VARIATIONS ON THE MONOTONE SUBSEQUENCE THEME OF ERDŐS AND SZEKERES*

J. MICHAEL STEELE†

Abstract. A review is given of the results on the length of the longest increasing subsequence and related problems. The review covers results on random and pseudo-random sequences as well as deterministic ones. Although most attention is given to previously published research, some new proofs and new results are given. In particular, some new phenomena are demonstrated for the monotonic subsequences of *sections* of sequences. A number of open problems from the literature are also surveyed.

Key words. Monotone subsequence, unimodal subsequence, partial ordering, limit theory, irrational numbers, derandomization, pseudo-random permutations.

AMS(MOS) subject classifications. 60C05, 06A10

1. Introduction. The main purpose of this article is to review a number of developments that spring from the classic theorem of Erdős and Szekeres (1935) which tells us that from a sequence of n^2+1 distinct real numbers we can always extract a *monotonic* subsequence of length at least $n+1$. Although the Erdős-Szekeres theorem is purely deterministic, the subsequent work is accompanied by a diverse collection of results that make contact with randomness, pseudo-randomness, and the theory of algorithms.

Central to the stochastic work that evolved from the discovery of Erdős and Szekeres is the theorem that tells us that for independent random variables X_i, $1 \le i < \infty$, with a continuous distribution, the length I_n of the longest increasing subsequence in $\{X_1, X_2, \ldots, X_n\}$,

$$I_n = \max\{k : X_{i_1} < X_{i_2} < \ldots < X_{i_k} \text{ with } 1 \le i_1 < i_2 < \ldots < i_k \le n\},$$

satisfies

$$\lim_{n\to\infty} I_n/\sqrt{n} = 2 \quad \text{with probability one .} \tag{1.1}$$

The suggestion that 2 might be the right limiting constant was first put forth by Baer and Brock (1968) who had engaged in an interesting Monte Carlo study motivated by S. Ulam (1961), but the first rigorous progress is due to Hammersley (1972) who showed $I_n/\sqrt{n}$ converges in probability to a constant $C > 0$. The constant C proved to be difficult to determine, but Logan and Shepp (1977) first showed $C \ge 2$ by a sustained argument of the calculus of variations and then $C \le 2$ was established by Vershik and Kerov

* Research supported in part by NSF Grant Number DMS 92-11634 and Army Research Office Grant DAAL03-91-G-0110.

† Department of Statistics, University of Pennsylvania, The Wharton School, 3000 Steinberg Hall-Dietrich Hall, Philadelphia, PA 19104-6302.

(1977) using information about the Plancherel measure on Young-tableaux. The later result was subsequently simplified by Pilpel (1986) who also gave the bound valid for all n,

$$EI_n \leq \sum_{j=1}^{n} 1/\sqrt{j}.$$

More recently, Aldous and Diaconis (1993) have given an insightful proof that $C = 2$ by building on elementary aspects of the theory of interacting particle systems.

In the next section we sketch six (or more) proofs of the Erdős-Szekeres theorem. The central intention of reviewing these proofs is to see what each of the methods can tell us about combinatorical technique, but the very multiplicity of proofs of the Erdős-Szekeres theorem offers an indication that the result is not so special as one might suspect. Quite to the contrary, the Erdős-Szekeres theorem offers us a touchstone for understanding a variety of useful ideas. The third section looks at generalizations of (1.1) to $d \geq 2$ and to structures like unimodality. Section four then develops the theory of monotone subsequences for pseudo-random sequences. That section gives particularly detailed information about the Weyl sequences, $x_n = n\alpha \bmod 1$ for irrational α. Section five develops the theory of monotone subsequences for sections of an infinite sequence.

The final section comments briefly on open problems and on some broader themes that seem to be emerging in the relationship between probability and combinatorial optimization.

2. Six or more proofs. Perhaps the most widely quoted proof of the Erdős-Szekeres theorem is that of Hammersley (1972) which uses a visually compelling pigeon-hole argument. The key idea is to place the elements of the sequence $x_1, x_2, \ldots, x_m$ with $m = n^2 + 1$ into a set of ordered columns by the following rules:

(a) let x_1 start the first column, and, for $i \geq 1$,

(b) if x_i is greater than or equal to the value that is on top of a column, we put x_i on top of the first such column, and

(c) otherwise start a new column with x_i. The first point to notice about this construction is that the elements of any column correspond to an increasing subsequence. The second observation is that the only time we shift to a later column is when we have an item that is smaller than one of its predecessors. Thus, if there are k columns in the final structure, we can trace back from the last of these and find monotone decreasing subsequence of length k. Since $n^2 + 1$ numbers are placed into the column structure, one must either have more than n columns or some column that has height greater than n. Either way, we find a monotone subsequence of length $n + 1$.

Hammersley's proof is charming, but the original proof of Erdős and Szekeres (1935) can in some circumstances teach us more. To follow the original plan, we first let $f(n)$ denote the least integer such that any sequence of $f(n)$ real numbers must contain a monotone subsequence of length n. One clearly has $f(1) = 1$, $f(2) = 2$, and, with a moments reflection, $f(3) = 5$. By using the construction that extends the example $\{3, 2, 1, 6, 5, 4, 9, 8, 7\}$ we see $f(n) > (n-1)^2$, and the natural conjecture is that $f(n) = (n-1)^2 + 1$. The method we use to prove this identity calls on a modest bit of geometry, but the combinatorial technique that it teaches best could be called the *abundance principle*: In many situations if a structure of a certain size must have a special substructure, then a somewhat larger structure must have many of the special substructures.

The Erdős-Szekeres proof is quickly done at a blackboard, although a few more words are needed on paper. To show by induction that $f(n) = (n-1)^2 + 1$, we begin by considering an integer $b \geq 0$ and a set of $f(n) + b$ distinct points in $\mathbb{R}^2$. By applying the induction hypothesis $b + 1$ times we can find $b + 1$ distinct points that are terminal points of monotone sets of length n. For the moment we withhold our budget set B of b points, and we invoke the induction hypothesis on set P with cardinality $f(n)$. The induction hypothesis gives us a monotone subsequence of length n. We remove the last point of this sequence from P, and we add one of the budget points from B to get a new set P' of cardinality $f(n)$.

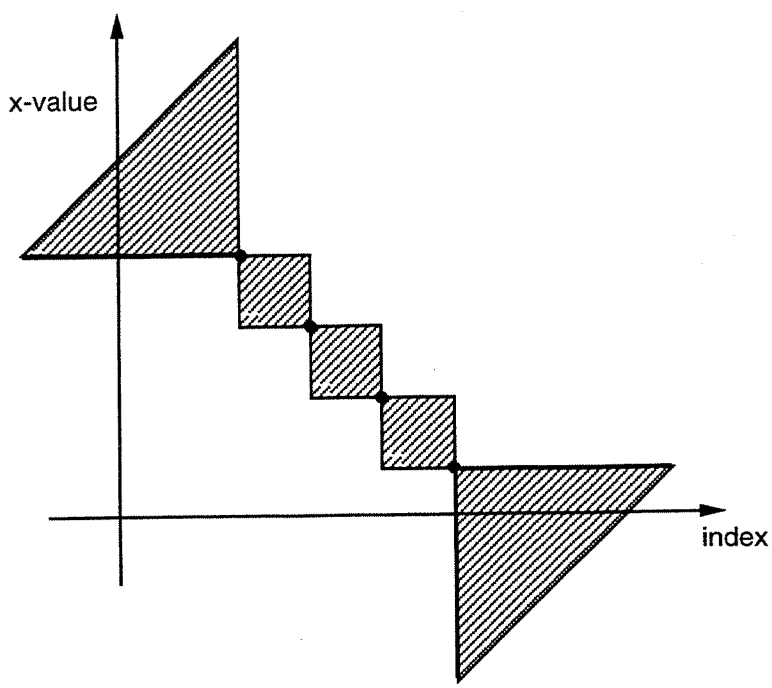

FIG. 2.1. *The dots denote the points of S^+, and the shaded region is the region D of points that are not comparable to the points of S^+ in the up-and-to-the-right order.*

