Erdos-Ko-Rado Theorem and Kruskal-Katona Theorem

A thesis submitted to

Indian Institute of Science Education and Research Pune

in partial fulfilment of the requirements for the

BS-MS Dual Degree Programme

Thesis Supervisor: Dr. Soumen Maity

by

Kumar Vasumitra

April, 2013

Indian Institute of Science Education and Research Pune

Sai Trinity Building, Pashan, Pune India 411021

Declaration

This is to certify that this thesis entitled "Erdos-Ko-Rado Theorem and Kruskal-Katona Theorem" submitted towards the partial fulfillment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research Pune, represents work carried out by Mr. Kumar Vasumitra under the supervision of Dr. Soumen Maity.

Coordinator of Mathematics Faculty

Committee:

Dr. Soumen Maity

Reader 1

Reader 2

Acknowledgements

I would like to thank Dr. Soumen Maity for guiding me persistently with patience and giving me the opportunity to explore the subject with encouragement to face new problems.

I would also like to thank my family and my friends for their support. I would also like to thank Ankur for helping me with my project and the thesis. I would like to thank Mr. Jeeten Patel in particular for constantly urging me to try new approaches towards problems and towards life. Also, I would like to thank IISER Pune for providing me with the opportunity to carry my research on.

Abstract

Given a finite set **X**, an important problem in hypergraph theory is how large or small a family of subsets of **X** can be when it satisfies certain restrictions. Naturally, these type of questions appear throughout mathematics and so, hypergraph theory can be applied in areas ranging from topology to theoretical computer science. Two important concepts in hypergraph theory are "intersecting families" and "sections". The principle result in for maximum size of "intersecting families" is Erdos-Ko-Rado theorem and the principle result for minimum size of "sections" is Kruskal-Katona theorem. By defining suitable notions of "intersecting family" and "sections" one can find remarkable analogues of these theorems for other structures as multisets. This thesis aims to further understanding of "sections" and "intersecting families" in sets and multisets.

Contents

Chapter 1

Introduction

Let $\mathbf{X}=\{x_1, x_2, ...x_n\}$ be a finite set. A hypergraph H=$(E_1, E_2, ..., E_m)$ on $\mathbf{X}$ is defined to be a family of subsets of $\mathbf{X}$ satisfying the following properties:

$$E_i \neq \phi, \qquad 1 \leq i \leq m \tag{1.1}$$

$$\bigcup_{i=1}^{m} E_i = \mathbf{X} \tag{1.2}$$

The subsets $E_1, E_2, ...E_m$ are called the edges of the hypergraph and the elements $x_1, x_2, ...x_n$ are called the vertices. Note that the first condition excludes all the empty subsets and the second condition excludes all the isolated vertices from further discussions on hypergraphs.

A hypergraph is also called a set system or a family of sets drawn from the universal set $\mathbf{X}$. The difference between a set system and a hypergraph is not well defined and depends on the questions being asked. Hypergraph theory tends to ask questions similar to those of graph theory, such as connectivity and colorability while the theory of set systems tends to ask non graph theoretic questions, such as Sperner theory. We now recall some relevant definitions and concepts from the literature.

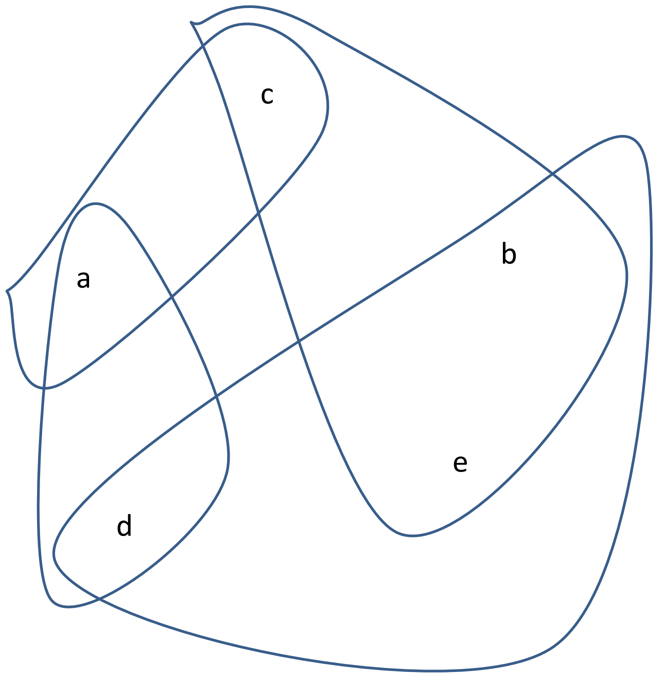

Figure 1.1: Hypergraph

A simple hypergraph *(Sperner family)* is one in which no edge is a subset of other. If there is a simple hypergraph H=$(E_1, E_2, ...E_m)$, then

$$E_i \subset E_j \Rightarrow i = j \tag{1.3}$$

Representation of Hypergraph by Figure

The vertices can be represented by points on a plane. Each edge is designated by a closed curve enclosing the vertices it contains.

For example, Figure (1.1) represents the hypergraph $H = (\{a, c\}, \{a, d\}, \{c, b, e\}, \{b, d, e\})$.

Representation by Incidence Matrix

A hypergraph can also be represented by its incidence matrix $A = ((a^i_j))$. In the incidence matrix of the hypergraph, the edges $E_1, E_2, ...E_m$ are represented by columns and the vertices $x_1, x_2, ...x_n$ are represented by rows. Further, $a^i_j = 0$ if the vertex x_i is not in the edge E_j and $a^i_j = 1$ if the vertex x_i lies in edge E_j.

1.1 Preliminaries

Definition 1 *The dual of a hypergraph* $H = (E_1, E_2, ...E_m)$ *on* $\boldsymbol{X}$ *is a hypergraph* $H^* = (X_1, X_2, ..., X_n)$ *whose vertices* $e_1, e_2, ..., e_m$ *correspond to the edges of* H *and with edges* $X_1, X_2, ...X_n$

$$X_i = \{e_j / x_i \in E_j \quad in \quad H\} \tag{1.4}$$

Note that the incidence matrix of H^* is the transpose of the incidence matrix of H. Hence, the dual of dual of a hypergraph is the hypergraph itself, i.e.

$$((H^*)^*) = H \tag{1.5}$$

Definition 2 *The order of a hypergraph* H *is defined as the number of elements of* $\boldsymbol{X}$ *and is denoted by* $n(H)$.

The number of edges of a hypergraph H is denoted by $m(H)$.

Definition 3 *The rank* $r(H)$ *of a hypergraph* H *is*

$$r(H) = \max_j \mid E_j \mid \tag{1.6}$$

Definition 4 *The anti-rank* $s(H)$ *of a hypergraph* H *is defined as*

$$s(H) = \min_j \mid E_j \mid \tag{1.7}$$

Further, if $r(H) = s(H)$, all the edges have the same cardinality and the hypergraph is said to be $uniform$.

Definition 5 *Let $J \subset \{1, 2, ..., m\}$. Then the family $H^{'} = (E_j/j \in J)$ is called the partial hypergraph of H generated by J.*

Note that a partial hypergraph contains some of the edges of the hypergraph.

Definition 6 *Let $A \subset \boldsymbol{X}$. Then, the family*

$$H_A = (E_j \cap A, 1 \leq j \leq m, \mid E_j \cap A \mid \neq 0)$$

is called the subhypergraph of H induced by the set A.

Definition 7 *For $x \in \boldsymbol{X}$, the star of x, $H(x)$ is defined as the partial hypergraph formed by edges containing x.*

The number of edges in $H(x)$, denoted by $m(H(x))$, is called degree of x.
$d_H(x) = m(H(x))$
The maximum degree of the hypergraph H is always denoted by $\triangle(H)$.
Thus, $\triangle(H) = \max_{x \in \mathbf{X}} d_H(x)$

Definition 8 *A hypergraph in which all vertices have the same degree is said to be regular.*

Also note that $\triangle(H) = r(H^*)$, and that the dual of a regular hypergraph is uniform.

Definition 9 *Let r,n be integers, $1 \leq r \leq n$. Then, the r-uniform complete hypergraph of order n (or the r-complete hypergraph) is defined to be a hypergraph, denoted by K_r^n and containing all the r subsets of the set $\boldsymbol{X}$ of cardinality n.*

Definition 10 *Let H be a simple hypergraph on $\boldsymbol{X}$ of rank r and let $k \leq r$ be an integer. The $k-$ section of the hypergraph, $[H]_k$ is defined to a hypergraph with edges $F \subset \boldsymbol{X}$ satisfying either $|F| = k$ and $F \subset E$, for $some$ $E \in H$ or $|F| < k$ and $F = E$, for $some$ $E \in H$. Note that $[H]_k$ is a simple hypergraph of rank k.*

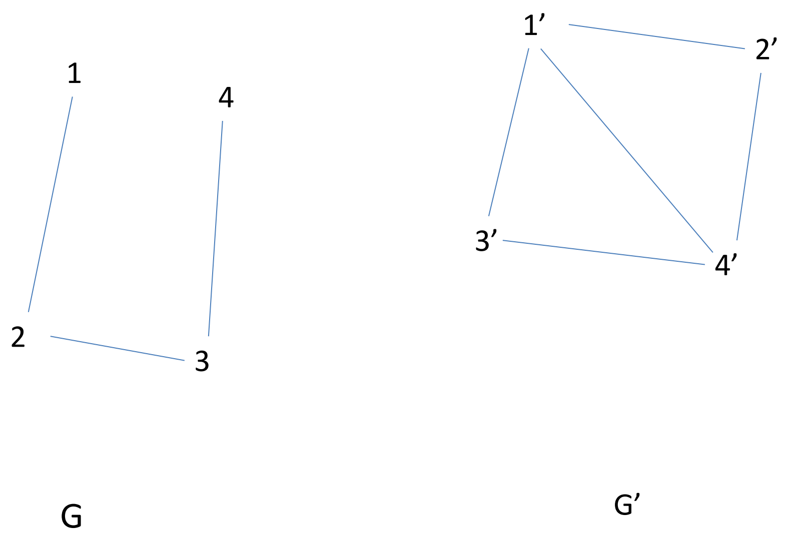

Figure 1.2: Graph Homomorphism

For details on these topics, see [2].

