Since $Y \leq m$ we conclude that

$$\Pr(Y \geq m^{1-t\delta^2/2}) \geq m^{-t\delta^2/2} \ . \tag{5.3}$$

One can check that for $t = \lfloor 1+1/(\delta-\delta^2/4-\delta^3/2) \rfloor$, $m^{1-t\delta^2/2} > 2^{n/2}$ and the right-hand side of (5.3) is greater than the right-hand side of (5.2). Thus, with positive probability, $|A_1 \cup A_2 \cup \ldots \cup A_t| > n/2$ and still this union is disjoint to more than $2^{n/2}$ members of F. This contradiction implies inequality (5.1). □

5. Disjoint pairs.

The probabilistic method is most striking when it is applied to prove theorems whose statement does not seem to suggest at all the need for probability. Most of the examples given in the previous sections are simple instances of such statements. In this section we describe a (slightly) more complicated result, due to Alon and Frankl (1985), which solves a conjecture of Daykin and Erdös.

Let $\mathcal{F}$ be a family of m distinct subsets of $X = \{1, 2, \ldots, n\}$. Let $d(\mathcal{F})$ denote the number of disjoint pairs in F, i.e.,

$$d(\mathcal{F}) = \left|\left\{(F, F') : F, F' \in \mathcal{F} \ , \ F \cap F' = \emptyset\right\}\right| \ .$$

Daykin and Erdös conjectured that if $m = 2^{\left(\frac{1}{2}+\delta\right)n}$, then, for every fixed $\delta > 0$, $d(\mathcal{F}) = o(m^2)$, as n tends to infinity. This result follows from the following theorem, which is a special case of a more general result.

Theorem 5.1. *Let $\mathcal{F}$ be a family of $m = 2^{\left(\frac{1}{2}+\delta\right)n}$ subsets of $X = \{1, 2, \ldots, n\}$, where $\delta > 0$. Then*

$$d(\mathcal{F}) < m^{2-\frac{\delta^2}{2}} \ . \tag{5.1}$$

Proof: Suppose (5.1) is false and pick independently t members $A_1, A_2, \ldots, A_t$ of $\mathcal{F}$ with repetitions at random, where t is a large positive integer, to be chosen later. We will show that with positive probability $|A_1 \cup A_2 \cup \ldots \cup A_t| > n/2$ and still this union is disjoint to more than $2^{n/2}$ distinct subsets of X. This contradiction will establish (5.1). In fact

$$\begin{aligned} \Pr\left(|A_1 \cup A_2 \cup \ldots \cup A_t| \le n/2\right) &\le \sum_{S \subset X, |S| \le n/2} \Pr(A_i \subset S, i = 1, \ldots, t) \\ &\le 2^n (2^{n/2}/2^{((1/2)+\delta)n})^t = 2^{n(1-\delta t)} \ . \end{aligned} \tag{5.2}$$

Define

$$v(B) = \left|\{A \in \mathcal{F} : B \cap A = \emptyset\}\right| \ .$$

Clearly

$$\sum_{B \in \mathcal{F}} v(B) = 2d(\mathcal{F}) \ge 2m^{2-\delta^2/2} \ .$$

Let Y be a random variable whose value is the number of members $B \in \mathcal{F}$ which are disjoint to all the A_i $(1 \le i \le t)$. By the convexity of z^t the expected value of Y satisfies

$$E(Y) = \sum_{B \in \mathcal{F}} \left(v(B)/m\right)^t = \frac{1}{m^t} \cdot m \left(\frac{\sum v(B)^t}{m}\right) \ge \frac{1}{m^t} \cdot m \left(\frac{2d(\mathcal{F})}{m}\right)^t > 2m^{1-t\delta^2/2} \ .$$

Theorem 3.3. *If* $\mathcal{F} = \{(A_i, B_i)\}_{i=1}^{h}$ *is a* (k,ℓ)*-system then* $h \leq \binom{k+\ell}{k}$.

