Graph parameters and invariants of the orthogonal group

Regts, G.

Citation for published version (APA):
Regts, G. (2013). Graph parameters and invariants of the orthogonal group.
Chapter 7

Edge-reflection positive partition functions of vertex-coloring models

Recall from Section 5.1 that the partition function of a vertex-coloring model is also the partition function of an edge-coloring model. In this chapter we characterize, using some fundamental results from geometric invariant theory, for which vertex-coloring models their partition functions are edge-reflection positive, i.e., for which vertex-coloring models their partition functions are partition functions of real edge-coloring models.

This chapter is based on [54] except for Section 7.2, which is based on joint work with Jan Draisma [20, Section 6].

7.1 Introduction

In his paper [66] (see also [67]) on the characterization of partition functions of real edge-coloring models, Szegedy gave an explicit way to construct from a vertex-coloring model \((a, B)\) over \(\mathbb{C}\) an edge-coloring model \(h\) over \(\mathbb{C}\) such that \(p_{a,B}(H) = p_h(H)\) for every \(H \in \mathcal{G}\). We will now describe this construction.

Let \((a, B)\) be an \(n\)-color vertex-coloring model over \(\mathbb{C}\). As \(B\) is symmetric we can write \(B = U^T U\) for some \(n \times k\) (complex) matrix \(U\), for some \(k\). Let \(u_1, \ldots, u_n \in \mathbb{C}^k\) be the columns of \(U\). Define the edge-coloring model \(h\) by \(h := \sum_{i=1}^n a_i \text{ev}_{u_i}\), where for \(u \in \mathbb{C}^k\), \(\text{ev}_u \in R(\mathbb{C})^*\) is the linear map defined by
p \mapsto p(u)$ for $p \in R(C)$. (Recall that $R(C) = \mathbb{C}[x_1, \ldots, x_k].$)

**Lemma 7.1** (Szegedy [66]). Let $(a, B)$ and $h$ be as above. Then $p_{a,B} = p_h$.

**Proof.** Let $G = (V, E) \in \mathcal{G}$. Then $p_h(G)$ is equal to

$$
\sum_{\phi: E \rightarrow \{0, 1\}} \prod_{e \in V} h(\prod_{e \in \delta(v)} x_{\phi(e)}) = \sum_{\phi: E \rightarrow \{0, 1\}} \prod_{e \in V} \left( \sum_{i=1}^n a_i \prod_{e \in \delta(v)} u_i(\phi(e)) \right)
$$

(7.1)

$$
= \sum_{\psi: V \rightarrow \{0, 1\}} \prod_{e \in V} \left( \sum_{e \in \delta(v)} a_{\psi(v)} \prod_{e \in \delta(v)} u_{\psi(v)}(\phi(e)) \right)
$$

$$
= \sum_{\psi: V \rightarrow \{0, 1\}} \prod_{e \in V} a_{\psi(v)} \cdot \sum_{\phi: E \rightarrow \{0, 1\}} \prod_{e \in \delta(v)} u_{\psi(v)}(\phi(e))
$$

$$
= \sum_{\psi: V \rightarrow \{0, 1\}} \prod_{e \in V} a_{\psi(v)} \cdot \prod_{e \in \delta(v)} u_{\psi(v)}(\phi(e)) u_{\psi(v)}(\phi(vw))
$$

$$
= \sum_{\psi: V \rightarrow \{0, 1\}} \prod_{e \in V} a_{\psi(v)} \cdot \prod_{vw \in E} \prod_{i=1}^k u_{\psi(v)}(i) u_{\psi(v)}(i) = p_{a,B}(G).
$$

Where the last line follows from the fact that $U^T U = B$. This completes the proof.

Note that the proof of Lemma 7.1 also shows that if an edge-coloring model $h$ is of the form $h = \sum_{i=1}^n a_i e v_i$ for certain nonzero $a_i \in \mathbb{C}^k$ and certain vectors $u_i \in \mathbb{C}$, then the partition function of $h$ is equal to the partition function of $(a, B)$ (on $\mathcal{G}$), where $a = (a_1, \ldots, a_n)$ and $B_{i,j} = u_i^T u_j$. We will sometimes abuse notation and call $h$ a vertex-coloring model.

Let $(a, B)$ be a vertex-coloring model. If $B$ is positive semidefinite, then $h$ can be taken to be real valued, that is, in view of Theorem 5.2, $p_{a,B}$ is edge-reflection positive. Szegedy [66] moreover observed that for $B = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$, with $a, b \geq 0$, $p_{a,B}$ is also edge-reflection positive. Clearly, for $a = 0$ and $b = 1$, this matrix is not positive semidefinite. This made him ask the question, which partition of vertex-coloring models are edge-reflection positive (cf. [66, Question 3.2]).

In this chapter we give a complete characterization of edge-reflection positive partition functions of vertex-coloring models over $\mathbb{C}$. Let $h = \sum_{i=1}^n a_i e v_i$, for nonzero $a_i$ and distinct vectors $u_i \in \mathbb{C}^k$. We start by giving a simple characterization in terms of the $u_i$ and $a_i$ for $p_h$ to be edge-reflection positive.

