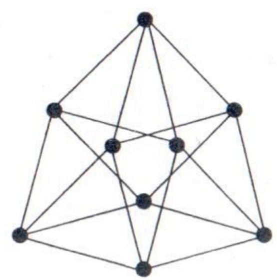
BÉLA BOLLOBÁS

Combinatorics

SET SYSTEMS, HYPERGRAPHS, FAMILIES OF VECTORS, AND COMBINATORIAL PROBABILITY



Prove the theorem of Shearer and Kleitman (1979) stating that for $n \geq 2$ there exist two orthogonal partitions of $\mathcal{P}(X)$ into chains.

(Give explicit constructions for $n = 2$ and 3. For $n \geq 4$ take the partition into symmetric chains given in the proof of Theorem 1 and characterized in Exercise 4. Replacing every set by its complement, obtain another partition. Prove that these partitions are almost orthogonal: only $\emptyset$ and X belong to the same chain in both partitions. Change one of these partitions slightly by transferring $\emptyset$ to a suitable chain to obtain two orthogonal partitions.)

6. Consider a partition of $\mathcal{P}(X)$ into symmetric chains and sum the cardinalities of the largest sets in the chains. Show that the sum is $\frac{n-1}{2}\binom{n}{\lfloor n/2 \rfloor} + 2^{n-1}$.

7. Let $p = 2n$, $X = [n]$, $Z = [p]$, $Y = Z \setminus X = \{n+1, \ldots, p\}$, $m = \binom{n}{\lfloor n/2 \rfloor}$ and let $\mathcal{P}(X) = \sum_{i=1}^{m} C_i$ and $\mathcal{P}(Y) = \sum_{i=1}^{m} C_i'$ be partitions into symmetric chains. For a chain $C_i = \{A_k, A_{k+1}, \ldots, A_{n-k}\}$, $A_k \subset A_{k+1} \subset \ldots \subset A_{n-k}$, $|A_j| = j$, let α_i be a sequence *string* of length $n-k$ whose first k terms form A_k and for $j > k$ the jth term is x_j where $A_j \setminus A_{j-1} = \{x_j\}$; define β_i similarly for C_i'. Let $\overline{\alpha}_i$ be the string α_i written in the reverse order and let γ be the following concatenation of strings:

$$\gamma = \overline{\alpha}_1\beta_1\overline{\alpha}_1\beta_2 \ldots \overline{\alpha}_1\beta_m\overline{\alpha}_2\beta_1\overline{\alpha}_2\beta_2 \ldots \overline{\alpha}_2\beta_m \ldots$$

$$\ldots \overline{\alpha}_m\beta_1\overline{\alpha}_m\beta_2 \ldots \overline{\alpha}_m\beta_m.$$

Show that every subset of $Z = [p]$ is the set of some consecutive terms of γ. Make use of the result of Exercise 6 to prove that γ has length

$$(n-1)m^2 + 2^n m \sim \frac{2}{\pi}2^{2n} = \frac{2}{\pi}2^p.$$

(Lipski (1978))

8. For $n = 3$ write out explicitly the string γ in Exercise 7.

§5. SHADOWS

Given a hypergraph $\mathcal{A} \subset X^{(r)}$, the *lower shadow* of $\mathcal{A}$ is

$$\partial_l(\mathcal{A}) = \{B \in X^{(r-1)} : B \subset A \quad \text{for some } A \in \mathcal{A}\},$$

and the *upper shadow* of $\mathcal{A}$ is

$$\partial_u(\mathcal{A}) = \{B \in X^{(r+1)} : B \supset A \quad \text{for some } A \in \mathcal{A}\}.$$

Usually we shall consider only the lower shadow, so when there is no danger of confusion, we shall write ∂ for ∂_l. If $r = 0$ and $\emptyset \neq \mathcal{A} \subset X^{(r)} = X^{(0)} = \{\emptyset\}$ then $\mathcal{A} = \{\emptyset\}$ so $\partial_l(\mathcal{A}) = \emptyset$ and $\partial_u(\mathcal{A}) = X^{(1)}$, and if $r = n$ and $\emptyset \neq \mathcal{A} \subset X^{(r)} = X^{(n)} = \{X\}$ then $\mathcal{A} = \{X\}$ so $\partial_l(\mathcal{A}) = X^{(n-1)}$ and $\partial_u(\mathcal{A}) = \emptyset$. Therefore, when considering shadows of a hypergraph $\mathcal{A}$, we shall assume that $\mathcal{A} \subset X^{(r)}$, $1 \leq r \leq n-1$.

The local LYM inequality (Theorem 3.3) states that

$$|\partial\mathcal{A}| = |\partial_l\mathcal{A}| \geq \frac{|\mathcal{A}|}{\binom{n}{r}}\binom{n}{r-1}. \tag{1}$$

Can we do better than (1)? Even more, what is the greatest lower bound for $|\partial\mathcal{A}|$, the cardinality of the lower shadow of a hypergraph $\mathcal{A} \subset X^{(r)}$, in terms of n, r and $|\mathcal{A}|$? This important problem was solved by Kruskal (1963). The result was rediscovered by Katona in 1968 and, in a more general form, by Clements and Lindström in 1969, and has come to be known as the Kruskal-Katona theorem. Our aim in this section is to prove this fundamental theorem.

Before launching into the theorem, we shall prepare the ground at a leisurely pace. After our preliminary work the proof itself will need very little effort.

