Orthogonal Representations

1 Introductory example: Shannon capacity

The following problem in information theory was raised by Claude Shannon. Consider a noisy channel through which we are sending messages composed of a finite alphabet $V$. There is an output alphabet $U$, and each $v \in V$, when transmitted through the channel, can come out as any element in a set $U_v \subseteq U$. Usually there is a probability distribution specified on each set $U_v$, telling us the probability with which $v$ produces a given $u \in U_v$, but for the problem we want to discuss, these probabilities do not matter. As a matter of fact, the output alphabet will play no role either, except to tell us which pairs of input characters can be confused: those pairs $(v, v')$ for which $U_v \cap U_{v'} \neq \emptyset$.

One way to model the problem is as follows: We consider $V$ as the set of nodes of a graph, and connect two of them by an edge if they can be confused. This way we obtain a graph $G$, which we call the confusion graph of the alphabet. The maximum number of non-confusable messages of length 1 is the maximum number of nonadjacent nodes (the maximum size of a stable set) in the graph $G$, which we denote by $\alpha(G)$.

Now we consider longer messages, say of length $k$. We want to select as many of them as possible so that no two of them can possibly be confused. This means that for any two of these selected messages, there should be a position, where the two characters are not confusable. As we shall see, the number of words we can select grows as $\Theta^k$ for some $\Theta \geq 1$, which is called the Shannon zero-error capacity of the channel.

A simple and natural way to create such a set of words is to pick a non-confusable subset of the alphabet, and use only those words composed from this set. So if we have $\alpha$ non-confusable characters in our alphabet, then we can create $\alpha^k$ non-confusable messages of length $k$. But, as we shall see, making use of other characters in the alphabet we can create more! How much more, is the issue in this discussion.

Let us look at two simple examples (Figure 1).

![Figure 1: Two confusion graphs. In the alphabet \{p, q, b, d\} two letters that are related by a reflection in a horizontal or vertical line are confusable, but not if they are related by two such reflection. The confusability graph of the alphabet \{m, n, u, v, w\} is only convincing a little in handwriting, but this graph plays an important role in this book.](image-url)
Example 1. Consider the simple alphabet \((p, q, d, b)\), where the pairs \(\{p, q\}\), \(\{q, d\}\), \(\{d, b\}\) and \(\{b, p\}\) are confusable (Figure 1, left). We can just keep \(p\) and \(d\) (which are not confusable), which allows us to create \(2^k\) non-confusable messages of length \(k\). On the other hand, if we use a word, then all the \(2^k\) words obtained from it by replacing some occurrences of \(p\) and \(q\) by the other, as well as some occurrences of \(b\) and \(d\) by the other, are excluded. Hence the number of messages we can use is at most \(4^k / 2^k = 2^k\). ♦

Example 2. [5-cycle] If we switch to alphabets with 5 characters, then we get a much more difficult problem. Let \(V = \{m, n, u, v, w\}\) be our alphabet, with confusable pairs \(\{m, n\}\), \(\{n, u\}\), \(\{u, v\}\), \(\{v, w\}\) and \(\{w, m\}\) (Figure 1, right; we refer to this example as the “pentagon”). Among any three characters there are two that can be confused, so we have only two non-confusable characters. Restricting the alphabet to two such characters (say, \(m\) and \(v\)), we get \(2^k\) non-confusable messages of length \(k\).

But we can do better: the following 5 messages of length two are non-confusable: \(mm, nu, uw, vn\) and \(wv\). This takes some checking: for example, \(mm\) and \(nu\) cannot be confused, because their second characters, \(m\) and \(u\), cannot be confused. If \(k\) is even, then we can construct \(5^k / 2\) non-confusable messages, by concatenating any \(k/2\) of the above 5. This number grows like \((\sqrt{5})^k \approx 236^k\) instead of \(2^k\), a substantial gain! ♦

Can we do better by looking at longer messages (say, messages of length 10 or 1000), and by some \textit{ad hoc} method finding among them more that \(5^5\) non-confusable messages? We are going to show that we cannot, which means that the set of words composed of the above 5 messages of length 2 is optimal.

The trick is to represent the alphabet in a different way. Let us assign to each character \(i \in V\) a vector \(u_i\) in some Euclidean space \(\mathbb{R}^d\). If two characters are not confusable, then we represent them by orthogonal vectors. If a subset of characters \(S\) is non-confusable, then the vectors \(u_i (i \in S)\) are mutually orthogonal unit vectors, and hence for every unit vector \(c\),

\[
\sum_{i \in S} (c^T u_i)^2 \leq 1.
\]

Hence \(|S| \min_{i \in S} (c^T u_i)^2 \leq 1\), or

\[
|S| \leq \max_{i \in V} \frac{1}{(c^T u_i)^2} \leq \max_{i \in V} \frac{1}{(c^T u_i)^2}.
\]

So if we find a representation \(u\) and a unit vector \(c\) for which the squared products \((c^T u_i)^2\) are all large (which means that the angels \(\angle(c, u_i)\) are all small), then we get a good upper bound on \(|S|\).

For the alphabet in Example 2 (the pentagon), we use the 3-dimensional vectors in Figure 2. To describe these, consider an umbrella with 5 ribs of unit length. Open it up to the point when nonconsecutive ribs are orthogonal. This way we get 5 unit vectors \(u_m, u_n, u_u, u_v, u_w\), assigned to the nodes of the pentagon so that each \(u_i\) forms the same angle with the “handle” \(c\) and any
two nonadjacent nodes are labeled with orthogonal vectors. With some effort, one can compute that \((c^Tu_i)^2 = 1/\sqrt{5}\) for every \(i\), and so we get that \(|S| \leq \sqrt{5}\) for every non-confusable set \(S\). Since \(|S|\) is an integer, this implies that \(|S| \leq 2\).

This is ridiculously much work to conclude that the 5-cycle does not contain 3 nonadjacent nodes! But the vector representation is very useful for handling longer messages. We define the tensor product of two vectors \(u = (u_1, \ldots, u_n) \in \mathbb{R}^n\) and \(v = (v_1, \ldots, v_m) \in \mathbb{R}^m\) as the vector

\[
u \circ v = (u_1v_1, \ldots, u_1v_m, u_2v_1, \ldots, u_2v_m, \ldots, u_nv_1, \ldots, u_nv_m)^T \in \mathbb{R}^{nm}.
\]

It is easy to see that \(|u \circ v| = |u||v|\), and (more generally) if \(u, x \in \mathbb{R}^n\) and \(v, y \in \mathbb{R}^m\), then \((u \circ v)^T(x \circ y) = (u^Tx)(v^Ty)\). For a \(k \geq 1\), if we represent a message \(i_1 \ldots i_k\) by the vector \(u_{i_1} \circ \cdots \circ u_{i_k}\), then non-confusable messages will be represented by orthogonal vectors. Indeed, if \(i_1 \ldots i_k\) and \(j_1 \ldots j_k\) are not confusable, then there is at least one subscript \(r\) for which \(i_r\) and \(j_r\) are not confusable, hence \(u^Tu_{i_r} = 0\), which implies that

\[
(u_{i_1} \circ \cdots \circ u_{i_k})^T(u_{j_1} \circ \cdots \circ u_{j_k}) = (u_{i_1}^Tu_{j_1}) \cdots (u_{i_k}^Tu_{j_k}) = 0.
\]

Using \(c \circ \cdots \circ c\) (\(k\) factors) as the “handle”, we get that for any set \(S\) of non-confusable messages of length \(k\),

\[
|S| \leq \max_{i_1, \ldots, i_k} \frac{1}{((c \circ \cdots \circ c)^T(u_{i_1} \circ \cdots \circ u_{i_k}))^2} = \max_{i_1, \ldots, i_k} \frac{1}{(c^Tu_{i_1})^2 \cdots (c^Tu_{i_k})^2} = (\sqrt{5})^k.
\]

So every set of non-confusable messages of length \(k\) has at most \((\sqrt{5})^k\) elements. We have seen that this bound can be attained, at least for even \(k\). Thus we have established that the Shannon zero-error capacity of the pentagon is \(\sqrt{5}\).

2 The theta-function

2.1 Orthogonal representations

An orthogonal representation of a simple graph \(G\) in \(\mathbb{R}^d\) assigns to each \(i \in V\) a vector \(u_i \in \mathbb{R}^d\) such that \(u_i^Tu_j = 0\) whenever \(ij \in \overline{E}\). An orthonormal representation is an orthogonal representation in which all the representing vectors

Figure 2: An umbrella representing the pentagon.
have unit length. Clearly we can always scale the nonzero vectors in an orthogonal representation to unit length, and zero vectors (which are orthogonal to everything) can be either excluded or ignored in most cases.

There is something arbitrary in requiring that nonadjacent nodes be represented by orthogonal vectors; why not adjacent nodes? We use here the more standard convention, but sometimes we need to talk about an orthogonal representation of the complement, which we also call a dual orthogonal representation.

Note that we did not insist that adjacent nodes are mapped onto nonorthogonal vectors.

Example 3. For \( d = 1 \), the vector labels are just real numbers \( u_i \), and the constraints \( u_i u_j = 0 \) (\( ij \in E \)) mean that no two nodes labeled by nonzero numbers are adjacent; in other words, the support of \( u \) is a stable set of nodes. Scaling the nonzero numbers to 1, we obtain incidence vectors of stable sets. (The connection between stable sets and orthonormal representations in higher dimension will be the central topic of the next chapter.)

Since very simple problems about stable sets are NP-hard (for example, their maximum size), this example should warn us that orthogonal representations can be very complex.

Example 4. Every graph has a trivial orthonormal representation in \( \mathbb{R}^V \), in which node \( i \) is represented by the standard basis vector \( e_i \). This representation is not faithful unless the graph has no edges. However, it is easy to perturb this representation to make it faithful. Of course, we are interested in “nontrivial” orthogonal representations, which are more “economical” than this trivial one.

Example 5. Every graph \( G \) has a faithful orthogonal representation in \( \mathbb{R}^E \), in which we label a node by the indicator vector of the set of edges incident with it. It is perhaps natural to expect that this simple representation will be rather “uneconomical” for most purposes.

Example 6. Figure 3 below shows a simple orthogonal representation in 2 dimensions of the graph obtained by adding a diagonal to the pentagon.

![Figure 3: An (almost) trivial orthogonal representation](image-url)
Example 7. The previous example can be generalized. In an orthogonal representation, we can label a set of nodes with the same nonzero vector if and only if these nodes form a clique. Let $k = \chi(\overline{G})$ (the chromatic number of the complement of $G$), then there is a family $\{B_1, \ldots, B_k\}$ of disjoint complete subgraphs covering all the nodes. Mapping every node of $B_i$ to $e_i$ ($i = 1, \ldots, k$) is an orthonormal representation.

Example 8. In the introduction we have seen an orthogonal representation with a more interesting geometric content. The previous example gives an orthogonal representation of $C_5$ in 3-space (Figure 4, left). The “umbrella” representation defined in the introduction gives another orthogonal representation of the graph $C_5$ in 3-space.

![Figure 4: The graph $C_5$ and two orthonormal representations of it.](image)

2.2 Definition of the theta-function

Let us start with the definition:

$$\vartheta(G) = \min_{u, c} \max_{i \in V} \frac{1}{(c^T u_i)^2}$$

(3)

where the minimum is taken over all orthonormal representations $(u_i : i \in V)$ of $G$ and all unit vectors $c$. We call $c$ the “handle” of the representation (for the origin of the name, see Example 2). If $\varphi$ is the half-angle of the smallest rotational cone with axis in the direction of $c$, containing all vectors in an orthonormal representation, then clearly

$$\vartheta(G) = \frac{1}{(\cos \varphi)^2}.$$  

(4)

Of course, we could fix $c$ to be (say) the standard basis vector $e_1$, but this is not always convenient. We may always assume that the dimension $d$ of the ambient space is at most $n$. We could require that $c^T u_i \geq 0$ for all nodes, since we can replace any $u_i$ by its negative. With just a little more complicated argument, we may even require that

$$c^T u_i = \frac{1}{\sqrt{\vartheta(G)}}$$

(5)
for every node $i$ without changing the optimum value in $|3|$. Indeed, let $\alpha_i = 1/\sqrt{\vartheta(G)c^T u_i}$, then $0 \leq \alpha_i \leq 1$, and so we can replace $u_i$ by the $(d + n)$-dimensional vector

$$u'_i = \left( \frac{\alpha_i u_i}{\sqrt{1 - \alpha_i^2}} \right) \sqrt{1 - \alpha_i^2} e_i.$$  

Extending $c$ with $n$ zeroes to get a unit vector $c'$, we get an orthonormal representation in which $(c')^T u'_i = 1/\sqrt{\vartheta(G)}$ for all $i$.

