

TOPICS IN ALGEBRAIC COMBINATORICS

Richard P. Stanley

Course notes for 18.312 (Algebraic Combinatorics)

M.I.T., Spring 2007

PRELIMINARY (INCOMPLETE) VERSION OF JANUARY, 2007

Acknowledgment. I am grateful to Sergey Fomin for his careful reading of the manuscript and for several helpful suggestions.

1 Walks in graphs.

Given a finite set S and integer $k \geq 0$, let $\binom{S}{k}$ denote the set of k -element subsets of S , and let $\left(\binom{S}{k}\right)$ denote the set of k -element multisubsets (sets with repeated elements) on S . For instance, if $S = \{1, 2, 3\}$ then (using abbreviated notation),

$$\binom{S}{2} = \{12, 13, 23\}, \quad \left(\binom{S}{2}\right) = \{11, 22, 33, 12, 13, 23\}.$$

A (finite) *graph* G consists of a *vertex set* $V = \{v_1, \dots, v_p\}$ and *edge set* $E = \{e_1, \dots, e_q\}$, together with a function $\varphi : E \rightarrow \left(\binom{V}{2}\right)$. We think that if $\varphi(e) = uv$ (short for $\{u, v\}$), then e connects u and v or equivalently e is *incident* to u and v . If there is at least one edge incident to u and v then we say that the vertices u and v are *adjacent*. If $\varphi(e) = vv$, then we call e a *loop* at v . If several edges e_1, \dots, e_j ($j > 1$) satisfy $\varphi(e_1) = \dots = \varphi(e_j) = uv$, then we say that there is a *multiple edge* between u and v . A graph without loops or multiple edges is called *simple*. In this case we can think of E as just a subset of $\binom{V}{2}$ [why?].

The *adjacency matrix* of the graph G is the $p \times p$ matrix $\mathbf{A} = \mathbf{A}(G)$, over the field of complex numbers, whose (i, j) -entry a_{ij} is equal to the number of edges incident to v_i and v_j . Thus \mathbf{A} is a real symmetric matrix (and hence has real eigenvalues) whose trace is the number of loops in G .

A *walk* in G of length ℓ from vertex u to vertex v is a sequence $v_1, e_1, v_2, e_2, \dots, v_\ell, e_\ell, v_{\ell+1}$ such that:

- each v_i is a vertex of G
- each e_j is an edge of G
- the vertices of e_i are v_i and v_{i+1} , for $1 \leq i \leq \ell$
- $v_1 = u$ and $v_{\ell+1} = v$.

1.1 Theorem. For any integer $\ell \geq 1$, the (i, j) -entry of the matrix $\mathbf{A}(G)^\ell$ is equal to the number of walks from v_i to v_j in G of length ℓ .

Proof. This is an immediate consequence of the definition of matrix multiplication. Let $\mathbf{A} = (a_{ij})$. The (i, j) -entry of $\mathbf{A}(G)^\ell$ is given by

$$(\mathbf{A}(G)^\ell)_{ij} = \sum a_{ii_1} a_{i_1 i_2} \cdots a_{i_{\ell-1} j},$$

where the sum ranges over all sequences $(i_1, \dots, i_{\ell-1})$ with $1 \leq i_k \leq p$. But since a_{rs} is the number of edges between v_r and v_s , it follows that the summand $a_{ii_1} a_{i_1 i_2} \cdots a_{i_{\ell-1} j}$ in the above sum is just the number (which may be 0) of walks of length ℓ from v_i to v_j of the form

$$v_i, e_1, v_{i_1}, e_2, \dots, v_{i_{\ell-1}}, e_\ell, v_j$$

(since there are a_{ii_1} choices for e_1 , $a_{i_1 i_2}$ choices for e_2 , etc.) Hence summing over all $(i_1, \dots, i_{\ell-1})$ just gives the total number of walks of length ℓ from v_i to v_j , as desired. \square

We wish to use Theorem 1.1 to obtain an explicit formula for the number $(\mathbf{A}(G)^\ell)_{ij}$ of walks of length ℓ in G from v_i to v_j . The formula we give will depend on the eigenvalues of $\mathbf{A}(G)$. The eigenvalues of $\mathbf{A}(G)$ are also called simply the *eigenvalues of G* . Recall that a real symmetric $p \times p$ matrix M has p linearly independent real eigenvectors, which can in fact be chosen to be orthonormal (i.e., orthogonal and of unit length). Let u_1, \dots, u_p be real orthonormal unit eigenvectors for M , with corresponding eigenvalues $\lambda_1, \dots, \lambda_p$. All vectors u will be regarded as $p \times 1$ *column* vectors. We let t denote transpose, so u^t is a $1 \times p$ *row* vector. Thus the dot (or scalar

or inner) product of the vectors u and v is given by $u^t v$ (ordinary matrix multiplication). In particular, $u_i^t u_j = \delta_{ij}$ (the Kronecker delta). Let $U = (u_{ij})$ be the matrix whose columns are u_1, \dots, u_p , denoted $U = [u_1, \dots, u_p]$. Thus U is an orthogonal matrix and

$$U^t = U^{-1} = \begin{bmatrix} u_1^t \\ \cdot \\ \cdot \\ \cdot \\ u_p^t \end{bmatrix},$$

the matrix whose rows are u_1^t, \dots, u_p^t . Recall from linear algebra that the matrix U diagonalizes M , i.e.,

$$U^{-1} M U = \text{diag}(\lambda_1, \dots, \lambda_p),$$

where $\text{diag}(\lambda_1, \dots, \lambda_p)$ denotes the diagonal matrix with diagonal entries $\lambda_1, \dots, \lambda_p$.

