On the Subsequence Theorem of Erdős and Szekeres

Jan van Leeuwen

Technical Report UU-CS-2016-004
April 2016

Department of Information and Computing Sciences
Utrecht University, Utrecht, The Netherlands
www.cs.uu.nl

ISSN: 0924-3275

Department of Information and Computing Sciences
Utrecht University
P.O. Box 80.089
3508 TB Utrecht
The Netherlands

On the Subsequence Theorem of Erdős and Szekeres

Jan van Leeuwen

Dept. of Information and Computing Sciences, Utrecht University
Princetonplein 5, 3584 CC Utrecht, The Netherlands.
j.vanleeuwen1@uu.nl

Abstract. In the early 1950s, Kruskal generalized the well-known subsequence theorem of Erdős and Szekeres to finite sequences over all domains satisfying his requirement of 'relation spaces'. We give a Ramsey-theoretic perspective on the subsequence theorem and generalize it to sequences over all domains with a finite number of binary relations, without constraints. We show that Kruskal's generalization is equivalent to this more general Ramsey version, giving an elegant mathematical basis to the former. The critical bounds in the results are proved to be best possible, even for domains with partial orders only. The Ramsey version subsumes various earlier generalizations of the subsequence theorem and allows for many new applications. For example, the theorem of Erdős and Szekeres is a special case of various interesting graph-theoretic properties.

Keywords: monotone subsequences, Erdős-Szekeres theorem, Ramsey's theorem, binary relations, posets, De Bruijn's theorem, graphs, tournaments.

"The Basic Theorem is not in essence a statement about the real number system."
J.B. Kruskal, 1952 [22]

1 Introduction

In 1935, Erdős and Szekeres [10] proved the following elegant result in their study of convex configurations among n points in the plane: *if* $\sigma = \sigma_1, \cdots, \sigma_n$ *is a sequence of* n^2+1 *distinct numbers, then* σ *contains an increasing subsequence of* $n+1$ *numbers or a decreasing subsequence of* $n+1$ *numbers (or both).* The result is an early gem in extremal combinatorics.

Several appealing proofs of the theorem have since appeared (cf. [36]). The result also follows from Dilworth's Lemma for partially ordered sets P: *if* $|P| \geq w \cdot s + 1$, *then* P *has an anti-chain of size at least* $w+1$ *or a chain of length at least* $s+1$, *or both* [7]. Just consider $P = \{\sigma_1, \cdots, \sigma_n\}$ with $\sigma_i \preceq \sigma_j$ if and only if $i \leq j$ and $\sigma_i \leq \sigma_j$, and apply the lemma with $w = s = n$ (cf. [20]).

In the nineteen fifties, J.B. Kruskal [22] noted that the subsequence theorem is *not* just a theorem about numbers or partial orders. In a first generalization he considered a geometric notion of monotonicity and the occurrence of monotone subsequences in finite-dimensional vector spaces. In a second generalization, Kruskal argued that the subsequence theorem can be seen as a result for finite sequences in all, what he called, *relation spaces* with finitely many binary relations.

Kruskal's second generalization is commonly applied to finite sequences of vectors, but the result still appears to be less general than desired. One reason may be that Kruskal imposed a strong assumption ('axiom') on the relations in a 'relation space', restricting its apparent generality. We will argue that the limiting assumption can be removed and that Kruskal's generalization is essentially a much broader result.

For a proper perspective, we turn to Ramsey theory [16, 19]. This theory does not deal with sequences *per se*. However, it is well-known that Ramsey's Theorem can be used to prove the 'infinite version' of the subsequence theorem: *any infinite sequence of distinct real numbers contains an infinite increasing subsequence or an infinite decreasing subsequence, or both.* In Section 2 we show that this fact neatly generalizes to sequences over arbitrary domains X with finitely many binary relations, *without* restriction.

The proper generalization of the Erdős-Szekeres theorem for all finite sequences that are sufficiently long is now immediate from the same approach, using the finite version of Ramsey's Theorem. However, for obtaining the precise 'Ramsey number' depending on the various parameters in the result, a different proof is needed. In Section 2 we provide the missing link. The proof builds on Seidenberg's combinatorial proof of the Erdős-Szekeres theorem, thus showing the full potential of this proof [34].

The 'Ramsey version' of the Erdős-Szekeres theorem holds for all domains with multiple binary relations, without limiting assumptions. Nevertheless, we will argue in Section 3 that Kruskal's generalization is *equivalent* to the Ramsey version of the subsequence theorem. It puts Kruskal's generalization in a new perspective and gives it the broader scope we claimed. We will show that several other restricted versions are equivalent to the general Ramsey version as well.

The Ramsey version unifies and extends many other generalizations of the Erdős-Szekeres theorem. In Section 2 we show this for partially ordered sets, in Section 4) for tuple spaces $X = X_1 \times \cdots \times X_d$ ($d \geq 1$). We show e.g. that Kruskal's generalization of De Bruijn's theorem is a direct consequence of the general theorem as well. In Section 5 we apply the Ramsey version to derive various 'subsequence theorems' for graphs like rooted trees and tournaments.

In Section 6 we show that the Ramsey bound in the generalized version of the Erdős-Szekeres theorem is best possible for all parameter values. In fact, we prove the that the result is already best possible for spaces with *multiple partial orders* only. In Section 7 we give some conclusions.

The subsequence theorem of Erdős and Szekeres has been generalized in various contexts, geometric and otherwise [1, 4, 5, 8, 11, 13, 24, 26, 28, 32, 33, 35, 38]. In the present paper we attempt to record the full generality of Kruskal's result from the broader, Ramsey-theoretic perspective.

2 Generalizing the Subsequence Theorem

In this Section we first formulate and prove an analogue of the subsequence theorem for *infinite* sequences over spaces with multiple binary relations. Next we prove the corresponding result for *finite* sequences of the same generality. Then we digress on the interesting case of spaces with multiple partial orders, which is historically and practically relevant. Optimality of the bounds will be shown in Section 6.

Terminology In order to state the results, we need some terminology. Let X be an arbitrary non-empty set and $R \subseteq X \times X$ a binary relation on X. The inverse of R is $R^{-1} = \{(b, a) \mid (a, b) \in R\}$. We write $(a, b) \in R$ as aRb. We say that $a, b \in X$ are *related* in R if aRb or bRa (or both), otherwise we say that a and b are *unrelated* in R. The elements of $S \subseteq X$ are said to be (mutually) unrelated in R if ($|S| = 1$ or) when any two elements $a, b \in S$ with $a \neq b$ are unrelated in R. Given a sequence σ of distinct elements from X, a subsequence $\sigma_{i_1}, \cdots, \sigma_{i_k}$ ($1 \leq i_1 < \cdots < i_k \leq n$) is called *ascending* in R if $\sigma_{i_1} R \cdots R \sigma_{i_k}$. It is called *descending* in R if

$\sigma_{i_1} R^{-1} \cdots R^{-1} \sigma_{i_k}$ or, equivalently, if $\sigma_{i_k} R \cdots R \sigma_{i_1}$. The terms are consistent with those for partial orders. (NB: Ascending and descending subsequences in a relation R need not be cycle-free when R is not a partial order.)

2.1 Generalization for Infinite Sequences

The infinite version of the subsequence theorem is usually stated for sequences of reals. This version is well-known from the proof of various convergence results in real analysis such as the *Bolzano-Weierstraß* theorem.

Burkill and Mirsky [3] noted in 1973 that the infinite version can be seen as an application of Ramsey's Theorem. Newman and Parsons [27] later observed the same fact and noted that their argument did not make use of any special property of the reals, except that they form an (infinite) linearly ordered set. Hence they concluded the following, more general fact: *any infinite sequence of elements from any linearly ordered set has a monotone subsequence.*

Remark. The result goes further than the corresponding 'set version' which asserts that *every infinite linear order contains an infinite ascending sequence or an infinite descending sequence (or both).* This fact is known as the *Ascending or Descending Sequence principle* (**ADS**) [18, 23]. We return to this principle below.

For a more general result, we make the step from linearly ordered sets to sets with any finite number of binary relations. The following fact may be observed, which expands on [27] and [16] (p. 17-18) and is 'typically Ramsey'.

Theorem 0. *Let σ be an infinite sequence of distinct elements from domain X and let $\{R_1, \cdots, R_q\}$ be a collection of binary relations over X. Then σ contains an infinite subsequence whose elements are unrelated in each of the relations $R_1, \cdots, R_q$ or σ contains an infinite subsequence that is ascending or descending in at least one of the R_i $(1 \leq i \leq q)$.*

Proof. Let σ and $\{R_1, \cdots, R_q\}$ be as given. Put the 2-subsets $[i, j]$ with $i \neq j$ into bins $B_0, B_1, C_1, \cdots, B_q, C_q$ as follows:

- if σ_i and σ_j are unrelated under each of the relations $\{R_1, \cdots, R_q\}$ then put $[i, j]$ into bin B_0.
- if σ_i and σ_j are related under any of the relations $\{R_1, \cdots, R_q\}$ then assume that the elements of $[i, j]$ are ordered such that $i < j$ and do the following:
  - if $\sigma_i R_k \sigma_j$ and k is the smallest index $(1 \leq k \leq q)$ such that this holds, then put $[i, j]$ into B_k.
  - if $\sigma_i R_k \sigma_j$ does *not* hold for any k $(1 \leq k \leq q)$, then let $\sigma_i R_l^{-1} \sigma_j$ where l is the smallest index $(1 \leq l \leq q)$ such that the latter holds. (Such an l must exist.) Then put $[i, j]$ into C_l.