**Lemma 7.2.** Let $u_1, \ldots, u_n \in \mathbb{C}^k$ be distinct vectors, let $a \in (\mathbb{C}^*)^n$ and let $h := \sum_{i=1}^n a_i e v_i$. Then $h$ is an edge-coloring model over $\mathbb{R}$ if and only if the set \{\(\begin{pmatrix} u_i \\ a_i \end{pmatrix}\) | $i = 1, \ldots, n\} is closed under complex conjugation.
7.1. Introduction

Proof. Suppose first that the set \( \{ \begin{pmatrix} u_i \\ a_i \end{pmatrix} \mid i = 1, \ldots, n \} \) is closed under complex conjugation. Then for \( p \in R(\mathbb{R}) \), \( h(p) = \sum_{i=1}^{n} a_i p(u_i) = \sum_{i=1}^{n} a_i p(u_i) = h(p) \). Hence, \( h(p) \in \mathbb{R} \). So \( h \) is real valued.

Now the 'only if' part. By possibly adding some vectors to \( \{u_1, \ldots, u_n\} \) and extending the vector \( a \) with zero's, we may assume that \( \{u_1, \ldots, u_n\} \) is closed under complex conjugation. We must show that \( u_i = \pi_j \) implies \( a_i = \pi_j \). We may assume that \( u_1 = \pi_2 \). Using interpolating polynomials (cf. [17, Lemma 2.9]) we find \( p \in R(\mathbb{C}) \) such that \( p(u_j) = 1 \) if \( j = 1, 2 \) and \( 0 \) otherwise. Let \( p' := 1/2(p + \overline{p}) \). Then \( p' \in R(\mathbb{R}) \) and consequently, \( h(p') = \sum_{i=1}^{n} a_i p(u_i) = a_1 + a_2 \in \mathbb{R} \). Similarly, there exists \( q \in R(\mathbb{C}) \) such that \( q(u_1) = i \), \( q(u_2) = -i \) and \( q(u_i) = 0 \) if \( j > 2 \). Setting \( q' := 1/2(q + \overline{q}) \) and applying \( h \) to it, we find that \( i(a_1 - a_2) \in \mathbb{R} \). So we conclude that \( a_1 = \overline{a_2} \). Continuing this way proves the lemma.

Lemma 7.2 clearly explains why for \( B = \begin{pmatrix} a & b \\ b & a \end{pmatrix} \), with \( a, b \geq 0 \), we have that \( p_{1, B} \) is edge-reflection positive. Here is another example.

Example 7.1.

Let \( B = \begin{pmatrix} 0 & 2 & 0 & 2 \\ 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 4 \\ 2 & 0 & 4 & 0 \end{pmatrix} \) and let \( U = \begin{pmatrix} 1 & 1 & 1 & 1 \\ i & -i & i & -i \\ 0 & 0 & 1 & 1 \\ 0 & 0 & i & -i \end{pmatrix} \). (7.2)

Then \( U^T U = B \), and so by Lemma 7.2, \( p_{1, B} \) is equal to the partition function of a real edge-coloring model.

One might think that Lemma 7.2 already gives the answer to Szegedy’s question, but the only thing it says is that if \( h := \sum_{i=1}^{n} a_i e_{u_i} \) for certain \( a \in (\mathbb{C}^*)^n \) and distinct \( u_1, \ldots, u_n \in \mathbb{C}^k \), then it is easy to check whether \( h \) is real. In case \( h \) is not real valued, it does not rule out the possibility that there is another real-valued edge-coloring model \( h' \) (with possible more than \( k \) colors) such that \( p_{h}(H) = p_{h'}(H) \) for all graphs \( H \). Yet, surprisingly, a certain converse to Lemma 7.2 holds. We need however a few more definitions to state it.

For a \( k \times n \) matrix \( U \) we denote its columns by \( u_1, \ldots, u_n \). By \( U^* \) we denote the conjugate transpose of \( U \). Let, for any \( k, (\cdot, \cdot) \) denote the standard bilinear form on \( \mathbb{C}^k \). We call the matrix \( U \) nondegenerate if the span of \( u_1, \ldots, u_n \) is nondegenerate with respect to \( (\cdot, \cdot) \). In other words, if \( \text{rk}(U^T U) = \text{rk}(U) \). We think of vectors in \( \mathbb{C}^k \) as vectors in \( \mathbb{C}^l \) for any \( l \geq k \). We can now state the main result of this chapter. The proof will be given in Section 7.3.
Theorem 7.3. Let \((a, B)\) be a twin-free \(n\)-color vertex-coloring model. Let \(U\) be a nondegenerate \(k \times n\) matrix such that \(U^T U = B\). Then the following are equivalent:

(i) \(p_{a, B} = p_y\) for some real edge-coloring model \(y\),

(ii) there exist \(l \geq k\) and \(g \in O_l(C)\) such that the set \(\{ \begin{pmatrix} gu_i \\ a_i \end{pmatrix} \mid i = 1, \ldots, n \}\) is closed under complex conjugation,

(iii) there exist \(l \geq k\) and \(g \in O_l(C)\) such that \(\sum_{i=1}^{n} a_i \text{ev}_{gu_i}\) is real.

If moreover, \(UU^* \in \mathbb{R}^{k \times k}\), then we can take \(g\) equal to the identity in (ii) and (iii).

Observe that if the set of columns of \(gU\) is closed under complex conjugation, then \(gU(gU)^*\) is real. So the existence of a nondegenerate matrix \(U\) such that \(U^T U = B\) and \(UU^*\) is real, is a necessary condition for \(p_{a, B}\) to be the partition function of an edge-coloring model over \(\mathbb{R}\).

In case \(B\) is real, there is an easy way to obtain a \(k \times n\) rank \(k\) matrix \(U\), where \(k = \text{rk}(B)\), such that \(UU^* \in \mathbb{R}^{k \times k}\) and \(U^T U = B\), using the spectral decomposition of \(B\). So by Theorem 7.3, we get the following characterization of partition functions of real vertex-colorings that are partition functions of real edge-coloring models. We will state it as a corollary.

