim_{n \to +\infty} N(n) = +\infty$, and if $A(n)$ denotes the number of elements of $S_n$ having the property $A$, and if $\lim_{n \to +\infty} \frac{A(n)}{N(n)} = 1$, we say that in the limit for $n \to +\infty$ "almost all" elements of $S$ possess the property $A$.

In proving such assertions, probabilistic methods are usually the natural tool, in spite of the fact that the problem considered has nothing to do with chance.

It often happens that the easiest (or the only available) way to prove that $S$ contains at least one element having the property $A$ for sufficiently large values of $n$, is to prove that $\liminf_{n\to+\infty} \frac{A(n)}{N(n)} > 0$; in this way the existence of elements of $S_n$ having property $A$ can be proved while the actual construction of an element of $S_n$ having the property $A$ cannot be carried out. Thus probabilistic considerations lead often to proofs of existence concerning finite systems. While investigations of the above described type have been frequently made in number theory (see for instance [1], where further references are given) and graph theory (we mention here for instance our papers [2], [3], [4], [5]), up to now such methods were only exceptionally (see [6]) applied to the study of finite algebraic systems.

In the present paper we shall give an example of applying probabilistic methods to the study of finite groups.

Let $G_n$ be a finite Abelian group of order $n$. (We use the additive notation for the group operation). Let us choose $k$ arbitrary elements of $G_n$, and denote them by $a_1, a_2, \cdots, a_k$. Let us consider all possible $2^k$ sums $\varepsilon_1 a_1 + \varepsilon_2 a_2 + \cdots + \varepsilon_k a_k$ where $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_k$ are equal either to 0 or to 1. In other words we consider the set of all possible sums $a_{i_1} + a_{i_2} + \cdots + a_{i_r}$ where $1 \leqq i_1 < i_2 < \cdots < i_r \leqq k$ and $0 \leqq r \leqq k$. The question is now, whether all elements $b$ of $G_n$ can be represented in the form $b = \sum_{i=1}^{k} \varepsilon_i a_i$?

Of course this is possible only if $2^k \geqq n$, i.e. if $k \geqq \frac{\log n}{\log 2}$. We shall prove in §2 (see Theorem 2) that if we choose the elements $a_1, \cdots, a_k$ of $G_n$ at random and if $k \geqq \frac{\log n + \log\log n + \omega_n}{\log 2}$ where $\omega_n$ tends to $+\infty$ for $n \to +\infty$ arbitrarily slowly, then every $b \in G_n$ can be represented in the form

$$b = \sum_{i=1}^{k} \varepsilon_i a_i \tag{1}$$

with probability tending to 1 for $n \to +\infty$. It is natural to ask also, how much larger the value of $k$ has to be chosen that each $b \in G_n$ should have approximately the same number of representations in the form (1). We shall prove in §1 that if $k \geqq \frac{2\log n + c}{\log 2}$ where $c$ is a sufficiently large positive number then every

$b \in G_n$ has approximately the same number of representations in the form (1) with probability near to 1 (see Theorem 1). More exactly, if

$$k \geqq \frac{2\log n + 2\log \frac{1}{\varepsilon} + \log \frac{1}{\delta}}{\log 2}$$

then with probability $\geqq 1 - \delta$ the number of representations of $b$ in the form (1) is contained between $(1-\varepsilon)\frac{2^k}{n}$ and $(1+\varepsilon)\frac{2^k}{n}$ for all $b \in G_n$; here $\varepsilon$ and $\delta$ are arbitrary small positive numbers (Theorem 2). We conjecture but could not prove up to now that the factor 2 of $\log n$ in Theorem 1 cannot be replaced by a smaller number.

The method which we apply to prove Theorem 1 consists of two steps: The first step consists simply in the evaluation of the expectation of the mean square deviation of the random distribution $\frac{V_k(b)}{2^k}$, where $V_k(b)$ denotes the number of representations of $b$ in the form (1). This idea has been first applied to a particular problem of number theory by *P. Turán* [7]. The idea has been recently developed by *Ju. V. Linnik* [8] into a powerful method in number theory, called by him the "dispersion method". The second step may be characterized as utilizing the smoothing effect of random choice: if most of the numbers $V_k(b)$ are almost equal, but a few of them may be considerably smaller or larger than the average, then by choosing a relatively small number of further elements $a_{k+1}, \cdots, a_{k+j}$ at random, the distribution gets smoothed out, i.e. the distribution $\{2^{-k+r} V_{k+r}(b)\}(b \in G_n)$ is usually much more uniform than the distribution $\{2^{-k}V_k(b)\}$.

