Shifting shadows: the Kruskal–Katona Theorem

Shagnik Das

Introductory remarks

As we have seen[1], antichains and intersecting families are fundamental to Extremal Set Theory. The two central theorems, Sperner's Theorem and the Erdős–Ko–Rado Theorem, have inspired decades of research since their discovery[2], helping establish Extremal Set Theory as a vibrant and rapidly growing area of Discrete Mathematics.

One must, then, pay a greater than usual[3] amount of respect to the Kruskal–Katona Theorem, as it builds on both Sperner's Theorem and the Erdős–Ko–Rado Theorem. Indeed, as you will prove for your homework assignment, the Kruskal–Katona Theorem provides a precise and refined characterisation of the number of sets of different sizes that an antichain can contain. On the other hand, the Erdős–Ko–Rado Theorem follows almost immediately as a consequence of the Kruskal–Katona Theorem.

There is no doubt, then, that the Kruskal–Katona Theorem is truly a gem of Extremal Set Theory. In this note we introduce the theorem and give a short proof via the shifting technique, a powerful method that can be used to prove many results in this field.

The statement of the theorem(s)

Shadows The Kruskal–Katona Theorem concerns the size of shadows of set families, so in order to state it, we should first define what the shadow of a family is.

Definition 1 (Shadow)**.** *Given a k-uniform set family $\mathcal{F}$, the shadow $\partial \mathcal{F}$ is defined as*

$$\partial \mathcal{F} = \{ E : E = F \setminus \{x\} \textit{ for some } F \in \mathcal{F} \textit{ and } x \in F \} .$$

The shadow $\partial \mathcal{F}$ is therefore the $(k-1)$-uniform set family consisting of all $(k-1)$-sized subsets of sets in $\mathcal{F}$. Note that the shadow is intrinsic to $\mathcal{F}$, and does not depend on the underlying set over which $\mathcal{F}$ is defined.

Extremal problem Now that we know what the shadow is, we[4] wish to know more about it. The most fundamental property of a shadow is its size — how many sets can it contain? Clearly, the size of the shadow $\partial \mathcal{F}$ will depend on the size of $\mathcal{F}$: if $\mathcal{F} \subseteq \mathcal{F}'$, then $\partial \mathcal{F} \subseteq \partial \mathcal{F}'$. A reasonable question, then, is to fix the size of $\mathcal{F}$, and then ask how small/large the shadow of $\mathcal{F}$ can be.

The maximisation problem turns out to not be so interesting: clearly every k-set contains k subsets of size $(k-1)$, so we immediately have the upper bound $|\partial \mathcal{F}| \leq k\, |\mathcal{F}|$. It is also not difficult to see that we get equality when, for instance, the sets in $\mathcal{F}$ are pairwise disjoint. We therefore turn our attention to the minimisation problem.

[1] Or, rather more accurately, as we have been told, for our temporal budget did not allow for an exhaustive review of all that comprises the field of Extremal Set Theory.

[2] And, fortunately for us, these wells show no signs of running dry.

[3] For one ought to respect every mathematical theorem, proposition and lemma.

[4] For we are creatures of immense and unsettling curiosity.

Question 2. *Given a k-uniform set family $\mathcal{F}$ of m sets, how small can $\partial\mathcal{F}$ be?*

Brief meditation should suggest that one would be best served by taking the sets over as few elements as possible. In particular, if $m = \binom{n}{k}$ for some n, it seems entirely reasonable for the smallest shadow to be attained by the family $\mathcal{F} = \binom{[n]}{k}$ (in which case $\partial\mathcal{F} = \binom{[n]}{k-1}$ has size $\binom{n}{k-1}$). Now Mathematics is strewn with problems where the reasonable answer is completely wrong, but this is not one of them. The Kruskal–Katona Theorem proves that this intuitive solution is indeed correct, and in fact precisely determines the minimum size of the shadow for every k and m. However, as an exact result, its statement is a little involved, and requires some preparation.

Cascade notation In order to state the precise result, it is convenient to represent the number of sets, m, in a form related to our clique-based construction. This representation is known as *cascade notation*, where a natural number is represented as a sum of binomial coefficients.[5]

Definition 3 (k-cascade representation)**.** *Given natural numbers k and m, the* <u>*k-cascade representation*</u> *of m is*

$$m = \binom{a_k}{k} + \binom{a_{k-1}}{k-1} + \binom{a_{k-2}}{k-2} + \ldots + \binom{a_s}{s},$$

where $a_k > a_{k-1} > a_{k-2} > \ldots > a_s \geq s \geq 1$.

