4 The Sperner property.

In this section we consider a surprising application of certain adjacency matrices to some problems in extremal set theory. An important role will also be played by finite groups. In general, extremal set theory is concerned with finding (or estimating) the most or least number of sets satisfying given set-theoretic or combinatorial conditions. For example, a typical easy problem in extremal set theory is the following: What is the most number of subsets of an n-element set with the property that any two of them intersect? (Can you solve this problem?) The problems to be considered here are most conveniently formulated in terms of partially ordered sets, or posets for short. Thus we begin with discussing some basic notions concerning posets.

4.1 Definition. A *poset* (short for partially ordered set) P is a finite set, also denoted P, together with a binary relation denoted $\leq$ satisfying the following axioms:

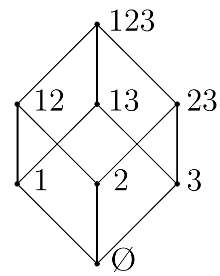
(P1) (reflexivity) $x \leq x$ for all $x \in P$

(P2) (antisymmetry) If $x \leq y$ and $y \leq x$, then $x = y$.

(P3) (transitivity) If $x \leq y$ and $y \leq z$, then $x \leq z$.

One easy way to obtain a poset is the following. Let P be any collection of sets. If $x, y \in P$, then define $x \leq y$ in P if $x \subseteq y$ as sets. It is easy to see that this definition of $\leq$ makes P into a poset. If P consists of *all* subsets of an n-element set S, then P is called a (finite) *boolean algebra* of *rank* n and is denoted by B_S. If $S = \{1, 2, \ldots, n\}$, then we denote B_S simply by B_n. Boolean algebras will play an important role throughout this section.

There is a simple way to represent small posets pictorially. The *Hasse diagram* of a poset P is a planar drawing, with elements of P drawn as dots. If $x < y$ in P (i.e., $x \leq y$ and $x \neq y$), then y is drawn "above" x (i.e., with a larger vertical coordinate). An edge is drawn between x and y if y *covers* x, i.e., $x < y$ and no element z is in between, i.e., no z satisfies $x < z < y$. By the transitivity property (P3), all the relations of a finite poset are determined by the cover relations, so the Hasse diagram determines P. (This is not true for infinite posets; for instance, the real numbers $\mathbb{R}$ with their usual order is a poset with no cover relations.) The Hasse diagram of the boolean algebra B_3 looks like



We say that two posets P and Q are *isomorphic* if there is a bijection (one-to-one and onto function) $\varphi : P \to Q$ such that $x \le y$ in P if and only if $\varphi(x) \le \varphi(y)$ in Q. Thus one can think that two posets are isomorphic if they differ only in the names of their elements. This is exactly analogous to the notion of isomorphism of groups, rings, etc. It is an instructive exercise to draw Hasse diagrams of the one poset of order (number of elements) one (up to isomorphism), the two posets of order two, the five posets of order three, and the sixteen posets of order four. More ambitious readers can try the 63 posets of order five, the 318 of order six, the 2045 of order seven, the 16999 of order eight, the 183231 of order nine, the 2567284 of order ten, the 46749427 of order eleven, the 1104891746 of order twelve, the 33823827452 of order thirteen, and the 1338193159771 of order fourteen. Beyond this the number is not currently known.

A *chain* C in a poset is a totally ordered subset of P, i.e., if $x, y \in C$ then either $x \le y$ or $y \le x$ in P. A finite chain is said to have *length* n if it has $n+1$ elements. Such a chain thus has the form $x_0 < x_1 < \cdots < x_n$. We say that a finite poset is *graded of rank* n if every maximal chain has length n. (A chain is *maximal* if it's contained in no larger chain.) For instance, the boolean algebra B_n is graded of rank n [why?]. A chain $y_0 < y_1 < \cdots < y_j$ is said to be *saturated* if each y_{i+1} covers y_i. Such a chain need not be maximal since there can be elements of P smaller than y_0 or greater than y_j. If P is graded of rank n and $x \in P$, then we say that x has *rank* j, denoted $\rho(x) = j$, if some (or equivalently, every) saturated chain of P with top element x has length j. Thus [why?] if we let $P_j = \{x \in P : \rho(x) = j\}$, then P is a *disjoint* union $P = P_0 \cup P_1 \cup \cdots \cup P_n$, and every maximal chain of P has the form $x_0 < x_1 < \cdots < x_n$ where $\rho(x_j) = j$. We write $p_j = |P_j|$, the number of elements of P of rank j. For example, if $P = B_n$ then $\rho(x) = |x|$ (the cardinality of x as a set) and

$$p_j = \#\{x \subseteq \{1, 2, \cdots, n\} : |x| = j\} = \binom{n}{j}.$$

(Note that we use both $|S|$ and $\#x$ for the cardinality of the finite set S.)