This prescription partitions the collection of all 2-subsets of $\mathbb{N}$ into $2q + 1$ bins. By Ramsey's Theorem [29], there must exist an infinite subset $S \subseteq \mathbb{N}$ such that the 2-subsets of S are all contained in the same bin. Let $s_1 < s_2 < \cdots$ be a listing of S, in sorted order. Now the following cases can arise:

- all 2-subsets of S are contained in B_0: then the elements of $\sigma_{s_1}, \sigma_{s_2}, \cdots$ are unrelated in each of the relations R_i $(1 \leq i \leq q)$.
- all 2-subsets of S are contained in B_k for some k with $1 \leq k \leq q$: then $\sigma_{s_1}, \sigma_{s_2}, \cdots$ is ascending in R_k.
- all 2-subsets of S are contained in C_l for some l with $1 \leq l \leq q$: then $\sigma_{s_1}, \sigma_{s_2}, \cdots$ is descending in R_l.

Hence $\sigma_{s_1}, \sigma_{s_2}, \cdots$ is an infinite subsequence of σ as required. □

Theorem 0 can be formulated also in case the elements of σ are not all distinct, but we will not digress on this. We mention the following application, which extends the observation from [27]: *every infinite sequence of distinct elements from a poset X contains an infinite subsequence whose elements form an anti-chain in X or an infinite subsequence whose elements form a chain in X, or both.*

Remark. The observation for posets is noted in the literature (cf. [6], p. 242). The result goes further than the corresponding, but more commonly stated 'set version' which only asserts that *every infinite poset contains an infinite anti-chain or an infinite chain (or both).* The latter is known as the *Chain-AntiChain principle* (**CAC**). Clearly, **Ramsey's Theorem** $\Rightarrow$ **CAC** $\Rightarrow$ **ADS**, but it is known that the reverse implications are problematic [18, 23].

2.2 Generalization for Finite Sequences

Theorem 0 gives the template for the generalization of the Erdős-Szekeres theorem for finite sequences. Here we formulate and prove the generalization for arbitrary sets with any finite number of binary relations defined on them. We implicitly assume that all numeric parameters are integers ≥ 1.

Theorem 1. *Let σ be a sequence of n distinct elements and $\{R_1, \cdots, R_q\}$ a collection of binary relations over X. If $n \geq w \cdot s_1 \cdots s_q \cdot r_1 \cdots r_q + 1$, then σ contains a subsequence of length $w+1$ whose elements are unrelated in* each *of the relations $R_1, \cdots, R_q$ or there is an i $(1 \leq i \leq q)$ such that σ contains an ascending subsequence of length $s_i + 1$ or a descending subsequence of length $r_i + 1$ in R_i.*

Proof. If σ has a subsequence of size $w+1$ whose elements are mutually unrelated in each of the R_i's, then we are done. Thus, for the remainder of the proof we may assume that no subsequence of σ can have more than w elements that are mutually unrelated in each of the binary relations. Write $\sigma = \sigma_1, \cdots, \sigma_n$.

For every position i of σ, define the $2q$-tuple $[u_{i1}, d_{i1}, \cdots, u_{iq}, d_{iq}] \in \mathbb{N}^{2q}$ with u_{it} equal to the length of the longest ascending subsequence according to relation R_t ending at i, and d_{it} equal to the length of the longest descending subsequence according to R_t beginning at i $(1 \leq t \leq q)$. Clearly $u_{it} \geq 1$ and $d_{it} \geq 1$ for every i and t.

Suppose that $u_{it} \leq s_t$ and $d_{it} \leq r_t$ for all i $(1 \leq i \leq n)$ and t $(1 \leq t \leq q)$. As there can be at most $s_1 \cdots s_q \cdot r_1 \cdots r_q$ different $2q$-tuples with these constraints, it follows by the pigeonhole principle that there must at least $w+1$ positions of σ that all have the same tuple. Say these positions are $i_1, \cdots, i_{w+1}$ with $i_1 < \cdots < i_{w+1}$.

By assumption, the set $\{\sigma_{i_1}, \cdots, \sigma_{i_{w+1}}\}$ cannot all consist of elements that are unrelated in every one of the relations $R_1, \cdots, R_q$. Hence there must be a t $(1 \leq t \leq q)$ and positions $i_j < i_k$ such that σ_{i_j} and δ_{i_k} are *related* in the relation P_t. Assume that $\delta_{i_j} R_t \delta_{i_k}$. Then the longest ascending subsequence according to relation R_t ending at position i_k must be longer than the one ending at position i_j by at least one (namely by the term δ_{i_k}), contradicting the fact that the $2q$-tuples at both positions are equal. The case where $\delta_{i_k} R_t \sigma_{i_j}$ similarly leads to a contradiction.

It follows that there must be a position i where at least one of the components of the associated $2q$-tuple does not stay within the bound, i.e. there must be a t $(1 \leq t \leq q)$ such that $u_{it} > s_t$ or $d_{it} > r_t$. This completes the proof. □

In keeping with the Ramsey-theoretic nature of the result, the bound on n in Theorem 1 does not depend on any characteristic of X or the binary relations. The proof of Theorem 1 generalizes the elegant proof of the Erdős-Szekeres theorem by Seidenberg [34] and the extension of the latter by Kruskal to relation spaces [22].

The original Erdős-Szekeres theorem can be retrieved by considering the case of *ordered* sets, i.e. sets with a single relation R that is a total order. These sets clearly have no anti-chains of size greater than 1. The Erdős-Szekeres theorem then follows from Theorem 1 by taking $q = w = 1$.

2.3 Partially Ordered Sets

Theorem 1 takes an attractive form when applied to sets with *multiple partial orders*. In the terminology of posets, we obtain to the following result.

Theorem 1′. *Let σ be a sequence of n distinct elements and $\{P_1, \cdots, P_q\}$ a collection of partial orders over X. If $n \geq w \cdot s_1 \cdots s_q \cdot r_1 \cdots r_q + 1$, then σ contains a subsequence of length $w+1$ whose elements form an anti-chain in* each *of the partial orders $P_1, \cdots, P_q$ or there is an i $(1 \leq i \leq q)$ such that σ contains an ascending subsequence of length $s_i + 1$ or a descending subsequence of length $r_i + 1$ in P_i.*

Note that the most general formulation of the Erdős-Szekeres theorem found in most textbooks is the following: *Let σ be a sequence of n distinct numbers. If $n \geq sr+1$, then σ contains an increasing subsequence of length $s+1$ or a decreasing subsequence of length $r+1$* (cf. [20],[22]). Theorem 1′ implies the following stronger form for posets.

Corollary 1. *Let σ be a sequence of n distinct elements from poset P. If $n \geq wsr + 1$, then σ contains an anti-chain of size $w + 1$, or an ascending subsequence of length $s + 1$, or a descending subsequence of length $r + 1$.*

Corollary 1 generalizes the Erdős-Szekeres theorem in an appreciable way. It leads to easy proofs of some further consequences as well. Recall that the *width* of a poset is the size of its largest anti-chain.

Corollary 2. *Let P have width at most w, and let σ be a sequence of n distinct elements from P. Then σ contains a monotone subsequence of length $\lceil \sqrt{\frac{n}{w}} \rceil$.*

Proof. Let $t = \lceil \sqrt{\frac{n}{w}} \rceil - 1$. The result follows directly from Corollary 1 once we show that $n \geq wt^2 + 1$. To prove it, write $\sqrt{\frac{n}{w}} = a + \epsilon$, for integer a and $0 \leq \epsilon < 1$. If $\epsilon = 0$, then $a \geq 1$ and $t = a - 1$, hence $n = w \cdot a^2 \geq wa^2 - 2aw + w + 1 = wt^2 + 1$. If $\epsilon > 0$, then $t = a$, and $n = w(a + \epsilon)^2 > w \cdot a^2$. As n is integer, it follows that $n \geq w \cdot a^2 + 1 = wt^2 + 1$ again. □

Corollary 1 also implies Dilworth's Lemma. The lemma is usually derived from Dilworth's Decomposition Theorem [7], but it can also be proved directly [20].

Corollary 3. *Let S be any subset of poset P with at least n elements. If $n \geq ws+1$, then S contains an anti-chain of size $w + 1$ or a chain of length $s + 1$.*

Proof. Assume w.l.o.g. that S has n elements. Sort S topologically, and let the resulting sequence be σ. By construction, σ does not contain decreasing subsequences of length 2. Now apply Corollary 1 with $r = 1$. □

Finally, we note that Theorem 1′ can be proved fully within the theory of posets, just like the original Erdős-Szekeres theorem. The proof is non-elementary, as it uses Dilworth's Decomposition Theorem [7] and the Erdős-Szekeres theorem itself as prerequisites. We induct on q, with Corollary 1 as the base case.

Alternative proof of Corollary 1. For $q = 1$, assume σ does not contain anti-chains of size greater than w. By Dilworth's Theorem, the elements of σ (viewed as a poset) can be partitioned into at most w disjoint chains. By the pidgeonhole principle it follows that σ must contain a subsequence of length at least $\lceil \frac{1}{w}(wsr + 1) \rceil \geq sr + 1$ whose terms all belong to the same chain i.e. a total order. The corollary now follows by an application of the original Erdős-Szekeres theorem to the latter. □

Alternative proof of Theorem 1′. Assume by way of induction that the theorem holds for up to $q-1$ partial orders, for some $q > 1$. Now consider a sequence σ of n distinct elements from a domain with q partial orders $P_1, \cdots, P_{q-1}, P_q$, and assume that $n \geq w \cdot s_1 \cdots s_q \cdot r_1 \cdots r_q + 1$. By the induction base (Corollary 1) it follows that σ contains a subsequence of length $n' \geq w \cdot s_1 \cdots s_{q-1} \cdot r_1 \cdots r_{q-1} + 1$ whose elements form an anti-chain in P_q, or σ contains an ascending subsequence of length $s_q + 1$ or a descending subsequence of length $r_q + 1$ in P_q. In the latter two cases we are done. In the former case, apply the induction hypothesis to the subsequence of length n' and notice that any anti-chain in it is also an anti-chain in P_q. This completes the induction. □

Theorem 1′ is the general Ramsey version of the Erdős-Szekeres theorem for the case of sets with multiple partial orders. In Section 6 we will show that it is an optimal result, in the sense that the bound in it is tight.

