Abstract

We prove that any large enough collection of distinct sets behaves locally as singletons, complementaries of singletons or as a collection completely ordered by inclusion. We give bounds to specify what we mean by “large enough”; they are tight for particular cases.

Résumé:

Nous montrons que toute collection suffisamment grande d’ensembles distincts se comporte localement comme une collection de singletons, de complémentaires de singletons ou comme une collection totalement ordonnée par l’inclusion. Nous donnons des bornes pour préciser ce que l’on entend par “suffisamment grand” ; elle sont serrées dans des cas particuliers.

Introduction

We give here results about the local structure of any large enough family of distinct sets. By local structure, we mean a description of some of the sets for some of the elements of their union. Our Lemma 1 states that in any large enough collection of distinct sets, one can find an “increasing” or a “decreasing” sequence, in a weak sense described below.

Bauslaugh [1] originally gave an infinite version of that lemma and used it to find in any infinite twinless digraph some special induced subdigraph, thus giving a
counter-example to a property of compactness for list-colouring. But the proof in [1] has an error\(^1\), and our proof (section 1) may be considered as an erratum to [1].

Using Lemma 1 and Ramsey theory, we prove that in any large enough family of distinct sets, we can find a very precise substructure (Theorem 2, section 2). While we cannot give an exact bound in general, we provide lower and upper bounds and we give an exact bound for special cases.

Füredi and Tuza [3] gave a theorem that is more precise than Lemma 1 in the case where the sets under consideration are “small”. We then use the ideas introduced in section 2 and the ideas of Füredi and Tuza to get a new result (section 3).

For any non negative integer \(n\), we denote by \([n]\) the set \(\{1, 2, \ldots, n\}\) (\([0]\) = \(\emptyset\)). If \(E\) is a finite set, \(|E|\) denotes its cardinality and for each \(k\), \(\binom{E}{k}\) denotes the collection of all subsets of \(E\) of size \(k\).

1 Increasing and decreasing sequence in large enough collections of sets

We first define increasing and decreasing sequences of sets.

Definition 1 Let \((H_1, H_2, \ldots, H_k)\) be a finite sequence of sets.

- The sequence is \((k-1)\)-increasing if \(H_1 \subsetneq H_1 \cup H_2 \subsetneq \cdots \subsetneq H_1 \cup H_2 \cup \cdots \cup H_k\).

  Similarly, a countably infinite sequence of sets \((H_1, H_2, \ldots)\) is increasing if and only if \(H_1 \subsetneq H_1 \cup H_2 \subsetneq \cdots \subsetneq \cdots\).

- The sequence is \((k-1)\)-decreasing if \(H_1 \supsetneq H_1 \cap H_2 \supsetneq \cdots \supsetneq H_1 \cap H_2 \cap \cdots \cap H_k\).

  Similarly, a countably infinite sequence of sets \((H_1, H_2, \ldots)\) is decreasing if and only if \(H_1 \supsetneq H_1 \cap H_2 \supsetneq \cdots \supsetneq \cdots\).

How large should be a collection of sets to contain a \((k-1)\) increasing or decreasing sequence? This is answered by our first lemma.

Lemma 1 Let \(k, l\) and \(m\) be in \(\mathbb{N}\), and \(\{H_1, H_2, \ldots, H_m\}\) be a collection of \(m\) distinct sets. If \(m > \binom{k+l}{l}\) then at least one of the following two statements holds:

- Among the \(H_i\)'s, one can find a \((k + 1)\)-increasing sequence of sets \((H_{i_1}, H_{i_2}, \ldots, H_{i_{k+2}})\).

\(^1\)In [1], in the proof of Lemma 7, it is claimed (page 21, line 18) that “\(|S\setminus(S \cap A_i)|\) must take on arbitrary large finite values . . .”. However, the instance \(A_i = \mathbb{N}\setminus\{i\}\) with \(A = S = \mathbb{N}\) satisfies all the requirements while \(|S\setminus(S \cap A_i)|\) takes only value 1.
• Among the $H_i$’s, one can find an $(l + 1)$-decreasing sequence of sets $(H_{i_1}, H_{i_2}, \ldots, H_{i_{l+2}})$.

Proof:
We proceed by induction on $k, l$. If $k = 0$ or $l = 0$, the lemma is clear. Assume now $k > 0$ and $l > 0$, and let $\{H_1, \ldots, H_m\}$ be a collection of $m$ distinct sets with $m > \binom{k+l}{l}$. So $m \geq 2$, and there exists $x \in (H_1 \cup H_2 \cup \cdots \cup H_m) \setminus (H_1 \cap H_2 \cap \cdots \cap H_m)$. Let $m_1$ (resp. $m_2$) be the number of sets among $H_1, H_2, \ldots, H_m$ that contain $x$ (resp. that do not contain $x$). So $m_1$ and $m_2$ are positive and $m = m_1 + m_2$. Since $\binom{k+l}{l} = \binom{k+l-1}{l-1} + \binom{k-l}{l}$, at least one of the two following cases holds:

• $m_1 > \binom{k+l-1}{l-1}$. By the induction hypothesis we find among the sets that contain $x$ a $(k + 1)$-increasing sequence or an $l$-decreasing sequence. In the first case we are done. In the second one it suffices to append any set without $x$ to the $l$-decreasing sequence to get an $(l + 1)$-decreasing sequence.

• $m_2 > \binom{k-l}{l}$. Similarly, we find an $(l + 1)$-decreasing sequence, or a $(k + 1)$-increasing sequence by appending any set with $x$ to a $k$-increasing sequence of sets without $x$.

The tightness of the bound $\binom{k+l}{l}$ in Lemma 1 is established by considering the collection of sets $\binom{k+l}{l}$.

