The Sandwich Theorem

Donald E. Knuth

Abstract: This report contains expository notes about a function $\vartheta(G)$ that is popularly known as the Lovász number of a graph G. There are many ways to define $\vartheta(G)$, and the surprising variety of different characterizations indicates in itself that $\vartheta(G)$ should be interesting. But the most interesting property of $\vartheta(G)$ is probably the fact that it can be computed efficiently, although it lies "sandwiched" between other classic graph numbers whose computation is NP-hard. I have tried to make these notes self-contained so that they might serve as an elementary introduction to the growing literature on Lovász's fascinating function.

The Sandwich Theorem DEK notes last revised 6 December 1993

The Sandwich Theorem

It is NP-complete to compute $\omega(G)$, the size of the largest clique in a graph G, and it is NP-complete to compute $\chi(G)$, the minimum number of colors needed to color the vertices of G. But Grötschel, Lovász, and Schrijver proved [5] that we can compute in polynomial time a real number that is "sandwiched" between these hard-to-compute integers:

$$\omega(G) \le \vartheta(\overline{G}) \le \chi(G)\,. \tag{$*$}$$

Lovász [13] called this a "sandwich theorem." The book [7] develops further facts about the function $\vartheta(G)$ and shows that it possesses many interesting properties. Therefore I think it's worthwhile to study $\vartheta(G)$ closely, in hopes of getting acquainted with it and finding faster ways to compute it.

Caution: The function called $\vartheta(G)$ in [13] is called $\vartheta(\overline{G})$ in [7] and [12]. I am following the latter convention because it is more likely to be adopted by other researchers—[7] is a classic book that contains complete proofs, while [13] is simply an extended abstract.

In these notes I am mostly following [7] and [12] with minor simplifications and a few additions. I mention several natural problems that I was not able to solve immediately although I expect (and fondly hope) that they will be resolved before I get to writing this portion of my forthcoming book on Combinatorial Algorithms. I'm grateful to many people—especially to Martin Grötschel and László Lovász—for their comments on my first drafts of this material.

These notes are in numbered sections, and there is at most one Lemma, Theorem, Corollary, or Example in each section. Thus, "Lemma 2" will mean "the lemma in section 2".

0. Preliminaries. Let's begin slowly by defining some notational conventions and by stating some basic things that will be assumed without proof. All vectors in these notes will be regarded as column vectors, indexed either by the vertices of a graph or by integers. The notation $x \ge y$, when x and y are vectors, will mean that $x_v \ge y_v$ for all v. If A is a matrix, A_v will denote column v, and A_{uv} will be the element in row u of column v. The zero vector and the zero matrix and zero itself will all be denoted by 0.

We will use several properties of matrices and vectors of real numbers that are familiar to everyone who works with linear algebra but not to everyone who studies graph theory, so it seems wise to list them here:

(i) The *dot product* of (column) vectors a and b is

$$a \cdot b = a^T b\,; \tag{0.1}$$

the vectors are *orthogonal* (also called perpendicular) if $a \cdot b = 0$. The *length* of vector a is

$$\|a\| = \sqrt{a \cdot a}\,. \tag{0.2}$$

Cauchy's inequality asserts that

$$a \cdot b \le \|a\|\,\|b\|\,; \tag{0.3}$$

equality holds iff a is a scalar multiple of b or $b = 0$. Notice that if A is any matrix we have

$$(A^T A)_{uv} = \sum_{k=1}^{n} (A^T)_{uk} A_{kv} = \sum_{k=1}^{n} A_{ku} A_{kv} = A_u \cdot A_v\,; \tag{0.4}$$

in other words, the elements of $A^T A$ represent all dot products of the columns of A.

(ii) An *orthogonal matrix* is a square matrix Q such that $Q^T Q$ is the identity matrix I. Thus, by (0.4), Q is orthogonal iff its columns are unit vectors perpendicular to each other. The transpose of an orthogonal matrix is orthogonal, because the condition $Q^T Q = I$ implies that Q^T is the inverse of Q, hence $QQ^T = I$.

(iii) A given matrix A is *symmetric* (i.e., $A = A^T$) iff it can be expressed in the form

$$A = QDQ^T \tag{0.5}$$

where Q is orthogonal and D is a diagonal matrix. Notice that (0.5) is equivalent to the matrix equation

$$AQ = QD\,, \tag{0.6}$$

which is equivalent to the equations

$$AQ_v = Q_v \lambda_v$$

for all v, where $\lambda_v = D_{vv}$. Hence the diagonal elements of D are the eigenvalues of A and the columns of Q are the corresponding eigenvectors.

Properties (i), (ii), and (iii) are proved in any textbook of linear algebra. We can get some practice using these concepts by giving a constructive proof of another well known fact:

Lemma. *Given k mutually perpendicular unit vectors, there is an orthogonal matrix having these vectors as the first k columns.*

Proof. Suppose first that $k = 1$ and that x is a d-dimensional vector with $\|x\| = 1$. If $x_1 = 1$ we have $x_2 = \cdots = x_d = 0$, so the orthogonal matrix $Q = I$ satisfies the desired condition. Otherwise we let

$$y_1 = \sqrt{(1 - x_1)/2}\,, \qquad y_j = -x_j/(2y_1) \quad \text{for } 1 < j \le d\,. \tag{0.7}$$

Then

$$y^T y = \|y\|^2 = y_1^2 + \frac{x_2^2 + \cdots + x_d^2}{4y_1^2} = \frac{1 - x_1}{2} + \frac{1 - x_1^2}{2(1 - x_1)} = 1\,.$$

And x is the first column of the Householder [8] matrix

$$Q = I - 2yy^T\,, \tag{0.8}$$

which is easily seen to be orthogonal because

$$Q^T Q = Q^2 = I - 4yy^T + 4yy^T yy^T = I\,.$$

Now suppose the lemma has been proved for some $k \geq 1$; we will show how to increase k by 1. Let Q be an orthogonal matrix and let x be a unit vector perpendicular to its first k columns. We want to construct an orthogonal matrix Q' agreeing with Q in columns 1 to k and having x in column $k+1$. Notice that

$$Q^T x = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ y \end{pmatrix}$$

by hypothesis, where there are 0s in the first k rows. The $(d-k)$-dimensional vector y has squared length

$$\|y\|^2 = Q^T x \cdot Q^T x = x^T Q Q^T x = x^T x = 1\,,$$

so it is a unit vector. (In particular, $y \neq 0$, so we must have $k < d$.) Using the construction above, we can find a $(d-k) \times (d-k)$ orthogonal matrix R with y in its first column. Then the matrix

$$Q' = Q \begin{pmatrix} 1 & & & \\ & \ddots & & 0 \\ & & 1 & \\ & 0 & & R \end{pmatrix}$$

does what we want. □

1. Orthogonal labelings. Let G be a graph on the vertices V. If u and v are distinct elements of V, the notation $u \text{ — } v$ means that they are adjacent in G; $u \not\text{—}\, v$ means they are not.

An assignment of vectors a_v to each vertex v is called an *orthogonal labeling* of G if $a_u \cdot a_v = 0$ whenever $u \not\text{—}\, v$. In other words, whenever a_u is not perpendicular to a_v in the labeling, we must have $u \text{ — } v$ in the graph. The vectors may have any desired dimension d; the components of a_v are a_{jv} for $1 \leq j \leq d$. Example: $a_v = 0$ for all v always works trivially.

The *cost* $c(a_v)$ of a vector a_v in an orthogonal labeling is defined to be 0 if $a_v = 0$, otherwise

$$c(a_v) = \frac{a_{1v}^2}{\|a_v\|^2} = \frac{a_{1v}^2}{a_{1v}^2 + \cdots + a_{dv}^2}\,.$$

Notice that we can multiply any vector a_v by a nonzero scalar t_v without changing its cost, and without violating the orthogonal labeling property. We can also get rid of a zero vector by increasing d by 1 and adding a new component 0 to each vector, except that the zero vector gets the new component 1. In particular, we can if we like assume that all vectors have unit length. Then the cost will be a_{1v}^2.

Lemma. *If $S \subseteq V$ is a stable set of vertices (i.e., no two vertices of S are adjacent) and if a is an orthogonal labeling then*

$$\sum_{v\in S} c(a_v) \le 1\,. \tag{1.1}$$

Proof. We can assume that $\|a_v\| = 1$ for all v. Then the vectors a_v for $v \in S$ must be mutually orthogonal, and Lemma 0 tells us we can find a $d \times d$ orthogonal matrix Q with these vectors as its leftmost columns. The sum of the costs will then be at most $q_{11}^2 + q_{12}^2 + \cdots + q_{1d}^2 = 1$. $\square$

Relation (1.1) makes it possible for us to study stable sets geometrically.

2. Convex labelings. An assignment x of real numbers x_v to the vertices v of G is called a *real labeling* of G. Several families of such labelings will be of importance to us:

The *characteristic labeling* for $U \subseteq V$ has $x_v = \begin{cases} 1 & \text{if } v \in U; \\ 0 & \text{if } v \notin U. \end{cases}$

A *stable labeling* is a characteristic labeling for a stable set.

A *clique labeling* is a characteristic labeling for a clique (a set of mutually adjacent vertices).

$\mathtt{STAB}(G)$ is the smallest convex set containing all stable labelings, i.e., $\mathtt{STAB}(G) = \text{convex hull}\,\{\, x \mid x \text{ is a stable labeling of } G\,\}$.

$\mathtt{QSTAB}(G) = \{\, x \ge 0 \mid \sum_{v\in Q} x_v \le 1 \text{ for all cliques } Q \text{ of } G\,\}\,.$

$\mathtt{TH}(G) = \{\, x \ge 0 \mid \sum_{v\in V} c(a_v)x_v \le 1 \text{ for all orthogonal labelings } a \text{ of } G\,\}\,.$

Lemma. TH *is sandwiched between* STAB *and* QSTAB*:*

$$\mathtt{STAB}(G) \subseteq \mathtt{TH}(G) \subseteq \mathtt{QSTAB}(G)\,. \tag{2.1}$$

Proof. Relation (1.1) tells that every stable labeling belongs to $\mathtt{TH}(G)$. Since $\mathtt{TH}(G)$ is obviously convex, it must contain the convex hull $\mathtt{STAB}(G)$. On the other hand, every clique labeling is an orthogonal labeling of dimension 1. Therefore every constraint of QSTAB(G) is one of the constraints of TH(G). $\square$

Note: QSTAB first defined by Shannon [18], and the first systematic study of STAB was undertaken by Padberg [17]. TH was first defined by Grötschel, Lovász, and Schrijver in [6].

3. Monotonicity. Suppose G and G' are graphs on the same vertex set V, with $G \subseteq G'$ (i.e., u — v in G implies u — v in G'). Then

every stable set in G' is stable in G, hence STAB$(G) \supseteq$ STAB(G');

every clique in G is a clique in G', hence QSTAB$(G) \supseteq$ QSTAB(G');

every orthogonal labeling of G is an orthogonal labeling of G',

hence TH$(G) \supseteq$ TH(G').

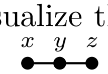
In particular, if G is the empty graph $\overline{K}_n$ on $|V| = n$ vertices, all sets are stable and all cliques have size ≤ 1, hence

$$\texttt{STAB}(\overline{K}_n) = \texttt{TH}(\overline{K}_n) = \texttt{QSTAB}(\overline{K}_n) = \{\, x \mid 0 \le x_v \le 1 \text{ for all } v \,\}, \text{ the } n\text{-cube.}$$

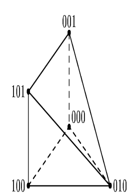
If G is the complete graph K_n, all stable sets have size ≤ 1 and there is an n-clique, so

$$\texttt{STAB}(K_n) = \texttt{TH}(K_n) = \texttt{QSTAB}(K_n) = \{\, x \ge 0 \mid \textstyle\sum_v x_v \le 1 \,\}, \text{ the } n\text{-simplex.}$$

Thus all the convex sets STAB(G), TH(G), QSTAB(G) lie between the n-simplex and the n-cube.

Consider, for example, the case $n = 3$. Then there are three coordinates, so we can visualize the sets in 3-space (although there aren't many interesting graphs). The QSTAB of 

is obtained from the unit cube by restricting the coordinates to $x + y \le 1$ and $y + z \le 1$; we can think of making two cuts in a piece of cheese:



The vertices $\{000, 100, 010, 001, 101\}$ correspond to the stable labelings, so once again we have STAB$(G) =$ TH$(G) =$ QSTAB(G).

4. The theta function. The function $\vartheta(G)$ mentioned in the introduction is a special case of a two-parameter function $\vartheta(G, w)$, where w is a nonnegative real labeling:

$$\vartheta(G, w) = \max\{\, w \cdot x \mid x \in \texttt{TH}(G) \,\}\,; \tag{4.1}$$

$$\vartheta(G) = \vartheta(G, \mathbb{1}) \text{ where } \mathbb{1} \text{ is the labeling } w_v = 1 \text{ for all } v. \tag{4.2}$$

This function, called the *Lovász number* of G (or the *weighted Lovász number* when $w \neq \mathbb{1}$), tells us about 1-dimensional projections of the n-dimensional convex set $\mathtt{TH}(G)$.

Notice, for example, that the monotonicity properties of §3 tell us

$$G \subseteq G' \;\Rightarrow\; \vartheta(G,w) \geq \vartheta(G',w) \tag{4.3}$$

for all $w \geq 0$. It is also obvious that ϑ is monotone in its other parameter:

$$w \leq w' \;\Rightarrow\; \vartheta(G,w) \leq \vartheta(G,w')\,. \tag{4.4}$$

The smallest possible value of ϑ is

$$\vartheta(K_n,w) = \max\{w_1,\ldots,w_n\}\,; \quad \vartheta(K_n) = 1\,. \tag{4.5}$$

The largest possible value is

$$\vartheta(\overline{K}_n,w) = w_1 + \cdots + w_n\,; \quad \vartheta(\overline{K}_n) = n\,. \tag{4.6}$$

Similar definitions can be given for STAB and QSTAB:

$$\alpha(G,w) = \max\{\, w\cdot x \mid x \in \mathtt{STAB}(G)\,\}\,, \qquad \alpha(G) = \alpha(G,\mathbb{1})\,; \tag{4.7}$$

$$\kappa(G,w) = \max\{\, w\cdot x \mid x \in \mathtt{QSTAB}(G)\,\}\,, \qquad \kappa(G) = \kappa(G,\mathbb{1})\,. \tag{4.8}$$

Clearly $\alpha(G)$ is the size of the largest stable set in G, because every stable labeling x corresponds to a stable set with $\mathbb{1}\cdot x$ vertices. It is also easy to see that $\kappa(G)$ is at most $\overline{\chi}(G)$, the smallest number of cliques that cover the vertices of G. For if the vertices can be partitioned into k cliques $Q_1,\ldots,Q_k$ and if $x \in \mathtt{QSTAB}(G)$, we have

$$\mathbb{1}\cdot x = \sum_{v\in Q_1} x_v + \cdots + \sum_{v\in Q_k} x_v \leq k\,.$$

Sometimes $\kappa(G)$ is less than $\overline{\chi}(G)$. For example, consider the cyclic graph C_n, with vertices $\{0,1,\ldots,n-1\}$ and $u \mathrel{—} v$ iff $u \equiv v \pm 1 \pmod 1$. Adding up the inequalities $x_0 + x_1 \leq 1, \ldots, x_{n-2} + x_{n-1} \leq 1, x_{n-1} + x_0 \leq 1$ of QSTAB gives $2(x_0 + \cdots + x_{n-1}) \leq n$, and this upper bound is achieved when all x's are $\frac{1}{2}$; hence $\kappa(C_n) = \frac{n}{2}$, if $n > 3$. But $\overline{\chi}(G)$ is always an integer, and $\overline{\chi}(C_n) = \lceil \frac{n}{2} \rceil$ is greater than $\kappa(C_n)$ when n is odd.

Incidentally, these remarks establish the "sandwich inequality" $(*)$ stated in the introduction, because

$$\alpha(G) \leq \vartheta(G) \leq \kappa(G) \leq \overline{\chi}(G) \tag{4.9}$$

and $\omega(G) = \alpha(\overline{G})$, $\overline{\chi}(G) = \chi(\overline{G})$.

5. Alternative definitions of ϑ. Four additional functions ϑ_1, ϑ_2, ϑ_3, ϑ_4 are defined in [7], and they all turn out to be identical to ϑ. Thus, we can understand ϑ in many different ways; this may help us compute it.