Proof: Put $X = \bigcup_{i=1}^{h} (A_i \cup B_i)$ and consider a random order π of X. For each i, $1 \leq i \leq k$, let X_i be the event that all the elements of A_i precede all those of B_i in this order. Clearly $\mathrm{Prob}(X_i) = 1/\binom{k+\ell}{k}$. It is also easy to check that the events X_i are pairwise disjoint. Indeed, assume this is false and let π be an order in which all the elements of A_i precede those of B_i and all the elements of A_j precede those of B_j. Without loss of generality we may assume that the last element of A_i does not appear after the last element of A_j. But in this case, all elements of A_i precede all those of B_j, contradicting the fact that $A_i \cap B_j \neq \emptyset$. Therefore, all the events X_i are pairwise disjoint, as claimed. It follows that $1 \geq \Pr\left(\bigcup_{i=1}^{h} X_i\right) = \sum_{i=1}^{h} \Pr(X_i) = h \cdot 1/\binom{k+\ell}{k}$, completing the proof. □

Theorem 3.3 is sharp, as shown by the family $\mathcal{F} = \{(A, X \setminus A) : A \subset X \ , \ |A| = k\}$, where $X = \{1, 2, \ldots, k+\ell\}$.

4. Combinatorial Number Theory.

A subset A of an abelian group G is called *sum-free* if $(A + A) \cap A = \emptyset$, i.e., if there are no $a_1, a_2, a_3 \in A$ such that $a_1 + a_2 = a_3$.

Theorem 4.1 (Erdös (1965)). *Every set* $B = \{b_1, \ldots, b_n\}$ *of* n *nonzero integers contains a sum-free subset* A *of size* $|A| > \frac{1}{3}n$.

Proof: Let $p = 3k + 2$ be a prime, which satisfies $p > 2 \max_{1 \leq i \leq n} |b_i|$ and put $C = \{k+1, k+2, \ldots, 2k+1\}$. Observe that C is a sum-free subset of the cyclic group Z_p and that $\frac{|C|}{p-1} = \frac{k+1}{3k+1} > \frac{1}{3}$. Let us choose at random an integer x, $1 \leq x < p$, according to a uniform distribution on $\{1, 2, \ldots, p-1\}$, and define $d_1, \ldots, d_n$ by $d_i \equiv x b_i (\mathrm{mod}\, p)$, $0 \leq d_i < p$. Trivially, for every fixed i, $1 \leq i \leq n$, as x ranges over all numbers $1, 2, \ldots, p-1$, d_i ranges over all nonzero elements of Z_p and hence $\Pr(d_i \in C) = \frac{|C|}{p-1} > \frac{1}{3}$. Therefore, the expected number of elements b_i such that $d_i \in C$ is more than $n/3$. Consequently, there is an x, $1 \leq x < p$ and a subsequence A of B of cardinality $|A| > \frac{n}{3}$, such that $xa (\mathrm{mod}\, p) \in C$ for all $a \in A$. This A is clearly sum-free, since if $a_1 + a_2 = a_3$ for some $a_1, a_2, a_3 \in A$ then $xa_1 + xa_2 \equiv xa_3 (\mathrm{mod}\, p)$, contradicting the fact that C is a sum-free subset of Z_p. This completes the proof. □

In Alon-Kleitman (1988) it is shown that the constant $\frac{1}{3}$ in Theorem 4.1 cannot be replaced by $\frac{12}{29}$ (or any bigger constant). The best possible constant is not known.

for notational convenience. Now let $S_1, \ldots, S_m$ be uniformly and independently chosen n-sets, m to be determined. For each coloring χ let A_χ be the event that none of the S_i are monochromatic. By the independence of the S_i

$$Pr(A_\chi) \leq (1-p)^m$$

There are 2^v colorings so

$$Pr(\cup_\chi A_\chi) \leq 2^v(1-p)^m$$

When this quantity is less than 1 there exist $S_1, \ldots, S_m$ so that no A_χ holds, i.e., $S_1, \ldots, S_m$ is not 2-colorable - and hence $m(n) \leq m$.