In any infinite collection of distinct sets, we can find by Lemma 1 an arbitrarily long increasing or decreasing sequence. But this does not immediately imply that there is an infinite increasing or decreasing sequence. This is why we recall and prove here the infinite lemma originally stated by Bauslaugh [1]. One could try to find a compactness argument (see [7]) to establish a link between the finite lemma (Lemma 1) and the infinite lemma below:

Lemma 2 Let $\mathcal{H} = \{H_1, H_2, \ldots\}$ be an infinite collection of distinct sets. One of the two following propositions holds:

• Among the $H_i$’s, one can find an infinite increasing sequence $(H_{i_1}, H_{i_2}, \ldots)$.

• Among the $H_i$’s, one can find an infinite decreasing sequence $(H_{i_1}, H_{i_2}, \ldots)$.

Proof: We claim that there exists an infinite sequence $(x_1, H_{k_1}), (x_2, H_{k_2}), \ldots$, such that for every $i \geq 1$ one of the following two properties holds:

1. $x_i \notin H_{k_i}$ and for every $j > i$, $x_i \in H_{k_j}$. 


2. \( x_i \in H_{k_i} \) and for every \( j > i \), \( x_i \notin H_{k_j} \).

We establish the claim by induction on \( i \). For \( i = 1 \), pick any \( x_1 \) which lies in at least one \( H_k \) but not in all of them.

If \( x_1 \) lies in infinitely many \( H_k \)'s, then, let \( H_{k_1} \) be one \( H_k \) that does not contain \( x_1 \). Continue with the (infinite) collection of all \( H_k \)'s that contain \( x_1 \). If \( x_1 \) lies in only finitely many \( H_k \)'s, then, let \( H_{k_1} \) be one of them. Continue with the (infinite) collection of all \( H_k \)'s that do not contain \( x_1 \). The proof is entirely similar for each \( i \geq 1 \). So the claim is proved.

Now, one the two properties 1, 2, holds for infinitely many pairs \((x_i, H_{k_i})\). If it is property 1, we find an increasing sequence, and if it is property 2, we find a decreasing sequence.

\[ \square \]

Note that in Lemma 1 very little is required of the sets: they do not have to be subsets of a given set, or to be of a given size, or even to be finite. But the lemma does not tell much about the structure one may hope to find in a sufficiently large family of distinct sets, and one may suspect that a better result is hidden behind our lemma. Before going further, we introduce some definitions.

It will be convenient to work with incidence matrices. For any collection of sets \( \mathcal{H} \), and any 0-1 matrix \( N \) with \( a \) rows and \( b \) columns, we say that \( N \) can be found in \( \mathcal{H} \) if we can find distinct sets \( H_1, H_2, \ldots, H_a \in \mathcal{H} \) and distinct elements \( e_1, e_2, \ldots, e_b \in \bigcup_{H \in \mathcal{H}} H \) such that \( N \) is the incidence matrix of the sets \( H_1, H_2, \ldots, H_a \) over the elements \( e_1, e_2, \ldots, e_b \) (ie \( N_{\alpha,\beta} = 1 \) if and only if \( e_\beta \in H_\alpha \)).

We say that a 0-1 matrix \( N \) is a \( k \)-increasing matrix if it has \( k \) columns, \( k+1 \) rows and satisfies: \( N_{i+1,i} = 1 \) for every \( i \in [k] \) and \( N_{i,j} = 0 \) for every \( 1 \leq i \leq j \leq k \). We say that \( N \) is a \( k \)-decreasing matrix if it has \( k \) columns, \( k+1 \) rows and satisfies: \( N_{i,i} = 0 \) for every \( i \in [k] \) and \( N_{i,j} = 1 \) for every \( 1 \leq j < i \leq k \).

\[
\begin{bmatrix}
0 & \cdots & \cdots & 0 \\
1 & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \vdots \\
? & \ddots & \ddots & 0 \\
? & \cdots & ? & 1
\end{bmatrix}
\begin{bmatrix}
0 & ? & \cdots & ? \\
1 & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \vdots \\
? & \ddots & \ddots & 0 \\
1 & \cdots & \cdots & 1
\end{bmatrix}
\]

Increasing  
Decreasing

Figure 1: Increasing and decreasing matrices

Lemma 1 can be rephrased as follows: Let \( \mathcal{H} = \{H_1, H_2, \ldots, H_m\} \) be a collection of \( m \) distinct sets. If \( m > \binom{k+l}{l} \) then one can find in \( \mathcal{H} \) a \((k+1)\)-increasing matrix or an \((l+1)\)-decreasing matrix.
2 Finding more specific matrices

As noted by Bauslaugh in his study of infinite digraphs, Ramsey’s famous theorem may be combined with Lemma 2. In our finite extremal set-theoretic context, this gives a more precise idea of the kind of local structure that can be found in any large enough collection of sets.

For any integer \( n \geq 1 \) we call \( n \)-singleton matrix the 0-1 matrix \( S^n \) with \( n \) columns and \( n + 1 \) rows defined by \( S^n_{i,j} = 1 \) if and only if \( i = j + 1 \). We call \( n \)-co-singleton matrix the 0-1 matrix \( \bar{S}^n \) with \( n \) columns and \( n + 1 \) rows defined by \( \bar{S}^n_{i,j} = 0 \) if and only if \( i = j \). We call \( n \)-monotone matrix the 0-1 matrix \( M^n \) with \( n \) columns and \( n + 1 \) rows defined by \( M^n_{i,j} = 1 \) iff \( i \geq j + 1 \).

\[
\begin{pmatrix}
0 & \cdots & \cdots & 0 \\
1 & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & 1
\end{pmatrix}
\quad \begin{pmatrix}
0 & 1 & \cdots & 1 \\
1 & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & 1 \\
1 & \cdots & 1 & 0 \\
1 & \cdots & 1 & 1
\end{pmatrix}
\quad \begin{pmatrix}
0 & \cdots & \cdots & 0 \\
1 & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 \\
1 & \cdots & \cdots & 1
\end{pmatrix}
\]

Singleton \quad Co-singleton \quad Monotone

Figure 2: Singleton, co-singleton and monotone matrices

Notice that \( S^1 = M^1 = \bar{S}^1 \). If \( n > 1 \), then \( S^n, \bar{S}^n \) and \( M^n \) are distinct. Every singleton matrix is increasing and every co-singleton matrix is decreasing. The matrices which are both increasing and decreasing are the monotone matrices. We call complementary of a matrix \( N \) the matrix obtained from \( N \) by swapping 0 and 1. Up to rearrangements of the rows and/or the columns, the complementary of a cosingleton matrix is a singleton matrix and the complementary of a monotone matrix is a monotone matrix.