Corollary 7.4. Let \((a, B)\) be a twin-free \(n\)-color vertex-coloring model over \(\mathbb{R}\). Then \(p_{a, B} = p_h\) for some real edge-coloring model \(h\) if and only if for each \(i \in [n]\) there exists \(j \in [n]\) such that

(i) \(a_i = a_j\),

(ii) for each eigenvector \(v\) of \(B\) with eigenvalue \(\lambda\):

\[
\begin{align*}
\lambda > 0 & \Rightarrow v_i = v_j, \\
\lambda < 0 & \Rightarrow v_i = -v_j.
\end{align*}
\]

We will now give some examples to illustrate Theorem 7.3 and Corollary 7.4.

Example 7.2. Let \(G\) be the graph on two nodes \(x_1\) and \(x_2\) with node weights equal to 1; the loop at \(x_1\) has weight 1; the loop at \(x_2\) has weight 0 and the edge \(x_1x_2\) has weight 1. Then \(\text{hom}(H, G)\) is equal to the number of independent sets of \(H\). Using Theorem 7.3, it is easy to see that the partition function of any real edge-coloring model can not be equal to \(\text{hom}(\cdot, G)\). This can also be easily seen using Theorem 5.2.

Example 7.3. For any \(n \in \mathbb{N}\) with \(n \geq 2\) consider \(K_n\), the complete graph on \(n\) vertices. Then \(\text{hom}(H, K_n)\) is equal to the number of proper \(n\)-colorings of
7.2 Orbits of vertex-coloring models

The corresponding vertex-coloring model is \((I, J - I)\), where \(I\) denotes the all-ones vector, \(J\) the all-ones matrix and \(I\) the identity matrix. The eigenvalue \(-1\) of \(J - I\) has multiplicity \(n - 1\). Using that the eigenspace corresponding to \(-1\) is equal to \(\mathbb{1}^\perp\), it is easy to see, using Corollary 7.4, that \(\text{hom}(\cdot, K_n)\) is equal to the partition function of a real edge-coloring model if and only if \(n = 2\). We do not know whether it is easy to deduce this from Theorem 5.2.

In view of Theorem 5.2, Example 7.3 shows that for each \(n \geq 3\) there exists \(k, t \in \mathbb{N}\), \(k\)-fragments \(F_1, \ldots, F_t\) and \(\lambda \in \mathbb{R}^t\) such that \(\sum_{i,j=1}^t \lambda_i \lambda_j \text{hom}(F_i * F_j, K_n) < 0\). It would be interesting to characterize for which (twin-free) graphs \(G\) the invariant \(\text{hom}(\cdot, G)\) is edge-reflection positive. By Corollary 7.4, this depends on spectral properties of \(G\).

The remainder of this chapter is devoted to proving Theorem 7.3. The proof is based on a well-known generalization of the Hilbert-Mumford criterion, a fundamental result in geometric invariant theory. In the next section we use this criterion to characterize when the \(O_k(C)\)-orbit of a vertex-coloring model is closed. In Section 7.3 we then use this result combined with some ideas of Kempf and Ness to give a proof of Theorem 7.3.

7.2 Orbits of vertex-coloring models

In this section we will work with a general algebraically closed field \(\mathbb{F}\) of characteristic zero. Let \(k \in \mathbb{N}\) and let \(V\) be a \(k\)-dimensional vectorspace over \(\mathbb{F}\) equipped with a nondegenerate symmetric bilinear form \((\cdot, \cdot)\). Identify \(V\) with \(\mathbb{F}^k\) through the bilinear form. Let \(u_1, \ldots, u_n \in V\) be distinct, and let \(a \in (\mathbb{F}^*)^n\). Define the edge-coloring model \(h\) by \(h := \sum_{i=1}^n a_i \text{ev}_{u_i}\). In this section we will consider the \(O_k\)-orbit \(O_k h_{\leq e} \subset \mathbb{F}^{N_{\leq e}}\) for \(e \in \mathbb{N}\) (recall that \(h_{\leq e}\) denotes the restriction of \(h\) to the space of polynomials of degree at most \(e\)); we will characterize in terms of the \(u_i\) when this orbit is closed. Our main tool will be a well-known generalization of the Hilbert-Mumford criterion.

7.2.1 The one-parameter subgroup criterion

There is a beautiful criterion for closedness of orbits involving one-parameter subgroups of \(O_k\), i.e., homomorphisms \(\lambda : \mathbb{F}^* \to O_k\) of algebraic groups. We call a basis \(v_1, \ldots, v_k\) of \(V\) such that \((v_i, v_j) = \delta_{k+1,i+j}\) for all \(i, j\), (i.e. so that the Gram matrix of the basis has zeroes everywhere except ones on the longest anti-diagonal) a canonical basis. Let \(\lambda : \mathbb{F}^* \to O_k\) be a one-parameter subgroup. Then there exists a canonical basis \(v_1, \ldots, v_k\) of \(V\) such that \(\lambda(t) v_i = t^{d_i} v_i\) for
each \( t \in F^* \), for some integral weights \( d_1 \geq \cdots \geq d_k \) satisfying \( d_i = -d_{k+1-i} \) for all \( i \). This follows, for instance, from [25, Section 2.1.2] or [4, §23.4] (ignoring the subtle rationality issues there as \( F \) is algebraically closed) and the fact that all maximal tori are conjugate [4, §11.3]. Conversely, given a canonical basis \( v_1, \ldots, v_k \) and such a sequence of \( d_i \)'s, the homomorphism \( \lambda : F^* \to O_k \) defined by \( \lambda(t)v_i = t^{d_i}v_i \) is a one-parameter subgroup of \( O_k \).