1.2 Corollary. *Given the graph G as above, fix the two vertices v_i and v_j . Let $\lambda_1, \dots, \lambda_p$ be the eigenvalues of the adjacency matrix $\mathbf{A}(G)$. Then there exist real numbers c_1, \dots, c_p such that for all $\ell \geq 1$, we have*

$$(\mathbf{A}(G)^\ell)_{ij} = c_1 \lambda_1^\ell + \dots + c_p \lambda_p^\ell.$$

In fact, if $U = (u_{rs})$ is a real orthogonal matrix such that $U^{-1} \mathbf{A} U = \text{diag}(\lambda_1, \dots, \lambda_p)$, then we have

$$c_k = u_{ik} u_{jk}.$$

Proof. We have [why?]

$$U^{-1} \mathbf{A}^\ell U = \text{diag}(\lambda_1^\ell, \dots, \lambda_p^\ell).$$

Hence

$$\mathbf{A}^\ell = U \cdot \text{diag}(\lambda_1^\ell, \dots, \lambda_p^\ell) U^{-1}.$$

Taking the (i, j) -entry of both sides (and using $U^{-1} = U^t$) gives [why?]

$$(\mathbf{A}^\ell)_{ij} = \sum_k u_{ik} \lambda_k^\ell u_{jk},$$

as desired. \square

In order for Corollary 1.2 to be of any use we must be able to compute the eigenvalues $\lambda_1, \dots, \lambda_p$ as well as the diagonalizing matrix U (or eigenvectors u_i). There is one interesting special situation in which it is not necessary to compute U . A *closed walk* in G is a walk that ends where it begins. The number of closed walks in G of length ℓ starting at v_i is therefore given by $(\mathbf{A}(G)^\ell)_{ii}$, so the *total* number $f_G(\ell)$ of closed walks of length ℓ is given by

$$\begin{aligned} f_G(\ell) &= \sum_{i=1}^p (\mathbf{A}(G)^\ell)_{ii} \\ &= \text{tr}(\mathbf{A}(G)^\ell), \end{aligned}$$

where tr denotes trace (sum of the main diagonal entries). Now recall that the trace of a square matrix is the sum of its eigenvalues. If the matrix M has eigenvalues $\lambda_1, \dots, \lambda_p$ then [why?] M^ℓ has eigenvalues $\lambda_1^\ell, \dots, \lambda_p^\ell$. Hence we have proved the following.

1.3 Corollary. *Suppose $\mathbf{A}(G)$ has eigenvalues $\lambda_1, \dots, \lambda_p$. Then the number of closed walks in G of length ℓ is given by*

$$f_G(\ell) = \lambda_1^\ell + \dots + \lambda_p^\ell.$$

We now are in a position to use various tricks and techniques from linear algebra to count walks in graphs. Conversely, it is sometimes possible to count the walks by combinatorial reasoning and use the resulting formula to determine the eigenvalues of G . As a first simple example, we consider the *complete graph* K_p with vertex set $V = \{v_1, \dots, v_p\}$, and one edge between any two *distinct* vertices. Thus K_p has p vertices and $\binom{p}{2} = \frac{1}{2}p(p-1)$ edges.

1.4 Lemma. *Let J denote the $p \times p$ matrix of all 1's. Then the eigenvalues of J are p (with multiplicity one) and 0 (with multiplicity $p-1$).*

Proof. Since all rows are equal and nonzero, we have $\text{rank}(J) = 1$. Since a $p \times p$ matrix of rank $p-m$ has at least m eigenvalues equal to 0, we conclude that J has at least $p-1$ eigenvalues equal to 0. Since $\text{tr}(J) = p$ and the trace is the sum of the eigenvalues, it follows that the remaining eigenvalue of J is equal to p . \square

1.5 Proposition. *The eigenvalues of the complete graph K_p are as follows: an eigenvalue of -1 with multiplicity $p - 1$, and an eigenvalue of $p - 1$ with multiplicity one.*

Proof. We have $\mathbf{A}(K_p) = J - I$, where I denotes the $p \times p$ identity matrix. If the eigenvalues of a matrix M are μ_1, \dots, μ_p , then the eigenvalues of $M + cI$ (where c is a scalar) are $\mu_1 + c, \dots, \mu_p + c$ [why?]. The proof follows from Lemma 1.4. \square

1.6 Corollary. *The number of closed walks of length ℓ in K_p from some vertex v_i to itself is given by*

$$(\mathbf{A}(K_p)^\ell)_{ii} = \frac{1}{p}((p - 1)^\ell + (p - 1)(-1)^\ell). \quad (1)$$

(Note that this is also the number of sequences (i_1, \dots, i_ℓ) of numbers $1, 2, \dots, p$ such that $i_1 = i$, no two consecutive terms are equal, and $i_\ell \neq i_1$ [why?].)

