CHAPTER 1

Probabilistic Methods in Combinatorial Mathematics

ALFRED RÉNYI, *Mathematical Institute of the Hungarian Academy of Sciences*

1. THE RENAISSANCE OF COMBINATORIAL MATHEMATICS

It is beyond doubt that we are witnesses of a renaissance of combinatorial problems and methods in mathematics. This process started slowly and became evident only gradually in the last two decades, but its origins can be traced back to the 1920's. Without exaggeration one can say that all main branches of mathematics contributed to some extent to this development.

Probability theory and *statistics* was even at the time of stagnation of combinatorial mathematics in the nineteenth century and the beginning of our century the main source of problems and the main consumer of results of a combinatorial character. The rapid development of probability theory, which has been going on with increasing speed since the 1930's, meeting successfully the constantly growing challenge of its applications in practically every branch of human knowledge, was also the source of many problems of a combinatorial character. New types of combinatorial problems were raised in statistics, concerning the design of experiments and in the theory of order statistics. Another source of combinatorial problems was the arc-sine law, which led to the development of an important combinatorial theory[1]. Recently information theory became a rich source of combinatorial problems of a quite new type.

As a second source of the revival of interest in combinatorial mathematics, *graph theory* has to be mentioned. Before the second world war Hungary was one of the few countries where graph theory was taken seriously. The general opinion of mathematicians before the war is reflected in that some mathematicians called graph theory the "slums" of topology. However, in the years after the war it has been quickly realized all over the world that graph theory is a basic and independent chapter of combinatorial mathematics, having important applications, e.g., in operations research, in chemistry, in statistical physics, etc.. (Some such applications will be mentioned in what follows.)

A main cause of the revival of interest in combinatorial mathematics was the recent development of *numerical analysis*. The appearance of modern high speed computers led to a strong shift to finite mathematics in general, and thus to combinatorics, in particular.

Besides these main sources, stimulation came also from *algebra* (finite groups, Galois fields, matrices, lattices, etc.), *geometry* (finite geometries, discrete geometry), *number theory*, (difference-bases, combinatorial methods of P. Erdös), *set theory, topology* and *mathematical logic* (Sperner's theorem, Ramsey's therem) and also from *statistical mechanics* (Ising models), *genetics*, etc.. The wide spectrum of modern combinatorial mathematics and its various applications is exhibited by [1].

2. GENERAL METHODS AND SCOPE OF COMBINATORIAL MATHEMATICS

Characteristic for the present state of affairs of combinatorial mathematics is the wealth of particular problems and the lack of a systematic theory. Only a small number of basic methods are available. Among these the method of generating functions is to be mentioned first, which also plays a role in several other methods: in the counting method of G. Polya, generalized by N. G. de Bruijn, in symbolic methods, in the

[1] These problems were discussed in detail at a meeting held in Aarhus in 1962.

operator-method of G. C. Rota, etc.. The application of algebraic tools (finite fields, group theory, the Pfaffian, etc.) can also be considered as one of the general methods in combinatorial mathematics; in the analysis of networks, methods and concepts of Boolean algebra and mathematical logic are successfully applied. Another general method in combinatorial mathematics is the application of probability theory as a tool. We shall deal with this method in detail in the following sections of the present paper. In spite of the availability of these general methods of approach, if one is confronted with a particular problem of an unusual type, usually one has to attack it in an *ad hoc* way, and there is no general theory which would guide in the solution. The main reason for this seems to be that the field of combinatorial mathematics is still far from being systematized.

It is rather difficult to define at all what combinatorial mathematics really is, and to tell what its main chapters are, or what basic types of problems can be distinguished.

One possible and often used definition can be formulated as follows. Combinatorial mathematics is the theory of finite sets: it deals with relations and functions, and further, with sets of functions defined on finite sets, especially with problems of enumeration, construction and existence. It should be added that there are certainly some results concerning infinite sets (e.g., infinite graphs) which are undoubtedly of a combinatorial character. Thus in this respect the above definition seems to be a bit too restricted. However, in another respect the definition is not enough restricted, as clearly not every investigation concerning finite sets is of combinatorial character. For instance, though the study of the symmetric group of all permutations of a finite set is certainly a part of combinatorial mathematics, and in spite of the fact that every finite group is isomorphic to a subgroup of a symmetric group, nevertheless the main body of the theory of finite groups belongs to algebra and is not of combinatorial character.