One can show, by means of a simple induction[6], that the k-cascade representation always exists, and is unique.

The theorem Having dealt with the preliminaries, we are now in position to state the Kruskal–Katona Theorem, which was proven independently by the two eponymous combinators.

Theorem 4 (Kruskal, 1963; Katona, 1968)**.** *Any k-uniform set family $\mathcal{F}$ of size $m = \binom{a_k}{k} + \binom{a_{k-1}}{k-1} + \ldots + \binom{a_s}{s}$, where $a_k > a_{k-1} > \ldots > a_s \geq s \geq 1$, must have*

$$|\partial\mathcal{F}| \geq \binom{a_k}{k-1} + \binom{a_{k-1}}{k-2} + \ldots + \binom{a_s}{s-1}.$$

The colexicographic order It should not surprise[7] you to learn that Theorem 4 is best possible for every choice of k and m — how could a theorem with such a precise statement be anything but? The optimal construction follows our earlier intuition by building these cliques one set at a time, but is best described by way of the colexicographic order on k-subsets of $\mathbb{N}$.

[5] One could think of the k-cascade notation as a binomial version of the binary expansion, which expresses a number as a sum of powers of two. While the binary expansion is well-suited for many purposes, the k-cascade notation is perfect for Kruskal–Katona.

[6] Choose a_k to be the largest natural number such that $m \geq \binom{a_k}{k}$, and if we do not have equality, induct with $k-1$ and $m - \binom{a_k}{k}$. The uniqueness requires a little more work, but not much – if one chooses a_{k-i} to be too small for some $i \geq 0$, then one cannot 'make up' for it with the later values.

[7] Not least of all because we mentioned this earlier.

Definition 5 (Colexicographic order)**.** *In the colexicographic order on* $\binom{\mathbb{N}}{k}$*, we have* $A < B$ *if and only if* $\max(A \Delta B) \in B$.

Informally, sets with larger elements come later in the colexicographic order. For instance, if $k = 3$, the first few sets in the colexicographic order are

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

The first m sets in the colexicographic order give equality in Theorem 4, showing that the bound cannot be improved. Indeed, if the k-cascade representation of m is $\binom{a_k}{k} + \binom{a_{k-1}}{k-1} + \ldots + \binom{a_s}{s}$, then $\mathcal{C}(m,k)$, the family of the first m sets in the colexicographic order on $\binom{\mathbb{N}}{k}$, is the union of all sets in $\binom{[a_k]}{k}$, the sets obtained by appending $a_k + 1$ to any set in $\binom{[a_{k-1}]}{k-1}$, the sets obtained by appending $\{a_{k-1}+1, a_k+1\}$ to any set in $\binom{[a_{k-2}]}{k-2}$, and so on, until we get the sets formed by appending $\{a_{s+1}+1, a_{s+2}+1, \ldots, a_{k-1}+1, a_k+1\}$ to any set in $\binom{[a_s]}{s}$. The shadow of this family is then all sets in $\binom{[a_k]}{k-1}$, the sets obtained by appending $a_k + 1$ to any set in $\binom{[a_{k-1}]}{k-2}$, the sets obtained by appending $\{a_{k-1}+1, a_k+1\}$ to any set in $\binom{[a_{k-2}]}{k-3}$, and so on, until we reach the sets formed by appending $\{a_{s+1}+1, a_{s+2}+1, \ldots, a_{k-1}+1, a_k+1\}$ to any set in $\binom{[a_s]}{s-1}$. Hence we find $|\partial \mathcal{C}| = \binom{a_k}{k-1} + \binom{a_{k-1}}{k-2} + \ldots + \binom{a_s}{s-1}$.

A more convenient truth As alluded to in the introductory remarks, the Kruskal–Katona Theorem is not only true but also useful.[8] However, when it comes to applications, it can be a trying task to carry out calculations with the k-cascade representation of numbers. Indeed, there is a good reason for our[9] having designed computers to work in binary rather than binomials.

Fortunately, Lovász found a slightly weaker, but computationally much friendlier, form of Theorem 4. To present this version, we extend the definition of the binomial coefficient to real numbers, setting, for any $x \in \mathbb{R}$ and $k \in \mathbb{N}$,

$$\binom{x}{k} = \frac{x(x-1)\ldots(x-k+1)}{k!}.$$

On the domain $x \in [k, \infty)$, the polynomial is strictly increasing with range $[1, \infty)$, and so it follows that for any $m \geq 1$, there is a unique x such that $\binom{x}{k} = m$.