We will show, following [7], that if w is any fixed nonnegative real labeling of G, the inequalities

$$\vartheta(G,w) \le \vartheta_1(G,w) \le \vartheta_2(G,w) \le \vartheta_3(G,w) \le \vartheta_4(G,w) \le \vartheta(G,w) \tag{5.1}$$

can be proved. Thus we will establish the theorem of [7], and all inequalities in our proofs will turn out to be equalities. We will introduce the alternative definitions ϑ_k one at a time; any one of these definitions could have been taken as the starting point. First,

$$\vartheta_1(G,w) = \min_a \, \max_v \bigl(w_v/c(a_v)\bigr)\,, \text{ over all orthogonal labelings } a. \tag{5.2}$$

Here we regard $w_v/c(a_v) = 0$ when $w_v = c(a_v) = 0$; but the max is ∞ if some $w_v > 0$ has $c(a_v) = 0$.

Lemma. $\vartheta(G,w) \le \vartheta_1(G,w)$.

Proof. Suppose $x \in \mathtt{TH}(G)$ maximizes $w \cdot x$, and suppose a is an orthogonal labeling that achieves the minimum value $\vartheta_1(G,w)$. Then

$$\vartheta(G,w) = w \cdot x = \sum_v w_v x_v \le \left(\max_v \frac{w_v}{c(a_v)}\right) \sum_v c(a_v) x_v \le \max_v \frac{w_v}{c(a_v)} = \vartheta_1(G,w)\,. \ \Box$$

Incidentally, the fact that all inequalities are exact will imply later that every nonzero weight vector w has an orthogonal labeling a such that

$$c(a_v) = \frac{w_v}{\vartheta(G,w)} \quad \text{for all } v. \tag{5.3}$$

We will restate such consequences of (5.1) later, but it may be helpful to keep that future goal in mind.

6. Characterization via eigenvalues. The second variant of ϑ is rather different; this is the only one Lovász chose to mention in [13].

We say that A is a *feasible matrix* for G and w if A is indexed by vertices and

$$\begin{aligned} &A \text{ is real and symmetric;} \\ &A_{vv} = w_v \text{ for all } v \in V; \\ &A_{uv} = \sqrt{w_u w_v} \text{ whenever } u \not\to v \text{ in } G \end{aligned} \tag{6.1}$$

The other elements of A are unconstrained (i.e., they can be anything between $-\infty$ and $+\infty$).

If A is any real, symmetric matrix, let $\Lambda(A)$ be its maximum eigenvalue. This is well defined because all eigenvalues of A are real. Suppose A has eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$; then $A = Q\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n)Q^T$ for some orthogonal Q, and $\|Qx\| = \|x\|$ for all vectors x, so there is a nice way to characterize $\Lambda(A)$:

$$\Lambda(A) = \max\{\, x^T A x \mid \|x\| = 1 \,\}\,. \tag{6.2}$$

Notice that $\Lambda(A)$ might not be the largest eigenvalue in absolute value. We now let

$$\vartheta_2(G, w) = \min\{\, \Lambda(A) \mid A \text{ is a feasible matrix for } G \text{ and } w \,\}\,. \tag{6.3}$$

Lemma. $\vartheta_1(G, w) \le \vartheta_2(G, w)$.

Proof. Note first that the trace $\mathrm{tr}\, A = \sum_v w_v \ge 0$ for any feasible matrix A. The trace is also well-known to be the sum of the eigenvalues; this fact is an easy consequence of the identity

$$\mathrm{tr}\, XY = \sum_{j=1}^{m} \sum_{k=1}^{n} x_{jk} y_{kj} = \mathrm{tr}\, YX \tag{6.4}$$

valid for any matrices X and Y of respective sizes $m \times n$ and $n \times m$. In particular, $\vartheta_2(G, w)$ is always ≥ 0, and it is $= 0$ if and only if $w = 0$ $\big($when also $\vartheta_1(G, w) = 0\big)$.

So suppose $w \ne 0$ and let A be a feasible matrix that attains the minimum value $\Lambda(A) = \vartheta_2(G, w) = \lambda > 0$. Let

$$B = \lambda I - A\,. \tag{6.5}$$

The eigenvalues of B are λ minus the eigenvalues of A. $\big($For if $A = Q\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n)Q^T$ then $B = Q\,\mathrm{diag}(\lambda - \lambda_1, \ldots, \lambda - \lambda_n)Q^T$.$\big)$ Thus they are all nonnegative; such a matrix B is called *positive semidefinite*. By (0.5) we can write

$$B = X^T X\,, \qquad \text{i.e.,} \quad B_{uv} = x_u \cdot x_v\,, \tag{6.6}$$

when $X = \mathrm{diag}(\sqrt{\lambda - \lambda_1}, \ldots, \sqrt{\lambda - \lambda_n}\,)Q^T$.

Let $a_v = (\sqrt{w_v}\,, x_{1v}, \ldots, x_{rv})^T$. Then $c(a_v) = w_v / \|a_v\|^2 = w_v/(w_v + x_{1v}^2 + \cdots + x_{rv}^2)$ and $x_{1v}^2 + \cdots + x_{rv}^2 = B_{vv} = \lambda - w_v$, hence $c(a_v) = w_v/\lambda$. Also if $u \not\longrightarrow v$ we have $a_u \cdot a_v = \sqrt{w_u w_v} + x_u \cdot x_v = \sqrt{w_u w_v} + B_{uv} = \sqrt{w_u w_v} - A_{uv} = 0$. Therefore a is an orthogonal labeling and $\max_v w_v/c(a_v) = \lambda \ge \vartheta_1(G, w)$. □

7. A complementary characterization. Still another variation is based on orthogonal labelings of the complementary graph $\overline{G}$.

In this case we let b be an orthogonal labeling of $\overline{G}$, normalized so that $\sum_v \|b_v\|^2 = 1$, and we let

$$\vartheta_3(G,w) = \max\left\{ \sum_{u,v} (\sqrt{w_u}\, b_u)\cdot(\sqrt{w_v}\, b_v) \;\middle|\; b \text{ is a normalized orthogonal labeling of } \overline{G} \right\}. \tag{7.1}$$

A normalized orthogonal labeling b is equivalent to an $n\times n$ symmetric positive semidefinite matrix B, where $B_{uv} = b_u \cdot b_v$ is zero when $u \frown v$, and where $\operatorname{tr} B = 1$.

Lemma. $\vartheta_2(G,w) \le \vartheta_3(G,w)$.

This lemma is the "heart" of the proof that all ϑs are equivalent, according to [7]. It relies on a fact about positive semidefinite matrices that we will prove in §9.

Fact. *If A is a symmetric matrix such that $A\cdot B \ge 0$ for all symmetric positive semidefinite B with $B_{uv} = 0$ for $u \frown v$, then $A = X + Y$ where X is symmetric positive semidefinite and Y is symmetric and $Y_{vv} = 0$ for all v and $Y_{uv} = 0$ for $u \not\frown v$.*

Here $C\cdot B$ stands for the dot product of matrices, i.e., the sum $\sum_{u,v} C_{uv}B_{uv}$, which can also be written $\operatorname{tr} C^T B$. The stated fact is a duality principle for quadratic programming.

Assuming the Fact, let W be the matrix with $W_{uv} = \sqrt{w_u w_v}$, and let $\vartheta_3 = \vartheta_3(G,w)$. By definition (7.1), if b is any nonzero orthogonal labeling of $\overline{G}$ (not necessarily normalized), we have

$$\sum_{u,v} (\sqrt{w_u}\, b_u)\cdot(\sqrt{w_v}\, b_v) \le \vartheta_3 \sum_v \|b_v\|^2\,. \tag{7.2}$$

In matrix terms this says $W\cdot B \le (\vartheta_3 I)\cdot B$ for all symmetric positive semidefinite B with $B_{uv} = 0$ for $u \frown v$. The Fact now tells us we can write

$$\vartheta_3 I - W = X + Y \tag{7.3}$$

where X is symmetric positive semidefinite, Y is symmetric and diagonally zero, and $Y_{uv} = 0$ when $u \not\frown v$. Therefore the matrix A defined by

$$A = W + Y = \vartheta_3 I - X$$

is a feasible matrix for G, and $\Lambda(A) \le \vartheta_3$. This completes the proof that $\vartheta_2(G,w) \le \vartheta_3(G,w)$, because $\Lambda(A)$ is an upper bound on ϑ_2 by definition of ϑ_2. $\square$

8. Elementary facts about cones. A *cone* in N-dimensional space is a set of vectors closed under addition and under multiplication by nonnegative scalars. (In particular, it is convex: If c and c' are in cone C and $0 < t < 1$ then tc and $(1-t)c'$ are in C, hence $tc + (1-t)c' \in C$.) A *closed cone* is a cone that is also closed under taking limits.

F1. If C is a closed convex set and $x \notin C$, there is a hyperplane separating x from C. This means there is a vector y and a number b such that $c \cdot y \le b$ for all $c \in C$ but $x \cdot y > b$.

Proof. Let d be the greatest lower bound of $\|x - c\|^2$ for all $c \in C$. Then there's a sequence of vectors c_k with $\|x - c_k\|^2 < d + 1/k$; this infinite set of vectors contained in the sphere $\{\, y \mid \|x - y\|^2 \le d + 1 \,\}$ must have a limit point c_∞, and $c_\infty \in C$ since C is closed. Therefore $\|x - c_\infty\|^2 \ge d$; in fact $\|x - c_\infty\|^2 = d$, since $\|x - c_\infty\| \le \|x - c_k\| + \|c_k - c_\infty\|$ and the right-hand side can be made arbitrarily close to d. Since $x \notin C$, we must have $d > 0$. Now let $y = x - c_\infty$ and $b = c_\infty \cdot y$. Clearly $x \cdot y = y \cdot y + b > b$. And if c is any element of C and ϵ is any small positive number, the vector $\epsilon c + (1-\epsilon)c_\infty$ is in C, hence $\big\|x - \big(\epsilon c + (1-\epsilon)c_\infty\big)\big\|^2 \ge d$. But

$$\begin{aligned}\big\|x - \big(\epsilon c + (1-\epsilon)c_\infty\big)\big\|^2 - d &= \|x - c_\infty - \epsilon(c - c_\infty)\|^2 - d \\ &= -2\epsilon y \cdot (c - c_\infty) + \epsilon^2 \|c - c_\infty\|^2\end{aligned}$$

can be nonnegative for all small ϵ only if $y \cdot (c - c_\infty) \le 0$, i.e., $c \cdot y \le b$. $\square$

If A is any set of vectors, let $A^* = \{\, b \mid a \cdot b \ge 0 \text{ for all } a \in A \,\}$.

The following facts are immediate:

F2. $A \subseteq A'$ implies $A^* \supseteq A'^*$.

F3. $A \subseteq A^{**}$.

F4. A^* is a closed cone.

From F1 we also get a result which, in the special case that $C = \{\, Ax \mid x \ge 0 \,\}$ for a matrix A, is called Farkas's Lemma:

F5. If C is a closed cone, $C = C^{**}$.

Proof. Suppose $x \in C^{**}$ and $x \notin C$, and let (y, b) be a separating hyperplane as in F1. Then $(y, 0)$ is also a separating hyperplane; for we have $x \cdot y > b \ge 0$ because $0 \in C$, and we cannot have $c \cdot y > 0$ for $c \in C$ because $(\lambda c) \cdot y$ would then be unbounded. But then $c \cdot (-y) \ge 0$ for all $c \in C$, so $-y \in C^*$; hence $x \cdot (-y) \ge 0$, a contradiction. $\square$

If A and B are sets of vectors, we define $A + B = \{\, a + b \mid a \in A \text{ and } b \in B \,\}$.

F6. If C and C' are closed cones, $(C \cap C')^* = C^* + C'^*$.

Proof. If A and B are arbitrary sets we have $A^* + B^* \subseteq (A \cap B)^*$, for if $x \in A^* + B^*$ and $y \in A \cap B$ then $x \cdot y = a \cdot y + b \cdot y \geq 0$. If A and B are arbitrary sets including 0 then $(A + B)^* \subseteq A^* \cap B^*$ by F2, because $A + B \supseteq A$ and $A + B \supseteq B$. Thus for arbitrary A and B we have $(A^* + B^*)^* \subseteq A^{**} \cap B^{**}$, hence

$$(A^* + B^*)^{**} \supseteq (A^{**} \cap B^{**})^* .$$

Now let A and B be closed cones; apply F5 to get $A^* + B^* \supseteq (A \cap B)^*$. $\square$

F7. If C and C' are closed cones, $(C + C')^* = C^* \cap C'^*$. (I don't need this but I might as well state it.) **Proof.** F6 says $(C^* \cap C'^*)^* = C^{**} + C'^{**}$; apply F5 and $*$ again. $\square$

F8. Let S be any set of indices and let $A_S = \{\, a \mid a_s = 0 \text{ for all } s \in S \,\}$, and let $\overline{S}$ be all the indices not in S. Then

$$A_S^* = A_{\overline{S}} .$$

Proof. If $b_s = 0$ for all $s \notin S$ and $a_s = 0$ for all $s \in S$, obviously $a \cdot b = 0$; so $A_{\overline{S}} \subseteq A_S^*$. If $b_s \neq 0$ for some $s \notin S$ and $a_t = 0$ for all $t \neq s$ and $a_s = -b_s$ then $a \in A_S$ and $a \cdot b < 0$; so $b \notin A_S^*$, hence $A_{\overline{S}} \supseteq A_S^*$. $\square$

9. Definite proof of a semidefinite fact. Now we are almost ready to prove the result needed in the proof of Lemma 7.

Let D be the set of real symmetric positive semidefinite matrices (called "spuds" henceforth for brevity), considered as vectors in N-dimensional space where $N = \frac{1}{2}(n+1)n$. We use the inner product $A \cdot B = \operatorname{tr} A^T B$; this is justified if we divide off-diagonal elements by $\sqrt{2}$. For example, if $n = 3$ the correspondence between 6-dimensional vectors and 3×3 symmetric matrices is

$$(a, b, c, d, e, f) \;\leftrightarrow\; \begin{pmatrix} a & d/\sqrt{2} & e/\sqrt{2} \\ d/\sqrt{2} & b & f/\sqrt{2} \\ e/\sqrt{2} & f/\sqrt{2} & c \end{pmatrix}$$

preserving sum, scalar product, and dot product. Clearly D is a closed cone.

F9. $D^* = D$.

Proof. If A and B are spuds then $A = X^T X$ and $B = Y^T Y$ and $A \cdot B = \operatorname{tr} X^T X\, Y^T Y = \operatorname{tr} X Y^T Y X^T = (Y X^T) \cdot (Y X^T) \geq 0$; hence $D \subseteq D^*$. (In fact, this argument shows that $A \cdot B = 0$ iff $AB = 0$, for any spuds A and B, since $A = A^T$.)

If A is symmetric but has a negative eigenvalue λ we can write

$$A = Q\,\mathrm{diag}\,(\lambda, \lambda_2, \ldots, \lambda_n)\,Q^T$$

for some orthogonal matrix Q. Let $B = Q\,\mathrm{diag}\,(1, 0, \ldots, 0)Q^T$; then B is a spud, and

$$A \cdot B = \mathrm{tr}\; A^T B = \mathrm{tr}\; Q\,\mathrm{diag}\,(\lambda, 0, \ldots, 0)\,Q^T = \lambda < 0\,.$$

So A is not in D^*; this proves $D \supseteq D^*$. □

Let E be the set of all real symmetric matrices such that $E_{uv} = 0$ when $u \,—\, v$ in a graph G; let F be the set of all real symmetric matrices such that $F_{uv} = 0$ when $u = v$ or $u \not— v$. The Fact stated in Section 7 is now equivalent in our new notation to

Fact. $(D \cap E)^* \subseteq D + F$.

But we know that

$$\begin{aligned} (D \cap E)^* &= D^* + E^* && \text{by F6} \\ &= D + F && \text{by F9 and F8. } \square \end{aligned}$$

10. Another characterization. Remember ϑ, ϑ_1, ϑ_2, and ϑ_3? We are now going to introduce yet another function

$$\vartheta_4(G, w) = \max \left\{ \sum_v c(b_v) w_v \;\middle|\; b \text{ is an orthogonal labeling of } \overline{G} \right\}. \tag{10.1}$$

Lemma. $\vartheta_3(G, w) \le \vartheta_4(G, w)$.