The asymptotics provide a fairly typical example of those encountered when employing the probabilistic method. We first use the inequality $1 - p \leq e^{-p}$. This is valid for all positive p and the terms are quite close when p is small. When

$$m = \lceil \frac{v \ln 2}{p} \rceil$$

then $2^v(1-p)^m < 2^v e^{-pm} \leq 1$ so $m(n) \leq m$. Now we need find v to minimize v/p. We may interpret p as the probability of picking n white balls from an urn with $v/2$ white and $v/2$ black balls, sampling without replacement. It is tempting to estimate p by 2^{-n+1}, the probability for sampling with replacement. This approximation would yield $m \sim v2^{n-1}(\ln 2)$. As v gets smaller, however, the approximation becomes less accurate and, as we wish to minimize m, the tradeoff becomes essential. We use a second order approximation

$$p = \frac{2\binom{v/2}{n}}{\binom{v}{n}} = 2^{1-n} \prod_{i=0}^{n-1} \frac{v-2i}{v-i} \sim 2^{1-n} e^{-n^2/2v}$$

as long as $v \gg n$, estimating $\frac{v-2i}{v-i} \sim 1 - \frac{i}{v} \sim e^{-i/v}$. Elementary calculus gives $v = n^2/2$ for the optimal value. The evenness of v may require a change of at most 2 which turns out to be asymptotically negligible. This yields the following result of Erdős (1964):

Theorem 3.2.

$$m(n) < (1 + o(1)) \frac{e \ln 2}{4} n^2 2^n$$

.

Let $\mathcal{F} = \{(A_i, B_i)\}_{i=1}^h$ be a family of pairs of subsets of an arbitrary set. We call $\mathcal{F}$ a (k,ℓ)-*system* if $|A_i| = k$ and $|B_i| = \ell$ for all $1 \leq i \leq k$, $A_i \cap B_i = \emptyset$ and $A_i \cap B_j \neq \emptyset$ for all $1 \leq i, j \leq k$. B. Bollobás (1965) proved the following result, which has many interesting extensions and applications.

separates u_1 from u_i. Each of these problems can be solved by solving a standard network flow problem in time $O(n^{8/3})$, (see, e.g., Tarjan (1983).) By Lemma 2.3 the edge connectivity of G is simply the minimum between δ and $\min_{2 \le i \le k} s_i$. The total time of the algorithm is $O(n^{8/3} \log n)$, as claimed.

3. Combinatorics.

A *hypergraph* is a pair $H = (V, E)$, where V is a finite set whose elements are called *vertices* and E is a family of subsets of V, called *edges*. It is *n-uniform* if each of its edges contains precisely n vertices. We say that H has *property B*, or that it it *2-colorable* if there is a 2-coloring of v such that no edge is monochromatic. Let $m(n)$ denote the minimum possible number of edges of an n-uniform hypergraph that does not have property B.

Proposition 3.1 (Erdös (1963b)). *Every n-uniform hypergraph with less than 2^{n-1} edges has property B. Therefore $m(n) \ge 2^{n-1}$.*

Proof: Let $H = (V, E)$ be an n-uniform hypergraph with less than 2^{n-1} edges. Color V randomly by 2 colors. For each edge $e \in E$, let A_e be the event that e is monochromatic. Clearly $\text{Prob}(A_e) = 2^{1-n}$. Therefore

$$\text{Prob}\left(\bigcup_{e \in E} A_e \right) \le \sum_{e \in E} \text{Prob}(A_e) < 1$$

and there is a 2-coloring without monochromatic edges. □

In Chapter 3, Section 5 we present a more delicate argument, due to J. Beck, that shows that $m(n) \ge \Omega(n^{\frac{1}{3} - o(1)} 2^n)$.