Now, just to get started, suppose we took $b = 2n$, then since any terminal point must be associated with an decreasing sequence or an increasing sequence, we can suppose without loss of generality that there are $n + 1$ points S that are terminal points of increasing sequence of length n. If some two elements of S were in increasing order, we could extend one of the length n increasing sequences to a sequence of length $n + 1$. On the other hand if no two elements of S are in increasing order, then S constitutes a set of size $n + 1$ in decreasing order. We have therefore shown $f(n+1) \leq f(n) + 2n$, which is enough to show that $f(n) \leq n^2 + 1$, but we can do better by exercising a little more care.

This time we add just $b = 2n-1$ new points and get $2n$ terminal points. If we have $n+1$ of either of the two types of terminals, then we can proceed as before to find a monotone sequence of length $n + 1$. Therefore we can suppose that there are exactly n terminal points of each type. If S^+ and S^- denote the terminals of the increasing and decreasing sequences, then by the argument of the previous paragraph there is no loss in assuming that S^+ forms a decreasing set (see Figure 2.1). Now if any point $x \in S^-$ were to be up-and-to-the-right of any point of $y \in S^+$, then x would combine with the increasing sequence sending y to form an increasing sequence of length $n + 1$. Therefore all points of S^- are in the region D. These are an increasing sequence of length n, and the last of there points is majorized by some $y \in S^+$, telling us $S^+ \cup \{y\}$ would give the required subsequence.

More will be said later about the virtue of the Erdős-Szekeres proof, but first we want to recall what is perhaps the slickest and most systematic proof, the one due to Seidenberg (1959) and which is naturally suggested by dynamic programming. We take $S = \{p_1, p_2, \ldots, p_m\} \subset \mathbb{R}^2$ distinct points ordered by their x-coordinates, and we define $\varphi : S \to \mathbb{Z} \times \mathbb{Z}$ by letting $\varphi(p) = (s,t)$ if s is the length of the longest increasing subsequence terminating at p, and t is the length of the longest decreasing subsequence terminating at p. We are not too concerned with algorithms in this review, but at this point one may as well note that there is no problem in calculating $\varphi(p_k)$ in time $0(k)$ given $\varphi(p_1)$, $\varphi(p_2), \ldots, \varphi(p_{k-1})$ so the complete computation of φ on S can be determined in time $0(m^2)$, (cf. Friedman (1975)). Now, if S contains no monotone subsequence of length n, then $\varphi(S) \subseteq \{1, 2, \ldots, n-1\} \times \{1, 2, \ldots, n-1\}$. But φ is injective, since, if $p, q \in S$ and q follows p in x-coordinate order, then $\varphi(q)$ might have at least one coordinate larger than the corresponding coordinate of $\varphi(p)$. Hence we see that if $m \geq (n-1)^2 + 1$ we find that S must contain a monotone subsequence of length n. In other words, we have $f(n) \geq (n-1)^2 + 1$ and thus complete our third proof of the Erdős-Szekeres Theorem.

The fourth proof we consider is one due to Blackwell (1971) that is not so systematic as that of Seidenberg (1959) or as general the original proof of Erdős and Szekeres, but it serves well to make explicitly the connection to greedy algorithms.

If $S = \{x_1, x_2, \ldots, x_r\}$ is our set of $r > nm$ distinct real numbers, we say a monotone decreasing subsequence S' is *leftmost* if $x'_1 = x_1$ and each term x'_i of S' is equal to the next term of S' which is smaller than x'_{i-1}. Thus S' is the consequence of applying a greedy algorithm to the sequence S.

If we successively apply this greedy process to the points that remain after removal of the leftmost decreasing subsequence we obtain a decomposition of S into $S_1, S_2, \ldots, S_t$ where each S_i is a decreasing subsequence. The observation about this decomposition is that we can construct an increasing subsequence $\{a_1, 2, \ldots, a_t\}$ of S by the following backward moving process:

1. Select a_t arbitrarily from S_t
2. For any $j = t$ down to 1, select a_{j-1} as any term in S_{j-1} that is smaller than a_j.

Because of the definition of the S_j, $1 \leq j \leq t$ we can always complete the steps required in this process. We then find either an increasing set $\{a_1, a_2, \ldots, a_t\}$ with $t > n$ or else one of the decreasing subsequences S_j has cardinality bigger than m. In retrospect one can see that Blackwell's proof is almost isomorphic to Hammersley, though the associated picture and algorithmic feel are rather different.

The fifth proof is closely related to the one just given, but it still offers some useful distinctions. In the solution of Exercise 14.25, Lovász (1979) suggests that given a set S of $n^2 + 1$ real numbers $\{x_1, x_2, \ldots, x_{n^2+1}\}$ one can define a useful partition $A_1, A_2, \ldots$ of S by taking A_k to be the set of all x_j with $1 \leq j \leq n^2 + 1$ for which the longest increasing subsequence beginning with x_j has length exactly equal to k. One can easily check from this definition that each of the sets $A_k = \{i_1 < i_2 < \ldots < i_s\}$ gives rise to k monotone decreasing subsequence $x_{i_1} > x_{i_2} > \ldots > x_{i_s}$, and from this observation the Erdős-Szekeres theorem follows immediately. This last proof has the benefit of showing that any digraph with no directed path of length greater than k has chromatic number bounded by k.

The fifth proof is one that deserves serious consideration, but which would lead us too far afield for us to develop in detail. The central idea is that of the construction of Schensted (1961) that provides a one-to-one correspondence between pairs of standard Young tableaux and the set of permutations. This correspondence as well as its application to the theorem of Erdős and Szekeres and to algorithms for the determination of the longest monotone subsequence of a permutation are well described in Stanton and White (1986). The work of Schensted (1961) was substantially extended by Knuth (1970) to objects that go well beyond permutations. We do not know the extent to which the correspondence provided by Knuth (1970) might contribute to the central problems of this review.

Our final observation on the proofs is just to note that the Erdős-Szekeres theorem also follows from the well known decomposition theorem of Dilworth (1950) which says that any finite partially ordered set can be partitioned into k chains $C_1, C_2, \ldots, C_k$ where k is the maximum cardinality of all anti-chains in S. Using our previous S with the up-and-to-the-right ordering, we see that if there is no decreasing subsequence of length n, then $k < n$ and

$$|C_1| + |C_2| + \ldots + |C_k| = |S| \geq (n-1)^2 + 1$$

implies that for some $|C_j|$ we have $|C_j| \geq n$. Since C_j corresponds to an increasing sequence, we have the final proof.

Dilworth's theorem has itself been the subject of many further developments, some of which are directly connected to the issues engaged by this review. For a recent survey of the work related to Dilworth's theorem and sketches of several proofs of Dilworth's theorem, one should consult Bogart, Greene, and Kung (1990). A result that generalizes both the Dilworth decomposition theorem and the digraph theorem of the previous paragraph is the theorem of Gallai and Milgram (1960) which says that if α is the largest number of points in a digraph G that are not connected by any edge, then G can be covered by α directed paths.