Somewhat vaguely, problems concerning finite sets which are of a combinatorial character can be characterized by the fact that these problems are usually independent of the labelling of the elements of the basic set, i.e. invariant under any permutation of these elements. This is a common property of most (but not all) of those problems which we feel to be of combinatorial character. Problems in which the elements of the basic finite set have individual character and the set possesses an algebraic structure, belong usually to some other field of mathematics, e.g. algebra, number theory, etc.; nevertheless such problems may also exhibit combinatorial features.

3. PROBABILITY THEORY AS A TOOL OF COMBINATORIAL MATHEMATICS

Most classical results of combinatorial mathematics have been developed—since the time of Pascal and Fermat—with the aim of being applied in probability theory. However this relation between probability theory (as the field of application) and combinatorics (as a tool) can be reversed, and this led to interesting and even surprising results. Many important combinatorial problems can be attacked only in this way, or at least this is the easiest way of approach. For instance, in order to prove that combinatorial objects having certain properties exist, often the only available method is to show that by introducing a probability measure in a certain (finite) set of combinatorial objects, the subset of objects possessing the required properties has a positive probability and thus cannot be empty. This method of proof of course does not lead to an effective construction, only to *proofs of existence.* However, often such a proof shows not only that there exist objects of the required type but also that most of the objects of a certain set have the required properties, i.e., it is an exception if an object of the given type does not have the required properties.

As a matter of fact, the probabilistic approach is particularly useful if one wants to study *typical properties* of members of a class of combinatorial objects (like graphs, matrices, partitions, permutations, etc.), i.e., properties possesed by a majority of objects belonging to the class considered, as contrasted with properties which are possessed only by a negligibly small minority of these objects. In some cases in this way *problems of enumeration* can be solved also, at least asymptotically. In the following sections we shall illustrate these statements with some examples.

The aim of this paper is to present a few characteristic examples for the different ways in which probabilistic methods can be used in combinatorial mathematics. No attempt is made to achieve completeness in any respect and the selection of examples is of course strongly biased in favor of examples connected with my own work, done mainly in collaboration with P. Erdös.

4. PROBABILISTIC EXISTENCE PROOFS

One of the earliest examples of the application of probabilistic ideas in graph theory is the following theorem of P. Erdös [3]:

For every sufficiently large n there exist graphs G_n, having n vertices, such that neither G_n nor its complementary graph $\bar{G}$ contains a complete subgraph with more than $2\log n/\log 2$ vertices.

The proof can be told in two different ways: its probabilistic nature can be disguised or be emphasized. We shall choose the second way, and prove slightly more than stated above, namely the following:

For every fixed p with $0 < p < 1$ and for every $n \geqq n_0(p)$ there exists a graph G_n having n vertices such that G_n does not contain a complete graph with more than $2\log n/\log(1/p)$ vertices and $\bar{G}_n$ does not contain a complete graph of more than $2\log n/\log(1/(1-p))$ vertices.

This can be proved as follows: Let H_n be a set having n elements. Let Γ_n be the random graph obtained by connecting any pair of points of H_n by an edge with probability p, independently for each pair. Let $\bar{\Gamma}_n$ denote the complementary graph of Γ_n (i.e., two points P and Q of H_n are connected by an edge in $\bar{\Gamma}_n$ iff they are not connected in Γ_n). Let $v(G, k)$ denote the number of complete sub-graphs of order k in the graph G and let $\mathbf{E}_p$ denote the expectation with respect to the probability measure introduced. We shall show that for sufficiently large n,

$$(1)\qquad \mathbf{E}_p\left(v\left(\Gamma_n, \left[\frac{2\log n}{\log 1/p}\right]+1\right)\right. \\ \left. + v\left(\bar{\Gamma}_n, \left[\frac{2\log n}{\log 1/(1-p)}\right]+1\right)\right) < 1 .$$

As a matter of fact

$$\mathbf{E}_p(v(\Gamma_n, k)) = \binom{n}{k} p^{\binom{k}{2}} < n^k p^{\binom{k}{2}}/k!$$

and thus

$$\mathbf{E}_p\left(v\left(\Gamma_n, \left[\frac{2\log n}{\log 1/p}\right] + 1\right)\right) = \sigma(1)\,. \tag{2}$$