The best known upper bound to $m(n)$ is found by turning the probabilistic argument "on its head". Basically, the sets become random and each coloring defines an event. Fix V with v points, where we shall later optimize v. Let χ be a coloring of V with a points in one color, $b = v - a$ points in the other. Let $S \subset V$ be a uniformly selected n-set. Then

$$Pr(S \textit{ is monochromatic under } \chi) = \frac{\binom{a}{n} + \binom{b}{n}}{\binom{v}{n}}$$

Let us assume v is even for convenience. As $\binom{y}{n}$ is convex, this expression is minimized when $a = b$. Thus

$$Pr(S \textit{ is monochromatic under } \chi) \ge p$$

where we set

$$p = \frac{2\binom{v/2}{n}}{\binom{v}{n}}$$

$C(v)$ the set consisting of v together with all its neighbours. Suppose that during the process of picking vertices the number of vertices u that do not lie in the union of the sets $C(v)$ of the vertices chosen so far is r. By the assumption, the sum of the cardinalities of the sets $C(u)$ over all such uncovered vertices u is at least $r(\delta+1)$, and hence, by averaging, there is a vertex v that belongs to at least $r(\delta+1)/n$ such sets $C(u)$. Adding this v to the set of chosen vertices we observe that the number of uncovered vertices is now at most $r(1-\frac{\delta+1}{n})$. It follows that in each iteration of the above procedure the number of uncovered vertices decreases by a factor of $1-(\delta+1)/n$ and hence after $ln(\delta+1)$ steps there will be at most $n/(\delta+1)$ yet uncovered vertices which can now be added to the set of chosen vertices to form a dominating set of size at most the one in the conclusion of Theorem 2.2.

Combining this with some of the ideas of Matula (1987), we can obtain a very efficient algorithm to decide if a given undirected graph on n vertices is, say, $\frac{n}{2}$-edge connected. A *cut* in a graph $G=(V,E)$ is a partition of the set of vertices V into two nonempty disjoint sets $V=V_1 \cup V_2$. If $v_1 \in V_1$ and $v_2 \in V_2$ we say that the cut *separates* v_1 and v_2. The *size* of the cut is the number of edges of G having one end in V_1 and another end in V_2. In fact, we sometimes identify the cut with the set of these edges. The *edge-connectivity* of G is the minimum size of a cut of G. The following lemma is due to Matula.

Lemma 2.3. *Let $G=(V,E)$ be a graph with minimum degree δ and let $V=V_1 \cup V_2$ be a cut of size smaller than δ in G. Then every dominating set U of G has vertices in V_1 and in V_2.*

Proof: Suppose this is false and $U \subseteq V_1$. Choose, arbitrarily, a vertex $v \in V_2$ and let $v_1, v_2, \ldots, v_\delta$ be δ if its neighbors. For each i, $1 \leq i \leq \delta$, define an edge e_i of the given cut as follows; if $v_i \in V_1$ then $e_i = \{v, v_i\}$, otherwise, $v_i \in V_2$ and since U is dominating there is at least one vertex $u \in U$ such that $\{u, v_i\}$ is an edge; take such a u and put $e_i = \{u, v_i\}$. The δ edges $e_1, \ldots, e_\delta$ are all distinct and all lie in the given cut, contradicting the assumption that its size is less than δ. This completes the proof. □

Let $G=(V,E)$ be a graph on n vertices, and suppose we wish to decide if G is $n/2$ edge-connected, i.e., if its edge connectivity is at least $n/2$. Matula showed, by applying Lemma 2.3, that this can be done in time $O(n^3)$. By the remark following the proof of Theorem 2.2, we can slightly improve it and get an $O(n^{8/3}\log n)$ algorithm as follows. We first check if the minimum degree δ of G is at least $n/2$. If not, G is not $n/2$-edge connected, and the algorithm ends. Otherwise, by Theorem 2.2 there is a dominating set $U=\{u_1, \ldots, u_k\}$ of G, where $k=O(\log n)$, and it can in fact be found in $O(n^2)$-time. We now find, for each i, $2 \leq i \leq k$, the minimum size s_i of a cut that