The one-parameter subgroup criterion says the following: let \( W \) be a finite-dimensional \( O_k \)-module, and let \( w \in W \). Consider the orbit \( O_kw \subseteq W \). By Theorem [4.7], the Zariski closure of this orbit contains a unique closed orbit \( C \). Then there exists a one-parameter subgroup \( \lambda \) such that \( \lim_{t \to 0} \lambda(t)w \) exists and is contained in \( C \) (the Hilbert-Mumford criterion considers the special case where \( C = \{0\} \)). Here the existence of the limit by definition means that the morphism \( F^* \to W, \ t \mapsto \lambda(t)w \) extends to \( F \). It then does so in a unique manner, and the value at 0 is declared the limit. Put differently, just like \( V \), the \( \lambda \)-module \( W \) decomposes into a direct sum of weight spaces (cf. [25, Lemma 1.6.4]), and the condition is that all components of \( w \) in \( \lambda \)-weight spaces corresponding to negative weights are zero, and the component of \( w \) in the zero weight space is the limit. We record the one-parameter subgroup criterion as a theorem.

**Theorem 7.5.** Let \( W \) be a finite dimensional \( O_k \)-module, let \( w \in W \) and let \( C \) be the unique closed orbit contained in \( \overline{O_kw} \). Then there exists a one-parameter subgroup \( \lambda : F^* \to O_k \) such that the limit \( \lim_{t \to 0} \lambda(t)w \) exists and is contained in \( C \).

For a proof of Theorem 7.5 see e.g. [3, Theorem 4.2] or [32, Theorem 1.4].

**Example 7.4.** Recall the edge-coloring model \( h \) from Example 6.1 in Section 6.1, \( h \in k[x_1, x_2]^* \) is zero on all polynomials of degree at least 2. The restriction of \( h \) to the space of polynomials of degree at most 1 is an element of \( (V^*)^* = V \), namely, equal to \( v_1 := e_1 + ie_2 \). This is an isotropic vector relative to the bilinear form, and so is its complex conjugate \( v_2 := e_1 - ie_2 \). So the sequence \( 1/\sqrt{2}v_1, 1/\sqrt{2}v_2 \) forms a canonical basis of \( V \). The linear map \( V \to V \) scaling \( v_1 \) with \( t \in F \) and \( v_2 \) with \( t^{-1} \) is an element of the orthogonal group. Explicitly, this gives the one-parameter subgroup

\[
\lambda(t) = \frac{1}{2t} \begin{bmatrix} 1 + t^2 & i - it^2 \\ -i + it^2 & 1 + t^2 \end{bmatrix} \in O_2
\]

(7.3)

with the property that \( \lim_{t \to 0} \lambda(t)h_{\leq e} = 0 \) for all \( e \).
7.2. Orbits of vertex-coloring models

7.2.2 Application to vertex-coloring models

Here we will use the one-parameter subgroup criterion to characterize when the $O_k$-orbit of $h_{\leq \epsilon}$ is closed.

We will need the following well-known result.

**Proposition 7.6.** Let $u_1, \ldots, u_n \in V$ be nonzero. If $w_1, \ldots, w_n \in V$ are nonzero vectors such that

$$ (u_i, u_j) = (w_i, w_j) \text{ for all } i, j = 1, \ldots, n, $$

(7.4) then there exists $g \in O_k$ such that $gu_i = v_i$ for all $i \in [n]$.

For completeness we will sketch the proof.

**Proof.** Let $U$ denote the span of the $u_i$ and $W$ the span of the $w_i$. If $U = V$, we can just define a linear map $g : V \rightarrow V$ by $u_i \mapsto w_i$ for each $i$. It is easy to see that $g$ is well defined and that $g$ preserves the bilinear form, that is $g \in O_k$.

Next, if the bilinear form restricted to $U$ is nondegenerate, then we can reduce to the previous case by adding an orthonormal basis for $U^\perp$ to $\{u_1, \ldots, u_n\}$ and an orthonormal basis for $W^\perp$ to $\{w_1, \ldots, w_n\}$.

Finally, if $U$ is degenerate we can find $i \in [n]$ such that $(u_i, u_j) = 0$ for all $j \in [n]$. Let $U' \subset U$ and $W' \subset W$ be complements to $u_i$ and $w_i$ respectively. Then we may choose $u \in U'^\perp$ such that $(u_i, u) = 1$ and such that $(u, u) = 0$ (cf. [36] XV, §9]). Similarly, we may choose $w \in W'^\perp$ such that $(w_i, w) = 1$ and such that $(w, w) = 0$. Now add $u$ to $\{u_1, \ldots, u_n\}$ and $w$ to $\{w_1, \ldots, w_n\}$ and note that the dimension of $U$ (and of $W$) increases by one. Now just proceed until $U$ becomes nondegenerate so that we can reduce to the previous case. \qed

**Theorem 7.7.** Let $F = \mathbb{F}$, let $u_1, \ldots, u_n \in V$ be distinct and let $a \in (F^*)^n$. Let $h := \sum_{i=1}^n a_i ev_{u_i}$ and let $\epsilon \geq 3n$. Then the orbit $O_k h_{\leq \epsilon}$ is closed if and only if the restriction of the bilinear form to the span of the $u_i$ is nondegenerate.