Proof. Suppose b is a normalized orthogonal labeling of $\overline{G}$ that achieves the maximum ϑ_3; and suppose the vectors of this labeling have dimension d. Let

$$x_k = \sum_v b_{kv} \sqrt{w_v}\,, \qquad \text{for } 1 \le k \le d\,; \tag{10.2}$$

then

$$\vartheta_3(G, w) = \sum_{u,v} \sqrt{w_u}\, b_u \cdot b_v \sqrt{w_v} = \sum_{u,v,k} \sqrt{w_u w_v}\, b_{ku} b_{kv} = \sum_k x_k^2\,.$$

Let Q be an orthogonal $d \times d$ matrix whose first row is $(x_1/\sqrt{\vartheta_3}, \ldots, x_d/\sqrt{\vartheta_3})^T$, and let $b'_v = Q b_v$. Then $b'_u \cdot b'_v = b_u^T Q^T Q b_v = b_u^T b_v = b_u \cdot b_v$, so b' is a normalized orthogonal labeling of $\overline{G}$. Also

$$\begin{aligned} x'_k = \sum_v b'_{kv} \sqrt{w_v} &= \sum_{v,j} Q_{kj} b_{jv} \sqrt{w_v} \\ &= \sum_j Q_{kj} x_j = \begin{cases} \sqrt{\vartheta_3}\,, & k = 1; \\ 0\,, & k > 1. \end{cases} \end{aligned} \tag{10.3}$$

Hence by Cauchy's inequality

$$\vartheta_3(G,w) = \Bigl(\sum_v b'_{1v}\sqrt{w_v}\Bigr)^2 \le \Bigl(\sum_v \|b'_v\|^2\Bigr)\Bigl(\sum_{\substack{v\\ b'_v\ne 0}} \frac{b'^2_{1v}}{\|b'_v\|^2}\, w_v\Bigr)$$
$$= \sum_v c(b'_v)w_v \le \vartheta_4(G,w) \qquad (10.4)$$

because $\sum_v \|b'_v\|^2 = \sum_v \|b_v\|^2 = 1$. □

11. The final link. Now we can close the loop:

Lemma. $\vartheta_4(G,w) \le \vartheta(G,w)$.

Proof. If b is an orthogonal labeling of $\overline{G}$ that achieves the maximum ϑ_4, we will show that the real labeling x defined by $x_v = c(b_v)$ is in $\mathtt{TH}(G)$. Therefore $\vartheta_4(G,w) = w\cdot x$ is $\le \vartheta(G,w)$.

We will prove that if a is any orthogonal labeling of G, and if b is any orthogonal labeling of $\overline{G}$, then

$$\sum_v c(a_v)c(b_v) \le 1\,. \qquad (11.1)$$

Suppose a is a labeling of dimension d and b is of dimension d'. Then consider the $d\times d'$ matrices

$$A_v = a_v b_v^T \qquad (11.2)$$

as elements of a vector space of dimension dd'. If $u \ne v$ we have

$$A_u\cdot A_v = \operatorname{tr} A_u^T A_v = \operatorname{tr} b_u a_u^T a_v b_v^T = \operatorname{tr} a_u^T a_v b_v^T b_u = 0\,, \qquad (11.3)$$

because $a_u^T a_v = 0$ when $u \not\!\!-\, v$ and $b_v^T b_u = 0$ when $u \,-\, v$. If $u = v$ we have

$$A_v\cdot A_v = \|a_v\|^2\,\|b_v\|^2\,.$$

The upper left corner element of A_v is $a_{1v}b_{1v}$, hence the "cost" of A_v is $(a_{1v}b_{1v})^2/\|A_v\|^2 = c(a_v)c(b_v)$. This, with (11.3), proves (11.1). (See the proof of Lemma 1.) □

12. The main theorem. Lemmas 5, 6, 7, 10, and 11 establish the five inequalities claimed in (5.1); hence all five variants of ϑ are the same function of G and w. Moreover, all the inequalities in those five proofs are equalities (with the exception of (11.1)). We can summarize the results as follows.

Theorem. *For all graphs G and any nonnegative real labeling w of G we have*

$$\vartheta(G,w)=\vartheta_1(G,w)=\vartheta_2(G,w)=\vartheta_3(G,w)=\vartheta_4(G,w)\,. \tag{12.1}$$

Moreover, if $w\neq 0$, there exist orthogonal labelings a and b of G and $\overline{G}$, respectively, such that

$$c(a_v)=w_v/\vartheta\,; \tag{12.2}$$

$$\sum c(a_v)c(b_v)=1\,. \tag{12.3}$$

Proof. Relation (12.1) is, of course, (5.1); and (12.2) is (5.3). The desired labeling b is what we called b' in the proof of Lemma 10. The fact that the application of Cauchy's inequality in (10.4) is actually an equality,

$$\vartheta=\Bigl(\sum_v b_{1v}\sqrt{w_v}\Bigr)^2=\Bigl(\sum_v \|b_v\|^2\Bigr)\Bigl(\sum_{\substack{v\\ b_v\neq 0}}\frac{b_{1v}^2}{\|b_v\|^2}\,w_v\Bigr)\,, \tag{12.4}$$

tells us that the vectors whose dot product has been squared are proportional: There is a number t such that

$$\|b_v\|=t\,\frac{b_{1v}\sqrt{w_v}}{\|b_v\|}\,,\quad \text{if } b_v\neq 0\,;\qquad \|b_v\|=0\quad\text{iff}\quad b_{1v}\sqrt{w_v}=0\,. \tag{12.5}$$

The labeling in the proof of Lemma 10 also satisfies

$$\sum_v \|b_v\|^2=1\,; \tag{12.6}$$

hence $t=\pm 1/\sqrt{\vartheta}$.

We can now show

$$c(b_v)=\|b_v\|^2\,\vartheta/w_v\,,\quad \text{when } w_v\neq 0\,. \tag{12.7}$$

This relation is obvious if $\|b_v\|=0$; otherwise we have

$$c(b_v)=\frac{b_{1v}^2}{\|b_v\|^2}=\frac{\|b_v\|^2}{t^2w_v}$$

by (12.5). Summing the product of (12.2) and (12.7) over v gives (12.3). $\square$

13. The main converse. The nice thing about Theorem 12 is that conditions (12.2) and (12.3) also provide a *certificate* that a given value ϑ is the minimum or maximum stated in the definitions of ϑ, ϑ_1, ϑ_2, ϑ_3, and ϑ_4.

Theorem. *If a is an orthogonal labeling of G and b is an orthogonal labeling of $\overline{G}$ such that relations (12.2) and (12.3) hold for some ϑ and w, then ϑ is the value of $\vartheta(G,w)$.*

Proof. Plugging (12.2) into (12.3) gives $\sum w_v c(b_v) = \vartheta$, hence $\vartheta \le \vartheta_4(G,w)$ by definition of ϑ_4. Also,

$$\max_v \frac{w_v}{c(a_v)} = \vartheta\,,$$

hence $\vartheta \ge \vartheta_1(G,w)$ by definition of ϑ_1. $\square$

14. Another look at TH. We originally defined $\vartheta(G,w)$ in (4.1) in terms of the convex set TH defined in section 2:

$$\vartheta(G,w) = \max\{\, w\cdot x \mid x \in \mathtt{TH}(G)\,\}\,, \qquad \text{when } w \ge 0\,. \tag{14.1}$$

We can also go the other way, defining TH in terms of ϑ:

$$\mathtt{TH}(G) = \{\, x \ge 0 \mid w\cdot x \le \vartheta(G,w) \text{ for all } w \ge 0\,\}\,. \tag{14.2}$$

Every $x \in \mathtt{TH}(G)$ belongs to the right-hand set, by (14.1). Conversely, if x belongs to the right-hand set and if a is any orthogonal labeling of G, not entirely zero, let $w_v = c(a_v)$, so that $w\cdot x = \sum_v c(a_v)x_v$. Then

$$\vartheta_1(G,w) \le \max_v\bigl(w_v/c(a_v)\bigr) = 1$$

by definition (5.2), so we know by Lemma 5 that $\sum c(a_v)x_v \le 1$. This proves that x belongs to $\mathtt{TH}(G)$.

Theorem 12 tells us even more.

Lemma. $\mathtt{TH}(G) = \{\, x \ge 0 \mid \vartheta(\overline{G},x) \le 1\,\}$.

Proof. By definition (10.1),

$$\vartheta_4(\overline{G},w) = \max\left\{\, \sum_v c(a_v)w_v \mid a \text{ is an orthogonal labeling of } G \,\right\}. \tag{14.3}$$

Thus $x \in \mathtt{TH}(G)$ iff $\vartheta_4(\overline{G},x) \le 1$, when $x \ge 0$. $\square$

Theorem. $\mathtt{TH}(G) = \{\, x \mid x_v = c(b_v)$ *for some orthogonal labeling b of $\overline{G}$* $\}$.

Proof. We already proved in (11.1) that the right side is contained in the left.

Let $x \in \mathtt{TH}(G)$ and let $\vartheta = \vartheta(\overline{G},x)$. By the lemma, $\vartheta \le 1$. Therefore, by (12.2), there is an orthogonal labeling b of $\overline{G}$ such that $c(b_v) = x_v/\vartheta \ge x_v$ for all v. It is easy to reduce the cost of any vector in an orthogonal labeling to any desired value, simply by increasing the dimension and giving this vector an appropriate nonzero value in the new component while all other vectors remain zero there. The dot products are unchanged, so the new labeling is still orthogonal. Repeating this construction for each v produces a labeling with $c(b_v) = x_v$. $\square$

This theorem makes the definition of ϑ_4 in (10.1) identical to the definition of ϑ in (4.1).

15. Zero weights. Our next result shows that when a weight is zero, the corresponding vertex might as well be absent from the graph.

Lemma. *Let U be a subset of the vertices V of a graph G, and let $G' = G|U$ be the graph induced by U (i.e., the graph on vertices U with u — v in G' iff u — v in G). Then if w and w' are nonnegative labelings of G and G' such that*

$$w_v = w'_v \qquad \text{when} \quad v \in U\,, \qquad w_v = 0 \qquad \text{when} \quad v \notin U\,, \tag{15.1}$$

we have

$$\vartheta(G, w) = \vartheta(G', w')\,. \tag{15.2}$$

Proof. Let a and b satisfy (12.2) and (12.3) for G and w. Then $c(a_v) = 0$ for $v \notin U$, so $a|U$ and $b|U$ satisfy (12.2) and (12.3) for G' and w'. (Here $a|U$ means the vectors a_v for $v \in U$.) By Theorem 13, they determine the same ϑ. $\square$

16. Nonzero weights. We can also get some insight into the significance of nonzero weights by "splitting" vertices instead of removing them.

Lemma. *Let v be a vertex of G and let G' be a graph obtained from G by adding a new vertex v' and new edges*

$$u — v' \quad \text{iff} \quad u — v\,. \tag{16.1}$$

Let w and w' be nonnegative labelings of G and G' such that

$$w_u = w'_u\,, \qquad \text{when } u \neq v\,; \tag{16.2}$$

$$w_v = w'_v + w'_{v'}\,. \tag{16.3}$$

Then

$$\vartheta(G, w) = \vartheta(G', w')\,. \tag{16.4}$$

Proof. By Theorem 12 there are labelings a and b of G and $\overline{G}$ satisfying (12.2) and (12.3). We can modify them to obtain labelings a' and b' of G' and $\overline{G'}$ as follows, with the vectors of a' having one more component than the vectors of a:

$$a'_u = \binom{a_u}{0}, \qquad b'_u = b_u\,, \qquad \text{when } u \neq v\,; \tag{16.5}$$

$$a'_v = \binom{a_v}{\alpha}, \qquad a'_{v'} = \binom{a_v}{-\beta}, \qquad \alpha = \sqrt{\frac{w'_{v'}}{w'_v}}\,\|a_v\|\,, \qquad \beta = \sqrt{\frac{w'_v}{w'_{v'}}}\,\|a_v\|\,; \tag{16.6}$$

$$b'_v = b'_{v'} = b_v\,. \tag{16.7}$$

(We can assume by Lemma 15 that w'_v and $w'_{v'}$ are nonzero.) All orthogonality relations are preserved; and since $v \not\to v'$ in G', we also need to verify

$$a'_v \cdot a'_{v'} = \|a_v\|^2 - \alpha\beta = 0\,.$$

We have

$$c(a'_v) = \frac{c(a_v)\,\|a_v\|^2}{\|a_v\|^2 + \alpha^2} = \frac{c(a_v)}{1 + w'_{v'}/w'_v} = \frac{c(a_v)w'_v}{w_v} = \frac{w'_v}{\vartheta}\,,$$

and similarly $c(a'_{v'}) = w'_{v'}/\vartheta$; thus (12.2) and (12.3) are satisfied by a' and b' for G' and w'. □

Notice that if all the weights are integers we can apply this lemma repeatedly to establish that

$$\vartheta(G, w) = \vartheta(G')\,, \tag{16.8}$$

where G' is obtained from G by replacing each vertex v by a cluster of w_v mutually nonadjacent vertices that are adjacent to each of v's neighbors. (Recall that $\vartheta(G') = \vartheta(G', \mathbb{1})$, by definition (4.2).) In particular, if G is the trivial graph K_2 and if we assign the weights M and N, we have $\vartheta\big(K_2, (M, N)^T\big) = \vartheta(K_{M,N})$ where $K_{M,N}$ denotes the complete bipartite graph on M and N vertices.

A similar operation called "duplicating" a vertex has a similarly simple effect:

Corollary. *Let G' be constructed from G as in the lemma but with an additional edge between v and v'. Then $\vartheta(G, w) = \vartheta(G', w')$ if w' is defined by (16.2) and*

$$w_v = \max(w'_v, w'_{v'})\,. \tag{16.9}$$

Proof. We may assume that $w_v = w'_v$ and $w'_{v'} \neq 0$. Most of the construction (16.5)–(16.7) can be used again, but we set $\alpha = 0$ and $b'_{v'} = 0$ and

$$\beta = \sqrt{\frac{w_v - w'_{v'}}{w'_{v'}}}\,\|a_v\|\,.$$

Once again the necessary and sufficient conditions are readily verified. □

If the corollary is applied repeatedly, it tells us that $\vartheta(G)$ is unchanged when we replace the vertices of G by cliques.

17. Simple examples. We observed in section 4 that $\vartheta(G, w)$ always is at least

$$\vartheta_{\min} = \vartheta(K_n, w) = \max\{w_1, \ldots, w_n\} \tag{17.1}$$

and at most

$$\vartheta_{\max} = (\overline{K}_n, w) = w_1 + \cdots + w_n \,. \tag{17.2}$$

What are the corresponding orthogonal labelings?

For K_n the vectors of a have no orthogonal constraints, while the vectors of b must satisfy $b_u \cdot b_v = 0$ for all $u \neq v$. We can let a be the two-dimensional labeling

$$a_v = \begin{pmatrix} \sqrt{w_v} \\ \sqrt{\vartheta - w_v} \end{pmatrix}, \qquad \vartheta = \vartheta_{\min} \tag{17.3}$$

so that $\|a_v\|^2 = \vartheta$ and $c(a_v) = w_v/\vartheta$ as desired; and b can be one-dimensional,

$$b_v = \begin{cases} (1)\,, & \text{if } v = v_{\max} \\ (0)\,, & \text{if } v \neq v_{\max} \end{cases} \tag{17.4}$$

where $v_{\max}$ is any particular vertex that maximizes w_v. Clearly

$$\sum_v c(a_v) c(b_v) = \frac{c(a_{v_{\max}})}{\vartheta} = \frac{w_{v_{\max}}}{\vartheta} = 1\,.$$

For $\overline{K}_n$ the vectors of a must be mutually orthogonal while the vectors of b are unrestricted. We can let the vectors a be the columns of any orthogonal matrix whose top row contains the element

$$\sqrt{w_v/\vartheta}\,, \qquad \vartheta = \vartheta_{\max} \tag{17.5}$$

in column v. Then $\|a_v\|^2 = 1$ and $c(a_v) = w_v/\vartheta$. Once again a one-dimensional labeling suffices for b; we can let $b_v = (1)$ for all v.

18. The direct sum of graphs. Let $G = G' + G''$ be the graph on vertices

$$V = V' \cup V'' \tag{18.1}$$

where the vertex sets V' and V'' of G' and G'' are disjoint, and where u — v in G if and only if $u, v \in V'$ and u — v in G', or $u, v \in V''$ and u — v in G''. In this case

$$\vartheta(G, w) = \vartheta(G', w') + \vartheta(G'', w'')\,, \tag{18.2}$$

where w' and w'' are the sublabelings of w on vertices of V' and V''. We can prove (18.2) by constructing orthogonal labelings (a, b) satisfying (12.2) and (12.3).