We say that a graded poset P of rank n (always assumed to be finite) is *rank-symmetric* if $p_i = p_{n-i}$ for $0 \leq i \leq n$, and *rank-unimodal* if $p_0 \leq p_1 \leq \cdots \leq p_j \geq p_{j+1} \geq p_{j+2} \geq \cdots \geq p_n$ for some $0 \leq j \leq n$. If P is both rank-symmetric and rank-unimodal, then we clearly have

$$p_0 \leq p_1 \leq \cdots \leq p_m \geq p_{m+1} \geq \cdots \geq p_n, \text{ if } n = 2m$$

$$p_0 \leq p_1 \leq \cdots \leq p_m = p_{m+1} \geq p_{m+2} \geq \cdots \geq p_n, \text{ if } n = 2m + 1.$$

We also say that the sequence $p_0, p_1, \ldots, p_n$ itself or the polynomial $F(q) = p_0 + p_1 q + \cdots + p_n q^n$ is *symmetric* or *unimodal*, as the case may be. For instance, B_n is rank-symmetric and rank-unimodal, since it is well-known (and easy to prove) that the sequence $\binom{n}{0}, \binom{n}{1}, \ldots, \binom{n}{n}$ (the nth row of Pascal's triangle) is symmetric and unimodal. Thus the polynomial $(1+q)^n$ is symmetric and unimodal.

A few more definitions, and then finally some results! An *antichain* in a poset P is a subset A of P for which no two elements are comparable, i.e., we can never have $x, y \in A$ and $x < y$. For instance, in a graded poset P the "levels" P_j are antichains [why?]. We will be concerned with the problem of finding the largest antichain in a poset. Consider for instance the boolean algebra B_n. The problem of finding the largest antichain in B_n is clearly equivalent to the following problem in extremal set theory: Find the largest collection of subsets of an n-element set such that no element of the collection contains another. A good guess would be to take all the subsets of cardinality $\lfloor n/2 \rfloor$ (where $\lfloor x \rfloor$ denotes the greatest integer $\leq x$), giving a total of $\binom{n}{\lfloor n/2 \rfloor}$ sets in all. But how can we actually prove there is no larger collection? Such a proof was first given by Emmanuel Sperner in 1927 and is known as *Sperner's theorem*. We will give two proofs of Sperner's theorem in this section; one proof uses linear algebra and will be applied to certain other situations, while the other proof is an elegant combinatorial argument due to David Lubell in 1966, which we present for its "cultural value." Our extension of Sperner's theorem to certain other situations will involve the following crucial definition.

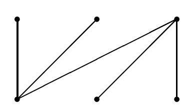
4.2 Definition. Let P be a graded poset of rank n. We say that P has the *Sperner property* or is a *Sperner poset* if

$$\max\{|A| : A \text{ is an antichain of } P\} = \max\{|P_i| : 0 \le i \le n\}.$$

In other words, no antichain is larger than the largest level P_i.

Thus Sperner's theorem is equivalent to saying that B_n has the Sperner property. Note that if P has the Sperner property there may still be antichains of maximum cardinality other than the biggest P_i; there just can't be any bigger antichains.

4.3 Example. A simple example of a graded poset that fails to satisfy the Sperner property is the following:



We now will discuss a simple combinatorial condition which guarantees that certain graded posets P are Sperner. We define an *order-matching* from P_i to P_{i+1} to be a *one-to-one* function $\mu : P_i \to P_{i+1}$ satisfying $x < \mu(x)$ for all $x \in P_i$. Clearly if such an order-matching exists then $p_i \le p_{i+1}$ (since μ is one-to-one). Easy examples show that the converse is false, i.e., if $p_i \le p_{i+1}$ then there need not exist an order-matching from P_i to P_{i+1}. We similarly define an order-matching from P_i to P_{i-1} to be a one-to-one function $\mu : P_i \to P_{i-1}$ satisfying $\mu(x) < x$ for all $x \in P_i$.