**Proof.** Let $U \subset V$ denote the space spanned by the $u_i$. Suppose first that the bilinear form restricted to $U$ is degenerate. Then we may assume that $(u_1, u_i)$ is 0 for all $i \in [n]$. Define $h' = \sum_{i=2}^n a_i ev_{u_i}$. By Proposition 7.6 there exists for each $\epsilon > 0$, $g \in O_k$ such that $gu_1 = \epsilon u_1$ and $gu_i = u_i$ for $i \geq 2$. This implies that $h'_{\leq \epsilon}$ is contained in the closure of the orbit of $h_{\leq \epsilon}$. We will now show that $h'_{\leq \epsilon}$ is not contained in the orbit $h_{\leq \epsilon}$.

Let $I(h) \subset R$ be the set of polynomials $p$ of degree at most $n$ such that $h(pq) = 0$ for all polynomials $q$ of degree at most $n - 1$. Then

$$ I(h) = \{ p \in R \mid \deg(p) \leq n, p(u_i) = 0 \text{ for } i = 1, \ldots, n \}. $$

(7.5)
The inclusion ‘⊇’ is clear. To see the other inclusion, let \( p_1, \ldots, p_n \) be interpolating polynomials at the \( u_i \), i.e., the \( p_i \) are polynomials of degree \( n - 1 \) such that \( p_i(u_j) = \delta_{i,j} \) for all \( i, j = 1, \ldots, n \) (cf. [17] Lemma 2.9). Then for a polynomial \( p \) of degree at most \( n \) we have that \( \deg(pp_i) \leq 2n - 1 \leq e \) and \( h(pp_i) = 0 \) if and only if \( p(u_i) = 0 \). This shows (7.5), which in turn implies

\[
\{ u \in V \mid p(u) = 0 \text{ for all } p \in I(h) \} = \{ u_1, \ldots, u_n \}.
\]

But since the \( u_i \) are distinct, (7.6) applied to \( h' \) implies that \( gh_{\leq e} \neq h'_{\leq e} \) for any \( g \in O_k \), showing that the orbit of \( h_{\leq e} \) is not closed.

For the converse, assume that the bilinear form restricted to \( U \) is nondegenerate. We will prove that the \( O_k \)-orbit of \( h_{\leq e} \) is closed. Let \( \lambda : \mathbb{F}^* \to O_k \) be a one-parameter subgroup such that \( \lim_{t \to 0} \lambda(t)h_{\leq e} \) exists. We will show that it lies in the orbit of \( h_{\leq e} \). Let \( v_1, \ldots, v_k \) be a canonical basis of \( V \) with \( \lambda(t)v_j = t^{d_j}v_j \) for weights \( d_1 \geq \cdots \geq d_k \). Let \( x_1, \ldots, x_k \) be the basis of \( V^* \) dual to \( v_1, \ldots, v_k \). For any monomial \( x^\alpha, \alpha \in \mathbb{N}^k \), we have

\[
(\lambda(t)h)(x^\alpha) = h(\lambda(t)^{-1}x^\alpha) = h(t^{a_1d_1+\cdots+a_kd_k}x^\alpha) = t^\alpha \sum_{i=1}^{n} a_i x^\alpha(u_i),
\]

where \( \alpha \cdot d := a_1d_1 + \cdots + a_kd_k \). By assumption, if \( x^\alpha \) is a monomial of degree at most \( e \), the limit for \( t \to 0 \) in (7.7) exists. Note that this implies for \( |\alpha| \leq e \):

\[
\alpha \cdot d < 0 \Rightarrow h(x^\alpha) = \sum_{i=1}^{n} a_i x^\alpha(u_i) = 0.
\]

(7.8)

In what follows, we exclude the trivial cases where \( k = 0 \) and where \( k = 1 \) and \( u_1 \) is the zero vector; in both of these cases the orbit of \( h \) is just a single point.

Next let \( b \in \{1, \ldots, k\} \) be the maximal index with \( x_b(U) \neq \{0\} \), and order the \( u_i \) such that \( x_b(u_1), \ldots, x_b(u_1) \neq 0 \) (\( l > 0 \)) and \( x_b(u_{i+1}), \ldots, x_b(u_n) = 0 \). By maximality of \( b \), \( U \) is contained in the span of \( v_1, \ldots, v_b \). So if \( d_b \) is nonnegative, then \( \lim_{t \to 0} \lambda(t)(u_1, \ldots, u_n) \) exists, and is by Proposition 7.6 contained in the orbit of \( (u_1, \ldots, u_n) \). (Since \( U \) is nondegenerate, the equations describing the orbit are given by (7.4).) Then also \( h_{\leq e} \) and \( \lim_{t \to 0} \lambda(t)h_{\leq e} \) lie in the same orbit. So we may assume that \( d_b < 0 \). (In particular, \( b > k/2 \).

