2002

A Partial Order on the Orthogonal Group

Colum Watt

Technological University Dublin, colum.watt@dit.ie

Thomas Brady

Dublin City University, thomas.brady@dcu.ie

Follow this and additional works at: https://arrow.tudublin.ie/scschmatart

Part of the Mathematics Commons

Recommended Citation


This Article is brought to you for free and open access by the School of Mathematics at ARROW@TU Dublin. It has been accepted for inclusion in Articles by an authorized administrator of ARROW@TU Dublin. For more information, please contact yvonne.desmond@tudublin.ie, arrow.admin@tudublin.ie, brian.widdis@tudublin.ie.

This work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 License
A partial order on the Orthogonal Group

Thomas Brady  
School of Mathematical Sciences  
Dublin City University  
Glasnevin, Dublin 9  
Ireland  
tom.brady@dcu.ie

Colum Watt  
School of Mathematics  
Trinity College  
Dublin 2  
Ireland  
colum@maths.tcd.ie

Abstract: We define a natural partial order on the orthogonal group and completely describe the intervals in this partial order. The main technical ingredient is that an orthogonal transformation induces a unique orthogonal transformation on each subspace of the orthogonal complement of its fixed subspace.

Keywords: orthogonal group, partial order.

Let $V$ be an $n$-dimensional vector space over a field $F$ and let $O(V)$ be the orthogonal group of $V$ with respect to a fixed anisotropic symmetric bilinear form $\langle , \rangle$. In this note we will define a natural partial order on $O(V)$ and completely describe the intervals in this partial order. The main technical ingredient is that an orthogonal transformation $A$ on $V$ induces a unique orthogonal transformation on each subspace of the orthogonal complement of the fixed subspace of $A$.

Recall that $A \in O(V)$ if $A : V \to V$ is linear and satisfies $\langle A(\vec{v}), A(\vec{w}) \rangle = \langle \vec{v}, \vec{w} \rangle$ for all $\vec{v}, \vec{w} \in V$. For standard results on symmetric bilinear forms and their associated orthogonal groups see [1], but note that we are making the further assumption that the form is anisotropic.

For each $A \in O(V)$, we define two subspaces of $V$, $F(A) = \ker(A - I)$ and $M(A) = \im(A - I)$, where $I$ is the identity operator on $V$. We note that $F(A)$ is the $+1$-eigenspace of $A$, sometimes called the fixed subspace of $A$. We will write $V = V_1 \perp V_2$ whenever $V$ is the orthogonal direct sum of
Proposition 1  \( V = F(A) \perp M(A) \)

Proof. Since the dimensions of \( F(A) \) and \( M(A) \) are complementary and the form is anisotropic, it suffices to show that these subspaces are orthogonal.

So let \( \vec{x} \in F(A) \) and \( \vec{y} \in M(A) \). Then \( \vec{x} = A(\vec{x}) \) and \( \vec{y} = (A - I)\vec{z} \) for some \( \vec{z} \in V \). Thus

\[
\langle \vec{x}, \vec{y} \rangle = \langle \vec{x}, (A - I)\vec{z} \rangle = \langle \vec{x}, A(\vec{z}) \rangle - \langle \vec{x}, \vec{z} \rangle = \langle A(\vec{x}), A(\vec{z}) \rangle - \langle \vec{x}, \vec{z} \rangle = 0.
\]

q.e.d.

We will be concerned with how the dimensions of these subspaces behave when we take products in \( O(V) \). For notational convenience we will write \( |U| \) for \( \dim(U) \).

Proposition 2  \(|M(AB)| \leq |M(A)| + |M(B)|\) for \( A, B \in O(V) \).

Proof. Using the identities \(|U| + |V| = |U + V| + |U \cap V|\), \( F(A) \cap F(B) \subseteq F(AB) \) and \( F(A) + F(B) \subseteq V \) we find that

\[
|F(A)| + |F(B)| = |F(A) + F(B)| + |F(A) \cap F(B)| \leq n + |F(AB)|,
\]

from which the result follows. q.e.d.