## §1. Finite Abelian groups

Let $G_n$ be an Abelian group of order $n$. The elements of $G_n$ will be denoted by $a, b, c, \cdots$ with or without indices. The group operation will be written as addition; accordingly we denote by 0 the unit-element of the group, and by $-a$ the element for which $a + (-a) = 0$; for any $a \in G_n$ we put $1 \cdot a = a$ and $0 \cdot a = 0$. Let $a_1, a_2, \cdots, a_k$ be $k$ elements of $G$ chosen at random, independently of each other, so that each $a_j$ may be equal to an arbitrary element of $G_n$ with the same probability $\frac{1}{n}$. We denote by $V_k(b)$ the number of representations of an element $b$ of $G_n$ in the form

$$b = \varepsilon_1 a_1 + \varepsilon_2 a_2 + \cdots + \varepsilon_k a_k \tag{1.1}$$

where each of the numbers $\varepsilon_j$ may have the value 0 or 1. Then for each $b \in G$ $V_k(b)$ is a random variable.

If the values of $a_1, \cdots, a_k$ are fixed, clearly

$$\sum_{b \in G_n} V_k(b) = 2^k$$

and thus $\left\{\frac{V_k(b)}{2^k}\right\}$ is a probability distribution,. Let $P(\cdots)$ denote the probability of the event in the brackets and let $E(\cdots)$ denote the expectation of the random variable in the brackets.

In what follows we shall often use the following elementary inequality, called usually Markoff's inequality: if $\xi$ is any nonnegative random variable and $\lambda$ a real number, $\lambda > 1$, then

$$P(\xi \geqq \lambda E(\xi)) \leqq \frac{1}{\lambda}. \tag{1.2}$$

If $A$ and $B$ are events, $\xi$ and $\eta$ random variables, we shall denote by $P(A \mid B)$ the conditional probability of the event $A$ under the condition $B$, by $E(\xi \mid B)$ the conditional expectation of $\xi$ under the condition $B$ and by $E(\xi \mid \eta)$ the conditional expectation of $\xi$ given $\eta$.

We prove first the following

**Lemma.**

$$D_k^2 = E\left(\sum_{b \in G}\left(V_k(b) - \frac{2^k}{n}\right)^2\right) = 2^k\left(1 - \frac{1}{n}\right) \tag{1.3}$$

**Proof of the Lemma.** We have clearly

$$D_k^2 = \sum_{b \in G} E(V_k^2(b)) - \frac{2^{2k}}{n} \tag{1.4}$$

Now

$$V_k(b) = \sum_{\varepsilon_1 a_1 + \dots + \varepsilon_k a_k = b} 1 \tag{1.5}$$

where the summation has to be extended over all $2^k$ $k$-tuples $(\varepsilon_1, \cdots, \varepsilon_k)$ of zeros and ones. Thus we obtain

$$\sum_b E(V_k^2(b)) = \sum \sum P(\varepsilon_1 a_1 + \cdots + \varepsilon_k a_k = \varepsilon'_1 a_1 \cdots + \varepsilon'_k a_k) \tag{1.6}$$

where $(\varepsilon_1, \cdots, \varepsilon_k)$ and $(\varepsilon'_1, \cdots, \varepsilon'_k)$ run independently over all $k$-tuples of zeros and ones. For the sake of brevity, let us put $\varepsilon = (\varepsilon_1, \cdots, \varepsilon_k)$, $\varepsilon' = (\varepsilon'_1, \cdots, \varepsilon'_k)$ further $a = (a_1, \cdots, a_k)$, $(\varepsilon, a) = \sum_{j=1}^{k} \varepsilon_j a_j$ and $(\varepsilon', a) = \sum_{j=1}^{k} \varepsilon'_j a_j$. Clearly if $\varepsilon = \varepsilon'$ then $P((\varepsilon,a) = (\varepsilon', a)) = 1$. Now let us suppose that $\varepsilon$ and $\varepsilon$ are not identical. Then there is value of $h$ $(1 \leqq h \leqq k)$ such that $\varepsilon_h \neq \varepsilon'_h$, i.e. either $\varepsilon_h = 1$ and $\varepsilon'_h = 0$, or $\varepsilon_h = 0$ and $\varepsilon'_h = 1$. Then if the values of $\varepsilon_j$ for $j$ different from $h$ are fixed the equation $(\varepsilon, a) = (\varepsilon', a)$ has exactly one solution for $a_h$ and the probability that $a_h$ is equal to this unique value is clearly $\frac{1}{n}$; thus if $\varepsilon \neq \varepsilon'$ we obtain $P((\varepsilon, a) = (\varepsilon', a)) = \frac{1}{n}$. Thus we get