3. Higher dimensions. The charm of the Erdős-Szekeres monotonicity theorem goes beyond the call for innovative proofs. There are several useful extensions and generalizations, though amazingly almost all of these have grown up a good many years after the original stimulus.

One natural issue concerns the generalization of the monotonicity theorem to d-dimensions. If we define a partial order on $\mathbb{R}^d$ by saying $y << w$ if $y = (y_1, y_2, \ldots, y_d)$ and $w = (w_1, w_2, \ldots, w_d)$ satisfy $y_1 \leq w_1$, $y_2 \leq w_2, \ldots, y_d \leq w_d$, then the natural variables of interest are $\Lambda^+(y_1, y_2, \ldots, y_n)$, the length of the longest chain in the set $\{y_1, y_2, \ldots, y_n\} \subset \mathbb{R}^d$ and correspondingly $\Lambda^-(y_1, y_2, \ldots, y_n)$, the cardinality of the largest anti-chain in $\{y_1, y_2, \ldots, y_n\}$. After tentative steps in Steele (1977) which were nevertheless enough to settle a conjecture of Robertson and Wright (1974), the definitive result was established by Bollobás and Winkler (1988). Their main result is that for $\{X_i : 1 \leq i < \infty\}$, independent and uniformly distributed on $[0,1]^d$, one has

$$\lim_{n \to \infty} \Lambda^+(X_1, X_2, \ldots, X_n)/n^{1/d} = c_d > 0$$

where the limit holds with probability one and where c_d is a constant that depends on the dimension $d \geq 1$. Further, Bollobás and Winkler (1988) also showed that the constants c_d satisfy the interesting relation

$$\lim_{d \to \infty} c_d = \sum_{n=0}^{\infty} \frac{1}{n!} = e.$$

Since it is trivial that $c_1 = 1$ and since we know from Logan and Shepp (1977) and Vershik and Kerov (1977) that $c_2 = 2$, we thus have in hand three instances that support the humorous but feasible speculation due to

Aldous (personal communication) that perhaps

$$c_d = \sum_{n=0}^{d} \frac{1}{n!}$$

for all $d \geq 2$. The correctness of this interpolation and also the one based on $c_d = (1 + 1/(d+1))^{d-1}$ are placed on slippery ground by simulations of P. Winkler and R. Silverstein (cf. Silverstein (1988)) that suggest that c_3 is approximately 2.35, a value that does not agree well with either of the two candidates of 5/2 and 9/4. Still, as the investigators note the convergence in the simulations was very slow, and the value of 2.35 is the result of heuristic curve fitting. The determination of c_d remains of interest from several points of view.

4. Totals and functionals of totals. Lifschitz and Pittel (1981) have studied the total number T_n of increasing subsequences of n independent uniformly distributed random variables. Among other results they found

$$ET_n \sim \alpha n^{-1/4} e^{2\sqrt{n}}$$

where $\alpha = (2\sqrt{\pi e})^{-1}$, and

$$E(T_n^2) \sim \beta_0 n^{-1/4} \exp(\beta_1 n^{1/2})$$

where $\beta_1 = 2(2 + \sqrt{5})^{1/2}$ and β_0 is equal to $(20\pi(2 + \sqrt{5})^{1/2} \exp(2 + \sqrt{5}))^{-1/2} \sim 0.016$. These asymptotic results were obtained by analytic methods beginning with the easy formula

$$E(T_n) = \sum_{k=0}^{n} \frac{1}{k!} \binom{n}{k}$$

and the trickier

$$E(T_n^2) = \sum_{k+\ell \leq n} r^{\ell} \left\{ \frac{1}{(k+\ell)!} \right\} \binom{n}{k+\ell} \binom{(k+1)/2 + \ell - 1}{\ell}.$$

Lifschitz and Pittel (1981) also proved that there is a constant γ such that as $n \to \infty$

$$n^{-\frac{1}{2}} \log T_n \to \gamma$$

in probability and in mean. The exact determination of γ remains an open problem, though one has the bounds

$$2 \log 2 \leq \gamma \leq 2.$$

5. Unimodal subsequences. The most natural variation on the theme of monotone subsequences is perhaps that of unimodal subsequences, which are those that either increase to a certain point and then decrease, or else decrease to a point and increase thereafter. Formally, given a sequence $S = \{x_1, x_2 \ldots, x_n\}$ we are concerned with $U_n(S)$ defined by $U_n(S) = \max(U_n^+, U_n^-)$ where

$$U_n^+ = \max\{k : \exists 1 \le i_1 < i_2 < \ldots < i_k \text{ with } x_{i_1} < x_{i_2} < \ldots < x_{i_j} > x_{i_{j+1}} > \ldots > x_{i_k} \text{ for some } j\}$$

and

$$U_n^- = \max\{k : \exists 1 \le i_1 < i_2 < \ldots < i_k \text{ with } x_{i_1} > x_{i_2} > \ldots x_{i_j} < x_{i_{j+1}} < \ldots < x_{i_k} \text{ for some } j\}$$

In a remarkable *tour-de-force*, Chung (1980) established the result that for any sequence S of n distinct reals, we have

$$U_n \ge \lceil (3n - 3/4)^{1/2} - 1/2 \rceil, \tag{5.1}$$

and, moreover, this result is best possible in the sense that for any $n \ge 1$ there is a sequence of distinct reals for which we have equality in (5.1). The complexity of the proof of this result is of a different order than that of the Erdős-Szekeres theorem, although there are important qualitative insights that Chung's proof shares with the dynamic programming of the Erdős-Szekeres theorem that was given in Section two. The proof provided by Chung calls instead on four functions parallel to the two components φ of Section two, although instead of considering only one simple injective image contained in $[1, n] \times [1, n]$, Chung must consider several such images that are contained in more complex domains.

With the deterministic problem resolved, Chung (1980) posed the natural problem of determining the asymptotic behavior of $U_n(X_1, X_2, \ldots, X_n)$ when the X_i are independent and random variables with a common continuous distribution. This turned out to be much easier to resolve than the problem solved by Chung, and in Steele (1981) it was proved that we have that

$$U_n/\sqrt{n} \to 2\sqrt{2}$$

with probability one. The corresponding constant when one permits k changes in the sense of the monotonicity was also found to be $2\sqrt{k}$.

6. Concentration inequalities. The length I_n of the longest increasing subsequence of n independent uniformly distributed random variables has the property of being rather tightly concentrated about its mean. One way to see this in terms of variance is to call on the Efron-Stein inequality which in the form given by Steele (1986) tells us that if $F(y_1, y_2, \ldots, y_{n-1})$

is any function of $n-1$ variables then by introducing n random variables by applying F to the variables $X_1, X_2, \ldots, X_n$ with the i'th variable withheld $F_i = F(X_1, X_2 \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$ and setting $\overline{F} = \frac{1}{n}(F_1 + F_2 + \ldots + F_n)$, we have

$$\text{Var}\ (F(X_1, X_2, \ldots, X_{n-1})) \leq E \sum_{i=1}^{n} (F_i - \overline{F})^2.$$

When we focus on I_{n-1} we first note that

$$I(X_1, X_2, \ldots, X_n) \leq 1 + I(X_1, X_2, \ldots, X_{i-1}, X_{i+1}, \ldots X_n)$$

and

$$I(X_1, X_2, \ldots X_{i-1}, X_{i+1}, \ldots X_n) \leq I(X_1, X_2, \ldots, X_n).$$

If we let A_n denote the number of sample points X_i that are in *all* of the increasing subsequence having maximum length L_n, we then find

$$\sum_{i=1}^{n} (I(X_1, X_2, \ldots, X_n) - I(X_1, X_2, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n))^2 \leq A_n$$

Since $A_n \leq I_n$ and since the quadratic sum is decreased by replacing $I(X_1, X_2, \ldots, X_n)$ by $n^{-1}(I(X_2, X_3, \ldots, X_n) + I(X_1, X_2, \ldots, X_n) + \ldots + I(X_1, X_2, \ldots, X_{n-1}) = \overline{I}$ we have

$$\sum_{i=1}^{n} (I(X_1, X_2, \ldots, X_{i-1}, X_{i+2}, \ldots X_n) - \overline{I})^2 \leq I_n.$$

Taking expectations and applying the Efron-Stein inequality gives

$$\text{Var}\ I_{n-1} \leq EI_n \leq C\sqrt{n}$$

where for any $n \geq n(\varepsilon)$ we can take $C \leq 2 + \varepsilon$.