Since the coordinates \( x_{b+1}, \ldots, x_k \) vanish identically on \( U \), it follows that \( U \) is contained in the subspace of \( V \) perpendicular to \( v_1, \ldots, v_{k-b} \). As \( U \) is nondegenerate, it does not contain a nonzero linear combination of \( v_1, \ldots, v_{k-b} \). This means, in particular, that the coordinates \( x_{k-b+1}, \ldots, x_b \) together separate the points \( u_1, \ldots, u_l \). Then so do the monomials \( x_{k-b+1}^2, \ldots, x_{b-1}^2x_b^3 \). Note that the dot product \( \alpha \cdot d \) is negative for each of these (e.g., for the first, it

76
equals $d_{k-b+1} + 2d_b = d_b < 0$ and from there the dot product decreases weakly to the right). It follows that there exists a linear combination $p$ of those cubic monomials for which $p(u_1), \ldots, p(u_l)$ are distinct and nonzero. Then, by (7.8) and the fact that $p(u_{l+1}) = \cdots = p(u_n) = 0$, the vector $(a_1, \ldots, a_l)^T$ is in the kernel of the Vandermonde matrix

$$
\begin{bmatrix}
p(u_1) & \cdots & p(u_l) \\
p(u_1)^2 & \cdots & p(u_l)^2 \\
\vdots & \vdots & \vdots \\
p(u_1)^l & \cdots & p(u_l)^l
\end{bmatrix}
$$

(7.9)

since the degree of $p^l$ is $3l \leq e$. This implies that $a_1, \ldots, a_l$ are all zero, contrary to the assumption that all $a_i$ are nonzero. This proves that the orbit of $h_{\leq e}$ is closed for $e \geq 3n$.

### 7.3 Proof of Theorem 7.3

In this section we give a proof of Theorem 7.3 using Theorem 7.7. We first need some preparations.

Let $W \in \mathbb{C}^{l \times n}$ be any matrix and consider the function $f_W : O_l(\mathbb{C}) \to \mathbb{R}$ defined by

$$
g \mapsto \text{tr}(W^* g^* g W) = \text{tr}((gW)^* g W),
$$

(7.10)

for $g \in O_l(\mathbb{C})$, where $\text{tr}(M)$ denotes the trace of a matrix $M$ and $M^*$ the conjugate transpose of $M$. This function was introduced by Kempf and Ness [33] in the context of connected reductive linear algebraic groups acting on finite dimensional vector spaces. Note that $f_W$ is left-invariant under $O_l(\mathbb{R})$ and right-invariant under $\text{Stab}(W) := \{g \in O_l(\mathbb{C}) \mid gW = W\}$. Let $e \in O_l(\mathbb{C})$ denote the identity. We are interested in the situation that the infimum of $f_W$ over $O_l(\mathbb{C})$ is equal to $f_W(e)$.

**Lemma 7.8.** The function $f_W$ has the following properties:

(i) $\inf_{g \in O_l(\mathbb{C})} f_W(g) = f_W(e)$ if and only if $WW^* \in \mathbb{R}^{l \times l}$,

(ii) If $WW^* \in \mathbb{R}^{l \times l}$, then $f_W(e) = f_W(g)$ if and only if $g \in O_l(\mathbb{R}) \cdot \text{Stab}(W)$.

**Proof.** We start by showing that $f_W$ has a critical point at $e$ if and only if $WW^* \in \mathbb{R}^{l \times l}$.

(7.11)
By definition, a critical point of \( f_W \) is a point \( g \in O_l(C) \) such that \( (Df_W)_g(X) = 0 \) for all \( X \in T_g(O_l(C)) \), where \( T_g(O_l(C)) \) is the tangent space of \( O_l(C) \) at \( g \) and where \( (Df_W)_g \) is the derivative of \( f_W \) at \( g \). It is well known that the tangent space of \( O_l(C) \) at \( e \) is the space of skew-symmetric matrices, i.e., \( T_e(O_l(C)) = \{ X \in C^{l \times l} \mid X^T + X = 0 \} \). It is easy to see that the derivative of \( f_W \) at \( e \) is the real-linear map \( (Df_W)_e \in \text{Hom}_R(C^{l \times l}, R) \) defined by \( Z \mapsto \text{tr}(W^*(Z + Z^*)W) \). Now let \( Z \) be skew-symmetric and write \( Z = X + iY \), with \( X, Y \in \mathbb{R}^{l \times l} \). Note that \( Z \) is skew-symmetric if and only if both \( X \) and \( Y \) are skew-symmetric. Write \( W = V + iT \) with \( V, T \in \mathbb{R}^{l \times l} \). Then \( (Df_W)_e(Z) \) is equal to

\[
\begin{align*}
\text{tr}((V^T - iT^T)(X + iY + X^T - iT))(V + iT)) \\
= 2\text{tr}((V^T - iT^T)iY(V + iT)) \\
= 2\text{tr}(T^TYV) - 2\text{tr}(V^TYT) = 4\text{tr}(T^TYV),
\end{align*}
\]

(7.12)

where we use that \( X \) and \( Y \) are skew symmetric, and standard properties of the trace. So \( Df_e(Z) = 0 \) for all skew symmetric \( Z \) if and only if \( V^T = TV^T \). That is, if and only if \( WW^* \in \mathbb{R}^{l \times l} \). This shows (7.11).

By a result of Kempf and Ness (cf. [33, Theorem 0.1]) we can now conclude that (i) and (ii) hold. However, we will give an independent and elementary proof.

First the proof of (i). Note that (7.11) immediately implies that \( f_W \) does not attain a minimum at \( e \) if \( WW^* \notin \mathbb{R}^{l \times l} \). (This follows easily from the method of Lagrange multipliers.) Conversely, suppose \( WW^* \in \mathbb{R}^{l \times l} \). Since \( WW^* \) is real and positive semidefinite there exists a real matrix \( V \) such that \( WW^* = VV^T \). Now note that, by the cyclic property of the trace, \( f_W(g) = \text{tr}(g^*gWW^*) \). So we have \( f_W = f_V \). Let \( I \) denote the identity matrix. Take any \( g = X + iY \in O_l(C) \), where \( X, Y \in \mathbb{R}^{l \times l} \). Using that \( X^TY - Y^TX = I \), and the fact that \( f_W \) is real valued, we find that

\[
f_W(g) = \text{tr}((X^TY + Y^TX)VV^T) = \text{tr}(VV^T) + 2\text{tr}(Y^TYVV^T) \\
= 2\text{tr}(VV^T) + \text{tr}(YV(YV)^T) \geq \text{tr}(VV^T) = f_W(e).
\]

(7.13)

This proves (i).