$$\sum_{\varepsilon} \sum_{\varepsilon'} P((\varepsilon, a) = (\varepsilon', a)) = 2^k + \frac{2^k(2^k - 1)}{n} \tag{1.7}$$

which proves our Lemma.

We can deduce from the Lemma immediately

**Theorem 1.** If

$$k \geqq \frac{2 \log n + 2 \log \frac{1}{\varepsilon} + \log \frac{1}{\delta}}{\log 2} \tag{1.8}$$

*where* $\varepsilon > 0$ *and* $\delta > 0$ *are arbitrary small positive numbers then*

$$(1.9) \qquad P\left( \operatorname*{Max}_{b \in G} \left| V_k(b) - \frac{2^k}{n} \right| \leqq \varepsilon \frac{2^k}{n} \right) > 1 - \delta$$

**Proof of Theorem 1.** Clearly

$$(1.10) \qquad \operatorname{Max} \left| V_k(b) - \frac{2^k}{n} \right|^2 \leqq \sum_{b \in G_n} \left( V_k(b) - \frac{2^k}{n} \right)^2$$

and thus by Markoff's well known inequality and the Lemma

$$(1.11) \qquad P\left( \operatorname*{Max}_{b \in G_n} \left| V_k(b) - \frac{2^k}{n} \right| > \varepsilon \frac{2^k}{n} \right) < \frac{n^2}{2^k \varepsilon^2}$$

Thus if (1.8) holds

$$(1.12) \qquad P\left( \operatorname*{Max}_{b \in G_n} \left| V_k(b) - \frac{2^k}{n} \right| > \varepsilon \frac{2^k}{n} \right) < \delta$$

which proves (1.8).

It follows from Theorem 1 that there exists in every Abelian group o order $n$ for each $k \geqq \dfrac{2 \log n + 2 \log \frac{1}{\varepsilon} + \log \frac{1}{\delta}}{\log 2}$ $k$ elements $a_1, \cdots, a_k$ such that each element $b$ of $G_n$ can be represented in the form $b = \varepsilon_1 a_1 + \cdots + \varepsilon_k a_k$ where $\varepsilon_j = 0$ or $1$ $(j = 1, 2, \cdots, k)$ $\dfrac{2^k}{n}(1 + \varepsilon_b)$ times where $|\varepsilon_b| \leqq \varepsilon$.

An interesting special case is obtained if $G_n$ is the additive group of residues mod $n$.

Now we proceed to prove

**Theorem 2.** *For any* $\delta > 0$ *if*

$$(1.13) \qquad k \geqq \frac{\log n + 2 \log \frac{1}{\delta} + \log \frac{\log n}{\log 2}}{\log 2} + 5$$

*then*

$$P\left( \operatorname*{Min}_{b \in G_n} V_k(b) > 0 \right) > 1 - \delta. \tag{1.14}$$

**Proof of Theorem 2.** Put

$$k_1 = \left[ \frac{\log n}{\log 2} \right] + d + 1 \tag{1.15}$$

where $d$ is a positive integer, $d \geqq \dfrac{2 \log \frac{1}{\delta}}{\log 2} + 2$.

Then by Lemma 1, denoting by $N_{k_1}$ the number of elements $b$ of $G_n$ for which $V_{k_1}(b) = 0$, we have

$$E(N_{k_1}) \leqq \frac{n^2}{2^{k_1}}. \tag{1.16}$$

Thus it follows by Markoff's inequality that for any $\lambda > 1$

$$P\left( N_{k_1} > \lambda \frac{n^2}{2^{k_1}} \right) \leqq \frac{1}{\lambda}. \tag{1.17}$$