This bound on Var I_{n-1} was easily won, but despite its simplicity is enables us to circumvent some rather heavy analysis. In particular, the variance bound, monotonicity of I_n, and the asymptotics of the mean $EI_n \sim 2\sqrt{n}$ give us an easy proof that $I_n/\sqrt{n} \to 2$ with probability one just by following the usual Chebyshev and subsequence arguments. This development may seem surprising since the strong law for I_n has served on several occasions as a key example of the effectiveness of subadditive ergodic theory (Durrett (1991) and Kingman (1973)). Here we should also note that Aldous (1993) gives a scandalously simple proof of the Efron-Stein inequality, making it easy enough to cover in even a first course in probability theory.

Lower bounds on Var I_n have been obtained recently by B. Bollobás and R. Pemantle who independently established that there is a $c > 0$ for which

$$\text{Var } I_n \geq cn^{1/8}$$

for all $n \geq 3$. In view of the bounds just reviewed one suspects that

$$\lim_{n \to \infty} \log(\text{Var } I_n)/\log n = \alpha$$

for some $1/8 \leq \alpha \leq 1/2$. The existence of the limit may not be too difficult to establish, but substantial new insight will be required to determine the value of α.

The tail probabilities of I_n were first studied in Frieze (1991) using the bounded difference method. The bound obtained by Frieze was subsequently improved by Bollobás and Brightwell (1992) who established that for all $\epsilon > 0$ there is a $\beta = \beta(\epsilon) > 0$ such that for $n \geq n(\varepsilon)$, we have

$$P\left(|I_n - EI_n| \geq n^{1/4+\varepsilon}\right) \leq \exp(-n^{\beta}).$$

The work of Bollobás and Brightwell (1992) also considered the d-dimensional increasing subsequences $\Lambda_d^+(X_1, X_2, \ldots, X_n) = \Lambda_{d,n}^+$ for X_i independent and uniformly distributed on the unit d-cube.

THEOREM 6.1. *For every $d \geq 2$, there is a constant A_d such that for all $n \geq n(d)$ one has*

$$P\left(|\Lambda_{d,n}^+ - E\Lambda_{d,n}^+| \geq \lambda A_d n^{1/2d} \log n/\log\log n\right) \leq 80\lambda^2 \exp(-\lambda^2)$$

for all λ with $2 < \lambda < n^{1/2d}/\log\log n$.

Large deviation results like the last one have many consequences but a particularly valuable consequence in the present case is that one can extract a rate of convergence result for certain means. We let N have the Poisson distribution with mean n and define constants $c_{n,d}$ by

$$n^{-1/d} E\Lambda_d^+(X_1, X_2, \ldots, X_N) = c_{n,d}.$$

The main result is that

$$c_d - c_{n,d} = 0\left(n^{-1/2d} \log^{3/2} n/\log\log n\right)$$

where $c_2 = 2$ and the other constants c_d are those of the Bollobás-Winkler theorem for $d > 2$ as viewed in Section 3.

The most precise result on the deviations of the length longest increasing subsequence I_n is due to Talagrand (1993) and emerges from the theory of abstract isoperimetric theorems for product measures developed in Talagrand (1988, 1991, 1993).

THEOREM 6.2. *If M_n denotes the median of I_n, then for all $u > 0$ we have*

$$P(I_n \geq M_n + u) \leq 2\exp(-u^2/(4(M_n - u)))$$

and

$$P(I_n \leq M_n - u) \leq 2\exp(-u^2/(4M_n)).$$

7. Pseudo-random sequences. If $0 < \alpha < 1$ is an irrational number and $\{X\}$ denotes the fractional part of X, then the sequence of values determined by the fractional parts of the multiples of α given by $\{\alpha\}, \{2\alpha\}, \{3\alpha\}, \ldots, \{n_\alpha\}$ are uniformly distributed in $[0,1]$ in the sense that if $[a,b] \subset [0,1]$ then the number of the integers k satisfying $\{k\alpha\} \in [a,b]$ for $1 \leq k \leq n$ is asymptotic to $(b-a)n$ as $n \to \infty$. Bohl (1909) was evidently the first to show sequence of points $\{n\alpha\}$ share this property with the independent uniformly distributed random variables X_n, $1 \leq n < \infty$, though the subsequent importance of this observation was certainly driven home by H. Weyl and G. Hardy. The project of exploring which properties of the X_n are shared by the $\{n\alpha\}$ is a natural one, of which the results of Kesten (1960) and Beck (1991) are telling examples. The survey of Niederreiter (1978) provides an extensive review of the ways that pseudo-random numbers can parallel those that are honestly random and also articulates ways where pseudo-random sequences can be even more useful.

There are some very simple properties that make $\{X_n\}$ and $\{n\alpha\}$ seem quite different, so there is particular charm to the fact that they can behave quite similarly in the context of such a non-standard issue as the length of the longest increasing subsequence. If we let $\ell_n^+(\alpha)$ and $\ell_n^-(\alpha)$ denote respectively the longest increasing and longest decreasing subsequences of $\{\alpha\}, \{2\alpha\}, \ldots, \{n\alpha\}$, then ℓ_n^+ and ℓ_n^- turn out to have behavior that echos closely the behavior of their stochastic cousins. The first investigation of this issue is given in Del Junco and Steele (1978) where it is proved using discrepancy estimates that

$$\frac{\log \ell_n^+(\alpha)}{\log n} \to \frac{1}{2} \text{ and } \frac{\log \ell_n^-(\alpha)}{\log n} \to \frac{1}{2}$$

for all algebraic irrationals α and for a set of irrationals of $[0,1]$ measure one.

In Boyd and Steele (1978) a more precise understanding of $\ell_n^+(\alpha)$ was obtained by using the continued fraction expansion of α. It turns out that $n^{1/2}$ is the correct order of ℓ_n^+ and ℓ_n^- if and only if α has bounded partial quotients. Moreover, when the partial quotient sequence of α is known one can determine the precise range of $\ell_n^+/\sqrt{n}$ and $\ell_n^-/\sqrt{n}$. For example, in the eternally favorite special case of the golden ratio $\alpha_0 = (1+\sqrt{5})/2 = [1; 1, 1, 1, \ldots]$, we have

$$\liminf_{n\to\infty} \ \ell_n^+(\alpha_0)/\sqrt{n} = 2/5^{1/4} = 1.337481\ldots$$

and

$$\limsup_{n\to\infty} \ \ell_n^-(\alpha_0)\sqrt{n} = 5^{1/4} = 1.495349\ldots .$$

There is a non-zero gap $\Delta(\alpha_0)$ between these two limits, and in fact there is no α for which one has a limit theorem precisely like that for independent random variables, but nevertheless we see there is a close connection between $\{n\alpha\}$ and the genuinely random case.