Proof. By Corollary 1.3 and Proposition 1.5, the total number of closed walks in K_p of length ℓ is equal to $(p - 1)^\ell + (p - 1)(-1)^\ell$. By the symmetry of the graph K_p , the number of closed walks of length ℓ from v_i to itself does not depend on i . (All vertices “look the same.”) Hence we can divide the total number of closed walks by p (the number of vertices) to get the desired answer. \square

What about non-closed walks in K_p ? It’s not hard to diagonalize explicitly the matrix $\mathbf{A}(K_p)$ (or equivalently, to compute its eigenvectors), but there is an even simpler special argument. We have

$$(J - I)^\ell = \sum_{k=0}^{\ell} (-1)^{\ell-k} \binom{\ell}{k} J^k, \quad (2)$$

by the binomial theorem. Now for $k > 0$ we have $J^k = p^{k-1}J$ [why?], while $J^0 = I$. (It is not clear *a priori* what is the “correct” value of J^0 , but in order for equation (2) to be valid we must take $J^0 = I$.) Hence

$$(J - I)^\ell = \sum_{k=1}^{\ell} (-1)^{\ell-k} \binom{\ell}{k} p^{k-1}J + (-1)^\ell I.$$

Again by the binomial theorem we have

$$\begin{aligned} (J - I)^\ell &= \frac{1}{p}((p-1)^\ell J - (-1)^\ell J) + (-1)^\ell I \\ &= \frac{1}{p}(p-1)^\ell J + \frac{(-1)^\ell}{p}(pI - J). \end{aligned} \quad (3)$$

Taking the (i, j) -entry of each side when $i \neq j$ yields

$$(\mathbf{A}(K_p)^\ell)_{ij} = \frac{1}{p}((p-1)^\ell - (-1)^\ell). \quad (4)$$

If we take the (i, i) -entry of (3) then we recover equation (1). Note the curious fact that if $i \neq j$ then

$$(\mathbf{A}(K_p)^\ell)_{ii} - (\mathbf{A}(K_p)^\ell)_{ij} = (-1)^\ell.$$

We could also have deduced (4) from Corollary 1.6 using

$$\sum_{i=1}^p \sum_{j=1}^p (\mathbf{A}(K_p)^\ell)_{ij} = p(p-1)^\ell,$$

the total number of walks of length ℓ in K_p . Details are left to the reader.

We now will show how equation (1) itself determines the eigenvalues of $\mathbf{A}(K_p)$. Thus if (1) is proved without first computing the eigenvalues of $\mathbf{A}(K_p)$ (which in fact is what we did two paragraphs ago), then we have another means to compute the eigenvalues. The argument we will give can be applied to any graph G , not just K_p . We begin with a simple lemma.

1.7 Lemma. *Suppose $\alpha_1, \dots, \alpha_r$ and β_1, \dots, β_s are nonzero complex numbers such that for all positive integers ℓ , we have*

$$\alpha_1^\ell + \dots + \alpha_r^\ell = \beta_1^\ell + \dots + \beta_s^\ell. \quad (5)$$

Then $r = s$ and the α 's are just a permutation of the β 's.

Proof. We will use the powerful method of *generating functions*. Let x be a complex number whose absolute value is close to 0. Multiply (5) by x^ℓ

and sum on all $\ell \geq 1$. The geometric series we obtain will converge, and we get

$$\frac{\alpha_1 x}{1 - \alpha_1 x} + \cdots + \frac{\alpha_r x}{1 - \alpha_r x} = \frac{\beta_1 x}{1 - \beta_1 x} + \cdots + \frac{\beta_s x}{1 - \beta_s x}. \quad (6)$$

This is an identity valid for sufficiently small (in modulus) complex numbers. By clearing denominators we obtain a polynomial identity. But if two polynomials in x agree for infinitely many values, then they are the same polynomial [why?]. Hence equation (6) is actually valid for *all* complex numbers x (ignoring values of x which give rise to a zero denominator).