Similarly we obtain

$$\mathbf{E}_p\left(v\left(\bar{\Gamma}_n, \left[\frac{2\log n}{\log 1/(1-p)}\right] + 1\right)\right) = \sigma(1)\,, \tag{3}$$

because

$$\mathbf{E}_p(v(\Gamma_n, k)) = \mathbf{E}_{1-p}(v(\bar{\Gamma}_n, k))\,,$$

and thus (3) follows from (2), replacing p by $1-p$. Thus, for sufficiently large n, (1) holds, i.e. there exists at least one graph G_n having n vertices, for which

$$v\left(G_n, \left[\frac{2\log n}{\log 1/p}\right] + 1\right) + v\left(\bar{G}_n, \left[\frac{2\log n}{\log 1/(1-p)}\right] + 1\right) = 0$$

as was to be proved.

Clearly for $p = 1/2$ we obtain the theorem of Erdös mentioned above. It should be added that no method of construction is known for the graphs, the existence of which was shown above.

As another, more involved, example of the same type let us mention the probabilistic proof [8] of the fact that most of the graphs with n vertices are, for large values of n, almost as asymmetric as possible.

We consider only non-directed graphs without multiple edges and without loops. We call such a graph *symmetric*, if there exists a non-identical permutation of its vertices which leaves the graph invariant. In other words a graph is called symmetric if the group of its automorphisms has order greater than 1. A graph which is not symmetric will be called *asymmetric*. The degree of symmetry of a symmetric graph is evidently measured by the order of its group of automorphisms. The question which led us to the results mentioned is the following: how can we measure the degree of asymmetry of an asymmetric graph?

Evidently any asymmetric graph can be made symmetric by deleting certain of its edges and by adding certain new edges connecting its vertices. We shall call such a transformation of the graph its *symmetrization*. For each symmetrization of the graph let us take the sum of the number of deleted edges—say r—and the number of new edges—say s. It is reasonable to define the degree of asymmetry $A[G]$ of a graph G, as the minimum of $r + s$ where the minimum is taken over all possible symmetrizations of the graph G. Clearly the asymmetry of a symmetric graph is according to this definition equal to 0, while the asymmetry of any asymmetric graph is a positive integer.

The question arises: how large can the degree of asymmetry be for a graph of order n? We shall denote by $\mathbf{A}(n)$ the maximum of $A[G]$ for all graphs G of order n.

We have shown that the asymmetry of a graph of order n cannot exceed $(n - 1)/2$ if n is odd, while if n is even the asymmetry cannot exceed $(n/2) - 1$; further, that this estimate is asymptotically best possible, that is, for any $\varepsilon > 0$ there can be found an integer n_0 such that for any $n \geqq n_0$ there exists a graph G_n of order n for which $A[G_n] > (n/2)\cdot(1 - \varepsilon)$. In other words,

$$\lim_{n\to\infty} \frac{\mathbf{A}(n)}{n} = \frac{1}{2} .$$

Our proof is not constructive, only a proof of existence. It uses probabilistic considerations. This method gives, however, more than stated above: it shows that for large values of n most graphs of order n are asymmetric, the degree of asymmetry of most of them being larger than $(n/2)\cdot(1 - \varepsilon)$, where $\varepsilon > 0$ is arbitrary.

I would like to add that no method is known to me to construct effectively a graph G_n with degree of asymmetry k for any k, even if no restriction on the number n of vertices is made.

As a third example for a probabilistic proof of existence I mention the proof of the existence of codes with given properties by the method of "random codes" (see e.g. [26]). The basic idea of this method—due to Shannon—is similar to that of the proof of Theorem 1; an additional complication arises because one has to distinguish between the average error and the maximal error of a code.