We are going to find an appropriate singleton, cosingleton or monotone matrix in any large enough collection of sets. We first recall Ramsey’s theorem.

Theorem 1 (Ramsey, see [4]) For any positive integer \( r \) there exists a positive integer \( n \) such that for every partition \((A_0, A_1)\) of \([n]\), one can find a subset \( A' \) of \([n] \) such that: \(|A'| \geq r\) and either \( \binom{A'}{2} \subseteq A_0 \) or \( \binom{A'}{2} \subseteq A_1 \).

We denote by \( R(r) \) the Ramsey number, i.e., the smallest integer \( n \) that satisfies the claim of the Ramsey theorem (for instance, \( R(3) = 6 \)). The exact value of \( R(r) \) is not known in general, even for small values of \( r \), although some lower and upper bounds have been given (see [4]).
Theorem 2 For every non negative integers $k$ and $l$, there exists a number $S$ such that for any collection of sets $\mathcal{H}$, $|\mathcal{H}| > S$ implies that at least one the following three propositions holds:

- The $(k + 1)$-singleton matrix can be found in $\mathcal{H}$.
- The $(l + 1)$-cosingleton matrix can be found in $\mathcal{H}$.
- The $\min(k + 1, l + 1)$-monotone matrix can be found in $\mathcal{H}$.

We denote by $S(k, l)$ the smallest integer that satisfies the claim. We have:

$$S(k, l) = S(l, k) \leq \left( R(k + 1) + R(l + 1) - 2 \right).$$

Proof: Let $k$ and $l$ be in $\mathbb{N}$, and consider a collection $\mathcal{H}$ of distinct sets such that $|\mathcal{H}| > \left( \frac{R(k + 1) + R(l + 1) - 2}{R(k + 1) - 1} \right)$. By Lemma 1, we find in $\mathcal{H}$ an $R(k + 1)$-increasing matrix $N$ or an $R(l + 1)$-decreasing matrix $N'$.

In the first case, let $A_0$ (resp. $A_1$) be the subset of $\left( \binom{[R(k + 1)]}{2} \right)$ consisting of the $(i, j)$'s such that $i > j$ and $N_{i+1,j} = 0$ (resp. $N_{i+1,j} = 1$). By Ramsey's theorem, we can find a subset of $[R(k + 1)]$, say without loss of generality, the subset $[k + 1]$ such that all the pairs in $\binom{[k + 1]}{2}$ are in $A_0$ or in $A_1$. If they are in $A_0$, we have found in $\mathcal{H}$ a $(k + 1)$-singleton matrix. If they are in $A_1$, we have found a $(k + 1)$-monotone matrix. The second case is similar.

Thus $S$ exists and we have $S(k, l) \leq \left( \frac{R(k + 1) + R(l + 1) - 2}{R(k + 1) - 1} \right)$. The claim $S(k, l) = S(l, k)$ is clear by complementation.

Note that an analogue of Theorem 2 with only two of the three cases considered would be false. To see this, it suffices to consider the situation when $\mathcal{H}$ itself is a collections of singletons, or a collection of co-singletons, or a collection of sets completely ordered by inclusion.

2.1 Some exact values for $S(k, l)$

Since the exact value of the Ramsey number is not known in general, it could seem hopeless to try to determine $S(k, l)$ exactly. Nevertheless, for small values of $k$ and $l$, we can give the exact value of $S(k, l)$. It appears that the upper bound for $S(k, l)$ using the Ramsey number is quite generous (for instance, it says $S(2, 2) \leq \binom{10}{5} = 252$).

The collection $\binom{[k+l]}{l}$ shows $S(k, l) \geq \binom{k+l}{l}$. Actually, for $l = 0$, we do have $S(k, l) = \binom{k+l}{l} = 1$. This simply says that if at least two distinct sets are given, the matrix $\binom{l}{l}$ can be found in them.

If $l = 1$, the situation is also simple:
Lemma 3 If \( l = 1 \), \( S(k, l) = \binom{k+l}{l} = k + 1 \) for every \( k \) in \( \mathbb{N} \).

Proof: The proof is easy by a direct induction on \( k \). We give here another proof: Let \( \mathcal{H} \) be a collection of sets. If \(|\mathcal{H}| > k + 1\), we want to find in \( \mathcal{H} \) the \((k+1)\)-singleton matrix, the \(2\)-cosingleton matrix or the \(2\)-monotone matrix. By Lemma 1 we find in \( \mathcal{H} \) a \((k+1)\)-increasing matrix (case 1) or a \(2\)-decreasing matrix (case 2). In case 1, if by fluke the \((k+1)\)-increasing matrix is the \((k+1)\)-singleton matrix we are done. If not, we find in \( \mathcal{H} \) the \(2\)-monotone matrix. In case 2, we are done since a \(2\)-decreasing matrix is either the \(2\)-cosingleton matrix or the \(2\)-monotone matrix. \( \square \)

As it is true for \( l = 0 \) and \( l = 1 \), one could think that \( S(k, l) = \binom{k+l}{l} \) in general. But this is false for \( k = l = 2 \): the matrix \( F \) below shows that \( S(2, 2) \geq 8 \). Indeed, \( F \) should be seen as the incidence matrix of eight distinct sets over four elements. The point is that \( S^3, \overline{S}^3 \) or \( M^3 \) are not submatrices of \( F \) even after rearranging the rows and the columns.

\[
F = \begin{pmatrix}
1000 \\
0100 \\
1101 \\
1110 \\
0011 \\
0110 \\
1001 \\
1100
\end{pmatrix}
\]

We are now going to prove that \( S(2, 2) = 8 \). Our proof is long and requires several lemmas, some of which may give ideas for more general results. It will be convenient to work with “reduced” collection of sets, in a sense that we define now.