3 Kruskal's Generalization Revisited

In this Section we will show that Kruskal's second generalization of the Erdős-Szekeres theorem is equivalent to the Ramsey version (Theorem 1). The result sheds an interesting light on Kruskal's generalization, as the Ramsey version does not require any special constraints.

We prove the equivalence in Section 3.2. In Section 3.3 we show that the equivalence can be extended to several other cases of restricted relation spaces as well. In Section 3.4 we discuss a further case, based on [37].

3.1 Definitions

Before we can state Kruskal's generalization, we need some basic concepts. These concepts all deal with special constraints on the set of relations.

Definition 1. *Consider a collection of binary relations $\{R_1, \cdots, R_q\}$ over X. The collection is called*

- r-complete*: if for every two (distinct) elements $x, y \in X$ there is an i $(1 \leq i \leq q)$ such that $x\ R_i\ y$.*
- r-total*: if for every two (distinct) elements $x, y \in X$ there is an i $(1 \leq i \leq q)$ such that $x\ R_i\ y$ or $y\ R_i\ x$.*
- r-symmetric*: if for every R in the collection, also R^{-1} belongs to it.*

Example. Consider the following relations on $\mathbb{R}^2$: R_a defined by $\langle u, v\rangle R_a \langle x, y\rangle$ iff $u \leq x$ and $v \leq y$, and R_b defined by $\langle u, v\rangle R_b \langle x, y\rangle$ iff $u \leq x$ and $v \geq y$. The collection $\{R_a, R_a^{-1}, R_b, R_b^{-1}\}$ is r-complete, the collection $\{R_a, R_b\}$ is r-total, and the collection $\{R_a, R_a^{-1}\}$ is r-symmetric.

The example shows that the concepts all differ. Observe that a collection of binary relations that is r-symmetric and r-total, is also r-complete. An r-complete collection is also r-total.

A collection of binary relations $\{R_1, \cdots, R_q\}$ is called *strictly r-total* if for every two distinct elements $x, y \in X$ there is an i $(1 \leq i \leq q)$ such that either $x\ R_i\ y$ or $y\ R_i\ x$. The given example of an r-total collection is in fact strictly r-total.

3.2 Kruskal's Extended Basic Theorem

In Kruskal's definition of relation spaces, a collection of relations is required to be *r-complete* ([22], p. 272). His generalization of the subsequence theorem is captured in the following *Extended Basic Theorem*. We state the theorem for arbitrary length parameters. (Kruskal only formulated the theorem for the case in which all parameters are equal.)

Theorem 2. *Let σ be a sequence of n distinct elements from X and $\{R_1, \cdots, R_q\}$ an* r-complete *collection of binary relations. If $n \geq s_1 \cdots s_q + 1$, then there is an i ($1 \leq i \leq q$) such that σ contains an ascending (descending) subsequence of length $s_i + 1$ according to R_i.*

We claim that this generalization of the Erdős-Szekeres theorem is equivalent to the general Ramsey-version of the result. (By equivalence we mean that one theorem may be proved from the other by merely manipulating relations.)

Theorem 3. *As generalizations of the subsequence theorem, Theorem 1 (Ramsey version) and Theorem 2 (Kruskal version) are equivalent.*

Proof. We show how one theorem can be transformed into the other. We distinguish the two obvious cases.

(i) *Theorem 1 ⇒ Theorem 2.* Let $\sigma = \sigma_1, \cdots, \sigma_n$ be a sequence of n distinct elements from X and let $\{R_1, \cdots, R_q\}$ be r-complete. Restrict X to the elements of σ: $X = \{\sigma_i \mid 1 \leq i \leq n\}$. Restrict the relations accordingly and modify them further, to obtain the collection $\{R'_1, \cdots, R'_q\}$ where

$$R'_k = R_k - \{(\sigma_j, \sigma_i) \mid i < j \text{ and } \sigma_j R_k \sigma_i\} \quad (1 \leq k \leq q).$$

The effect of the modification is that all descending subsequences of length greater than 1 in any of the relations $R_1, \cdots, R_q$ are broken. However, the r-completeness of $\{R_1, \cdots, R_q\}$ implies that every two elements σ_i and σ_j of X ($i \neq j$) remain related under at least one of the relations $\{R'_1, \cdots, R'_q\}$.

Now apply Theorem 1 to σ and $\{R'_1, \cdots, R'_q\}$ with $w = 1$, and $s_i = s_i$ and $r_i = 1$ for $1 \leq i \leq q$. The condition $n \geq s_1 \cdots s_q + 1$ exactly matches the requirement of Theorem 1, implying that there must be a k ($1 \leq k \leq q$) such that σ contains an ascending subsequence of length $s_k + 1$ in R'_k and thus, in R_k. (To obtain a descending subsequence, modify the relations $\{R_1, \cdots, R_q\}$ so as to break all ascending subsequences of length greater than 1, in a similar way.)

(ii) *Theorem 2 ⇒ Theorem 1.* Let $\sigma = \sigma_1, \cdots, \sigma_n$ be a sequence of n distinct elements from X and let $\{R_1, \cdots, R_q\}$ be an arbitrary collection of q relations. Restrict X to the elements of σ: $X = \{\sigma_i \mid 1 \leq i \leq n\}$. Define a further relation R_{q+1} on X by

$$R_{q+1} = \{(\sigma_i, \sigma_j) \mid i \neq j \text{ and there is no } 1 \leq k \leq q \text{ such that } \sigma_i R_k \sigma_j \text{ or } \sigma_j R_k \sigma_i\}$$

Thus, R_{q+1} consists precisely of those pairs of elements that are *unrelated* in any of the relations R_1 to R_q. Observe that the collection $\{R_1, \cdots, R_q, R_1^{-1}, \cdots, R_q^{-1}, R_{q+1}\}$ is r-complete. Now apply Theorem 1 to σ and the latter collection with $s_i = s_i$ and $s_{q+i} = r_i$ ($1 \leq i \leq q$), and $s_{2q+1} = w$. The condition $n \geq w \cdot s_1 \cdots s_q \cdot r_1 \cdots r_q + 1 = s_1 \cdots s_{2q+1} + 1$ exactly matches the requirement of Theorem 2, implying the conclusions of Theorem 1 by the properties of the relations in the collection. □

3.3 Further Basic Theorems

The Extended Basic Theorem applies to spaces with r-complete collections of relations. It is of interest to explore similar 'restricted' generalizations. The following important cases present themselves.

Theorem 2′. *Let σ be a sequence of n distinct elements from X and $\{R_1, \cdots, R_q\}$ an* r-total *collection of binary relations. If $n \geq s_1 \cdots s_q \cdot r_1 \cdots r_q + 1$, then there is an i ($1 \leq i \leq q$) such that σ contains an ascending subsequence of length $s_i + 1$ or a descending subsequence of length $r_i + 1$ in R_i.*

Theorem 2″. *Let σ be a sequence of n distinct elements from X and $\{R_1, \cdots, R_q\}$ an* r-symmetric *collection of binary relations. If $n \geq w \cdot s_1 \cdots s_q + 1$, then σ contains a subsequence of length $w + 1$ whose elements are unrelated in* each *of the relations $R_1, \cdots, R_q$ or there is an i $(1 \leq i \leq q)$ such that σ contains an ascending (descending) subsequence of length $s_i + 1$ in R_i.*

We show that the equivalence theorem can be extended to include Theorems 2′ and 2″ as well. We only indicate the proof ideas.

Theorem 3′. *As generalizations of the subsequence theorem, Theorem 1 (Ramsey version) and Theorems 2′ and 2″ are equivalent.*

Proof. We show that the conditions of the respective theorems can be transformed to match those of the other theorems.

(i) *Theorem 1 ⇒ Theorem 2′.* Let $\{R_1, \cdots, R_q\}$ be r-total. It follows that no two distinct elements of X are unrelated in each of the R_i $(1 \leq i \leq q)$. Now apply Theorem 1 with $w = 1$. Alternatively, notice that r-totality implies that $\{R_1, \cdots, R_q, R_1^{-1}, \cdots, R_q^{-1}\}$ is r-complete, and apply Theorem 2.

(ii) *Theorem 2′ ⇒ Theorem 1.* To prove that Theorem 1 can be derived from Theorem 2′, let $\{R_1, \cdots, R_q\}$ be an arbitrary collection of q relations. Define relation R_{q+1} by

$$R_{q+1} = \{(\sigma_i, \sigma_j) \mid i < j \text{ and there is no } 1 \leq k \leq q \text{ such that } \sigma_i R_k \sigma_j \text{ or } \sigma_j R_k \sigma_i\}$$

Observe that $\{R_1, \cdots, R_q, R_{q+1}\}$ is r-total, and that there are no descending subsequences of length greater than 1 in R_{q+1}. Now apply Theorem 2′ with $s_{q+1} = w$ and $r_{q+1} = 1$.

(iii) *Theorem 1 ⇒ Theorem 2″.* Let $\{R_1, \cdots, R_q\}$ be r-symmetric. Define relation R_{q+1} as in the proof of Theorem 3, i.e.

$$R_{q+1} = \{(\sigma_i, \sigma_j) \mid i \neq j \text{ and there is no } 1 \leq k \leq q \text{ such that } \sigma_i R_k \sigma_j \text{ or } \sigma_j R_k \sigma_i\}$$

Notice that $\{R_1, \cdots, R_q, R_{q+1}\}$ is necessarily r-complete. Now apply theorem 2 with $s_{q+1} = w$.