The following rather easy bounds will be very important.

**Theorem 9.** For every graph $G$,

$$\alpha(G) \leq \vartheta(G) \leq \chi(G).$$

**Proof.** First, let $S \subseteq V$ be a maximum stable set of nodes in $G$. Then in every orthonormal representation $(u_i)$, the vectors $\{u_i : i \in S\}$ are mutually orthogonal unit vectors. Hence

$$1 = \|c\|^2 \geq \sum_{i \in S} (c^T u_i)^2 \geq |S| \min_i (c^T u_i)^2,$$

and so

$$\max_{i \in V} \frac{1}{(c^T u_i)^2} \geq |S| = \alpha(G).$$

This implies the first inequality. The second follows from the orthonormal representation constructed in Example 7 using $c = (1/\sqrt{m}) \mathbb{I}$ as the handle.

Our examples in Section ?? give further upper bounds on $\vartheta$. Example 5 leads to the following construction: Assuming that there are no isolated nodes, we assign to each node $i \in V$ the vector $u_i \in \mathbb{R}^E$ defined by

$$(u_i)_e = \begin{cases} \frac{1}{\sqrt{\deg(i)}} & \text{if } e \text{ is incident with } i, \\ 0, & \text{otherwise}, \end{cases}$$

and define the handle $c = (1/\sqrt{m}) \mathbb{I}$. Then $c^T u_i = \sqrt{\deg(i)/m}$, and so we get the bound

$$\vartheta(G) \leq \max_i \frac{m}{\deg(i)} = \frac{m}{d_{\text{min}}}.$$  

(The upper bound $m/d_{\text{min}}$ for the stability number $\alpha(G)$ is easy to prove by counting edges.) From Example 8 we get by elementary trigonometry that

$$\vartheta(C_5) \leq \sqrt{5}.\tag{7}$$

Soon we will see that equality holds here.

It is clear that if $G'$ is an induced subgraph of $G$, then $\vartheta(G') \leq \vartheta(G)$ (an optimal orthonormal representation of $G$, restricted to $V(G')$, is an orthonormal representation of $G'$). It is also clear that if $G'$ is a spanning subgraph of $G$ (i.e., $V(G') = V(G)$ and $E(G') \subseteq E$), then $\vartheta(G') \geq \vartheta(G)$ (every orthonormal representation of $G'$ is an orthonormal representation of $G$).
3 Duality for theta

3.1 Alternative definitions

The graph parameter $\vartheta$ has many equivalent definitions. We are going to state some, which lead to an important dual formulation of this quantity.

**Vector chromatic number.** The following geometric definition was proposed by Karger, Motwani and Sudan. In terms of the complementary graph, this value is called the “vector chromatic number”. As a motivation for this name, consider a $t$-colorable graph $(t \geq 2)$, and let us color its nodes by $t$ unit vectors $f_1, \ldots, f_t \in \mathbb{R}^{t-1}$, pointing to the vertices of a regular simplex centered at the origin. We get a vector labeling $w$ (not an orthogonal representation of $G$, but closely related). It is not hard to compute that $f_k^T f_l = -1/(t-1)$ for $k \neq l$, and so $w_i^T w_j = -1/(t-1)$ for $ij \in E$.

Now let us forget about the condition that $w_i$ must be one of the vectors $f_k$: define a (strict) vector $t$-coloring $(t > 1)$ of the graph $G$ as a vector labeling $i \mapsto w_i \in \mathbb{R}^n$ such that $|w_i| = 1$ for all $i \in V$, and $w_i^T w_j = -1/(t-1)$ for all $ij \in E$. (The dimension $n$ in the definition above is just chosen to be large enough; allowing a higher dimension would not make any difference.)

If $G$ has at least one edge, then the last condition implies that $t \geq 2$. The smallest $t$ for which the graph $G$ has a vector $t$-coloring is called its (strict) vector chromatic number, and is denoted by $\chi_{vec}$. If $E = \emptyset$, so $G$ is edgeless and $\chi(G) = 1$, then we define $\chi_{vec} = 1$. For $t = 2$, the endpoints of any edge must be labeled by antipodal unit vectors. It follows that every connected component of $G$ is labeled by two antipodal vectors only, and so $G$ is bipartite. Conversely, every bipartite graph has a strict vector 2-coloring.

It is clear from the construction above that $\chi_{vec}(G) \leq \chi(G)$ for every graph $G$. Equality does not hold in general. Vector-labeling the nodes of a pentagon by the vertices of a regular pentagon inscribed in the unit circle, so that the edges are mapped onto the diagonals, we see that $\chi_{vec}(C_5) \leq \sqrt{5} < \chi(C_5) = 3$.

**Semidefinite matrices.** Next, we give a couple of formulas for $\vartheta$ in terms of semidefinite matrices. Let

$$\vartheta_{\text{diag}} = \min \{ 1 + \max_{i \in V} Y_{ii} : Y \in \mathbb{R}^{V \times V}, Y \succeq 0, Y_{ij} = -1 (ij \in E) \}$$

(8)

and

$$\vartheta_{\text{sum}} = \max \left\{ \sum_{i,j \in V} Z_{ij} : Z \in \mathbb{R}^{V \times V}, Z \succeq 0, Z_{ij} = 0 (ij \in E), \text{tr}(Z) = 1 \right\}.$$

(9)

It will turn out that these two values are equal. This equality is in fact a special case of the Duality Theorem of semidefinite optimization. It is not hard to check that (8) and (9) are dual semidefinite programs, and the first one has a strictly feasible solution. So the Duality Theorem ?? of semidefinite
programming applies, and asserts that the two programs have the same objective value. However, we are going to include a proof, to make our treatment self-contained.

**Dual orthogonal representation.** We use orthonormal representations of the complementary graph to define

\[ \vartheta_{\text{dual}} = \max \sum_{i \in V} (d^T v_i)^2, \]  

where the maximum extends over all orthonormal representations \((v_i : i \in V)\) of the complementary graph \(G\) and all unit vectors (handles) \(d\).

### 3.2 The main duality result

The main theorem of this section asserts that all these definitions lead to the same value.

**Theorem 10.** For every graph \(G\),

\[ \vartheta(G) = \chi_{\text{vec}}(G) = \vartheta_{\text{diag}}(G) = \vartheta_{\text{sum}}(G) = \vartheta_{\text{dual}}(G). \]

**Proof.** We prove the circle of inequalities

\[ \vartheta(G) \leq \chi_{\text{vec}}(G) \leq \vartheta_{\text{diag}}(G) \leq \vartheta_{\text{sum}}(G) \leq \vartheta_{\text{dual}}(G) \leq \vartheta(G). \]  

To prove the first inequality, let \( t = \chi_{\text{vec}}(G) \), and let \((w_i : i \in V)\) be an optimal vector \(t\)-coloring. Let \(c\) be a vector orthogonal to all the \(w_i\) (we increase the dimension of the space if necessary). Let \(u_i = \frac{1}{\sqrt{t}} c + \sqrt{t - 1} w_i\).

Then \(|u_i| = 1\) and \(u_i^T u_j = 0\) for \(ij \in \overline{E}\), so \((u_i)\) is an orthonormal representation of \(G\). Furthermore, with handle \(c\) we have \(c^T u_i = 1/\sqrt{t}\), which implies that \(\vartheta(G) \leq \chi_{\text{vec}}(G)\).

Second, let \(Y\) be an optimal solution of \([\text{8}].\) We may assume that all diagonal entries \(Y_{ii}\) are the same number \(t\), since we can replace all of them by the largest without violating the other constraints. The matrix \(1/(t-1)Y\) is positive semidefinite, and so it can be written as \(\text{Gram}(w_i : i \in V)\) with appropriate vectors \((w_i \in \mathbb{R}^n)\). These vectors form a strict vector \(t\)-coloring. Since \(\chi_{\text{vec}}(G)\) is the smallest \(t\) for which this exists, this proves that \(\chi_{\text{vec}}(G) \leq \vartheta_{\text{diag}}\).

The main step in the proof is to show that \(\vartheta_{\text{diag}} \leq \vartheta_{\text{sum}}\). Fix any \(t > \vartheta_{\text{sum}}\); it is easy to see that \(\vartheta_{\text{sum}} \geq 1\) and hence \(t > 1\). Let \(L\) denote the linear space of symmetric \(V \times V\) matrices satisfying \(Z_{ij} = 0\) \((ij \in E)\) and \((tI - J) \cdot Z = 0\), and let \(P\) denote the cone of positive semidefinite \(V \times V\) matrices.

We claim that \(P \cap L = \{0\}\). Suppose, to the contrary, that there is a symmetric matrix \(Z \neq 0\) such that \(Z \in P \cap L\). Clearly \(\text{tr}(Z) > 0\), and so, by
scaling, we may assume that $\text{tr}(Z) = 1$. Then $Z$ satisfies the conditions in the definition of $\vartheta_{\text{sum}}$, and so $\vartheta_{\text{sum}} \geq J \cdot Z = tI \cdot Z = t$, contradicting the choice of $t$.

So the linear space $L$ touches the convex cone $P$ at its apex only, and hence there is a hyperplane $H$ such that $L \subseteq H$ and $H \cap P = \{0\}$. Let $Y \cdot X = 0$ be the equation of $H$ (where $Y \neq 0$ is a symmetric $V \times V$ matrix); we may assume that $Y \cdot X \geq 0$ for all $X \in P$. This means that $Y$ is in the dual cone of $P$. This dual cone is $P$ itself, which means that $Y \succeq 0$. Furthermore, $L \subseteq H$ means that the equation of $H$ is a linear combination of the equations defining $L$, i.e., there are real numbers $a_{ij}$ ($ij \in E$) and $b$ such that

$$ Y = \sum_{ij \in E} a_{ij} E_{ij} + b(tI - J). $$

Considering a positive diagonal entry of $Y$, we see that $b > 0$, and since we are free to scale $Y$ by positive scalars, we may assume that $b = 1$. But this means that $Y$ satisfies the conditions in the definition of $\vartheta_{\text{diag}}$, and so $\vartheta_{\text{diag}} \leq 1 + \max_i Y_{ii} = t$. Since this holds for every $t > \vartheta_{\text{sum}}$, this implies that $\vartheta_{\text{diag}} \leq \vartheta_{\text{sum}}$.

To prove the fourth inequality in (11), let $Z$ be an optimum solution of the program (9) with objective function value $\vartheta_{\text{sum}}$. We can write $Z$ as

$$ Z = \text{Gram}(z_i : i \in V) $$

where $z_i \in \mathbb{R}^n$. Let us rescale the vectors $z_i$ to get the unit vectors $v_i = z_i / \|z_i\|$ (if $z_i = 0$, then we take a unit vector orthogonal to everything else as $v_i$). Define $d = (\sum_i z_i)$. By the properties of $Z$, the vectors $v_i$ form an orthonormal representation of $G$, and hence

$$ \vartheta_{\text{dual}} \geq \sum_i (d^T v_i)^2. $$

To estimate the right side, we use the equations

$$ \sum_i |z_i|^2 = \sum_i z_i^T z_i = \text{tr}(Z) = 1, \quad \sum_i z_i^2 = \sum_{i,j} z_i^T z_j = \sum_{i,j} Z_{ij} = \vartheta_{\text{sum}}, $$

and the Cauchy–Schwarz Inequality:

$$ \sum_i (d^T v_i)^2 = \left( \sum_i |z_i|^2 \right) \left( \sum_i (d^T v_i)^2 \right) \geq \left( \sum_i |z_i| (d^T v_i) \right)^2 $$

$$ = \left( \sum_i d^T z_i \right)^2 = \left( d^T \sum_i z_i \right)^2 = \left( \sum_i z_i \right)^2 = \vartheta_{\text{sum}}. $$

This proves that $\vartheta_{\text{dual}} \geq \vartheta_{\text{sum}}$.