Definition 2 We say that a collection \( \mathcal{H} = \{H_1, H_2, \ldots, H_m\} \) of \( m \) distinct sets is reduced if every element is useful to make the sets distinct, that is for every \( x \in H_1 \cup \cdots \cup H_m \), we can find \( i \) and \( j \) such that \( i \neq j \) and \( H_i \setminus \{x\} = H_j \setminus \{x\} \).

Note that in a reduced collection of sets, there cannot be any universal element, i.e., there is no element in \( H_1 \cap H_2 \cap \cdots \cap H_m \). Also there are no duplicated elements, that is for every \( x \) and \( y \) in \( H_1 \cup H_2 \cup \cdots \cup H_m \) with \( x \neq y \), we can find \( i \) and \( j \) such that \( H_i \cap \{x, y\} \neq H_j \cap \{x, y\} \).

From a collection \( \mathcal{H} \) of distinct sets, we can get a reduced collection \( \mathcal{H}' \) of the same cardinality by deleting useless elements as long as there are any (the resulting \( \mathcal{H}' \) may depend on the choice of the arbitrary order of the deletion of the useless elements). We say that \( \mathcal{H}' \) is obtained from \( \mathcal{H} \). If a singleton, co-singleton or monotone matrix
is found in \( H' \), then it can be found in \( H \). This is why, when computing \( S(k, l) \), we can suppose that the collections of sets we consider are reduced.

The following two lemmas give answers to natural questions: If elements are picked in a reduced collection of distinct sets, how many distinct sets of the collection may we hope over them? How many elements is there in a reduced collection of sets? The lemma below is implicitly stated in an article of Kogan [6]. That article gives an interesting characterization of the structure of special reduced collections of sets.

**Lemma 4** ([6]) Let \( H = \{H_1, H_2, \ldots, H_m\} \) be a reduced collection of \( m \) sets. Then \( H_1 \cup \cdots \cup H_m \) has at most \( m - 1 \) elements.

**Proof:** Easy induction on \( m \). \( \Box \)

**Lemma 5** Let \( H = \{H_1, H_2, \ldots, H_m\} \) be a reduced collection of sets. If \( e_1, e_2, \ldots, e_k \) are distinct elements of \( H_1 \cup H_2 \cup \cdots \cup H_m \) then we can find \( k + 1 \) \( H_i \)'s that are distinct over \( e_1, e_2, \ldots, e_k \), i.e., sets \( H_{i_1}, H_{i_2}, \ldots, H_{i_{k+1}} \) such that the sets \( H_{i_1} \cap \{e_1, e_2, \ldots, e_k\}, H_{i_2} \cap \{e_1, e_2, \ldots, e_k\}, \ldots, H_{i_{k+1}} \cap \{e_1, e_2, \ldots, e_k\} \) are distinct.

**Proof:** Easy induction on \( k \). \( \Box \)

From now on, for simplicity we will make no difference between a collection of sets and its incidence matrix, in which we can rearrange rows and columns. When a matrix is given, we call \( r_1, r_2, \ldots \) its rows and \( c_1, c_2, \ldots \) its columns. The incidence matrix of a reduced collection of sets is a 0-1 matrix where all rows are distinct, all columns are distinct, and for each column, there exist two rows that become identical if one erases the column. This implies that each column contains at least one 1 and one 0. We use the notation: \( r_i \mapsto r_j[c_k] \) to express the facts that \( r_i \) and \( r_j \) are identical except in column \( c_k \), and row \( r_j \) (resp. \( r_i \)) has a 1 (resp. 0) at column \( c_k \).

We need seven more lemmas to show that \( S(2, 2) = 8 \).

**Lemma 6** Let \( H \) be a collection of nine distinct sets over four elements. We can find \( S^3 \), \( S^3 \) or \( M^3 \) in \( H \).

**Proof:** Note that \( H \) is necessarily reduced since we cannot have nine distinct sets over three elements. First remark that if we have seven distinct sets over three elements, then obviously, we find in them \( S^3 \) or \( S^3 \). Let \( H \) be a reduced collection of nine distinct sets over four elements \( c_1, c_2, c_3, c_4 \). Because of the preceding remark, at most six of the rows of \( H \) are distinct over \( c_1, c_2, c_3 \). At least five of those rows are distinct over \( c_1, c_2, c_3 \) (if not, we cannot have nine distinct rows just with the column \( c_4 \)). Furthermore it is impossible that three rows of \( H \) are equal over \( c_1, c_2, c_3 \).

Hence, we can find in \( H \) six rows \( r_1, \ldots, r_6 \) such that \( r_1, r_2, r_3 \) are distinct over \( c_1, c_2, c_3 \), and \( r_1 \mapsto r_4[c_4], r_2 \mapsto r_5[c_4], r_3 \mapsto r_6[c_4] \). By Lemma 4, we can suppose
without loss of generality that $r_1$, $r_2$ and $r_3$ are distinct over $c_1, c_2$. Forget $c_3$. Since there are only four possible sets over $c_1, c_2$, we have only two cases to consider (the other two are equivalent by complementation):

- We find in $\mathcal{H}$ the matrix:

\[
\begin{pmatrix}
000 \\
010 \\
110 \\
001 \\
011 \\
111
\end{pmatrix}
\]

Here, we find $M^3$ in $\mathcal{H}$.