Definition 11 *Let $G = (V, E)$ be a graph with vertex set V and edge set E. A graph homomorphism f from $G = (V, E)$ to a graph $G' = (V', E')$ is a mapping $f : V \to V'$, from the vertex set of G to the vertex set of G' such that*

$$\{u, v\} \in E \Rightarrow \{f(u), f(v)\} \in E'.$$

Also, notice that if $f : G \to G'$ is a graph homomorphism from $G \to G'$ then we have

$$\alpha(G) \geq \alpha(G'),$$

where $\alpha(G)$ is the size of largest independent family of G.

In Fig (1.2), the graph G is homomorphic to the graph G' under the mapping f described as:

$$f(1) = 1'; f(2) = 2'; f(3) = 3'; f(4) = 4'.$$

1.2 Outline of the Thesis

We now give an outline of this thesis. In Chapter 1, we give an introduction to hypergraphs, intersecting families, section of a hypergraph and graph homomorphism. In Chapter 2, we review three important results from hypergraph theory: Sperner theorem, Erdos-Ko-Rado theorem and Kruskal-Katona theorem. In Chapter 3, we introduce multisets and an extension of Erdos-Ko-Rado theorem for multisets is given. Finally, an extension of Kruskal-Katona theorem for multisets is presented in Chapter 4.

Chapter 2

Some Results on Hypergraphs

In this chapter, we review three important results from hypergraph theory: Sperner theorem, Erdos-Ko-Rado theorem and Kruskal-Katona theorem.

2.1 Sperner Theorem

Theorem 2.1.1 [8] (Sperner[1928]: Proof by Yamamoto, Meshalkin, Lubell, Bollobas) *Every simple hypergraph H of order n satisfies*

$$\sum_{E \in H} \binom{n}{|E|}^{-1} \leq 1 \tag{2.1}$$

Further, the number of edges $m(H)$ satisfies

$$m(H) \leq \binom{n}{[\frac{n}{2}]} \tag{2.2}$$

Proof : For this proof, a simple construction is needed. Let $\mathbf{X}$ be a finite set of cardinality n, on which the hypergraph H is defined. Form all the subsets from the set $\mathbf{X}$. Now we designate these subsets by vertices and arrange them as:

The ϕ set comes in the first column. Next, all the vertices corresponding to subsets of cardinality 1 come in second column. The vertices corresponding to subsets of cardinality i come in the $(i+1)^{th}$ column.

On these vertices, we define a directed graph G as follows.

There is an arc between $A \subset \mathbf{X}$ and $B \subset \mathbf{X}$ if $A \subset B$ and $\mid A \mid = \mid B \mid -1$. Also, notice that:

(1) If $E \subset \mathbf{X}$, then the number of paths in G from ϕ to E is $\mid E \mid!$

(2) Since the hypergraph is simple, a path from ϕ to $\mathbf{X}$ via $E \in H$ cannot pass through $E' \in H$, where $E \neq E'$. For, if there is a path between E' and E via $E_1, E_2 ... E_k$; then we have

$$E' \subset E_1 \subset ... E_{k-1} \subset E_k \subset E$$

which is not possible as H is a simple hypergraph.

Now the number of paths in G from ϕ to $\mathbf{X}$ is $n!$. The number of paths from ϕ to $\mathbf{X}$ passing through E is $\mid E \mid !(n - |E|)!$. Thus we have,

$$n! \geq \sum_{E \in H} \mid E \mid !(n - |E|)! \tag{2.3}$$

or,

$$1 \geq \frac{\sum_{E \in H} \mid E \mid !(n - |E|)!}{n!} \tag{2.4}$$

(2.4), when rearranged, gives (2.1).

For the second part, note that

$$\binom{n}{[n/2]} \geq \binom{n}{\mid E \mid} \tag{2.5}$$

Hence, we get

$$1 \geq \sum_{E \in H} \binom{n}{\mid E \mid}^{-1} \geq m(H) \binom{n}{[n/2]}^{-1} \tag{2.6}$$

(2.6), when rearranged, gives (2.2), thereby completing the proof.

2.1.1 Linear Hypergraphs

A hypergraph $H = (E_1, E_2..., E_m)$ is said to be linear, if $\mid E_i \cap E_j \mid \leq 1$ for $i \neq j$.

Proposition 2.1.2 *The dual of a linear hypergraph is also linear.*

Proof Let $H = (E_1, E_2...E_m)$ be a linear hypergraph and let the dual, $H^* = (X_1, X_2...X_n)$ be a non linear hypergraph. Since, we assume H^* to be non linear, $\exists \;\; i, j \in \{1, 2, ...m\}$ such that $|X_i \cap X_j| \geq 2$.

Let $X_i \cap X_j = \{e_1, e_2, ...e_k\}$. Then in H, we get, $E_i \cap E_j = \{x_1, x_2\}$, where $i, j \in \{1, 2, ...k\}$ which is a contradiction of the starting assumption that H is linear.

Theorem 2.1.3 *For every linear hypergraph $H = (E_1, E_2...E_m)$ of order n we have*

$$\sum_{E \in H} \binom{|E|}{2} \leq \binom{n}{2} \tag{2.7}$$

In addition, if H is $r-uniform$, the number of edges $m(H)$ satisfies

$$m(H) = \frac{n(n-1)}{r(r-1)} \tag{2.8}$$

Proof The number of pairs (x, y) that are in a given edge E of H is $\binom{|E|}{2}$. Note that the pairs that are in $E \in H$ cannot be in $E' \in H$, where $E \neq E'$, as this would mean that $|E \cap E'| \geq 2$. Also the total number of pairs (x, y) that can be formed from the vertex set $\mathbf{X}$ is $\binom{n}{2}$. Hence, we get

$$\sum_{E \in H} \binom{|E|}{2} \leq \binom{n}{2} \tag{2.9}$$

Thus, we get the inequality in (2.7). Further, if H is $r-uniform$, the (2.7) reduces to

$$m(H)\binom{r}{2} \leq \binom{n}{2}, \tag{2.10}$$

or,

$$m(H)\frac{r(r-1)}{2} \leq \frac{n(n-1)}{2}, \tag{2.11}$$

which when rearranged gives (2.8).

2.2 Intersecting Families and Erdos-Ko-Rado Theorem

Given a hypergraph H, an intersecting family is defined as the set of edges having non empty pair wise intersection. If $E_1, E_2 ... E_k \in \mathcal{A}$, where $\mathcal{A}$ is an intersecting family, then $|E_i \cap E_j| \geq 1 \ \forall i, j \in \{1, 2...k\}$. For any vertex x in a hypergraph H, the star of x, $H(x)$ is an example of an intersecting family. The size of the largest intersecting family of a hypergraph H is always denoted by $\triangle_0(H)$ and satisfies

$$\triangle_0(H) \geq \max_{x \in X} |H(x)| = \triangle(H) \tag{2.12}$$

Theorem 2.2.1 *Every hypergraph H of order n, with no repeated edge satisfies*

$$\triangle_0(H) \leq 2^{(n-1)} \tag{2.13}$$

Further, every maximal intersecting family of a hypergraph of subsets of an n set has cardinality $2^{(n-1)}$.

Proof Let $\mathcal{A}$ be a maximal intersecting family of subsets of a set $\mathbf{X}$, where $|\mathbf{X}| = n$.

If $B_1 \notin \mathcal{A}$, then $\exists \ A_1 \in \mathcal{A}$, such that $A_1 \cap B_1 = \phi$ (This follows from the maximality of $\mathcal{A}$, else we could add B_1 to $\mathcal{A}$ and get a bigger intersecting family). Thus, we have $A_1 \subset \mathbf{X} - B_1$ and hence, $A_1 \cap (\mathbf{X} - B_1) \neq \phi$. Again, the maximality of $\mathcal{A}$ ensures that $(\mathbf{X} - B_1) \in \mathcal{A}$.

Further, if $\mathbf{X} - B_1 \in \mathcal{A}$, then $B_1 \notin \mathcal{A}$.

