Different Approaches to Unimodality

1 Combinatorial Way

We know that giving combinatorial proofs of equalities $a = b$ involves setting up two sets A and B with cardinalities a and b, respectively, and then setting up a bijection between the two sets. Similarly, showing that $a \leq b$, we can set up two sets A and B with the respective cardinalities and give an injection from A to B. Here we try this approach for the standard binomial coefficients and the coefficients of the **Gaussian Polynomial**:

$$G(q) = \sum_i a_i q^i := \frac{(q^n - 1)(q^{n-1} - 1)\cdots(q-1)}{(q^{n-k} - 1)(q^{n-k-1} - 1)\cdots(q-1) \times (q^k - 1)(q^{k-1} - 1)\cdots(q-1)}. \tag{1}$$

Note that the right hand side is the q-binomial coefficient $\binom{n}{k}_q$ counting the number of vector spaces of dimension k in an n-dimensional vector space over a finite field with q elements. Here we are treating q as a variable.

For the combinatorial proofs, we will be working with finite posets and chains in them, so let's recall some definitions. Let $(P, \leq)$ be a poset. A **chain** in P is a totally ordered sub-poset of P; the length of a finite chain is one less than the size of the chain. A chain $x_0 \leq x_1 \leq \cdots \leq x_k$ is **saturated** if x_{i+1} covers x_i, that is, if there is no intermediate element that can be added to the chain preserving the total order. A **graded poset** of rank n is a poset in which all the maximal chains have the same length. In such a poset, we can associate a unique rank function $\rho : P \to \{0, \ldots, n\}$ such that $\rho(\mathbf{0}) = \mathbf{0}$ and $\rho(y) = \rho(x) + 1$, if y covers x; thus, rank of $\mathbf{0}$ is zero, of atoms one and so on. The length of a finite poset is the length of the longest chain in P. In a graded poset, a **saturated symmetric chain** $x_0 \leq x_1 \leq \cdots \leq x_k$ is one where the ranks of x_0 and x_k add up to the rank of the poset; e.g., the boolean lattice, the division lattice of a number. A **saturated symmetric chain decomposition** (SSCD) of a graded poset is a partitioning of P in terms of disjoint saturated symmetric chains. Not every graded poset has such a decomposition (see Figure 1). An **antichain** A is a subset of P such that any two distinct elements in A are incomparable.

We can now prove Sperner's Theorem: The size of the largest antichain in the boolean lattice B_n is $\binom{n}{n/2}$.

To establish unimodality of the boolean lattice B_n of n elements, we need to give an injection from a set S of size k to a set of size $(k+1)$, for $k < n/2$. Let $\binom{n}{k}$ also denote the set of all subsets of size k. By repeating this injuction, we will get a maximal chain that ends at the middle rank, and by symmetry (working with the complement sets) we can extend it to a symmetric chain. By recursively applying this with elements with increasing ranks, we can construct a SSCD of the boolean lattice. Therefore, any injuction yields a SSCD. The converse is also true, so we will try to construct a SSCD for the boolean lattice B_n.

One way is to do induction on n. Suppose we have a symmetric chain for B_{n-1}, say $C_r \subset C_{r+1} \subset \cdots \subset C_{n-1-r}$. Now we can do two things: keep the chain as is since it is also valid in B_n or add the remaining element $\{n\}$ to get $C_r \cup \{n\} \subset C_{r+1} \cup \{n\} \subset \cdots \subset C_{n-1-r} \cup \{n\}$. But are these chains symmetric? No! Because the ranks of all the elements in the first chain have not increased and in the second chain they have increased by one, therefore, sum of the ranks of the first and last element doesn't add to n in both cases. One fix is to take the last element and add it to the first chain to get the following two chains:

$$C_r \subset C_{r+1} \subset \cdots \subset C_{n-1-r} \cup \{n\} \tag{2}$$

and

$$C_r \cup \{n\} \subset C_{r+1} \cup \{n\} \subset \cdots \subset C_{n-2-r} \cup \{n\}. \tag{3}$$