Theorem 6 (Lovász, 1979)**.** *Any k-uniform set family $\mathcal{F}$ of size $m = \binom{x}{k}$, where $x \geq k$, must have*

$$|\partial \mathcal{F}| \geq \binom{x}{k-1}.$$

As one might imagine, this is much easier to work with from the computational point of view. Note that when $m = \binom{n}{k}$ for some integer n, Theorems 4 and 6 agree, and so this is also tight infinitely often. For $\binom{n}{k} < m < \binom{n+1}{k}$, the lower bound in Theorem 6 is slightly smaller than that in 4, but usually not by enough to disrupt any applications.

[8] While one can forgive a theorem for not being useful, especially since it is hard to judge what will be useful 593 years in the future, one should always be wary of a theorem that is not true.

[9] This 'our' is a very plural 'our,' referring to all mankind. I cannot in good conscience claim any personal credit for the design of computers.

Shifting

To prove Theorems 4 and 6, we shall make use of the shifting technique, a powerful method in Extremal Set Theory.[10] In this section we will define the shifting operation and collect a few useful facts regarding shifting and shadows.

The operation If we believe the theorem to be true, then we know that the initial segments of the colexicographic order, $\mathcal{C}(m,k)$, minimise the size of the shadow. As we have noted earlier, these initial segments consist of sets that avoid large elements. The heuristic behind the shifting operation is to try to make our family look more like $\mathcal{C}(m,k)$ by replacing large elements by smaller ones.

Definition 7 (Shift operator). *Let $\mathcal{F} \subset \binom{\mathbb{N}}{k}$ for some $k \geq 1$, and fix some $i \geq 2$. The shift operator $\mathcal{S}_i$ provides a new family $\mathcal{S}_i(\mathcal{F}) = \{S_i(F) : F \in \mathcal{F}\}$, where*

$$S_i(F) = \begin{cases} F \setminus \{i\} \cup \{1\} & \text{if } i \in F, 1 \notin F, \text{and } F \setminus \{i\} \cup \{1\} \notin \mathcal{F}, \\ F & \text{otherwise} \end{cases}.$$

In other words, for every set $F \in \mathcal{F}$, we try to replace the element i by the element 1. The only reasons why we may not be able to do this are if $i \notin F$, $1 \in F$, or the set we would obtain is already in the family $\mathcal{F}$. In the latter case, we say that F was *blocked* from shifting.

Key properties The first (almost immediate) fact we observe is that shifting families preserves their size.

Claim 8. *For any finite $\mathcal{F} \subset \binom{\mathbb{N}}{k}$ and any $i \geq 2$, $|\mathcal{S}_i(\mathcal{F})| = |\mathcal{F}|$.*

Proof. The only way this could fail to hold is if $S_i(F) = S_i(F')$ for two different sets $F, F' \in \mathcal{F}$. Clearly at least one of the sets has to shift for this to happen, so we may assume $F \neq S_i(F) = F \setminus \{i\} \cup \{1\}$. For F to have shifted, it cannot have been blocked, so we cannot have $F' = S_i(F)$. Thus $F' \neq S_i(F') = F' \setminus \{i\} \cup \{1\}$. However, we then have $F = S_i(F) \setminus \{1\} \cup \{i\} = S_i(F') \setminus \{1\} \cup \{i\} = F'$, a contradiction. Thus the map $S_i : \mathcal{F} \to \mathcal{S}_i(\mathcal{F})$ must in fact be a bijection, and so $|\mathcal{F}| = |\mathcal{S}_i(\mathcal{F})|$. □

The next claim shows that shifting behaves nicely with shadows. There is some sort of subcommutativity — the shadow of a shifted family is contained in the shift of the shadow.

Claim 9. *For any finite $\mathcal{F} \subset \binom{\mathbb{N}}{k}$ and any $i \geq 2$, $\partial \mathcal{S}_i(\mathcal{F}) \subseteq \mathcal{S}_i(\partial \mathcal{F})$.*

Proof. Suppose $E \in \partial \mathcal{S}_i(\mathcal{F})$, and so $E = S_i(F) \setminus \{x\}$ for some $F \in \mathcal{F}$ and $x \in S_i(F)$. A simple case analysis shows $E \in \mathcal{S}_i(\partial \mathcal{F})$, which implies the claim.