Finally, to prove the last inequality in (11), it suffices to prove that if $(u_i : i \in V)$ is an orthonormal representation of $G$ in $\mathbb{R}^n$ with handle $c$, and $(v_i : i \in V)$ is an orthonormal representation of $G$ in $\mathbb{R}^m$ with handle $d$, then

$$ \sum_{i \in V} (d^T v_i)^2 \leq \max_{i \in V} \frac{1}{(c^T u_i)^2}. \quad (12) $$
The tensor product vectors $u_i \circ v_i$ ($i \in V$) are mutually orthogonal unit vectors. Indeed, $(u_i \circ v_i)^T(u_j \circ v_j) = (u_i^T u_j)(v_i^T v_j) = 0$, since either $u_i$ is orthogonal to $u_j$ or $v_i$ is orthogonal to $v_j$. Hence

$$\sum_i (c^T u_i)^2(d^T v_i)^2 = \sum_i ((c \circ d)^T(u_i \circ v_i))^2 \leq 1. \quad (13)$$

On the other hand,

$$\sum_i (c^T u_i)^2(d^T v_i)^2 \geq \min_i (c^T u_i)^2 \sum_i (d^T v_i)^2,$$

which implies that

$$\sum_i (d^T v_i)^2 \leq \frac{1}{\min_i (c^T u_i)^2} = \max_i \frac{1}{(c^T u_i)^2}.$$

This proves (12) and completes the proof of Theorem 10.

We can state the theorem more explicitly as the following sequence of formulas.

$$\vartheta(G) = \min \left\{ \max_{i \in V} \frac{1}{(c^T u_i)^2} : u \text{ ONR of } G, |c| = 1 \right\} \quad (14)$$

$$= \min \left\{ t > 1 : |w_i| = 1, w_i^T w_j = -\frac{1}{t-1} \ (ij \in E) \right\} \quad (15)$$

$$= \min \left\{ 1 + \max_{i \in V} Y_{ii} : Y \succeq 0, Y_{ij} = -1 \ (ij \in \overline{E}) \right\} \quad (16)$$

$$= \max \left\{ \sum_{i,j \in V} Z_{ij} : Z \succeq 0, Z_{ij} = 0 \ (ij \in E), \ \text{tr}(Z) = 1 \right\} \quad (17)$$

$$= \max \left\{ \sum_{i \in V} (d^T v_i)^2 : v \text{ ONR of } \overline{G}, |d| = 1 \right\}. \quad (18)$$

From this form it is clear (and we have seen this in the proof as well) that the powerful step in this sequence of formulas is the equality (16)=(17), where an expression as a minimum switches to an expression as a maximum. Note that before this equality we have conditions on the edges of $G$, which then get replaced by conditions on the edges of $\overline{G}$ for the last two rows.

### 3.3 Consequences of duality

From the fact that equality holds in (11), it follows that equality holds in all of the arguments above. Let us formulate some consequences. From the fact that equality must hold in (13), and from the derivation of this inequality, we see that for an optimal orthonormal representation $(u, c)$ and an optimal dual orthonormal representation $(v, d)$, the unit vector $c \circ d$ must be a linear
combination of the mutually orthogonal unit vectors $\mathbf{u}_i \circ \mathbf{v}_i$. The coefficients are easy to figure out, and we get

$$\mathbf{c} \circ \mathbf{d} = \sum_i (\mathbf{c}^T \mathbf{u}_i)(\mathbf{d}^T \mathbf{v}_i)(\mathbf{u}_i \circ \mathbf{v}_i),$$

(19)

or in a matrix form:

$$\mathbf{c} \mathbf{d}^T = \sum_i (\mathbf{c}^T \mathbf{u}_i)(\mathbf{d}^T \mathbf{v}_i)\mathbf{u}_i \mathbf{v}_i^T.$$  

(20)

This equation gives (upon multiplying by $\mathbf{d}$ from the right and by $\mathbf{c}^T$ from left, respectively) two equations expressing the handles as linear combinations of the (primal and dual) orthonormal representations:

$$\mathbf{c} = \sum_i (\mathbf{c}^T \mathbf{u}_i)(\mathbf{d}^T \mathbf{v}_i)^2 \mathbf{u}_i, \quad \mathbf{d} = \sum_i (\mathbf{c}^T \mathbf{u}_i)^2 (\mathbf{d}^T \mathbf{v}_i) \mathbf{v}_i.$$  

(21)

Using an optimal orthonormal representation with $\mathbf{c}^T \mathbf{u}_i = 1/\sqrt{\vartheta}$, we get the simpler formulas

$$\mathbf{c} = \frac{1}{\sqrt{\vartheta(G)}} \sum_{i \in V} (\mathbf{d}^T \mathbf{v}_i)^2 \mathbf{u}_i, \quad \mathbf{d} = \frac{1}{\vartheta(G)} \sum_{i \in V} (\mathbf{d}^T \mathbf{v}_i) \mathbf{v}_i.$$  

(22)

There is a simple relationship between the theta value of a graph and of its complement.

**Proposition 11.** For every graph $G$, we have $\vartheta(G) \vartheta(G^c) \geq n$.

**Proof.** Let $(\mathbf{u}, \mathbf{c})$ be an optimal orthogonal representation of $G$. Then applying (13) to the complementary graph, we get

$$\vartheta(G^c) \geq \sum_i (\mathbf{c}^T \mathbf{u}_i)^2 \geq n \min_i (\mathbf{c}^T \mathbf{u}_i)^2 = \frac{n}{\vartheta(G)}.$$  

Equality does not hold in (11) in general, but it does when $G$ has a node-transitive automorphism group. To prove this, we need an important fact about the symmetries of orthogonal representations. We say that an orthonormal representation $(\mathbf{u}_i, \mathbf{c})$ in $\mathbb{R}^d$ of a graph $G$ is automorphism invariant, if every automorphism $\gamma \in \text{Aut}(G)$ can be lifted to an orthogonal transformation $O_\gamma$ of $\mathbb{R}^d$ such that $O_\gamma \mathbf{c} = \mathbf{c}$ and $\mathbf{u}_{\gamma(i)} = O_\gamma \mathbf{u}_i$ for every node $i$. An optimal orthonormal representation (say, in the sense of (14)) is not necessarily invariant under automorphisms, but there is always one that is (see Figure 5). The representation on the left is also optimal with respect to minimizing the dimension, and it is not hard to see that $C_4$ has no automorphism invariant orthonormal representation in $\mathbb{R}^2$. So minimizing the dimension and minimizing the cone behave differently from this point of view.

**Theorem 12.** Every graph $G$ has an optimal orthonormal representation and an optimal dual orthonormal representation that are both automorphism invariant.
Figure 5: An optimal orthonormal representation of $C_4$ that is not invariant under its automorphisms, and one that is.

Proof. We give the proof for the dual orthonormal representation. The optimum solutions of the semidefinite program in (8) form a bounded convex set, which is invariant under the transformations $Z \mapsto P^T \alpha Z P \alpha$, where $P \alpha$ is the permutation matrix defined by an automorphism $\alpha$ of $G$. If $Z$ is an optimizer in (17), then so is $P^T \alpha Z P \alpha$ for every automorphism $\alpha \in \text{Aut}(G)$, and hence also

$$\hat{Z} = \frac{1}{|\text{Aut}(G)|} \sum_{\alpha \in \text{Aut}(G)} P^T \alpha Z P \alpha.$$

This matrix satisfies $P^T \alpha Z P \alpha = Z$ for all automorphisms $\alpha$.

The construction of an orthonormal representation of $G$ in the proof of $\vartheta_{\text{diag}} \leq \vartheta_{\text{sum}}$ in Theorem 10 can be done in a canonical way: we choose the columns of $Z^{1/2}$ as the vectors $z_i$, and use them to construct the dual orthonormal representation with $v_i = z_i^0$ and $d = (\sum_i z_i)^0$. The optimal dual orthonormal representation constructed this way will be invariant under the automorphism group of $G$.

**Corollary 13.** If $G$ has a node-transitive automorphism group, then

$$\vartheta(G) \vartheta(\overline{G}) = n.$$

Proof. It follows from Theorem 12 that $\overline{G}$ has an orthonormal representation $(v_i, d)$ in $\mathbb{R}^n$ such that $\sum_i(d^T v_i)^2 = \vartheta(G)$, and $d^T v_i$ is the same for every $i$. So $(d^T v_i)^2 = \vartheta(G)/n$ for all nodes $i$, and hence

$$\vartheta(\overline{G}) \leq \max_i \frac{1}{(d^T v_i)^2} = \frac{n}{\vartheta(G)}.$$

Since we already know the reverse inequality (Proposition 11), this proves the Corollary.

**Corollary 14.** If $G$ is a self-complementary graph with a node-transitive automorphism group, then $\vartheta(G) = \sqrt{n}$. In particular, $\vartheta(C_5) = \sqrt{5}$.  

12
A further important feature of the theta-function is its nice behavior with respect to graph product; we will see that this is what underlies its applications in information theory.

There are many different ways of multiplying two simple graphs $G$ and $H$, of which we need one in this chapter. The strong product $G \boxtimes H$ is defined on the underlying set $V(G) \times V(H)$. Two nodes $(u_1, v_1)$ and $(u_2, v_2)$ are adjacent if and only if either $ij \in E$ and $uv \in E(H)$, or $ij \in E$ and $u = v$, or $i = j$ and $uv \in E(H)$. It is easy to see that this multiplication is associative and commutative (up to isomorphism). The product of two complete graphs is a complete graph.

**Theorem 15.** For any two graphs $G$ and $H$, we have $\vartheta(G \boxtimes H) = \vartheta(G) \vartheta(H)$.

**Proof.** Let $(u_i : i \in V)$ be an optimal orthogonal representation of $G$ with handle $c$ ($u_i, c \in \mathbb{R}^n$), and let $(v_j : j \in V(H))$ be an optimal orthogonal representation of $H$ with handle $d$ ($v_j, d \in \mathbb{R}^m$). It is easy to check, using (2), that the vectors $u_i \circ v_j$ for $((i, j) \in V(G) \times V(H))$ form an orthogonal representation of $G \boxtimes H$. Furthermore, taking $c \circ d$ as its handle, we have by (2) again that

$$\left((c \circ d)^T(u_i \circ v_j)\right)^2 = (c^T u_i)^2 (d^T v_j)^2 \geq \frac{1}{\vartheta(G)} \cdot \frac{1}{\vartheta(H)},$$

and hence

$$\vartheta(G \boxtimes H) \leq \max_{i,j} \left(\frac{1}{(c \circ d)^T(u_i \circ v_j)}\right)^2 \leq \vartheta(G) \vartheta(H).$$

To prove that equality holds, we use the duality established in Section 3. Let $(v_i, d)$ be an orthonormal representation of $G$ which is optimal in the sense that $\sum_i (d^T v_i)^2 = \vartheta(G)$, and let $(w_j, e)$ be an orthonormal representation of $H$ such that $\sum_j (e^T w_j)^2 = \vartheta(H)$. It is easy to check that the vectors $v_i \circ w_j$ form an orthonormal representation of $G \boxtimes H$, and so using handle $d \circ e$ we get

$$\vartheta(G \boxtimes H) \geq \sum_{i,j} \left((d \circ e)^T(v_i \circ w_j)\right)^2 = \sum_{i,j} (d^T v_i)^2 (e^T w_j)^2 = \vartheta(G) \vartheta(H).$$

We already know the reverse inequality, which completes the proof. □

### 3.4 Eigenvalues and theta

To motivate the identities and inequalities to be proved in this section, let us survey some of the classical results which use spectral properties of graphs, or more generally linear algebra techniques, to bound quantities like the stability number $\alpha = \alpha(G)$, the clique number $\omega = \omega(G)$, or the chromatic number $\chi = \chi(G)$ in terms of the eigenvalues of the adjacency matrix $A = A_G$. It turns out that several of these results could be used to define $\vartheta(G)$, if generalized appropriately.
Let us start with an almost trivial inequality:

\[ \omega \leq 1 + \lambda_{\text{max}}(A). \]  

(23)

In terms of the complementary graph,

\[ \alpha \leq 1 + \lambda_{\text{max}}(A^\top). \]  

(24)

Indeed, the matrix \( A + I \) contains an \( \omega \times \omega \) submatrix \( J_\omega \) of 1’s, so \( 1 + \lambda_{\text{max}}(A) = \lambda_{\text{max}}(A + I) \geq \lambda_{\text{max}}(J_\omega) = \omega \). Note that in this argument, only the matrix entries in adjacent and diagonal positions are used. We could consider any symmetric matrix \( A' \) obtained from \( A \) by substituting arbitrary real numbers for the remaining entries (which are originally zeroes), to minimize the bound \( \lambda_{\text{max}}(A') + 1 \). (We keep \( A' \) symmetric, to have real eigenvalues.) What is the best bound on \( \omega \) we can obtain this way?