Suppose a' is an orthogonal labeling of G' such that

$$\|a'_v\|^2 = \vartheta' \qquad a'_{1v} = \sqrt{w'_v}\,, \tag{18.3}$$

and suppose a'' is a similar orthogonal labeling of G''. If a' has dimension d' and a'' has dimension d'', we construct a new labeling a of dimension $d = d' + d''$ as follows, where j' runs from 2 to d' and j'' runs from 2 to d'':

$$\begin{array}{rlrl} & \text{if } v \in V' & & \text{if } v \in V'' \\ a_{1v} &= \sqrt{w'_v} = a'_{1v}\,, & a_{1v} &= \sqrt{w''_v} = a''_{1v}\,, \\ a_{j'v} &= \sqrt{\vartheta/\vartheta'}\, a'_{j'v}\,, & a_{j'v} &= 0\,, \\ a_{(d'+1)v} &= \sqrt{\vartheta'' w'_v/\vartheta'}\,, & a_{(d'+1)v} &= -\sqrt{\vartheta' w''_v/\vartheta''}\,, \\ a_{(d'+j'')v} &= 0\,, & a_{(d'+j'')v} &= \sqrt{\vartheta/\vartheta''}\, a''_{j''v}\,. \end{array} \tag{18.4}$$

Now if $u, v \in V'$ we have

$$a_u \cdot a_v = \sqrt{w'_u w'_v} + \frac{\vartheta}{\vartheta'}\left(a'_u \cdot a'_v - \sqrt{w'_u w'_v}\,\right) + \frac{\vartheta''}{\vartheta'}\sqrt{w'_u w'_v} = \frac{\vartheta}{\vartheta'}\, a'_u \cdot a'_v\,; \tag{18.5}$$

thus $a_u \cdot a_v = 0$ when $a'_u \cdot a'_v = 0$, and

$$\|a_v\|^2 = \frac{\vartheta}{\vartheta'}\,\|a'_v\|^2 = \vartheta\,. \tag{18.6}$$

It follows that $c(a_v) = w_v/\vartheta$ as desired. A similar derivation holds for $u, v \in V''$. And if $u \in V'$, $v \in V''$, then

$$a_u \cdot a_v = \sqrt{w'_u w''_v} - \sqrt{w'_u w''_v} = 0\,. \tag{18.7}$$

The orthogonal labeling b of $\overline{G' + G''}$ is much simpler; we just let $b_v = b'_v$ for $v \in V'$ and $b_v = b''_v$ for $v \in V''$. Then (12.2) and (12.3) are clearly preserved. This proves (18.2).

There is a close relation between the construction (18.4) and the construction (16.6), suggesting that we might be able to define another operation on graphs that generalizes both the splitting and direct sum operation.

19. The direct cosum of graphs. If G' and G'' are graphs on disjoint vertex sets V' and V'' as in section 18, we can also define

$$G = G' \mp G'' \iff \overline{G} = \overline{G'} + \overline{G''}\,. \tag{19.1}$$

This means u — v in G if and only if either u — v in G' or u — v in G'' or u and v belong to opposite vertex sets. In this case

$$\vartheta(G,w) = \max\bigl(\vartheta(G',w'),\ \vartheta(G'',w'')\bigr) \tag{19.2}$$

and again there is an easy way to construct (a,b) from (a',b') and (a'',b'') to prove (19.2). Assume "without lots of generality" that

$$\vartheta(G',w') \ge \vartheta(G'',w'') \tag{19.3}$$

and suppose again that we have (18.3) and its counterpart for a''. Then we can define

$$\begin{array}{rlrl} & \text{if } v \in V' & & \text{if } v \in V'' \\ a_{1v} &= \sqrt{w_{v'}} = a'_{1v}\,, & a_{1v} &= \sqrt{w_{v''}} = a''_{1v}\,, \\ a_{j'v} &= a'_{j'v}\,, & a_{j'v} &= 0\,, \\ a_{(d'+1)v} &= 0\,, & a_{(d'+1)v} &= \sqrt{(\vartheta' - \vartheta'')w''_v/\vartheta''}\,, \\ a_{(d'+j'')v} &= 0\,, & a_{(d'+j'')v} &= \sqrt{\vartheta'/\vartheta''}\, a''_{j''v}\,. \end{array} \tag{19.4}$$

Now a_v is essentially unchanged when $v \in V'$; and when $u,v \in V''$ we have

$$a_u \cdot a_v = \sqrt{w''_u w''_v} + \left(\frac{\vartheta'}{\vartheta''} - 1\right)\sqrt{w''_u w''_v} + \frac{\vartheta'}{\vartheta''}\bigl(a''_u \cdot a''_v - \sqrt{w''_u w''_v}\,\bigr) = \frac{\vartheta'}{\vartheta''}\, a''_u \cdot a''_v\,. \tag{19.5}$$

Again we retain the necessary orthogonality, and we have $c(a_v) = w_v/\vartheta$ for all v.

For the b's, we let $b_v = b'_v$ when $v \in V'$ and $b_v = 0$ when $v \in V''$.

20. A direct product of graphs. Now let G' and G'' be graphs on vertices V' and V'' and let V be the $n = n'n''$ ordered pairs

$$V = V' \times V''\,. \tag{20.1}$$

We define the 'strong product',

$$G = G' * G'' \tag{20.2}$$

on V by the rule

$$\begin{aligned} &(u',u'') \text{ — } (v',v'') \quad \text{or} \quad (u',u'') = (v',v'') \quad \text{in } G \\ &\iff (u' \text{ — } v' \text{ or } u' = v' \text{ in } G') \quad \text{and} \quad (u'' \text{ — } v'' \text{ or } u'' = v'' \text{ in } G'')\,. \end{aligned} \tag{20.3}$$

In this case we have, for example, $K_{n'} * K_{n''} = K_{n'n''}$ and $\overline{K}_{n'} * \overline{K}_{n''} = \overline{K}_{n'n''}$. More generally, if G' is regular of degree r' and G'' is regular of degree r'', then $G' * G''$ is regular of degree $(r'+1)(r''+1) - 1 = r'r'' + r' + r''$.

I don't know the value of $\vartheta(G,w)$ for arbitrary w, but I do know it in the special case

$$w_{(v',v'')} = w'_{v'} w''_{v''}\,. \tag{20.4}$$

Lemma. *If G and w are given by (20.2) and (20.4), then*

$$\vartheta(G,w) = \vartheta(G',w')\,\vartheta(G'',w'')\,. \tag{20.5}$$

Proof. [12] Given orthogonal labelings (a',b') and (a'',b'') of G' and G'', we let a be the Hadamard product

$$a_{(j',j'')(v',v'')} = a'_{j'v'} a''_{j''v''}\,, \qquad 1 \le j' \le d'\,, \quad 1 \le j'' \le d''\,, \tag{20.6}$$

where d' and d'' are the respective dimensions of the vectors in a' and a''. Then

$$\begin{aligned} a_{(u',u'')} \cdot a_{(v',v'')} &= \sum_{j',j''} a'_{j'u'} a''_{j''u''} a'_{j'v'} a''_{j''v''} \\ &= (a'_{u'} \cdot a'_{v'})(a''_{u''} \cdot a''_{v''})\,. \end{aligned} \tag{20.7}$$

Thus $\|a_{(v',v'')}\|^2 = \|a'_{v'}\|^2\, \|a''_{v''}\|^2$ and

$$c(a_{(v'v'')}) = c(a'_{v'})c(a''_{v''})\,. \tag{20.8}$$

The same construction is used for b in terms of b' and b''.

All necessary orthogonalities are preserved, because we have

$$\begin{aligned} &(u',u'') \mathrel{—} (v',v'') \text{ and } (u',u'') \ne (v',v'') \text{ in } G \\ &\qquad \Rightarrow (u' \mathrel{—} v' \text{ and } u' \ne v' \text{ in } G') \text{ or } (u'' \mathrel{—} v'' \text{ and } u'' \ne v'' \text{ in } G'') \\ &\qquad \Rightarrow b_{(u',u'')} \cdot b_{(v',v'')} = 0\,; \\ &(u',u'') \not\mathrel{—} (v',v'') \text{ and } (u',u'') \ne (v',v'') \text{ in } G \\ &\qquad \Rightarrow (u' \not\mathrel{—} v' \text{ and } u' \ne v' \text{ in } G') \text{ or } (u'' \not\mathrel{—} v'' \text{ and } u'' \ne v'' \text{ in } G'') \\ &\qquad \Rightarrow a_{(u',u'')} \cdot a_{(v',v'')} = 0\,. \end{aligned}$$

(In fact one of these relations is $\Leftrightarrow$, but we need only $\Rightarrow$ to make (20.7) zero when it needs to be zero.) Therefore a and b are orthogonal labelings of G that satisfy (12.2) and (12.3). $\square$

21. A direct coproduct of graphs. Guess what? We also define

$$G = G' \mathbin{\overline{*}} G'' \iff \overline{G} = \overline{G'} * \overline{G''}\,. \tag{21.1}$$

This graph tends to be "richer" than $G' * G''$; we have

$$\begin{aligned}&(u',u'') \mathrel{—} (v',v'') \text{ and } (u',u'') \neq (v',v'') \text{ in } G\\ &\qquad \iff (u' \mathrel{—} v' \text{ and } u' \neq v' \text{ in } G') \text{ or } (u'' \mathrel{—} v'' \text{ and } u'' \neq v'' \text{ in } G'')\,. \end{aligned} \tag{21.2}$$

Now, for instance, if G' is regular of degree r' and G'' is regular of degree r'', then

$$G' \mathbin{\overline{*}} G'' \text{ is regular of degree } n'n'' - (n'-r')(n''-r'') = r'n'' + r''n' - r'r''\,.$$

(This is always $\geq r'r'' + r' + r''$, because $r'(n''-1-r'') + r''(n'-1-r') \geq 0$.) Indeed, $G' \mathbin{\overline{*}} G'' \supseteq G' * G''$ for all graphs G' and G''. The Hadamard product construction used in section 20 can be applied word-for-word to prove that

$$\vartheta(G,w) = \vartheta(G',w')\,\vartheta(G'',w'') \tag{21.3}$$

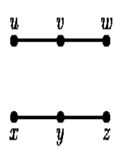
when G satisfies (21.1) and w has the special factored form (20.4).

It follows that many graphs have identical ϑ's:

Corollary. *If $G' * G'' \subseteq G \subseteq G' \mathbin{\overline{*}} G''$ and w satisfies (20.4), then (21.3) holds.*

Proof. This is just the monotonicity relation (4.3). The reason it works is that we have $a_{(u',v')} \cdot a_{(u'',v'')} = b_{(u',v')} \cdot b_{(u'',v'')} = 0$ for all pairs of vertices (u',u'') and (v',v'') whose adjacency differs in $G' * G''$ and $G' \mathbin{\overline{*}} G''$. □

Some small examples will help clarify the results of the past few sections. Let P_3 be the path of length 2 on 3 vertices, •—•—•, and consider the four graphs we get by taking its strong product and coproduct with $\overline{K}_2$ and K_2:

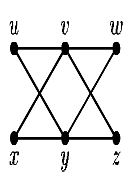
$\overline{K}_2 * P_3 =$ 

$$\vartheta = \max(u+w,v) + \max(x+z,y)$$

(Since P_3 may be regarded as $\overline{K}_2 \mathbin{\overline{+}} K_1$ and $\overline{K}_2$ is $K_1 + K_2$, this graph is

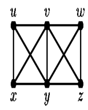
$$\big((K_1+K_1)\mathbin{\overline{+}}K_1\big) + \big((K_1+K_1)\mathbin{\overline{+}}K_1\big)$$

and the formula for ϑ follows from (18.2) and (19.2).)

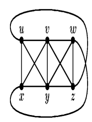
$\overline{K}_2 \mathbin{\overline{*}} P_3 =$ 

$$\vartheta = \max(u+w+x+z, v+y)$$

(This graph is $\overline{K}_2 \mathbin{\overline{+}} \overline{K}_4$; we could also obtain it by applying Lemma 16 three times to P_3.)

$K_2 * P_3 =$ 

$$\vartheta = \max\bigl(\max(u,x)+\max(w,z),\max(v,y)\bigr)$$

$K_2 \mathbin{\overline{*}} P_3 =$ 

$$\vartheta = \max\bigl(\max(u+w,x+z),\max(v,y)\bigr)$$

If the weights satisfy $u=\lambda x$, $v=\lambda y$, $w=\lambda z$ for some parameter λ, the first two formulas for ϑ both reduce to $(1+\lambda)\max(u+w,v)$, in agreement with (20.5) and (21.3). Similarly, the last two formulas for ϑ reduce to $\max(1,\lambda)\max(u+w,v)$ in such a case.

22. Odd cycles. Now let $G=C_n$ be the graph with vertices $0,1,\ldots,n-1$ and

$$u \mathrel{—} v \iff u-v\equiv \pm 1 \pmod{n}\,, \tag{22.1}$$

where n is an *odd* number. A general formula for $\vartheta(C_n,w)$ appears to be very difficult; but we can compute $\vartheta(C_n)$ without too much labor when all weights are 1, because of the cyclic symmetry.

It is easier to construct orthogonal labelings of $\overline{C}_n$ than of C_n, so we begin with that. Given a vertex v, $0\le v<n$, let b_v be the three-dimensional vector

$$b_v = \begin{pmatrix} \alpha \\ \cos v\varphi \\ \sin v\varphi \end{pmatrix}, \tag{22.2}$$

where α and φ remain to be determined. We have

$$\begin{aligned} b_u\cdot b_v &= \alpha^2+\cos u\varphi\ \cos v\varphi+\sin u\varphi\ \sin v\varphi \\ &= \alpha^2+\cos(u-v)\varphi\,. \end{aligned} \tag{22.3}$$

Therefore we can make $b_u\cdot b_v=0$ when $u\equiv v\pm 1$ by setting

$$\alpha^2 = -\cos\varphi\,, \qquad \varphi=\frac{\pi(n-1)}{n}\,. \tag{22.4}$$

This choice of φ makes $n\varphi$ a multiple of 2π, because n is odd. We have found an orthogonal labeling b of $\overline{C}_n$ such that

$$c(b_v) = \frac{\alpha^2}{1+\alpha^2} = \frac{\cos\pi/n}{1+\cos\pi/n}\,. \tag{22.5}$$

Turning now to orthogonal labelings of C_n, we can use $(2n-1)$-dimensional vectors

$$a_v = \begin{pmatrix} \alpha_0 \\ \alpha_1 \cos v\varphi \\ \alpha_1 \sin v\varphi \\ \alpha_2 \cos 2v\varphi \\ \alpha_2 \sin 2v\varphi \\ \vdots \\ \alpha_{n-1} \cos(n-1)v\varphi \\ \alpha_{n-1} \sin(n-1)v\varphi \end{pmatrix}, \tag{22.6}$$

with $\varphi = \pi(n-1)/n$ as before. As in (22.3), we find

$$a_u \cdot a_v = \sum_{k=0}^{n-1} \alpha_k^2 \cos k(u-v)\varphi\,; \tag{22.7}$$

so the result depends only on $(u-v) \bmod n$. Let $\omega = e^{i\varphi}$. We can find values of α_k such that $a_u \cdot a_v = x_{(u-v) \bmod n}$ by solving the equations

$$x_j = \sum_{k=0}^{n-1} \alpha_k^2 \omega^{jk}\,. \tag{22.8}$$

Now ω is a primitive nth root of unity; i.e., $\omega^k = 1$ iff k is a multiple of n. So (22.9) is just a finite Fourier transform, and we can easily invert it: For $0 \le m < n$ we have

$$\sum_{j=0}^{n-1} \omega^{-mj} x_j = \sum_{k=0}^{n-1} \alpha_k^2 \sum_{j=0}^{n-1} \omega^{j(k-m)} = n\,\alpha_m^2\,.$$

In our case we want a solution with $x_2 = x_3 = \cdots = x_{n-2} = 0$, and we can set $x_0 = 1$, $x_{n-1} = x_1 = x$, so we find

$$n\,\alpha_k^2 = x_0 + \omega^{-k} x_1 + \omega^k x_{n-1} = 1 + 2x \cos k\varphi\,.$$

We must choose x so that these values are nonnegative; this means $2x \le -1/\cos\varphi$, since $\cos k\varphi$ is most negative when $k = 1$. Setting x to this maximum value yields

$$c(a_v) = \alpha_0^2 = \frac{1}{n}\left(1 - \frac{1}{\cos\varphi}\right) = \frac{1 + \cos\pi/n}{n \cos\pi/n}\,. \tag{22.9}$$

So (22.5) and (22.9) give

$$\sum_v c(a_v)c(b_v) = \sum_v 1/n = 1\,. \tag{22.10}$$

This is (12.3), hence from (12.2) we know that $\vartheta(C_n) = \lambda$. We have proved, in fact, that

$$\vartheta(C_n, \mathbb{1}) = \frac{n \cos \pi/n}{1 + \cos \pi/n}\,; \tag{22.11}$$

$$\vartheta(\overline{C}_n, \mathbb{1}) = \frac{1 + \cos \pi/n}{\cos \pi/n}\,. \tag{22.12}$$

When $n = 3$, $C_n = K_n$ and these values agree with $\vartheta(K_3) = 1$, $\vartheta(\overline{K_3}) = 3$; when $n = 5$, $\overline{C}_5$ is isomorphic to C_5 so $\vartheta(C_5) = \sqrt{5}$; when n is large,

$$\vartheta(C_n) = \frac{n}{2} - \frac{\pi^2}{8n} + O(n^{-3})\,; \qquad \vartheta(\overline{C}_n) = 2 + \frac{\pi^2}{2n^2} + O(n^{-4})\,. \tag{22.13}$$

Instead of an explicit construction of vectors a_v as in (22.6), we could also find $\vartheta(C_n)$ by using the matrix characterization ϑ_2 of section 6. When all weights are 1, a feasible A has 1 everywhere except on the superdiagonal, the subdiagonal, and the corners. This suggests that we look at "circulant" matrices; for example, when $n = 5$,

$$A = \begin{pmatrix} 1 & 1+x & 1 & 1 & 1+x \\ 1+x & 1 & 1+x & 1 & 1 \\ 1 & 1+x & 1 & 1+x & 1 \\ 1 & 1 & 1+x & 1 & 1+x \\ 1+x & 1 & 1 & 1+x & 1 \end{pmatrix} = J + xP + xP^{-1}\,, \tag{22.14}$$

where J is all 1's and P is the permutation matrix taking j into $(j+1) \bmod n$. It is well known and not difficult to prove that the eigenvalues of the circulant matrix $a_0 I + a_1 P + \cdots + a_{n-1}P^{n-1}$ are

$$\sum_{0 \le j < n} \omega^{kj} a_j\,, \qquad 0 \le k < n\,, \tag{22.15}$$

where $\omega = e^{2\pi i/n}$. (Indeed, it suffices to find the eigenvalues of P itself. This ω is a different primitive root of unity from the ω we used in (22.8).) Hence the eigenvalues of (22.14) are

$$n + 2x\,, \quad x(\omega + \omega^{-1})\,, \quad x(\omega^2 + \omega^{-2})\,, \;\ldots\,, \quad x(\omega^{n-1} + \omega^{1-n})\,. \tag{22.16}$$

We minimize the maximum of these values if we choose x so that

$$n + 2x = -2x \cos \pi/n\,;$$

then

$$\Lambda(A) = -2x \cos \pi/n = \frac{n \cos \pi/n}{1 + \cos \pi/n} \tag{22.17}$$

is the value of $\vartheta(G)$.