First suppose $1, i \notin S_i(F)$. Since $1 \notin S_i(F)$, we must have $S_i(F) = F$, and hence $E \subset F$. Thus $E \in \partial \mathcal{F}$, and as $i \notin E$, $S_i(E) = E$. Hence $E \in \mathcal{S}_i(\partial \mathcal{F})$, as desired.

[10] For instance, the Erdős–Ko–Rado Theorem can also be proven by shifting, although that proof is somewhat longer than the elegant proof of Katona we saw in lecture.

Now suppose $1, i \in S_i(F)$. As $i \in S_i(F)$, we have $S_i(F) = F$, and so $E \in \partial\mathcal{F}$ as before. If $x \neq 1$, then $1 \in E$, and so $E = S_i(E) \in \mathcal{S}_i(\mathcal{F})$. If $x = 1$, then $E' = E \setminus \{i\} \cup \{1\} \subset F$, and so $E' \in \partial\mathcal{F}$. This means E is blocked from shifting, and so $S_i(E) = E$, implying $E \in \mathcal{S}_i(\partial\mathcal{F})$.

For the third case, suppose $S_i(F) \cap \{1, i\} = \{i\}$. As $i \in S_i(F)$, we must have $S_i(F) = F$. However, as $i \in F$ and $1 \notin F$, F must have been blocked from shifting by $F' = F \setminus \{i\} \cup \{1\} \in \mathcal{F}$. Since $E \subset S_i(F) = F$, $E \in \partial\mathcal{F}$. If $x = i$, then $i \notin E$, and so $E = S_i(E) \in \mathcal{S}_i(\partial\mathcal{F})$. If $x \neq i$, then E would be blocked from shifting by $E' = F' \setminus \{x\} \in \partial\mathcal{F}$, and so $E = S_i(E) \in \mathcal{S}_i(\partial\mathcal{F})$ in this case as well.

The final case is when $S_i(F) \cap \{1, i\} = \{1\}$. Observe that $i \notin E$, and so $S_i(E) = E$. Hence if $E \in \partial\mathcal{F}$, $E = S_i(E) \in \mathcal{S}_i(\partial\mathcal{F})$, as required. If F did not shift, then $F = S_i(F)$ and $E \in \partial\mathcal{F}$. If F did shift, then $F = S_i(F) \setminus \{i\} \cup \{1\}$. If $x = 1$, then $E \subset F$, and so as before $E \in \partial\mathcal{F}$. If $x \neq 1$, let $E' = E \setminus \{1\} \cup \{i\}$, and observe that $E' \subset F$, and so $E' \in \partial\mathcal{F}$. Then either $E \in \partial\mathcal{F}$ as well, or E' is not blocked from shifting, and $E = S_i(E') \in \mathcal{S}_i(\mathcal{F})$. This completes the case analysis. ☐

Stable families Shifting thus seems to make things better, or, at the very least, not make them worse.[11] If a family cannot be shifted, we call it stable.

Definition 10 (Stable). *We say a set family $\mathcal{F}$ is <u>stable</u> if $\mathcal{S}_i(\mathcal{F}) = \mathcal{F}$ for every $i \geq 2$.*

The following claim then puts the results from Claims 8 and 9 together.

Claim 11. *For any finite family $\mathcal{F} \subset \binom{\mathbb{N}}{k}$, there is a stable family $\mathcal{G} \subset \binom{\mathbb{N}}{k}$ such that $|\mathcal{G}| = |\mathcal{F}|$ and $|\partial\mathcal{G}| \leq |\partial\mathcal{F}|$.*

Proof. If $\mathcal{F}$ is stable, we may take $\mathcal{G} = \mathcal{F}$. Otherwise there is some $i \geq 2$ for which $\mathcal{F}' = \mathcal{S}_i(\mathcal{F}) \neq \mathcal{F}$. By Claim 8, $|\mathcal{F}'| = |\mathcal{F}|$, and by Claim 9, $|\partial\mathcal{F}'| \leq |\mathcal{S}_i(\partial\mathcal{F})| = |\partial\mathcal{F}|$. We now repeat this process with $\mathcal{F}'$: if it is stable, we are done, and otherwise we can shift it again.