4.4 Proposition. *Let P be a graded poset of rank n. Suppose there exists an integer $0 \le j \le n$ and order-matchings*

$$P_0 \to P_1 \to P_2 \to \cdots \to P_j \leftarrow P_{j+1} \leftarrow P_{j+2} \leftarrow \cdots \leftarrow P_n. \tag{17}$$

Then P is rank-unimodal and Sperner.

Proof. Since order-matchings are one-to-one it is clear that

$$p_0 \leq p_1 \leq \cdots \leq p_j \geq p_{j+1} \geq p_{j+2} \geq \cdots \geq p_n.$$

Hence P is rank-unimodal.

Define a graph G as follows. The vertices of G are the elements of P. Two vertices x, y are connected by an edge if one of the order-matchings μ in the statement of the proposition satisfies $\mu(x) = y$. (Thus G is a subgraph of the Hasse diagram of P.) Drawing a picture will convince you that G consists of a disjoint union of paths, including single-vertex paths not involved in any of the order-matchings. The vertices of each of these paths form a chain in P. Thus we have partitioned the elements of P into disjoint chains. Since P is rank-unimodal with biggest level P_j, all of these chains must pass through P_j [why?]. Thus the number of chains is exactly p_j. Any antichain A can intersect each of these chains at most once, so the cardinality $|A|$ of A cannot exceed the number of chains, i.e., $|A| \leq p_j$. Hence by definition P is Sperner. $\square$

It is now finally time to bring some linear algebra into the picture. For any (finite) set S, we let $\mathbb{R}S$ denote the real vector space consisting of all formal linear combinations (with real coefficients) of elements of S. Thus S is a basis for $\mathbb{R}S$, and in fact we could have simply defined $\mathbb{R}S$ to be the real vector space with basis S. The next lemma relates the combinatorics we have just discussed to linear algebra and will allow us to prove that certain posets are Sperner by the use of linear algebra (combined with some finite group theory).

4.5 Lemma. *Suppose there exists a linear transformation $U : \mathbb{R}P_i \to \mathbb{R}P_{i+1}$ (U stands for "up") satisfying:*

- *U is one-to-one.*
- *For all $x \in P_i$, $U(x)$ is a linear combination of elements $y \in P_{i+1}$ satisfying $x < y$. (We then call U an* order-raising operator.*)*

Then there exists an order-matching $\mu : P_i \to P_{i+1}$.

Similarly, suppose there exists a linear transformation $U : \mathbb{R}P_i \to \mathbb{R}P_{i+1}$ satisfying:

- *U is onto.*
- *U is an order-raising operator.*

Then there exists an order-matching $\mu : P_{i+1} \to P_i$.

Proof. Suppose $U : \mathbb{R}P_i \to \mathbb{R}P_{i+1}$ is a one-to-one order-raising operator. Let $[U]$ denote the matrix of U with respect to the bases P_i of $\mathbb{R}P_i$ and P_{i+1} of $\mathbb{R}P_{i+1}$. Thus the columns of $[U]$ are indexed by the elements $x_1, \dots, x_{p_i}$ of P_i (in some order) and the rows by the elements $y_1, \dots, y_{p_{i+1}}$ of P_{i+1}. Since U is one-to-one, the rank of $[U]$ is equal to p_i (the number of columns). Since the column rank of a matrix equals its row rank, $[U]$ must have p_i linearly independent rows. Say we have labelled the elements of P_{i+1} so that the first p_i rows of $[U]$ are linearly independent.

Let $A = (a_{ij})$ be the $p_i \times p_i$ matrix whose rows are the first p_i rows of $[U]$. (Thus A is a square submatrix of $[U]$.) Since the rows of A are linearly independent, we have

$$\det(A) = \sum \pm a_{\pi(1),1} \cdots a_{\pi(p_i),p_i} \neq 0,$$

where the sum is over all permutations π of $1, \dots, p_i$. Thus some term $\pm a_{\pi(1),1} \cdots a_{\pi(p_i),p_i}$ of the above sum in nonzero. Since U is order-raising, this means that [why?] $x_k < y_{\pi(k)}$ for $1 \leq k \leq p_i$. Hence the map $\mu : P_i \to P_{i+1}$ defined by $\mu(x_k) = y_{\pi(k)}$ is an order-matching, as desired.