5. PROBABILISTIC PROOF OF TYPICAL PROPERTIES AND ASYMPTOTIC ENUMERATION PROBLEMS

We start with the following simple example: in a tournament in which every player plays against every other exactly once and no game can end in a tie, the expected number of cyclic triples is exactly equal to one fourth of the total number of triples. (See e.g. [15]). This can be shown simply as follows: select any three players A, B and C. We can suppose that B has defeated A. In this case the triple A, B, C is cyclic if and only if C has defeated B and A has defeated C. If the chances of every game are equal (i.e. $1/2$) for each player, and independent from the outcome of every other game, then the probability of C defeating B and A defeating C is equal to $1/4$, which proves our assertion. It should be noted that it has been shown by Kendall and Smith that the maximal number of cyclic triples is asymptotically the same: if the number of players is denoted by n, the maximum is $(n^3 - n)/24$ if n is odd and $(n^3 - 4n)/24$ if n is even; thus the result can be stated as follows: "in most tournaments the number of cyclic triples is asymptotically maximal."

Many results of this type have been proved by P. Erdös and A. Rényi, in the theory of the evolution of random graphs ([4], [5], [6], [7], [12]). We mention here the first of these results: if n is large most graphs having n vertices and N edges are connected provided that $(2N - n \log n)/2n$ is a large positive number, while only a small minority of such graphs is connected if $(2N - n \log n)/2n$ is a large negative number. More recently we have shown that in case $(2N - n \log n)/2n$ is a large positive number the majority of graphs having n vertices and N edges contains a factor of degree 1 (see [12]). The theory of the evolution of random graphs has been recently applied also in chemistry, see e.g. [2].

Another recent result, due to A. Rényi and G. Szekeres [25], states that the order of magnitude of the diameter of most trees of order n is $\sqrt{n}$. Such results can usually be sharpened to results concering asymptotic distributions. For instance the result on random trees containing the above mentioned statement is as follows: let us consider the set $\mathcal{T}_n$ of all labelled trees of order n, with vertices $P_1, P_2, \ldots, P_n$ and let $h(T_n)$ denote the height of the tree $T_n \in \mathcal{T}_n$ over the point P_1 (i.e., the length of the longest path in T_n starting in P_1). Let $F_n(x)$ denote the probability that choosing at random a tree T_n

from the set $\mathcal{T}_n$ (with uniform distribution) one has $h(T_n) < x\sqrt{2n}$. Then the limit distribution

$$\lim_{n\to\infty} F_n(x) = F(x)$$

exists and is given by the formula (1.c [25])

$$F(x) = \frac{4\pi^{5/2}}{x^3}\sum_{n=1}^{\infty} n^2 e^{-n^2\pi^2/x^2} \qquad (0 < x < +\infty) .$$

Hence the expectation of $h(T_n)$ is $\sim\sqrt{2\pi n}$.

Note that in this problem it is trivial that $h(T_n)$ can take on every value between 1 and $n-1$; the question is not one of existence, but one of (asymptotic) enumeration.

Let us mention in this direction a recent result of J. Komlós [18]: the great majority of all n by n zero-one matrices are non-singular if n is large, and the same holds for matrices with elements ± 1. (Note that though the two questions are closely related they are not identical: the probability of a random n by n matrix with elements ± 1 being non-singular is the same as the probability of a random n by n zero-one matrix being non-singular provided all elements of its first row are equal to 1).

A similar result (see [10]) which we obtained with Erdös is that most of the n by n zero-one matrices containing N ones and $n^2 - N$ zeros have a positive permanent provided that $c = (N - n\log n)/n$ is large. More exactly we have proved that if we select one among the $\binom{n^2}{N}$ such matrices at random (with uniform distribution), the probability of its permanent being positive is $\sim e^{-2e^{-c}}$. We have also proved that if $(N - n\log n - (k-1)n\log\log n)/n$ is large then the permanent of the majority of such matrices is $> k$ for $k = 1, 2, \ldots$.

Recently P. Erdös and P. Turán have studied random permutations (see [11]). They have shown that if $O(\Pi_n)$ denotes the order of a permutation Π_n of n elements, then except for $\sigma(n!)$ permutations $\log O(\Pi_n)$ lies between the limits $-(1/2)\log^2 n \pm (\log n)^{3/2+\varepsilon}$ where $\varepsilon > 0$ is arbitrary small.