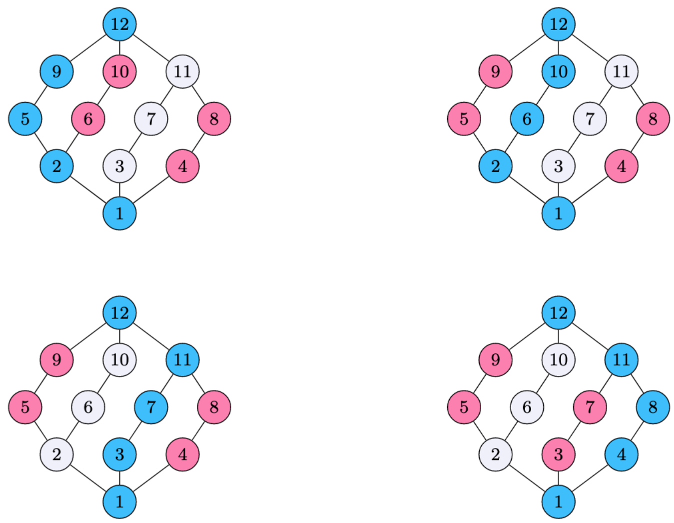

Figure 1: Graded poset with no SSCD

Therefore, every SSC in B_{n-1} gives us at most two chains in B_n (at most because when (2) has only one element, we get one chain). It is not hard to see that the set of chains form a SSCD of B_n.

The ith Whitney number of a graded poset is the number of elements with rank i. There is also the unimodality of these numbers. A poset is said to have the Sperner property if the the size of the largest antichain is the same as the largest Whitney number, or Whitney rank of the poset. There can be posets that violate the Sperner property, see Figure **??**.

Claim: A poset with SSCD has the Sperner property.

2 Linear Algebraic Paradigm

In the section above, we didn't explicitly give the injection. One can try to unravel it from the SSCD, but there is a linear algebraic way of coming up with the injection : Set up two vectors spaces $\mathcal{A}$ and $\mathcal{B}$ such that $\dim(\mathcal{A}) = a$ and $\dim(\mathcal{B}) = b$ and either show a linear transformation $T : \mathcal{A} \to \mathcal{B}$ that is injective, or equivalently give an injective map from a basis of $\mathcal{A}$ to a basis of $\mathcal{B}$. Let's see this approach for the case of binomial coefficients.

Since we want a vector space with dimension $\binom{n}{k}$, we consider the vector space V_k of all *formal sums* over subsets of size k of $[n]$, namely

$$\sum_{|S|=k:S\subseteq[n]} a_S S$$

where the scalars a_S are from some large enough field, say rationals. The product of the scalar-zero with any formal sum is the zero element of V_k. Note that V_k is not an inner-product space. The map that we

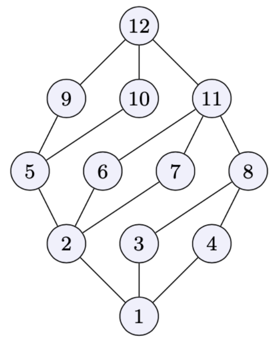

Figure 2: Graded Poset not satisfying sperner property

intend to show as an injection is the following [1]: For $S \in V_k$ define

$$H(S) := \sum_{j \notin S} S \cup j;$$

additionally, $H([n]) := 0$, since the sum is empty. Our main claim will be that $H : V_k \to V_{k+1}$ is injective as long as $k < n/2$. It is not easy to show because we have to argue that none of the formal sums in V_k are in the kernel of H, except zero. The map extends linearly to all formal sums i.e.,

$$H\left(\sum_{|S|=k : S \subseteq [n]} a_S S\right) = \sum_{|S|=k : S \subseteq [n]} a_S H(S).$$

In order to prove that H is injective, we need another companion operator $F : V_k \to V_{k-1}$ defined as [2]

$$F(S) := \sum_{j \in S} S \setminus j.$$

Just as with H, define $F\emptyset := 0$. Our first claim is the following:
Claim: The linear map $HF - FH : V_k \to V_k$ satisfies the following for any set S of size k

$$HF(S) - FH(S) = (2k - n)S =: \mu(k)S.$$

Proof. By linearity of the two operators, the sum

$$HF(S) = \sum_{j \in S} \sum_{i \notin S \setminus j} (S \setminus j) \cup i$$

and

$$FH(S) = \sum_{i \notin S} \sum_{j \in S \cup i} (S \cup i) \setminus j.$$

If i and j are different in the first summation then they cancel out with the corresponding term in the second summation. The only terms left uncanceled are when i and j are the same in both the sums. In the first sum, there are k such occurrences of S for each element in S and in the second sum there are $(n - k)$ such occurrences of S for each element not in S. Therefore, we get the right hand side. **Q.E.D.**

[1]Borrowing Zeilberger's analogy, H is for hiring-a-new-faculty.
[2]In Zeilberger's analogy, F is for firing-an-existing-faculty.