The case when U is onto rather than one-to-one is proved by a completely analogous argument. □

We now want to apply Proposition 4.4 and Lemma 4.5 to the boolean algebra B_n. For each $0 \leq i < n$, we need to define a linear transformation $U_i : \mathbb{R}(B_n)_i \to \mathbb{R}(B_n)_{i+1}$, and then prove it has the desired properties. We simply define U_i to be the simplest possible order-raising operator, namely, for $x \in (B_n)_i$, let

$$U_i(x) = \sum_{\substack{y \in (B_n)_{i+1} \\ y > x}} y. \tag{18}$$

Note that since $(B_n)_i$ is a basis for $\mathbb{R}(B_n)_i$, equation (18) does indeed define a unique linear transformation $U_i : \mathbb{R}(B_n)_i \to \mathbb{R}(B_n)_{i+1}$. By definition U_i is order-raising; we want to show that U_i is one-to-one for $i < n/2$ and onto for $i \geq n/2$. There are several ways to show this using only elementary linear algebra; we will give what is perhaps the simplest proof, though it is quite tricky. The idea is to introduce "dual" operators $D_i : \mathbb{R}(B_n)_i \to (B_n)_{i-1}$ to the U_i's (D stands for "down"), defined by

$$D_i(y) = \sum_{\substack{x \in (B_n)_{i-1} \\ x < y}} x, \tag{19}$$

for all $y \in (B_n)_i$. Let $[U_i]$ denote the matrix of U_i with respect to the bases $(B_n)_i$ and $(B_n)_{i+1}$, and similarly let $[D_i]$ denote the matrix of D_i with respect to the bases $(B_n)_i$ and $(B_n)_{i-1}$. A key observation which we will use later is that

$$[D_{i+1}] = [U_i]^t, \tag{20}$$

i.e., the matrix $[D_{i+1}]$ is the transpose of the matrix $[U_i]$ [why?]. Now let $I_i : \mathbb{R}(B_n)_i \to \mathbb{R}(B_n)_i$ denote the identity transformation on $\mathbb{R}(B_n)_i$, i.e., $I_i(u) = u$ for all $u \in \mathbb{R}(B_n)_i$. The next lemma states (in linear algebraic terms) the fundamental combinatorial property of B_n which we need. For this lemma set $U_n = 0$ and $D_0 = 0$ (the 0 linear transformation between the appropriate vector spaces).

4.6 Lemma. *Let $0 \leq i \leq n$. Then*

$$D_{i+1}U_i - U_{i-1}D_i = (n - 2i)I_i. \tag{21}$$

(Linear transformations are multiplied right-to-left, so $AB(u) = A(B(u))$.)

Proof. Let $x \in (B_n)_i$. We need to show that if we apply the left-hand side of (21) to x, then we obtain $(n-2i)x$. We have

$$D_{i+1}U_i(x) = D_{i+1}\left(\sum_{\substack{|y|=i+1 \\ x \subset y}} y\right)$$

$$= \sum_{\substack{|y|=i+1 \\ x\subset y}} \sum_{\substack{|z|=i \\ z\subset y}} z.$$

If $x, z \in (B_n)_i$ satisfy $|x \cap z| < i-1$, then there is no $y \in (B_n)_{i+1}$ such that $x \subset y$ and $z \subset y$. Hence the coefficient of z in $D_{i+1}U_i(x)$ when it is expanded in terms of the basis $(B_n)_i$ is 0. If $|x \cap z| = i-1$, then there is one such y, namely, $y = x \cup z$. Finally if $x = z$ then y can be any element of $(B_n)_{i+1}$ containing x, and there are $n-i$ such y in all. It follows that

$$D_{i+1}U_i(x) = (n-i)x + \sum_{\substack{|z|=i \\ |x\cap z|=i-1}} z. \tag{22}$$

By exactly analogous reasoning (which the reader should check), we have for $x \in (B_n)_i$ that

$$U_{i-1}D_i(x) = ix + \sum_{\substack{|z|=i \\ |x\cap z|=i-1}} z. \tag{23}$$

Subtracting (23) from (22) yields $(D_{i+1}U_i - U_{i-1}D_i)(x) = (n-2i)x$, as desired. $\square$