If n is even, the graph C_n is bipartite. We will prove later that bipartite graphs are perfect, hence $\vartheta(C_n) = n/2$ and $\vartheta(\overline{C}_n) = 2$ in the even case.

23. Comments on the previous example. The cycles C_n provide us with infinitely many graphs G for which $\vartheta(G)\vartheta(\overline{G}\,) = n$, and it is natural to wonder whether this is true in general. Of course it is not: If $G = \overline{K}_m + K_{n-m}$ then $\overline{G} = K_m \mp \overline{K_{n-m}}$, hence we know from Lemmas 18 and 19 that

$$\vartheta(G) = m + 1\,, \qquad \vartheta(\overline{G}\,) = \max(1, n-m)\,. \tag{23.1}$$

In particular, we can make $\vartheta(G)\vartheta(\overline{G}\,)$ as high as $n^2/4 + n/2$ when $m = \lfloor n/2 \rfloor$.

We can, however, prove without difficulty that $\vartheta(G)\vartheta(\overline{G}\,) \ge n$:

Lemma.

$$\vartheta(G, w)\vartheta(\overline{G}, w') \ge w \cdot w'\,. \tag{23.2}$$

Proof. By Theorem 12 there is an orthogonal labeling a of G and an orthogonal labeling b of $\overline{G}$ such that

$$c(a_v) = w_v/\vartheta(G, w)\,, \qquad c(b_v) = w'_v/\vartheta(\overline{G}, w')\,. \tag{23.3}$$

By (11.1) we have

$$\sum_v c(a_v)c(b_v) \le 1\,. \tag{23.4}$$

QED. □

24. Regular graphs. When each vertex of G has exactly r neighbors, Lovász and Hoffman observed that the construction in (22.14) can be generalized. Let B be the adjacency matrix of G, i.e., the $n \times n$ matrix with

$$B_{uv} = \begin{cases} 1\,, & \text{if } u \text{ — } v; \\ 0\,, & \text{if } u = v \text{ or } u \not\text{—}\, v. \end{cases} \tag{24.1}$$

Lemma. *If G is a regular graph,*

$$\vartheta(G) \le \frac{n\Lambda(-B)}{\Lambda(B) + \Lambda(-B)}\,. \tag{24.2}$$

Proof. Let A be a matrix analogous to (22.14),

$$A = J + xB\,. \tag{24.3}$$

Since G is regular, the all-1's vector $\mathbb{1}$ is an eigenvector of B, and the other eigenvectors are orthogonal to $\mathbb{1}$ so they are eigenvectors also of A. Thus if the eigenvalues of B are

$$r = \Lambda(B) = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n = -\Lambda(-B)\,, \tag{24.4}$$

the eigenvalues of A are

$$n + rx\,,\ x\lambda_2\,,\ \ldots\,,\ x\lambda_n\,. \tag{24.5}$$

(The Perron-Frobenius theorem tells us that $\lambda_1 = r$.) We have $\lambda_1 + \cdots + \lambda_n = \mathrm{tr}(B) = 0$, so $\lambda_n < 0$, and we minimize the maximum of (24.5) by choosing $n + rx = x\lambda_n$; thus

$$\Lambda(A) = x\lambda_n = \frac{-n\lambda_n}{r - \lambda_n}\,,$$

which is the right-hand side of (24.2). By (6.3) and Theorem 12 this is an upper bound on ϑ. $\square$

Incidentally, we need to be a little careful in (24.2): The denominator can be zero, but only when $G = \overline{K}_n$.

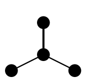
25. Automorphisms. An automorphism of a graph G is a permutation p of the vertices such that

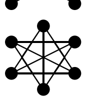
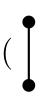
$$p(u) — p(v) \qquad \text{iff } u — v\,. \tag{25.1}$$

Such permutations are closed under multiplication, so they form a group.

We call G *vertex-symmetric* if its automorphism group is vertex-transitive, i.e., if given u and v there is an automorphism p such that $p(u) = v$. We call G *edge-symmetric* if its automorphism group is edge-transitive, i.e., if given $u — v$ and $u' — v'$ there is an automorphism p such that $p(u) = u'$ and $p(v) = v'$ or $p(u) = v'$ and $p(v) = u'$.

Any vertex-symmetric graph is regular, but edge-symmetric graphs need not be regular. For example,

 is edge-symmetric, not vertex-symmetric;

 is vertex-symmetric, not edge-symmetric. ( is a maximal clique)

The graph $\overline{C}_n$ is not edge-symmetric for $n > 7$ because it has more edges than automorphisms. Also, $\overline{C}_7$ has no automorphism that takes $0 — 2$ into $0 — 3$.

Lemma. *If G is edge-symmetric and regular, equality holds in Lemma 24.*

Proof. Say that A is an optimum feasible matrix for G if it is a feasible matrix with

$$\Lambda(A) = \vartheta(G)$$

as in section 6. We can prove that optimum feasible matrices form a convex set, as follows. First, $tA+(1-t)B$ is clearly feasible when A and B are feasible. Second,

$$\Lambda\bigl(tA+(1-t)B\bigr) \le t\Lambda(A)+(1-t)\Lambda(B)\,, \qquad 0\le t\le 1 \tag{25.2}$$

holds for all symmetric matrices A and B, by (6.2); this follows because there is a unit vector x such that $\Lambda\bigl(tA+(1-t)B\bigr) = x^T\bigl(tA+(1-t)B\bigr)x = tx^TAx+(1-t)x^TBx \le t\Lambda(A)+(1-t)\Lambda(B)$. Third, if A and B are optimum feasible matrices, the right side of (25.2) is $\vartheta(G)$ while the left side is $\ge \vartheta(G)$ by (6.3). Therefore equality holds.

If A is an optimum feasible matrix for G, so is $p(A)$, the matrix obtained by permuting rows and columns by an automorphism p. (I mean $p(A)_{uv} = A_{p(u)p(v)}$.) Therefore the average, $\bar{A}$, over all p is also an optimal feasible matrix. Since $p(\bar{A}) = \bar{A}$ for all automorphisms p, and since G is edge-symmetric, $\bar{A}$ has the form $J+xB$ where B is the adjacency matrix of G. The bound in Lemma 24 is therefore tight. □

(Note: If p is a permutation, let $P_{uv}=1$ if $u=p(v)$, otherwise 0. Then $(P^TAP)_{uv} = \sum P^T_{uj}A_{jk}P_{kv} = A_{p(u)p(v)}$, so $p(A)=P^TAP$.)

The argument in this proof shows that the set of all optimum feasible matrices A for G has a common eigenvector x such that $Ax=\vartheta(G)x$. The argument also shows that, if G has an edge automorphism taking u — v into u' — v', we can assume without loss of generality that $A_{uv} = A_{u'v'}$ in an optimum feasible matrix. This simplifies the computation of $\Lambda(A)$, and justifies our restriction to circulant matrices (22.14) in the case of cyclic graphs.

Theorem. *If G is vertex-symmetric, $\vartheta(G)\vartheta(\overline{G}) = n$.*

Proof. Say that b is an optimum normalized labeling of $\overline{G}$ if it is a normalized orthogonal labeling of $\overline{G}$ achieving equality in (7.1) when all weights are 1:

$$\vartheta = \sum_{u,v} b_u\cdot b_v\,, \qquad \sum_v \|b_v\|^2 = 1\,, \qquad b_u\cdot b_v = 0 \text{ when } u — v\,. \tag{25.3}$$

Let B be the corresponding spud; i.e., $B_{uv} = b_u\cdot b_v$ and $\vartheta = \sum_{u,v} B_{uv}$. Then $p(B)$ is also equivalent to an optimum normalized labeling, whenever p is an automorphism; and such matrices B form a convex set, so we can assume as in the lemma that $B=p(B)$ for all automorphisms p. Since G is vertex-symmetric, we must have $B_{vv}=1/n$ for all vertices v. Thus there is an optimum normalized labeling b with $\|b_v\|^2 = 1/n$, and the arguments of Lemma 10 and Theorem 12 establish the existence of such a b with

$$c(b_v) = \vartheta(G)/n \tag{25.4}$$

for all v. But b is an orthogonal labeling of $\overline{G}$, hence

$$\vartheta_1(\overline{G}, \mathbb{1}) \le n/\vartheta(G)$$

by the definition (5.2) of ϑ_1. Thus $\vartheta(\overline{G})\vartheta(G) \le n$; we have already proved the reverse inequality in Lemma 23. □

26. Consequence for eigenvalues. A curious corollary of the results just proved is the following fact about eigenvalues.

Corollary. *If the graphs G and $\overline{G}$ are vertex-symmetric and edge-symmetric, and if the adjacency matrix of G has eigenvalues*

$$\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n\,, \tag{26.1}$$

then

$$(\lambda_1 - \lambda_n)(n - \lambda_1 + \lambda_2) = -\lambda_n(\lambda_2 + 1)n\,. \tag{26.2}$$

Proof. By Lemma 25 and Theorem 25,

$$\frac{n\Lambda(-B)}{\Lambda(B) + \Lambda(-B)}\;\frac{n\Lambda(-\overline{B})}{\Lambda(\overline{B}) + \Lambda(-\overline{B})} = n\,, \tag{26.3}$$

where B and $\overline{B}$ are the adjacency matrices of G and $\overline{G}$, and where we interpret $0/0$ as 1. We have

$$\overline{B} = J - I - B\,. \tag{26.4}$$

If the eigenvalues of B are given by (26.1), the eigenvalues of $\overline{B}$ are therefore

$$n - 1 - \lambda_1 \ge -1 - \lambda_n \ge \cdots \ge -1 - \lambda_2\,. \tag{26.5}$$

(We use the fact that G is regular of degree λ_1.) Formula (26.2) follows if we plug the values $\Lambda(B) = \lambda_1$, $\Lambda(-B) = -\lambda_n$, $\Lambda(\overline{B}) = n - 1 - \lambda_1$, $\Lambda(-\overline{B}) = 1 + \lambda_2$ into (26.3). □

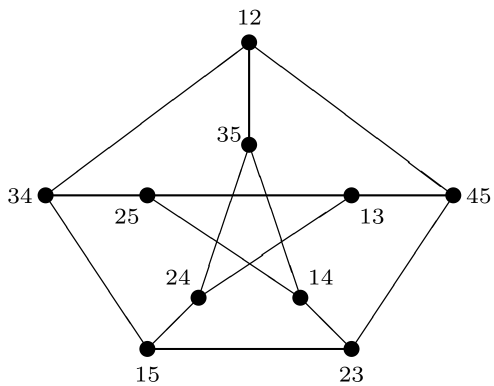
27. Further examples of symmetric graphs. Consider the graph $P(m, t, q)$ whose vertices are all $\binom{m}{t}$ subsets of cardinality t of some given set S of cardinality m, where

$$u \,—\, v \iff |u \cap v| = q\,. \tag{27.1}$$

We want $0 \le q < t$ and $m \ge 2t - q$, so that the graph isn't empty. In fact, we can assume that $m \ge 2t$, because $P(m, r, q)$ is isomorphic to $P(m, m - t, m - 2t + q)$ if we map each subset u into the set difference $S \setminus u$:

$$|(S \setminus u) \cap (S \setminus v)| = |S| - |u \cup v| = |S| - |u| - |v| + |u \cap v|\,. \tag{27.2}$$

The letter P stands for Petersen, because $P(5,2,0)$ is the well known "Petersen graph" on 10 vertices,



(27.3)

These graphs are clearly vertex-symmetric and edge-symmetric, because every permutation of S induces an automorphism. For example, to find an automorphism that maps u — v into u' — v', let $u = (u \cap v) \cup \bar{u}$, $v = (u \cap v) \cup \bar{v}$, $u' = (u' \cap v') \cup \bar{u}'$, $v' = (u' \cap v') \cup \bar{v}'$, and apply any permutation that takes the q elements of $u \cap v$ into the q elements of $u' \cap v'$, the $t - q$ elements of $\bar{u}$ into the $t - q$ elements of u', and the $t - q$ elements of $\bar{v}$ into $\bar{v}'$. Thus we can determine $\vartheta\bigl(P(m,t,q)\bigr)$ from the eigenvalues of the adjacency matrix. Lovász [12] discusses the case $q = 0$, and his discussion readily generalizes to other values of q. It turns out that $\vartheta\bigl(P(m,t,0)\bigr) = \binom{m-1}{t-1}$. This is also the value of $\alpha\bigl(P(m,t,0)\bigr)$, because the $\binom{m-1}{t-1}$ vertices containing any given point form a stable set.

The special case $t = 2$, $q = 0$ is especially interesting because those graphs also satisfy the condition of Corollary 26. We have

$$n = \binom{m}{2}, \qquad \lambda_1 = \binom{m-2}{2}, \qquad \lambda_2 = 1\,, \qquad \lambda_n = 3 - m\,, \tag{27.4}$$

and (26.2) does indeed hold (but not "trivially"). It is possible to cover $P(m,2,0)$ with disjoint maximum cliques; hence $\kappa\bigl(P(m,2,0)\bigr) = \binom{m}{2} \big/ \lfloor \frac{m}{2} \rfloor = 2\lceil \frac{m}{2} \rceil - 1$. In particular, when G is the Petersen graph we have $\alpha(G) = \vartheta(G) = 4$, $\kappa(G) = 5$; also $\alpha(\overline{G}) = 2$, $\vartheta(\overline{G}) = \kappa(\overline{G}) = \frac{5}{2}$.

28. A bound on ϑ. The paper [12] contains one more result about ϑ that is not in [7], so we will wrap up our discussion of [12] by describing [12, Theorem 11].

Theorem. *If G has an orthogonal labeling of dimension d with no zero vectors, we have $\vartheta(G) \le d$.*

Proof. Given a non-zero orthogonal labeling a of dimension d, we can assume that $\|a_v\|^2 = 1$ for all v. (The hypothesis about zeros is important, since there is trivially an orthogonal labeling of any desired dimension if we allow zeros. The labeling needn't be optimum.) Then we construct an orthogonal labeling a'' of dimension d^2, with $c(a''_v) = 1/d$ for all v, as follows:

Let a'_v have d^2 components where the (j,k) component is $a_{jv}a_{kv}$. Then

$$a'_u \cdot a'_v = (a_u \cdot a_v)^2 \tag{28.1}$$

as in (20.7). Let Q be any orthogonal matrix with d^2 rows and columns, such that the (j,k) entry in row $(1,1)$ is $1/\sqrt{d}$ for $j=k$, 0 otherwise. Then we define

$$a''_v = Qa'_v\,. \tag{28.2}$$

Once again $a''_u \cdot a''_v = (a_u \cdot a_v)^2$, so a'' is an orthogonal labeling. We also have first component

$$a''_{(1,1)v} = \sum_{j,k} \frac{[j=k]}{\sqrt{d}}\, a'_{(j,k)v} = \sum_k \frac{a^2_{kv}}{\sqrt{d}} = \frac{1}{\sqrt{d}}\,; \tag{28.3}$$

hence $c(a''_v) = 1/d$. This proves $\vartheta(G) \le d$, by definition of ϑ_1. $\square$

This theorem improves the obvious lower bound $\alpha(G)$ on the dimension of an optimum labeling.