(iv) *Theorem 2″ ⇒ Theorem 1.* To prove that Theorem 1 can be derived from Theorem 2″, let $\{R_1, \cdots, R_q\}$ be an arbitrary collection of q relations. Obviously, $\{R_1, \cdots, R_q, R_1^{-1}, \cdots, R_q^{-1}\}$ is r-symmetric. Applying Theorem 2″ to the latter, immediately gives the result of Theorem 1. □

Theorems 2 to 2″ give a powerful range of special cases of Theorem 1. By the equivalences, the proof of Theorem 1 suffices for all. The proofs of the other theorems then follow by mere algebraic manipulation.

However, direct proofs of Theorems 2, 2′ and 2″ may be obtained as well, by suitably modifying the proof of Theorem 1. We show this for Theorem 2″.

Alternative proof of Theorem 2″. If $\{R_1, \cdots, R_q\}$ is r-symmetric, a subsequence of σ that is descending in one relation is ascending in another, and vice versa. Thus, we can restrict ourselves in the proof of Theorem 1 to tuples that consist of the lengths of the longest ascending (equivalently, descending) subsequences ending at positions of σ only. When the proof comes to arguing uniqueness of the tuples, modify the argument as follows. When considering a t $(1 \leq t \leq q)$ and positions $i_j < i_k$ such that $\sigma_{i_j} R_t \sigma_{i_k}$ or $\sigma_{i_k} R_t \sigma_{i_j}$, then by r-symmetry there must be an s $(1 \leq s \leq q)$ such that $R_t^{-1} \equiv R_s$. Hence we have $\sigma_{i_j} R_t \sigma_{i_k}$ or $\sigma_{i_j} R_s \sigma_{i_k}$, and the proof can be completed as for Theorem 1 by just considering the ascending case. □

3.4 Subbarao's Generalization

In the nineteen sixties M.V. Subbarao [37] observed, independently and apparently unaware of Kruskal's paper, that the subsequence theorem can be generalized to a theorem for finite sequences of items from any domain with a binary relation which is asymmetric and total. Albeit in a different form, he essentially proved Theorem $2'$ for the case of a *strictly r-total* collection consisting of a *single* relation ($q = 1$).

In general, the strictly r-total case of Theorem $2'$ can be very useful. However, the following result shows that the strictly r-total case is as general as the other cases and thus cannot give better bounds, in general.

Corollary 4. *As generalizations of the subsequence theorem, Theorem 1 (Ramsey version) and the strictly r-total case of Theorem $2'$ are equivalent.*

Proof. We show again how the theorems can be transformed into each other. We have two cases.

(i) *Theorem 1 $\Rightarrow$ Theorem $2'$ for strictly r-total collections.* As any strictly r-total collection is r-total, this follows from the same argument as in Theorem $3'$.

(ii) *Theorem $2'$ for strictly r-total collections $\Rightarrow$ Theorem 1.* Let $\sigma = \sigma_1, \cdots, \sigma_n$ be a sequence of n distinct elements from X, and let $\{R_1, \cdots, R_q\}$ be an arbitrary collection of q binary relations. Define relations $\widetilde{R}_k$ $(1 \le k \le q)$, $\widetilde{R}_{q+k}$ $(1 \le k \le q)$, and $\widetilde{R}_{2q+1}$ as follows:

- $\widetilde{R}_k \quad = \{(\sigma_i, \sigma_j) \mid i < j \text{ and } \sigma_i R_k \sigma_j\}$,
- $\widetilde{R}_{q+k} = \{(\sigma_i, \sigma_j) \mid i < j \text{ and } \sigma_j R_k \sigma_i\}$, and
- $\widetilde{R}_{2q+1} = \{(\sigma_i, \sigma_j) \mid i < j \text{ and there is no } 1 \le k \le q \text{ such that } \sigma_i R_k \sigma_j \text{ or } \sigma_j R_k \sigma_i\}$

Observe that the collection $\{\widetilde{R}_1, \cdots, \widetilde{R}_q, \widetilde{R}_{q+1}, \cdots, \widetilde{R}_{2q}, \widetilde{R}_{2q+1}\}$ is strictly r-total. Also notice that in σ, ascending subsequences according to R_k correspond to ascending subsequences in $\widetilde{R}_k$ and descending subsequences according to R_k correspond to ascending subsequences in $\widetilde{R}_{q+k}$ $(1 \le k \le q)$. Furthermore, $\widetilde{R}_k$ and $\widetilde{R}_{q+k}$, and $\widetilde{R}_{2q+1}$, have no descending subsequences of length greater than 1.

Now define $\widetilde{s}_k = s_k, \widetilde{s}_{q+k} = r_k, \widetilde{r}_k = 1$, and $\widetilde{r}_{q+k} = 1$ for $1 \le k \le q$, $\widetilde{s}_{2q+1} = w$ and $\widetilde{r}_{2q+1} = 1$. Apply the strictly r-total case of Theorem $2'$ to the collection $\{\widetilde{R}_1, \cdots, \widetilde{R}_q, \widetilde{R}_{q+1}, \cdots, \widetilde{R}_{2q}, \widetilde{R}_{2q+1}\}$ with these parameters. Then one sees that for $n \ge \widetilde{s}_1 \cdots \widetilde{s}_{2q+1} \widetilde{r}_1 \cdots \widetilde{r}_{2q+1} + 1 = w \cdot s_1 \cdots s_q \cdot r_1 \cdots r_q + 1$ the desired conclusion for Theorem 1 follows. □

4 Generalizations in Tuple Spaces

Kruskal's second generalization seems to have been designed especially for dealing with sequences in *tuple spaces* like $\mathbb{R}^d$ $(d \ge 1)$. We will argue that most subsequence theorems for tuple spaces generalize to the broader setting of Sections 2 and 3.

4.1 Basic Generalization

The Erdős-Szekeres theorem has been generalized to higher-dimensional spaces in several ways. Generalizations of the theorem to $\mathbb{R}^d$ are generally presented in terms of bounds on the cardinality of *sets* of d-tuples that guarantee the existence of monotone (ascending or descending) sequences in the set. We stay with the framework of arbitrary subsequences and general tuple spaces.

Consider an arbitrary tuple space $X = X_1 \times \cdots X_d$, where the component spaces X_i $(1 \le i \le d)$ are arbitrary non-empty sets. Tuples $x \in X$ are denoted as $x = (x_1, \cdots, x_d)$.

Definition 2. *For any signature $c \in \{-1,1\}^d$ we define the binary relation R_c on X by xR_cy if and only if for each i ($1 \le i \le d$), $x_iR_iy_i$ (if $c_i = 1$) or $y_iR_ix_i$ (if $c_i = -1$).*

This gives 2^d binary relations on X. (The definition generalizes the corresponding one for $\mathbb{R}^d$ in [33].) Let $c_1, \cdots, c_{2^d}$ be a listing of all signatures $c \in \{-1,1\}^d$.

Notice that, if all relations R_i are total orders, then every R_c is a partial order on X. Moreover, the collection of partial orders $\{R_c \mid c \in \{-1,1\}^d\}$ in this case is (r-symmetric and) *r-complete*.

Theorem 4. *Let σ be a sequence of n distinct tuples from $X_1 \times \cdots \times X_d$ and let $\{R_c \mid c \in \{-1,1\}^d\}$ be r-complete. Let $l_i \ge 1$ be arbitrary integers, for $1 \le i \le 2^d$. If $n \ge s_1 \cdots s_{2^d} + 1$, then there is a i ($1 \le i \le 2^d$) such that σ contains an ascending (descending) subsequence of length $s_i + 1$ according to R_{c_i}.*

Proof. The result follows as an instance of Theorem 2. □

Theorem 4 corresponds to a generalization of the Erdős-Szekeres theorem for subsets of $\mathbb{R}^d$ by Saxton, in the 'non-strictly monotone' case ([33], Proposition 3). He proved the bound on n to be best possible, and also that subsets that satisfy the bound must contain sequences of the indicated length that are 'strictly monotone' in all coordinates. In contrast, Theorem 4 applies to arbitrary *given* sequences. It is easy to obtain variants of Theorem 4 for other cases, e.g. for r-total collections.

4.2 De Bruijn's Theorem

According to [22], a first result for sequences in tuple spaces was obtained by N.G. de Bruijn (undated). De Bruijn's theorem is obtained from Theorem 4 by taking $X = \mathbb{R}^d$ and setting all s_i's equal, say equal to m.

Generalizing this to arbitrary tuple spaces, Theorem 4 implies that under the stated conditions, *every sequence of n distinct elements from $X = X_1 \times \cdots \times X_d$ with $n \ge m^{2^d} + 1$ must have an ascending (descending) subsequence of length at least $m + 1$ according to some relation R_c.*

Remark. Kalmanson [21] rediscovered De Bruijn's theorem in the set version, proved it best possible, and generalized it to finite-dimensional metric spaces with a polyhedral unit sphere. More general results for the set version were obtained in [33]. A proof that the stated bound in De Bruijn's theorem is best possible was given earlier also in [2].

The following corollary corresponds to an observation by Heinrich-Litan [17] for $X = \mathbb{R}^d$ in the set case. She proved that her result is essentially best possible, by constructing an n-element subset $S \subseteq \mathbb{R}^d$ which does not contain any monotone sequences of length greater than $\lceil n^{\frac{1}{2^{d-1}}} \rceil$.