This result is proved in a more general setting in [2]. However, from the proof above we see that equality occurs if and only if

\[
F(A) \cap F(B) = F(AB) \quad \text{and} \quad F(A) + F(B) = V.
\]

Thus, using the identities \([U + V]^{\perp} = U^{\perp} \cap V^{\perp}\) and \( U^{\perp} + V^{\perp} = [U \cap V]^{\perp}\) we get the following characterization.

Corollary 1  \(|M(AB)| = |M(A)| + |M(B)|\) if and only if \( M(AB) = M(A) \oplus M(B) \).

Definition 1  We will write \( A \leq C \) if \( |M(C)| = |M(A)| + |M(A^{-1}C)| \).

Proposition 3  The relation \( \leq \) is a partial order on \( O(V) \) and satisfies

\[
A \leq B \leq C \Rightarrow A^{-1}B \leq A^{-1}C.
\]
Proof. Reflexivity is immediate. To establish antisymmetry suppose $A \leq C$ and $C \leq A$. Then

$$|M(C)| = |M(A)| + |M(A^{-1}C)| = |M(C)| + |M(C^{-1}A)| + |M(A^{-1}C)|$$

giving $F(C^{-1}A) = F(A^{-1}C) = V$ or $A = C$.

To establish transitivity, suppose $A \leq B$ and $B \leq C$. Then

$$|M(C)| \leq |M(A)| + |M(A^{-1}B)| + |M(B^{-1}C)|$$

$$= |M(A)| + |M(B)| - |M(A)| + |M(C) - |M(B)||$$

$$= |M(C)|$$

So both of the inequalities are actually equalities. The first line gives $A \leq C$ and $\leq$ is transitive. The third line gives the second assertion above. q.e.d.

The association of the subspace $M(A)$ to an element $A \in O(V)$ defines a map $M$ from $O(V)$ to the set of subspaces of $V$. The next sequence of lemmas shows that the restriction of $M$ to the interval $[I, C] = \{A \in O(V) \mid A \leq C\}$ is a bijection onto the set of subspaces of $M(C)$.

In what follows we fix $C$ and a subspace $W$ of $M(C)$ and we suppose that $A \in O(V)$ satisfies $M(A) = W$. We define $U$ to be the unique subspace of $M(C)$ which satisfies $|U| = |W|$ and $(C - I)U = W$. This is possible since $C - I$ is invertible when restricted to $M(C)$.

**Lemma 1** If $W \subseteq M(C)$ then $V = W^\perp \oplus U$.

**Proof.** Since the subspaces have complementary dimensions it suffices to show that their intersection is trivial. So let $\vec{x} \in W^\perp \cap U$. Then $\vec{x} \in W^\perp$ and $(C - I)\vec{x} = \bar{w}$ for some $\bar{w} \in W$. Thus $C\vec{x} = \vec{x} + \bar{w}$, with $\vec{x} \in W^\perp$ and $\bar{w} \in W$ so that

$$\langle \vec{x}, \vec{x} \rangle = \langle C\vec{x}, C\vec{x} \rangle = \langle \vec{x} + \bar{w}, \vec{x} + \bar{w} \rangle = \langle \vec{x}, \vec{x} \rangle + \langle \bar{w}, \bar{w} \rangle.$$

Thus $\bar{w} = 0$ since $\langle , \rangle$ is anisotropic and $\vec{x} = 0$ since $C - I$ is an isomorphism on $M(C)$. q.e.d.

**Lemma 2** $F(A^{-1}C) \subseteq F(C) \perp U$.

3
Lemma 5

Indeed the case. Since $F$ is an identity on $V$, $\text{Proj}_V$ gives $\vec{x}$. Let $\vec{z}$.