Next, suppose that \( f_W(g) = f_W(e) \) for some \( g \in O_l(C) \). Again, since \( WW^* \) is real and positive semidefinite there exists a real matrix \( V \) such that \( WW^* = VV^T \). Moreover, the span of the columns of \( V \) is equal to the span of the columns of \( W \). This implies that \( \text{Stab}(V) = \text{Stab}(W) \). Now write \( g = X + iY \), with \( X, Y \in \mathbb{R}^{l \times l} \). As, by (7.13), \( f_W(g) = f_W(e) \) if and only if \( YV = 0 \), it follows that \( gV = XV + iYV = XV \) is a real matrix. Let \( v_1, \ldots, v_n \) be the columns of \( V \). Then, since by definition of the orthogonal group, \( (gv_i, gv_j) = (v_i, v_j) \) for
all \( i, j \), and since the \( g v_i \) and the \( v_i \) are real, there exists \( g_1 \in O_l(\mathbb{R}) \) such that \( g_1 g V = V \). This implies that \( g \in O_l(\mathbb{R}) \cdot \text{Stab}(V) \). This finishes the proof of (ii). \( \square \)

For any \( e \), let \( \langle \cdot, \cdot \rangle \) denote the Hermitian inner product on \( \mathbb{C}^{N_{\leq e}} \) induced by the standard Hermitian inner product on \( \bigoplus_{i=1}^{e} (\mathbb{C}^l)^{\otimes i} \), by viewing elements of \( \mathbb{C}^{N_{\leq e}} \) as symmetric tensors. The next proposition has as conclusion a special case of Theorem 0.2 in [33].

**Proposition 7.9.** Let \( h \) be any \( l \)-color edge-coloring model. Let \( C \) be the unique closed orbit in \( O_l(\mathbb{C}) h_{\leq e} \). Then there exists \( h'_{\leq e} \in C_e \) such that

\[
\inf_{g \in O_l(\mathbb{C})} \langle gh_{\leq e}, gh_{\leq e} \rangle \geq \langle h'_{\leq e}, h'_{\leq e} \rangle. \tag{7.14}
\]

Moreover, the infimum is attained if and only if \( h_{\leq e} \in C_e \).

**Proof.** Clearly, the infimum is attained at some \( g \in O_l(\mathbb{C}) \) if \( h_{\leq e} \in C_e \). So we can take \( h' = gh \).

Now assume that \( h_{\leq e} \not\in C_e \). Fix any \( g \in O_l(\mathbb{C}) \), write \( y := gh \) and, as in the proof of Theorem 7.7, let \( \lambda : C^* \to O_l(\mathbb{C}) \) be a one-parameter subgroup such that \( \lim_{t \to 0} \lambda(t) y_{\leq e} = y'_{\leq e} \in C_e \). Let \( v_1, \ldots, v_l \) be a canonical basis of \( \mathbb{C}^l \) with \( \lambda(t) v_j = t^{d_j} v_j \) for weights \( d_1 \geq \cdots \geq d_l \). Let \( x_1, \ldots, x_l \) be the basis of \( (\mathbb{C}^l)^* \) dual to \( v_1, \ldots, v_l \). Recall from (7.7) that for any monomial \( x^a \), \( a \in \mathbb{N}^l \), we have \( (\lambda(t)) y(x^a) = t^{a \cdot d} y(x^a) \). Since, by assumption, the limit \( \lim_{t \to 0} t^{a \cdot d} y(x^a) \) exists for \( |a| \leq e \), this implies:

\[
y'_{\leq e}(x^a) = \begin{cases} 0 (= y(x^a)) & \text{if } a \cdot d < 0, \\
y(x^a) & \text{if } a \cdot d = 0, \\
0 & \text{if } a \cdot d > 0. \end{cases} \tag{7.15}
\]

For \( e' \leq e \) and \( \phi : [e'] \to [l] \) let \( \phi \cdot d := a \cdot d \), for \( a \in \mathbb{N}^l \) such that \( x^a = x_{\phi(1)} \cdots x_{\phi(l)} = x^\phi \). Then, as \( y_{\leq e} \neq y'_{\leq e} \), by (7.15),

\[
\langle y_{\leq e}, y_{\leq e} \rangle = \sum_{\phi : [e'] \to [l], \phi \cdot d \geq 0} y(x^\phi) y(x^\phi) > \sum_{\phi : [e'] \to [l], \phi \cdot d = 0} y(x^\phi) y(x^\phi) = \langle y'_{\leq e}, y'_{\leq e} \rangle. \tag{7.16}
\]

So for each \( g \in O_l(\mathbb{C}) \) we can find \( y'_{\leq e} \in C_e \) such that \( \langle gh_{\leq e}, gh_{\leq e} \rangle > \langle y'_{\leq e}, y'_{\leq e} \rangle \), proving the first statement. This moreover implies that the infimum is not attained if \( h_{\leq e} \not\in C_e \), finishing the proof. \( \square \)
We need one more lemma before we can prove Theorem 7.3.