Hence, $B \to \mathbf{X} - B$ is a bijection between $P(\mathbf{X}) - \mathcal{A}$ and $\mathcal{A}$, where $P(\mathbf{X})$ is the power set of $\mathbf{X}$. Also, the bijection ensures that

$|P(\mathbf{X}) - \mathcal{A}| = |\mathcal{A}|$

Thus, we have

$$|\mathcal{A}| = \frac{|P(\mathbf{X})|}{2} = 2^{n-1} \tag{2.14}$$

Now, any hypergraph H is a partial hypergraph of $\mathcal{A}$. Hence,

$$\triangle_0(H) \leq m(H) \leq m(\mathcal{A}) \tag{2.15}$$

Equations (2.14) and (2.15) give us (2.13)

Lemma 2.2.2 [6](Greene, Katona, Kleitman) *Let $x_1, x_2...x_n$ be points on a circle in that order and let $A = (A_1, A_2...A_m)$ be a family of circular intervals of points satisfying the following properties*

(1) $|A_i| \leq \frac{n}{2} \quad \forall \;\; i \;\; \in \;\; \{1, 2...m\}$

(2) $|A_i \cap A_j| \neq 0 \quad \forall \;\; i, j \;\; \in \;\; \{1, 2...m\}$

(3) $A_i \not\subset A_j \quad \forall \;\; i, j \;\; \in \;\; \{1, 2...m\}$

then,

(4) $m \leq \min_i |A_i|$

and

(5) $\sum_i |A_i|^{-1} \quad \leq 1$

Equality in (5) *is attained iff A is a family of circular intervals of cardinality m and each having a point in common.*

Proof : Let A_1 be a circular interval of minimum size. Then,

(a) From *(2)*, $|A_1 \cap A_i| \;\; \neq \;\; 0 \;\; \forall \;\; i \neq 1$.

(b) And, from *(1)* and *(3)*, all other intervals have only one of their ends coinciding with an end of A_1.

(c) Also, from *(3)*, the intervals $A_1 \cap A_i$ are all different.

Thus, the number of possible intervals of this form is $m - 1 \leq (2|A_1| - 1)$.

We claim two sets $A_i \cap A_1$ and $A_j \cap A_1$, $i \neq j$, $i \neq 1$ and $j \neq 1$ cannot constitute a partition of A_1. As, if they constitute a partition of A_1, they will have to coincide with opposite sides of A_1. Else, they will violate *(3)*. But if they coincide with opposite sides of A_1, then $|A_i \cap A_j| = 0$, (*(1)* requires $|A_i| \leq \frac{n}{2}$)and thus, they will violate *(2)*. Hence, A_i and A_j cannot constitute a partition of A_1.

Thus, out of all total cases, only half of them are possible, i.e. $m - 1 \leq |A_1| - 1$. Hence, we get $m \leq |A_1|$, which completes the proof for (4).

Also, we have

$$\sum_{A_i \in A} |A_i|^{-1} \leq \frac{m}{|A_1|} \leq 1$$

which gives us (5).

Also, equality in *(5)* implies

$$1 = \sum_{1 \leq i \leq m} |A_i|^{-1} \quad \leq \quad \frac{m}{|A_1|} \quad \leq \quad 1 \tag{2.16}$$

So we have, $|A_i| = |A_1| = m$, $1 \leq i \leq m$. Thus, the A_i are intervals of length whose initial end points are m successive points on the circle.

Conversely, if the A_i satisfy *(1),(2),(3)* and have length m, then obviously we have an equality in *(5)*.

Theorem 2.2.3 **[6](Erdos-Ko-Rado)** *Let H be a simple intersecting hypergraph of order n and rank $r \leq \frac{n}{2}$. Then*

$$\sum_{E \in H} \binom{n-1}{|E|-1}^{-1} \leq 1 \tag{2.17}$$

and

$$m(H) \leq \binom{n-1}{r-1} \tag{2.18}$$

Further, we have equality in (2) when H is a star of K_r^n (and if ($r < \frac{n}{2}$).

Proof : Let $\mathbf{X} = \{x_1, x_2 ... x_n\}$ be the vertex set of H and for any permutation π of $1, 2, ... n$, denote by H_π the set of edges of H which are circular intervals of the circular sequence $x_{\pi_1}, x_{\pi_2} ... x_{\pi_n}, x_{\pi_1}$

Also, for $E \in H$, put

$$\beta(E) = |\{\pi / E \in H_\pi\}| \tag{2.19}$$

Also note that from lemma (2.2.2).

$$\sum_{E \in H_\pi} \frac{1}{|E|} \leq 1 \tag{2.20}$$

We then have,

$$\sum_{E \in H} \frac{\beta(E)}{|E|} = \sum_{E \in H} \sum_{\pi | E \in H_\pi} \frac{1}{|E|} = \sum_{\pi} \sum_{E \in H_\pi} \frac{1}{|E|} \leq n! \tag{2.21}$$

Let E_0 be an edge of H with $|E_0| = h$ and let x_0 be an element of E_0. Since E_0 is also an edge of the hypergraph $H^{'} = K_n^h(x_0)$ and from lemma (2.2.2) we have equality in (2.20) for $H^{'}$, we also have equality in (2.21) for $H^{'}$. Thus, we have

$$\frac{\beta(E_0)}{|E_0|} = \frac{1}{m(H')} \sum_{E' \in H'} \frac{\beta(E')}{|E'|} = \frac{n!}{m(H')} = n! \binom{n-1}{|E_0| - 1}^{-1} \tag{2.22}$$

or,

$$\sum_{E \in H} \binom{n-1}{|E| - 1}^{-1} = \frac{1}{n!} \sum_{E \in H} \frac{\beta(E)}{|E|} \leq \frac{n!}{n!} = 1 \tag{2.23}$$

Thus we have (2.17).

For the second part, note that every $E \in H$ satisfies $|E| \leq r \leq \frac{n}{2}$. Thus, we have

$$m(H) \binom{n-1}{r-1}^{-1} \leq \sum_{E \in H} \binom{n-1}{|E| - 1}^{-1} \leq 1 \tag{2.24}$$

(2.24) gives (2.18), thereby completing the proof.

The intersecting family is further generalized by the concept of t-intersecting family. For a hypergraph $H = (E_1, E_2 ... E_m)$, a t-intersecting family $\mathcal{A}$ is a set of edges that intersect in t or more vertices. Thus, if $\mathcal{A}$ is a t- intersecting family, then we have

$$|E_i \cap E_j| \geq t, \qquad \forall \quad E_i,\ E_j \in \quad H \tag{2.25}$$

For example, the family $\mathcal{A} = \{\{1,2,3,4\},\{2,3,4,5\},\{1,2,3,5\}\}$, is a 2-intersecting family on $\mathbf{X} = \{1,2,3,4,5\}$.

Erdos-Ko-Rado theorem for t-intersecting families is an important result, which is stated below without giving a proof.

Theorem 2.2.4 [9](Erdos-Ko-Rado Theorem for t-intersecting families) *Let $n \geq k \geq t \geq 1$, and let $\mathcal{A}$ be a family of $k-uniform$, $t-intersecting$ subsets of the set $[n] = \{1,2,...n\}$. If $n \geq (k-t+1)(t+1)$, then,*

$$|\mathcal{A}| \leq \binom{n-t}{k-t} \tag{2.26}$$

Moreover, if $n > (k-t+1)(t+1)$, then this bound is achieved by a trivially $t-intersecting$ system, that is by a family $\mathcal{A}$ containing all the $k-subsets$ of the set $[n]$ that contain a fixed $t-subset$ from the set $[n]$.

2.3 Section of a Hypergraph and Kruskal-Katona Theorem

The Kruskal-Katona theorem gives a tight lower bound on the size of $r-1$ section of an r-uniform hypergraph.

Theorem 2.3.1 [5](Kruskal, Katona) *Let H be an r-uniform hypergraph with*

$$m(H) = m = \binom{a_r}{r} + \binom{a_{r-1}}{r-1} + \binom{a_{r-2}}{r-2} + ... + \binom{a_s}{s} \tag{2.27}$$

and

$$a_r > a_{r-1} > ... > a_s \geq s \geq 1 \tag{2.28}$$

Then,

$$m([H]_{r-1}) \geq \binom{a_r}{r-1} + \binom{a_{r-1}}{r-2} + \binom{a_{r-2}}{r-3} + ... + \binom{a_s}{s-1} \tag{2.29}$$

The proof of Kruskal-Katona theorem presented here was given by Frankl [5]. This proof requires two lemmas which are stated and proved first before starting with the proof of the theorem. We now prove a lemma that demonstrates that every positive integer m has an r-binomial representation.

Lemma 2.3.2 : *Let m and r be positive integers. Then there exist integers $a_r, a_{r-1}, ...a_s$ such that*

$$m = \binom{a_r}{r} + \binom{a_{r-1}}{r-1} + \binom{a_{r-2}}{r-2} + ... + \binom{a_s}{s} \tag{2.30}$$

and

$$a_r > a_{r-1} > ... > a_s \geq s \geq 1 \tag{2.31}$$

Further, the a_i's are uniquely defined by (2.30) *and* (2.31) *and a_r is the largest integer such that*

$$m - \binom{a_r}{r} \geq 0 \tag{2.32}$$

Proof : The proof proceeds by induction on r. For any given m, with $r = 1$ the decomposition exists trivially and is unique, as $m = \binom{m}{1}$. We assume that for any $m > 0$, the decomposition exists with $r-1$ and is unique. Let a_r be the largest integer such that $m - \binom{a_r}{r} \geq 0$. Then from our assumption, a decomposition of $m - \binom{a_r}{r}$ with $r-1$ exists, i.e.

$$m - \binom{a_r}{r} = \binom{a_{r-1}}{r-1} + \binom{a_{r-2}}{r-2} + ... + \binom{a_s}{s} \tag{2.33}$$

with

$$a_{r-1} > a_{r-2}... > a_s \geq s \geq 1 \tag{2.34}$$

We must have $a_r > a_{r-1}$, else we would have

$$m \geq \binom{a_r}{r} + \binom{a_{r-1}}{r-1} = \binom{a_r+1}{r} \tag{2.35}$$

which is not in accordance with our assumption . Hence, the existence of decomposition is proved.