The following lower bound on the chromatic number of a graph is more difficult to prove:

\[ \chi \geq 1 - \frac{\lambda_{\text{max}}(A)}{\lambda_{\text{min}}(A)} \]  

(25)

(note that \( \lambda_{\text{min}}(A) < 0 \) if \( G \) has at least one edge, which we may assume). We will not go through the proof; but if you do, you realize that it uses only the 0’s in the adjacency matrix, so we can play with the 1’s to get the sharpest possible lower bound. What is the best bound on \( \chi \) we can obtain this way?

Hoffman (unpublished) proved the following upper bound on \( \alpha \), somewhat analogous to the bound (25): If \( G \) is a \( d \)-regular graph, then

\[ \alpha(G) \leq \frac{-n\lambda_{\text{min}}(A)}{d - \lambda_{\text{min}}(A)} = \frac{-n\lambda_{\text{min}}(A)}{\lambda_{\text{max}}(A) - \lambda_{\text{min}}(A)}. \]  

(26)

Looking at the proof, one realizes that we use here where the 0’s of \( A \) are, and also the fact that all row sums are the same; but not the actual values of the entries corresponding to edges. What is the best bound we can obtain by playing with the entries in adjacent positions?

Perhaps it is not surprising that the answer to the first two questions posed above is \( \vartheta(G) \), and the answer to the third is closely related. This will follow from the next identities.

**Proposition 16.** For every graph \( G \),

\[ \vartheta(G) = \min_U \lambda_{\text{max}}(U), \]  

where \( U \) ranges over all \( V \times V \)-matrices with \( U_{ij} = 1 \) for \( ij \in E \) and also for \( i = j \). Furthermore,

\[ \vartheta(G) = \max_W \lambda_{\text{max}}(W), \]  

where \( W \) ranges over all positive semidefinite \( V \times V \)-matrices with \( W_{ij} = 0 \) for \( ij \in E \) and \( W_{ii} = 1 \) for \( i \in V \). Furthermore,

\[ \vartheta(G) = 1 + \max_T \frac{\lambda_{\text{max}}(T)}{|\lambda_{\text{min}}(T)|}, \]  

(27)
where $T$ ranges over all symmetric nonzero $V \times V$-matrices with $T_{ij} = 0$ for $ij \in E$ and also for $i = j$.

Note that (28) can be written as

$$\vartheta(G) = \max_v \lambda_{\text{max}}(\text{Gram}(v)), \quad (30)$$

where $v$ ranges over dual orthonormal representations of $G$ (see Exercise 13 for the connection of this formula with the definition 18 of $\vartheta$).

**Proof.** Let $Y$ be a minimizer matrix in (16), let $D$ be the diagonal matrix obtained from $Y$ by changing all off-diagonal entries to 0, and define $U = I + D - Y$. Then $U$ satisfies the conditions in the Proposition, and (using that $Y \succeq 0$)

$$\lambda_{\text{max}}(U) \leq \lambda_{\text{max}}(I + D) = 1 + \max_{i \in V} Y_{ii} = \vartheta(G).$$

The reverse inequality follows similarly, by starting with a minimizer in (27), and considering $Y = \lambda_{\text{max}}(U)I - U$.

To prove (28), let $Z$ be a minimizer in (17), and define a $V \times V$ matrix $B$ by

$$W_{ij} = \frac{Z_{ij}}{\sqrt{Z_{ii}Z_{jj}}}.$$

Then $W$ satisfies the conditions in (28). Define a vector $x \in \mathbb{R}^V$ by $x_i = \sqrt{Z_{ii}}$, then $x$ is a unit vector and

$$\lambda_{\text{max}}(B) \geq x^T W x = \sum_{i,j} Z_{ij} = \vartheta(G).$$

The reverse inequality follows similarly, by starting with an optimizer $W$ in (28), and scaling its rows and columns by the entries of an eigenvector belonging to $\lambda_{\text{max}}(W)$.

Finally, to prove (29), consider an optimizer $T$ in it, then $W = I - \frac{1}{\lambda_{\text{min}}(T)} T$ is positive semidefinite, has 0’s in adjacent positions and 1’s on the diagonal. Hence by (28), we have

$$\vartheta(G) \geq \lambda_{\text{max}}(W) = 1 - \frac{\lambda_{\text{max}}(T)}{\lambda_{\text{min}}(T)}.$$

The reverse inequality follows by a similar argument. \hfill \Box

Proposition 16 can be combined with different known estimates for the largest eigenvalue of a matrix. As an example, using that the largest eigenvalue of a matrix is bounded above by the largest $\ell_1$-norm of its rows, we get that for every optimal dual orthogonal representation $v$,

$$\vartheta(G) \leq \max_i \sum_j |v_i^T v_j| \quad (31)$$

Finally, we show how a strengthening of Hoffman’s bound (26) can be derived from Proposition 16.
Proposition 17. Let $G$ be a $d$-regular graph. Then for every symmetric nonzero $V \times V$-matrix $M$ such that $M_{ij} = 0$ for $ij \in E$ and also for $i = j$, and $M$ has equal row-sums, we have

$$\vartheta(G) \leq \frac{-n\lambda_{\min}(M)}{\lambda_{\max}(M) - \lambda_{\min}(M)}.$$  

If the automorphism group of $G$ is transitive on the nodes, then there is such a matrix $M$ attaining equality. If the automorphism group is transitive on the edges, then equality holds for $M = A_G$.

Proof. Let $M$ have eigenvalues $\lambda_1 = d \geq \lambda_2 \geq \cdots \geq \lambda_n$. The matrix $J - tM$ satisfies the conditions in (27) for every value of $t$. Using the condition that all row-sums of $M$ are the same, we see that $1$ is a common eigenvector of $J$ and $M$, and it follows that all eigenvectors of $M$ are eigenvectors of $J$ as well. Hence the eigenvalues of $J - tM$ are $n - td, -t\lambda_2, \ldots, -t\lambda_n$. From the fact that the trace of $M$ is zero, and hence $\lambda_n < 0$, it follows that the largest eigenvalue of $J - tM$ is either $n - td$ or $-t\lambda_n$, and we get the best bound if we choose $t$ so that these two are equal: $t = n/(d - \lambda_n)$, giving the bound in the Proposition.

We can see just as in the proof of Theorem 12 that there is an optimizing matrix $U$ in (27) that is invariant under the automorphisms. So if $G$ has a node-transitive automorphism group, then the row-sums of this matrix are equal, and the same holds for $M = J - U$. This matrix $M$ satisfies the conditions in the Proposition, and attains equality.

If $G$ has an edge-transitive automorphism group, then all nonzero entries of $M$ are the same, and hence $M = tA_G$ for some $t \neq 0$. The value of $t$ cancels from the formula, so $M = A_G$ also provides equality.

4 Computing the theta-function

Perhaps the most important consequence of the formulas proved in Section 3 is that the value of $\vartheta(G)$ is polynomial time computable. More precisely,

Theorem 18. There is an algorithm that computes, for every graph $G$ and every $\varepsilon > 0$, a real number $t$ such that

$$|\vartheta(G) - t| < \varepsilon.$$  

The running time of the algorithm is polynomial in $n$ and $\log(1/\varepsilon)$.

Algorithms proving this theorem can be based on almost any of our formulas for $\vartheta$. The simplest is to refer to Theorem 10 giving a formulation of $\vartheta(G)$ as the optimum of a semidefinite program $\vartheta$, and the polynomial time solvability of semidefinite programs.

The significance of this fact is underlined if we combine it with Theorem 9. The two important graph parameters $\alpha(G)$ and $\chi(G)$ are both NP-hard, but they have a polynomial time computable quantity sandwiched between them.
Computability of the theta-function in an approximate sense, with an arbitrary small error, is an important fact, but not always satisfactory. Often we want explicit algebraic expressions, or at least explicit bounds. For example, it is important to know that \( \vartheta(C_5) = \sqrt{5} \), not just that it is between 2.23606 and 2.23607. In the rest of this section we compute the theta-function of several classes of graphs, to illustrate the use of the formulas in the previous sections.

**Example 19.** [Cycles] Even cycles are trivial: If \( n \) is even, then \( \alpha(C_n) = \vartheta(C_n) = \chi(C_n) = n/2 \) and \( \alpha(C_n) = \vartheta(C_n) = \chi(C_n) = 2 \). To derive the theta-function on odd cycles, we can use Proposition 17: The eigenvalues of \( C_n \) are \( 2 \cos(2k\pi/n) (k = 0, 1, \ldots, n - 1) \), of which \( k = 0 \) gives the largest (which is 2) and \( k = (n - 1)/2 \) gives the smallest one (which is \( 2 \cos((n - 1)\pi/n) = -2 \cos(\pi/n) \)). Hence

\[
\vartheta(C_n) = \frac{n \cos(\pi/n)}{1 + \cos(\pi/n)}.
\]

Since \( C_n \) has a node-transitive automorphism group, this implies that

\[
\vartheta(C_n) = 1 + \frac{1}{\cos(\pi/n)}.
\]

In particular, \( \vartheta(C_n) \to 2 \) as \( n \to \infty \).

**Example 20.** [Kneser graphs] The Kneser graph \( K_k^n \) is defined on node set \( \binom{[n]}{k} \), by connecting two \( k \)-sets if and only if they are disjoint \((1 \leq k \leq n)\). Let us assume that \( n \geq 2k \) to exclude the trivial case of a graph with no edges. The set of \( k \)-sets containing any fixed element of \([n]\) is stable, hence \( \alpha(K_k^n) \geq \binom{n-1}{k-1} \). The Erdős–Ko–Rado Theorem asserts that this is the exact value; this fact will follow from our considerations below.

To compute the theta-function of this graph, we use the eigenvalues of its adjacency matrix. These are well known from coding theory:

\[
\lambda_t = (-1)^t \binom{n - k - t}{k - t}, \quad (t = 0, 1, \ldots, k).
\]

The multiplicity of eigenvalue \( \lambda_t \) is \( \binom{n}{t} - \binom{n}{t-1} \) (but this will not be important here). The largest eigenvalue in \( \binom{n-k}{k} \) (the degree of each node), while the smallest is the next one, \( -\binom{n-k-1}{k-1} \).