Of the several natural orders on $X^{(r)}$, the set of all r-subsets of $X = [n] = \{1,2,\dots,n\}$, two tend to be particularly popular. Write $A, B, \dots \in X^{(r)}$ as $A = \{a_1, a_2, \dots, a_r\}$, $B = \{b_1, b_2, \dots, b_r\}, \dots$ with $a_1 < a_2 < \dots < a_r$, $b_1 < b_2 < \dots < b_r, \dots$. In the *lexicographic* (or simply *lex*) *order* $A < B$ if either $a_1 < b_1$, or $a_1 = b_1$ and $a_2 < b_2$, or $a_1 = b_1$, $a_2 = b_2$ and $a_3 < b_3$, or $\dots$ $a_1 = b_1, a_2 = b_2, \dots, a_{r-1} = b_{r-1}$ and $a_r < b_r$. In the *colexicographic* (or *colex*) *order* we have $A < B$ if $A \neq B$ and for $s = \max\{t : a_t \neq b_t\}$ we have $a_s < b_s$. Thus $A < B$ in the colex order if either $a_r < b_r$, or $a_r = b_r$ and $a_{r-1} < b_{r-1}$, or $a_r = b_r$, $a_{r-1} = b_{r-1}$ and $a_{r-2} < b_{r-2}$, or $\dots$ $a_r = b_r, a_{r-1} = b_{r-1}, \dots, a_2 = b_2$ and $a_1 < b_1$.

Note that $\{3,4,7,9\} < \{3,5,6,9\}$ in the lex order and $\{3,5,6,9\} < \{3,4,7,9\}$ in the colex order. It is immediate that both lex and colex are indeed *orders* on $X^{(r)}$, i.e. if $A, B, C \in X^{(r)}$ then precisely one of $A = B, A < B$ and $B < A$ holds, and if $A < B$ and $B < C$ then $A < C$. As customary, by $A \leq B$ we mean that $A < B$ or $A = B$.

Throughout the book we shall work with the colex order so an unspecified order on $X^{(r)}$ will always mean the colex order; as we shall see, this order is extremely useful in the study of set systems. Let us explore some properties of the colex order. Just as a matter of curiosity, what is the relationship between lex and colex orders? To get the colex order, take the lex order with respect to the reverse order on $[n] = \{1,2,\dots,n\}$, and reverse it. Thus, using the reverse order on $[9]$, in the lex order we have $\{9,7,4,3\} < \{9,6,5,3\}$ so, reversing it, we find that $\{3,5,6,9\} < \{3,4,7,9\}$ in the colex order. Let us list the first 10 elements of $[8]^{(3)}$ in increasing (colex) order: 123, 124, 134, 234, 125, 135, 235, 145, 245, 345. We try to keep the large elements as small as possible and increase them only when we have run out of other options. Thus the 11th, 12th and 13th elements of $[8]^{(3)}$ are 126, 136 and 236. What is the 14th element? We can get away without increasing 6 but we have to increase 3 to 4; having increased 3 we can take 1 instead of 2: 146. What is the set of the first $\binom{m_r}{r}$ elements of $X^{(r)}$? A set $A = \{a_1, \dots, a_r\}$ with $a_r \leq m_r$ is less than any set $B = \{b_1, \dots, b_r\}$ with $b_r \geq m_r + 1$. Since there are precisely $\binom{m_r}{r}$ sets $A = \{a_1, \dots, a_r\}$ with $a_r \leq m_r$, the set of the first $\binom{m_r}{r}$ elements of $X^{(r)}$ is precisely $[m_r]^{(r)}$. What comes after these sets? The r-sets ending in $m_r + 1$ (we are trying to keep the last element small!) with the $(r-1)$-sets of the other elements being the first few sets in the colex order.

To formalize this argument as a theorem, for $\mathcal{A} \subset \mathcal{P}(X)$ and $B \subset X$ set

$$\mathcal{A} + B = \{A \cup B : A \in \mathcal{A}\}$$

and if $B \subset A$ for all $A \in \mathcal{A}$, set

$$\mathcal{A} - B = \{A \setminus B : A \in \mathcal{A}\}.$$

Furthermore, for $0 \leq m_s < m_{s+1} < \dots < m_r$ let

$$\mathcal{B}^{(r)}(m_r, \dots, m_s) = \bigcup_{j=s}^{r} \left([m_j]^{(j)} + \{m_{j+1}+1, \dots, m_r+1\}\right) \subset X^{(r)}.$$

The condition $0 \leq m_s < m_{s+1} < \dots < m_r$ guarantees that we define a system of r-sets. Though the definition makes sense for all $0 \leq m_s < m_{s+1} < \dots < m_r$ usually we shall assume that $s \leq m_s$ since then the parameters $m_r, m_{r-1}, \dots, m_s$ are determined by the set system. Set

$$b^{(r)}(m_r, \dots, m_s) = \sum_{j=s}^{r} \binom{m_j}{j}$$

so that $|\mathcal{B}^{(r)}(m_r, \dots, m_s)| = b^{(r)}(m_r, \dots, m_s)$. Furthermore

$$\mathcal{B}^{(r)}(m_r, \dots, m_s) = [m_r]^{(r)} \cup \left(\mathcal{B}^{(r-1)}(m_{r-1}, \dots, m_s) + \{m_r + 1\}\right) \quad (1)$$

and so

$$b^{(r)}(m_r, \dots, m_s) = b^{(r)}(m_r) + b^{(r-1)}(m_{r-1}, \dots, m_s). \quad (2)$$