Every time we shift the family, we strictly increase the number of sets containing 1. This process must therefore terminate within $|\mathcal{F}|$ steps, at which point we would have the desired stable family. ☐

If the only stable families were the initial segments of the colexicographic order, $\mathcal{C}(m, k)$, then Theorem 4 would follow directly from Claim 11. Sadly[12], this is not the case. However, we can still derive enough structural information about the shadows of stable families to prove Theorems 4 and 6.

Given any set family $\mathcal{F} \subset \binom{\mathbb{N}}{k}$, we have a partition $\mathcal{F} = \mathcal{F}_0 \cup \mathcal{F}_1$, where $\mathcal{F}_0 = \{F \in \mathcal{F} : 1 \notin F\}$ and $\mathcal{F}_1 = \{F \in \mathcal{F} : 1 \in F\}$. Since 1 belongs to every set in $\mathcal{F}_1$, we can define the $(k-1)$-uniform family $\mathcal{F}_1' = \{F \setminus \{1\} : F \in \mathcal{F}_1\}$. The following claim shows that this set $\mathcal{F}_1'$ cannot be too small.

Claim 12. $|\mathcal{F}_1'| \geq |\partial\mathcal{F}_0|$.

[11] Much like with the Hippocratic Oath, the basic tenet of shifting is to 'first do no harm.'

[12] But not too sadly, for if this were true, then shifting could only be used for problems whose extremal families were of the form $\mathcal{C}(m, k)$, and would not be so widely useful.

Proof. In fact, we shall show $\partial\mathcal{F}_0 \subseteq \mathcal{F}_1'$. Indeed, suppose $E \in \partial\mathcal{F}_0$. Then we must have $E = F \setminus \{x\}$ for some $F \in \mathcal{F}_0$ and $x \in F$. As $F \in \mathcal{F}_0$, $x \geq 2$. Since $\mathcal{F}$ is stable, $\mathcal{S}_x(\mathcal{F}) = \mathcal{F}$, and thus $S_x(F) = F$. This means F was blocked from shifting, so $F' = F \setminus \{x\} \cup \{1\} \in \mathcal{F}$, and in particular is in $\mathcal{F}_1$. Hence $E = (F \setminus \{x\} \cup \{1\}) \setminus \{1\} \in \mathcal{F}_1'$, as required. □

Once we know $\mathcal{F}_1'$ is reasonably large, we will be in place to apply this next claim, which shows that $|\partial\mathcal{F}|$ is controlled by $\mathcal{F}_1'$.

Claim 13. $|\partial\mathcal{F}| = |\mathcal{F}_1'| + |\partial\mathcal{F}_1'|$.

Proof. We clearly have $\partial\mathcal{F} = \partial\mathcal{F}_0 \cup \partial\mathcal{F}_1$. In Claim 12, we saw $\partial\mathcal{F}_0 \subseteq \mathcal{F}_1'$. We will now show that $\partial\mathcal{F}_1$ consists of those sets in $\mathcal{F}_1'$ together with the sets obtained by appending $\{1\}$ to any set in $\partial\mathcal{F}_1'$. As sets in $\mathcal{F}_1'$ do not contain 1, these two kinds of sets are disjoint, and the claim follows.

That $\mathcal{F}_1' \subseteq \partial\mathcal{F}_1$ follows from its definition, as for every $F' \in \mathcal{F}_1'$, we have $F' = F \setminus \{1\}$ for some $F \in \mathcal{F}_1$.

The second kind of set is of the form $E = F' \setminus \{x\} \cup \{1\}$ for some $F' \in \mathcal{F}_1'$ and $x \in F'$. Then $E = (F' \cup \{1\}) \setminus \{x\}$, where $F = F' \cup \{1\} \in \mathcal{F}_1$ by definition of $\mathcal{F}_1'$. Hence $E \in \partial\mathcal{F}_1$, and thus every set in $\partial\mathcal{F}_1'$ gives a different set in $\partial\mathcal{F}_1$.

To see that every set in $\partial\mathcal{F}_1$ is covered, note that if $E = F \setminus \{x\}$ for some $F \in \mathcal{F}_1$ and $x \in F$, then if $x = 1$ we have $F \in \mathcal{F}_1'$. If $x \neq 1$, then $E = E' \cup \{1\}$, where $E' = F' \setminus \{x\}$, where $F' = F \setminus \{1\} \in \mathcal{F}_1'$. Hence E is obtained by adding 1 to the set $E' \in \partial\mathcal{F}_1'$. □

The proof

We now have all the ingredients required to prove Theorems 4 and 6. The proof we give here is a short and unified proof given by Frankl in 1984.[13] The proof is almost identical for both theorems, differing only in some calculations. Hence we shall present the proof for Theorem 4, and when different calculations are required for Theorem 6, we shall write those calculations in brackets and in blue text, [like this].