We apply the formula in Proposition 17 and get

\[
\vartheta(K_k^n) \leq \frac{n \binom{n-k-1}{k-1}}{\binom{n-k}{k} - \binom{n-k-1}{k-1}} = \binom{n-1}{k-1}.
\]

Comparing this upper bound with the lower bound on \( \alpha(K_k^n) \), we see that they are equal, and so

\[
\vartheta(K_k^n) = \alpha(K_k^n) = \binom{n-1}{k-1}.
\]
In particular, the Petersen graph $K^{5}_{2}$ has $\vartheta(K^{5}_{2}) = 4$.

Since $K^{n}_{k}$ has a node-transitive automorphism group, it follows that

$$\vartheta(K^{n}_{k}) = \left(\frac{n}{k}\right) / \left(\frac{n-1}{k-1}\right) = \frac{n}{k}.$$  

This is quite close to $\alpha(K^{n}_{k}) = \lfloor n/k \rfloor$, but can be arbitrarily far from the chromatic number of $K^{n}_{k}$, which is known to be $n - 2k + 2$. ♦

**Example 21.** [Paley graphs] The Paley graph $\text{Pal}_p$ is defined for a prime $p \equiv 1 \pmod{4}$. We take the $\{0, 1, \ldots, p - 1\}$ as nodes, and connect two of them if their difference is a quadratic residue. It is clear that these graphs have a node-transitive automorphism group, and it is easy to see that they are self-complementary. So Corollary 14 applies, and gives that $\vartheta(\text{Pal}_p) = \sqrt{p}$. To determine the stability number of Paley graphs is an unsolved number-theoretic problem; it is conjectured that $\alpha(\text{Pal}_p) = O((\log p)^2)$. ♦

**Example 22.** [Cycles with diagonals] For graphs with a node-transitive automorphism group, the existence of automorphism-invariant optima can be very useful. We illustrate this on the graph $W_n$ obtained of an even cycle $C_n$ with its longest diagonals added. The graph $W_6$ is just the Kuratowski graph $K_{3,3}$, $W_8$ is the Wagner graph from Example 20. In fact, we can restrict our attention to the case when $n = 4k$ is a multiple of 4, since otherwise $W_n$ is bipartite and $\vartheta(W_n) = \alpha(W_n) = n/2$. We can observe the easy bounds $2k - 1 \leq \vartheta(W_n) \leq 2k$, since $W_n$ has $2k - 1$ nonadjacent nodes and can be covered by $2k$ edges.

Let $V(W_n) = \{0, 1, \ldots, n - 1\}$, where the nodes are labeled in the order of the original cycle. There exists an optimizing matrix $U$ in (27) that is invariant under rotation, which means that it has only three different entries:

$$U_{ij} = \begin{cases} 1 + a, & \text{if } j - i \equiv \pm 1 \pmod{4k}, \\ 1 + b, & \text{if } j - i \equiv 2k \pmod{4k}, \\ 1, & \text{otherwise.} \end{cases}$$

It is easy to see that for every $(4k)$-th root of unity $\varepsilon$, the vector $(1, \varepsilon, \varepsilon^2, \ldots, \varepsilon^{n-1})$ is an eigenvector of $U$. (This is a complex vector, so if we want to stay in the real field, we have to consider its real and imaginary parts; but it is more convenient here to do computations with complex vectors.) The eigenvalue $\lambda_r$ corresponding to $\varepsilon = e^{2\pi ir/n}$ is easy to compute:

$$\lambda_0 = 4k + 2a + b,$$

and

$$\lambda_r = 1 + a\varepsilon + \varepsilon^2 + \cdots + b\varepsilon^{2k} + \varepsilon^{2k+1} + \cdots + a\varepsilon^{n-1} = a(\varepsilon + \bar{\varepsilon}) + b\varepsilon^{2k}$$

$$= 2a \cos \frac{r\pi}{2k} + b(-1)^r \quad (r > 0).$$
We want to minimize \( \max_r \lambda_r \). Each \( \lambda_r = \lambda_r(a, b) \) is a linear function of \( a \) and \( b \), so we can find \( \vartheta \) as the optimum of the linear program in 3 variables \( a, b \) and \( t \):

\[
\vartheta(W_n) = \min_{a, b} \max_r \lambda_r(a, b) = \begin{cases} 
\text{minimize} & t \\
\text{subject to} & \lambda_r(a, b) \leq t, \ r = 1, \ldots, n - 1.
\end{cases}
\]

There are many ways to do a back-of-the-envelope computation here; one gets that \( a = -k, b = -k + k \cos(\pi/k) \) is an optimal solution, giving

\[
\vartheta(W_n) = k + k \cos \frac{\pi}{k}.
\]

The main point in this example is to illustrate that for graphs with a node-transitive automorphism group, the value of the theta-function can be computed by a linear program, analogous to (35), where the number of unknowns is the number of orbits of the automorphism group on the edges. This may or may not lead to simple formula like in this case, but the computation is easy to perform, often even by hand.

Example 23. [Self-polar polytopes] A polytope \( P \subseteq \mathbb{R}^d \) is called \textit{self-polar}, if \( P^* = -P \). Note that this condition implies that for each vertex \( v \), the inequality \((-v)^T x \leq 1 \) defines a facet \( F_v \) of \( P \), and we obtain all facets this way. We call two vertices \( v \) and \( v' \) of \( P \) opposite, if \( v' \) lies on \( F_v \). In other words, \( v^T v' = -1 \), which shows that this is a symmetric relation. We call the polytope \textit{strongly self-polar}, if it is inscribed in a ball centered at the origin, in other words, there is an \( r > 0 \) such that \( |v| = r \) for all vertices \( v \). For two opposite vertices we have \( 1 = v^T u < |v||u| = r^2 \), and hence \( r > 1 \). It also follows that the distance of any facet from the origin is \( 1/r \), so the sphere with radius \( 1/r \) about the origin is contained in \( P \) and touches every facet.

In dimension 2, regular polygons with an odd number of vertices, with appropriate edge length, are strongly self-polar. It is known that for every dimension \( d \) and \( \varepsilon > 0 \) there exist strongly self-polar polytopes inscribed in a sphere with radius \( r < 1 + \varepsilon \).

Let \( P \) be a strongly self-polar polytope in \( \mathbb{R}^d \), and let \( G \) be the graph on \( V = V(P) \), in which two vertices are connected if and only if they are opposite. It is known that \( \chi(G) \geq d + 1 \).

We can estimate \( \vartheta(G) \) and \( \vartheta(G) \) as follows. Let us label each vertex \( v \) of \( P \) with the vector

\[
\mathbf{u}_v = \frac{1}{\sqrt{r^2 + 1}} \begin{pmatrix} v^T \\ 1 \end{pmatrix}.
\]

This is trivially a unit vector in \( \mathbb{R}^{d+1} \), and \( \mathbf{u}_v \perp \mathbf{u}_{v'} \) for opposite vertices \( v \) and \( v' \). So \( \mathbf{u} \) is an orthonormal representation of \( G \). Using the vector \( e_{d+1} \) as handle, we see that

\[
\vartheta(G) \leq \max_v \frac{1}{(e_{d+1}^T \mathbf{u}_v)^2} = r^2 + 1,
\]
and
\[ \vartheta(G) \geq \sum_{\nu} (\mathbf{e}_{\nu}^T \mathbf{u}_{\nu})^2 = \frac{n}{r^2 + 1}. \]

In particular, we see that \( \chi(G) \) can be arbitrarily large while \( \vartheta(G) \) can be arbitrarily close to 2.

Example 24. [Random graphs] Consider the most basic random graph \( G(n, 1/2) \) (as usual, \( G(n, p) \) denotes a random graph on \( n \) nodes with edge probability \( p \)). It is a nontrivial problem to determine the theta-function of a random graph, and not completely solved. To start with a heuristic, recall Corollary 14 of a self-complementary graph \( G \) with a node-transitive automorphism group has \( \vartheta(G) = \sqrt{n} \). As a special case, the Paley graph \( \text{Pal}_p \) is quasirandom (informally, it behaves like a random graph in many respects) with edge-density 1/2 and \( \vartheta(\text{Pal}_p) = \sqrt{p} \).

The graph \( G(n, 1/2) \) is, of course, not self-complementary, and its automorphism group is trivial, with high probability. However, its distribution is invariant under complementation and also under all permutations of the nodes. Informally, it is difficult to distinguish it from its complement (as it is difficult to distinguish any two random graphs with the same edge-density), and apart from a little variance in the degrees, it is difficult to distinguish any two nodes. So perhaps it does behave like a self-complementary graph with a node-transitive automorphism group would!

This heuristic predicts the right order of magnitude of \( \vartheta(G(n, 1/2)) \), namely it is of the order \( \sqrt{n} \). However, no proof is known that would build on the heuristic above. It is nontrivial to prove that
\[ \frac{1}{3} \sqrt{n} < \vartheta(G(n, 1/2)) < 3\sqrt{n}. \]  \hfill (37)

The result extends to estimating the theta-function for random graphs with other edge-densities. If \( p \) is a constant and \( n \to \infty \), then with high probability,
\[ \frac{1}{3} \sqrt{\frac{p}{1 - p}} \leq \vartheta(G(n, p)) \leq 3 \sqrt{\frac{(1 - p)n}{p}}. \]

We sketch the proof in the case \( p = 1/2 \). First, consider the upper bound. The proof uses the following bound on the eigenvalues of random matrices. Let \( A \) be the matrix defined by
\[ A_{ij} = \begin{cases} -1, & \text{if } ij \in E, \\ 1, & \text{otherwise}. \end{cases} \]

Note that \( E(A_{ij}) = 0 \) and \( A_{ij}^2 = 1 \) for \( i \neq j \). The matrix \( A \) satisfies the conditions in Proposition 16 and hence
\[ \vartheta(G) \leq \lambda_{\max}(A) \leq 3\sqrt{n}. \]
To prove the lower bound, it suffices to invoke Lemma 11 and apply the upper bound to the complementary graph:

\[ \vartheta(G(n, \frac{1}{2})) = \vartheta(G(n, \frac{1}{2})) \geq \frac{n}{\vartheta(G(n, \frac{1}{2}))} \geq \frac{n}{3\sqrt{n}} \geq \frac{1}{3} \sqrt{n} \]

with high probability.

5 Stable sets

5.1 Stability number and theta

We have introduced \( \vartheta(G) \) as an upper bound on the stability number \( \alpha(G) \). How good an approximation of the stability number is obtained this way? We have seen in Theorem 9 that

\[ \alpha(G) \leq \vartheta(G) \leq \chi(G) \]

But \( \alpha(G) \) and \( \chi(G) \) can be very far apart, and unfortunately, the approximation of \( \alpha \) by \( \vartheta \) can be quite as poor. We have seen that for a random graph \( G \) with edge density \( \frac{1}{2} \), we have \( \alpha(G) = O(\ln n) \), but \( \vartheta(G) = \Theta(\sqrt{n}) \) (with high probability as \( n \to \infty \)). Even worse examples can be constructed: sequences of graphs \( G \) for which \( \alpha(G) = n^{o(1)} \) and \( \vartheta(G) = n^{1-o(1)} \); in other words, \( \vartheta/\alpha \) can be larger than \( n^{1-\varepsilon} \) for every \( \varepsilon > 0 \). (The existence of such graphs also follows from the \( P \neq NP \) hypothesis and the theorem that it is NP-hard to determine \( \alpha(G) \) with a relative error less than \( n^{1-\varepsilon} \).) This also implies that \( \vartheta(G) \) does not approximate the chromatic number within a factor of \( n^{1-\varepsilon} \).

These constructions do leave a little room for something interesting, namely in the cases when either \( \alpha \) is very small, or if \( \vartheta \) is very large. There are indeed (rather weak, but useful) results in both cases.

First, consider the case when \( \alpha \) is very small. One can construct a graph with \( \alpha(G) = 2 \) and \( \vartheta(G) = \Omega(n^{1/3}) \). This is as large as \( \vartheta(G) \) can be:

**Theorem 25.** Let \( G \) be a graph with \( n \) nodes and \( \alpha(G) = k \). Then

\[ \vartheta(G) \leq 16n^{k-1}. \]

To prove this theorem, we need a lemma that facilitates recurrence when bounding the theta-function.