When enumerating the first m elements of $X^{(r)}$ in the colex order, the value $n = |X|$ has no effect. Indeed, the first $\binom{n-1}{r}$ elements of $[n]^{(r)}$ form precisely $[n-1]^{(r)}$ and for $m \geq n$ the restriction of the colex order on $[m]^{(r)}$ to $[n]^{(r)}$ is independent of m. Putting this slightly differently, the colex orders on $[r]^{(r)} \subset [r+1]^{(r)} \subset [r+1]^{(r)} \subset \dots$ are just initial segments of the colex order on $\mathbf{N}^{(r)}$, the set of all r-sets on $\mathbf{N}$. Note that this is false for the lex order. The lex order on $\mathbf{N}^{(3)}$ is 123, 124, 125, 126, …, 134, 135, 136, 137, …, 234, 235, 236, 237, …, etc., a not too useful enumeration of $\mathbf{N}^{(3)}$! In the lex order the $\binom{6}{3} = 20$th element of $[6]^{(3)}$ is $\{4,5,6\}$ but the 20th element of $[25]^{(3)}$ is $\{1,2,22\}$.

By now the alert reader may have realized that the binary representation of natural numbers can be used to define the colex order on $\mathbf{N}^{(r)}$ — better still, on $\mathbf{N}^{(<\omega)}$, the set of all finite subsets of $\mathbf{N}$. Map $\mathbf{N}^{(<\omega)}$ into $\mathbf{N}$ by sending a set $A = \{a_1, \dots, a_r\}$ into $\varphi(A) = \sum_{i=1}^{r} 2^{a_i}$ (and so the empty set $\emptyset$ into $\varphi(\emptyset) = 1$). The map φ sets up a one to one correspondence between $\mathbf{N}^{(<\omega)}$ and the subset $\{1\} \cup 2\mathbf{N}$ of $\mathbf{N}$; use this correspondence to transfer the order from $\mathbb{N}$ onto $\mathbb{N}^{(<\omega)}$. The restriction of this order on $\mathbb{N}^{(<\omega)}$ to $\mathbb{N}^{(r)}$ is precisely the colex order on $\mathbb{N}^{(r)}$: for $A, B \in \mathbb{N}^{(r)}$ we have $A < B$ iff $\varphi(A) < \varphi(B)$.

Theorem 1. *For $m \in \mathbb{N}$ the set of the first m elements of $\mathbb{N}^{(r)}$ in the colex order is $\mathcal{B}^{(r)}(m_r, \dots, m_s)$ where the numbers $m_s, m_{s+1}, \dots, m_r$ are the unique natural numbers such that $s \le m_s < m_{s+1} < \dots < m_r$ and*

$$m = b^{(r)}(m_r, \dots, m_s) = \sum_{j=s}^{r} \binom{m_j}{j}. \tag{3}$$

Proof. Let us see first by induction on m that $m_r, m_{r-1}, \dots, m_s$ exist and are determined by m. This is certainly true for $m = 1$: we must have $m_r = r$ and $s = r$. Suppose then that $m \ge 2$. Since for $k \ge r$ we have

$$\sum_{l=0}^{r-1} \binom{k-l}{r-l} = \binom{k+1}{r} - 1 < \binom{k+1}{r},$$

we must have

$$m_r = \max\left\{k : \binom{k}{r} \le m\right\}.$$

But then either $m = b^{(r)}(m_r) = \binom{m_r}{r}$ or else there are unique integers $s \le m_s < m_{s+1} < \dots < m_{r-1}$ such that

$$m - b^{(r)}(m_r) = b^{(r-1)}(m_{r-1}, m_{r-2}, \dots, m_s)$$

and so, by (2),

$$m = b^{(r)}(m_r, \dots, m_s),$$

with $m_r, m_{r-1}, \dots, m_s$ uniquely determined by m.

The argument above very nearly proves that $\mathcal{B}^{(r)}(m_r, \dots, m_s)$ is the set of the first n elements of $\mathbb{N}^{(r)}$. The set of the first $b^{(r)}(m_r) \le m$ elements is precisely $[m_r]^{(r)} = \mathcal{B}^{(r)}(m_r)$ and the next $m' = m - b^{(r)}(m_r) = b^{(r-1)}(m_{r-1}, \dots, m_s)$ elements all end in $m_r + 1$ and start in the first m' elements on $\mathbb{N}^{(r-1)}$. Hence, by induction on m (or on r, for that matter), these m' elements form $\mathcal{B}^{(r-1)}(m_{r-1}, \dots, m_s) + \{m_r + 1\}$. Relation (1) shows that the set of the first m elements of $\mathbb{N}^{(r)}$ is $\mathcal{B}^{(r)}(m_r, \dots, m_s)$. ∎

What is the mth element of $\mathbb{N}^{(r)}$? If $m = b^{(r)}(m_r, \dots, m_s)$, as in (3), then it is just the last element of $\mathcal{B}^{(r)}(m_r, \dots, m_s)$, so it is the last element of $[m_s]^{(s)}$ together with $m_{s+1} + 1, m_{s+2} + 1, \dots, m_r + 1$:

$\{m_s - s + 1, m_s - s + 2, \dots, m_s, m_{s+1} + 1, m_{s+2} + 1, \dots, m_r + 1\}$. Conversely, let $A = \{a_1, a_2, \dots, a_r\} \in \mathbb{N}^{(r)}$ with $a_1 < a_2 < \dots < a_r$. Define $s = \max\{t : 1 \le t \le r, a_t = a_{t-1} + 1\}$. Then

$$\{B \in \mathbb{N}^{(r)} : B \le A\} = \mathcal{B}^{(r)}(a_r - 1, a_{r-1} - 1, \dots, a_{s+1} - 1, a_s).$$