For proving uniqueness, lets assume two distinct decompositions exist:

$$m = \binom{a_r}{r} + \binom{a_{r-1}}{r-1} + \binom{a_{r-2}}{r-2} + ... + \binom{a_s}{s} = \binom{b_r}{r} + \binom{b_{r-1}}{r-1} + \binom{b_{r-2}}{r-2} + ... + \binom{b_s}{s} \tag{2.36}$$

where, the b_i's also satisfy equation (2.34). Now, observe that

$$m \leq \binom{a_r}{r} + \binom{a_r-1}{r-1} + \binom{a_r-2}{r-2} + ... + \binom{a_r-r-1}{s} = \binom{a_r+1}{r} \tag{2.37}$$

If $a_r < b_r$, then

$$m \leq \binom{a_r+1}{r} \leq \binom{b_r}{r} \leq m \tag{2.38}$$

This implies $m = \binom{a_r+1}{r}$, which violates the definition of r. Thus, $a_r = b_r$ and hence, the decomposition is unique. This decomposition of m is also called the r-binomial representation of m. It is of some importance and will be explored in detail in later chapters.

Lemma 2.3.3 *Let H be an r-uniform hypergraph on $\boldsymbol{X} = \{x_1, x_2, ...x_n\}$. Let $H(x_1)$ be the star of the vertex x_1. Then there exists an r-uniform hypergraph $[H']$ on X with $m(H') = m(H)$, $m([H']_{r-1}) \leq m([H]_{r-1})$ which satisfies*

$$F \in [H' - H'(x_1)]_{r-1} \Rightarrow F \cup \{x_1\} \in H' \tag{2.39}$$

Proof For a vertex $x_j \in \mathbf{X}$, $x_j \neq x_1$, put

$$\sigma_{x_j}(E) = \begin{cases} (E - \{x_j\}) \cup \{x_1\} \quad if \quad x_j \in E, \quad x_1 \notin E \quad and \quad (E - x_j) \cup \{x_1\} \notin H \\ E \qquad\qquad\qquad\qquad\qquad\qquad otherwise \end{cases} \tag{2.40}$$

Also, $\sigma_{x_j}(H) = \{\sigma_{x_j}(E)/E \in H\}$. We claim, $[\sigma_{x_j}(H)]_{r-1} \subset \sigma_{x_j}[H]_{r-1}$.

We have to show that $A \in [\sigma_{x_j}(H)]_{r-1} \Rightarrow A \in \sigma_{x_j}[H]_{r-1}$.

First, suppose that $A = \sigma_{x_j}(A)$.

If $B \in [\sigma_{x_j}(A)]_{r-1}$, then $B \in [A]_{r-1}$ (as $A = \sigma_{x_j}(A)$). Thus, $A = B \cup \{x_i\}$, for some $i \leq n$. Now, it suffices to prove that $\sigma_{x_j}(B) = B$, as this would imply $B \in \sigma_{x_j}[H]_{r-1}$.

case 1. If $i = j$, then, $x_j \notin B$ and hence, $B = \sigma_{x_j} B$.

case 2. If $i = j$, $x_j \in B$, then, $(B - \{x_j\}) \cup \{x_1\} = A - \{x_j\} \in [A]_{r-1}$. Thus, $\sigma_{x_j}(B) = B$.

case 3. If $i = 1$, $\;x_j \notin B$, then, $\sigma_{x_j}(B) = B$.

case 4. If $i \neq 1$, $i \neq j$, then $B = \sigma_{x_j}(B)$, unless $x_j \in B$ and $x_1 \notin B$. But in that case, $x_j \in A$ and $x_1 \notin A$. Also, we have $\sigma_{x_j}(A) = A$, so we must have $(A - \{x_j\}) \cup \{x_1\} \in H$. Thus, $(B - \{x_j\}) \cup \{x_1\} \in [H]_{r-1}$, and so, $\sigma_{x_j}(B) = B$.

Next, assume that $A \neq \sigma_{x_j}(A)$. Then, $x_j \in A$, $x_1 \notin A$ and $\sigma_{x_j}(A) = (A - \{x_j\}) \cup \{x_1\}$. Now if, $B \in [\sigma_{x_j}(A)]_{r-1}$ and $x_1 \notin B$, then $B = A - \{x_j\}$ and hence, $\sigma_{x_j}(B) = B$. If $B \in [\sigma_{x_j}(A)]_{r-1}$ and $x_1 \in B$, then $B^* = (B - \{x_1\}) \cup \{x_j\} \subset A$ and so $B^* \in [H]_{r-1}$. If $B \notin [H]_{r-1}$, then $\sigma_{x_j}(B^*) = B$ so that $B \in \sigma_{x_j}[H]_{r-1}$. If finally, $B \in [H]_{r-1}$, $x_j \notin B$ so that $B = \sigma_{x_j}(B)$. Thus, the proof is complete.

Now we can move on to the proof of *Kruskal-Katona Theorem.*

2.4 Proof of Kruskal-Katona Theorem

Assume that H satisfies

$$F \in [H - H(x_1)]_{r-1} \Rightarrow F \cup \{x_1\} \in H \tag{2.41}$$

Also let, $H_1 = \{E - \{x_1\}/E \in H(x_1)\}$. Then,

$$m([H]_{r-1}) \geq m(H_1) + m([H_1]_{r-2}) \tag{2.42}$$

The theorem holds trivially for $r = 1$ and $m = 1$. We proceed by induction on m and r. Suppose that

$$m(H_1) \geq \binom{a_r - 1}{r - 1} + ... + \binom{a_s - 1}{s - 1} \tag{2.43}$$

From the induction hypothesis, for the hypergraph H_1, we get

$$m([H_1]_{r-2}) \geq \binom{a_r - 1}{r - 2} + ... + \binom{a_s - 1}{s - 2} \tag{2.44}$$

Thus, from (2.43)

$$m([H]_{r-1}) \geq \binom{a_r - 1}{r - 1} + ... + \binom{a_s - 1}{s - 1} + \binom{a_r - 1}{r - 2} + ... + \binom{a_s - 1}{s - 2} \tag{2.45}$$

or,

$$m([H]_{r-1}) = \binom{a_r}{r - 1} + ... + \binom{a_s}{s - 1} \tag{2.46}$$

Now, suppose that

$$m(H_1) < \binom{a_r - 1}{r - 1} + \binom{a_{r-1} - 1}{r - 2} + ... + \binom{a_s - 1}{s - 1} \tag{2.47}$$

Thus,

$$m(H - H(x_1)) = m(H) - m(H_1) > \binom{a_r}{r} + ... + \binom{a_s}{s} - \binom{a_r - 1}{r - 1} - \binom{a_{r-1} - 1}{r - 2} - ... - \binom{a_s - 1}{s - 1} \tag{2.48}$$

or,

$$m(H - H(x_1)) > \binom{a_r - 1}{r} + ... + \binom{a_s - 1}{s} \tag{2.49}$$

But, we have

$$m(H_1) \geq m(H - H(x_1)) \geq \binom{a_r - 1}{r - 1} + ... + \binom{a_s - 1}{s - 1} \tag{2.50}$$

which violates (2.48). This completes our proof.

Corollary 2.4.1 *Let H be an r-uniform hypergraph and let k be an integer with $r > k \geq 2$. If a is the largest integer such that $m(H) \geq \binom{a}{r}$, then*

$$m([H]_k) \geq \binom{a}{k} \tag{2.51}$$

Proof Let H_1 be a partial hypergraph of H with $m(H_1) = \binom{a}{r}$. From theorem (2.4)

$$m([H_1]_{r-1}) \geq \binom{a}{r - 1} \tag{2.52}$$

Further, let H_2 be a partial hypergraph of $[H_1]_{r-1}$ with $m(H_2) = \binom{a}{r-1}$. From Theorem 2.3.1

$$m([H_2]_{r-2}) \geq \binom{a}{r - 2} \tag{2.53}$$

Continuing, we get

$$m([H_{r-k}]_k) \geq \binom{a}{k} \tag{2.54}$$

Since, $[H_{r-k}]_k \subset [H_k]$, we have

$$m([H_k]) \geq \binom{a}{k} \tag{2.55}$$

which completes our proof.

Chapter 3

Erdos-Ko-Rado Theorem for Multisets

3.1 Some Basics of Multisets

Multisets are generalizations of sets in which an element is allowed to appear more than once. As with sets, the order of elements in a multiset is irrelevant. The cardinality of a multiset is the number of elements including repetitions. Also, k-multiset system on an m set is a collection of multisets of cardinality k containing elements from m.

Also, we represent the set $\{1, 2, ...m\}$ by $[m]$.