- We find in $\mathcal{H}$ the matrix:

\[
\begin{pmatrix}
000 \\
100 \\
010 \\
001 \\
101 \\
011
\end{pmatrix}
\]

Here, we find $S^3$ in $\mathcal{H}$.

The following lemma will be used extensively in the sequel.

**Lemma 7** Let $\mathcal{H}$ be a reduced collection of sets in which we can find the matrix $\begin{pmatrix} 000 \\ 111 \end{pmatrix}$. Then, we can find $S^3$, $\bar{S}^3$ or $M^3$ in $\mathcal{H}$.

**Proof:** By Lemma 5 we can find in $\mathcal{H}$ a matrix $M$ with four distinct rows, and among them 000 and 111. Suppose that $M$ is not $S^3$, $\bar{S}^3$ or $M^3$. Then there are only two cases to consider (the other cases are equivalent by permuting rows or columns or swapping 0 and 1):

- We find in $\mathcal{H}$ the matrix $M = \begin{pmatrix} 000 \\ 100 \\ 010 \\ 001 \\ 101 \\ 011 \end{pmatrix}$.

There is no possibility to add a fifth row, different from the four above, to this 3-column matrix without finding $S^3$ or $M^3$. If we delete the element corresponding to the last column, then two sets of $\mathcal{H}$ must become equal. There are then four subcases: we can add to $M$ the row 001 (and then we find $S^3$), or 101 ($\rightarrow M^3$), or 011 ($\rightarrow M^3$) or 110 ($\rightarrow M^3$).
• We find in \( \mathcal{H} \) the matrix \( M = \begin{pmatrix} 000 \\ 100 \\ 011 \\ 111 \end{pmatrix} \).

Again, there is no possibility to add a fifth row to \( M \). If we delete the element corresponding to the second column, then two sets of \( \mathcal{H} \) must become equal. There are then four subcases: we can add to \( M \) the row 010 \( \rightarrow M^3 \), or 110 \( \rightarrow M^3 \), or 001 \( \rightarrow M^3 \) or 101 \( \rightarrow M^3 \).

\[ \square \]

**Lemma 8** Let \( \mathcal{H} \) be a reduced collection of sets over at least five elements. If we find in \( \mathcal{H} \) the matrix
\[
\begin{pmatrix}
0000 \\
1000 \\
0100 \\
1010 \\
0101
\end{pmatrix}
\]
then we can find \( S^3 \), \( \bar{S}^3 \) or \( M^3 \) in \( \mathcal{H} \).

**Proof:** Assume that we can find in \( \mathcal{H} \) this matrix and that we cannot find \( S^3 \), \( \bar{S}^3 \) or \( M^3 \) in \( \mathcal{H} \). There exist \( i \) and \( j \) such that \( r_i \leftrightarrow r_j[c_3] \). \( i \neq 1 \), otherwise we find \( S^3 \). By Lemma 7, we can assume w.l.o.g. that \( i = 2 \), and let \( j = 4 \). Similarly, there exist \( i' \) and \( j' \) s.t. \( r_{i'} \leftrightarrow r_{j'}[c_4] \), and since we assumed that \( S^3 \) cannot be found in \( \mathcal{H} \) we can put w.l.o.g. \( i' = 3 \) and let \( j' = 5 \). Thus, we have found in \( \mathcal{H} \) the matrix \( M \):
\[
M = \begin{pmatrix}
0000 \\
1000 \\
0100 \\
1010 \\
0101
\end{pmatrix}
\]

There is by assumption a fifth element (or column) \( c_5 \). We consider three cases (the number in a square will always represent the current hypothesis):

**First case:** We find in \( \mathcal{H} \):
\[
\begin{pmatrix}
0000 & 0 \\
1000 & 0^2 \\
0100 & 0^3 \\
1010 & 0^4 \\
0101 & 0^4 \\
\end{pmatrix}
\]

\(^1\)by Lemma 7
\(^2\)because \( r_2 \leftrightarrow r_4[c_3] \)
\(^3\)because \( r_3 \leftrightarrow r_5[c_4] \)
\(^4\)each column must have a 1 somewhere

By Lemma 7, we are allowed to put 1 only twice in the last row. In all cases, we find \( S^3 \).
Second case: We find in $\mathcal{H}$:

\[
\begin{pmatrix}
0000 & 1 \\
1000 & 0 \\
0100 & 0^3 \\
1010 & 0^2 \\
0101 & 0^1 \\
\end{pmatrix}
\]

1 by Lemma 7

2 because $r_2 \mapsto r_4[c_3]$

3 because $r_3 \mapsto r_5[c_4]$

We find $S^3$.

Third case: We find in $\mathcal{H}$:

\[
\begin{pmatrix}
0000 & 1 \\
1000 & 1 \\
0100 & 1^1 \\
1010 & 1^2 \\
0101 & 1^3 \\
0^4 \\
\end{pmatrix}
\]

1 because of the second case by symmetry

2 because $r_2 \mapsto r_4[c_3]$

3 because $r_3 \mapsto r_5[c_4]$

4 let $r_6$ be such that $r_6 \mapsto r_i[c_5]$ for some $i$

By Lemma 7, $r_6$ cannot have three or four 1’s on the first four columns. If $r_6$ has at most one 1 on these columns, then necessarily we are done by Lemma 7. Thus necessarily $r_6$ has exactly two 1’s on the first four columns.

One of these 1 has to be in the first or second column, otherwise we find $S^3$. By symmetry of the first two columns, we can assume that $r_6$ has a 1 in column $c_1$. If the other 1 is in column $c_4$, again we find $S^3$. If it is in column $c_3$, we are done by Lemma 7. Thus we are left with only one possibility, and we find:

\[
\begin{pmatrix}
0000 & 1 \\
1000 & 1 \\
0100 & 1 \\
1010 & 1 \\
0101 & 1 \\
1100 & 0 \\
1100 & 1^1 \\
\end{pmatrix}
\]

1 there exists $i$ s.t. $r_6 \mapsto r_i[c_5]$

We find $\bar{S}^3$.

\[\square\]

Lemma 9 Let $\mathcal{H}$ be a reduced collection of sets over at least five elements. If we find in $\mathcal{H}$ the matrix $\begin{pmatrix} 0000 \\ 1000 \\ 0110 \end{pmatrix}$ then we can find $S^3$, $\bar{S}^3$ or $M^3$ in $\mathcal{H}$.