For example, what is the 1000th element of $\mathbb{N}^{(5)}$? We need $1000 = b^{(5)}(m_5, \dots)$. Now m_5 is the largest natural number satisfying $\binom{m_5}{5} \le 1000$ so $m_5 = 12$. As $1000 - \binom{12}{5} = 208 > 0$, m_4 is the largest natural number satisfying $\binom{m_4}{4} \le 208$ so $m_4 = 9$. Then we get $208 - \binom{9}{4} = 82$, giving $m_3 = 8$ and $82 - \binom{8}{3} = 26$. From this, $m_2 = 7$ and $26 - \binom{7}{2} = 5$. Finally, $m_1 = 5$. Therefore the first 1000 elements form $\mathcal{B}^{(5)}(12, 9, 8, 7, 5)$ and the 1000th element is $\{5, 8, 9, 10, 13\}$. What are then the 1001st, 1002nd and 1003rd elements? They are $\{6, 8, 9, 10, 13\}$, $\{7, 8, 9, 10, 13\}$ and $\{1, 2, 3, 11, 13\}$. Finally, what are the sets of the first 1001, 1002 and 1003 elements? By applying our rules, we see that they are $\mathcal{B}^{(5)}(12, 9, 8, 7, 6)$, $\mathcal{B}^{(5)}(12, 10)$ and $\mathcal{B}^{(5)}(12, 10, 3)$. Let us note another recursion formula for $b^{(r)}(m_r, \dots, m_s)$:

$$\begin{aligned} b^{(r)}(m_r, \dots, m_s) = {} & b^{(r)}(m_r - 1, \dots, m_s - 1) \\ & + b^{(r-1)}(m_r - 1, \dots, m_s - 1) \end{aligned} \tag{4}$$

This relation is immediate from

$$\binom{m_j - 1}{j} + \binom{m_j - 1}{j - 1} = \binom{m_j}{j},$$

but a more combinatorial argument goes as follows. For $\mathcal{A} \subset X^{(r)}$ set

$$\mathcal{A}_0 = \{A \in \mathcal{A} : 1 \notin A\}$$

and

$$\mathcal{A}_1 = \{A \in \mathcal{A} : 1 \in A\}.$$

Then $\mathcal{A}$ is partitioned as $\mathcal{A}_0 \cup \mathcal{A}_1$ and so

$$|\mathcal{A}| = |\mathcal{A}_0| + |\mathcal{A}_1| = |\mathcal{A}_0| + |\mathcal{A}_1 - \{1\}|. \tag{5}$$

Now if $\mathcal{A} = \mathcal{B}^{(r)}(m_r, \dots, m_s)$ then (5) turns into (4). Thus (4) is just a natural consequence of a natural partition of $\mathcal{B}^{(r)}(m_r, \dots, m_s)$.

Though we could dwell quite a bit longer on the fascinating properties of the colex order, it is time to consider shadows. What is the (lower) shadow of $\mathcal{B}^{(r)}(m_r, \ldots, m_s)$, where $s \le m_s < m_{s+1} < \ldots < m_r$? Clearly

$$\partial \mathcal{B}^{(r)}(m_r, \ldots, m_s) = \mathcal{B}^{(r-1)}(m_r, \ldots, m_s).$$

This shows that the shadow of a system of $b^{(r)}(m_r, \ldots, m_s)$ r-sets need not have more than $b^{(r-1)}(m_r, \ldots, m_s)$ elements. The celebrated Kruskal-Katona theorem claims that this example is best possible: the shadow of a system of $b^{(r)}(m_r, \ldots, m_s)$ r-sets (with $s \le m_s < m_{s+1} < \ldots < m_r$) contains at least $b^{(r-1)}(m_r, \ldots, m_s)$ $(r-1)$-sets. The original proofs have been greatly simplified over the years: thanks to Kleitman (1966b), Daykin, Godfrey and Hilton (1974), Daykin (1974) and Frankl (1984), there is a very elegant and simple proof, based on properties of the colex order and the compression operators $\tilde{R}_{ij}$.

For $i, j \in X$, $i \ne j$, and $A \subset X$ define

$$R_{ij}(A) = \begin{cases} (A \setminus \{j\}) \cup \{i\} & \text{if } i \notin A, j \in A \\ A & \text{otherwise.} \end{cases}$$

Thus R_{ij} replaces the element j by the element i whenever possible. By definition $|R_{ij}(A)| = |A|$ and if $i < j$ then $R_{ij}(A) \le A$ in the colex order. Note that R_{ij} is not a $1-1$ map: for $B \subset X \setminus \{i,j\}$ one has $R_{ij}(B \cup \{i\}) = R_{ij}(B \cup \{j\}) = B \cup \{i\}$. To see the action of R_{ij}, let

$$\mathcal{P}_{ij}(X) = \{A \subset X : i \in A, j \notin A\}.$$

Then R_{ij} gives a $1-1$ correspondence between $\mathcal{P}_{ji}(X)$ and $\mathcal{P}_{ij}(X)$, and on $\mathcal{P}(X) \setminus \mathcal{P}_{ji}(X)$ the operator R_{ij} is simply the identity. Let us associate with R_{ij} a map $\tilde{R}_{ij}$ sending a set system into another set system: for $\mathcal{A} \subset \mathcal{P}(X)$ let

$$\tilde{R}_{ij}(\mathcal{A}) = \{R_{ij}(A) : A \in \mathcal{A}\} \cup \{A : A, R_{ij}(A) \in \mathcal{A}\}.$$