29. Compatible matrices. There's another way to formulate the theory we've been developing, by looking at things from a somewhat higher level, following ideas developed by Lovász and Schrijver [15] a few years after the book [7] was written. Let us say that the matrix A is *λ-compatible with G and w* if A is an $(n+1)\times(n+1)$ spud indexed by the vertices of G and by a special value 0, having the following properties:

- $A_{00} = \lambda$;
- $A_{vv} = A_{0v} = w_v$ for all vertices v;
- $A_{uv} = 0$ whenever $u \not\multimap v$ in G.

Lemma. *There exists an orthogonal labeling a for G with costs $c(a_v) = w_v/\lambda$ if and only if there exists a matrix A that is λ-compatible with G and w.*

Proof. Given such an orthogonal labeling, we can normalize each vector so that $\|a_v\|^2 = w_v$. Then when $w_v \ne 0$ we have

$$\frac{w_v}{\lambda} = c(a_v) = \frac{a^2_{1v}}{w_v},$$

so we can assume that $a_{1v} = w_v/\sqrt{\lambda}$ for all v. Add a new vector a_0, having $a_{10} = \sqrt{\lambda}$ and $a_{j0} = 0$ for all $j > 1$. Then the matrix A with $A_{uv} = a_u \cdot a_v$ is easily seen to be λ-compatible with G and w.

Conversely, if such a matrix A exists, there are $n+1$ vectors $a_0, \ldots, a_n$ such that $A_{uv} = a_u \cdot a_v$; in particular, $\|a_0\|^2 = \lambda$. Let Q be an orthogonal matrix with first row $a_0^T/\sqrt{\lambda}$, and define $a'_v = Qa_v$ for all v. Then $a'_{10} = \sqrt{\lambda}$ and $a'_{j0} = 0$ for all $j > 1$. Also $a'_u \cdot a'_v = a_u \cdot a_v = A_{uv}$ for all u and v. Hence $\sqrt{\lambda}a'_{1v} = a'_0 \cdot a'_v = A_{0v} = w_v$ and $\|a'_v\|^2 = a'_v \cdot a'_v = A_{vv} = w_v$, for all $v \in G$, proving that $c(a'_v) = w_v/\lambda$. Finally a' is an orthogonal labeling, since $a'_u \cdot a'_v = A_{uv} = 0$ whenever $u \not- v$. □

Corollary. *$x \in \mathtt{TH}(G)$ iff there exists a matrix 1-compatible with $\overline{G}$ and x.*

Proof. Set $\lambda = 1$ in the lemma and apply Theorem 14. □

The corollary and definition (4.1) tell us that $\vartheta(G, w)$ is $\max(w_1x_1 + \cdots + w_nx_n)$ over all x that appear in matrices that are 1-compatible for $\overline{G}$ and x. Theorem 12 tells us that $\vartheta(G, w)$ is also the minimum λ such that there exists a λ-compatible matrix for G and w. The "certificate" property of Theorem 13 has an even stronger formulation in matrix terms:

Theorem. *Given a nonnegative weight vector $w = (w_1, \ldots, w_n)^T$, let A be λ-compatible with G and w, where λ is as small as possible, and let B be 1-compatible with $\overline{G}$ and x, where $w_1x_1 + \cdots + w_nx_n$ is as large as possible. Then*

$$ADB = 0\,, \tag{29.1}$$

where D is the diagonal matrix with $D_{00} = -1$ and $D_{vv} = +1$ for all $v \neq 0$. Conversely, if A is λ-compatible with G and w and if B is 1-compatible with $\overline{G}$ and x, then (29.1) implies that $\lambda = w_1x_1 + \cdots + w_nx_n = \vartheta(G, w)$.

Proof. Assume that A is λ-compatible with G and w, and B is 1-compatible with $\overline{G}$ and x. Let $B' = DBD$, so that B' is a spud with $B'_{00} = 1$, $B'_{0v} = B'_{v0} = -x_v$, and $B'_{uv} = B_{uv}$ when u and v are nonzero. Then the dot product $A \cdot B'$ is

$$\lambda - w_1x_1 - \cdots - w_nx_n - w_1x_1 - \cdots - w_nx_n + w_1x_1 + \cdots + w_nx_n = \lambda - (w_1x_1 + \cdots + w_nx_n)\,,$$

because $A_{uv}B_{uv} = 0$ when u and v are vertices of G. We showed in the proof of F9 in section 9 that the dot product of spuds is nonnegative; in fact, that proof implies that the dot product is zero if and only if the ordinary matrix product is zero. So $\lambda = w_1x_1 + \cdots + w_nx_n = \vartheta(G, w)$ iff $AB' = 0$, and this is equivalent to (29.1). □

Equation (29.1) gives us further information about the orthogonal labelings a and b that appear in Theorems 12 and 13. Normalize those labelings so that $\|a\|^2 = w_v$ and

$\|b\|^2 = x_v$. Then we have

$$\sum_{t \in G} w_t \,(b_t \cdot b_v) = \vartheta \, x_v \,, \tag{29.2}$$

$$\sum_{t \in G} x_t \,(a_t \cdot a_v) = w_v \,, \tag{29.3}$$

$$\sum_{t \in G} (a_t \cdot a_u)(b_t \cdot b_v) = w_u x_v \,, \tag{29.4}$$

for all vertices u and v of G. (Indeed, (29.2) and (29.4) are respectively equivalent to $(AB')_{0v} = 0$ and $(AB')_{vv} = 0$; (29.3) is equivalent to $(B'A)_{0v} = 0$.) Notice that if $\widehat{A}$ and $\widehat{B}$ are the $n \times n$ spuds; obtained by deleting row 0 and column 0 from optimum matrices A and B, these equations are equivalent to

$$\widehat{B} w = \vartheta x \,, \qquad \widehat{A} x = w \,, \qquad \widehat{A}\widehat{B} = w x^T \,. \tag{29.5}$$

Equation (29.1) is equivalent to (29.5) together with the condition $w_1 x_1 + \cdots + w_n x_n = \vartheta$.

Since $AB' = 0$ iff $B'A = 0$ when A and B' are symmetric matrices, the optimum matrices A and B' commute. This implies that they have common eigenvectors: There is an orthogonal matrix Q such that

$$A = Q \operatorname{diag}(\lambda_0, \ldots, \lambda_n)\, Q^T \,, \qquad B' = Q \operatorname{diag}(\mu_0, \ldots, \mu_n)\, Q^T \,. \tag{29.6}$$

Moreover, the product is zero, so

$$\lambda_0 \mu_0 = \cdots = \lambda_n \mu_n = 0 \,. \tag{29.7}$$

The number of zero eigenvalues λ_k is $n+1-d$, where d is the smallest dimension for which there is an orthogonal labeling a with $A_{uv} = a_u \cdot a_v$. A similar statement holds for B', since the eigenvalues of B and B' are the same; y is an eigenvector for B iff Dy is an eigenvector for B'. In the case $G = C_n$, studied in section 22, we constructed an orthogonal labeling (22.3) with only three dimensions, so all but 3 of the eigenvalues μ_k were zero. When all the weights w_v are nonzero and $\vartheta(G)$ is large, Theorem 28 implies that a large number of λ_k must be nonzero, hence a large number of μ_k must be zero.

The "optimum feasible matrices" A studied in section 6 are related to the matrices $\widehat{A}$ of (29.5) by the formula

$$\vartheta \widehat{A} = w w^T - \vartheta \operatorname{diag}(w_1, \ldots, w_n) - \operatorname{diag}(\sqrt{w_1}, \ldots, \sqrt{w_n})\, A \operatorname{diag}(\sqrt{w_1}, \ldots, \sqrt{w_n}) \,, \tag{29.8}$$

because of the construction following (6.6). If the largest eigenvalue $\Lambda(A) = \vartheta$ of A occurs with multiplicity r, the rank of $\vartheta I - A$ will be $n - r$, hence $\widehat{A}$ will have rank $n - r$ or $n - r + 1$, and the number of zero eigenvalues λ_k in (29.6) will be $r + 1$ or r.

30. Antiblockers. The convex sets $\mathtt{STAB}$, $\mathtt{TH}$, and $\mathtt{QSTAB}$ defined in section 2 have many special properties. For example, they are always nonempty, closed, convex, and nonnegative; they also satisfy the condition

$$0 \le y \le x \quad \text{and} \quad x \in X \;\Rightarrow\; y \in X\,. \tag{30.1}$$

A set X of vectors satisfying all five of these properties is called a *convex corner*.

If X is any set of nonnegative vectors we define its *antiblocker* by the condition

$$\operatorname{abl} X = \{\, y \ge 0 \mid x \cdot y \le 1 \text{ for all } x \in X \,\}\,. \tag{30.2}$$

Clearly $\operatorname{abl} X$ is a convex corner, and $\operatorname{abl} X \supseteq \operatorname{abl} X'$ when $X \subseteq X'$.

Lemma. *If X is a convex corner we have* $\operatorname{abl}\operatorname{abl} X = X$.

Proof. (Compare with the proof of F5 in section 8.) The relation $X \subseteq \operatorname{abl}\operatorname{abl} X$ is obvious by definition (30.2), so the lemma can fail only if there is some $z \in \operatorname{abl}\operatorname{abl} X$ with $z \notin X$. Then there is a hyperplane separating z from X, by F1; i.e., there is a vector y and a number b such that $x \cdot y \le b$ for all $x \in X$ but $z \cdot y > b$. Let y' be the same as y but with all negative components changed to zero. Then (y', b) is also a separating hyperplane. [*Proof:* If $x \in X$, let x' be the same as x but with all components changed to zero where y has a negative entry; then $x' \in X$, and $x \cdot y' = x' \cdot y \le b$. Furthermore $z \cdot y' \ge z \cdot y > b$.] If $b = 0$, we have $\lambda y' \in \operatorname{abl} X$ for all $\lambda > 0$; this contradicts $z \cdot \lambda y' \le 1$. We cannot have $b < 0$, since $0 \in X$. Hence $b > 0$, and the vector $y'/b \in \operatorname{abl} X$. But then $z \cdot (y'/b)$ must be ≤ 1, a contradiction. □

Corollary. *If G is any graph we have*

$$\mathtt{STAB}(\overline{G}) = \operatorname{abl} \mathtt{QSTAB}(G)\,, \tag{30.3}$$

$$\mathtt{TH}(\overline{G}) = \operatorname{abl} \mathtt{TH}(G)\,, \tag{30.4}$$

$$\mathtt{QSTAB}(\overline{G}) = \operatorname{abl} \mathtt{STAB}(G)\,. \tag{30.5}$$

Proof. First we show that

$$\operatorname{abl} X = \operatorname{abl} \text{convex hull } X. \tag{30.6}$$

The left side surely contains the right. And any element $y \in \operatorname{abl} X$ will satisfy

$$(\alpha_1 x^{(1)} + \cdots + \alpha_k x^{(k)}) \cdot y \le 1$$

when the α's are nonnegative scalars summing to 1 and the $x^{(j)}$ are in X. This proves (30.6), because the convex hull of X is the set of all such $\alpha_1 x^{(1)} + \cdots + \alpha_k x^{(k)}$.

Now (30.6) implies (30.5), because the definitions in section 2 say that

$$\begin{aligned}\mathtt{QSTAB}(\overline{G}) &= \mathrm{abl}\,\{\, x \mid x \text{ is a clique labeling of } \overline{G}\,\} \\ &= \mathrm{abl}\,\{\, x \mid x \text{ is a stable labeling of } G\,\}\,, \\ \mathtt{STAB}(G) &= \text{convex hull}\,\{\, x \mid x \text{ is a stable labeling of } G\,\}\,.\end{aligned}$$

And (30.5) is equivalent to (30.3) by the lemma, because $\mathtt{STAB}(G)$ is a convex corner. (We must prove (30.1), and it suffices to do this when y equals x in all but one component; and in fact by convexity we may assume that y is 0 in that component; and then we can easily prove it, because any subset of a stable set is stable.)

Finally, (30.4) is equivalent to Theorem 14, because $\mathtt{TH}(G) = \mathrm{abl}\,\{\, x \mid x_v = c(a_v)$ for some orthogonal labeling of $G\,\}$. □

The sets STAB and QSTAB are polytopes, i.e., they are bounded and can be defined by a finite number of inequalities. But the antiblocker concept applies also to sets with curved boundaries. For example, let

$$X = \{\, x \geq 0 \bigm| \|x\| \leq 1 \,\} \tag{30.7}$$

be the intersection of the unit ball and the nonnegative orthant. Cauchy's inequality implies that $x \cdot y \leq 1$ whenever $\|x\| \leq 1$ and $\|y\| \leq 1$, hence $X \subseteq \mathrm{abl}\, X$. And if $y \in \mathrm{abl}\, X$ we have $y \in X$, since $y \neq 0$ implies $\|y\| = y \cdot (y/\,\|y\|) \leq 1$. Therefore $X = \mathrm{abl}\, X$.

In fact, the set X in (30.7) is the only set that equals its own antiblocker. If $Y = \mathrm{abl}\, Y$ and $y \in Y$ we have $y \cdot y \leq 1$, hence $Y \subseteq X$; this implies $\mathrm{abl}\, Y \supseteq X$.

31. Perfect graphs. Let $\omega(G)$ be the size of a largest clique in G. The graph G is called *perfect* if every induced subgraph G' of G can be colored with $\omega(G')$ colors. (See section 15 for the notion of induced subgraph. This definition of perfection was introduced by Claude Berge in 1961.)

Let G^+ be G with vertex v duplicated, as described in section 16. This means we add a new vertex v' with the same neighbors as v and with v — v'.

Lemma. *If G is perfect, so is G^+.*

Proof. Any induced subgraph of G^+ that is not G^+ itself is either an induced subgraph of G (if it omits v or v' or both), or has the form G'^+ for some induced subgraph G' of G (if it retains v and v'). Therefore it suffices to prove that G^+ can be colored with $\omega(G^+)$ colors.

Color G with $\omega(G)$ colors and suppose v is red. Let G' be the subgraph induced from G by leaving out all red vertices except v. Recolor G' with $\omega(G')$ colors, and assign a new color to the set $G^+\backslash G'$, which is stable in G^+. This colors G^+ with $\omega(G')+1$ colors, hence $\omega(G^+) \le \omega(G')+1$.

We complete the proof by showing that $\omega(G^+) = \omega(G')+1$. Let Q be a clique of size $\omega(G')$ in G'.

Case 1. $v \in Q$. Then $Q \cup \{v'\}$ is a clique of G^+.

Case 2. $v \notin Q$. Then Q contains no red element.

In both cases we can conclude that $\omega(G^+) \ge \omega(G')+1$. □

Theorem. *If G is perfect,* $\mathtt{STAB}(G) = \mathtt{QSTAB}(G)$.

Proof. It suffices to prove that every $x \in \mathtt{QSTAB}(G)$ with *rational* coordinates is a member of $\mathtt{STAB}(G)$, because $\mathtt{STAB}(G)$ is a closed set.

Suppose $x \in \mathtt{QSTAB}(G)$ and qx has integer coordinates. Let G^+ be the graph obtained from G by repeatedly duplicating vertices until each original vertex v of G has been replaced by a clique of size qx_v. Call the vertices of that clique the *clones* of v.

By definition of $\mathtt{QSTAB}(G)$, if Q is any clique of G we have

$$\sum_{v\in Q} x_v \le 1\,.$$

Every clique Q' of G^+ is contained in a clique of size $\sum_{v\in Q} qx_v$ for some clique Q of G. (Including all clones of each element yields this possibly larger clique.) Thus $\omega(G^+) \le q$, and the lemma tells us that G^+ can be colored with q colors because G^+ is perfect.

For each color k, where $1 \le k \le q$, let $x_v^{(k)} = 1$ if some clone of v is colored k, otherwise $x_v^{(k)} = 0$. Then $x^{(k)}$ is a stable labeling. Hence

$$\frac{1}{q}\sum_{k=1}^{q} x^{(k)} \in \mathtt{STAB}(G)\,.$$

But every vertex of G^+ is colored, so $\sum_{k=1}^{q} x_v^{(k)} = qx_v$ for all v, so $q^{-1}\sum_{k=1}^{q} x^{(k)} = x$. □

32. A characterization of perfection. The converse of Theorem 31 is also true; but before we prove it we need another fact about convex polyhedra.