Corollary 5. *Let σ be a sequence of n distinct tuples from $X_1 \times \cdots \times X_d$ and let $\{R_c \mid c \in \{-1,1\}^d\}$ be r-complete. Then σ must contain an ascending (descending) subsequence of length at least $\lceil n^{\frac{1}{2^d}} \rceil$ according to some R_c.*

Proof. Consider Theorem 4 with $s_i = m$ ($1 \le i \le 2^d$). If m can be chosen such that $n \ge m^{2^d} + 1$ and $m \ge \lceil n^{\frac{1}{2^d}} \rceil - 1$, then the result follows by De Bruijn's theorem. To see that it can be done, consider the following two cases. If $n^{\frac{1}{2^d}} = a$ is integer, then $n = a^{2^d}$ and one can take $m = a - 1$. If $n^{\frac{1}{2^d}} = a + \epsilon$ for integer a and $0 < \epsilon < 1$, then $n \ge a^{2^d} + 1$ and $m = a$ will do. □

As an application of Theorem 4, consider finite sequences of arguments of some 1-1 function $f : X_1 \times \cdots \times X_d \to Y$. Let R be a binary relation on Y, and assume that $\{R\}$ is r-total. (For example, $Y = \mathbb{R}$ and R the $\leq$-relation.) We show that all sufficiently long sequences will contain monotone subsequences on which f is monotone as well.

Theorem 5. *Let σ be a sequence of n distinct tuples from $X_1 \times \cdots \times X_d$ and let $\{R_c \mid c \in \{-1,1\}^d\}$ be r-complete. If $n \geq m^{2^{d+1}} + 1$, then there is an i $(1 \leq i \leq 2^d)$ such that σ contains a subsequence $\sigma_{i_1}, \sigma_{i_2}, \cdots$ of length $m+1$ with the following property: the subsequence is ascending according to R_{c_i} and the image sequence $f(\sigma_{i_1}), f(\sigma_{i_2}), \cdots$ is monotone according to R.*

Proof. Consider the space $X_1 \times \cdots \times X_d \times Y$. For each $c \in \{-1,1\}^d$ and $c' \in \{-1,1\}$, let $R_{c,c'}$ be defined such that $x R_{c,c'} y$ if and only if $(x_1, \cdots, x_d)\, R_c\, (y_1, \cdots, y_d)$ and furthermore, $x_{d+1}\, R\, y_{d+1}$ (if $c' = 1$) or $y_{d+1}\, R\, x_{d+1}$ (if $c' = -1$). It easily follows from the assumptions that the collection $\{R_{c,c'} \mid c \in \{-1,1\}^d$ *and* $c' \in \{-1,1\}\}$ is r-complete. Let $\sigma = \sigma_1, \cdots, \sigma_n$. The result now follows by applying Theorem 4 to the sequence $\sigma' = \langle \sigma_1, f(\sigma_1)\rangle, \langle \sigma_2, f(\sigma_2)\rangle, \cdots, \langle \sigma_n, f(\sigma_n)\rangle$. □

4.3 Generalization of De Bruijn's Theorem

Kruskal [22] generalized De Bruijn's theorem to a result for finite sequences in 'joint relation spaces'. These spaces are obtained by allowing more than a single relation to occur in each of the coordinate spaces of a tuple space. We derive a general formulation which extends Kruskal's result to a fully parameterized one.

Suppose that for every i $(1 \leq i \leq d)$ one is given a *finite collection* of binary relations $R_j^{(i)}$ $(1 \leq j \leq k_i)$ on X_i. In the tuple space, one can now define joint relations as follows.

Definition 3. *For any $h = (h_1, \cdots, h_d)$ with $1 \leq h_i \leq k_i$ $(1 \leq i \leq d)$, let R_h be the relation over $X_1 \times \cdots \times X_d$ defined by $x\, R_h\, y$ if and only if for each i $(1 \leq i \leq d)$, $x_i R_{h_i}^{(i)} y_i$.*

The definition leads to $K = k_1 \cdots k_d$ binary relations on $X = X_1 \times \cdots \times X_d$. Denote the index set of the relations by $H = \{(h_1, \cdots, h_d) \mid 1 \leq h_i \leq k_i\ (1 \leq i \leq d)\}$. Let $e : \{1, \cdots, K\} \to H$ be an arbitrary 1-1 enumeration of the index set, i.e. of the relations R_h.

Theorem 6. *Let σ be a sequence of n distinct tuples from $X_1 \times \cdots \times X_d$ and let each of the d collections $\{R_j^{(i)} \mid 1 \leq j \leq k_i\}$ be r-complete $(1 \leq i \leq d)$. If $n \geq s_1 \cdots s_K + 1$, then there is a t $(1 \leq t \leq K)$ such that σ contains a subsequence of length $s_t + 1$ that is ascending (descending) according to the joint relation $R_{e(t)}$.*

Proof. Observe that $\{R_h \mid h \in H\}$ is r-complete, by the r-completeness of the component collections. The result now follows from Theorem 2. □

Kruskal's generalization of De Bruijn's theorem ([22], Theorem 3) is obtained from Theorem 6 by taking $X_i = \mathbb{R}$ for all $1 \leq i \leq d$ and all s_t equal.

5 Subsequence Theorems in Graph Theory

The Ramsey generalization of the subsequence theorem and its equivalents can be applied elegantly in graph theory. We show here that various seemingly special properties have the succinct flavor of the Erdős-Szekeres theorem. We consider this for a number of different graph classes.

5.1 Undirected Graphs

Let $G = \langle V, E \rangle$ be an arbitrary undirected graph. The set of edges E is a binary relation, with uEv if and only if $(u, v) \in E$ for all $u, v \in V$. The sets of vertices from V that are 'unrelated' in E are precisely the *independent sets* in G.

General The classical Erdős-Szekeres theorem may be phrased graph-theoretically as follows. Let $\sigma = \sigma_1, \cdots, \sigma_n$ be a sequence of n distinct numbers. Define $G\langle V, E \rangle$ such that V consists of nodes $v_1, \cdots, v_n$ corresponding to the successive positions of σ and E satisfies $v_i E v_j$ if $\sigma_{\min(i,j)} > \sigma_{\max(i,j)}$ $(1 \le i, j \le n)$. G is the *inversion graph* associated with σ ([12]).

In the linear layout of G, with nodes from left to right, increasing subsequences of σ correspond to independent sets in G and decreasing subsequences correspond to right-going simple paths. (The nodes on a right-going path even form a clique, as is easily verified.)

Thus, if $n \ge ws + 1$, then the Erdős-Szekeres theorem amounts to the claim that the graph G we constructed must contain an independent set of size $w + 1$ or a right-going simple path of length $s + 1$. This property turns out to hold for all undirected graphs.

Theorem 7. *Let G be an arbitrary undirected graph, and let σ be a sequence of n distinct vertices of G. If $n \ge ws + 1$, then σ contains a subsequence whose elements form an independent set of size $w + 1$ in G or a subsequence that is a (simple) path of length $s + 1$ in G (or both).*

Proof. Notice that the collection consisting of the single relation E is r-symmetric. The result now follows from Theorem $2''$. □

Note that Theorem 7 differs from the well-known result in Ramsey theory that any graph with a least $R(w + 1, s + 1)$ vertices contains either an independent set of size $w + 1$ or a clique of size $s + 1$, for a suitable Ramsey number $R(s + 1, w + 1)$. For general graphs, this Ramsey number is considerably larger than the bound in Theorem 7 [16, 20].

For an application, consider graphs as intersection graphs [12]. Given a set U, an *intersection graph* over U is any finite graph whose vertices x correspond to non-empty subsets $S_x \subseteq U$ such that there is an edge between x and y if and only if $S_x \cap S_y \ne \emptyset$. Graph G is said to have *intersection number* k if k is the minimum size of any set U such that G is an intersection graph over U [9, 30].

Corollary 6. *Let G have intersection number k, and let σ be any sequence of n distinct vertices of G. Then σ contains a subsequence that is a simple path of length at least $\lceil \frac{n}{k} \rceil$ in G.*

Proof. Consider G as an intersection graph over a set U with $|U| = k$. Clearly, G cannot have an independent set of size greater than k. Let $s = \lceil \frac{n}{k} \rceil - 1$. The result follows directly from Theorem 7 once we show that $n \ge ks + 1$. To see it, write $\lceil \frac{n}{k} \rceil = \frac{n}{k} + \epsilon$ for some $0 \le \epsilon < 1$. Now $ks + 1 = k(\frac{n}{k} + \epsilon - 1) + 1 = n - (1 - \epsilon)k + 1 < n + 1$, and thus by integrality $ks + 1 \le n$. □

Rooted Trees The Erdős-Szekeres theorem can also be phrased in terms of rooted trees, as follows. Let σ be a sequence of n distinct numbers. Without loss of generality, assume that σ is a permutation of the numbers 1 to n. Let $v_1, \cdots, v_n$ be n nodes with v_i corresponding to i $(1 \le i \le n)$. Assume that the nodes are laid out in the given order as a rooted tree T consisting of a single path from leaf node v_1 to root node v_n.

View σ as a sequence of nodes in T, identifying the numbers in σ with the corresponding nodes in T. (Thus σ jumps up and down in T.) For the 'one-path' tree T, the Erdős-Szekeres theorem asserts the following: if $n \geq sr+1$, then σ contains a subsequence of length $s+1$ in which each node is an ancestor of the node preceding it, or it has a subsequence of length $r+1$ in which each node is a descendant of the node preceding it. This turns out to hold for all rooted trees.

Theorem 8. *Let T be an arbitrary rooted tree, and σ a sequence of n distinct nodes of T. If $n \geq wsr+1$, then σ contains a subsequence of $w+1$ nodes that are neither ancestors nor descendants of each other in T, or a subsequence of length $s+1$ in which each node is an ancestor of the node preceding it, or a subsequence of length $r+1$ in which each node is a descendant of the node preceding it.*

Proof. Define the binary relation E by xEy if and only if x is a descendant of y in T. Now apply Theorem 1 or, alternatively, Theorem $1'$. (E is a partial order.) □

In theorem 8 we may equally well assume that T is an arbitrary forest of rooted trees.