We call $\tilde{R}_{ij}$ a *compression operator*. Clearly

$$|\tilde{R}_{ij}(\mathcal{A})| = |\mathcal{A}|$$

for all $\mathcal{A} \subset \mathcal{P}(X)$, and

$$\tilde{R}_{ij}(\mathcal{A}) \subset X^{(r)} \quad \text{if} \quad \mathcal{A} \subset X^{(r)}.$$

Call a set system $\mathcal{A} \subset \mathcal{P}(X)$ *left compressed* or simply *compressed* if $\tilde{R}_{ij}(\mathcal{A}) = \mathcal{A}$ whenever $1 \le i < j \le n$. Thus $\mathcal{A}$ is left compressed if $A \in \mathcal{A}$, $j \in A$, $i \notin A$ and $i < j$ imply $R_{ij}(A) = (A \setminus \{j\}) \cup \{i\} \in \mathcal{A}$. Equivalently, $\mathcal{A}$ is left compressed if whenever $A_1 = \{a_s, a_{s+1}, \ldots, a_t\} \subset A$, $a_s < a_{s+1} < \ldots < a_t$, $B_1 = \{b_s, b_{s+1}, \ldots, b_t\} \subset X \setminus A$, $b_s < b_{s+1} < \ldots < b_t$ and $b_s < a_s$, $b_{s+1} < a_{s+1}, \ldots, b_t < a_t$, the set $(A \setminus A_1) \cup B_1$ also belongs to $\mathcal{A}$.

At this stage one cannot help wondering whether a compressed set of m r-sets is not simply the set of the first m r-sets in the colex order. In fact, this cannot be true for every m because $R_{ij}(A) \le A$ for $i < j$ not only in the colex order but also in the lex order. For example, $\mathcal{A} = \{1\,2\,3,\ 1\,2\,4,\ 1\,2\,5,\ 1\,2\,6\}$ is compressed but it is rather far from $\mathcal{B}^{(3)}(4) = \{1\,2\,3,\ 1\,2\,4,\ 1\,3\,4,\ 2\,3\,4\}$, the set of the first 4 elements in the colex order.

For us the most important property of a compression operator is that it does not increase a shadow of the system.

Lemma 2. *(i) For $\mathcal{A} \subset X^{(r)}$ and $1 \le i < j \le n$ we have*

$$|\partial \mathcal{A}| \ge |\partial \tilde{R}_{ij}(\mathcal{A})|. \tag{6}$$

(ii) For all $\mathcal{A} \subset X^{(r)}$ there is a left compressed set system $\mathcal{A}' \subset X^{(r)}$ such that

$$|\mathcal{A}| = |\mathcal{A}'| \quad \text{and} \quad |\partial \mathcal{A}| \ge |\partial \mathcal{A}'|. \tag{7}$$

Proof. (i) Let $B \in \partial \tilde{R}_{ij}(\mathcal{A}) \setminus \partial \mathcal{A}$. Then $B \in \mathcal{P}_{ij}(X)$ and $R_{ji}(B) \in \partial \mathcal{A} \setminus \partial \tilde{R}_{ij}(\mathcal{A})$. Since $R_{ji} : \mathcal{P}_{ij}(X) \to \mathcal{P}_{ji}(X)$ is $1-1$,

$$|\partial \tilde{R}_{ij}(\mathcal{A}) \setminus \partial \mathcal{A}| = |R_{ji}(\partial \tilde{R}_{ij}(\mathcal{A}) \setminus \partial \mathcal{A})| \le |\partial \mathcal{A} \setminus \partial \tilde{R}_{ij}(\mathcal{A})|$$

so (6) does hold.

In words: (6) holds because every set B in $\partial \tilde{R}_{ij}(\mathcal{A}) \setminus \partial \mathcal{A}$ contains i but not j; replacing i by j we obtain a set $R_{ji}(B) \in \partial \mathcal{A} \setminus \partial \tilde{R}_{ij}(\mathcal{A})$.

(ii) Let us construct a sequence of systems $\mathcal{A}_0, \mathcal{A}_1, \ldots$ as follows. Set $\mathcal{A}_0 = \mathcal{A}$. Suppose we have constructed $\mathcal{A}_0, \mathcal{A}_1, \ldots \mathcal{A}_k$. If $\mathcal{A}_k$ is compressed then stop the sequence, otherwise pick an operator $\tilde{R}_{ij}$, $1 \le i < j \le n$, for which $\tilde{R}_{ij}(\mathcal{A}_k) \ne \mathcal{A}_k$ and set $\mathcal{A}_{k+1} = \tilde{R}_{ij}(\mathcal{A}_k)$. This sequence has to end in some system $\mathcal{A}_l$ since with

$$w(\mathcal{A}_i) = \sum_{A \in \mathcal{A}_i} \sum_{a \in A} a$$

we have $w(\mathcal{A}_0) > w(\mathcal{A}_1) > \ldots$. The system $\mathcal{A}' = \mathcal{A}_l$ is compressed and $|\mathcal{A}'| = |\mathcal{A}|$.