Proof. It is not at all clear from this formula that $A$ gives $\vec{x}$ uniquely as $\vec{x} = \vec{y} + \vec{z}$ with $\vec{y} \in F(C)$ and $\vec{z} \in M(C)$. Thus

$$(C - I)\vec{z} = (C - I)\vec{x} = (A - I)\vec{x} \in M(A) = W,$$

giving $\vec{z} \in U$. This gives $F(A^{-1}C) \subseteq F(C) + U$ and the orthogonality of the subspaces follows since $U \subseteq M(C)$. q.e.d.

Lemma 3 If $M(A) = W$ and $A \subseteq C$ then $F(A^{-1}C) = F(C) \perp U$.

Proof. Since $A \subseteq C$ we have $|M(A^{-1}C)| = |M(C)| - |M(A)|$ so that

$$|F(A^{-1}C)| = n - |M(A^{-1}C)| = n - |M(C)| + |W| = |F(C)| + |U|.$$ 

This dimension calculation can now be combined with Lemma 2. q.e.d.

It is now possible to give a formula for $A$. If $V = V_1 \oplus V_2$ we define the projection $\text{Proj}_{V_2}^V$ to be the linear transformation which coincides with the identity on $V_1$ and with the zero transformation on $V_2$.

Lemma 4 If $A \subseteq C$ and $M(A) = W$ then $A = I + (C - I)\text{Proj}_U^W$.

Proof. If $M(A) = W$ then $F(A) = W^\perp$ so that $A$ coincides with $I$ on $W^\perp$. Since $F(A^{-1}C)$ contains $U$ by Lemma 3, $A$ coincides with $C$ on $U$. Thus $A - I$ coincides with the zero transformation on $W^\perp$ and with $C - I$ on $U$, giving $A - I = (C - I)\text{Proj}_U^W$, by Lemma 1. q.e.d.

It is not at all clear from this formula that $A$ is orthogonal. However this is indeed the case.

Lemma 5 $A = I + (C - I)\text{Proj}_U^W \in O(V)$.

Proof. Let $\vec{x}, \vec{y} \in V$ and use Lemma 1 to express $\vec{x} = \vec{x}_1 + \vec{x}_2$, $\vec{y} = \vec{y}_1 + \vec{y}_2$, with $\vec{x}_1, \vec{y}_1 \in U$ and $\vec{x}_2, \vec{y}_2 \in W^\perp$. Then, using the fact that $A$ coincides with $I$ on $W^\perp$ and with $C$ on $U$,

$$\langle A(\vec{x}), A(\vec{y}) \rangle = \langle C(\vec{x}_1) + \vec{x}_2, C(\vec{y}_1) + \vec{y}_2 \rangle = \langle C\vec{x}_1, C\vec{y}_1 \rangle + \langle C\vec{x}_1, \vec{y}_2 \rangle + \langle \vec{x}_2, C\vec{y}_1 \rangle + \langle \vec{x}_2, \vec{y}_2 \rangle = \langle \vec{x}_1, \vec{y}_1 \rangle + \langle C\vec{x}_1, \vec{y}_2 \rangle + \langle \vec{x}_2, C\vec{y}_1 \rangle + \langle \vec{x}_2, \vec{y}_2 \rangle = \langle \vec{x}, \vec{y} \rangle + \langle (C - I)\vec{x}_1, \vec{y}_2 \rangle + \langle \vec{x}_2, (C - I)\vec{y}_1 \rangle = \langle \vec{x}, \vec{y} \rangle,$$
since both \((C - I)\vec{x}_1\) and \((C - I)\vec{y}_1\) lie in \(W\). q.e.d.

We will call \(A\) the transformation induced by \(C\) on \(W\). Combining the above lemmas we get the following result.

**Theorem 1** If \(C \in O(V)\) and \(W\) is a subspace of \(M(C)\) then there exists a unique \(A \in O(V)\) satisfying \(A \leq C\) and \(M(A) = W\).