**Lemma 26.** Let \( G \) be a graph, and for \( i \in V \), let \( G_i = G[N(i)] \). Then

\[ \vartheta(G) \leq 1 + \max_{i \in V} \sqrt{|N(i)|} \vartheta(G_i). \]

**Proof.** Let \( (v_i, d) \) be an optimal dual orthonormal representation of \( G \). By inequality (31),

\[ \vartheta(G) \leq \max_i \sum_j |v_i^T v_j| = 1 + \max_i \sum_{j \in N(i)} |v_i^T v_j|. \]
For any given \( i \), we use the Cauchy–Schwarz Inequality:

\[
\left( \sum_{j \in \mathcal{N}(i)} |v_j^\top v_i| \right)^2 \leq |\mathcal{N}(i)| \sum_{j \in \mathcal{N}(i)} (v_j^\top v_j)^2 \leq |\mathcal{N}(i)| \vartheta(G_i),
\]

since we can consider \( (v_j : j \in \mathcal{N}(i)) \) as a dual orthonormal representation of \( G_i \) with handle \( v_i \). Combining these two inequalities, we get the Lemma. \( \square \)

**Proof of Theorem 25**. The case \( k = 1 \) is trivial, so assume that \( k > 1 \). We use induction on the number of nodes. Let \( \Delta = \max_i |\mathcal{N}(i)| \) and \( G_i = G[\mathcal{N}(i)] \), then \( \alpha(G_i) \leq k - 1 \) for every node \( i \), and hence \( \vartheta(G_i) \leq 16\Delta^{(k-2)/k} \). So Lemma 26 implies that

\[
\vartheta(G) \leq 1 + \max_i \sqrt{|\mathcal{N}(i)| \vartheta(G_i)} \leq 1 + \sqrt{\Delta \cdot (16\Delta^{k-2/k})} = 1 + 4\Delta^{k-1/k}.
\]

If \( \Delta \leq 3n^{k/(k+1)} \), then

\[
1 + 4\Delta^{k-1/k} \leq 1 + 4 \cdot 3^{k-1/n} n^{k+1/k} - \Delta^{k-1} < 1 + 12n^{k-1/k} < 16n^{k-1/k},
\]

and we are done. If \( \Delta > 3n^{k/(k+1)} \), then let \( i \) be a node with \( |\mathcal{N}(i)| = \Delta \), and consider the partition of the nodes into the sets \( S_1 = \mathcal{N}(i) \) and \( S_2 = \{i\} \cup \mathcal{N}(i) \). Note that \( \alpha(G[S_2]) \leq \alpha(G) - 1 = k - 1 \). So by induction on the number of nodes,

\[
\vartheta(G) \leq \vartheta(G[S_1]) + \vartheta(G[S_2]) \leq 16\Delta^{k-2/k} + 16(n - \Delta)\Delta^{k-1/k}
\]

\[
= 16\Delta^{k-2/k} + 16n^{k-1/k} \left( 1 - \frac{\Delta}{n} \right)^{k-1/k} < 16\Delta^{k-2/k} + 16n^{k-1/k} \left( 1 - \frac{k-1}{k+1} \cdot \frac{\Delta}{n} \right)
\]

\[
= 16n^{k-1/k} + 16\Delta \left( \Delta^{\frac{k-1}{k+1}} - \frac{k-1}{k+1} n^{-\frac{2}{k+1}} \right)
\]

\[
\leq 16n^{k-1/k} + 16\Delta n^{-\frac{2}{k+1}} \left( 3^{\frac{k-1}{k+1}} - \frac{k-1}{k+1} \right) \leq 16n^{k-1/k}.
\]

\( \square \)

To motivate the next theorem, let us think of the case when the chromatic number is small, say \( \chi(G) \leq t \), where we think of \( t \) as a constant. Trivially, \( \alpha(G) \geq n/t \), and nothing better can be said. From the weaker assumption that \( \omega(G) \leq t \) we get a much weaker bound: using the theory of off-diagonal Ramsey numbers, one gets (essentially) a power of \( n \) as a lower bound:

\[
\alpha(G) = \Omega \left( \left( \frac{n}{\log n} \right)^{\frac{1}{t+1}} \right). \tag{38}
\]

The inequality \( \vartheta(G) \leq t \) is a condition that is between the previous two, and the following theorem does give a lower bound on \( \alpha(G) \) that is better than \( \Omega \) (for \( t > 2 \)). In addition, the proof provides a polynomial time randomized algorithm to construct a stable set of the appropriate size, whose use we will explain later.
Theorem 27. Let $G$ be a graph and $t = \vartheta(G)$. Then

$$\alpha(G) \geq \frac{n^{3/10}}{\sqrt{\ln n}}.$$  

Proof. The case $t = 2$ is easy (cf. Exercise 1), so we assume that $t > 2$. By Theorem 10, $t$ is the strict vector chromatic number of $G$, and hence there are unit vectors $u_i \in \mathbb{R}^k$ (in some dimension $k$) such that $u_i^T u_j = -1/(t-1)$ whenever $ij \in E$.

Let $w$ be a random point in $\mathbb{R}^d$ whose coordinates are independent standard Gaussian random variables. Fix an $s > 0$, and consider the set $S = \{ i : w^T u_i \geq s \}$. The inner product $w^T u_i$ has standard Gaussian distribution, and hence the probability that a given node belongs to $S$ is

$$Q(s) = \frac{1}{\sqrt{2\pi}} \int_s^{\infty} e^{-x^2/2} dx,$$

and the expected size of $S$ is $E|S| = Q(s)n$.

Next we show that $S$ does not induce too many edges. Let $ij \in E$, then $|u_i + u_j|^2 = (2t-4)/(t-1)$, and so the probability that both nodes $u_i$ and $u_j$ belong to $S$ can be estimated as follows:

$$P(w^T u_i \geq s, \; w^T u_j \geq s) \leq P(w^T (u_i + u_j) \geq 2s) = Q\left(\sqrt{\frac{2t-2}{t-2}s}\right).$$

Hence the expected number of edges spanned by $S$ satisfies

$$E|E[S]| \leq Q\left(\sqrt{\frac{2t-2}{t-2}s}\right)m,$$

where $m = |E|$. We can delete at most $|E[S]|$ nodes from $S$ (one from each edge it induces) to get a stable set $T$ with expected size

$$E|T| \geq E(|S| - |E[S]|) = E|S| - E|E[S]| \geq Q(s)n - Q\left(\sqrt{\frac{2t-2}{t-2}s}\right)m.$$

We want to choose $s$ so that it maximizes the right hand side. By elementary computation we get that

$$s = \sqrt{\frac{2t-4}{t-1} \ln \frac{m}{n}}$$

is an approximately optimal choice. Using the well known estimates

$$\frac{1}{\sqrt{2\pi}} e^{-s^2/2} < Q(s) < \frac{1}{\sqrt{2\pi}} \frac{s}{s^2 + 1} e^{-s^2/2}, \quad (39)$$

we get

$$\alpha(G) \geq \frac{1}{10\sqrt{\ln(m/n)}} m^{-\frac{t-2}{2}} n^{\frac{2t-2}{t-2}}.$$
If $m < n^{(2t+1)/(t+1)}$, then this proves the theorem.

If $m \geq n^{(2t+1)/(t+1)}$, then there is a node $i$ with degree $\Delta \geq 2m/n \geq 2n^{1/(t+1)}$. Clearly $\vartheta(G[N(i)]) \leq t-1$ (see Exercise 4), and hence by induction on $n$, $G[N(i)]$ has a stable set of size at least $\Delta^{3/t}/(10\sqrt{\ln \Delta}) > n^{3/(t+1)}/(10\sqrt{\ln n})$. This proves the theorem in this case as well.

5.2 The stable set polytope

Stable sets and cliques give rise to important polyhedra associated with graphs. After summarizing some basic properties of these polyhedra, we show that orthogonal representations provide an interesting related convex body, with nice duality properties.

The stable set polytope $\text{STAB}(G)$ of a graph $G$ is the convex hull of incidence vectors of all stable sets. This gives us a polytope in $\mathbb{R}^V$. The stability number $\alpha(G)$ can be obtained by maximizing the linear function $\sum_{i \in V} x_i$ over this polytope, which suggests that methods from linear programming can be used here.

With this goal in mind, we have to find a system of linear inequalities whose solution set is exactly the polytope $\text{STAB}(G)$. It would be best to find a minimal such system, which is unique. If we can find this system, then the task of computing the stability number $\alpha(G)$ of $G$ reduces to maximizing $\sum_{i \in V} x_i$ subject to these constraints, which means solving a linear program. Unfortunately, this system of linear inequalities is in general exponentially large and very complicated. But if we find at least some linear inequalities valid for the stable set polytope, then solving the linear program we get an upper bound on $\alpha(G)$, and for special graphs, we get the exact value.

So we want to find linear inequalities (constraints) valid for the incidence vector of every stable set. We start with the trivial nonnegativity constraints:

$$x_i \geq 0 \quad (i \in V). \quad (40)$$

The fact that the set is stable is reflected by the edge constraints:

$$x_i + x_j \leq 1 \quad (ij \in E). \quad (41)$$

Inequalities (40) and (41) define the fractional stable set polytope $\text{FSTAB}(G)$. Integral points in $\text{FSTAB}(G)$ are exactly the incidence vectors of stable sets, but $\text{FSTAB}(G)$ may have other (nonintegral) vertices, and is in general larger than $\text{STAB}(G)$ (cf. Exercise 17). The case of equality has a nice characterization.

Proposition 28. $\text{STAB}(G) = \text{FSTAB}(G)$ if and only if $G$ is bipartite. \hfill \square

Let $\alpha^f(G)$ denote the maximum of $\sum_i x_i$ over $x \in \text{FSTAB}(G)$. Trivially $\alpha(G) \leq \alpha^f(G)$, and $\alpha^f(G)$ is computable in polynomial time (since (40) and (41) describe a linear program defining $\alpha^f$).

We can strengthen the edge constraints if the graph has larger cliques. Every clique $B$ gives rise to a clique constraint:

$$\sum_{i \in B} x_i \leq 1. \quad (42)$$
Inequalities \[40\] and \[42\] define a polytope \(Q_{\text{STAB}}(G)\), the \textit{clique-constrained fractional stable set polytope} of \(G\). Since cliques in \(G\) correspond to stable sets in \(\overline{G}\) and vice versa, it is easy to see that \(Q_{\text{STAB}}(G)\) is just the antiblocker of \(\text{STAB}(\overline{G})\) (cf. Section ??).

Again, we can introduce a corresponding relaxation of the stability number, namely the quantity \(\alpha^*(G)\) defined as the maximum of \(\sum x_i\) over \(x \in Q_{\text{STAB}}(G)\). This quantity is a sharper upper bound on \(\alpha(G)\) than \(\alpha^f\), but it is NP-hard to compute.

The polytope \(Q_{\text{STAB}}(G)\) is contained in \(\text{FSTAB}(G)\), but is still larger than \(\text{STAB}(G)\) in general. The case of equality leads us to an interesting and rich class of graphs, of which we give a very brief survey.

5.3 \textbf{Perfect graphs}

Recall again the basic inequality

\[\alpha(G) \leq \vartheta(G) \leq \chi(G).\]

For graphs with \(\alpha(G) = \chi(G)\), we have equality here, so \(\vartheta\) is an integer. To know this is useful; just to mention one consequence, the approximation in Theorem ?? it suffices to compute the result with an error bound of \(1/3\).

But which graphs have this nice property? It turns out that the condition \(\alpha(G) = \chi(G)\) does not say much about the structure of \(G\), but a strengthened version of it leads to a very interesting class of graphs. A graph \(G\) is called \textit{perfect}, if for every induced subgraph \(G'\) of \(G\), we have \(\omega(G') = \chi(G')\). Every bipartite graph is perfect, since their induced subgraphs are also bipartite. Figure 6 shows some perfect and nonperfect graphs.