In Boyd and Steele (1978) it is further proved that the gap $\Delta(\alpha)$ is minimized precisely for $\alpha_0 = (1+\sqrt{5})/2$. It is also proved there that for all irrational α we have

$$\limsup_{n\to\infty} \ \ell_n^+(\alpha)\ell_n^-(\alpha)/n = 2,$$

and for α with unbounded partial quotients

$$\liminf_{n\to\infty} \ \ell_n^+(\alpha)\ell_n^-(\alpha)/n = 1.$$

These results contrast with the analog for a sequence of independent uniformly distributed random variables for which we have

$$\lim_{n\to\infty} \ I_n^+ I_n^- = 4 \quad \text{a.s.}$$

The behavior of the pseudo-random Weyl sequence $\{\alpha\}$, $\{2\alpha\}$, $\{3\alpha\},\ldots,$ $\{n\alpha\}$ is not unique in the world of pseudo-random sequences. There is a remarkable sequence due to van der Corput (cf. Hammersley and Handscomb (1964)) that was invented in order to provide a sequence $\{x_i \in [0,1] : 1 \le i < \infty\}$ that has an especially small discrepancy

$$D_n = \sup_{0\le x\le 1} \Big|\sum_{i=1}^{n} 1_{[0,x]}(x_i) - x\Big|.$$

To define this sequence we first note that for any integer $n \ge 0$ there is a unique representation $n = \sum_{i=0}^{\infty} a_i 2^i$ where $a_i \in \{0,1\}$. The n-th element $\varphi_2(n)$ of the van der Corput sequence of base 2 is given by "reflecting the expansion of n in the decimal point," that is we have

$$\varphi_2(n) = \sum_{i=0}^{\infty} a_i 2^{-i-1}.$$

The sequence is thus given by $\{1/2, 1/4, 3/4, 1/8, 5/8, 3/8, 7/8, \ldots\}$, and one can see that the sequence does indeed disperse itself in a charmingly uniform fashion. Béjian and Faure (1977) have shown that there are sequences that are still more uniform but even there new sequences are conceptually quite close to the original idea.

Asymptotic equidistribution is perhaps the most basic feature of a sequence of independent uniformly distributed sequence, but one can check that the van der Corput sequence is far more uniform than a random sequence. This divergence from the behavior of independent uniform random variables is one of the facts that adds zest to the study of the behavior of the longest increasing subsequence of $\{\varphi_2(k) : 1 \leq k \leq n\}$.

The basic results are the following:

$$\limsup_{n\to\infty} \ \ell^+(\varphi_2(1), \varphi_2(2), \ldots, \varphi_2(n))/\sqrt{n} = 3/2$$

and

$$\liminf_{n\to\infty} \ \ell^+(\varphi_2(1), \varphi_2(2), \ldots, \varphi_2(n))/\sqrt{n} = \sqrt{2}.$$

For any integer p, one can define a *base* p van der Corput sequence by expanding n in base p and letting $\varphi_p(n)$ be the analogous "reflection in the decimal point" for $p > 2$ we have

$$\limsup_{n\to\infty} \ \ell^+(\varphi_p(1), \varphi_p(2), \ldots, \varphi_p(n))/\sqrt{n} = p^{1/2}$$

and

$$\liminf_{n\to\infty} \ \ell^+(\varphi_p(1), \varphi_p(2), \ldots, \varphi_p(n))/\sqrt{n} = 2(1 - p^{-1})^{1/2}.$$

Because of the interest that has been devoted to the hard won constant $c_2 = 2$ for the limit of $\ell_n^+/\sqrt{n}$ for the case of independent uniform variables, we should record that the last limit tells us that

$$\lim_{p\to\infty} \liminf_{n\to\infty} \ \ell^+(\varphi_p(1), \varphi_p(2), \ldots, \varphi_p(n))/\sqrt{n} = 2.$$

8. Theory of subsequences of sections. Many combinatorial problems exhibit new behaviors when they are imbedded in an infinite sequence of nested problems. The most classical instance of this phenomenon is offered by the theory of irregularity of distribution (cf. Beck and Chen (1987)).

To see how this phenomenon is manifested in the context of the Erdős-Szekeres theorem we suppose we are given an infinite sequence of distinct reals $S = \{x_1, x_2, \ldots, x_n, \ldots\}$, we focus on $M(x_1, x_2, \ldots, x_n) = M_n(S)$ which denote the cardinality of the largest monotone subsequence of $\{x_1, x_2, \ldots, x_n\}$. The Erdős-Szekeres theorem tells us that $M_n \geq \sqrt{n}$ for all $n \geq 1$, and, this is the best one can say for any fixed n. Still, for any infinite sequence S of distinct reals we can do better than $\sqrt{n}$ on infinitely many blocks of length n.

THEOREM 8.1. *There is a constant* $\gamma > 1$ *such that*

$$\limsup_{i,n\to\infty} \ M(x_{i+1}, x_{i+2}, \ldots, x_{i+n})/\sqrt{n} \geq \gamma. \tag{8.1}$$

Before embarking on the proof of this result we provide an example that shows that a doubly indexed result is what one required if the intention is to beat the Erdős-Szekeres theorem by a factor exceeding 1.We can construct an infinite sequence of integers $\{x_1, x_2, \ldots\}$ such that for all of the finite sections $\{x_1, x_2, \ldots, x_n\}$ we have $M(x_1, x_2, \ldots, x_n) = \lceil \sqrt{n} \rceil$. This sequence can be constructed as a concatenation of blocks

$$B_k = \{(-1)^{k+j}3^k + (-1)^{k+j}(2k - j) : j = 0, 1, 2, \ldots, 2k\}$$

for $k = 0, 1, 2, \ldots$. We note that $|B_k| = 2k + 1$, so $|B_0| + |B_1| + \ldots + |B_k| = (k+1)^2$. Also, we note that $B_0 = \{0\}$, $B_1 = \{-5, 4, -3\}$, $B_2 = \{13, -12, 11, -10, 9\}$ and $B_3 = \{-33, 32, -31, 30, -29, 28, -27\}$.

The proof of Theorem 8.1 depends on results that tell us that if $M(x_1, x_2, \ldots, x_n)$ and $M(x_1, x_2, \ldots, x_{2n})$ are both small, then $M(x_{n+1}, x_{n+2}, \ldots, x_{2n})$ must be exceptionally large. We will provide one such result as a consequence of the next lemma, but we first require some notation. Given any fixed sequence S, we let

$$\begin{aligned}
a(k) &= \max\{t : \exists x_{i_1} > x_{i_2} > \ldots > x_{i_t} \text{ with } 1 \le i_1 < i_2 < \ldots < i_t = k\} \\
b(k) &= \max\{t : \exists x_{i_1} < x_{i_2} < \ldots < x_{i_t} \text{ with } 1 \le i_1 < i_2 < \ldots < i_t = k\} \\
c_n(k) &= \max\{t : \exists x_{i_1} > x_{i_2} > \ldots > x_{i_t} \text{ with } k = i_1 < i_2 < \ldots < i_t \le n\} \\
d_n(k) &= \max\{t : \exists x_{i_1} < x_{i_2} < \ldots < x_{i_t} \text{ with } k = i_1 < i_2 < \ldots < i_{i_t} \le n\}.
\end{aligned}$$

To illuminate these definitions, we note $a(k)$ is the length of the longest decreasing sequence terminating with x_k, and $c_n(k)$ is the length of the longest decreasing subsequence beginning with x_k and ending before the nth element of the sequence. Thus, the sum $a(k) + c_n(k)$ is the length of the longest decreasing subsequence of $\{x_1, x_2, \ldots, x_n\}$ that contains x_k.