Fix a complex number $\gamma \neq 0$. Multiply (6) by $1 - \gamma x$ and let $x \rightarrow 1/\gamma$. The left-hand side becomes the number of α_i 's which are equal to γ , while the right-hand side becomes the number of β_j 's which are equal to γ [why?]. Hence these numbers agree for all γ , so the lemma is proved. \square

1.8 Example. Suppose that G is a graph with 12 vertices, and that the number of closed walks of length ℓ in G is equal to $3 \cdot 5^\ell + 4^\ell + 2(-2)^\ell + 4$. Then it follows from Corollary 1.3 and Lemma 1.7 [why?] that the eigenvalues of $\mathbf{A}(G)$ are given by 5, 5, 5, 4, -2, -2, 1, 1, 1, 1, 0, 0.

2 Cubes and the Radon transform.

Let us now consider a more interesting example of a graph G , one whose eigenvalues have come up in a variety of applications. Let \mathbb{Z}_2 denote the cyclic group of order 2, with elements 0 and 1, and group operation being addition modulo 2. Thus $0 + 0 = 0$, $0 + 1 = 1 + 0 = 1$, $1 + 1 = 0$. Let \mathbb{Z}_2^n denote the direct product of \mathbb{Z}_2 with itself n times, so the elements of \mathbb{Z}_2^n are n -tuples (a_1, \dots, a_n) of 0's and 1's, under the operation of component-wise addition. Define a graph C_n , called the n -cube, as follows: The vertex set of C_n is given by $V(C_n) = \mathbb{Z}_2^n$, and two vertices u and v are connected by an edge if they differ in exactly one component. Equivalently, $u + v$ has exactly one nonzero component. If we regard \mathbb{Z}_2^n as consisting of *real* vectors, then these vectors form the set of vertices of an n -dimensional cube. Moreover, two vertices of the cube lie on an edge (in the usual geometric sense) if and only if they form an edge of C_n . This explains why C_n is called the n -cube. We also see that walks in C_n have a nice geometric interpretation — they are simply walks along the edges of an n -dimensional cube.

We want to determine explicitly the eigenvalues and eigenvectors of C_n . We will do this by a somewhat indirect but extremely useful and powerful technique, the finite Radon transform. Let \mathcal{V} denote the set of all functions $f : \mathbb{Z}_2^n \rightarrow \mathbb{R}$, where \mathbb{R} denotes the field of real numbers. (NOTE: For groups other than \mathbb{Z}_2^n it is necessary to use complex numbers rather than real numbers. We could use complex numbers here, but there is no need to do so.) Note that \mathcal{V} is a vector space over \mathbb{R} of dimension 2^n [why?]. If $u = (u_1, \dots, u_n)$ and $v = (v_1, \dots, v_n)$ are elements of \mathbb{Z}_2^n , then define their *dot product* by

$$u \cdot v = u_1 v_1 + \dots + u_n v_n,$$

where the computation is performed modulo 2. Thus we regard $u \cdot v$ as an element of \mathbb{Z}_2 . The expression $(-1)^{u \cdot v}$ is defined to be the *real number* $+1$ or -1 , depending on whether $u \cdot v = 0$ or 1 , respectively. Since for integers k the value of $(-1)^k$ depends only on $k \pmod{2}$, it follows that we can treat u and v as integer vectors without affecting the value of $(-1)^{u \cdot v}$. Thus, for instance, formulas such as

$$(-1)^{u \cdot (v+w)} = (-1)^{u \cdot v + u \cdot w} = (-1)^{u \cdot v} (-1)^{u \cdot w}$$

are well-defined and valid.

We now define two important bases of the vector space \mathcal{V} . There will be one basis element of each basis for each $u \in \mathbb{Z}_2^n$. The first basis, denoted B_1 , has elements f_u defined as follows:

$$f_u(v) = \delta_{uv}, \quad (7)$$

the Kronecker delta. It is easy to see that B_1 is a basis, since any $g \in \mathcal{V}$ satisfies

$$g = \sum_{u \in \mathbb{Z}_2^n} g(u) f_u \quad (8)$$

[why?]. Hence B_1 spans \mathcal{V} , so since $|B_1| = \dim \mathcal{V} = 2^n$, it follows that B_1 is a basis. The second basis, denoted B_2 , has elements χ_u defined as follows:

$$\chi_u(v) = (-1)^{u \cdot v}.$$

In order to show that B_2 is a basis, we will use an inner product on \mathcal{V} (denoted $\langle \cdot, \cdot \rangle$) defined by

$$\langle f, g \rangle = \sum_{u \in \mathbb{Z}_2^n} f(u) g(u).$$

Note that this inner product is just the usual dot product with respect to the basis B_1 .