5.2 Directed Graphs

Tournaments A tournament on n nodes is any directed graph G that is obtained from the complete graph K_n by giving each edge in it a unique orientation. Tournaments are well-studied [25]. Rédei's theorem learns that all tournaments have a directed hamiltonian path.

The Erdős-Szekeres theorem can be linked to tournaments as follows. Let $\sigma = \sigma_1, \cdots, \sigma_n$ be a sequence of n distinct numbers. Consider the graph G obtained from the K_n by identifying the nodes $v_1, \cdots, v_n$ with the respective positions of σ and orienting edges such that $v_i E v_j$ if and only if $\sigma_i < \sigma_j$ $(1 \leq i, j \leq n)$.

Observe that G is a transitive, i.e. acyclic tournament. Increasing subsequences of σ correspond to directed paths in G, and decreasing subsequences to 'reverse directed' paths. Thus, if $n \geq sr+1$, then the Erdős-Szekeres theorem implies that, if we traverse the nodes of G in some arbitrary order $\pi = \pi_1, \cdots, \pi_n$ (a permutation), then G will include a directed path of length $s+1$ or a reverse directed path of length $r+1$, with the nodes on the path occurring in the same order as in π.

As every transitive tournament is the tournament of a set of numbers as we defined it (cf. [25], Thm. 9), the stated property holds for all *acyclic* tournaments. Interestingly, it is a property of *all* tournaments.

Theorem 9. *Let G be an arbitrary tournament, and σ a sequence of n distinct vertices of G. If $n \geq sr+1$, then σ contains a subsequence that is a directed path of length $s+1$ or a subsequence that is a reverse directed path of length $r+1$ in G.*

Proof. Define E by uEv if and only $\langle u, v \rangle$ is an arc in G $(u, v \in G)$. Notice that the collection just consisting of relation E is r-total. The result now follows from Theorem $2'$. □

Edge-weighted digraphs An edge-weighted digraph is any directed graph $G = \langle V, E, w \rangle$ with V and E as usual, and $w : E \to \mathbb{R}$ a scalar function that assigns weights to the edges.

We first introduce some terminology. A *forward path with k hops* in G is any sequence of k edges $\langle u_1, v_1 \rangle, \cdots, \langle u_k, v_k \rangle$ such that for each i $(1 \leq i \leq k-1)$, there is directed path from v_i to u_{i+1}. A *backward path with k hops* is any sequence of k edges $\langle u_1, v_1 \rangle, \cdots, \langle u_k, v_k \rangle$ for which $\langle u_k, v_k \rangle, \cdots, \langle u_1, v_1 \rangle$ is a forward path with k hops. A forward or backward path is called *increasing* if the weights on the successive edges in the path are increasing, and it is called *decreasing* if the weights are decreasing. Finally, a pair of distinct arcs $\langle u, v\rangle$ and $\langle x, y\rangle$ will be called *diametric* if there is no directed path from v to x and no directed path from y to u. A *diametric set* (of arcs) is any subset of E whose elements are pairwise diametric.

The Erdős-Szekeres theorem can be phrased in terms of edge-weighted graphs as follows. Let $\sigma = \sigma_1, \cdots, \sigma_n$ be a sequence of n distinct numbers. Let G' be the graph with vertices $v_1, \cdots, v_{n+1}$ and arcs $\langle v_i, v_{i+1}\rangle$ $(1 \leq i \leq n)$. Assign weights to edges by $w(\langle v_i, v_{i+1}\rangle) = \sigma_i$ $(1 \leq i \leq n)$. Consider the sequence of n edges $\kappa = \langle v_1, v_2\rangle, \cdots, \langle v_n, v_{n+1}\rangle$ in G'. The theorem of Erdős and Szekeres now amounts to the claim that, if $n \geq s_1 s_2 + 1$, then κ contains a subsequence that is an increasing forward path with $s_1 + 1$ hops or a subsequence that is a decreasing forward path with $s_2 + 1$ hops (or both).

The stated property seems hard to generalize: G' is a line and thus a very special graph, and κ lists edges in precisely the right order. However, the property is an instance of a general one that holds for *all* edge-weighted graphs.

Theorem 10. *Let G be an arbitrary edge-weighted graph, and μ a sequence of n (distinct) edges from G with distinct weights. If $n \geq w s_1 s_2 r_1 r_2 + 1$, then μ contains a subsequence μ' for which at least one of the following holds:* (a) *the edges of μ' form a diametric set of size $w + 1$,* (b) *μ' is an increasing forward path with $s_1 + 1$ hops,* (c) *μ' is a decreasing forward path with $s_2 + 1$ hops,* (d) *μ' is an increasing backward path with $r_1 + 1$ hops, or* (e) *μ' is a decreasing backward path with $r_2 + 1$ hops.*

Proof. Define the binary relations E_1, E_2 over the edges of G as follows. We set $\langle u, v\rangle E_1 \langle x, y\rangle$ if and only if there is a directed path from v to x and $w(\langle u, v\rangle) \leq w(\langle x, y\rangle)$. Next, we set $\langle u, v\rangle E_2 \langle x, y\rangle$ if and only if there is a directed path from v to x and $w(\langle u, v\rangle) \geq w(\langle x, y\rangle)$.

Now apply Theorem 1 to μ and the collection of binary relations $\{E_1, E_2\}$. Interpreting the theorem, we obtain the result in terms of forward ('ascending') and backward ('descending') paths. The sets of edges that are 'unrelated' in each of the relations E_1, E_2 are precisely the diametric sets. □

Returning to the initial example of the Erdős-Szekeres theorem, it is now clear why G' and κ are special: G' has no diametric sets of size greater than 1, and κ contains no backward paths, increasing or decreasing. Hence, it is precisely an instance of Theorem 10 with $w = r_1 = r_2 = 1$.

Directed Acyclic Graphs Directed acyclic graphs naturally arise when considering partial orders. Let $G = \langle V, E\rangle$ be an arbitrary directed acyclic graph. Colour the edges of G using the color set $\{c_1, \cdots, c_q\}$ arbitrarily with one color per edge. The following observation can be made, expanding on Corollary 3.

Theorem 11. *Let σ be an arbitrary sequence of n distinct vertices from G. If $n \geq w \cdot h^q + 1$, then G has a monochromatic directed path of length $h+1$ or G contains a subset of at least $w + 1$ vertices of which no two are connected by a monochromatic directed path.*

Proof. Each color c_i $(1 \leq i \leq q)$ induces a partial order P_i on the nodes of G, with $u \preceq_i v$ if and only if there is a (directed) path from u to v of which the edges (if any) are colored c_i only.

Sort G topologically, and let σ be the resulting sequence of vertices. Clearly σ has no descending subsequences of length greater than 1, in any of the partial orders. The result now follows from Theorem 1′, setting $s_1 = \cdots = s_q = h$ and $r_1 = \cdots = r_q = 1$. □

6 Optimality of the Bound on n

In this section we consider the optimality of the bound on n in Theorems 1, 2, 2′, and 2″. By the equivalence, if the bound is best possible for one of the theorems, then it is for all of them.

Kruskal [22] outlined an argument that Theorem 2 is optimal, i.e. the case of r-complete collections of relations with $s_1 = \cdots = s_q$. In this section we focus on Theorem 1 and give an inductive argument that it is best possible for any choice of parameters. More precisely, we show that it is already best possible when all binary relations are restricted to be *partial orders* on a tuple space, i.e. in the case of Theorem 1′.

We begin with a self-contained proof that the bound in Corollary 1 is best possible. Then we give a proof for the general result.

6.1 Optimality of Corollary 1

Consider Corollary 1. We know that the corollary is best possible for any s and r and $w = 1$ (i.e. the original theorem). This can be extended for arbitrary $w \geq 1$ as follows. Let $\delta = \delta_1, \cdots, \delta_n$ be a sequence of distinct integers of length $n = sr$ in which every monotone increasing subsequence has length at most s and every monotone decreasing subsequence has length at most r. Such sequences exist (cf. [20], Ex. 4.13). Now proceed as follows.

Let S be the set of all integers occurring in δ and let $p_1, \cdots, p_w$ be w distinct prime numbers larger than any integer in S. Let P be the poset of all integers $u \cdot p_i$ with $u \in S$ and $1 \leq i \leq w$. For $a, b \in P$ we let $a \preceq b$ if and only if there exist an i and $u, v \in S$ such that $a = u \cdot p_i$, $b = v \cdot p_i$ and $u \leq v$. One easily sees that the width of P is w.

Let δ^p denote the sequence of products $\delta_1 \cdot p, \cdots, \delta_n \cdot p$, and consider the sequence σ obtained by concatenating all δ^{p_i}'s, i.e. $\sigma = \delta^{p_1}, \cdots, \delta^{p_w}$. Observe that σ has length wsr, that every ascending subsequence in it has length at most s and that every descending subsequence in it has length at most r. This shows that the bound in Corollary 1 is sharp.

6.2 Optimality of Theorem 2′ for Partial Orders

In order to prove the optimality of Theorem 1′ we will first show that the bound in Theorem 2′ is best possible for posets, i.e. the case with multiple partial orders and $w = 1$. Then we extend this result to multiple partial orders and arbitrary w.

Notation. *For arbitrary $s, r \geq 1$ we let $\delta[s,r] = \delta[s,r]_1, \cdots, \delta[s,r]_{sr}$ be any sequence of sr integers with the property that every increasing subsequence in it has length at most s and every decreasing subsequence in it has length at most r.*

Notation. *Let $\sigma = \delta_1, \cdots, \delta_n$ be a sequence and i an integer. Then $\langle \sigma, i \rangle$ denotes the sequence $\langle \delta_1, i \rangle, \cdots, \langle \delta_n, i \rangle$.*

In order to show that Theorem 2′ is best possible we prove the following lemma, which asserts that subsequences can be tightly packed just below the bound.