**Lemma 7.10.** Let $h := \sum_{i=1}^n a_i e v u_i \in R(C)^*$, with $a \in (C^*)^n$ and $u_1, \ldots, u_n \in C^k$ distinct. Suppose the bilinear form restricted to the span of the $u_i$ is nondegenerate. If $y$ is a real $l$-color edge-coloring model such that $p_h(H) = p_y(H)$ for all $H \in G$, then there exists $g \in O_l(C)$ such that $gh = y$.

**Proof.** We may assume that $l \geq k$. Recall that in case $l > k$ we add colors to $h$. This is done by appending the $u_i$'s with zero's. Note that the bilinear form restricted to the span of the $u_i$ remains nondegenerate. Then, by Theorem 7.7 for each $d \geq 3n$, the orbit $O_l h_{\leq d}$ is equal to the unique closed orbit $C_d$. We will now show that the orbit of $y_{\leq d}$ is also equal to $C_d$ for any $d$.

For any $e \leq d$, $O_l(C)$ embeds naturally into $O_l(C)$. Let $g \in O_l(C)$, and write $g = X + iY$, with $X, Y \in R^{F \times F}$. Then, using that $X^T X - Y^T Y = I$,

$$
\langle gy_e, gy_e \rangle = \langle Xy_e, Xy_e \rangle + \langle Yy_e, Yy_e \rangle = \langle y_e, y_e \rangle + 2 \langle Yy_e, Yy_e \rangle \geq \langle y_e, y_e \rangle.
$$

As this holds for any $e \leq d$, we can now conclude by Proposition 7.9 that the orbit of $y_{\leq d}$ is closed.

We now claim that there exists $g \in O_l(C)$ such that $gh = y$. Indeed, since $\text{Stab}(y_{\leq d}) = \bigcap_{d' \leq d} \text{Stab}(y_{\leq d'})$ and since $O_l(C)$ is Noetherian, there exists $d_1 \geq 3n$ such that $\text{Stab}(y_{\leq d_1}) = \bigcap_{d \in \mathbb{N}} \text{Stab}(y_{\leq d})$. Recall that we have a canonical bijection from $O_l(C)/\text{Stab}(y_{\leq d})$ to $C_d$ given by

$$g \text{Stab}(y_{\leq d}) \mapsto gy_{\leq d}$$

(cf. the proof of Theorem 6.11). This implies that for any $d \geq d_1$, if $g \in O_l(C)$ is such that $gy_{\leq d} = h_{\leq d}$, then also $gy = h$. This proves the lemma. \(\square\)

Now we can give a proof of Theorem 7.3.

**Theorem 7.3.** Let $(a,b)$ be a twin-free $n$-color vertex-coloring model. Let $U$ be a nondegenerate $k \times n$ matrix such that $U^T U = B$. Then the following are equivalent:

(i) $p_{a,B} = p_y$ for some real edge-coloring model $y$,

(ii) there exist $l \geq k$ and $g \in O_l(C)$ such that the set $\left\{ \left( \begin{array}{c} gu_i \\ a_i \end{array} \right) \right\}_{i=1 \ldots, n}$ is closed under complex conjugation,

(iii) there exist $l \geq k$ and $g \in O_l(C)$ such that $\sum_{i=1}^n a_i e v g u_i$ is real.

If moreover, $UU^* \in R^{k \times k}$, then we can take $g$ equal to the identity in (ii) and (iii).
Proof. Observe that since \((a, B)\) is twin free, the columns of \(U\) are distinct. Lemma 7.2 implies the equivalence of (ii) and (iii) for the same \(g\) and \(l\) in (ii) and (iii). Moreover, since \((gU)^T gU = U^T g^T gU = U^T U = B\), for any \(g \in O_l(C)\), Lemma 7.1 shows that (iii) implies (i).

Let \(u_1, \ldots, u_n\) be the columns of \(U\) and let \(h := \sum_{i=1}^{n} a_i e v_{u_i}\). We will now prove that (i) implies (iii). Let \(y\) be a real \(l\)-color edge-coloring model such that \(p_{a, B} = p_y\). Since \(U\) is nondegenerate, we may assume, by Lemma 7.10, that \(y = g h\) for some \(g \in O_l(C)\). Now note that \(gh = \sum_{i=1}^{n} a_i e v_{gu_i}\). This shows that (i) implies (iii).

Now assume that \(UU^* \in \mathbb{R}^{k \times k}\). We will show that (i) implies (iii) with \(g = e\). Let \(y\) be a real \(l\)-color edge-coloring model such that \(p_{a, B} = p_y\). Just as above, we may assume that \(y = \sum_{i=1}^{n} a_i e v_{gu_i}\), for some \(g \in O_l(C)\). Lemma 7.2 implies that the set \(\{gu_i\}\) is closed under complex conjugation. This implies that \(gU(gU)^* \in \mathbb{R}^{l \times l}\). So by Lemma 7.8 (i) the infimum of \(f_{gU}\) is attained at \(e\). Equivalently, the infimum of \(f_U\) is attained at \(g\). Since \(UU^* \in \mathbb{R}^{k \times k}\), this implies, by Lemma 7.8 (ii), that \(g \in O_l(\mathbb{R}) \cdot \text{Stab}(U)\). Hence \(g = g_1 \cdot s\) for some \(g_1 \in O_l(\mathbb{R})\) and \(s \in \text{Stab}(U)\). Now note that since \(sh = h\) we have that \(h = g_1^{-1} y\) and hence \(h\) is real. \(\square\)