![Figure 6: Some perfect graphs (first row) and some nonperfect graphs (second row).](image)

To be perfect is a rather strong structural property; nevertheless, many interesting classes of graphs are perfect (bipartite graphs, their complements and their linegraphs, interval graphs, comparability and incomparability graphs of posets, chordal graphs, split graphs, etc.).
The following deep characterization of perfect graphs was conjectured by
Berge in 1961 and proved by Chudnovski, Robertson, Seymour and Thomas in
2006:

**Theorem 29** (Strong Perfect Graph Theorem). A graph is perfect if and only
if neither the graph nor its complement contains a chordless odd cycle longer
than 3.

As a corollary we can state the “Weak Perfect Graph Theorem”:

**Theorem 30.** The complement of a perfect graph is perfect.

From this theorem it follows that in the definition of perfect graphs we could
replace the equation $\omega(G') = \chi(G')$ by $\alpha(G') = \chi(G')$. Perfectness can also be
characterized in terms of the stable set polytope:

**Theorem 31.** $\text{STAB}(G) = \text{QSTAB}(G)$ if and only if $G$ is perfect.

Based on our remark above, the condition $\text{STAB}(G) = \text{QSTAB}(G)$ is equiva-

tent to saying that $\text{STAB}(G)$ and $\text{QSTAB}(G)$ are antiblockers, which is a condition

symmetric in $G$ and $\overline{G}$. So Theorem 31 implies Theorem 30.

Turning to algorithms, Theorem 31 implies:

**Corollary 32.** The stability number and the chromatic number of a perfect
graph are polynomial time computable.

Using the algorithms of Corollary 32 one can compute more than just these
values: one can compute a maximum stable set and an optimal coloring in a
perfect graph in polynomial time (see Exercise 18).

It is interesting (and somewhat frustrating) that the only other way of com-
puting the stability number of a perfect graph in polynomial time is to use
the Strong Perfect Graph Theorem and the very advanced structure theory of
perfect graphs developed for it.

Theorem 31 extends to the weighted version of the theta-function. Maximiz-
ing a linear function over $\text{STAB}(G)$ or $\text{QSTAB}(G)$ is NP-hard; but, surprisingly,
$\text{TSTAB}$ behaves much better: Every linear objective function can be maximized
over $\text{TSTAB}(G)$ (with an arbitrarily small error) in polynomial time. This ap-
pplies in particular to $\vartheta(G)$, which is the maximum of $\sum x_i$ over $\text{TSTAB}(G)$.

### 5.4 Orthogonality constraints

For every orthonormal representation $(u_i, c)$ of $G$, we consider the linear con-
straint

$$\sum_{i \in V} (c^T u_i)^2 x_i \leq 1,$$

which we call an orthogonality constraint. The solution set of nonnegativity and
orthogonality constraints is denoted by $\text{TSTAB}(G)$. It is clear that $\text{TSTAB}$ is
a closed, full-dimensional, convex set. The orthogonality constraints are valid
if $x$ is the indicator vector of a stable set of nodes (cf. [0]), and therefore they
are valid for \( \text{STAB}(G) \). Furthermore, every clique constraint is an orthogonality constraint. Indeed, for every clique \( B \), the constraint \( \sum_{i \in B} x_i \leq 1 \) is obtained from the orthogonal representation

\[
i \mapsto \begin{cases} e_1, & i \in B, \\ e_i, & \text{otherwise}, \end{cases} \quad c = e_1.
\]

Hence

\[
\text{STAB}(G) \subseteq \text{TSTAB}(G) \subseteq \text{QSTAB}(G) \tag{44}
\]

for every graph \( G \).

There are several other characterizations of \( \text{TSTAB} \). These are based on an extension of the theta-function to the case when we are also given a weighting \( w : V \to \mathbb{R}_+ \). Generalizing the formulas in Remark \ref{rem:theta-function}, the quantity \( \vartheta(G,w) \) can be defined by any of the following formulas:

\[
\begin{align*}
\vartheta(G,w) &= \min \left\{ \max_{i \in V} \frac{w_i}{(c^T u_i)^2} : u \text{ ONR of } G, |c| = 1 \right\} \tag{45} \\
&= \min \left\{ t \geq 2 : |y_j|^2 = t - w_i, \ y^T_i y_j = -\sqrt{w_i w_j} (ij \in E) \right\} \tag{46} \\
&= \min \left\{ \max_{i \in V} (Y_{ii} + w_i) : Y \succeq 0, Y_{ij} = -\sqrt{w_i w_j} (ij \in E) \right\} \tag{47} \\
&= \max \left\{ \sum_{i,j \in V} w_i w_j Z_{ij} : Z \succeq 0, Z_{ii} = 0 (ij \in E), \sum_i Z_{ii} = 1 \right\} \tag{48} \\
&= \max \left\{ \sum_{i \in V} w_i (d^T v_i)^2 : v \text{ ONR of } G, |d| = 1 \right\}. \tag{49}
\end{align*}
\]

The equivalence of \((\ref{eq:theta-function}) - (\ref{eq:theta-function})\) can be obtained extending the proof Theorem \ref{thm:theta-function} to the node-weighted version (at the cost of a little more computation).

For every orthonormal representation \( u = (u_i : i \in V) \) with handle \( c \), we call the vector \((c^T u_i)^2 : i \in V\) the profile of the orthogonal representation. We can state two further characterizations of \( \text{TSTAB}(G) \):

**Proposition 33.** (a) \( x \in \text{TSTAB}(G) \) if and only if \( \vartheta(G,x) \leq 1 \).

(b) The body \( \text{TSTAB}(G) \) is exactly the set of profiles of dual orthonormal representations of \( G \).

**Proof.** (a) follows from \((\ref{eq:theta-function})\).

(b) The profile of every dual orthonormal representation belongs to \( \text{TSTAB}(G) \); this is equivalent to \((\ref{eq:theta-function})\). Conversely, let \( x \in \text{TSTAB}(G) \). Then \( \vartheta(G,x) \leq 1 \) by (a), so \((\ref{eq:theta-function})\) implies that there is a dual orthonormal representation \( v \) of \( G \) with handle \( d \) for which \( x_i \leq d^T v_i \) for all nodes \( i \in V \). Thus the vectors \( v'_i = (x_i/d^T v_i) v_i \) satisfy \( d^T v'_i = x_i \). The vectors \( v'_i \) are not of unit length, but the vectors

\[
\begin{align*}
v''_i &= \left( \frac{v'_i}{\sqrt{1 - |v'_i|^2}} \right) \\
d'' &= \begin{pmatrix} d \\ 0 \end{pmatrix}
\end{align*}
\]

form a dual orthonormal representation of \( G \) with profile \( x \). \( \square \)
The last characterization of TSTAB(G) is equivalent to the following duality result.

**Corollary 34.** TSTAB(G) is the antiblocker of TSTAB(G). □

Next we determine the vertices of TSTAB(G). Recall that a *vertex* of a convex body K is a boundary point v that is the unique point of intersection of all hyperplanes supporting K at v. This means that there is a pointed convex cone containing K with v as its vertex. This is to be distinguished from an *extreme point*, which is the unique point of intersection of a hyperplane supporting K with K.

**Theorem 35.** The vertices of TSTAB(G) are exactly the incidence vectors of stable sets in G.

Does this imply that TSTAB(G) = STAB(G)? Of course not, since TSTAB(G) (as every convex body) is the convex hull of its extreme points, but not necessarily of its vertices.

**Proof.** The vector 1_A, where A is a stable set of nodes, is the unique common point of n hyperplanes \( x_i = 1 \) (\( i \in A \)) and \( x_i = 0 \) (\( i \in V \setminus A \)), and so it is a vertex of TSTAB(G).

Conversely, let \( z = (z_i : i \in V) \) be a vertex of TSTAB(G). If \( z_i = 0 \) for some node \( i \), then we can delete \( i \): We get a graph \( G' \) for which TSTAB(G') = TSTAB(G) \( \cap \{ z_i = 0 \} \), and so \( z|_{V \setminus i} \) is a vertex of TSTAB(G'), and we can proceed by induction.

So we may assume that \( z_i > 0 \) for all \( i \). Since \( z \in TSTAB(G) \), we can write \( z_i = (d^T v_i)^2 \) for some dual orthonormal representation \( (v_i : i \in V) \) of G and unit vector \( d \).

Let \( a^T x \leq 1 \) be a hyperplane that supports TSTAB(G) at \( z \). Then \( a \in TSTAB(G) \) by Corollary 34 and hence there is an orthonormal representation \( (u_i : i \in V) \) of G and unit vector \( c \) such that \( a_i = (c^T u_i)^2 \). By the same argument as in the derivation of (21), we get

\[
d = \sum_i (c^T u_i)^2 (d^T v_i) v_i.
\]

Multiplying by any vector \( y \), we get

\[
d^T y = \sum_i (c^T u_i)^2 (d^T v_i)(v_i^T y) = \sum a_i (d^T v_i)(v_i^T y).
\]

Thus if \( y \) is not orthogonal to \( d \), then the point \( z' \) defined by \( z'_i = (d^T v_i)(v_i^T y)/(d^T y) \) is contained in the supporting hyperplane \( \sum a_i x_i = 1 \). This holds for every supporting hyperplane at \( z \). Since \( z \) is a vertex, the only common point of hyperplanes supporting TSTAB(G) at \( z \) is \( z \) itself (here we use that \( z \) is a vertex, not just an extreme point). Thus \( z' = z \), which means that \( v_i^T y = (d^T v_i)(d^T y) \) for all \( i \) (here we use that \( z_i \neq 0 \) for all \( i \)). This linear equation must also hold for the vectors \( y \) orthogonal to \( d \). So we get that
\( \mathbf{v}_i = (d^T \mathbf{v}_i) \mathbf{d} \), and since \( \mathbf{v}_i \) is a unit vector, we get \( \mathbf{v}_i = \mathbf{d} \). So no two vectors \( \mathbf{v}_i \) are orthogonal, and thus \( G \) has no edges. So \( z = \mathbb{1}_V \) is the incidence vector of a stable set as claimed.

**Example 36.** Consider the graph \( C_5 \), with node set \{1, \ldots, 5\}. The polytope \( \text{STAB}(C_5) \) has 11 vertices (the origin, the basic unit vectors, and the incidence vectors of nonadjacent pairs of nodes). The facets are defined by the nonnegativity constraints, edge constraints, and single further inequality

\[
x_1 + x_2 + x_3 + x_4 + x_5 \leq 2. \tag{50}
\]

Since \( C_5 \) has no triangles, we have \( \text{QSTAB} = \text{FSTAB} \). This polytope has a single vertex \((\frac{1}{\sqrt{5}}, \ldots, \frac{1}{\sqrt{5}})\) in addition to the incidence vectors of stable sets.

Turning to \( \text{TSTAB} \), we know by Theorem 35 that it has 11 vertices, the same vertices as \( \text{STAB} \). The umbrella construction in Example 8 gives a point \((\frac{1}{\sqrt{5}}, \ldots, \frac{1}{\sqrt{5}})^T \in \text{TSTAB}(C_5) \) (51) which is not in \( \text{STAB}(G) \) by (50). Applying the umbrella construction to the complement, and scaling, we get an orthogonality constraint

\[
x_1 + \cdots + x_5 \leq \sqrt{5}, \tag{52}
\]

showing that the special vertex of \( \text{FSTAB}(C_5) \) does not belong to \( \text{TSTAB}(C_5) \).

This example shows that not every orthogonality constraint follows from the clique constraints. In fact, the number of essential orthogonality constraints is infinite unless the graph is perfect.