Proof of Theorem 4 [Theorem 6]. We prove the theorem by induction on k.

The base case of $k = 1$ is trivial. Indeed, the 1-cascade representation of m is $\binom{m}{1}$. Moreover, the shadow of any non-empty 1-uniform family $\mathcal{F}$ is $\partial\mathcal{F} = \{\emptyset\}$, and so we have $|\partial\mathcal{F}| = 1 = \binom{m}{0}$, as required in both theorems.

For the induction step, we may assume $k \geq 2$, and that the theorem holds for $(k-1)$-uniform families of any size. We prove the induction step by an inner induction, this time on m.

> For the base case of this inner induction, suppose $m = 1 = \binom{k}{k}$. $\mathcal{F}$ then consists of a single set F of size k, and so $\partial\mathcal{F}$ consists of all k subsets formed by removing any one element from F. Thus $|\partial\mathcal{F}| = k = \binom{k}{k-1}$, as required in both theorems.

[13] I have seen some sources attribute this to Lovász, but Frankl's paper is available here: `http://www.renyi.hu/~pfrankl/1984-2.pdf`.

We now proceed with the induction step of the inner induction, and hence we may assume the theorem holds for any k-uniform family of size smaller than m. Suppose $\mathcal{F}$ is a k-uniform set family of size m that minimises the size of $\partial\mathcal{F}$. By Claim 11, we may assume $\mathcal{F}$ is stable.

By Claim 13, $|\partial\mathcal{F}| = |\mathcal{F}_1'| + |\partial\mathcal{F}_1'|$. In order to obtain the required lower bound on $|\partial\mathcal{F}|$, we need $\mathcal{F}_1'$ to be large enough. This is guaranteed by our final claim.

Claim 14. $|\mathcal{F}_1'| \geq \binom{a_k-1}{k-1} + \binom{a_{k-1}-1}{k-2} + \ldots + \binom{a_s-1}{s-1}$. $\quad [|\mathcal{F}_1'| \geq \binom{x-1}{k-1}.]$

Let us first see how this claim completes the proof. Recall that $\mathcal{F}_1'$ is a $(k-1)$-uniform family, and hence by the outer induction (on k), we may lower bound the size of its shadow. Indeed, by the induction hypothesis, we have

$$|\partial\mathcal{F}_1'| \geq \binom{a_k-1}{k-2} + \binom{a_{k-1}-1}{k-3} + \ldots + \binom{a_s-1}{s-2}. \qquad \left[|\partial\mathcal{F}_1'| \geq \binom{x-1}{k-2}.\right]$$

(Here we set $\binom{a_1-1}{-1} = 0$ if $s = 1$.) Putting this together, Claim 13 gives

$$\begin{aligned}
|\partial\mathcal{F}| &= |\mathcal{F}_1'| + |\partial\mathcal{F}_1'| \qquad \left[\geq \binom{x-1}{k-1} + \binom{x-1}{k-2} = \binom{x}{k-1}.\right] \\
&\geq \left[\binom{a_k-1}{k-1} + \binom{a_{k-1}-1}{k-2} + \ldots + \binom{a_s-1}{s-1}\right] \\
&\quad + \left[\binom{a_k-1}{k-2} + \binom{a_{k-1}-1}{k-3} + \ldots + \binom{a_s-1}{s-2}\right] \\
&= \left[\binom{a_k-1}{k-1} + \binom{a_k-1}{k-2}\right] + \left[\binom{a_{k-1}-1}{k-2} + \binom{a_{k-1}-1}{k-3}\right] \\
&\quad + \ldots + \left[\binom{a_s-1}{s-1} + \binom{a_s-1}{s-2}\right] \\
&= \binom{a_k}{k-1} + \binom{a_{k-1}}{k-2} + \ldots + \binom{a_s}{s-1}.
\end{aligned}$$

This gives the desired lower bound on $|\partial\mathcal{F}|$, completing the induction step of the inner induction, modulo Claim 14.