Lemma 1. *Let $q \geq 1$ and $s_1, \cdots, s_q, r_1, \cdots, r_q \geq 1$ be arbitrary integers, and let $n = s_1 \cdots s_q \cdot r_1 \cdots r_q$. Then there are a domain D_q with $|D_q| = n$ and an r-total set of partial orders $P_1, \cdots, P_q$ on D_q such that the following holds: there is a sequence $\sigma = \sigma_q$ consisting of the n distinct elements from D_q such that for every i $(1 \leq i \leq q)$, every ascending subsequence of σ according to P_i has length at most s_i and every descending subsequence in σ according to P_i has length r_i.*

Proof. The proof proceeds by induction on q, constructing σ_{q+1} by composition from its 'q-dimensional' version.

For $q = 1$, the lemma trivially holds, by optimality of the original Erdős-Szekeres theorem. More concretely, given any $s_1 \geq 1$ and $r_1 \geq 1$, any sequence $\sigma_1 = \delta[s_1, r_1]$ will do. Here, D_1 consists of the $s_1 \cdot r_1$ elements of the sequence, and P_1 is the natural total order between integers (which is r-total by itself).

Assume the lemma holds for q (some $q \geq 1$). Now consider the case $q + 1$. Let $s_1, \cdots, s_q, s_{q+1}, r_1, \cdots, r_q, r_{q+1} \geq 1$ be arbitrary integers. Set $n = s_1 \cdots s_q \cdot r_1 \cdots r_q$. Let D_q and partial orders $P_1, \cdots, P_q$ on D_q and sequence $\sigma_q = \delta_1, \cdots, \delta_n$ be as implied by the induction hypothesis for q and $s_1, \cdots, s_q, r_1, \cdots, r_q$.

Now construct $D_{q+1} = \{\langle x, k\rangle \mid x \in D_q \ \ and\ 1 \leq k \leq s_{q+1}r_{q+1}\}$. Thus by induction, the elements of D_{q+1} are exactly the tuples $\langle \delta_j, k\rangle$ with $1 \leq j \leq n$ and $1 \leq k \leq s_{q+1}r_{q+1}$. We extend $P_1, \cdots, P_q$ to partial orders $P'_1, \cdots, P'_q$ on D_{q+1} as follows, where we use $\preceq_i$ and $\preceq_{i'}$ to denote relationships according to P_i and P'_i respectively: $\langle x, k\rangle \preceq_{i'} \langle y, l\rangle$ if and only $k = l$ and $x \preceq_i y$.

Define partial order P_{q+1} on D_{q+1} as follows: $\langle x, k\rangle \preceq_{q+1} \langle y, l\rangle$ if and only if $\langle x, k\rangle = \langle y, l\rangle$ or $k \neq l$ and $\delta[s_{q+1}, r_{q+1}]_k < \delta[s_{q+1}, r_{q+1}]_l$. i.e. the partial ordering is determined by the ordering of the integers in the 'Erdős-Szekeres sequence' $\delta[s_{q+1}, r_{q+1}]$. (The condition $k \neq l$ is not necessary as it is already implied by the distinctness of the integers in $\delta[s_{q+1}, r_{q+1}]$, but we include it for clarity.)

Claim. The set $\{P'_1, \cdots, P'_q, P_{q+1}\}$ is an r-total set of partial orders on D_{q+1}.

Proof. One easily verifies that $P'_1, \cdots, P'_q, P_{q+1}$ are partial orders on D_{q+1}, using that $P_1, \cdots, P_q$ are partial orders on D_q. By induction the latter form an r-total set on D_q. We now consider the set $P'_1, \cdots, P'_q, P_{q+1}$.

Consider two arbitrary elements $\langle x, k\rangle, \langle y, l\rangle \in D_{q+1}$. If $k = l$ then it follows by the r-totality of $\{P_1, \cdots, P_q\}$ that there must be an i ($1 \leq i \leq q$) such that $\langle x, k\rangle \preceq_{i'} \langle y, l\rangle$ or $\langle y, l\rangle \preceq_{i'} \langle x, k\rangle$. If $k \neq l$, then the definition of P_{q+1} implies that $\langle x, k\rangle \preceq_{q+1} \langle y, l\rangle$ or $\langle y, l\rangle \preceq_{q+1} \langle x, k\rangle$. Hence $\{P'_1, \cdots, P'_q, P_{q+1}\}$ is r-total. □

Finally, define the sequence σ by $\sigma = \langle \sigma_q, 1\rangle, \cdots, \langle \sigma_q, s_{q+1}r_{q+1}\rangle$. Clearly σ consists of $ns_{q+1}r_{q+1} = s_1 \cdots s_{q+1} \cdot r_1 \cdots r_{q+1}$ distinct elements, namely the elements of D_{q+1} in 'reverse lexicographic order'.

Claim. Sequence σ satisfies all requirements for σ_{q+1}.

Proof. We will only show that for every i ($1 \leq i \leq q$) every ascending subsequence of σ with respect to the partial order P'_i has length at most s_i, and that every ascending subsequence of σ with respect to P_{q+1} has length at most s_{q+1}. The proof of the corresponding statement for descending subsequences and their length is similar.

First consider the case of ascending subsequences according to P'_i ($1 \leq i \leq q$). Let $\langle x_1, k_1\rangle \preceq_{i'} \cdots \preceq_{i'} \langle x_t, k_t\rangle$ be an ascending subsequence of σ. If $t = 1$ then $t \leq s_i$ and we are done. Thus assume that $t > 1$. By the definition of P'_i we must have $k_1 = \cdots = k_t$ and hence, that $x_1, \cdots, x_t$ is an ascending subsequence of σ_q with respect to P_i. By the induction hypothesis it follows that $t \leq s_i$.

Next consider P_{q+1}. Let $\langle x_1, k_1\rangle \preceq_{q+1} \cdots \preceq_{q+1} \langle x_t, k_t\rangle$ be an ascending subsequence of σ. If $t = 1$ then $t \leq s_{q+1}$ and we are done again. Thus assume that $t > 1$. By the property of being a subsequence we have that $k_1 < \cdots < k_t$. By the definition of P_{q+1} it then follows that $\delta[s_{q+1}, r_{q+1}]_{k_1} < \cdots < \delta[s_{q+1}, r_{q+1}]_{k_t}$ and hence that $\delta[s_{q+1}, r_{q+1}]_{k_1}, \cdots, \delta[s_{q+1}, r_{q+1}]_{k_t}$ is an ascending subsequence of the sequence $\delta[s_{q+1}, r_{q+1}]$. It follows that $t \leq s_{q+1}$, by the 'Erdős-Szekeres' property of $\delta[s_{q+1}, r_{q+1}]$. □

It follows that we can set $\sigma_{q+1} = \sigma$ for the induction step. This completes the proof of the induction and thus of the lemma. □

6.3 Optimality of Theorem 1′

We now extend Lemma 1 to establish the tightness of the bound on n in Theorem 1′. To this end we show the following.

Theorem 12. *Let $q \geq 1$ and $w, s_1, \cdots, s_q, r_1, \cdots, r_q \geq 1$ be arbitrary integers, and let $n = w \cdot s_1 \cdots s_q \cdot r_1 \cdots r_q$. Then there are a domain E_q with $|E_q| = n$ and partial orders $R_1, \cdots, R_q$ on E_q such that the following holds: there is a sequence $\delta = \delta_q$ consisting of the n distinct elements from E_q such that every subsequence of δ whose elements form an anti-chain in each of $R_1, \cdots, R_q$ has length at most w and for every i $(1 \leq i \leq q)$, every ascending subsequence of δ according to P_i has length at most s_i and every descending subsequence in δ according to P_i has length r_i.*

Proof. Let $q \geq 1$ and $s_1, \cdots, s_q, r_1, \cdots, r_q \geq 1$ be arbitrary integers and let $m = s_1 \cdots s_q \cdot r_1 \cdots r_q$. By Lemma 1 there are a finite D_q and an r-total set of partial orders $P_1, \cdots, P_q$ on D_q such that the following holds: there is a sequence $\sigma_q = \mu_1, \cdots, \mu_m$ consisting of m distinct elements from D_q such that for every i $(1 \leq i \leq q)$, every ascending subsequence of σ according to P_i has length at most s_i and every descending subsequence in σ according to P_i has length r_i.

Let $w \geq 1$ be an arbitrary integer. If $w = 1$, then Lemma 1 is actually sufficient for our purpose, by r-totality of the set of partial orders. Thus, assume from now on that $w > 1$. Let $E_q = \{\langle x, k\rangle \mid x \in D_q \;\; and \; 1 \leq k \leq w\}$. Thus, the elements of E_q are exactly the tuples $\langle \mu_j, k\rangle$ with $1 \leq j \leq m$ and $1 \leq k \leq w$. We extend the $P_1, \cdots, P_q$ to partial orders $R_1, \cdots, R_q$ on E_q as follows, where we will use $\preceq_i$ and $\preceq_{i'}$ to denote relationships according to P_i and R_i respectively: $\langle x, k\rangle \preceq_{i'} \langle y, l\rangle$ if and only $k = l$ and $x \preceq_i y$. As in the proof of Lemma 1, it is easily seen that $R_1, \cdots, R_q$ are partial orders.

Define the sequence δ by $\delta = \langle \sigma_q, 1\rangle, \cdots, \langle \sigma_q, w\rangle$. Clearly σ consists of $n = w \cdot m$ distinct elements, namely the elements of E_q in reverse lexicographic order again. Now consider the properties of δ.