2.1 Lemma. *The set $B_2 = \{\chi_u : u \in \mathbb{Z}_2^n\}$ forms a basis for \mathcal{V} .*

Proof. Since $|B_2| = \dim \mathcal{V} (= 2^n)$, it suffices to show that B_2 is linearly independent. In fact, we will show that the elements of B_2 are orthogonal. We have

$$\begin{aligned} \langle \chi_u, \chi_v \rangle &= \sum_{w \in \mathbb{Z}_2^n} \chi_u(w) \chi_v(w) \\ &= \sum_{w \in \mathbb{Z}_2^n} (-1)^{(u+v) \cdot w}. \end{aligned}$$

It is left as an easy exercise to the reader to show that for any $y \in \mathbb{Z}_2^n$, we have

$$\sum_{w \in \mathbb{Z}_2^n} (-1)^{y \cdot w} = \begin{cases} 2^n, & \text{if } y = \mathbf{0} \\ 0, & \text{otherwise.} \end{cases}$$

where $\mathbf{0}$ denotes the identity element of \mathbb{Z}_2^n (the vector $(0, 0, \dots, 0)$). Thus $\langle \chi_u, \chi_v \rangle = 0$ if and only if $u + v = \mathbf{0}$, i.e., $u = v$, so the elements of B_2 are orthogonal (and nonzero). Hence they are linearly independent as desired. \square

We now come to the key definition of the Radon transform.

2.2 Definition. Given a subset Γ of \mathbb{Z}_2^n and a function $f \in \mathcal{V}$, define a new function $\Phi_\Gamma f \in \mathcal{V}$ by

$$\Phi_\Gamma f(v) = \sum_{w \in \Gamma} f(v + w).$$

The function $\Phi_\Gamma f$ is called the (*discrete* or *finite*) *Radon transform* of f (on the group \mathbb{Z}_2^n , with respect to the subset Γ).

We have defined a map $\Phi_\Gamma : \mathcal{V} \rightarrow \mathcal{V}$. It is easy to see that Φ_Γ is a linear transformation; we want to compute its eigenvalues and eigenvectors.

2.3 Theorem. *The eigenvectors of Φ_Γ are the functions χ_u , where $u \in \mathbb{Z}_2^n$. The eigenvalue λ_u corresponding to χ_u (i.e., $\Phi_\Gamma \chi_u = \lambda_u \chi_u$) is given by*

$$\lambda_u = \sum_{w \in \Gamma} (-1)^{u \cdot w}.$$

Proof. Let $v \in \mathbb{Z}_2^n$. Then

$$\begin{aligned} \Phi_\Gamma \chi_u(v) &= \sum_{w \in \Gamma} \chi_u(v + w) \\ &= \sum_{w \in \Gamma} (-1)^{u \cdot (v+w)} \\ &= \left(\sum_{w \in \Gamma} (-1)^{u \cdot w} \right) (-1)^{u \cdot v} \\ &= \left(\sum_{w \in \Gamma} (-1)^{u \cdot w} \right) \chi_u(v). \end{aligned}$$

Hence

$$\Phi_\Gamma \chi_u = \left(\sum_{w \in \Gamma} (-1)^{u \cdot w} \right) \chi_u,$$

as desired. \square

Note that because the χ_u 's form a basis for \mathcal{V} by Lemma 2.1, it follows that Theorem 2.3 yields a complete set of eigenvalues and eigenvectors for Φ_Γ . Note also that the eigenvectors χ_u of Φ_Γ are independent of Γ ; only the eigenvalues depend on Γ .

Now we come to the payoff. Let $\Delta = \{\delta_1, \dots, \delta_n\}$, where δ_i is the i th unit coordinate vector (i.e., δ_i has a 1 in position i and 0's elsewhere). Note that the j th coordinate of δ_i is just δ_{ij} (the Kronecker delta), explaining our notation δ_i . Let $[\Phi_\Delta]$ denote the matrix of the linear transformation $\Phi_\Delta : \mathcal{V} \rightarrow \mathcal{V}$ with respect to the basis B_1 of \mathcal{V} given by (7).

2.4 Lemma. *We have $[\Phi_\Delta] = \mathbf{A}(C_n)$, the adjacency matrix of the n -cube.*

Proof. Let $v \in \mathbb{Z}_2^n$. We have

$$\begin{aligned} \Phi_\Delta f_u(v) &= \sum_{w \in \Delta} f_u(v+w) \\ &= \sum_{w \in \Delta} f_{u+w}(v), \end{aligned}$$

since $u = v + w$ if and only if $u + w = v$. There follows [why?]

$$\Phi_\Delta f_u = \sum_{w \in \Delta} f_{u+w}. \tag{9}$$

Equation (9) says that the (u, v) -entry of the matrix Φ_Δ is given by

$$(\Phi_\Delta)_{uv} = \begin{cases} 1, & \text{if } u + v \in \Delta \\ 0, & \text{otherwise.} \end{cases}$$

Now $u + v \in \Delta$ if and only if u and v differ in exactly one coordinate. This is just the condition for uv to be an edge of C_n , so the proof follows. \square

2.5 Corollary. *The eigenvectors E_u ($u \in \mathbb{Z}_2^n$) of $\mathbf{A}(C_n)$ (regarded as linear combinations of the vertices of C_n , i.e., of the elements of \mathbb{Z}_2^n) are given by*

$$E_u = \sum_{v \in \mathbb{Z}_2^n} (-1)^{u \cdot v} v. \quad (10)$$

The eigenvalue λ_u corresponding to the eigenvector E_u is given by

$$\lambda_u = n - 2\omega(u), \quad (11)$$

where $\omega(u)$ is the number of 1's in u . ($\omega(u)$ is called the Hamming weight or simply the weight of u .) Hence $\mathbf{A}(C_n)$ has $\binom{n}{i}$ eigenvalues equal to $n - 2i$, for each $0 \leq i \leq n$.