Representation by a Family of Vectors: A family of vectors can be used to represent multisets. Let $\mathbf{X}$ be a set with $|\mathbf{X}| = n$. Then, multisets on this set can be represented by vectors of dimension n, with i^{th} component of the vector designating the multiplicity of i^{th} element of $\mathbf{X}$

For example, let $\mathbf{X} = \{1, 2, 3, 4, 5\}$. Then, the vector $v = (2, 0, 1, 3, 2)$ represents the multiset $\{1, 1, 3, 4, 4, 4, 5, 5\}$.

The intersection of two multisets S and T, designated by $S \cap T$, contains all the elements which are in both S and T. If a given element appears more than once in S or T (or both), the intersection contains k copies of that element, where k is the smaller of the number of times the element appears in S or T. For example, if $S = \{0, 1, 1, 2, 2, 2\}$ and $T = \{1, 2, 2, 3\}$, the intersection $S \cap T$ is $\{1, 2, 2\}$. Two multisets are said to be intersecting if they have at least one element in common. A collection of multisets is intersecting if each pair of multisets in that collection is intersecting. Erdos-Ko-Rado theorem for multisets gives the maximum size of an intersecting collection of k-multisets of a m set.

3.2 Erdos-Ko-Rado Theorem for Multisets

Theorem 3.2.1 [7](Erdos-Ko-Rado theorem for multisets) *Let k, m be positive integers and with $m \geq k + 1$. If $\mathcal{A}$ is an intersecting collection of multisets of $[m]$, then*

$$|\mathcal{A}| \leq \binom{m+k-2}{k-1} \tag{3.1}$$

Moreover, if $m > k + 1$, equality in (3.1) is achieved iff $\mathcal{A}$ is a collection of all the k-multisets of $[m]$, each containing a fixed element from $[m]$.

Proof : The proof of this theorem uses a homomorphism from a Kneser graph to a graph whose vertices are the k-multisets of $[m]$.

A Kneser graph $K(n, k)$, over a set $[n]$ is defined to be a graph whose vertices are all the k-sets of the set $[n]$, denoted by $\binom{[n]}{k}$, and two vertices are adjacent iff the k-sets they correspond to are disjoint. We represent by $\alpha(K(n, k))$ the size of largest independent set in $K(n, k)$. Note that an independent set of vertices in $K(n, k)$ is an intersecting k-set system.

We now define a multiset analogue of the Kneser graph. Let k, m be positive integers. Then $M(m,k)$ is defined to be a graph with vertices the $k-multisets$ of the set $[m]$, denoted by $\left(\binom{[m]}{k}\right)$, and two vertices of this graph are adjacent iff the multisets they correspond to are disjoint. Thus an independent set in $M(m,k)$ is an intersecting family of k-multisets on the set $[m]$. We denote by $\alpha(M(m,k))$ the size of maximum intersecting family of $M(m,k)$. Also, the number of vertices in $M(m,k)$ is $\binom{m+k-1}{k}$.

Further, let $n = m+k-1$. Then $K(n,k)$ has the same number of vertices as $M(m,k)$ and $\forall \;\; B \in \binom{[n]}{k}$, $B \cap [m] \neq \phi$.

For a set $A \subseteq [m]$ of cardinality a, where $1 \leq a \leq k$, the number of $k-sets$,B, from $[n]$, such that $B \cap [m] = A$ will be equal to

$$\binom{n-m}{k-a} = \binom{k-1}{k-a} \tag{3.2}$$

Similarly, the number of k-multisets from $[m]$ which contain all of the elements of A and no others is

$$\binom{a+(k-a)-1}{k-a} = \binom{k-1}{k-a} \tag{3.3}$$

Hence, there exists a bijection from $f : K(n,k) \to M(m,k)$ such that for any $B \in V(K(n,k))$, the set of distinct elements in $f(B)$ is $B \cap [m]$, where $V(K(n,k))$ is the vertex set of $K(n,k)$.

If $A, B \in \binom{[n]}{k}$ are two adjacent vertices of $K(n,k)$, then $(A \cap [m]) \cap (B \cap [m]) = \phi$. Thus if A and B are adjacent, $f(A)$ and $f(B)$ are also adjacent. So, the bijection $f : \binom{[n]}{k} \to \left(\binom{[m]}{k}\right)$ is a graph homomorphism.

Thus, we have

$$\alpha(M(m,k)) \leq \alpha(K(n,k)) \tag{3.4}$$

From Erdos-Ko-Rado theorem, we have, if $n \geq 2k$,

$$\alpha(K(n,k)) \leq \binom{n-1}{k-1} \tag{3.5}$$

Thus, we have for $m \geq k+1$

$$\alpha(M(m,k)) \leq \binom{n-1}{k-1} = \binom{m+k-2}{k-1} \tag{3.6}$$

An intersecting collection of k-multisets from $[m]$ consisting of all k-multisets containing a fixed element from $[m]$ will have size

$$\binom{m+k-2}{k-1,}$$

which gives the upper bound on the size of $\mathcal{A}$ in theorem, which completes the proof of (3.1).

Let $m > k+1$ and let $\mathcal{A}$ be an intersecting multiset of size $\binom{m+k-2}{k-1}$. From the homomorphism defined above, the pre-image of $\mathcal{A}$ will be an independent set of $K(n,k)$ of size $\binom{n-1}{k-1}$. Using $m > k+1$ and $n = m = k-1$, we get $n > 2k$. From Erdos-Ko-Rado theorem, we get that $f^{-1}(\mathcal{A})$ will be a collection of $k-subsets$ of the set $[n]$, each containing a fixed element from $[n]$, say x_0. If $x_0 \in [m]$, then it follows from the definition of f that every multiset in $\mathcal{A}$ contains the element x_0. Thus $\mathcal{A}$ will be a family of multisets each containing a fixed element from $[m]$, as required.

If $x_0 \notin [m]$, then $f^{-1}(\mathcal{A})$ will contain the sets $A = \{1, m+1, m+2...n\}$ and $B = \{2, m+1, m+2...n\}$, as $m > k+1$ implies $m > 2$. But $f(A) \cap f(B) = 0$ which violates our initial assumption $\mathcal{A}$ is an intersecting collection of multisets.

Therefore, when $m > k+1$ and $\mathcal{A}$ is an intersecting family of maximum size, then $\mathcal{A}$ consists of all the k-multisets of the set $[m]$ containing a fixed element from $[m]$.

Theorem 3.2.2 *Let m, k be positive integers with $m \leq k$. If $\mathcal{A}$ is an intersecting family of multisets on the set $[m]$.*

Then, if m is odd:

$$|\mathcal{A}| \leq |\mathcal{M}_{(>\frac{m}{2})}| \tag{3.7}$$

and equality in holds (3.7) *iff* $\mathcal{A} = \mathcal{M}_{(>\frac{m}{2})}$*, where* $\mathcal{M}_{(>\frac{m}{2})}$ *is the collection of all the multisets that contain more than* $\frac{m}{2}$ *distinct elements from the set [m].*

And, if m is even:

$$|\mathcal{A}| \leq |\mathcal{M}_{(>\frac{m}{2})}| + \frac{1}{2}|\mathcal{M}_{(\frac{m}{2})}| \tag{3.8}$$

Equality in (3.8) *holds iff* $\mathcal{A}$ *consists of* $\mathcal{M}_{(>\frac{m}{2})}$ *and a maximal intersecting family from* $\mathcal{M}_{(\frac{m}{2})}$*, where* $\mathcal{M}_{(\frac{m}{2})}$ *is the collection of all the multisets that contain exactly* $\frac{m}{2}$ *distinct elements from the [m] set.*

Proof : First we have

$$|\mathcal{M}_{(\frac{m}{2})}| = \binom{m}{\frac{m}{2}}\binom{k-1}{k-\frac{m}{2}} \tag{3.9}$$

and

$$|\mathcal{M}_{(>\frac{m}{2})}| = \sum_{j=\lceil\frac{m+1}{2}\rceil}^{m} \binom{m}{j}\binom{k-1}{k-j} \tag{3.10}$$

For any multiset A on a set $[m]$, the support of A is the set $S_A \subset [m]$ consisting of all the distinct elements from $[m]$ that come in A.

Also, notice that two k-multisets A and B on $[m]$ will be intersecting iff $(S_A \cap S_B) \neq \phi$ and that each S_A will have a unique complement S'_A defined by $S'_A = [m]/S_A$ in $[m]$.

Let $\mathcal{A}$ be a family of intersecting multisets on set $[m]$ and $M \in \mathcal{A}$ be a multiset such that $|S_M| = \{min|S_A|/A \in \mathcal{A}\}$. The theorem holds trivially for $m = 2$, so we will assume $m > 2$.

Suppose $|S_M| < \frac{m}{2}$. Let $B_1 = \{A \in \mathcal{A}/S_A = S_M\}$ and let B_2 be family of k-multisets on the set $[m]$ such that, if $B \in B_2$, then $S_B = S'_M$. Thus, we have, $B_1 \subseteq \mathcal{A}$ and $B_2 \cap \mathcal{A} = \phi$.

Consider the family $A' = (A/B_1) \cup B_2$. We want to show that A' is an intersecting family larger than A. For this, notice that every multiset in the family $\mathcal{A}/B_1$ contains at least one element from $[m]/S_M$ and $[m]/S_M= S_B \; \forall \quad B \in B_2$. Thus A' is an intersecting collection of k-multisets.