By (6) the cardinalities of the shadows are non-increasing since $\mathcal{A}_{k+1} = \tilde{R}_{ij}(\mathcal{A}_k)$:

$$|\partial \mathcal{A}_0| \geq |\partial \mathcal{A}_1| \geq \ldots \geq |\partial \mathcal{A}_l| = |\partial \mathcal{A}'|,$$

proving (7). ■

We are about to state and prove the Kruskal-Katona theorem. First let us define a function $\partial^{(r)} : \mathbf{N} \to \mathbf{N}$ as follows: for $m \geq 1$ take the unique representation of m guaranteed by Theorem 1:

$$m = b^{(r)}(m_r, \ldots, m_s)$$

and define

$$\partial^{(r)}(m) = b^{(r-1)}(m_r, \ldots, m_s). \tag{8}$$

Thus $\partial^{(r)}(m)$ is the number of $(r-1)$-sets in the shadow of the first m r-sets in the colex order.

Theorem 3. *Let $r \geq 1$ and $\mathcal{A} \subset X^{(r)}$. Then*

$$|\partial \mathcal{A}| \geq \partial^{(r)}(|\mathcal{A}|) \tag{9}$$

i.e. the shadow of $\mathcal{A}$ is at least as large as the shadow of the first $|\mathcal{A}|$ r-sets in the colex order. If $|\mathcal{A}| = \binom{m_r}{r}$ for some $m_r \geq r$ then equality holds in (9) *iff $\mathcal{A} \simeq [m_r]^{(r)}$.*

Proof. In proving (9), by Lemma 2(i) we may assume that $\mathcal{A}$ is compressed. If we could say that $\mathcal{A}$ is just the set of the first $|\mathcal{A}|$ elements then (9) would be proved. As it is, we have to work a little to prove (9).

We shall apply a double induction: first on r and then on $m = |\mathcal{A}|$, and we shall use the analogue of the splitting in (5) to deduce (9). If either $r = 1$ or $m = 1$ then the inequality is trivial so suppose $r \geq 2$ and $m \geq 2$ and the inequality holds for $r-1$ and all values of m and for r and $1, 2, \ldots, m-1$. Let

$$m = b^{(r)}(m_r, \ldots, m_s)$$

where $s \leq m_s < m_{s+1} < \ldots < m_r$ and define

$$\mathcal{A}_0 = \{A \in \mathcal{A} : 1 \notin A\}$$

and

$$\mathcal{A}_1 = \{A \in \mathcal{A} : 1 \in A\} = \mathcal{A} \setminus \mathcal{A}_0.$$

Then

$$\partial \mathcal{A}_0 \subset \mathcal{A} - \{1\} \tag{10}$$

since if $B \in \partial \mathcal{A}_0$ then $B \cup \{j\} \in \mathcal{A}_0$ for some $j > 1$ and so $R_{1j}(B \cup \{j\}) = B \cup \{1\} \in \mathcal{A}_1$ and $B \in \mathcal{A}_1 - \{1\}$. By (10),

$$\left|\partial \mathcal{A}_0\right| \leq \left|\mathcal{A}_1 - \{1\}\right| = \left|\mathcal{A}_1\right| : \tag{11}$$

If $j \in A \in \mathcal{A}_1$ and $j > 1$ then

$$A \setminus \{j\} = \left((A \setminus \{1\}) \setminus \{j\}\right) \cup \{1\}$$

so $\partial \mathcal{A}_1$ has the following partition:

$$\partial \mathcal{A}_1 = (\mathcal{A}_1 - \{1\}) \cup \left(\partial(\mathcal{A}_1 - \{1\}) + \{1\}\right). \tag{12}$$

If $\mathcal{A}$ is $\mathcal{B}^{(r)}(m_r, \ldots, m_s)$, namely the set system we hope is extremal, then $|\mathcal{A}_1| = b^{(r-1)}(m_r-1, \ldots, m_s-1)$ and $|\mathcal{A}_0| = b^{(r)}(m_r-1, \ldots, m_s-1)$, i.e. we get the splitting corresponding to (4). Thus it is natural to distinguish two cases according to the relationship between $|\mathcal{A}_1|$ and $b^{(r-1)}(m_r-1, \ldots, m_s-1)$.

Suppose that $|\mathcal{A}_1| < b^{(r-1)}(m_r-1, \ldots, m_s-1)$. Then, by (4),

$$|\mathcal{A}_0| > b^{(r)}(m_r-1, \ldots, m_s-1)$$

so, by our induction hypothesis,

$$|\partial \mathcal{A}_0| \geq b^{(r-1)}(m_r-1, \ldots, m_s-1),$$

contradicting (11). Therefore $|\mathcal{A}_1| \geq b^{(r-1)}(m_r-1, \ldots, m_s-1)$.

By (12) and the induction hypothesis applied to $\partial(\mathcal{A}_1 - \{1\}) \subset X^{(r-1)}$, we find that

$$\begin{aligned}|\partial \mathcal{A}_1| &\geq |\mathcal{A}_1| + \partial^{(r-1)}(|\mathcal{A}_1|)\\ &\geq b^{(r-1)}(m_r-1, \ldots, m_s-1) + b^{(r-2)}(m_r-1, \ldots, m_s-1)\\ &= b^{(r-1)}(m_r, \ldots, m_s),\end{aligned}$$

where in the last step we used (4). Hence

$$|\partial \mathcal{A}| \geq |\partial \mathcal{A}_1| \geq b^{(r-1)}(m_r, \ldots, m_s) = \partial^{(r)}(|\mathcal{A}|),$$

proving (9).