Proof. For any function $g \in \mathcal{V}$ we have by (8) that

$$g = \sum_v g(v) f_v.$$

Applying this equation to $g = \chi_u$ gives

$$\chi_u = \sum_v \chi_u(v) f_v = \sum_v (-1)^{u \cdot v} f_v. \quad (12)$$

Equation (12) expresses the eigenvector χ_u of Φ_Δ (or even Φ_Γ for any $\Gamma \subseteq \mathbb{Z}_2^n$) as a linear combination of the functions f_v . But Φ_Δ has the same matrix with respect to the basis of the f_v 's as $\mathbf{A}(C_n)$ has with respect to the vertices v of C_n . Hence the expansion of the eigenvectors of Φ_Δ in terms of the f_v 's has the same coefficients as the expansion of the eigenvectors of $\mathbf{A}(C_n)$ in terms of the v 's, so equation (10) follows.

According to Theorem 2.3 the eigenvalue λ_u corresponding to the eigenvector χ_u of Φ_Δ (or equivalently, the eigenvector E_u of $\mathbf{A}(C_n)$) is given by

$$\lambda_u = \sum_{w \in \Delta} (-1)^{u \cdot w}. \quad (13)$$

Now $\Delta = \{\delta_1, \dots, \delta_n\}$, and $\delta_i \cdot u$ is 1 if u has a one in its i th coordinate and is 0 otherwise. Hence the sum in (13) has $n - \omega(u)$ terms equal to +1 and $\omega(u)$ terms equal to -1, so $\lambda_u = (n - \omega(u)) - \omega(u) = n - 2\omega(u)$, as claimed. \square

We have all the information needed to count walks in C_n .

2.6 Corollary. *Let $u, v \in \mathbb{Z}_2^n$, and suppose that $\omega(u + v) = k$ (i.e., u and v disagree in exactly k coordinates). Then the number of walks of length ℓ in C_n between u and v is given by*

$$(\mathbf{A}^\ell)_{uv} = \frac{1}{2^n} \sum_{i=0}^n \sum_{j=0}^k (-1)^j \binom{k}{j} \binom{n-k}{i-j} (n-2i)^\ell, \quad (14)$$

where we set $\binom{n-k}{i-j} = 0$ if $j > i$. In particular,

$$(\mathbf{A}^\ell)_{uu} = \frac{1}{2^n} \sum_{i=0}^n \binom{n}{i} (n-2i)^\ell. \quad (15)$$

Proof. Let E_u and λ_u be as in Corollary 2.5. In order to apply Corollary 1.2, we need the eigenvectors to be of *unit* length (where we regard the f_v 's as an orthonormal basis of \mathcal{V}). By equation (10), we have

$$|E_u|^2 = \sum_{v \in \mathbb{Z}_2^n} ((-1)^{u \cdot v})^2 = 2^n.$$

Hence we should replace E_u by $E'_u = \frac{1}{2^{n/2}} E_u$ to get an orthonormal basis. According to Corollary 1.2, we thus have

$$(\mathbf{A}^\ell)_{uv} = \frac{1}{2^n} \sum_{w \in \mathbb{Z}_2^n} E_{uw} E_{vw} \lambda_w^\ell.$$

Now E_{uw} by definition is the coefficient of f_w in the expansion (10), i.e., $E_{uw} = (-1)^{u \cdot w}$ (and similarly for E_v), while $\lambda_w = n - 2\omega(w)$. Hence

$$(\mathbf{A}^\ell)_{uv} = \frac{1}{2^n} \sum_{w \in \mathbb{Z}_2^n} (-1)^{(u+v) \cdot w} (n - 2\omega(w))^\ell. \quad (16)$$

The number of vectors w of Hamming weight i which have j 1's in common with $u + v$ is $\binom{k}{j} \binom{n-k}{i-j}$, since we can choose the j 1's in $u + v$ which agree with w in $\binom{k}{j}$ ways, while the remaining $i - j$ 1's of w can be inserted in the

$n - k$ remaining positions in $\binom{n-k}{i-j}$ ways. Since $(u + v) \cdot w \equiv j \pmod{2}$, the sum (16) reduces to (14) as desired. Clearly setting $u = v$ in (14) yields (15), completing the proof. \square

It is possible to give a direct proof of (15) avoiding linear algebra. Thus by Corollary 1.3 and Lemma 1.7 (exactly as was done for K_n) we have another determination of the eigenvalues of C_n . With a little more work one can also obtain a direct proof of (14). Later in Example 9.9.12, however, we will use the eigenvalues of C_n to obtain a combinatorial result for which no nonalgebraic proof is known.