Lemma. *Suppose P is the set $\{\, x \ge 0 \mid x\cdot z \le 1$ for all $z \in Z\,\} = \operatorname{abl} Z$ for some finite set Z and suppose $y \in \operatorname{abl} P$, i.e., y is a nonnegative vector such that $x \cdot y \le 1$ for all $x \in P$. Then the set*

$$Q = \{\, x \in P \mid x\cdot y = 1\,\} \tag{32.1}$$

is contained in the set $\{\, x \mid x\cdot z = 1\,\}$ for some $z \in Z$ (unless Q and Z are both empty).

Proof. This lemma is "geometrically obvious"—it says that every vertex, edge, etc., of a convex polyhedron is contained in some "facet"—but we ought also to prove it. The proof is by induction on $|Z|$. If Z is empty, the result holds because P is the set of all nonnegative x, hence y must be 0 and Q must be empty.

Suppose z is an element of Z that does not satisfy the condition; i.e., there is an element $x \in P$ with $x \cdot y = 1$ and $x \cdot z \neq 1$. Then $x \cdot z < 1$. Let $Z' = Z \setminus \{z\}$ and $P' = \operatorname{abl} Z'$. It follows that $x' \cdot y \leq 1$ for all $x' \in P'$. For if $x' \cdot y > 1$, a convex combination $x'' = \epsilon x + (1-\epsilon)x'$ will lie in P for sufficiently small ϵ, but $x'' \cdot y > 1$.

Therefore by induction, $Q' = \{\, x \in P' \mid x \cdot y = 1 \,\}$ is contained in $\{\, x \mid x \cdot z' = 1 \,\}$ for some $z' \in Z'$, unless Q' is empty, when we can take $z' = z$. And $Q \subseteq Q'$, since $P \subseteq P'$. □

Theorem. *G is perfect if and only if* $\mathtt{STAB}(G) = \mathtt{QSTAB}(G)$.

Proof. As in section 15, let $G|U$ be the graph induced from G by restriction to vertices U. If X is a set of vectors indexed by V and if $U \subseteq V$, let $X|U$ be the set of all vectors indexed by U that arise from the vectors of X when we suppress all components x_v with $v \notin U$. Then it is clear that

$$\mathtt{QSTAB}(G|U) = \mathtt{QSTAB}(G)|U\,, \tag{32.2}$$

because every $x \in \mathtt{QSTAB}(G|U)$ belongs to $\mathtt{QSTAB}(G)$ if we set $x_v = 0$ for $v \notin U$, and every $x \in \mathtt{QSTAB}(G)$ satisfies $\sum_{v\in Q} x_v \leq 1$ for every clique $Q \subseteq U$. Also

$$\mathtt{STAB}(G|U) = \mathtt{STAB}(G)|U\,, \tag{32.3}$$

because every stable labeling of $G|U$ is a stable labeling of G if we extend it with zeros, and every stable labeling of G is stable for $G|U$ if we ignore components not in U.

Therefore $\mathtt{STAB}(G) = \mathtt{QSTAB}(G)$ iff $\mathtt{STAB}(G') = \mathtt{QSTAB}(G')$ for all induced graphs. By Theorem 31 we need only prove that $\mathtt{STAB}(G) = \mathtt{QSTAB}(G)$ implies G can be colored with $\omega(G)$ colors.

Suppose $\mathtt{STAB}(G) = \mathtt{QSTAB}(G)$. Then by Corollary 30,

$$\mathtt{STAB}(\overline{G}) = \mathtt{QSTAB}(\overline{G})\,. \tag{32.4}$$

Let $P = \mathtt{STAB}(\overline{G})$, and let $y = \mathbb{1}/\omega(G)$. Then $x \cdot y \leq 1$ whenever x is a clique labeling of G, i.e., whenever x is a stable labeling of $\overline{G}$; so $x \cdot y \leq 1$ for all $x \in P$. Let Z be the set of all stable labelings of G, i.e., clique labelings of $\overline{G}$. Then $P = \mathtt{QSTAB}(\overline{G}) = \operatorname{abl} Z$ and Z is nonempty. So the lemma applies and it tells us that the set Q defined in (32.1) is contained in $\{\, x \mid x \cdot z = 1 \,\}$ for some stable labeling z of G. Therefore every maximum clique labeling x satisfies $x \cdot z = 1$; i.e., every clique of size $\omega(G)$ intersects the stable set S corresponding to z. So $\omega(G') = \omega(G) - 1$, where

$$G' = G|(V \setminus S)\,. \tag{32.5}$$

By induction on $|V|$ we can color the vertices of G' with $\omega(G')$ colors, then we can use a new color for the vertices of S; this colors G with $\omega(G)$ colors. □

Lovász states in [13] that he knows no polynomial time algorithm to test if G is perfect; but he conjectures ("guesses") that such an algorithm exists, because the results we are going to discuss next suggest that much more might be provable.

33. Another definition of ϑ. The following result generalizes Lemma 9.3.21 of [7].

Lemma. *Let a and b be orthogonal labelings of G and $\overline{G}$ that satisfy the conditions of Theorem 12, normalized so that*

$$\|a_v\|^2\|b_v\|^2 = w_v c(b_v)\,, \qquad a_{1v} \ge 0\,, \qquad \textit{and} \qquad b_{1v} \ge 0\,, \tag{33.1}$$

for all v. Then

$$\sum_v a_{jv} b_{kv} = \begin{cases} \sqrt{\vartheta(G,w)}\,, & \text{if } j = k = 1; \\ 0, & \text{otherwise.} \end{cases} \tag{33.2}$$

Proof. Let $a_0 = (\sqrt{\vartheta}, 0, \ldots, 0)^T$ and $b_0 = (-1, 0, \ldots, 0)^T$. Then the $(n+1) \times (n+1)$ matrices $A = a^T a$ and $B = b^T b$ are spuds, and $A \cdot B = 0$. (In the special case $\|a_v\|^2 = w_v$, $\|b_v\|^2 = c(b_v)$, matrix B is what we called B' in the proof of Theorem 29.) Therefore $0 = \operatorname{tr} A^T B = \operatorname{tr} a^T a b^T b = \operatorname{tr} b a^T a b^T = (ab^T) \cdot (ab^T)$, and we have $ab^T = 0$. In other words

$$a_{j0} b_{k0} + \sum_v a_{jv} b_{kv} = 0$$

for all j and k. □

We now can show that ϑ has yet another definition, in some ways nicer than the one we considered in section 6. (Someday I should try to find a simpler way to derive all these facts.) Call the matrix B *dual feasible* for G and w if it is indexed by vertices and

$$\begin{aligned} &B \text{ is real and symmetric;} \\ &B_{vv} = w_v \text{ for all } v \in V; \\ &B_{uv} = 0 \text{ whenever } u \not\frown v \text{ in } G; \end{aligned} \tag{33.3}$$

and define

$$\vartheta_6(G,w) = \max\{\, \Lambda(B) \mid B \text{ is positive semidefinite and dual feasible for } G \text{ and } w \,\}\,. \tag{33.4}$$

(Compare with the analogous definitions in (6.1) and (6.3).)

Theorem. $\vartheta(G,w) = \vartheta_6(G,w)$.

Proof. If B is positive semidefinite and dual feasible, and if λ is any eigenvalue of B, we can write $B = QDQ^T$ where Q is orthogonal and D is diagonal, with $D_{11} = \lambda$. Let $b = \sqrt{D}\,Q^T$; then b is an orthogonal labeling of $\overline{G}$ with $\|b_v\|^2 = w_v$ for all v. Furthermore $c(b_v) = b_{1v}^2/w_v = \lambda\, q_{v1}^2/w_v$, where $(q_{11},\ldots,q_{n1})$ is the first column of Q. Therefore $\sum_v c(b_v)w_v = \lambda\sum_v q_{v1}^2 = \lambda$, and we have $\lambda \le \vartheta_4(G,w)$ by (10.1). This proves that $\vartheta_6 \le \vartheta$.

Conversely, let a and b be orthogonal labellings of G and $\overline{G}$ that satisfy the conditions of Theorem 12. Normalize them so that $\|a_v\|^2 = c(b_v)$ and $\|b_v\|^2 = w_v$. Then $a_{1v}^2 = c(a_v)c(b_v) = w_v c(b_v)/\vartheta = b_{1v}^2/\vartheta$. The lemma now implies that $(b_{11},\ldots,b_{1n})^T$ is an eigenvector of b^Tb, with eigenvalue ϑ. This proves that $\vartheta \le \vartheta_6$. $\square$

Corollary. $\vartheta(G) = 1 + \max\{\,\Lambda(B)/\Lambda(-B) \mid B$ *is dual feasible for G and 0*$\}$.

Proof. If B is dual feasible for G and 0, its eigenvalues are $\lambda_1 \ge \cdots \ge \lambda_n$ where $\lambda_1 = \Lambda(B)$ and $\lambda_n = -\Lambda(-B)$. Then $B' = I + B/\Lambda(-B)$ has eigenvalues $1-\lambda_1/\lambda_n,\ldots,$ $1-\lambda_n/\lambda_n = 0$. Consequently B' is positive semidefinite and dual feasible for G and $\mathbb{1}$, and $1 + \Lambda(B)/\Lambda(-B) = \Lambda(B') \le \vartheta_6(G)$.

Conversely, suppose B' is positive semidefinite and dual feasible for G and $\mathbb{1}$, with $\Lambda(B') = \vartheta = \vartheta(G)$. Let $B = B' - I$. Then B is dual feasible for G and 0, and $0 \le \Lambda(-B) \le 1$ since the sum of the eigenvalues of B is $\operatorname{tr} B = 0$. It follows that $\vartheta - 1 = \Lambda(B) \le \Lambda(B)/\Lambda(-B)$. $\square$

34. Facets of TH. We know that $\mathtt{TH}(G)$ is a convex corner set in n-dimensional space, so it is natural to ask whether it might have $(n-1)$-dimensional facets on its nontrivial boundary—for example, a straight line segment in two dimensions, or a region of a plane in three dimensions. This means it has n linearly independent vectors $x^{(k)}$ such that

$$\sum_v x_v^{(k)} c(a_v) = 1 \tag{34.1}$$

for some orthogonal labeling a of G.

Theorem. *If $\mathtt{TH}(G)$ contains linearly independent solutions $x^{(1)},\ldots,x^{(n)}$ of (34.1), then there is a maximal clique Q of G such that*

$$c(a_v) = \begin{cases} 1\,, & v \in Q;\\ 0\,, & v \notin Q.\end{cases} \tag{34.2}$$

Proof. Theorem 14 tell us that every $x^{(k)} \in \mathtt{TH}(G)$ has $x_v^{(k)} = c(b_v^{(k)})$ for some orthogonal labeling of $\overline{G}$. Set $w_v = c(a_v)$; then $\vartheta(G,w) = 1$, by Theorem 13. We can normalize the labelings so that $\|a_v\|^2 = a_{1v} = w_v$ and $\|b_v^{(k)}\|^2 = b_{1v}^{(k)} = x_v^{(k)}$. Hence, by Lemma 33,

$$\sum_v x_v^{(k)} a_v = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = e_1 \,. \tag{34.3}$$

Let

$$Q = \{\, v \mid a_{1v} \neq 0 \,\} = \{\, v \mid c(a_v) \neq 0 \,\} \tag{34.4}$$

and suppose Q has m elements. Then (34.3) is equivalent to the matrix equation

$$Ax^{(k)} = e_1 \tag{34.5}$$

where A is a $d \times m$ matrix and $x^{(k)}$ has m components $x_v^{(k)}$, one for each $v \in Q$. By hypothesis there are m linearly independent solutions to (34.5), because there are n linearly independent solutions to (34.3). But then there are $m-1$ linearly independent solutions to $Ax = 0$, and it follows that A has rank 1: Every row of A must be a multiple of the top row (which is nonzero). And then (34.5) tells us that all rows but the top row are zero. We have proved that

$$c(a_v) \neq 0 \;\Rightarrow\; c(a_v) = 1 \,. \tag{34.6}$$

Therefore if u and v are elements of Q we have $a_u \cdot a_v \neq 0$, hence u — v; Q is a clique.

Moreover, Q is maximal. For if $v \notin Q$ is adjacent to all elements of Q, there is a k such that $x_v^{(k)} > 0$. But the characteristic labeling of $Q \cup \{v\}$ is an orthogonal labeling a' such that $\sum_u x_u^{(k)} c(a'_u) = 1 + x_v^{(h)} > 1$, hence $x^{(k)} \notin \mathtt{TH}(G)$. □

Conversely, it is easy to see that the characteristic labeling of any maximal clique Q does have n linearly independent vectors satisfying (34.1), so it does define a facet. For each vertex u we let $x_u^{(u)} = 1$, and $x_v^{(u)} = 0$ for all $v \neq u$ except for one vertex $v \in Q$ with $v \not\!\!— u$ (when $u \notin Q$). Then $x^{(u)}$ is a stable labeling so it is in $\mathtt{TH}(G)$. The point of the theorem is that a constraint $\sum_v x_v c(a_v) \leq 1$ of $\mathtt{TH}(G)$ that is not satisfied by all $x \in \mathtt{QSTAB}(G)$ cannot correspond to a facet of $\mathtt{TH}(G)$.

Corollary.

$$\begin{aligned} \mathtt{TH}(G) \textit{ is a polytope} &\iff \mathtt{TH}(G) = \mathtt{QSTAB}(G) \\ &\iff \mathtt{TH}(G) = \mathtt{STAB}(G) \iff G \text{ is perfect.} \end{aligned}$$

Proof. If $\mathtt{TH}(G)$ is a polytope it is defined by facets as in the theorem, which are nothing more than the constraints of $\mathtt{QSTAB}(G)$; hence $\mathtt{TH}(G) = \mathtt{QSTAB}(G)$. Also the antiblocker of a convex corner polytope is a polytope, so $\mathtt{TH}(\overline{G})$ is a polytope by (30.4); it must be equal to $\mathtt{QSTAB}(\overline{G})$. Taking antiblockers, we have $\mathtt{TH}(G) = \mathtt{STAB}(G)$ by (30.3). The converses are easy since STAB and QSTAB are always polytopes. The connection to perfection is an immediate consequence of Theorem 32 and Lemma 2. □

We cannot strengthen the corollary to say that $\vartheta(G) = \alpha(G)$ holds if and only if $\vartheta(G) = \kappa(G)$; the Petersen graph (section 27) is a counterexample.

35. Orthogonal labelings in a perfect graph. A perfect graph has

$$\vartheta(G, w) = \alpha(G, w) = \max\{x \cdot w \mid x \text{ is a stable labeling of } G\}\,, \tag{35.1}$$

and Theorem 12 tells us there exist orthogonal labelings of G and $\overline{G}$ such that (12.2) and (12.3) hold. But it isn't obvious what those labelings might be; the proof was not constructive.

The problem is to find vectors a_v such that $a_u \cdot a_v = 0$ when $u \not\to v$ and such that (12.2) holds; then it is easy to satisfy (12.3) by simply letting b be a stable labeling where the maximum occurs in (35.1).

The following general construction gives an orthogonal labeling (not necessarily optimum) in any graph: Let $g(Q)$ be a nonnegative number for every clique Q, chosen so that

$$\sum_{v \in Q} g(Q) = w_v\,, \qquad \text{for all } v\,. \tag{35.2}$$

Furthermore, for each clique Q, let

$$a_{Qv} = \begin{cases} \sqrt{g(Q)}, & \text{if } v \in Q; \\ 0, & \text{otherwise.} \end{cases} \tag{35.3}$$

Then

$$a_u \cdot a_v \;=\; \sum_{\{u,v\} \subseteq Q} g(Q)\,,$$

hence $a_u \cdot a_v = 0$ when $u \not\to v$. If we also let $a_{Q0} = \sqrt{g(Q)}$ for all Q, $a_{00} = 0$, we find

$$a_0 \cdot a_v = a_v \cdot a_v = \sum_{v \in Q} g(Q) = w_v\,.$$

We have constructed a matrix A that is λ-compatible with G and w, in the sense of section 29, where

$$\lambda = a_0 \cdot a_0 = \sum_{Q} g(Q)\,. \tag{35.4}$$

An orthogonal labeling with costs $c(a'_v) = w_v/\lambda$ can now be found as in the proof of Lemma 29.

The duality theorem of linear programming tells us that the minimum of (35.4) subject to the constraints (35.2) is equal to the maximum value of $w \cdot x$ over all x with $\sum_{v \in Q} x_v \le 1$ for all Q. When x maximizes $w \cdot x$, we can assume that $x \ge 0$, because a negative x_v can be replaced by 0 without decreasing $w \cdot x$ or violating a constraint. (Every subset of a clique is a clique.) Thus, we are maximizing $w \cdot x$ over $\mathtt{QSTAB}(G)$; the construction in the previous paragraph allows us to reduce λ as low as $\kappa(G, w)$. But $\kappa(G, w) = \vartheta(G, w)$ in a perfect graph, so this construction solves our problem, once we have computed $g(Q)$.