The proof of these statements consists of two parts. Let α_k denote the number of cycles of length k in the permutation Π_n. Then clearly $O(\Pi_n)$ is equal to the least common multiple of those numbers k for which $\alpha_k \geqq 1$. Now it is proved in [11] by a number theoretic argument (this is the hard part of the proof) that for the majority of the permutations $\log O(\Pi_n)$ is not too far from $\sum \alpha_k \log k = S_n$. On the other hand it is easy to show that the mean value of S_n is asymptotically equal to $-(1/2)\log^2 n$ and its variance to $(1/3)\log^3 n$; from this the result follows by Chebyschev's inequality. More exactly we have

$$\mathbf{E}(S_n) = \sum_{k \leqq n} \frac{\log k}{k} \tag{4}$$

and

$$D^2(S_n) = \sum_{k \leqq n} \frac{\log^2 k}{k} - \sum_{\substack{k+l>n \\ k \leqq n,\, l \leqq n}} \frac{\log k \log l}{kl} \,. \tag{5}$$

Formulas (4) and (5) can be deduced from the following remark, due to L. A. Shepp and S. P. Lloyd: if Π_n is chosen at random with uniform distribution on all $n!$ permutations of order n, the joint probability distribution of $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ is the same as the conditional joint distribution of the independent random variables $\beta_1, \beta_2, \ldots, \beta_n$ such that β_k has Poisson distribution with parameter z^k/k $(0 < z < 1;\ k = 1, 2, \ldots, n)$ subject to the condition $\sum_{k=1}^{n} k\beta_k = n$.

Erdös and Turán have proved also in another paper (in print) that $\log O(\Pi_n)$ is asymptotically normally distributed with mean $\mathbf{E}(S_n)$ and variance $D^2(S_n)$. To prove this, besides the number theoretic estimations mentioned, a possible starting point is the following explicit formula for the characteristic function of the random variable S_n:

$$\mathbf{E}(e^{itS_n}) = (1 - z) \exp\left[\sum_{k=1}^{\infty} \frac{z^k}{k^{1-it}}\right] . \tag{6}$$

(6) follows also from the mentioned remark of Shepp and Lloyd.

6. OTHER USES OF PROBABILISTIC METHODS

Probabilistic methods can also be applied in the theory of search; it turns out that under certain conditions a random strategy of search may be asymptotically almost as good as the (usually much more complicated) best systematic strategy (see [21], [22], [23], [24]).

In other cases the number of operations to be carried out when applying a random strategy of search is larger by a constant factor only, compared with the corresponding number for the best systematic strategy. This is the case, for example, for the much discussed problem of separation of the fair and the counterfeit coins (see [9] and [19]) and for a related problem where the factor is $\log_2 3 = 1.58\ldots$.

We have proved that the random search algorithm needs $(\log_2 9)\cdot n/\log_2 n$ weighings, while Lindström has shown that the best systematic algorithm consists of $2n/\log_2 n$ steps only.

It should be mentioned that in the asymptotic evaluation of combinatorial results, the analytic methods developed under the influence and in close connection with probability theory may be successfully used even without a probabilistic interpretation. In this context I would like to call attention to the results of W. K. Hayman [17], who has shown that the moduli of the terms around the maximal term of the power series of an entire function are asymptotically normally distributed provided the function belogs to a certain rather wide class- called by Hayman the class of *admissible* functions. This class contains, e.g., the function e^{c^x}; in which case one gets as a particular case of Hayman's result an asymptotic formula for the number $T(n)$ of partitions of a set of n elements, namely the formula

$$T(n) \sim \frac{\exp\left(n\left(\tau_n + \frac{1}{\tau_n} - 1\right) - 1\right)}{\sqrt{\log n}}$$

where τ_n is defined as the positive root of the equation $\tau_n e^{\tau_n} = x$ (Compare [20]).

Another way to deduce this result is by the following probabilistic interpretation of the Bell-numbers $T(n)$: let $x_1, x_2, \ldots, x_n \ldots$ be independent random variables each having a Poisson distribution with mean value 1. Let ν be a random variable, independent from the x_ks, having a Poisson distribution with mean value e. Then the random variable $y = x_1 + x_2 + \ldots + x_\nu$ has the distribution

$$P(y = n) = \frac{T(n)}{n!} e^{1-e} \qquad (n = 0, 1, \ldots)\,.$$

The asymptotically normal distribution of many combinatorial functions has been proved by V. Gončarov [14] without a direct probabilistic interpretation. Those results where a probabilistic interpretation is found are more elegant, like that given by W. Feller [13] concerning the asymptotically normal distribution of the number of inversions and cycles of a random permutation and that given recently by L. H. Harper [16] on Stirling numbers.