2.7 Example. Setting $k = 1$ in (14) yields

$$\begin{aligned} (\mathbf{A}^\ell)_{uv} &= \frac{1}{2^n} \sum_{i=0}^n \left[\binom{n-1}{i} - \binom{n-1}{i-1} \right] (n-2i)^\ell \\ &= \frac{1}{2^n} \sum_{i=0}^{n-1} \binom{n-1}{i} \frac{(n-2i)^{\ell+1}}{n-i}. \quad \square \end{aligned}$$

3 Random walks.

Let G be a finite graph. We consider a random walk on the vertices of G of the following type. Start at a vertex u . (The vertex u could be chosen randomly according to some probability distribution or could be specified in advance.) Among all the edges incident to u , choose one uniformly at random (i.e., if there are k edges incident to u , then each of these edges is chosen with probability $1/k$). Travel to the vertex v at the other end of the chosen edge and continue as before from v . Readers with some familiarity with probability theory will recognize this random walk as a special case of a finite state Markov chain. Many interesting questions may be asked about such walks; the basic one is to determine the probability of being at a given vertex after a given number ℓ of steps.

Suppose vertex u has *degree* d_u , i.e., there are d_u edges incident to u (counting loops at u once only). Let $\mathbf{M} = \mathbf{M}(G)$ be the matrix whose rows and columns are indexed by the vertex set $\{v_1, \dots, v_p\}$ of G , and whose (u, v) -entry is given by

$$M_{uv} = \frac{\mu_{uv}}{d_u},$$

where μ_{uv} is the number of edges between u and v (which for simple graphs will be 0 or 1). Thus M_{uv} is just the probability that if one starts at u , then the next step will be to v . An elementary probability theory argument (equivalent to Theorem 1.1) shows that if ℓ is a positive integer, then $(\mathbf{M}^\ell)_{uv}$ is equal to probability that one ends up at vertex v in ℓ steps given that one has started at u . Suppose now that the starting vertex is not specified, but rather we are given probabilities ρ_u summing to 1 and that we start at vertex u with probability ρ_u . Let P be the row vector $P = [\rho_{v_1}, \dots, \rho_{v_p}]$. Then again an elementary argument shows that if $P\mathbf{M}^\ell = [\sigma_{v_1}, \dots, \sigma_{v_p}]$, then σ_v is the probability of ending up at v in ℓ steps (with the given starting distribution). By reasoning as in Section 1, we see that if we know the eigenvalues and eigenvectors of \mathbf{M} , then we can compute the crucial probabilities $(\mathbf{M}^\ell)_{uv}$ and σ_u .

Since the matrix \mathbf{M} is not the same as the adjacency matrix \mathbf{A} , what does all this have to do with adjacency matrices? The answer is that in one important case \mathbf{M} is just a scalar multiple of \mathbf{A} . We say that the graph G

is *regular of degree d* if each $d_u = d$, i.e., each vertex is incident to d edges. In this case it's easy to see that $\mathbf{M}(G) = \frac{1}{d}\mathbf{A}(G)$. Hence the eigenvectors E_u of $\mathbf{M}(G)$ and $\mathbf{A}(G)$ are the same, and the eigenvalues are related by $\lambda_u(\mathbf{M}) = \frac{1}{d}\lambda_u(\mathbf{A})$. Thus random walks on a regular graph are closely related to the adjacency matrix of the graph.

3.1 Example. Consider a random walk on the n -cube C_n which begins at the “origin” (the vector $(0, \dots, 0)$). What is the probability p_ℓ that after ℓ steps one is again at the origin? Before applying any formulas, note that after an even (respectively, odd) number of steps, one must be at a vertex with an even (respectively, odd) number of 1's. Hence $p_\ell = 0$ if ℓ is odd. Now note that C_n is regular of degree n . Thus by (11), we have

$$\lambda_u(\mathbf{M}(C_n)) = \frac{1}{n}(n - 2\omega(u)).$$

By (15) we conclude that

$$p_\ell = \frac{1}{2^n n^\ell} \sum_{i=0}^n \binom{n}{i} (n - 2i)^\ell.$$

Note that the above expression for p_ℓ does indeed reduce to 0 when ℓ is odd.

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.

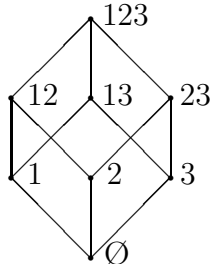
4.1 Definition. A *poset* (short for partially ordered set) P is a finite set, also denoted P , together with a binary relation denoted \leq satisfying the following axioms:

- (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 $S = \{1, 2, \dots, n\}$, 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) $\varphi : P \rightarrow Q$ such that $x \leq y$ in P if and only if $\varphi(x) \leq \varphi(y)$ 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, the 1338193159771 of order fourteen, the 68275077901156 of order fifteen, and the 4483130665195087 of order sixteen. 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 $x \leq y$ or $y \leq x$ in P . A finite chain is said to have *length* n if it has $n + 1$ elements. Such a chain thus has the form $x_0 < x_1 < \dots < x_n$. 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 $y_0 < y_1 < \dots < y_j$ 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

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

(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, \dots, p_n itself or the polynomial $F(q) = p_0 + p_1q + \cdots + p_nq^n$ 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 $\binom{n}{0}, \binom{n}{1}, \dots, \binom{n}{n}$ (the n th 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.