Let $|S_M| = i$. Then,

$$|B_1| = \binom{k-1}{k-i} \tag{3.11}$$

and

$$|B_2| = \binom{k-1}{k-m+i} \tag{3.12}$$

Now, to show that $|A'| > |A|$, it is sufficient to show that

$$\binom{k-1}{k-m+i} > \binom{k-1}{k-i} \tag{3.13}$$

or, equivalently

$$(k-i)!(i-1)! > (k-m+1)!(m-i-1)! \tag{3.14}$$

Since, $i < \frac{m}{2}$ and $m \leq k$, we have $k - i > k - \frac{m}{2} > k - m + 1 \geq 1$. Therefore,

$$(k-i)!(i-1)!$$

$$\begin{aligned}
&= (k-i)(k-i-1)...(k-m+i+1)(k-m+i)(i-1)! \\
&\geq (m-i)(m-i-1)...(i+1)(k-m+i)(i-1)! \\
&= \frac{m-i}{i}(m-i-1)!(k-m+i)! \\
&> (m-i-1)!(k-m+1)!
\end{aligned}$$

which is what we required.

Thus, if $\mathcal{A}$ is of maximum size, it cannot contain a multiset with less than $\frac{m}{2}$ distinct elements from $[m]$. Any k-multiset with more than $\frac{m}{2}$ distinct elements from $[m]$ will intersect with any other such k-multiset. This completes the proof for the case when m is odd. When m is even, we will have to take care of the multisets that contain exactly $m/2$

distinct elements. The multisets in $\mathcal{M}_{(\frac{m}{2})}$ intersect with any multiset that contain more than $\frac{m}{2}$ distinct elements. Further, $\mathcal{M}_{(\frac{m}{2})}$ is not an intersecting family. Since the size of maximal intersecting collection of $\frac{m}{2}$ subsets of $[m]$ is $\binom{m}{m/2}/2$, the maximum intersecting multiset family contains half of the elements from $\mathcal{M}_{(\frac{m}{2})}$. Thus, the proof for the theorem is complete.

Chapter 4

Kruskal-Katona Theorem for Multisets

For proving Kruskal-Katona theorem for multisets, we need to consider a different representation of the Kruskal-Katona Theorem. For this, we need some basic concepts.

4.1 Squashed Ordering of Sets

Consider a set $S = \{1, 2, ..n\}$ and the k-subsets of the set S. Given two k-subsets A and B of S, we define an inequality as: $A <_L B$ if the smallest element of the symmetric difference $A + B = (A \cap B') \cup (A' \cap B)$ is in A, where both A and B are k-subsets of the set S. The ordering of k-subsets thus obtained is called the lexicographic ordering of k-subsets of the set S.

For example, consider $S = \{1, 2, 3, 4, 5\}$ and let $k = 3$. Then the 3-subsets in lexicographic ordering are:

$$\{1,2,3\}$$

$$\{1,2,4\}$$
$$\{1,2,5\}$$
$$\{1,3,4\}$$
$$\{1,3,5\}$$
$$\{1,4,5\}$$
$$\{2,3,4\}$$
$$\{2,3,5\}$$
$$\{2,4,5\}$$
$$\{3,4,5\}$$

Note that lexicographic ordering is similar to dictionary ordering of words.

Given two k-subsets A and B of the set S, we say $A <_S B$ if the largest element of the symmetric difference $A + B$ is in B. Using this inequality, we can arrange the k-subsets in an ordering, called the squashed ordering. This squashed ordering plays an important role in Kruskal-Katona theorem. For example, the 3-subsets of the set $S = \{1, 2, 3, 4, 5\}$ in squashed ordering are:

$$\{1,2,3\}$$
$$\{1,2,4\}$$
$$\{1,3,4\}$$
$$\{2,3,4\}$$
$$\{1,2,5\}$$
$$\{1,3,5\}$$
$$\{2,3,5\}$$
$$\{1,4,5\}$$
$$\{2,4,5\}$$

{3,4,5}

Squashed Ordering and r-binomial Representation of a Number

Squashed ordering of k-subsets of a set S is related to the r-binomial representation of a number. Given m and r, consider the r-binomial representation of m.

$$m = \binom{a_r}{r} + \binom{a_{r-1}}{r-1} + \binom{a_{r-2}}{r-2} + ... + \binom{a_s}{s}.$$

In the squashed ordering of r-subsets of the set $S = \{1, 2...n\}$, first $\binom{a_r}{r}$ subsets are the r-subsets of $\{1, 2, ...a_r\}$.

The next $\binom{a_{r-1}}{r-1}$ subsets are those obtained by adjoining $a_r + 1$ to the $(r-1)$-subsets of $\{1, 2, ...a_{r-1}\}$ and so on till the final $\binom{a_s}{s}$ are those obtained by adjoining $\{a_{s+1}+1, ...a_r+1\}$ to the $s-subsets$ of $\{1, 2....a_s\}$. Thus, the m^{th} r-subset in squashed order is $\{a_r+1, a_{r-1}+1, ...a_{s+1}+1, a_s, a_s-1, ...a_s-s+1\}$. Note that this does not depend on n, the cardinality of the set S.

For example, let $m = 9$ and $r = 3$.

Then, we have, $9 = \binom{4}{3} + \binom{3}{2} + \binom{2}{1}$

Hence, the 9^{th} set in squashed ordering is $\{5, 4, 2\}$.

4.2 Restating Kruskal-Katona Theorem

Consider a hypergraph H on S, with $m(H) = m$, with the r-binomial representation of m given by

$$m = \binom{a_r}{r} + \binom{a_{r-1}}{r-1} + \binom{a_{r-2}}{r-2} + ... + \binom{a_s}{s}.$$

Also, consider the first m r-subsets of S in the squashed ordering. Note that the $(r-1)$ subsets contained in them are:

all $\binom{a_r}{r-1}$ $(r-1)$ subsets from $\{1, 2, ..a_r\}$.

all $\binom{a_{r-1}}{r-2}$ $(r-2)$ subsets from $\{1, 2...a_r - 1\}$ combined with $\{1 + a_r\}$.

$$\vdots$$

all $\binom{a_s}{s-1}$ $(s-1)$ subsets from $\{1, 2...a_s\}$ combined with $\{1 + a_r, 1 + a_{r-1}...1 + a_{s+1}\}$.

Thus, there are $\binom{a_r}{r-1} + \binom{a_{r-1}}{r-2} + ... + \binom{a_s}{s-1}$ $(r-1)$ subsets. Also, note that these $(r-1)$ subsets are in the squashed ordering.

Let $H = (E_1, E_2...E_m)$ be a family of k-subsets of the set $S = \{1, 2, ...n\}$. Then, the compression of H, denoted by $\mathcal{C}H$, is defined to be the collection of sets containing first $|H|$ k-subsets of S in the squashed ordering.

Now, Kruskal-Katona theorem can be restated as:

Theorem 4.2.1 [1] ***(Kruskal-Katona Theorem)*** *For an r-uniform hypergraph H,*

$$[\mathcal{C}H]_{r-1} \subseteq \mathcal{C}[H]_{r-1} \tag{4.1}$$

Essentially, Kruskal-Katona theorem states that for any given m, the number of $(r-1)$ subsets in $H = (E_1, E_2...E_m)$ over S is minimized by taking H to be a collection of first m k-subsets of S in the squashed ordering.

4.3 Kruskal-Katona Theorem for Multisets

For studying multisets over the set $[n]$, its easier to work with the vector representation of multisets. In this section, we will represent multisets with n-tuples as described earlier. A multiset family $S(k_1, k_2...k_n)$, by definition, contains all the vectors $x = (x_1, x_2, ...x_n)$, such that each x_i is an integer satisfying $0 \leq x_i \leq k_i$. The rank of x is defined to be $|x| = x_1 + x_2... + x_n$. The vectors in S of a given rank are arranged lexicographically as: If

$a = (a_1, ...a_n)$ and $b = (b_1, ...b_n)$, we say $a <_L b$ if $a_1 < b_1$ or if $a_1 = b_1, ..., a_{i-1} = b_{i-1}, a_i < b_i$. As an example, consider $S(2, 3, 4)$. The vectors of rank 3 in lexicographic order are

$$(0\ 0\ 3)$$
$$(0\ 1\ 2)$$
$$(0\ 2\ 1)$$
$$(0\ 3\ 0)$$
$$(1\ 0\ 2)$$
$$(1\ 1\ 1)$$
$$(1\ 2\ 0)$$
$$(2\ 0\ 1)$$
$$(2\ 1\ 0)$$

If $\mathcal{A}$ is a collection of m k-vectors of $S(k_1, k_2...k_n)$, then $[\mathcal{A}]_{r-1}$ is given by

$$[\mathcal{A}]_{r-1} = \{x = (x_1...x_n) : |x| = k - 1 : (x_1...x_{i-1}, x_i + 1, x_{i+1}...x_n) \in \mathcal{A} \ \ i \leq n\} \tag{4.2}$$

Theorem 4.3.1 *Let $k_1 \leq k_2... \leq k_n$ and let $\mathcal{A}$ be a collection of k-vectors of the multiset $S(k_1, k_2...k_n)$. Then*

$$[\mathcal{C}\mathcal{A}]_{r-1} \subset \mathcal{C}[\mathcal{A}]_{r-1} \tag{4.3}$$

This implies

$$|[\mathcal{C}\mathcal{A}]_{r-1}| \leq |[\mathcal{A}]_{r-1}| \tag{4.4}$$

The proof of this theorem was given by *Clements and Lindstrom* [4] and requires some lemmas.