LEMMA 8.1. *For any* $0 < \varepsilon < 1/2$ *and* $n \ge n_0(\varepsilon)$, *if we have the bound* $M(x_1, x_2, \ldots x_n) \le (1+\varepsilon)\sqrt{n}$, *then there exists* $1 \le k \le n$ *such that*

$$a(k) \ge (1 - \delta)\sqrt{n}$$

and

$$b(k) \ge (1 - \delta)\sqrt{n}$$

provided that $\delta^2 > 2\varepsilon$.

Proof. The mapping $k \to (a(k), b(k))$ is injective since if $k < k'$ then we have $a(k) < a(k')$ or $b(k) < b(k')$, or both, because $x_{k'}$ will extend at least one of the monotone sequences ending at x_k. The set of integers (s, t) with $1 \le s, t \le (1+\varepsilon)\sqrt{n}$ and $\min(s, t) < (1-\delta)\sqrt{n}$ has cardinality at most $(1-\delta)^2 n + 2(\varepsilon+\delta)(1-\delta)n < (1 - \delta^2 + 2\varepsilon)n$ and the n distinct points $(a(k), b(k))$ are among the lattice points $[1, M_n]^2 \subset [1, (1+\varepsilon)\sqrt{n}]^2$, so for $\delta^2 > 2\varepsilon$ the pigeonhole principle yields the lemma. □

PROPOSITION 8.1. *For $0 < \varepsilon < 1/2$ and all $n \geq n_0(\varepsilon)$, the inequalities*

$$\begin{aligned} M(x_1, x_2, \ldots, x_n) &\leq (1+\varepsilon)\sqrt{n} \\ M(x_1, x_2, \ldots, x_{2n}) &\leq (1+\varepsilon)\sqrt{2n} \end{aligned}$$

imply

$$M(x_{n+1}, x_{n+2}, \ldots, x_{2n}) \geq (1 - 25\delta)\sqrt{2n}$$

provided $\delta^2 > 2\varepsilon$.

Proof. Even before beginning, we note that the factor of 25 given above can be improved, but it suffices for our main point and allows for simple computations. By the preceding lemma we have $1 \leq k \leq n$ such that

$$a(k) \geq (1-\delta)\sqrt{n} \text{ and } b(k) \geq (1-\delta)\sqrt{n}.$$

By considering subsequences that go through x_k and continue from x_j with $n \leq j \leq 2n$ we see that

$$M(x_1, x_2, \ldots, x_{2n}) \geq (1-\delta)\sqrt{n} + \min\{c_{2n}(j), d_{2n}(j)\},$$

so, by the bound on $M(x_1, x_2, \ldots, x_{2n})$, we find

$$\min\{c_{2n}(j), d_{2n}(j)\} < (1+\varepsilon)\sqrt{2n} - (1-\delta)\sqrt{n} \tag{8.2}$$

for all $n < j \leq 2n$. Now, unless the conclusion of the proposition holds we also have

$$\max\{c_{2n}(j), d_{2n}(j)\} \leq (1-25\delta)\sqrt{2n}, \tag{8.3}$$

so the issue is to estimate the number of $n < j \leq 2n$ that can satisfy (5.2) and (5.3). Since the mapping $g \mapsto (c_{2n}(j), d_{2n}(j))$ is injective, the proposition follows if there are fewer than n solutions of (5.2) and (5.3).

We need to count the number of positive lattice points (i, j) with

$$\min\{i, j\} < \{(1+\varepsilon)\sqrt{2} - 1 + \delta\}\sqrt{n}$$

and

$$\max\{i, j\} < \sqrt{2}(1-25\delta)\sqrt{n}.$$

A calculation shows that the area of the L-shaped region defined by these inequalities is bounded by $n\{1-\delta\} < n$. □

To complete the proof of the main theorem of this section, we first note that setting $\delta^2 = 2\varepsilon$ and solving $1 + \varepsilon = (1 - 25\delta 0\sqrt{2}$ we are lead to a value of $\delta = 0.117118$. This tells us that in the theorem we can take any $j \leq 1 + \delta^2/2 \leq 1.00014$.

The key problem that remains open at this point is the determination of the best possible value of j.

9. Subsequences along cycles. We say that a sequence of integers $(i_1, i_2, \ldots, i_k)$ has d *descents* if it can be written as the concatination of d monotonic blocks but cannot be written as a concatination of fewer than d monotonic blocks.

For example, the sequence $(5, 9, 11, 2, 3, 8, 4, 1)$ can be written as the concatination of four monotonic increasing blocks $(5, 9, 11)$, $(2, 3, 8)$, (4), (1) and thus the sequence has $d = 4$ descents. We let

$$\ell^+(d; x_1, x_2, \ldots, x_n) =$$
$$\max\{k : x_{i_1} < x_{i_2} < \ldots < x_{i_k} \text{ where } (i_1, i_2, \ldots, i_k) \text{ has } d \text{ descents}\}.$$

We define $\ell^-(d; x_1, x_2, \ldots, x_n)$ as the corresponding maximum length d-descent monotone sequence of x_i's. The purpose of introducing these quantities is that they lead to a quite natural analog of the Erdős-Szekeres theorem.

THEOREM 9.1. *For any n distinct real numbers, we have*

$$\ell^+(d; x_1, x_2, \ldots, x_n)\ell^-(d; x_1, x_2, \ldots, x_n) \geq dn.$$

Proof. We first remark that the case $d = 1$ is the usual Erdős-Szekeres theorem. Further in the trivial case $d = n$, we have equality since $\ell^+ = \ell^- = n$. □

10. Common ascending subsequences. If π and σ are two permutations of $\{1, 2, \ldots, n\}$, we say they have a *common ascending subsequence of length r* if there are indices $1 \leq i_1 < i_2 < \ldots < i_r \leq n$ and $1 \leq j_1 < j_2 < \ldots < j_r \leq n$ such that $\pi(i_s) = \sigma(j_s)$ for all $1 \leq s \leq r$. The notion of the longest common ascending subsequence $\lambda(\pi, \sigma)$ was introduced in Alon (1990) for the purpose of derandomizing the randomized maximum flow algorithm of Cheriyan and Hagerup (1989). The connection between $\lambda(\pi, \sigma)$ and the central theme of this review is made most explicit by noting that $\lambda(\pi, \sigma)$ is equal to the length of the longest increasing subsequence of $\sigma^{-1}\pi(1), \sigma^{-1}\pi(2), \ldots, \sigma^{-1}\pi(n)$. The main result of Alon (1990) is the following theorem.

THEOREM 10.1. *For every two integers k and n with $k \geq n$, one can construct in time $0(kn)$, a sequence of permutations $\pi_1, \pi_2, \ldots, \pi_n$ of $\{1, 2, \ldots, n\}$ such that any permutation σ satisfies*

$$\frac{1}{k}\sum_{i=1}^{k} \lambda(\sigma, \pi_i) = 0(n^{2/3}).$$

The fact that this result is reasonably sharp can be deduced from the probabilistic results of the first section. For any fixed $\pi_1, \pi_2, \ldots, \pi_n$, we have for a random σ that

$$E\frac{1}{k}\sum_{i=1}^{k} \pi(\sigma, \pi_i) = EI_n \geq (2 - \varepsilon)\sqrt{n}$$

for any $\varepsilon > 0$ and $n \geq n(\varepsilon)$. Combining this observation with Alon's theorem with $k = n$ we find that there is a constant $c > 0$ such that the functional

$$A_n = \max_{\sigma} \min_{\{\pi_i\}} \frac{1}{n} \sum_{i=1}^{n} \lambda(\sigma, \pi_i)$$

satisfies

$$(2 - \varepsilon)\sqrt{n} \leq A_n \leq cn^{2/3}.$$

The main point of reviewing this information is to point out the problem of determining the true order of A_n.