Let us denote by $A_0$ the event $N_{k_1} \leqq \lambda \dfrac{n^2}{2^{k_1}} \leqq \dfrac{\lambda \cdot n}{2}$. Supposing that $A_0$ holds, we select an element $a_{k_1+1}$ at random. Let $N_{k_1+1}$ denote the number of elements $b$ of $G_n$ for which $V_{k_1+1}(b) = 0$. Clearly $V_{k_1+1}(b) = 0$ if and only if $V_{k_1}(b) = 0$ and $V_{k_1}(b') = 0$ where $b' = b - a_{k_1+1}$; and this has probability $\dfrac{N_{k_1}}{n}$. Thus it follows

$$E(N_{k_1+1} \mid N_{k_1}) = \frac{N_{k_1}^2}{n} \tag{1.18}$$

which implies

$$E(N_{k_1+1} \mid A_0) \leqq \frac{\lambda^2 n}{2^{2d}} \tag{1.19}$$

and thus

$$P\left( N_{k_1+1} > \frac{2\lambda^3 n}{2^{2d}} \right) \leqq \frac{1}{2\lambda} \tag{1.20}$$

Let $A_1$ denote the event $N_{k_1+1} \leqq \frac{2\lambda^3 n}{2^{2d}}$, and let us now suppose that both $A_0$ and $A_1$ hold. Choosing the element $a_{k_1+2}$ at random and repeating the same argument, we obtain

$$E(N_{k_1+2} \mid N_{k_1+1}) = \frac{N_{k_1+1}^2}{n} \tag{1.21}$$

and thus

$$E(N_{k_1+2} \mid A_0 A_1) \leqq \frac{4\lambda^6 n}{2^{4d}}. \tag{1.22}$$

(Here and in what follows the product of events denotes the joint occurence of these events). This implies

$$P\left( N_{k_1+2} > \frac{16\lambda^7 n}{2^{4d}} \,\middle|\, A_0 A_1 \right) \leqq \frac{1}{4\lambda}. \tag{1.23}$$

Let us continue this process; let the elements $a_{k_1+3}, \cdots a_{k_1+j}$ be chosen at random independently and with a uniform distribution in $G_n$. Let in general $A_{k_1+i}$ denote the event

$$N_{k_1+i} \leqq \frac{2^{2^{i}-1} \lambda^{2^{i+1}-1} \cdot n}{2^{2^i d}} = M_i \qquad (i = 0, 1, \cdots, j) \tag{1.24}$$

then we obtain, putting $B_i = A_0 A_1 \cdots A_i \quad (i = 0, 1, \cdots, j)$

$$P(\bar{A}_i \mid B_{i-1}) \leqq \frac{1}{2^i \lambda} \quad (i = 0, 1, \cdots, j). \tag{1.25}$$

Now clearly if $j$ is an integer for which

$$j \geqq \frac{\log \dfrac{\log n}{\log 2}}{\log 2} \tag{1.26}$$

then

$$M_j \leqq \frac{1}{2} n^{(2 \log \lambda / \log 2) + 2 - d} \tag{1.27}$$

On the other hand we have

$$P(\bar{B}_j) = P(\bar{A}_0) + \sum_{i=1}^{j} P(\bar{A}_i \cdot B_{i-1}) \tag{1.28}$$

and thus, in view of $P(CD) = P(C \mid D) P(D) \leqq P(C \mid D)$

$$P(\bar{B}_j) \leqq P(\bar{A}_0) + \sum_{i=1}^{j} P(\bar{A}_i \mid B_{i-1}) \tag{1.29}$$

Thus we obtain that

$$P(\bar{B}_j) \leqq \frac{1}{\lambda} + \frac{1}{2\lambda} + \frac{1}{4\lambda} + \cdots \leqq \frac{2}{\lambda}. \tag{1.30}$$

Now we shall choose

$$\lambda = 2^{(d-2)/2}. \tag{1.31}$$

Then we have $\dfrac{2 \log \lambda}{\log 2} = d - 2$, and thus by (1.27)

$$M_j \leqq \frac{1}{2} < 1. \tag{1.32}$$

It follows that if the event $B_j$ takes place, $N_{k_1 + j} = 0$. Thus if (1.26) holds,

$$P(\min_{b \in G_n} V_{k_1 + j}(b) > 0) \geqq 1 - 2^{1 - ((d-2)/2)} \tag{1.33}$$

and thus if $d \geqq \dfrac{2\log\dfrac{1}{\delta}}{\log\ 2} + 4$ then

$$P(\operatorname*{Min}_{b \in G_n} V_{k_1+j}(b) > 0) \geqq 1 - \delta. \tag{1.34}$$

As $k_1$ is defined by (1.15) and $j$ is an integer for which (1.26) holds, clearly $k = k_1 + j$ is any integer for which (1.13) holds.