Proof of Claim 14. Suppose the claim were false. As $m = |\mathcal{F}| = |\mathcal{F}_0| + |\mathcal{F}_1|$

and $|\mathcal{F}_1'| = |\mathcal{F}_1|$, we must have

$$\begin{aligned}
|\mathcal{F}_0| &= m - |\mathcal{F}_1'| && \left[> \binom{x}{k} - \binom{x-1}{k-1} = \binom{x-1}{k}.\right] \\
&> \left[\binom{a_k}{k} + \binom{a_{k-1}}{k-1} + \ldots + \binom{a_s}{s}\right] \\
&\quad - \left[\binom{a_k - 1}{k-1} + \binom{a_{k-1} - 1}{k-2} + \ldots + \binom{a_s - 1}{s-1}\right] \\
&= \left[\binom{a_k}{k} - \binom{a_k - 1}{k-1}\right] + \left[\binom{a_{k-1}}{k-1} - \binom{a_{k-1} - 1}{k-2}\right] \\
&\quad + \ldots + \left[\binom{a_s}{s} - \binom{a_s - 1}{s-1}\right] \\
&= \binom{a_k - 1}{k} + \binom{a_{k-1} - 1}{k-1} + \ldots + \binom{a_s - 1}{s}.
\end{aligned}$$

Recall Claim 12, which gives $|\mathcal{F}_1'| \geq |\partial \mathcal{F}_0|$. Since $\mathcal{F}_0$ is a k-uniform family of size smaller than m (since $\mathcal{F}$ is stable, $\mathcal{F}_1$ is non-empty), we can apply the inner induction hypothesis to lower bound $|\partial \mathcal{F}_0|$.[14] Thus we have

$$|\mathcal{F}_1'| \geq |\partial \mathcal{F}_0| \geq \binom{a_k - 1}{k-1} + \binom{a_{k-1} - 1}{k-2} + \ldots + \binom{a_s - 1}{s-1}, \quad \left[|\mathcal{F}_1'| \geq \binom{x-1}{k-1},\right]$$

contradicting the falsity of the claim. Hence $\mathcal{F}_1'$ must be as large as claimed, completing the proof. □

Thus the inner induction (on m) is proved, completing the outer induction (on k) and, with it, the proof of Theorem 4 [Theorem 6]. □

Concluding remarks

The idea behind this beautiful proof is that it is enough to consider the shadow cast by the sets containing 1 in a stable family. The proof is then double-inductive. We first show that $\mathcal{F}_1'$ cannot be small, using the induction on m to lower bound the size of $\partial \mathcal{F}_0$. Once we know $\mathcal{F}_1'$ is large, we can then show that the shadow of $\mathcal{F}$ has the desired size, using the induction on k to lower bound the size of $\partial \mathcal{F}_1'$.

As we noted earlier, had it been true that the only stable families are the initial segments of the colexicographic order $\mathcal{C}(m, k)$, then Claim 11 would already have completed the proof. While this is not true, one can proof Theorem 4 with shifting arguments alone. One needs to introduce *compressions*, which are more general shifting operators. It is then possible to compress a family repeatedly until it becomes $\mathcal{C}(m, k)$.

[14] The eagle-eyed reader may object here: "If $x \in [k, k+1)$, then $x - 1 < k$, in which case we cannot apply the induction hypothesis for the proof of Theorem 6!" A valid point, but you need not fear — in this case, the conclusion of Claim 14 is trivia. Indeed, since $|\mathcal{F}_0| > \binom{x-1}{k} \geq 0$, $\mathcal{F}_0$ is non-empty. Now observe that a single k-set in $\mathcal{F}_0$ contributes k sets to the shadow, and so, since $x - 1 \in [k-1, k)$, $|\partial \mathcal{F}_0| \geq k = \binom{k}{k-1} \geq \binom{x-1}{k-1}$, giving the desired lower bound without using the induction hypothesis.

However, one must be careful while doing this, as the analogue of Claim 9 does not hold in general — compressing a family could increase the size of its shadow. What one can show, though, is that if the compressions are applied in the correct order, then everything is well-behaved.

Finally, it is worth reiterating that the Kruskal–Katona Theorem is an incredibly useful theorem to have in one's arsenal. Aside from the extremal set theoretic applications you will encounter in the homework, the theorem is also used in Extremal Graph Theory, Discrete Geometry and Algebra, to name but three examples.

This concludes our exposition of this wonderful theorem; that's all there is, there isn't any more.