Therefore, with positive probability no event A_K occurs, i.e., there is a tournament on n vertices that has the property S_k. □

Let $f(k)$ denote the minimum possible number of vertices of a tournament that has the property S_k. Since $\binom{n}{k} < \left(\frac{en}{k}\right)^k$ and $(1-2^{-k})^{n-k} < e^{-(n-k)/2^k}$, Theorem 2.1 implies that $f(k) \leq k^2 \cdot 2^k \cdot (\ln 2)\big(1+o(1)\big)$. It is not too difficult to check that $f(1) = 3$ and $f(2) = 7$. As proved by Szekeres (cf. Moon(1968)), $f(k) \geq c_1 \cdot k \cdot 2^k$. Can one find an explicit construction of tournaments with at most c_2^k vertices having property S_k? Such a construction is known, but is not trivial; it is described in Chapter 9.

A *dominating set* of an undirected graph $G = (V, E)$ is a set $U \subseteq V$ such that every vertex $v \in V - U$ has at least one neighbor in U.

Theorem 2.2. *Let $G = (V, E)$ be a graph on n vertices, with minimum degree $\delta > 1$. Then G has a dominating set of at most $n\frac{1+\ln(\delta+1)}{\delta+1}$ vertices.*

Proof: Put $p = \frac{\ln(\delta+1)}{\delta+1}$ and let us pick, randomly and independently, each vertex of V with probability p. Let X be the (random) set of all vertices picked and let $Y = Y_X$ be the random set of all vertices in $V - X$ that do not have any neighbor in X. The expected value of $|X|$ is clearly np. For each fixed vertex $v \in V$, $\mathrm{Prob}(v \in Y) = \mathrm{Prob}(v \text{ and its neighbors are not in } X) \leq (1-p)^{\delta+1}$. Since the expected value of a sum of random variables is the sum of their expectations (even if they are not independent) and since the random variable $|Y|$ can be written as a sum of n indicator random variables χ_v ($v \in V$), where $\chi_v = 1$ if $v \in Y$ and $\chi_v = 0$ otherwise, we conclude that the expected value of $|X| + |Y|$ is at most $np + n(1-p)^{\delta+1} = n\frac{\ln(\delta+1)}{\delta+1} + n\left(1 - \frac{\ln(\delta+1)}{\delta+1}\right)^{\delta+1} \leq n\frac{1+\ln(\delta+1)}{\delta+1}$. Consequently, there is at least one choice of $X \subseteq V$ such that $|X| + |Y_X| \leq n\frac{1+\ln(\delta+1)}{\delta+1}$. The set $U = X \cup Y_X$ is clearly a dominating set of G whose cardinality is at most $n\frac{1+\ln(\delta+1)}{\delta+1}$. □

Two simple but important ideas are incorporated in the last proof. The first is the linearity of expectation; many applications of this simple, yet powerful principle appear in Chapter 2. The second is, maybe, more subtle, and is an example of the "alteration" principle which is discussed in Chapter 3. The random choice did not supply the required dominating set U immediately; it only supplied the set X, which has to be altered a little (by adding to it the set Y_X) to provide the required dominating set.

It can be easily deduced from the results in Alon (1990a) that the bound in Theorem 2.2 is nearly optimal. A non-probabilistic, algorithmic, proof of this theorem can be obtained by choosing the vertices for the dominating set one by one, when in each step a vertex that covers the maximum number of yet uncovered vertices is picked. Indeed, for each vertex v denote by