Lemma 4.3.2 *If $\mathcal{A}$ is a collection of k-vectors of $S(k_1, k_2...k_n)$ and if $\mathcal{A}$ is compressed, then $[\mathcal{A}]_{r-1}$ is also compressed.*

Let $\mathcal{A}$ be a collection of k-vectors of the multiset $S(k_1, ...k_n)$, denote by $\mathcal{A}_{i:d}$ the collection of those members of $\mathcal{A}$ whose i^{th} component is d, and let $\mathcal{CA}_{i:d}$ denote the first $|\mathcal{A}_{i:d}|$ with i^{th} component d. This is called i-compression and we say $\mathcal{A}$ is i-compressed if $\mathcal{CA}_{i:d} = \mathcal{A}_{i:d}$ for each $d = 0, 1, 2...k_i$.

Starting with any collection of k-vectors $\mathcal{A}$, define a sequence $\mathcal{A}^1, \mathcal{A}^2...$ as follows:

$\mathcal{A}^1 = \mathcal{A}$

$\mathcal{A}^2$= Union of all $\mathcal{CA}^1_{i:d}$, $(d = 0, 1, ...k_1)$

$\mathcal{A}^3$= Union of all $\mathcal{CA}^2_{2:d}$, $(d = 0, 1, ...k_2)$

and so on cyclically , such that

$\mathcal{A}^{m+1}$= Union of all $\mathcal{CA}^m_{r:d}$, $(d = 0, 1, ...k_r)$

where $r \equiv m(mod \quad n)$. If $A^j \neq A^{j+1}$, then at least one member of A^j is being replaced by an earlier vector in lexicographic order. Eventually, no more such replacements will be possible and hence we get the following lemma:

Lemma 4.3.3 *There exists a positive integer p such that $\mathcal{A}^p$ is i-compressed for all $i = 1, 2, ...n$.*

Lemma 4.3.4 *Let $n \geq 3$, $a = (a_1, a_2, ...a_n)$, $b = (b_1, ...b_n)$, $|a| = |b| = k$, $a < b$ and $b_n = 0$ or $a_n = k_n$. Then if $b \in \mathcal{A}$, where $\mathcal{A}$ is i-compressed, then $a \in \mathcal{A}$.*

Proof We shall find a sequence of k-vectors from a to b such that any two consecutive members of the sequence agree in the first, second or n^{th} component. It will then follow that all the members of the sequence, including a, are in $\mathcal{A}$. First we deal with the case when $a_n = b_n$.

If $a_1 = b_1$, then the sequence $a < b$ suffices, so now suppose that $a_1 < b_1$. First subcase to be considered is $a_i > 0$ for some i such that $2 \leq i \leq n-1$. In this case we have

$$a = (a_1, ...a_n) < (a_1 + 1, a_2'', ..., a_{n-1}'', a_n) \tag{4.5}$$

where $a_2^{''}, ..., a_{n-1}^{''}$ are chosen such that $a_2^{''} + ... + a_{n-1}^{''}$=$a_2 + ... + a_{n-1} - 1$, and so that $(a_2{''}...a_{n-1}^{''})$ is early in the lexicographic order as possible. If $a_1 + 1 = b_1$, we have $a < (b_1, a_2^{''}, ...a_{n-1}^{''}, a_n) \leq b$ as required. If $a_1 + 1 < b_1$ and $a_i^{''} > 0$ for some i such that $2 \leq i \leq n-1$, repeat this process. Either $b_1 - a_1$ applications of this process will give a sequence as required or, we enter the second subcase where we have $a < ... < a^{''}$, $a^{''} < b_1, a_2^{''} = ... = a_{n-1}^{''} = 0$. But then we have $a^{''} = (a_1^{''}, ...a_{n-1}^{''}, k_n) < (b_1, a_2^{''}, ...a_{n-1}^{''}, k_n - b_1 + a_1^{''}) \leq b$, as required.

We next consider the case when $b_n = 0$. The above argument can be applied to $b' = (k_1 - b_1, ..., k_n - b_n)$ and a'. The compliments of vectors from b' to a' give us a sequence from a to b.

Thus, our proof is complete.

Lemma 4.3.5 *Suppose that theorem* (4.3.1) *is true in* $(n-1)$ *dimensions and that* $\mathcal{B}$ *is a collection of* $(k-1)$ *vectors of* $S(k_1, ...k_n)$, $k_1 \leq k_2 \leq ...k_n$, *such that* $[\mathcal{A}]_{r-1} \subseteq \mathcal{B}$. *Then* $[\mathcal{A}^j]_{r-1} \subseteq \mathcal{B}^j \ \forall \ j \geq 1$.

Proof Since the lemma is true for $j = 1$, we use induction on j. Let S_k denote the set of all the k-vectors of $S(k_1, ...k_n)$.

Suppose that

$$[\mathcal{A}^j]_{r-1} \subseteq \mathcal{B}^j.$$

Also, notice that

$$\mathcal{A}^{j+1} = \cup_d \mathcal{C}((\mathcal{A}^|)_{i:d}) \quad (i \equiv j(mod \ \ n)) \tag{4.6}$$

If $d > 0$, the members of $[(\mathcal{A}^j)_{i:d}]_{r-i}$ are of two types, those whose i^{th} component is d and those whose i^{th} component is $(d-1)$. First, consider those whose i^{th} component is d. They constitute $[(\mathcal{A}^j)_{i:d}]_{r-i} \cap (S_{k-1})_{i:d}$ and thus, we have $[(\mathcal{A}^j)_{i:d}]_{r-i} \cap (S_{k-1})_{i:d} \subseteq (\mathcal{B}^j)_{i:d}$

$([\mathcal{A}^j]_{r-1} \subseteq \mathcal{B}^j)$. Now, $(\mathcal{A}^j)_{i:d}$ has $(n-1)$ effective components, so from theorem (4.3.1)

$$[\mathcal{C}((\mathcal{A}^j)_{i:d})]_{r-1} \cap (S_{k-1})_{i:d} \subseteq \mathcal{C}([(\mathcal{A}^j)_{i:d}])_{r-1} \cap (S_{k-1})_{i:d} \subseteq \mathcal{C}((\mathcal{B}^j)_{i:d}) \tag{4.7}$$

Next, consider the members of $[(\mathcal{A}^j)_{i:d}]_{r-1}$ whose i^{th} component is $(d-1)$. Also, notice that different members of $\mathcal{A}^j$ with i^{th} component d give rise to different members of $[\mathcal{A}^j]_{r-1}$ with i^{th} component $(d-1)$. Thus, we have

$$|(\mathcal{A}^j)_{i:d}| = |[(\mathcal{A}^j)_{i:d-1}]_{r-1}| \subseteq |(\mathcal{B}^j)_{i:d-1}| \tag{4.8}$$

for $d \geq 1$, so that

$$|\mathcal{C}((\mathcal{A}^j)_{i:d})| \leq |\mathcal{C}((\mathcal{B}^j)_{i:d-1})| \tag{4.9}$$

Previously, we proved the result for those members of $[\mathcal{C}((\mathcal{A}^j)_{i:d})]_{r-1}$ with i^{th} component d. The other members constitute $[\mathcal{C}(\mathcal{A}^j)_{i:d}]_{r-1} \cap (S_{k-1})_{i:d-1}$, which by lemma (4.3.2) consists of first $|\mathcal{C}((\mathcal{A}^j)_{i:d})|$ members of $(S_{k-1})_{i:d-1}$. Since $\mathcal{C}((\mathcal{B}^j)_{i:d-1})$ consists of first $|\mathcal{C}((\mathcal{B}^j)_{i:d-1})|$ members of $(S_{k-1})_{i:d-1}$, (4.9) yields

$$[\mathcal{C}((\mathcal{A}^j)_{i:d})]_{r-1} \cap (S_{k-1})_{i:d-1} \subseteq \mathcal{C}((\mathcal{B}^j)_{i:d-1}) \tag{4.10}$$

From (4.8) and (4.10) we obtain

$$[\mathcal{C}((\mathcal{A}^j)_{i:d})]_{r-1} \subseteq \mathcal{C}((\mathcal{B}^j)_{i:d}) \cup \mathcal{C}((\mathcal{B}^j)_{i:d-1}) \tag{4.11}$$

for each $d \geq 1$. Since, from our assumption, we also have

$$[\mathcal{C}((\mathcal{A}^j)_{i:0})]_{r-1} \subseteq \mathcal{C}([(\mathcal{A}^j)_{i:0})]_{r-1} \subseteq \mathcal{C}((\mathcal{B}^j)_{i:0}) \tag{4.12}$$

We finally obtain from (4.6)

$$[\mathcal{A}^{j+1}]_{r-1} = \cup_d [\mathcal{C}((\mathcal{A}^j)_{i:d})]_{r-1} \subseteq \cup_d \mathcal{C}((\mathcal{B}^j)_{i:d}) = \mathcal{B}^{j+1} \tag{4.13}$$

Thus, our proof is complete.

4.4 Proof of Kruskal-Katona Theorem

Consider a collection of k-vectors of the multiset $S(k_1, k_2, ...k_n)$, where $k_1 < k_2... < k_n$. The theorem holds trivially for $k = 2$. Assume that its true in $(n-1)$ dimensions, and consider the induction step from $(n-1)$ to n.