Next claim is about the commutativity of repeated firings with single hirings:
Claim: For $v \in V_k$, we have $(HF^r - F^rH)v = (\mu(k) + \cdots + \mu(k-r+1))F^{r-1}v$.
Proof. The proof is by induction over k. The term

$$\begin{aligned} HF^{r+1} - F^{r+1}H &= HF^{r+1} - F^rHF + F^rHF - F^{r+1}H \\ &= (HF^r - F^rH)F + F^r(HF - FH). \end{aligned}$$

So for $v \in V_k$ we have

$$\left(HF^{r+1} - F^{r+1}H\right)v = (HF^r - F^rH)Fv + F^r(HF - FH)v.$$

But note that $Fv \in V_{k-1}$, therefore, applying the induction hypothesis for $(k-1)$ to the first term on the right hand side and claim one to the second term we obtain

$$\begin{aligned} \left(HF^{r+1} - F^{r+1}H\right)v &= (\mu(k-1) + \cdots + \mu(k-r+1))F^{r-1}Fv + F^r\mu(k)v \\ &= (\mu(k) + \mu(k-1) + \cdots + \mu(k-r+1))F^rv \end{aligned}$$

as desired. **Q.E.D.**

Now we are ready to show the main claim that $H : V_k \to V_{k+1}$ is an injective map, that is, if $Hv = 0$ then $v = 0$ for $v \in V_k$. The proof is by contradiction. Suppose there is a non-zero $v \in V_k$ such that $Hv = 0$. For such a v and $r \geq 1$ we have from the second claim above that

$$HF^rv = (\mu(k) + \cdots + \mu(k-r+1))F^{r-1}v.$$

Applying H^{r-1} on both sides we get

$$H^rF^rv = (\mu(k) + \cdots + \mu(k-r+1))H^{r-1}F^{r-1}v,$$

and by repetition we obtain that

$$H^rF^rv = (\mu(k) + \cdots + \mu(k-r+1))(\mu(k) + \cdots + \mu(k-r+2))\cdots\mu(k)v.$$

Taking $r = k+1$ we obtain that the left hand size is zero, since $v \in V_k$, F^kv is a multiple of the empty-set and $F\emptyset = 0$ as the summation is an empty sum. However, the right hand side above is a non-zero multiple of v as long as $2k < n$, which gives us a contradiction. Therefore, $v = 0$ as desired.

Note that we showed that $H : V_K \to V_{k+1}$ is injective, but that doesn't obviously give us an *injection of sets from* $\binom{n}{k} \to \binom{n}{k+1}$. But this follows from a more general principle.

Lemma 1 *If $f : V \to W$ is an injective linear map between two vector spaces, where* $\dim(V) \leq \dim(W)$*, then there is a injective map between their basis as well.*

Proof. Let $v_1, \ldots, v_m$ (resptectively, $w_1, \ldots, w_n$) be the basis for V and W, then from f we can obtain an injective mapping of v_i to some w_j. Let A be the $n \times m$ matrix where the rows are indexed by w_j's and columns by v_i's. If $f(v_i) = \sum_{j=1}^n a_jw_j$ then the column entry of A is the vector $(a_1, \ldots, a_n)$, that is, the $n \times m$ matrix

$$[f(v_1)|f(v_2)|\cdots|f(v_m)] = [w_1|w_2|\cdots|w_n] \cdot A.$$

Since f is injective that matrix A is full rank, namely m. Therefore, it has an $m \times m$ minor, let's say, the top m rows, whose determinant doesn't vanish. In the Laplacian expansion of this minor into $m!$ terms, there must be one that doesn't vanish, say corresponding to a permutation π. Since the rows of A are indexed by w_j's, $\pi(j) \in [m]$, for $j \in [m]$. Therefore, match the vector $v_i \in V$ with the vector $w_{\pi^{-1}(i)}$ in W. This injective map can then be used to obtain a SSCD of B_n. **Q.E.D.**

2.1 A Proof Along Proctor's Line

Zeilberger's proof is a simplification of Proctor's proof. However, it misses giving the insight that Proctor's proof has: namely, highlighting the representation theoretic framerwork that is behind the definitions of the linear functions. The key idea is to "take advantage of one of nice features of linear algebra, the ability to change basis." This means that instead of showing that the kernel of the map H is trivial, it constructs a newer set of basis for each of the vector spaces V_k iteratively; the construction ensures some mirror symmetry around $n/2$. Moreover, Proctor's proof constructs a chain of disjoint vectors that form a basis for the complete space $V := \cup_k V_k$. Then using the idea in Lemma 1, we can construct a matching of vector spaces as desired.