We did not discuss in this paper the probabilistic proofs of combinatorial identities, because it has long been well known that this method of proof is available for a large number of identities.

References

1. Beckenbach, E. F. (ed.) *Applied Combinatorial Mathematics*, John Wiley and Sons, New York, 1964.
2. Bruneau, C. M. "Theorie des graphes Stochastiques .appliquée à la synthèse et à la degradation aléatoires de composés macromoleculaires multifonctionnelles," Ph. C. Thesis, Paris, 1966, pp. 1–32.
3. Erdös, P. "Some Remarks on the Theory of Graphs," *Bull. Amer. Math. Soc.*, **53** (1947), 292–294.
4. Erdös, P. and Rényi, A. "On Random Graps I," *Publicationes Mathematicae (Debrecen)*, **6** (1959), 290–297.
5. Erdös, P. and Rényi, A. "On the Evolution of Random Graphs," *Publ. Math. Inst. Hung. Acad. Sci.*, **5A** (1960), 17–61.
6. Erdös, P. and Rényi, A. "On the Strength of Connectedness of a Random Graph," *Acta Math. Acad. Sci. Hung.*, **12** (1961), 261–267.
7. Erdös, P. and Rényi, A. "On the Evolution of Random Graphs," *Bull. Inst. Internat. Statist.*, **38** (1961), 343-347.
8. Erdös, P. and Rényi, A. "Asymmetric Graphs," *Acta Math. Acad. Sci. Hung.*, **14** (1963), 297–315.
9. Erdös, P. and Rényi, A. "On Two Problems of Information Theory," *Publ. Math. Inst. Hung, Acad. Sci.*, **8** (1963), 241.
10. Erdös, P. and Rényi, A. "On Random Matrices," *Publ. Math. Inst. Hung. Acad. Sci.*, **8** (1963), 455–461.
11. Erdös, P. and Turán, P. "On some Problems of a Statistical Group Theory I," *Z. Wahrscheinlichkeitstheorie und Vers. Gebiete*, **4** (1965), 175–186.
12. Erdös, P. and Rényi, A. "On the Existence of a Factor of Degree One of a Connected Random Graph," *Acta Math. Acad. Sci. Hung.*, **17** (1966), 359–368.
13. Feller, W. *An Introduction to Probability Theory and Its Applications*, Vol. 1, John Wiley and Sons, New York, 1957, Ch. V, Sec. 6.
14. Gončarov, V. "On the Field of Combinatory Analysis," *Izvestia*

Akad. Nauk Ser. Math., 8 (1944), 3–48; see also Amer. Math. Soc. Translation Series, 1955.

15. Harary, F., Norman, R. Z. and Cartwright, D. *Structural Models*, John Wiley and Sons, New York, 1965.
16. Harper, L. H. "Stirling Behaviour is Asymptotically Normal," to be published.
17. Hayman, W. K. "A Generalization of Stirling's Formula," *J. reine und angew. Math.*, **196** (1956), 67–95.
18. Komlós, J. "On the Determinant of (0, 1) Matrices," to appear in *Studia Sci. Math. Hung.*
19. Lindström, B. "On a Combinatory Detection Problem," *Publ. Math. Inst. Hung. Acad. Sci.*, **9** (1964), 195–207.
20. Moser, L. and Wyman, M. "An Asymptotic Formula for the Bell Numbers," *Trans. Roy. Soc. Can.*, **49** (1955), 49–53.
21. Rényi, A. "On Random Generating Elements of a Finite Boolean Algebra," *Acta Sci. Math.* (*Szeged*), **22** (1961), 75–81.
22. Rényi, A. "On a Problem of Information Theory," *Publ. Math. Inst. Hung. Acad. Sci.*, **6** (1961), 505–516.
23. Rényi, A. "Statistical Laws of Accumulation of Information," *Bull. Inst. Internat. Statist.*, **39** (1962), 311–316.
24. Rényi, A. "On the Theory of Random Search," *Bull. Amer. Math. Soc.*, **71** (1965), 809–828.
25. Rényi, A. and Szekeres, G. "On the Height of Trees," to appear in *Austral. J. Math.*
26. Wolfowitz, J. *Coding Theorems of Information Theory*, 2nd ed., Springer, 1964, p. 96.