By lemma (4.3.3) $\exists$ a positive integer p such that $\mathcal{V} = \mathcal{A}^p$ is i-compressed for $i = 1, 2, ...n$. Let $\mathcal{W} = ([\mathcal{A}]_{r-1})^p$. If we take $\mathcal{B} = [\mathcal{A}]_{r-1}$, from lemma (4.3.5) we get $[\mathcal{V}]_{r-1} \subseteq \mathcal{W}$. Next we prove that $\mathcal{V}$ can be altered to $\mathcal{CA}$ and $\mathcal{W}$ to a subset of $\mathcal{C}([\mathcal{A}]_{r-1})$ in such a way that $[\mathcal{CA}]_{r-1} \subseteq \mathcal{C}[\mathcal{A}]_{r-1}$ is obtained.

First, we consider the case $\mathcal{V} = S_k$, i.e. when every k-vector is in $\mathcal{V}$. Then $|\mathcal{V}| = |\mathcal{A}^p| = |\mathcal{A}| = |\mathcal{CA}|$, so that $\mathcal{CA} = S_k$. Also, $[S]_{k-1} = [\mathcal{V}]_{r-1} \subseteq \mathcal{W}$. So, we must have $\mathcal{W} = [S]_k - 1$. Since $|\mathcal{W}| = |[\mathcal{A}]_{r-1}| = |\mathcal{C}[\mathcal{A}]_{r-1}|$, it follows that $\mathcal{C}[\mathcal{A}]_{r-1} = [S]_{k-1}$ and hence $[\mathcal{CA}]_{r-1} = [S]_{k-1}$, hence $[\mathcal{CA}]_{r-1} = [S]_{r-1} = \mathcal{C}[\mathcal{A}]_{r-1}$, as required.

Next assume that $\mathcal{V} \neq S_k$. Let a be the first vector of S_k which is not in $\mathcal{V}$, and let b be the last vector of $\mathcal{V}$. If $b < a$ then $\mathcal{V} = \mathcal{CA}$ and $[\mathcal{CA}]_{r-1} \subseteq \mathcal{W}$, where $|\mathcal{W}| = |[\mathcal{A}]_{r-1}|$. If $b > a$ and $b_n = 0$, then lemma (4.3.4) when applied to $\mathcal{V}$ would give $a \in \mathcal{V}$, which is a contradiction. Thus, we must have $b_n > 0$. Define

$$b^* = (b_1, b_2...b_{n-1}, b_n - 1) \tag{4.14}$$

and

$$a^* = (a_1, a_2...a_{n-1}, a_n - 1) \quad , \qquad if \quad a_n \quad > \quad 0 \tag{4.15}$$

Since, $[\mathcal{V}]_{r-1} \subseteq \mathcal{W}$, we have $b^* \in \mathcal{W}$. Now, all the k-vectors in $[b^*]_{r-1}$, other than b, must come after b in the ordering, therefore the vector b^* in $[\mathcal{V}]_{r-1}$ comes from b and from no other vector in $\mathcal{V}$. Next we alter $\mathcal{V}$ and $\mathcal{W}$ as follows.

Define

$$\mathcal{V}^* = (\mathcal{V} - \{b\}) - \{a\} \tag{4.16}$$

and

$$\mathcal{W}^* = (\mathcal{W} - \{b\}) \cup \{a\} \qquad (if \quad a_n > 0) \tag{4.17}$$

or,

$$\mathcal{W}^* = \mathcal{W} \qquad (otherwise) \tag{4.18}$$

We show that $[\mathcal{V}^*]_{r-1} \subseteq \mathcal{W}^*$. Since $b^* \notin [\mathcal{V} - \{b\}]_{r-1}$, it suffices to prove that $[\{a\}]_{r-1} \subseteq \mathcal{W}^*$. If $a_n > 0$, then $a^* \in [\{a\}]_{r-1}$ and thus $a^* \in \mathcal{W}^*$. For other members, first note that $a_n < k_n$, since otherwise lemma (4.3.4) would give $a \in \mathcal{V}$. If $a_i > 0$ for some $i \leq n-1$ then $(a_1...a_i - 1...a_n)$ is there in $[\mathcal{V}]_{r-1}$ as $(a_1, ..a_i - 1, ...a_n + 1)$ precedes a in ordering and a is the first vector of S_k not in $\mathcal{V}$. However, $[\mathcal{V}]_{r-1} \subseteq \mathcal{W}$, so we have $[\{a\}]_{r-1} \subseteq \mathcal{W}$. To show that $[\{a\}]_{r-1} \subseteq \mathcal{W}^*$, we must show that $b^* \notin [\{a\}]_{r-1}$. Suppose that

$$b^* = (a_1...a_i - 1...a_n + 1) \tag{4.19}$$

Then $b = (a_1...a_i - 1...a_n + 1)$ contradicting the fact $a < b$.

Now we have converted $\mathcal{V}$ and $\mathcal{W}$ to $\mathcal{V}*$ and $\mathcal{W}*$ with $[\mathcal{V}*]_{r-1} \subseteq \mathcal{W}*$. Now $\mathcal{V}*$ is i-compressed for all i, just as $\mathcal{V}$ was and so, we can repeat the process, each time replacing the last vector in $\mathcal{V}$ with an earlier vector. After a finite number of steps, $\mathcal{V}$ will be compressed to $\mathcal{CA}$ and $\mathcal{W}$ to a set $\mathcal{U}$ satisfying

$$[\mathcal{CA}]_{r-1} \subseteq \mathcal{U} \tag{4.20}$$

Now, $\mathcal{C}[\mathcal{A}]_{r-1}$ consists of first $|[\mathcal{A}]_{r-1}|$ members of S_{k-1} while, by lemma (4.3.2), $[\mathcal{CA}]_{r-1}$ consists of first $|[\mathcal{CA}]_{r-1}|$ members of S_{k-1}. Also

$$|[\mathcal{CA}]_{r-1}| \leq |\mathcal{U}| = |\mathcal{W}| = |[\mathcal{A}]_{r-1}| \tag{4.21}$$

Hence $[\mathcal{CA}]_{r-1} \subseteq \mathcal{C}[\mathcal{A}]_{r-1}$, as required.

Thus, the proof for Kruskal Katona theorem for multisets is complete. For details and further discussions on this topic, see [1].

Chapter 5

Conclusions

In this thesis, extensions of Erdos-Ko-Rado theorem and Kruskal-Katona theorem are given for multisets. Erdos-Ko-Rado theorem gives the size and structure of the largest collection of intersecting k-multisets. An obvious open problem is determining the size and structure of the largest collection of t-intersecting ($t \geq 2$) k-multisets, that is collection of multisets where the size of the intersection for every pair of multisets is at least t. The following conjecture was given by Brockman and Kay [3].

Conjecture 5.0.1 *Let k, m and t be positive integers, with $t \leq k$ and $m \geq t(k-t)+2$. If $\mathcal{A}$ is a collection of intersecting k-multisets of $[m]$, then*

$$|\mathcal{A}| \leq \binom{m+k-t-1}{k-t} \tag{5.1}$$

Moreover, if $m > t(k-t)+2$, equality holds iff $\mathcal{A}$ is a collection of all the k-multisets from $[m]$, each containing a fixed multiset from $[m]$.

The homomorphism from $K(n,k)$ to $M(m,k)$, as defined in the proof of Theorem 3.3.1, gave a straightforward way to show that the size of largest independent set in $M(m,k)$ is no larger than the size of largest independent set in $K(n,k)$. The ideas from the proof of Theorem 3.3.1 can be generalized as: let $K(n, k, t)$ be the graph whose vertices are the k-subsets of $[n]$ where two vertices A, B are adjacent if $|A \cap B| < t$. Also let $M(m, k, t)$ be the graph whose vertices are the k-multisets from $[m]$ where two vertices C, D are adjacent if $|C \cap D| < t$. If we could show the existence of a homomorphism from $K(n, k, t)$ to $M(n, k, t)$, it could be used to prove a bound on maximum size of a t-intersecting collection using Theorem 2.2.4 [9].

Bibliography

[1] Anderson, I., *Combinatorics of Finite Sets*, Dover publications, New York, 1987.

[2] Berge, C., *Hypergraphs: Combinatorics of Finite Sets*, volume 45 of North Holland Mathematical Library, Elsevier Science Publishers B.V., Amsterdam, 1989.

[3] Brockman, G., Kay, B., Elementary techniques for Erdos-Ko-Rado like theorems, *ArXiv e-prints*, August 2008.

[4] Clements, G.F., Lindstrm, B., A Generalization of a Combinatorial Theorem of Macaulay, *J. of Combinatorial Theory* 7, 230-238, 1969.

[5] Frankl, P., A new short proof for the Kruskal-Katona theorem, *Discrete Math.* 48, 327-329, 1984.

[6] Greene, C., Katona, G.O.H., and Kleitman, D.J., Extensions of the Erdos-Ko-Rado theorem, *SIAM J.* 55, 1-8, 1976.

[7] Meagher, K., Purdy, A., An Erdos-Ko-Rado theorem for multisets *The Electronic Journal of Combinatorics* 16, 2000.

[8] Lubell, D., A short proof of Sperner's lemma, *Journal of Combinatorial Theory* 1(2), p. 299, 1966.

[9] Wislon, R.M., The exact bound in Erdos-Ko-Rado theorem, *Combinatorica* 4(2-3), 247-257, 1984.