The second assertion is easily read out of the proof above. If we have equality in (9) for some $\mathcal{A}$ with $|\mathcal{A}| = m = \binom{m_r}{r}$, $m_r > r$, then $|\mathcal{A}_1| = b^{(r-1)}(m_r - 1)$ and $|\partial(\mathcal{A}_1 - \{1\})| = b^{(r-2)}(m_r - 1)$. Therefore, by the induction hypothesis applied to $\mathcal{A}_1 - \{1\}$, we see that $\mathcal{A}_1$ is the set of all $\binom{m_r-1}{r-1}$ r-sets containing 1 and contained in $[m_r]$. Then the shadow of $\mathcal{A}_1$ is $[m_r]^{(r-1)}$ and as $\partial\mathcal{A} = \partial\mathcal{A}_1$, we must have $\mathcal{A} = [m_r]^{(r)}$.

It seems that we are done, but not quite, since $\mathcal{A}$ is not our original set system but one obtained from the original by applying a sequence of compression operators. However, to overcome this difficulty, all we have to note is that if $\tilde{R}_{ij}(\mathcal{A}) = [Y]^{(r)}$ for some $Y \subset X$, $|Y| = m_r$, $i \in Y$ and $j \notin Y$, and $\mathcal{A} \neq [Y]^{(r)}$ then $\mathcal{A} = \tilde{R}_{ji}([Y]^{(r)}) = [R_{ji}(Y)]^{(r)}$. Hence $[m_r]^{(r)}$ was obtained from some system $[Y_1]^{(r)}$, that arose from some system $[Y_2]^{(r)}$, etc., where $|Y_1| = |Y_2| = \ldots = m_r$. Thus the system we started out with has to be $[Y]^{(r)}$ for some $Y \subset X$, $|Y| = m_r$, i.e. it is isomorphic to $[m_r]^{(r)} = \mathcal{B}^{(r)}(m_r)$, as claimed. ∎

According to the Kruskal-Katona theorem, the shadow of m r-sets cannot be smaller than the shadow of the first m r-sets in the colex order. In addition, the shadow of the first m r-sets in the colex order consists of precisely the first $\partial^{(r)}(m)$ $(r-1)$-sets in the colex order. This very fortunate phenomenon enables one to extend the LYM inequality (Theorem 3.2) to a characterization of the parameters of a Sperner family. This characterization was noted by Clements (1973) and by Daykin, Godfrey and Hilton (1974).

Theorem 4. *Let $f_0, f_1, \ldots, f_n$ be a sequence of non-negative integers. There is a Sperner family $\mathcal{F} \subset \mathcal{P}(X)$ such that $f_i = |\mathcal{F}_i| = |\mathcal{F} \cap X^{(i)}|$, $i = 0, 1, \ldots, n$, iff*

$$g_n = f_n \le \binom{n}{n}$$

$$g_{n-1} = \partial^{(n)}(g_n) + f_{n-1} \le \binom{n}{n-1},$$

$$g_{n-2} = \partial^{(n-1)}(g_{n-1}) + f_{n-2} \le \binom{n}{n-2},$$

$$\vdots$$

$$g_0 = \partial^1(g_1) + f_0 \le \binom{n}{0}.$$

Proof. Suppose $\mathcal{F}$ is a Sperner family and $f_i = |\mathcal{F}_i|$, $i = 1, 2, \ldots, n$. Define $\mathcal{H}_n \subset X^{(n)}, \mathcal{H}_{n-1} \subset X^{(n-1)}, \ldots$ by setting $\mathcal{H}_n = \mathcal{F}_n$ and

$$\mathcal{H}_j = \partial(\mathcal{H}_{j+1}) \cup \mathcal{F}_j, \qquad 0 \le j \le n-1.$$

Note that

$$\mathcal{H}_{j+1} = \partial^{n-(j+1)}\mathcal{F}_n \cup \partial^{(n-1)-(j+1)}\mathcal{F}_{n-1} \cup \ldots \cup \mathcal{F}_{j+1}$$

so

$$\partial(\mathcal{H}_{j+1}) \cap \mathcal{F}_j = \emptyset$$

and

$$|\partial(\mathcal{H}_{j+1})| + |\mathcal{F}_j| \le \binom{n}{j}.$$

Furthermore, with $h_j = |\mathcal{H}_j|$, by the Kruskal-Katona theorem,

$$h_n = g_n$$

and

$$h_j \ge \partial^{(j+1)}(h_{j+1}) + f_j, \qquad 0 \le j \le n-1.$$

Hence

$$g_j \le h_j \le \binom{n}{j}$$

for $j = n, n-1, \ldots, 0$, as claimed.

Conversely, suppose $g_0, g_1, \ldots, g_n$ satisfy the conditions. Denote by $\mathcal{C}^{(r)}(m)$ the set of the first m elements of $X^{(r)}$ in the colex order and set

$$\mathcal{G}_n = \mathcal{C}^{(n)}(g_n)$$

and

$$\mathcal{G}_j = \mathcal{C}^{(j)}(g_j) \setminus \mathcal{C}^{(j)}(\partial^{(j+1)}(g_{j+1})), \qquad 0 \le j \le n-1.$$

Since

$$\partial(\mathcal{C}^{(j+1)}(g_{j+1})) = \mathcal{C}^{(j)}(\partial^{(j+1)}(g_{j+1})),$$

the family $\mathcal{G} = \bigcup_{j=0}^{n} \mathcal{G}_j$ is a Sperner family, $|\mathcal{G}_n| = g_n = f_n$ and $|\mathcal{G}_j| = g_j - \partial^{(j+1)}(g_{j+1}) = f_j$, $0 \le j \le n-1$. ∎

To conclude this section, let us see how Theorem 3 can be used to give a best possible lower bound for the upper shadow of a set system $\mathcal{A} \subset X^{(r)}$.