The induced transformations are familiar objects for two special classes of subspace.

**Corollary 2** If \(W \subseteq M(C)\) is an invariant subspace of \(C\), then the induced transformation on \(W\) is the restriction of \(C\) to \(W\).

**Proof.** In this case, \(U = W\) and the projection in the formula for \(A\) becomes an orthogonal projection. q.e.d.

**Corollary 3** If \(\text{char}(F) \neq 2\) and \(W\) is a one-dimensional subspace of \(M(C)\) then the orthogonal transformation induced by \(C\) on \(W\) is always the orthogonal reflection in \(W^⊥\).

**Proof.** Since \(W\) is one-dimensional \(A\) must act on \(W\) by multiplication by a scalar \(\alpha\). The orthogonality of \(A\) forces \(\alpha^2 = 1\) and \(W = M(A)\) gives \(\alpha \neq 1\). q.e.d.

The poset \((O(V), \leq)\) is not a lattice, since distinct elements \(C_1\) and \(C_2\) with \(M(C_1) = M(C_2)\) cannot have a common upper bound. However the intervals are lattices and can be easily described.

**Theorem 2** If \(A \leq C\) in \(O(V)\) and \(|M(C)| - |M(A)| = m\) then the interval \([A, C] = \{B \in O(V) \mid A \leq B \leq C\}\) is isomorphic to the lattice of subspaces of \(F^m\) under inclusion.

**Proof.** The lattices of subspaces of \(F^m\) under inclusion is isomorphic to the interval \([M(A), M(C)]\) in the lattice of subspaces of \(V\). The function \(B \mapsto M(B)\) is a bijection from the interval \([A, C]\) to the latter interval by Theorem 1. This map respects the partial orders by Corollary 1. To see that the inverse map respects the partial orders suppose that \(M(A) \subseteq W_1 \subseteq W_2 \subseteq M(C)\). Let \(B_1, B_2\) be the transformations induced on \(W_1, W_2\) respectively by \(C\) and let \(B'_1\) be the transformation induced on \(W_1\) by \(B_2\). Then \(B'_1 \leq B_2 \leq C\) gives \(B'_1 \leq C\), but \(M(B_1) = M(B'_1) = W_1\) so the uniqueness part of Theorem 1 gives \(B_1 = B'_1\) and \(B_1 \leq B_2\). q.e.d.

Each chain in \(M(C)\) thus gives rise to a special factorization of \(C\).
**Corollary 4** If $C \in O(V)$ and $W_1 \subset W_2 \subset \cdots \subset W_k = M(C)$ is a chain of subspaces in $M(C)$ then $C$ factors uniquely as a product of $k$ transformations $C = B_1B_2 \ldots B_k$, with $B_1B_2 \ldots B_i \leq C$ and $M(B_1B_2 \ldots B_i) = W_i$.

*Proof.* If we define $C_i$ to be the transformation induced by $C$ on $W_i$ then $B_i = (C_{i-1})^{-1}C_i$. q.e.d.

The case where this chain is maximal gives a strong version of the Cartan-Dieudonné theorem.

**Corollary 5** If $\text{char}(F) \neq 2$, $C \in O(V)$ with $|M(C)| = k$ and $W_1 \subset W_2 \subset \cdots \subset W_k = M(C)$ is a maximal flag in $M(C)$ then $C$ factors uniquely as a product of $k$ reflections, $C = R_1R_2 \ldots R_k$, with $M(R_1R_2 \ldots R_i) = W_i$.

*Proof.* Here the transformation $B_i$ defined in Corollary 4 satisfies $|M(B_i)| = 1$ so that $B_i$ is a reflection by Corollary 3. q.e.d.

**Note 1** Using similar methods one can prove analogs of all the above results in the case of a unitary transformation over a finite-dimensional complex vector space. In this case we deal with complex linear subspaces, the induced transformations are unitary (and hence complex linear) and complex reflections replace the above reflections.

**References**