11. Optimal sequential selection. One of the intriguing themes of sequential selection is that one sometimes does surprisingly well in making selections even without knowledge of the future or recourse to change past choices. One illustration of the theme is the "secretary problem" of Gilbert and Mosteller (1966) that tells us that given $X_1, X_2, \ldots, X_n$ independent and identically distributed random variables, there is a stopping time $\tau = \tau_n$ such that

$$P(X_\tau = \max_{1 \leq i \leq n} X_i) > e^{-1}.$$

There turns out to be a parallel phenomenon that takes place in the theory of monotone subsequences.

The problem is to determine how well one can make a sequence of choices from a sequentially reveal set of independent random variables with a known common continuous distribution. To put this problem rigorously, we call a sequence of stopping times $1 \leq \tau_1 < \tau_2 < \ldots$ a *policy* if they are adapted to $X_1, X_2, \ldots$ and if we have $X_{\tau_1} < X_{\tau_2} < \ldots < X_{\tau_k} < \ldots$. We let $\mathcal{S}$ denote the set of all policies. The main result about such policies is given in Samuels and Steele (1981).

THEOREM 11.1. *For any sequence of independent random variables with continuous distribution F and associated set of policies $\mathcal{S}$, we have*

$$u_n = \sup_{\mathcal{S}} E(\max\{k : \tau_k \leq n\}) \sim \sqrt{2n}$$

as $n \to \infty$.

This theorem thus tells us that there is a policy by which we can make sequential selections from $X_1, X_2, \ldots, X_n$ and obtain a monotone increasing subsequence of length that is asymptotic in expectation to $\sqrt{2n}$. Since the best one could do with full knowledge of $\{X_1, X_2, \ldots, X_n\}$ is to obtain a subsequence with length that is asymptotic in expectation to $2\sqrt{n}$, we see that a "mortal" does worse than a "prophet" by only a factor of $\sqrt{2}$.

It was observed by Burgess Davis (cf. Samuels and Steele (1981), Section 7), that if one lets ℓ_n denote the expected value of length of the longest increasing subsequence that can be made sequentially from a random permutation of $\{1, 2, \ldots, \}$, then one has $\ell_n \sim u_n$ and consequently $\ell_n \sim \sqrt{2n}$. The importance of this remark comes from the fact that the "natural" selection studied in Baer and Brock (1968) is just "sequential" selection as studied here. Thus, the results of Samuels and Steele (1981) coupled with the key observation of Burgess Davis resolve the main problem posed in Baer and Brock (1968).

12. Open problems and concluding remarks. Erdős (see Chung (1980), p. 278) has raised the question of determining the optimum values associated with various *weighted* versions of the monotone and k-modal subsequence problems. Specifically, let $\mathcal{W} = \{(w_1, w_2, \ldots w_n) : w_i \geq 0 \text{ and } w_1 + w_2 + \ldots + w_n = 1\}$ and for $w \in \mathcal{W}$ and for $0 \leq k$ let $\mathcal{U}(w)$ denote the set of k-unimodal subsequences of w. The problem is to determine the values

$$\tau(n,k) = \min_{w \in \mathcal{W}} \max_{u \in \mathcal{U}(w)} \sum_{w_i \in u} w_i.$$

To illustrate, we note that it is easy to show that $\tau(n, 0) \leq n^{1/2}$, since by considering a perturbation of uniform weights that have the same ordering which yields a longest monotone subsequence of length $\lceil u^{1/2} \rceil$, we see $\tau(n, 0) \leq n^{-1} \lceil n^{1/2} \rceil$.

One surely suspects that $\tau(n, 0)\sqrt{n} \to 1$ as $n \to \infty$, but this has not yet been established, though it might be easy. By similar considerations using Chung's theorem, one sees that limsup $\tau(n, 1)\sqrt{n} \leq \sqrt{3}$ while we expect that actually $\tau(n, 1) \to \sqrt{3}$.

Erdős posed the problem of determining the largest integer $f(n)$ such that any sequence of $m = f(n)$ distinct real numbers $x_1, x_2, \ldots, x_m$ can be decomposed into n monotonic sequences. Hanani (1957) proved that

$$f(n) = n(n+3)/2.$$

A question posed by Erdős (1973) for which there seems to have been no progress is the following:

Given $x_1, x_2, \ldots, x_n$ distinct real numbers determine

$$\max_{M} \sum_{i \in M} x_i$$

where the maximum is over all subsets of indeces $1 \leq i_1 < i_2 < \ldots < i_k \leq n$ such that $x_{i_1}, x_{i_2}, \ldots, x_{i_k}$ is monotone.

Acknowledgement. I am pleased to thank D. Aldous, P. Diaconis and M. Talagrand for making available prepublication copies of their work. I also thank L. Lovász for comments that contributed to the discussion in Section nine, S. Janson for comments that contributed to Section six, and P. Winkler for comments that contributed to Section three. P. Erdős provided

the reference to the work of Hanani, as well as the inspiration to pursue this survey.

REFERENCES

[1] D. Aldous, *Reversible Markov chains and random walks on graphs*, Monograph Draft, version of January 25, 1993, Department of Statistics, University of California, Berkeley, CA (1993).

[2] D. Aldous and P. Diaconis, *Hammersley's interacting particle process and longest increasing subsequences*, Technical Report, Department of Statistics, U.C. Berkeley, Berkeley, CA (1993).

[3] N. Alon, *Generating pseudo-random permutations and maximum flow algorithms*, Information Processing Letters 35 (1990), pp. 201–204.

[4] N. Alon, B. Bollobás, G. Brightwell and S. Janson, *Linear extensions of a random partial order*, to appear in Ann. Appl. Prob. 3 (1993).

[5] R.M. Baer and P. Brock, *Natural sorting over permutation spaces*, Math. Comp. 22 (1968), pp. 385–510.

[6] J. Beck, *Randomness of* $n\sqrt{2} \bmod 1$ *and a Ramsey property of the hyperbola*, Colloquia Mathematica Societatis János Bolyai 60 Sets, Graphs, and Numbers, Budapest, Hungary (1991).

[7] J. Beck and W.W.L. Chen, *Irregularities of Distribution*, Cambridge University Press, 1987.

[8] R. Béjian and H. Faure, *Discrépance de la suite de van der Corput*, C.R. Acad. Sci. Paris Sér. A 285 (1977), pp. 313–316.

[9] P. Blackwell, *An alternative proof of a theorem of Erdős and Szekeres*, Amer. Mathematical Monthly 78 (1971), p. 273.

[10] K.P. Bogart, C. Greene, and J.P. Kung, *The impact of the chain decomposition theorem on classical combinatorics*, in The Dilworth Theorems: Selected Papers of Robert P. Dilworth (K.P. Bogart, R. Freese, and J.P. Kung, eds.), Birkhäuser Publishers, Boston (1990), pp. 19–29.