**Proposition 37.** \( \text{TSTAB}(G) \) is polyhedral if and only if the graph is perfect.

**Proof.** If \( G \) is perfect, then \( \text{STAB}(G) = \text{QSTAB}(G) \) by Theorem 31 and (44) implies that \( \text{TSTAB}(G) = \text{STAB}(G) = \text{QSTAB}(G) \), so \( \text{TSTAB}(G) \) is polyhedral. To prove the converse, suppose that \( \text{TSTAB}(G) \) is polyhedral, then Theorem 35 implies that \( \text{TSTAB}(G) = \text{STAB}(G) \). We can apply this argument to \( \overline{G} \), since the antiblocker \( \text{TSTAB}(G)^\text{abl} = \text{TSTAB}({\overline{G}}) \) is also polyhedral; we get that

\[
\text{TSTAB}(G) = \text{TSTAB}(\overline{G})^\text{abl} = \text{STAB}(\overline{G})^\text{abl} = \text{QSTAB}(G).
\]

So \( \text{STAB}(G) = \text{TSTAB}(G) = \text{QSTAB}(G) \), which implies that \( G \) is perfect by Theorem 31.

6 Applications

**6.1 Shannon capacity**

In the introduction, we have described how to use orthogonal representations to determine the Shannon zero-error capacity of the pentagon. What happens with other confusion graphs?
Let $V$ be an alphabet with confusion graph $G = (V, E)$. To describe the confusion graph of longer messages, we use the strong product of two graphs. In these terms, $\alpha(G \boxtimes k)$ is the maximum number of non-confusable words of length $k$: words composed of elements of $V$, so that for every two words there is at least one $i$ ($1 \leq i \leq k$) such that the $i$-th letters are different and nonadjacent in $G$, i.e., non-confusable. It is easy to see that

$$\alpha(G \boxtimes H) \geq \alpha(G)\alpha(H).$$

(53)

This implies that

$$\alpha(G^{\boxtimes(k+l)}) \geq \alpha(G^{\boxtimes k})\alpha(G^{\boxtimes l}),$$

(54)

and hence

$$\alpha(G^{\boxtimes k}) \geq \alpha(G)^k.$$  

(55)

The Shannon capacity (zero-error capacity, if we want to be pedantic) of a graph $G$ is the value

$$\Theta(G) = \lim_{k \to \infty} \alpha(G^{\boxtimes k})^{1/k}.$$  

(56)

Inequality (54) implies, via Fekete’s Lemma, that the limit exists, and (55) implies that

$$\Theta(G) \geq \alpha(G).$$

(57)

Rather little is known about this graph parameter for general graphs. For example, it is not known whether $\Theta(G)$ can be computed for all graphs by any algorithm (polynomial or not), although there are several special classes of graphs for which this is not hard. The behavior of $\Theta(G)$ and the convergence in (56) are rather erratic.

Let us describe a few facts we do know. First, let us generalize the argument from the Introduction bounding $\Theta(C_4)$. Let $\chi(G)$ denote the minimum number of complete subgraphs covering the nodes of $G$ (this is the same as the chromatic number of the complementary graph.) Trivially

$$\alpha(G) \leq \chi(G).$$

(58)

Any covering of $G$ by $\chi(G)$ cliques and of $H$ by $\chi(H)$ cliques gives a “product covering” of $G \boxtimes H$ by $\chi(G)\chi(H)$ cliques, and so

$$\chi(G \boxtimes H) \leq \chi(G)\chi(H).$$

(59)

Hence

$$\alpha(G^{\boxtimes k}) \leq \chi(G^{\boxtimes k}) \leq \chi(G)^k,$$

and thus

$$\Theta(G) \leq \chi(G).$$

(60)

It follows that if $\alpha(G) = \chi(G)$, then $\Theta(G) = \alpha(G)$; for such graphs, nothing better can be done than reducing the alphabet to the largest mutually non-confusable subset. In particular, this answers the Shannon capacity problem for perfect graphs.
Instead of $\overline{\chi}$, we can use $\vartheta$ to bound the Shannon capacity:

$$\alpha(G^k) \leq \vartheta(G^k) \leq \vartheta(G)^k,$$

which implies

**Proposition 38.** For every graph $G$,

$$\Theta(G) \leq \vartheta(G).$$

Since $\vartheta(C_5) = \sqrt{5}$, we get that equality holds in Example 2: $\Theta(C_5) = \sqrt{5}$.

This argument can be generalized to an infinite class of graphs:

**Corollary 39.** If $G$ is a self-complementary graph with a node-transitive automorphism group, then $\Theta(G) = \sqrt{n}$.

**Proof.** The diagonal in $G \boxtimes G$ is stable, so $\alpha(G \boxtimes G) = \alpha(G \boxtimes G) \geq n$, and hence $\Theta(G) \geq \sqrt{n}$. On the other hand, $\Theta(G) \leq \vartheta(G) = \sqrt{n}$ by Corollary 14.

**Example 40.** Paley graphs form a class of graphs to which this corollary applies, and whose Shannon capacity can be determined exactly: $\Theta(\text{Paley}_p) = \vartheta(\text{Paley}_p) = \sqrt{p}$. Assuming that the stability number of a Paley graph is polylogarithmic in $p$ (as conjectured), for this infinite family of graphs the Shannon capacity is much higher than the stability number.

The tensor product construction in the proof of Theorem 15 shows that if $G$ has an orthonormal representation in dimension $c$, and $H$ has an orthonormal representation in dimension $d$, the $G \boxtimes H$ has an orthonormal representation in dimension $cd$. It follows that the minimum dimension of any orthonormal representation is an upper bound on $\Theta(G)$. Exercise 3 shows that this bound is never better than $\vartheta(G)$. However, if we consider orthogonal representations over fields of finite characteristic, then the analogue of $\vartheta$ is not defined, but the minimum dimension may provide a better bound on the Shannon capacity than $\vartheta$.

### 6.2 Approximate coloring

Suppose that somebody gives us a graph and guarantees that the graph is 3-colorable, without telling us the 3-coloring itself. Can we find this 3-coloring? (This “hidden 3-coloring problem” may sound artificial, but this kind of situation does arise in cryptography and other data security applications; one can think of the hidden 3-coloring as a “watermark” that can be verified if we know where to look.)

It is easy to argue that knowing that the graph is 3-colorable does not help: it is still NP-hard to find the 3-coloration. But suppose that we would be satisfied with finding a 4-coloration, or 5-coloration, or $(\log n)$-coloration; is this easier? It is known that to find a 4-coloration is still NP-hard, but little is known above this. There is a polynomial time algorithm that, given a 3-colorable graph,
computes a coloring with $O(n^{0.19996})$ colors. As an application of the results in this chapter, we sketch a weaker version of that.

The proof of Theorem 27 can be turned into an algorithm that computes a stable set of size $n^{3/4}/(10\sqrt{\ln n})$. Coloring this set with one color, deleting its nodes, and repeating the procedure with the remaining graph $G_1$, we get a coloring with at most $40n^{1/4}/\sqrt{\ln n}$ colors. Indeed, the number of nodes in $G_1$ is

$$n_1 \leq n - 10 \frac{n^{3/4}}{10\sqrt{\ln n}} = n \left(1 - \frac{1}{10n^{1/4}\sqrt{\ln n}}\right),$$

then by induction, the number of colors we use is at most

$$1 + 40n^{1/4} \sqrt{\ln n_1} \leq 1 + 40n^{1/4} \left(1 - \frac{1}{10n^{1/4}\sqrt{\ln n}}\right)^{1/4} \sqrt{\ln n} = 40n^{1/4} \sqrt{\ln n}.$$

**Exercise 1.** If $\vartheta(G) = 2$, then $G$ is bipartite.

**Exercise 2.** (a) If $H$ is an induced subgraph of $G$, then $\vartheta(H) \leq \vartheta(G)$. (b) If $H$ is a spanning subgraph of $G$, then $\vartheta(H) \geq \vartheta(G)$.

**Exercise 3.** Prove that the minimum dimension in which a graph $G$ has an orthonormal representation is at least $\vartheta(G)$.

**Exercise 4.** Let $G$ be a graph and $v \in V$. (a) $\vartheta(G \setminus v) \geq \vartheta(G) - 1$. (b) If $v$ is an isolated node, then $\vartheta(G \setminus v) = \vartheta(G) - 1$. (c) If $v$ is adjacent to all other nodes, then $\vartheta(G \setminus v) = \vartheta(G)$.

**Exercise 5.** Let $G$ be a graph and let $V = S_1 \cup \cdots \cup S_k$ be a partition of $V$. (a) $\vartheta(G) \leq \sum_i \vartheta(G[S_i])$. (b) If no edge connects nodes in different sets $S_i$, then equality holds. (c) Suppose that any two nodes in different sets $S_i$ are adjacent. How can $\vartheta(G)$ be expressed in terms of the $\vartheta(G[S_i])$?

**Exercise 6.** Prove that the graph parameter $\vartheta^+(G)$ introduced in Remark ?? could be defined in any of the following ways: (a) tightening the conditions in [18] by requiring that $w_i^Tw_j \geq 0$ for all $i$ and $j$; (b) relaxing the conditions in [19] by requiring only $w_i^Tw_j \leq -1/(t-1)$ for $ij \in E$; (c) relaxing the orthogonality conditions in [14] by requiring only $u_i^Tu_j \leq 0$ for $ij \in E$.

**Exercise 7.** Formulate and prove the analogue of Exercise 6 for the graph parameter $\vartheta^- (G)$.

**Exercise 8.** Prove that $\vartheta^- (G) \vartheta^+(G) \geq n$ for every graph $G$.

**Exercise 9.** Prove the following weak converse to Proposition 11: Every graph $G$ has a nonempty subset $S \subseteq V$ such that

$$\vartheta(G) \leq \left(1 + \frac{1}{2} + \cdots + \frac{1}{n}\right) \frac{|S|}{\vartheta(G[S])}.$$
Exercise 11. Prove the formula (36) for even cycles with their longest diagonals added, by showing that the solution given above it is is optimal.

Exercise 12. Construct an optimal dual orthogonal representation for the Kneser graph $K_{n,k}$.

Exercise 13. Let $v_1, \ldots, v_n \in \mathbb{R}^n$ be any set of vectors, and let $A = \text{Gram}(v)$. Prove that
$$\lambda_{\text{max}}(A) = \max \left\{ \sum_i (d^T v_i)^2 : d \in \mathbb{R}^d, |d| = 1 \right\}.$$

Exercise 14. With the notation of Lemma 26, every graph $G$ satisfies
$$\vartheta(G) (\vartheta(G) - 1)^2 \leq \sum_{i \in V} \vartheta(G_i)^2.$$

Exercise 15. Prove that the Euclidean norm of the sum of $n$ unit vectors such that among any three of them some two are orthogonal is at most $16n^2/3$. Prove that this is best possible.

Exercise 16. (a) Show that any stable set $S$ provides a feasible solution of the dual program in (8). (b) Show that any $k$-coloring of $G$ provides a feasible solution of the primal program in (8). (c) Give a new proof of the Sandwich Theorem 9 based on (a) and (b).

Exercise 17. Prove that the vertices of $\text{FSTAB}(G)$ are half-integral, and show by an example that they are not always integral.

Exercise 18. Show that for a perfect graph, a maximum stable set and a coloring with minimum number of colors can be computed in polynomial time.

Exercise 19. The fractional chromatic number $\chi^*(G)$ is defined as the least real number $t$ for which there exists a family $(A_j : j = 1, \ldots, p)$ of stable sets in $G$, and nonnegative weights $(\tau_j : j = 1, \ldots, p)$ such that $\sum_j \tau_j = t$ and $\sum_j \tau_j 1_{A_j} \geq 1_V$. The fractional clique number $\omega^*(G)$ is the largest real number $s$ for which there exist nonnegative weights $(\sigma_i : i \in V)$ such that $\sum_i \sigma_i = s$ and $\sum_{i \in A} \sigma_i \leq 1$ for every stable set $A$.

(a) Prove that $\omega(G) \leq \omega^*(G)$ and $\chi(G) \geq \chi^*(G)$.

(b) Prove that $\chi^*(G) = \omega^*(G)$.

(b) Prove that $\vartheta(G) \leq \chi^*(G)$.