11
Proof: By Lemma 5, we can suppose that $r_1, \ldots, r_5$ are distinct over $c_1, \ldots, c_4$. Assume that we cannot find $S^3, \bar{S}^3$ or $M^3$ in $\mathcal{H}$. By Lemma 7 and Lemma 8, $r_4$ and $r_5$ have both exactly two 1's over $c_1, \ldots, c_4$. For the rows $r_4$ and $r_5$, 0011 or 0101 bring $S^3$ and 0110 is impossible because of $r_3$. Thus, for $r_4$ and $r_5$ the only possibilities are 1001, 1010 and 1100.

Since erasing $c_2$ make two rows equal, we can suppose that $r_4$ is 1100 over the first four columns. And since erasing $c_3$ make two rows equal, we can suppose that $r_5$ is 1010 over the first four columns. As each column must have at least one 1, we find:

$$\begin{pmatrix} 0000 \\ 1000 \\ 0110 \\ 1100 \\ 1010 \\ 1 \end{pmatrix}$$

The last line has exactly one 1 in the first three columns by Lemma 7 and Lemma 8. In each case, we find $S^3$.

\[\square\]

Lemma 10 Let $\mathcal{H}$ be a reduced collection of sets over at least five elements. If we find in $\mathcal{H}$ the matrix $\begin{pmatrix} 0000 \\ 1000 \end{pmatrix}$ then we can find $S^3, \bar{S}^3$ or $M^3$ in $\mathcal{H}$.

Proof: By Lemma 5, we can suppose that $r_1, \ldots, r_5$ are distinct over $c_1, \ldots, c_4$. Assume that we cannot find $S^3, \bar{S}^3$ or $M^3$ in $\mathcal{H}$. By Lemma 7, 8 and 9, we know that $r_3, r_4$ and $r_5$ must have exactly two 1 over $c_1, \ldots, c_4$, and one of them on $c_1$. Finally, we find in $\mathcal{H}$ the following matrix and then $S^3$.

$$\begin{pmatrix} 0000 \\ 1000 \\ 1100 \\ 1010 \\ 1001 \end{pmatrix}$$

\[\square\]

Lemma 11 Let $\mathcal{H}$ be a reduced collection of distinct sets over at least five elements. If we find in $\mathcal{H}$ the matrix $(0000)$ or the matrix $(1111)$, then we can find $S^3, \bar{S}^3$ or $M^3$ in $\mathcal{H}$.

12
Proof: Let us suppose that we find (0000) in \( H \), and assume that we cannot find \( S^3 \), \( \bar{S}^3 \) or \( M^3 \) in \( H \). By Lemma 10 we know that we cannot have \( r_1 \rightarrow r_i[c_1] \) for some \( i \). Thus, we may suppose that \( r_2 \rightarrow r_3[c_1] \) and we find in \( H \) the following matrix:

\[
\begin{pmatrix}
0000 \\
0 \\
1
\end{pmatrix}
\]

On the columns 2, 3, 4, if we complete the rows 2 and 3 by at most one 1 we are done by Lemma 10. If we complete by two or three 1’s, we are done by Lemma 7. The case where we find (1111) is similar by complementation.

Lemma 12 Let \( H \) be a reduced collection of sets over at least five elements. We can find \( S^3 \), \( \bar{S}^3 \) or \( M^3 \) in \( H \).

Proof: Assume that we cannot find \( S^3 \), \( \bar{S}^3 \) or \( M^3 \) in \( H \). We suppose w.l.o.g. that \( r_1 \leftrightarrow r_2[c_1] \). If \( H \) has at least six columns, then we are be done by Lemma 11. Hence there are exactly five columns, and we can find in \( H \) the matrix:

\[
\begin{pmatrix}
00011 \\
10011
\end{pmatrix}
\]

(1) We first suppose that \( H \) has at most nine sets. If we delete any one of the columns \( c_1, \ldots, c_5 \), then two rows must become equal. Since there are at most nine rows, we know that a row is involved twice in this process, say w.l.o.g. \( r_1 \) or \( r_2 \). We can assume it is \( r_1 \) by symmetry. If we can find \( r_3 \) such that \( r_3 \rightarrow r_1[c_j] \), with \( j = 4 \) or 5, we are done by Lemma 11. Hence we can assume w.l.o.g. that \( r_1 \rightarrow r_3[c_2] \). We find in \( H \) the matrix:

\[
\begin{pmatrix}
00011 \\
10011 \\
01011
\end{pmatrix}
\]

Let \( r_4 \) be such that \( r_i \rightarrow r_4[c_3] \) for some \( i \). By the argument of the beginning of this proof, we know that \( r_4 \) has exactly three 1’s and two 0’s. Exactly one of these 1’s is over \( c_1, c_2 \) (by Lemma 7 and to avoid \( S^3 \)). By symmetry between \( c_1 \) and \( c_2 \), and between \( c_4 \) and \( c_5 \), we obtain w.l.o.g. the following matrix:

\[
\begin{pmatrix}
00011 \\
10011 \\
01011
\end{pmatrix}
\]
We now consider two cases:

**First case:** We find in $\mathcal{H}$:

\[
\begin{pmatrix}
0 & 0 & 0 & 1 & 1 \\
1 & 0 & 0 & 1 & 1 \\
0 & 1 & 0 & 1 & 1 \\
1 & 0 & 1 & 0 & 1 \\
1 & 0 & 0 & 0 & 1
\end{pmatrix}
\]

1. if 1, we find $\bar{S}^3$ with the columns 1, 4, 5.
2. if 0, we are done by Lemma 7.
3. if 1, we are done by Lemma 7.