Let P be a graded poset of rank r with levels $P_0, \ldots, P_r$. With each level P_i we associate the vector space $\widetilde{P}_i$ freely generated by the elements of P_i, that is,

$$\widetilde{P}_i := \left\{ \sum_{x \in P_i} a_x x \right\}, \tag{4}$$

where a_x are scalars from some large field[3]. Let $\widetilde{P}$ be the union of these vector spaces; it is, therefore, a graded vector space. Just as was done earlier, we will associate three functions with $\widetilde{P}$, similar to H and F, and a third one called T (to extend Zielberger's analogy, it will be "transfer").

We will define the operators on elements of P_i, which form a basis for $\widetilde{P}_i$, and lift them by linearity to $\widetilde{P}_i$. An order raising operator H is such that $H\widetilde{P}_i \subset \widetilde{P}_{i+1}$. In our case it is given as follows: for $x \in P_i$

$$H(x) := \sum_{y \text{ cover } x} y, \tag{5}$$

and extended linearly to $\widetilde{P}_i$

$$H(\sum_{x \in P_i} a_x x) = \sum_{x \in P_i} a_x H(x).$$

The transfer operator $T : \widetilde{P}_i \to \widetilde{P}_i$ is given as for $x \in P_i$

$$T(x) := (2i - r)x, \tag{6}$$

and the order lowering operator $F : \widetilde{P}_{i+1} \to \widetilde{P}_i$ is defined for $y \in P_{i+1}$

$$F(y) := \sum_{y \text{ covers } x} c(x, y)x, \tag{7}$$

where $c(x, y)$ will depend on some properties of x and y. Just as for H, both the operators F and T can be extended linearly. Let's illustrate the three functions in detail for a special case

2.2 Tilted Shading Problem

We now see the proofs of the two problems that Proctor mentions in his paper. The first one, as mentioned in the heading, was settled by Sylvester in 1878 "by aid of a construction drawsn from the resources of Imagniative reason," in his words.

The problem is as follows: we are given a "tilted grid", as shown in Figure **??**, of $m \times n$ tiles. We have to tile the grid, however, a legal tiling is one, where the tiles don't slide down Let a_1 be the number of squares shaded in the top-right tilted row, a_2 in the row below it, and so on a_n in the bottom most tilted row. Then the legality of shading implies that

$$0 \le a_1 \le a_2 \le \cdots \le a_n$$

where $0 \le a_i \le m$. Let $L(m, n)$ be the ranked poset of valid shadings where $\boldsymbol{a} \le \boldsymbol{b}$ iff for all $i = 1, \ldots, n$, $a_i \le b_i$, and the rank is $\sum_i a_i$, that is the number of tiles shaded. Since all the valid shadings are mutually incomparable, the ith Whitney number is the possible number of valid shadings using i tiles. Let L_i denote

[3]We can also take it to be a ring, and then work with the module over the ring.

the ith level and $\widetilde{L}_i$ the corresponding vector space. Then the functions H and T are as described above. For $\boldsymbol{b} \in L_i$, the function T is given as

$$T(\boldsymbol{b}) = \sum_{\boldsymbol{a} \text{ covered by } \boldsymbol{b}} c(\boldsymbol{a}, \boldsymbol{b})\boldsymbol{a}$$

where, if $I = I(\boldsymbol{a}, \boldsymbol{b})$ is the index where $\boldsymbol{b}$ and $\boldsymbol{a}$ differ then

$$c(\boldsymbol{a}, \boldsymbol{b}) := (m + n - a_I - I)(a_I + I).$$

Let $[A, B] := AB - BA$ for two linear operators. The three fundamental relations relating these functions are:

1. $[T, H] = 2H$
2. $[T, F] = -2F$, and
3. $[H, F] = T$.