[11] P. Bohl, *Über ein in der Theorie der säkularen Störunger vorkommendes Problem*, J. Reine U. Angew. Math. 135 (1909), pp. 189–283.

[12] B. Bollobás and G. Brightwell, *The height of a random partial order: Concentration of measure*, Ann. Appl. Prob. 2 (1992), pp. 1009–1018.

[13] B. Bollobás and P.M. Winkler, *The longest chain among random points in Euclidean space*, Proc. Amer. Math. Soc. 103 (1988), pp. 347–353.

[14] G. Brightwell, *Models of Random Partial Orders in Surveys in Combinatorics*, 1993, London Mathematical Society Lecture Notes 187 (K. Walker, ed.), (1993).

[15] J. Cheriyan and T. Hagerup, *A randomized maximum-flow algorithm*, Proceedings of the IEEE-FOCS 1989 (1989), pp. 118–123.

[16] F.P.K. Chung, *On Unimodal Subsequences*, J. Combinatorial Theory (A) 29 (1980), pp. 267–279.

[17] J.G. van der Corput, *Verteilungs funktionen, I,II*, Nederl. Akad. Wetensch. Proc. 38 (1935), pp. 813–821, 1058–1066.

[18] A. Del Junco and J.M. Steele, *Hammersley's Law for the van der Corput Sequence: An Instance of Probability Theory for Pseudorandom Numbers*, Annals of Probability 7 (2) (1979), pp. 267–275.

[19] R. Durrett, *Probability: Theory and examples*, Wadsworth, Brooks/Cole, Pacific Grove, CA (1991).

[20] R.P. Dilworth, *A Decomposition Theorem for Partially Ordered Sets*, Annals of Mathematics 51 (1950), pp. 161–165.

[21] P. Erdős, Ko Chao, and R. Rado, *Intersection theorems for systems of finite sets*, Quarterly J. Mathematics Oxford Series (2) 12 (1961), pp. 313–320.

[22] P. Erdős, *Unsolved problems in Graph Theory and Combinatorial Analysis*, Combinatorial Mathematics and It's Applications (Proc. Conf. Oxford 1969), Academic Press, London (1971), pp. 97–109. Also in *Paul Erdős: The Art of Counting*, (J. Spencer, ed.), MIT Press, Cambridge, MA (1973).

[23] P. Erdős and R. Rado, *Combinatorial theorems on classifications of subsets of a given set*, Proc. London Math Soc. 3 (2) (1952), pp. 417–439.

[24] P. Erdős and Gy. Szekeres, *A combinatorial problem in geometry*, Compositio Math. 2 (1935), pp. 463–470.

[25] A.M. Frieze, *On the length of the longest monotone subsequence in a random permutation*, Ann. Appl. Prob. 1 (1991), pp. 301–305.

[26] T. Gallai and N. Milgram, *Verallgemeinerung eines graphentheoretischen Satzes von Reide*, Acta Sci. Math. 21 (1960), pp. 181–186.

[27] J.P. Gilbert, and F. Mosteller, *Recognizing the maximum of a sequence*, J. Amer. Statist. Soc. 61 (1966), pp. 228–249.

[28] J.M. Hammersley, *A few seedlings of research*, Proc. 6th Berkeley Symp. Math. Stat. Prob., U. of California Press (1972), pp. 345–394.

[29] J.M. Hammersley and D.C. Handscomb, *Monte Carlo Methods*, Methuen Publishers, London (1964).

[30] H. Hanani, *On the number of monotonic subsequence*, Bull. Res. Council Israel, Sec. F (1957), pp. 11–13.

[31] H. Kesten, *Uniform distribution* mod 1, Annals of Math. 71 (1960), pp. 445–471.

[32] J.F.C. Kingman, *Subadditive ergodic theory*, Ann. Prob. 1 (1973), pp. 883–909.

[33] D. Knuth, *Permutations, matrices, and generalized Young tableaux*, Pacific J. Mathematics 34 (1970), pp. 709–727.

[34] W. Kraskiewicz, *Reduced decompositions in hyperoctahedral groups*, C.R. Acad. Sci. Paris 309 (1989), pp. 903–907.

[35] V. Lifschitz and B. Pittel, *The number of increasing subsequences of the random permutation*, J. Comb. Theory Ser. A 31 (1981), pp. 1–20.

[36] B.F. Logan and L.A. Shepp, *A variational problem for Young tableaux*, Advances in Mathematics 26 (1977), pp. 206–222.

[37] L. Lovász, *Combinatorial Problems and Exercises*, North-Holland Publishing Company, Amsterdam (1979).

[38] H. Niederreiter, *Quasi-Monte Carlo methods and pseudo-random numbers*, Bull. Amer. Math. Soc. 84 (1978), pp. 957–1041

[39] S. Pilpel, *Descending subsequences of random permutations*, J. Comb. Theory, Ser. A 53 (1) (1990), pp. 96–116.

[40] S. Samuels and J.M. Steele, *Optimal Sequential Selection of a Monotone Subsequence from a random sample*, Ann. Prob. 9 (1981), 937–947.

[41] C. Schensted, *Longest increasing and decreasing subsequences*, Canadian J. Math. 13 (1961), pp. 179–191.

[42] A. Seidenberg, *A simple proof of a theorem of Erdős and Szekeres*, J. London Math. Soc. 34 (1959), p. 352.

[43] R.S. Silverstein, *Estimation of a constant in the theory of three dimensional random orders*, BS/MS Thesis in Mathematics (1988), Emory University.

[44] D. Stanton, and D. White, *Constructive Combinatorics*, Springer-Verlag, New York (1986).

[45] J.M. Steele, *Limit properties of random variables associated with a partial ordering of R^d*, Ann. Prob. 5 (1977), pp. 395–403.

[46] J.M. Steele, *Long unimodal subsequences: A problem of F.R.K. Chung*, Discrete Mathematics 33 (1981), pp. 223–225.

[47] J.M. Steele, *An Efron-Stein inequality for non-symmetric statistics*, Annals of Statistics 14 (1986), pp. 753–758.

[48] M. Talagrand, *An isoperimetric theorem on the cube and the Khintchin-Kahane inequalities*, Proc. Amer. Math. Soc. 104 (1988), pp. 905–909.

[49] M. Talagrand, *A new isoperimetric inequality for product measure and the tails of sums of independent random variables*, Geometric and Functional Analysis, 1 (2) (1991), pp. 211–223.

[50] M. Talagrand, *Concentration of measure and isoperimetric inequalities in a product space*, Technical Report, Department of Mathematics, Ohio State University, Columbus, OH (1993).

[51] W.T. Trotter, *Combinatorics and Partially Ordered Sets: Dimension Theory*, The Johns Hopkins University Press, Baltimore (1992), pp. xiv + 307.

[52] S.M. Ulam, *Monte Carlo calculations in problems of mathematical physics*, in Modern Mathematics for the Engineer (E.F. Beckenbach, ed.) McGraw-Hill, New York (1961).

[53] A.M. Veršik and S.V. Kerov, *Asymptotics of the Plancherel measure of the symmetric group and the limiting form of Young tableaux*, Dokl. Akad. Nauk. SSSR 233 (1977), pp. 1024–1028.

[54] A.M. Veršik and S.V. Kerov, *Asymptotics of the largest and the typical dimensions of irreducible representations of the symmetric group*, Functional Analysis and Its Applications 19 (1985), pp. 21–31.

[55] D.B. West, *Extremal Problems in partially ordered sets*, in Ordered Sets: Proceedings of the Banff Conference (I. Rival, eds.) Reidel, Dordrecht and Boston (1982), pp. 473–521.