The special case of a bipartite graph is especially interesting, because its cliques have only one or two vertices. Suppose all edges of G have the form u — v where $u \in U$ and $v \in V$, and consider the network defined as follows: There is a special source vertex s connected to all $u \in U$ by a directed arc of capacity w_u, and a special sink vertex t connected from all $v \in V$ by a directed arc of capacity w_v. The edges u — v of G are also present, directed from u to v with infinite capacity. Any flow from s to t in this network defines a suitable function g, if we let

$$\begin{aligned} g(\{u,v\}) &= \text{the flow in } u \to v\,, \\ g(\{u\}) &= w_u \text{ minus the flow in } s \to u\,, \\ g(\{v\}) &= w_v \text{ minus the flow in } v \to t\,, \end{aligned}$$

for all $u \in U$ and $v \in V$. Let S be a subset of $U \cup V$. If we cut the edges that connect s or t with vertices not in S, we cut off all paths from s to t if and only if S is a stable set. The minimum cut (i.e., the minimum sum of capacities of cut edges) is equal to the maximum flow; and it is also equal to

$$\sum_{u \in U} w_u + \sum_{v \in V} w_v - \max\{w \cdot x \mid x \text{ is a stable labeling}\} = \sum_{u \in U} w_u + \sum_{v \in V} w_v - \alpha(G, w)\,.$$

Thus the value of $\lambda = \sum_Q g(Q)$ is $\sum_{u \in U} w_u - \{\text{flow from } s\} + \sum_{v \in V} w_v - \{\text{flow to } t\} + \{\text{flow in } u \to v \text{ arcs}\} = \alpha(G, w) = \vartheta(G, w)$ as desired.

For general perfect graphs G, a solution to (35.4) with $\lambda = \vartheta(G, w)$ can be found in polynomial time as shown in equation (9.4.6) of [7]. However, the methods described in [7] are not efficient enough for practical calculation, even on small graphs.

36. The smallest non-perfect graph. The cyclic graph C_5 is of particular interest because it is the smallest graph that isn't perfect, and the smallest case where the function $\vartheta(G, w)$ is not completely known.

The discussion following Theorem 34 points out that $\mathtt{TH}(G)$ always has facets in common with $\mathtt{QSTAB}(G)$, when those facets belong also to $\mathtt{STAB}(G)$. It is not hard to see that

$\mathtt{QSTAB}(C_5)$ has ten facets, defined by $x_j = 0$ and $x_j + x_{j \bmod 5} = 1$ for $0 \le j < 5$; and $\mathtt{STAB}(C_5)$ has an additional facet defined by $x_0 + x_1 + x_2 + x_3 + x_4 = 2$. The weighted functions α and κ of section 4 are evaluated by considering the vertices of $\mathtt{STAB}$ and $\mathtt{QSTAB}$:

$$\alpha(C_5, \{w_0, \ldots, w_4\}^T) = \max(w_0 + w_2, w_1 + w_3, w_2 + w_4, w_3 + w_0, w_4 + w_1)\,; \tag{36.1}$$
$$\kappa(C_5, \{w_0, \ldots, w_4\}^T) = \max\bigl(\alpha(C_5, \{w_0, \ldots, w_4\}^T), (w_0 + \cdots + w_4)/2\bigr)\,. \tag{36.2}$$

Where these functions agree, they tell us also the value of ϑ.

For example, let $f(x) = \vartheta(C_5, \{x, 1, 1, 1, 1, 1\}^T)$. Relations (36.1) and (36.2) imply that $f(x) = x + 1$ when $x \ge 2$. Clearly $f(0) = 2$, and section 22 tells us that $f(1) = \sqrt{5}$. Other values of $f(x)$ are not yet known. Equation (23.2) gives the lower bound $f(x)^2 \ge x^2 + 4$. Incidentally, the a vectors

$$\begin{pmatrix} \sqrt{x} \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \begin{pmatrix} \sqrt{x} \\ 1 \\ \sqrt{x+1} \\ \sqrt{(x-2)(x+1)} \\ 0 \end{pmatrix} \quad \begin{pmatrix} 1 \\ -\sqrt{x} \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \begin{pmatrix} 1 \\ -\sqrt{x} \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \begin{pmatrix} \sqrt{x} \\ 1 \\ -\sqrt{x+1} \\ 0 \\ \sqrt{(x-2)(x+1)} \end{pmatrix}$$

and $b = (1)\,(0)\,(0)\,(0)\,(0)$ establish $f(x)$ for $x \ge 2$ in the fashion of Theorems 12 and 13.

Let $\phi = (1 + \sqrt{5}\,)/2$ be the golden ratio. The matrices A and B' of Theorem 29, when $G = C_5$ and $w = \mathbb{1}$, are

$$A = \begin{pmatrix} \sqrt{5} & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & \phi-1 & 0 & 0 & \phi-1 \\ 1 & \phi-1 & 1 & \phi-1 & 0 & 0 \\ 1 & 0 & \phi-1 & 1 & \phi-1 & 0 \\ 1 & 0 & 0 & \phi-1 & 1 & \phi-1 \\ 1 & \phi-1 & 0 & 0 & \phi-1 & 1 \end{pmatrix};$$

$$B' = \frac{1}{\sqrt{5}} \begin{pmatrix} \sqrt{5} & -1 & -1 & -1 & -1 & -1 \\ -1 & 1 & 0 & \phi-1 & \phi-1 & 0 \\ -1 & 0 & 1 & 0 & \phi-1 & \phi-1 \\ -1 & \phi-1 & 0 & 1 & 0 & \phi-1 \\ -1 & \phi-1 & \phi-1 & 0 & 1 & 0 \\ -1 & 0 & \phi-1 & \phi-1 & 0 & 1 \end{pmatrix}.$$

They have the common eigenvectors

$$\begin{pmatrix} \sqrt{5} \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad \begin{pmatrix} \sqrt{5} \\ -1 \\ -1 \\ -1 \\ -1 \\ -1 \end{pmatrix} \quad \begin{pmatrix} 0 \\ \phi \\ 1 \\ -1 \\ -\phi \\ 0 \end{pmatrix} \quad \begin{pmatrix} 0 \\ 0 \\ \phi \\ 1 \\ -1 \\ -\phi \end{pmatrix} \quad \begin{pmatrix} 0 \\ 1 \\ -\phi \\ \phi \\ -1 \\ 0 \end{pmatrix} \quad \begin{pmatrix} 0 \\ 0 \\ 1 \\ -\phi \\ \phi \\ -1 \end{pmatrix},$$

with respective eigenvalues $(\lambda_0, \ldots, \lambda_5) = (2\sqrt{5}, 0, \sqrt{5}/\phi, \sqrt{5}/\phi, 0, 0)$ and $(\mu_0, \ldots, \mu_5) = (0, 2, 0, 0, 1/\phi, 1/\phi)$. (Cf. (29.6) and (29.7).)

37. Perplexing questions. The book [7] explains how to compute $\vartheta(G, w)$ with given tolerance ϵ, in polynomial time using an ellipsoid method, but that method is too slow and numerically unstable to deal with graphs that have more than 10 or so vertices. Fortunately, however, new "interior-point methods" have been developed for this purpose, especially by Alizadeh [1,2], who has computed $\vartheta(G)$ when G has hundreds of vertices and thousands of edges. He has also shown how to find large stable sets, as a byproduct of evaluating $\vartheta(G, w)$ when w has integer coordinates. Calculations on somewhat smaller cyclically symmetric graphs have also been reported by Overton [16]. Further computational experience with such programs should prove to be very interesting.

Solutions to the following four concrete problems may also help shed light on the subject:

P1. Describe $\mathtt{TH}(C_5)$ geometrically. This upright set is isomorphic to its own antiblocker. (Namely, if $(x_0, x_1, x_2, x_3, x_4) \in \mathtt{TH}(C_5)$, then so are its cyclic permutations $(x_1, x_2, x_3, x_4, x_0)$, etc., as well as the cyclic permutations of $(x_0, x_4, x_3, x_2, x_1)$; $\mathtt{TH}(\overline{C}_5)$ contains the cyclic permutations of $(x_0, x_2, x_4, x_1, x_3)$ and $(x_0, x_3, x_1, x_4, x_2)$.) Can the values $f(x) = \vartheta(C_5, \{x, 1, 1, 1, 1\}^T)$, discussed in section 36, be expressed in closed form when $0 < x < 2$, using familiar functions?

P2. What is the probable value of $\vartheta(G, w)$ when G is a random graph on n vertices, where each of the $\binom{n}{2}$ possible edges is independently present with some fixed probability p? (Juhász [9] has solved this problem in the case $w = \mathbb{1}$, showing that $\vartheta(G)/\sqrt{(1-p)n/p}$ lies between $\frac{1}{2}$ and 2 with probability approaching 1 as $n \to \infty$.)

P3. What is the minimum d for which G almost surely has an orthogonal labeling of dimension d with no zero vectors, when G is a random graph as in Problem P2? (Theorem 28 and the theorem of Juhász [9] show that d must be at least of order $\sqrt{n}$. But Lovász tells me that he suspects the correct answer is near n. Theorem 29 and its consequences might be helpful here.)

P4. Is there a constant c such that $\vartheta(G) \le c\sqrt{n}\,\alpha(G)$ for all n-vertex graphs G? (This conjecture was suggested by Lovász in a recent letter. He knows no infinite family of graphs where $\vartheta(G)/\alpha(G)$ grows faster than $O(\sqrt{n}/\log n)$. The latter behavior occurs for random graphs, which have $\alpha(G) = \log_{1/p} n$ with high probability [4, Chapter XI].)

Another, more general, question is to ask whether it is feasible to study two- or three-dimensional projections of $\mathtt{TH}(G)$, and whether they have combinatorial significance. The function $\vartheta(G, w)$ gives just a one-dimensional glimpse.

Lovász and Schrijver have recently generalized the topics treated here to a wide variety of more powerful techniques for studying 0–1 vectors associated with graphs [15]. In particular, one of their methods can be described as follows: Let us say that a *strong orthogonal labeling* is a vector labeling such that $\|a_v\|^2 = c(a_v)$ and $a_u \cdot a_v \geq 0$, also satisfying the relation

$$c(a_u) + c(a_v) + c(a_w) - 1 \leq a_u \cdot a_v + a_v \cdot a_w \leq c(a_v) \tag{37.1}$$

whenever $u \not\!\!- w$. In particular, when $w = v$ this relation implies that $a_u \cdot a_v = 0$, so the labeling is orthogonal in the former sense.

Notice that every stable labeling is a strong orthogonal labeling of $\overline{G}$. Let S be a stable set and let u and w be vertices such that u — w. If u and w are not in S, condition (37.1) just says that $0 \leq c(a_v) \leq 1$, which surely holds. If u is in S, then $w \notin S$ and (37.1) reduces to $c(a_v) \leq c(a_v) \leq c(a_v)$; this holds even more surely.

Let

$$\mathtt{TH}_{-}(G) = \{\, x \mid x_v = c(b_v) \text{ for some strong orthogonal labeling of } \overline{G}\,\}. \tag{37.2}$$

(This set is called $N_+(\mathtt{FR}(G))$ in [15].) We also define

$$\vartheta_{-}(G, w) = \max\{\, w \cdot x \mid x \in \mathtt{TH}_{-}(G)\,\}\,. \tag{37.3}$$

The argument in the two previous paragraphs implies that

$$\mathtt{STAB}(G) \subseteq \mathtt{TH}_{-}(G) \subseteq \mathtt{TH}(G)\,,$$

hence

$$\alpha(G, w) \leq \vartheta_{-}(G, w) \leq \vartheta(G, w)\,. \tag{37.4}$$

The authors of [15] prove that $\vartheta_{-}(G, w)$ can be computed in polynomial time, about as easily as $\vartheta(G, w)$; moreover, it can be a significantly better approximation to $\alpha(G, w)$. They show, for example, that $\mathtt{TH}_{-}(G) = \mathtt{STAB}(G)$ when G is any cyclic graph C_n. In fact, they prove that if $x \in \mathtt{TH}_{-}(G)$ and if v_0 — v_1, v_1 — v_2, …, v_{2n} — v_0 is any circuit or multicircuit of G, then $x_{v_0} + x_{v_1} + \cdots + x_{v_{2n}} \leq n$. This suggests additional research problems:

P5. What is the smallest graph such that $\mathtt{STAB}(G) \neq \mathtt{TH}_{-}(G)$?

P6. What is the probable value of $\vartheta_{-}(G)$ when G is a random graph as in Problem P2?

A recent theorem by Arora, Lund, Motwani, Sudan, and Szegedy [3] proves that there is an $\epsilon > 0$ such that no polynomial algorithm can compute a number between $\alpha(G)$ and

$n^\epsilon\alpha(G)$ for all n-vertex graphs G, unless $P = NP$. Therefore it would be surprising if the answer to P6 turns out to be that $\vartheta_-(G)$ is, say, $O(\log n)^2$ with probability $\to 1$ for random G. Still, this would not be inconsistent with [3], because the graphs for which $\alpha(G)$ is hard to approximate might be decidedly nonrandom.

Lovász has called my attention to papers by Kashin and Konyagin [10, 11], which prove (in a very disguised form, related to (6.2) and Theorem 33) that if G has no stable set with 3 elements we have

$$\vartheta(G) \le 2^{2/3} n^{1/3}\,; \tag{37.5}$$

moreover, such graphs exist with

$$\vartheta(G) = \Omega(n^{1/3}/\sqrt{\log n}\,)\,. \tag{37.6}$$

Further study of methods like those in [15] promises to be exciting indeed. Lovász has sketched yet another approach in [14].

References

[1] Farid Alizadeh, "A sublinear-time randomized parallel algorithm for the maximum clique problem in perfect graphs," *ACM-SIAM Symposium on Discrete Algorithms* **2** (1991), 188–194.

[2] Farid Alizadeh, "Interior point methods in semidefinite programming with applications to combinatorial optimization," preprint, 1993, available from `alizadeh@icsi.berkeley.edu`.

[3] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy, "Proof verification and intractability of approximation problems," *Proceedings of the 33rd IEEE Symposium on Foundations of Computer Science* (1992), 14–23.

[4] Béla Bollobás, *Random Graphs* (Academic Press, 1985).

[5] Martin Grötschel, L. Lovász, and A. Schrijver, "The ellipsoid method and its consequences in combinatorial optimization," *Combinatorica* **1** (1981), 169–197.

[6] M. Grötschel, L. Lovász. and A. Schrijver, "Relaxations of vertex packing," *Journal of Combinatorial Theory* **B40** (1986), 330–343.

[7] Martin Grötschel, László Lovász, and Alexander Schrijver, *Geometric Algorithms and Combinatorial Optimization* (Springer-Verlag, 1988), §9.3.

[8] A. S. Householder, "Unitary triangularization of a nonsymmetric matrix," *Journal of the Association for Computing Machinery* **5** (1958), 339–342.

[9] Ferenc Juhász, "The asymptotic behaviour of Lovász' ϑ function for random graphs," *Combinatorica* **2** (1982), 153–155.

[10] B. S. Kashin and S. V. Koni͡agin, "O sistemakh vektorov v Gil'bertovom prostranstve," *Trudy Matematicheskogo Instituta imeni V. A. Steklova* **157** (1981), 64–67. English translation, "On systems of vectors in a Hilbert space," *Proceedings of the Steklov Institute of Mathematics* (American Mathematical Society, 1983, issue 3), 67–70.

[11] S. V. Koni͡agin, "O sistemakh vektorov v Evclidovom prostranstve i odnoĭ ékstremal'-noĭ zadache dli͡a mnogochlenov," *Matematicheskie Zametki* **29** (1981), 63–74. English translation, "Systems of vectors in Euclidean space and an extremal problem for polynomials," *Mathematical Notes of the Academy of Sciences of the USSR* **29** (1981), 33–39.

[12] L. Lovász, "On the Shannon capacity of a graph," *IEEE Transactions on Information Theory* **IT–25** (1979), 1–7.

[13] L. Lovász, An Algorithmic Theory of Numbers, Graphs, and Convexity. *CBMS Regional Conference Series in Applied Mathematics* (SIAM, 1986), §3.2.

[14] L. Lovász, "Stable sets and polynomials," to appear in *Discrete Mathematics*.

[15] L. Lovász and A. Schrijver, "Cones of matrices and set functions and 0–1 optimization," *SIAM Journal on Optimization* **1** (1991), 166–190.

[16] Michael Overton, "Large-scale optimization of eigenvalues," *SIAM Journal on Optimization* **2** (1992), 88–120.

[17] M. W. Padberg, "On the facial structure of set packing polyhedra," *Mathematical Programming* **5** (1973), 199–215.

[18] Claude E. Shannon, "The zero error capacity of a channel," *IRE Transactions on Information Theory* **2**, 3 (Sep 1956), 8–19.