4.7 Theorem. *The operator U_i defined above is one-to-one if $i < n/2$ and is onto if $i \geq n/2$.*

Proof. Recall that $[D_i] = [U_{i-1}]^t$. From linear algebra we know that a (rectangular) matrix times its transpose is *positive semidefinite* (or just *semidefinite* for short) and hence has nonnegative (real) eigenvalues. By Lemma 4.6 we have

$$D_{i+1}U_i = U_{i-1}D_i + (n-2i)I_i.$$

Thus the eigenvalues of $D_{i+1}U_i$ are obtained from the eigenvalues of $U_{i-1}D_i$ by adding $n-2i$. Since we are assuming that $n-2i > 0$, it follows that the eigenvalues of $D_{i+1}U_i$ are strictly positive. Hence $D_{i+1}U_i$ is invertible (since it has no 0 eigenvalues). But this implies that U_i is one-to-one [why?], as desired.

The case $i \geq n/2$ is done by a "dual" argument (or in fact can be deduced directly from the $i < n/2$ case by using the fact that the poset B_n is "self-dual," though we will not go into this). Namely, from the fact that

$$U_iD_{i+1} = D_{i+2}U_{i+1} + (2i+2-n)I_{i+1}$$

we get that $U_i D_{i+1}$ is invertible, so now U_i is onto, completing the proof. $\square$

Combining Proposition 4.4, Lemma 4.5, and Theorem 4.7, we obtain the famous theorem of Sperner.

4.8 Corollary. *The boolean algebra B_n has the Sperner property.*

It is natural to ask whether there is a less indirect proof of Corollary 4.8. In fact, several nice proofs are known; we give one due to David Lubell, mentioned before Definition 4.2.

Lubell's proof of Sperner's theorem. First we count the total number of maximal chains $\emptyset = x_0 < x_1 < \cdots < x_n = \{1, \ldots, n\}$ in B_n. There are n choices for x_1, then $n-1$ choices for x_2, etc., so there are $n!$ maximal chains in all. Next we count the number of maximal chains $x_0 < x_1 < \cdots < x_i = x < \cdots < x_n$ which contain a given element x of rank i. There are i choices for x_1, then $i-1$ choices for x_2, up to one choice for x_i. Similarly there are $n-i$ choices for x_{i+1}, then $n-2$ choices for x_{i+2}, etc., up to one choice for x_n. Hence the number of maximal chains containing x is $i!(n-i)!$.

Now let A be an antichain. If $x \in A$, then let C_x be the set of maximal chains of B_n which contain x. Since A is an antichain, the sets C_x, $x \in A$ are pairwise disjoint. Hence

$$\begin{aligned} |\bigcup_{x \in A} C_x| &= \sum_{x \in A} |C_x| \\ &= \sum_{x \in A} (\rho(x))!(n-\rho(x))! \end{aligned}$$

Since the total number of maximal chains in the C_x's cannot exceed the total number $n!$ of maximal chains in B_n, we have

$$\sum_{x \in A} (\rho(x))!(n-\rho(x))! \leq n!$$

Divide both sides by $n!$ to obtain

$$\sum_{x \in A} \frac{1}{\binom{n}{\rho(x)}} \leq 1.$$

Since $\binom{n}{i}$ is maximized when $i = \lfloor n/2 \rfloor$, we have

$$\frac{1}{\binom{n}{\lfloor n/2 \rfloor}} \leq \frac{1}{\binom{n}{\rho(x)}},$$

for all $x \in A$ (or all $x \in B_n$). Thus

$$\sum_{x \in A} \frac{1}{\binom{n}{\lfloor n/2 \rfloor}} \leq 1,$$

or equivalently,

$$|A| \leq \binom{n}{\lfloor n/2 \rfloor}.$$

Since $\binom{n}{\lfloor n/2 \rfloor}$ is the size of the largest level of B_n, it follows that B_n is Sperner.
□

In view of the above elegant proof of Lubell, the reader may be wondering what was the point of giving a rather complicated and indirect proof using linear algebra. Admittedly, if all we could obtain from the linear algebra machinery we have developed was just another proof of Sperner's theorem, then it would have been hardly worth the effort. But in the next section we will show how Theorem 4.7, when combined with a little finite group theory, can be used to obtain many interesting combinatorial results for which simple, direct proofs are not known.