2. Graph Theory.

A *tournament* on a set V of n players is an orientation $T = (V, E)$ of the edges of the complete graph on the set of vertices V. Thus, for every two distinct elements x and y of V either (x, y) or (y, x) are in E, but not both. The name tournament is natural, since one can think of the set V as a set of players in which each pair participates in a single match, where (x, y) is in the tournament iff x beats y. We say that T has the property S_k if for every set of k players there is one who beats them all. For example, a directed triangle $T_3 = (V, E)$, where $V = \{1, 2, 3\}$ and $E = \{(1, 2), (2, 3), (3, 1)\}$, has S_1. Is it true that for every finite k there is a tournament T (on more than k vertices) with the property S_k? As shown by Erdös (1963a), this problem, raised by Schütte, can be solved almost trivially by applying probabilistic arguments. Moreover, these arguments even supply a rather sharp estimate for the minimum possible number of vertices in such a tournament. The basic (and natural) idea is that if n is sufficiently large, as a function of k, then a *random* tournament on the set $V = \{1, \ldots, n\}$ of n players is very likely to have property S_k. By a random tournament we mean here a tournament T on V obtained by choosing, for each $1 \leq i < j \leq n$, independently, either the edge (i, j) or the edge (j, i), where each of these two choices is equally likely. Observe that in this manner, all the $2^{\binom{n}{2}}$ possible tournaments on V are equally likely, i.e., the probability space considered is symmetric. It is worth noting that we often use in applications symmetric probability spaces. In these cases, we shall sometimes refer to an element of the space as a *random element*, without describing explicitly the probability distribution. Thus, for example, in the proof of Proposition 1.1 random 2-edge-colorings of K_n were considered, i.e., all possible colorings were equally likely. Similarly, in the proof of the next simple result we study random tournaments on V.

Theorem 2.1. *If* $\binom{n}{k}(1 - 2^{-k})^{n-k} < 1$ *then there is a tournament on* n *vertices that has the property* S_k.

Proof: Consider a random tournament on the set $V = \{1, \ldots, n\}$. For every fixed subset K of size k of V, let A_k be the event that there is no vertex which beats all the members of K. Clearly $\mathrm{Prob}(A_K) = (1 - 2^{-k})^{n-k}$. This is because for each fixed vertex $v \in V - K$, the probability that v does not beat all the members of K is $1 - 2^{-k}$, and all these $n - k$ events corresponding to the various possible choices of v are independent. It follows that

$$\mathrm{Prob}\left(\bigcup_{\substack{K \subset V \\ |K|=k}} A_K\right) \leq \sum_{\substack{K \subset V \\ |K|=k}} \mathrm{Prob}(A_K) = \binom{n}{k}(1 - 2^{-k})^{n-k} < 1 \ .$$

most of the applications of the probabilistic method in discrete mathematics. Theoretically, this is, indeed, the case. However, in practice, the probability is essential. It would be hopeless to replace the applications of many of the tools appearing in this book, including, e.g., the second moment method, the Lovász Local Lemma and the concentration via martingales by counting arguments, even when these are applied to finite probability spaces.

The probabilistic method has an interesting algorithmic aspect. Consider, for example, the proof of Proposition 1.1 that shows that there is a two edge coloring of K_n without a monochromatic $K_{2\log_2 n}$. Can we actually find such a coloring? This question, as asked, may sound ridiculous; the total number of possible colorings is finite, so we can try them all until we find the desired one. However, such a procedure may require $2^{\binom{n}{2}}$ steps; an amount of time which is exponential in the size $(= \binom{n}{2})$ of the problem. Algorithms whose running time is more than polynomial in the size of the problem are usually considered unpractical. The class of problems that can be solved in polynomial time, usually denoted by P (see, e.g., Aho, Hopcroft and Ullmann (1974)), is, in a sense, the class of all solvable problems. In this sense, the exhaustive search approach suggested above for finding a good coloring of K_n is not acceptable, and this is the reason for our remark that the proof of Proposition 1.1 is non-constructive; it does not suply a constructive, efficient and deterministic way of producing a coloring with the desired properties. However, a closer look at the proof shows that, in fact, it can be used to produce, effectively, a coloring which is very likely to be good. This is because for large k, if $n = \lfloor 2^{k/2} \rfloor$ then $\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < \frac{2^{1+\frac{k}{2}}}{k!}\left(\frac{n}{2^{k/2}}\right)^k \leq \frac{2^{1+\frac{k}{2}}}{k!} \ll 1$. Hence, a random coloring of K_n is very likely not to contain a monochromatic $K_{2\log n}$. This means that if, for some reason, we *must* present a two coloring of the edges of K_{1024} without a monochromatic K_{20} we can simply produce a random two coloring by flipping a fair coin $\binom{1024}{2}$ times. We can then hand the resulting coloring safely; the probability that it contains a monochromatic K_{20} is less than $\frac{2^{11}}{20!}$; probably much smaller than our chances of making a mistake in any rigorous proof that a certain coloring is good! Therefore, in some cases the probabilistic, non-constructive method, does supply effective probabilistic algorithms. Moreover, these algorithms can sometimes be converted into deterministic ones. This topic is discussed in some detail in Chapter 15.