Theorem 5. *Let $1 \le r \le n-1$, $\mathcal{A} \subset X^{(r)}$ and let $\mathcal{B}$ be the set of the last $|\mathcal{A}|$ elements of $X^{(r)}$ in the colex order. Then*

$$|\partial_u \mathcal{A}| \ge |\partial_u \mathcal{B}|$$

Proof. For $\mathcal{F} \subset \mathcal{P}(X)$ let $\mathcal{F}^c = \{F^c : F \in \mathcal{F}\}$ be the family of complements of sets in $\mathcal{F}$. Then $\partial_u \mathcal{A} = \left(\partial_l(\mathcal{A}^c)\right)^c$ so

$$|\partial_u \mathcal{A}| = |\partial_l(\mathcal{A}^c)| \ge |\partial_l(\mathcal{C})| = |\partial_u(\mathcal{C}^c)|$$

where $\mathcal{C}$ is the set of the first $|\mathcal{A}|$ elements of $X^{(n-r)}$ in the colex order. Note that $\mathcal{C}^c$ is precisely the set of the last $|\mathcal{A}|$ elements of $X^{(r)}$ in the colex order so $\mathcal{C}^c = \mathcal{B}$. ■

Exercises

1. What is the value of $b^{(4)}(14,6,5,2)$? Describe the set $\mathcal{B}^{(4)}(14,6,4,2)$!

2. What is the 1001st element of $\mathbf{N}^{(4)}$? What are the 999th, 1000th, 1002nd, 1003rd and 1004th elements?

3. Check that if $\mathcal{A} \subset \mathbf{N}^{(2)}$, $|\mathcal{A}| = 4$ then $|\partial\mathcal{A}| \ge 4$. How many non-isomorphic systems $\mathcal{A}$ are there with $|\partial\mathcal{A}| = 4$?

4. Show that if $m = b^{(r)}(m_r, \dots, m_2, 1)$ where $r \ge 2$ and $m_2 \ge 3$ then there are several non-isomorphic families $\mathcal{A} \subset \mathbf{N}^{(r)}$ such that $|\mathcal{A}| = m$ and $|\partial\mathcal{A}|$ is as small as possible, namely $b^{(r-1)}(m_r, \dots, m_2)$.

5. Prove that if $r \ge 3$, $\mathcal{A} \subset \mathbf{N}^{(r)}$, $|\mathcal{A}| = b^{(r)}(m_r, m_{r-1})$ where $r \le m_{r-1} < m_r$, and $|\partial\mathcal{A}| \le b^{(r-1)}(m_r, m_{r-1})$ then $\mathcal{A} \cong \mathcal{B}^{(r)}(m_r, m_{r-1})$.

6. Deduce from Theorem 4 the following result of Daykin, Godfrey and Hilton (1974). If there is a Sperner system $\mathcal{F}$ with parameters $f_i = |\mathcal{F}_i| = |\mathcal{F} \cap X^{(i)}|$ then there is a Sperner system $\mathcal{F}'$ with parameters f_i' such that $f_i' = 0$ if $i > n/2$, $f'_{n/2} = f_{n/2}$ and $f_i' = f_i + f_{n-i}$ for $i < n$.

§6. RANDOM SETS

This section touches on a rather large subject, the theory of random graphs. We shall hardly do more than define some of the terms, prove a consequence of the Kruskal-Katona theorem and note another result about random sets. When talking about random graphs, we shall assume that the reader has encountered the basic concepts of graph theory like connectedness, complete graph, cycle, path and diameter. The reader unfamiliar with these concepts should just skip the remarks about graphs and pass on to random sets. As we wish to keep our convention that the ground set X has n elements, our notation concerning random graphs will be unconventional. For an extensive account of the theory of random graphs the reader should consult Bollobás (1985).

Let $V = \{x_1, x_2, \dots, x_t\}$ and let $\mathcal{G}^t$ be the set of all graphs on V. Consider the subset $\mathcal{G}(t,k)$ of $\mathcal{G}^t$ consisting of graphs with $k = k(t)$ edges. Setting $n = \binom{t}{2}$ we see that $\mathcal{G}(t,k)$ contains precisely $\binom{n}{k}$ graphs. Turn $\mathcal{G}(t,k)$ into a ***probability space*** by giving all its elements the same probability, and write $G_{t,k}$ for a random element of $\mathcal{G}(t,k)$; we call $G_{t,k}$ a ***random graph of order t and size k***. The ***probability*** in the space $\mathcal{G}(t,k)$ is denoted by $P_{t,k}$. A ***property*** Q of graphs is naturally identified with the set of graphs on V having Q. Then $Q_k = Q \cap \mathcal{G}(t,k)$ is the set of graphs in $\mathcal{G}(t,k)$ having property Q. Given a property Q of graphs, $P_{t,k}(Q) = P_{t,k}(Q_k) = |Q_k|/\binom{n}{k}$ is the probability that $G_{t,k}$ has Q. Most graph properties one considers are *monotone*: if $G, H \in \mathcal{G}^t$ and $G \in Q$ then either $H \supset G$ implies that $H \in Q$ (and the property Q is then *monotone increasing*) or else $H \subset G$ implies $H \in Q$ (and the property Q is then *monotone decreasing*). A property Q is monotone increasing iff its *negation*, $\neg Q = Q^c = \mathcal{G} \setminus Q$, is monotone decreasing. The properties of containing a triangle, being connected, having diameter at most d, containing a Hamiltonian cycle and having chromatic number at least s are all monotone increasing properties. It is easy to see (and intuitively obvious) that if Q is a monotone increasing property then $P_{t,k}(Q)$ is a