We find $S^3$ (columns 2, 3 and 4).

**Second case:** We find in $\mathcal{H}$:

\[
\begin{pmatrix}
0 & 0 & 0 & 1 & 1 \\
1 & 0 & 0 & 1 & 1 \\
0 & 1 & 0 & 1 & 1 \\
1 & 0 & 1 & 0 & 1 \\
1 & 0 & 0 & 0 & 1
\end{pmatrix}
\]

1. by Lemma 7.
2. by Lemma 7.

We find $S^3$ (columns 1, 2 and 3).

(2) So the lemma is proved unless $\mathcal{H}$ has more than nine sets. In this case, we pick nine of them. If they form a reduced collection then we are done. If they do not, we delete a useless element. We stay with nine distinct sets defined over four elements, and we are done by Lemma 6.

$\square$

Now, by Lemma 6 and Lemma 12 we obtain:

**Proposition 1** Let $\mathcal{H}$ be a collection of at least 9 distinct sets. We can find $S^3$, $\bar{S}^3$ or $M^3$ in $\mathcal{H}$. Hence $S(2, 2) = 8$.  

14
2.2 A lower bound for $S(l, l)$

The exact value of $S(k, l)$ in general seems difficult to determine. We already noted that $S(k, l) \geq \binom{k+1}{l}$. Better bounds can be found.

**Proposition 2** For $l \geq 2$, $S(l, l) \geq \left( \frac{2l}{l-1} \right) + \left( \frac{2l-3}{l-1} \right)$.

**Proof:** We consider the collection $\mathcal{H} = A_0 \cup A_1 \cup \cdots \cup A_{l-2} \cup B \cup C \cup D$, with:

\[
A_i = \left\{ H \in \binom{[2l]}{i+1} \right\} \text{ s.t. } 2l \in H \text{ and } H \cap \{2l\} = \binom{[l+i-1]}{i} \\
B = \left\{ H \in \binom{[2l]}{l} \right\} \text{ s.t. } 1 \in H \text{ and } 2l-1 \notin H \text{ and } 2l \in H \\
C = \left\{ \binom{[2l-1]}{l} \right\} \\
D = \left\{ H \in \binom{[2l]}{l+1} \right\} \text{ s.t. } 1 \in H \text{ and } 2l \in H \\
\]

The transpose of the incidence matrix of $\mathcal{H}$ for $l = 3$ is given in figure 3.

We have $|\mathcal{H}| = \left( \sum_{i=0}^{l-2} \binom{l+i-1}{i} \right) + \binom{2l-3}{l-2} + \binom{2l-1}{l} + \binom{2l-2}{l-1}$. Since $\sum_{i=0}^{l-2} \binom{l+i-1}{i} = \binom{2l-2}{l}$, we obtain that $|\mathcal{H}| = \binom{2l-3}{l-1} + \binom{2l}{l}$. We now prove that $S^{l+1}, S_{l+1}^C$ or $M^{l+1}$ cannot be found in $\mathcal{H}$.

Assume that we can find $M^{l+1}$ in $\mathcal{H}$. Then we can find sets $H_1$ and $H_2$ in $\mathcal{H}$ and an increasing sequence $c_1, \ldots, c_{l+1}$ in $[2l]$ s.t. for each $k = 1, \ldots, l+1, c_k \notin H_1$ and $c_k \in H_2$. $|H_2| \geq l+1$ gives $H_2 \in D$, and $c_{l+1} = 2l$. But $|H_1| \leq l-1$ gives $H_1 \in \cup_{i=0}^{l-2} A_i$, and we have a contradiction since $2l \in H_1$.

Assume now that we can find $S^{l+1}$ in $\mathcal{H}$. Denote by $H$ in $\mathcal{H}$ the set corresponding to the rows with all 1’s in $S^{l+1}$, and by $c_1 < c_2 < \ldots < c_{l+1}$ the elements in $[2l]$ corresponding to these 1’s. We have $|H| \geq l+1$, hence $H \in D$, $c_1 = 1$ and $c_{l+1} = 2l$. We then have a set $H'$ in $\mathcal{H}$ s.t. $|H'| \geq l$, $1 \notin H'$ and $2l \in H'$. But such a set does not exist.

Finally assume that we can find $S_{l+1}^C$ in $\mathcal{H}$. Then there exist sets $H, H_1, \ldots, H_{l+1}$ in $\mathcal{H}$, elements $c_1 < c_2 < \ldots < c_{l+1}$ in $[2n]$ such that $H \cap \{c_1, \ldots, c_{l+1}\} = \emptyset$ and for $j$ in $\{1, \ldots, l+1\}$, $H_j \cap \{c_1, \ldots, c_{l+1}\} = \{c_j\}$. We have $|H| \leq l-1$, thus $H \in \cup_{i=0}^{l-2} A_i$, $2l \in H$ and $c_{l+1} < 2l$. For each $j = 1, \ldots, l+1$, $H_j$ has at least $l$ 0’s in the columns $1, \ldots, 2l-1$, since $H_j \notin D$ and $H_j \notin C$. Since no set in $\cup_{i=0}^{l-2} A_i \cup B$ contains $2l-1$, this imply that $c_{l+1} \neq 2l-1$ and for each $j = 1, \ldots, l+1$, $H_j$ has at least $l$ 0’s in the columns $1, \ldots, 2l-2$, hence $H_j \notin B$. We have obtained that for each $j$ in $\{1, \ldots, l+1\}$, there exists a (necessarily unique) $i_j$ in $\{0, \ldots, l-2\}$ such that $H_j \in A_{i_j}$. Fix $j$ with...
maximum index $i_j$. We have $c_{i+1} \leq l + i_j - 1 \leq 2l - 3$. Hence $H_j$ has at least $l$ 0's in the columns $1, ..., l + i_j - 1$. But $H_j \setminus \{2l\} \in \binom{\binom{l+i_j}{2}-l}{i_j}$, hence a contradiction. 