First, let $\langle x_1, k_1\rangle, \cdots, \langle x_t, k_t\rangle$ be a subsequence of elements which forms an anti-chain in each of the $R_1, \cdots, R_q$. If $t = 1$, then $t \leq w$ and we are done. Thus assume that $t > 1$. If there are any two indices a, b with $1 \leq a \neq b \leq t$ such that $k_a = k_b$, then notice that by r-totality of $P_1, \cdots, P_q$ there must be an i such that $x_a \preceq_i x_b$ or $x_b \preceq_i x_a$. Hence, by definition of R_i, $\langle x_a, k_a\rangle \preceq_{i'} \langle x_b, k_b\rangle$ or $\langle x_b, k_b\rangle \preceq_{i'} \langle x_a, k_b\rangle$. This contradicts the assumption that all elements in the subsequence were incomparable under all partial orders. Thus we may assume that all $k_1, \cdots, k_t$ are different. But then $t \leq w$, by the definition of E_q. Thus, any subsequence of elements which forms an anti-chain in each of the $R_1, \cdots, R_q$ is of size at most w.

Next, consider any i $(1 \leq i \leq q)$. Let $\langle x_1, k_1\rangle \preceq_{i'} \cdots \preceq_{i'} \langle x_t, k_t\rangle$ be an ascending subsequence with respect to R_i. Again, if $t = 1$ we have $t \leq s_i$ and we are done. Thus, assume $t > 1$. As the elements in the subsequence are all comparable, we must have $k_1 = \cdots = k_t$. It follows that $x_1, \cdots, x_t$ must be an ascending subsequence in P_i. Hence $t \leq s_i$ in this case as well. The argument for descending subsequences is completely similar.

We conclude that $\delta_q = \delta$ satisfies the properties required for the theorem. □

We note that Lemma 1 is a special case of Theorem 12, as is also suggested by the proof. For, take $w = 1$ in Theorem 12. Let x, y be any two distinct elements appearing in δ, say occurring in this order. If x and y are incomparable with respect to every R_i $(1 \leq i \leq q)$, then the subsequence x, y of δ violates the theorem. Thus the set of partial orders $\{R_1, \cdots, R_q\}$ is necessarily r-total when $w = 1$.

Theorem 12 implies that the bound on n in Theorem 1′, and thus in Theorem 1, is best possible.

7 Discussion

The classical subsequence theorem of Erdős and Szekeres [10] has been in extended and generalized in multiple directions, since its inception in the nineteen thirties. We reflected on an early study by J.B. Kruskal [22] who generalized the result to sequences in all domains that satisfy the constraints of, what he called, 'relation spaces'.

To obtain a proper benchmark, we showed that the subsequence theorem can be generalized to arbitrary domains with a finite number of binary relations, without any constraints at all. The key to this result is an application of Ramsey's theorem which reveals the elegant, general principle behind the subsequence theorem.

We subsequently showed that Kruskal's generalization is equivalent to the general Ramsey version. We showed several further equivalences, all involving useful cases of the Ramsey version. Freeing Kruskal's generalization from the constraint of relation spaces, the new generalization unifies and extends a variety of results that all fit the pattern.

The general Ramsey version leads to a wide range of new applications that all generalize the classical Erdős-Szekeres theorem, from posets to graph theory. Applications can be imagined in all contexts where large streams of structured data are at play.

The Erdős-Szekeres theorem has given rise to many further studies concerning the length of the longest increasing or decreasing subsequences that can occur in a sequence of n numbers [31]. It would be interesting to study length distributions for subsequences in the general framework of this paper as well.

References

1. M.H. Albert, R.E.L. Aldred, M.D. Atkinson, C.C. Handley, D.A. Holton, D.J. McCaughan, B.E. Sagan, Monotonic sequence games, in: M.H. Albert, R.J. Nowakowski (Eds), *Games of No Chance 3*, Vol 56, MSRI Publications, Cambridge University Press, 2009, pp. 309-327.
2. N. Alon, Z. Furedi, M. Katchalski, Separating pairs of points by standard boxes, *Europ. J. Combinatorics* 6 (1985) 205-210.
3. H. Burkill, L. Mirsky, Monotonicity, *J. Math. Analysis and Applic.* 41 (1973) 391-410.
4. F.R.K. Chung, On unimodal subsequences, *J. Combin. Theory*, Series A, 29 (1980) 267-279.
5. V. Chvátal, M. Komlós, Some combinatorial theorems on monotonicity, *Canad. Math. Bull.* 14:2 (1971) 151-157.
6. A. Dasgupta, *Set Theory - With an Introduction to Real Point Sets*, Birkhäuser, Springer Verlag, New York, 2014.
7. R.P. Dilworth, A decomposition theorem for partially ordered sets, *Annals of Mathematics* 51:1 (1950) 161-165.
8. M. Eliáš, J. Matoušek, Higher-order Erdős-Szekeres theorems, *Adv. in Mathematics* 244 (2013) 1-15.
9. P. Erdős, A.W. Goodman, L. Pósa, The representation of graphs by set intersections, *Canad. Math. J.* 18 (1966) 106-112.
10. P. Erdős, G. Szekeres, A combinatorial problem in geometry, *Compositio Mathematica* 2 (1935) 463470.
11. J. Fox, J. Pach, B. Sudakov, A. Suk, Erdős-Szekeres-type theorems for monotone paths and convex bodies, *Proc. London Math. Soc.* 105 (2012) 953-982.
12. M.C. Golumbic, *Algorithmic Graph Theory and Perfect Graphs*, Academic Press, New York NY, 1980.
13. E. Gottlieb, M. Sheard, Erdős-Szekeres results for set partitions, *Integers* 15 (2015) ♯A29.
14. R.L. Graham, Roots of Ramsey theory, in E. Bolker et al, *Andrew M. Gleason, Glimpses of a Life in Mathematics*, Festschrift, 1992, pp. 39-47, UMASS, Boston, `http://www.math.ucsd.edu/~ronspubs/92_08_ramsey_roots.pdf`.

15. R.L. Graham, J. Nešetřil, Ramsey theory in the work of Paul Erdős, in: R.L. Graham, J. Nešetřil, S. Butler (Eds), *The Mathematics of Paul Erdős II*, 2nd Ed., Springer, New York, 2013, pp. 171-193, `http://www.math.ucsd.edu/~ronspubs/13_04_Ramsey.pdf`.
16. R.L. Graham, B.L. Rothschild, J.H. Spencer, *Ramsey Theory*, 2nd Ed., Wiley Series in Discrete Mathematics and Optimization, Wiley, Hoboken NJ, 2013.
17. L. Heinrich-Litan, Monotone subsequences in $\mathbb{R}^d$, Techn. Report B 00-19, Fachbereich Mathematik und Informatik, Freie Universität Berlin, Berlin, 2000.
18. D.R. Hirschfeldt, R.A. Shore, Combinatorial principles weaker than Ramsey's theorem for pairs, *J. Symbolic Logic* 72 (2007) 171-206.
19. T.J. Jech, *Set Theory*, 3rd Millennium ed, Springer Monographs in Mathematics, Springer, Berlin, 2002.
20. S. Jukna, *Extremal Combinatorics*, Springer-Verlag, Berlin, 2001.
21. K. Kalmanson, On a theorem of Erdős and Szekeres, *J. Combin. Theory*, Series A, 15 (1973) 343-346.
22. J.B. Kruskal, Jr., Monotonic subsequences, *Proc. AMS* 4 (1953), 264-274.
23. M. Lerman, R. Solomon, H. Towsner, Separating principles below Ramsey's Theorem for Pairs, *J. Math. Logic* 13:2 (2013) 1350007.
24. N. Linial, M. Simkin, Monotone subsequences in high-dimensional permutations, *arXiv*:1602.02719, 2016.
25. J.W. Moon, *Topics on Tournaments*, Holt, Rinehart and Winston, New York, 1968.
26. G. Moshkovitz, A. Shapira, Ramsey theory, integer partitions and a new proof of the Erdős-Szekeres theorem, *Advances in Mathematics* 262 (2014) 1107-1129.
27. D.J. Newman, T.D. Parsons, On monotone subsequences, *Amer. Math. Monthly* 95:1 (1988) 44-45.
28. A.M. Odlyzko, J.B. Shearer, R. Siders, Monotonic subsequences in dimensions higher than one, *The Electronic Journal of Combinatorics*, ♯ R14, 4:2 (1997) 1-8.
29. F.P. Ramsey, On a problem in formal logic, *Proc. London Math. Soc.* (Series 2) 30:1 (1930) 264-286
30. F.S. Roberts, Applications of edge coverings by cliques, *Discr. Appl. Mathematics* 10 (1985) 93-109.
31. D. Romik, *The Surprising Mathematics of Longest Increasing Subsequences*, Institute of Mathematics textbooks, Cambridge University Press, New York, 2015.
32. T. Sakai, J. Urrutia, On the heaviest increasing or decreasing subsequence of a permutation, and paths and matchings on weighted point sets, in: A. Márquez *et al* (Eds), *Computational Geometry*, XIV Spanish Meeting (EGC 2011), Lecture Notes in Computer Science Vol. 7579, Springer-Verlag, 2012, pp. 175-184.
33. D. Saxton, Strictly monotonic multidimensional sequences and stable sets in pillage games, *J. Combin. Theory*, Series A, 118 (2011) 510-524.
34. A. Seidenberg, A simple proof of a theorem of Erdős and Szekeres, *J. London Mathematical Society* 34 (1959) 352.
35. R. Siders, Monotone subsequences in any direction, *J. Combin. Theory*, Series A, 85 (1999) 243-253.
36. M.J. Steele, Variations on the monotone subsequence theme of Erdős and Szekeres, in: D. Aldous *et al*, *Discrete Probability and Algorithms*, Springer-Verlag, New York, 1995, pp. 111-131.
37. M.V. Subbarao, On a theorem of Erdős and Szekeres, *Can. Math. Bull.* 11 (1968) 597-598.
38. T. Szabó, G. Tardos, A multidimensional generalization of the Erdős-Szekeres Lemma on monotone subsequences, *Combinatorics, Probability and Computing* 10:6 (2001) 557-565.