This proves Theorem 2.

## §2. Some remarks

In order to obtain some further insight into the problem, it would be useful to compute the moments

$$\mu_r(n,k) = E(\sum_{b \in G_n} [V_k(b)]^r). \tag{2.1}$$

for $r = 3, 4, \cdots$. It is easy to show that

$$\mu_3(n,k) = 2^k + \frac{3.2^k(2^k-1)}{n} + \frac{2^k(2^k-1)(2^k-2)}{n^2}. \tag{2.2}$$

Formula (2.2) follows from the fact that if $\varepsilon^{(1)}$, $\varepsilon^{(2)}$ and $\varepsilon^{(3)}$ are any three different $k$-tuples of zeros and ones

$$P((\varepsilon^{(1)}, a) = (\varepsilon^{(2)}, a) = (\varepsilon^{(3)}, a)) = \frac{1}{n^2}.$$

However

$$P((\varepsilon^{(1)}, a) = (\varepsilon^{(2)}, a) = (\varepsilon^{(3)}, a) = (\varepsilon^{(4)}, a))$$

is not equal to $\frac{1}{n^3}$ for any four different $k$-tuples $\varepsilon^{(1)}, \varepsilon^{(2)}, \varepsilon^{(3)}, \varepsilon^{(4)}$; and this makes the computation of $\mu_r(n,k)$ for $r \geqq 4$ more dificult.

A surprising feature of our results is that they do not depend at all on the structure of the group $G_n$.

Let us mention that both theorems 1 and 2 can be generalized for arbitrary non-Abelian finite groups, in the following way: Let $G_n$ be any group of order $n$, let us choose the elements $a_1, \cdots, a_k$ of $G_n$ at random (with uniform distribution) and independently. Let $V_k(b)$ denote the number of representations of $b$ in the form $b = a_{i_1} a_{i_2} \cdots a_{i_r}$ (we now write the group-operation as multiplication) where $1 \leqq i_1 < i_2 < \cdots < i_r \leqq k$ and $0 \leqq r \leqq k$ (an empty product denotes the unity element). Then the statements of Theorems 1 and 2 are valid.

However if we consider all possible products of different elements $a_{i_1} a_{i_2} \cdots a_{i_r}$ which can be formed from the elements $a_1, a_2, \cdots, a_k$ chosen at random, and do not consider only such products in which $i_1 < i_2 < \cdots < i_r$, then the situation changes completely. In this case the structure of the group $G_n$ becomes relevant. We feel, that the supposition that only such products formed from the elements $a_1, \cdots, a_n$ chosen at random should be considered in which $a_i$ always preceeds $a_j$ if both occur and $i < j$, is unnatural. This is the reason why we restricted ourselves in § 1. to formulate our theorems for the case of Abelian groups.

## References


1. A. Rényi, Probabilistic methods in number theory, *Proceedings of the International Congress of Mathematicians*, Edinburgh, 1958, p. 529–539.

2. P. Erdös-A. Rényi, On random graphs, I. *Publicationes Mathematicae (Debrecen)* 6/1959/290–297.

3. P. Erdös-A. Rényi, On the evolution of random graphs, International Statistical Institute, 32. Session, Tokyo, 1960, 119.1–5.

4. P. Erdös-A. Rényi, On the evolution of random graphs, *MTA Mat. Kut. Int. Közleményei* 5/1960/17–61.

5. P. Erdös-A. Rényi, On the strength of connectedness of a random graph, *Acta Math. Acad. Sci. Hung.* 12/1961/261–267.

6. A. Rênyi, On random generating elements of a finite Boolean algebra, *Acta Sci. Math. Szeged*, 22/1961/75–81.

7. P. Turán, On a theorem of Hardy and Ramanujan, *Journal of the London Math. Soc.* 9/1934/274–176.

8. Yu. V. Linnik, The dispersion method in binary additive problems /in Russian/, University of Leningrad, 1961, 1–208.

MATHEMATICAL INSTITUTE
HUNGARIAN ACADEMY OF SCIENCES
BUDAPEST, HUNGARY

(Received October 31, 1963)