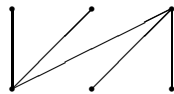
4.2 Definition. Let P be a graded poset of rank n . We say that P has the *Sperner property* or is a *Sperner poset* if

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

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 $\mu : P_i \rightarrow P_{i+1}$ satisfying $x < \mu(x)$ for all $x \in P_i$. Clearly if such an order-matching exists then $p_i \leq p_{i+1}$ (since μ is one-to-one). Easy examples show that the converse is false, i.e., if $p_i \leq p_{i+1}$ 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 $\mu : P_i \rightarrow P_{i-1}$ 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 $0 \leq j \leq n$ and order-matchings

$$P_0 \rightarrow P_1 \rightarrow P_2 \rightarrow \cdots \rightarrow P_j \leftarrow P_{j+1} \leftarrow P_{j+2} \leftarrow \cdots \leftarrow P_n. \quad (17)$$

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 $U : \mathbb{R}P_i \rightarrow \mathbb{R}P_{i+1}$ (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 $\mu : P_i \rightarrow P_{i+1}$.

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

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

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

Proof. Suppose $U : \mathbb{R}P_i \rightarrow \mathbb{R}P_{i+1}$ 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 rows of $[U]$ are indexed by the elements x_1, \dots, x_{p_i} of P_i (in some order) and the columns 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 rows). Since the row rank of a matrix equals its column rank, $[U]$ must have p_i linearly independent columns. Say we have labelled the elements of P_{i+1} so that the first p_i columns of $[U]$ are linearly independent.

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

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

where the sum is over all permutations π of $1, \dots, p_i$. Thus some term $\pm a_{1\pi(1)} \cdots a_{p_i\pi(p_i)}$ of the above sum is 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 $\mu : P_i \rightarrow P_{i+1}$ 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. \square

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 $U_i : \mathbb{R}(B_n)_i \rightarrow \mathbb{R}(B_n)_{i+1}$, 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

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

Note that since $(B_n)_i$ is a basis for $\mathbb{R}(B_n)_i$, equation (18) does indeed define a unique linear transformation $U_i : \mathbb{R}(B_n)_i \rightarrow \mathbb{R}(B_n)_{i+1}$. 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 $D_i : \mathbb{R}(B_n)_i \rightarrow \mathbb{R}(B_n)_{i-1}$ to the U_i ’s (D stands for “down”), defined by

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

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

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

i.e., the matrix $[D_{i+1}]$ is the transpose of the matrix $[U_i]$ [why?]. Now let $I_i : \mathbb{R}(B_n)_i \rightarrow \mathbb{R}(B_n)_i$ 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*

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

(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

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

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

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

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

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

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

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

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

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 $\emptyset = x_0 < x_1 < \cdots < x_n = \{1, \dots, n\}$ 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

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

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. \square

There is another nice way to show directly that B_n is Sperner, namely, by constructing an explicit order-matching $\mu : (B_n)_i \rightarrow (B_n)_{i+1}$ when $i < n/2$. We will define μ by giving an example. Let $n = 21$, $i = 9$, and $S = \{3, 4, 5, 8, 12, 13, 17, 19, 20\}$. We want to define $\mu(S)$. Let $(a_1, a_2, \dots, a_{21})$ be a sequence of ± 1 's, where $a_i = 1$ if $i \in S$, and $a_i = -1$ if $i \notin S$. For the set S above we get the sequence (writing $-$ for -1)

$$- - 111 - - 1 - - - 11 - - - 1 - 11 - .$$

Replace any two consecutive terms $1 -$ with 00 :

$$- - 1100 - 00 - - 100 - - 00100.$$

Ignore the 0's and replace any two consecutive terms $1 -$ with 00 :

$$- - 1000000 - - 0000 - 00100.$$

Continue:

$$- - 00000000 - 0000 - 00100.$$

At this stage no further replacement is possible. The nonzero terms consist of a sequence of $-$'s followed by a sequence of 1 's. There is at least one $-$ since $i < n/2$. Let k be the position (coordinate) of the last $-$; here $k = 16$. Define $\mu(S) = S \cup \{k\} = S \cup \{16\}$. The reader can check that this procedure

gives an order-matching. In particular, why is μ injective (one-to-one), i.e., why can we recover S from $\mu(S)$?

In view of the above elegant proof of Lubell and the explicit description of an order-matching $\mu : (B_n)_i \rightarrow (B_n)_{i+1}$, 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.