The probabilistic method is a powerful tool in Combinatorics and in Graph Theory. It is also extremely useful in Number Theory and in Combinatorial Geometry. More recently it has been applied in the development of efficient algorithmic techniques and in the study of various computational problems. In the rest of this chapter we present several simple examples that demonstrate some of the broad spectrum of topics in which this method is helpful. More complicated examples, involving various more delicate probabilistic arguments, appear in the rest of the book.

Chapter 1: The Basic Method

1. The probabilistic method.

The probabilistic method is a powerful tool in tackling many problems in discrete mathematics. Roughly speaking, the method works as follows: Trying to prove that a structure with certain desired properties exists, one defines an appropriate probability space of structures and then shows that the desired properties hold in this space with positive probability. The method is best illustrated by examples. Here is a simple one. The *Ramsey-number* $R(k,\ell)$ is the smallest integer n such that in any two-coloring of the edges of a complete graph on n vertices K_n by red and blue, either there is a red K_k (i.e., a complete subgraph on k vertices all of whose edges are colored red) or there is a blue K_ℓ. Ramsey (1929) showed that $R(k,\ell)$ is finite for any two integers k and ℓ. Let us obtain a lower bound for the diagonal Ramsey numbers $R(k,k)$.

Proposition 1.1. *If* $\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < 1$ *then* $R(k,k) > n$. *Thus* $R(k,k) > 2^{k/2}$ *for all* $k \geq 3$.

Proof: Consider a random two-coloring of the edges of K_n obtained by coloring each edge independently either red or blue, where each color is equally likely. For any fixed set R of k vertices, let A_R be the event that the induced subgraph of K_n on R is *monochromatic* (i.e., that either all its edges are red or they are all blue). Clearly, $\mathrm{Prob}(A_R) = 2^{1-\binom{k}{2}}$. Since there are $\binom{n}{k}$ possible choices for R, the probability that at least one of the events A_R occurs is at most $\binom{n}{k}2^{1-\binom{k}{2}} < 1$. Thus, with positive probability, no event A_R occurs and there is a two-coloring of K_n without a monochromatic K_k, i.e., $R(k,k) \geq n$. Note that if $k \geq 3$ and we take $n = \lfloor 2^{k/2} \rfloor$ then $\binom{n}{k}2^{1-\binom{k}{2}} < \frac{2^{1+\frac{k}{2}}}{k!} \cdot \frac{n^k}{2^{k^2/2}} < 1$ and hence $R(k,k) > 2^{k/2}$ for all $k \geq 3$. □

This simple example demonstrates the essence of the probabilistic method. To prove the existence of a good coloring we do not present one explicitly, but rather show, in a non-constructive way, that it exists. This example appeared in a paper of P. Erdös from 1947. Although Szele applied the probabilistic method to another combinatorial problem, mentioned in Chapter 2, already in 1943, Erdös was certainly the first one who understood the full power of this method and has been applying it successfully over the years to numerous problems. One can, of course, claim, that the probability is not essential in the proof given above. An equally simple proof can be described by counting; we just check that the total number of two-colorings of K_n is bigger than the number of those containing a monochromatic K_k. Moreover, since the vast majority of the probability spaces considered in the study of combinatorial problems are finite spaces, this claim applies to