\rightline{\Box}

3 An exact bound for subsets of $[k+l]$  

F"uredi and Tuza gave a theorem that, in a sense, is an improvement of Lemma 1. It states that if all the sets are “small”, a very special increasing matrix (a singleton matrix) can be found. In what follows, $k$ and $l$ are nonnegative integers.

Theorem 3 (F"uredi, Tuza [3]) Let $\mathcal{H}$ be a collection of distinct sets $H_1, H_2, \ldots, H_m$. If $m > \binom{k+l}{i}$ and if we have $|H_i| \leq l$ for every $i$, then we can find in $\mathcal{H}$ a $(k+1)$-singleton matrix.

The proof of that theorem is based on the following theorem proved independently by Frankl and Kalai:

Theorem 4 (Frankl [2]; Kalai [5]) Let $A_1, A_2, \ldots, A_m$ be sets of size at most $l$ and let $B_1, B_2, \ldots, B_m$ be sets of size at most $k$ with $A_i \cap B_i = \emptyset$. Suppose that $A_i \cap B_j \neq \emptyset$ for all $i > j$. Then $m \leq \binom{k+l}{i}$.

The ideas of F"uredi and Tuza can be be used to provide a new result which looks like Theorem 2 except that here, we do have an exact bound as proved by the collection $\binom{[k+l]}{i}$:

Theorem 5 Let $\mathcal{H}$ be a collection of distinct sets $H_1, H_2, \ldots, H_m$, all of them included in $[k+l]$. If $m > \binom{k+l}{i}$ then at least one of the following three conditions is true:
1. The \((k + 1)\)-singleton matrix can be found in the collection of the sets of \(\mathcal{H}\) that have at most \(l\) elements.

2. The \((l + 1)\)-cosingleton matrix can be found in the collection of the sets of \(\mathcal{H}\) that have at least \(l + 1\) elements.

3. For some \(i \neq j\), \(H_j \subset H_i\), \(|H_j| \leq l\) and \(|H_i| \geq l + 1\).

Proof:

Let \(\mathcal{H}\) be a collection of distinct sets \(H_1, H_2, \ldots, H_m\), all of them included in \([k + l]\), and such that none of the three conditions 1, 2, 3 hold. We are going to show \(m \leq \binom{k + l}{l}\), thus proving the theorem. If \(H\) is a subset of \([k + l]\), \(\overline{H}\) denotes the complementary of \(H\) in \([k + l]\).

Suppose w.l.o.g. that for every \(i \leq j\) we have \(|H_i| \leq |H_j|\). Let \(n\) be the integer such that: \(|H_1| \leq l, |H_2| \leq l, \ldots, |H_n| \leq l, |H_{n+1}| \leq k - 1, \ldots, |H_m| \leq k - 1\). Note that \(n\) may be \(0\) or \(m\).

Let: \(A_1 = H_1, A_2 = H_2, \ldots, A_n = H_n, B_{n+1} = \overline{H}_{n+1}, \ldots, B_m = \overline{H}_m\).

For every set \(A_i = H_i\), \(i \leq n\), we claim that we can construct a set \(B_i\) such that \(A_i \cap B_i = \emptyset\), \(|B_i| = k\) and for every \(j\), \((1 \leq i < j \leq n \Rightarrow A_j \cap B_i \neq \emptyset)\). Indeed, consider a smallest set \(B_i\) such that \(B_i \cap A_i = \emptyset\) and such that for every \(A_j\) (with \(j \leq n\)) not included in \(A_i\), \(B_i \cap A_j \neq \emptyset\). Note that \(B_i\) exists and that \(1 \leq i < j \leq n \Rightarrow A_j \cap B_i \neq \emptyset\). So, if \(|B_i| = k\), we are done. If \(|B_i| < k\), we are done easily by completing \(B_i\) with elements not in \(A_i\). If \(|B_i| \geq k + 1\), let \(B_i = \{e_1, \ldots, e_{k+1}, \ldots\}\). By minimality, for every \(h \in [k + 1]\), there exists a set \(A_{ih}\) such that \(B_i \cap A_{ih} \neq \emptyset\) and \((B_i \setminus \{e_{ih}\}) \cap A_{ih} = \emptyset\). Hence, the incidence matrix of the sets \(A_1, A_{i_1}, \ldots, A_{i_{k+1}}\) over the elements \(e_1, \ldots, e_{k+1}\) is the \(k + 1\) singleton matrix, contradicting the fact that condition 1 does not hold for \(\mathcal{H}\).

Finally, we claim that if \(m \geq i \geq n + 1 \geq j \geq 1\), then \(A_i \cap B_j \neq \emptyset\). Suppose not, and let us consider \(m \geq i \geq n + 1 \geq j \geq 1\) such that \(A_i \cap B_j = \emptyset\). Since \(|A_i| = l\) and \(|B_j| = k\), we know that \((A_i, B_j)\) is a partition of \([k + l]\). Since \(A_j \cap B_j = \emptyset = A_i \cap B_i\), we have \(A_j \subset A_i \subset \overline{B_i}\). Since \(H_j = A_j\) and \(H_i = \overline{B_i}\), we obtain \(H_j \subset H_i\), contradicting the fact that condition 3 does not hold for \(\mathcal{H}\).

Theorem 4, and the sets \(A_i\) and \(B_i\) imply \(m \leq \binom{k + l}{l}\).

\[\square\]

Note that in the case \(k = l\), Theorem 5 is not an immediate consequence of Sperner’s lemma, which states that in any collection of \(\binom{2n}{n}\) + 1 subsets of \([2n]\) one
can find a subset included in another one (see [8]). Indeed, Sperner’s lemma says nothing about the size of the two subsets.

References


