Clifford algebras and Lie groups

Eckhard Meinrenken

Lecture Notes, University of Toronto, Fall 2009. (Revised version of lecture notes from 2005)

CHAPTER 1

Symmetric bilinear forms

Throughout, $\mathbb{K}$ will denote a field of characteristic $\neq 2$. We are mainly interested in the cases $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, and sometimes specialize to those two cases.

1. Quadratic vector spaces

Suppose V is a finite-dimensional vector space over $\mathbb{K}$. For any bilinear form $B\colon V \times V \to \mathbb{K}$, define a linear map

$$B^\flat\colon V \to V^*,\ v \mapsto B(v, \cdot).$$

The bilinear form B is called *symmetric* if it satisfies $B(v_1, v_2) = B(v_2, v_1)$ for all $v_1, v_2 \in V$. Since $\dim V < \infty$ this is equivalent to $(B^\flat)^* = B^\flat$. The symmetric bilinear form B is uniquely determined by the associated quadratic form, $Q_B(v) = B(v, v)$ by the *polarization identity*,

$$B(v, w) = \tfrac{1}{2}\big(Q_B(v + w) - Q_B(v) - Q_B(w)\big). \tag{1}$$

The *kernel* (also called *radical*) of B is the subspace

$$\ker(B) = \{v \in V \mid B(v, v_1) = 0 \text{ for all } v_1 \in V\},$$

i.e. the kernel of the linear map $B^\flat$. The bilinear form B is called *non-degenerate* if $\ker(B) = 0$, i.e. if and only if $B^\flat$ is an isomorphism. A vector space V together with a non-degenerate symmetric bilinear form B will be referred to as a *quadratic vector space*. Assume for the rest of this chapter that (V, B) is a quadratic vector space.

Definition 1.1. A vector $v \in V$ is called *isotropic* if $B(v, v) = 0$, and *non-isotropic* if $B(v, v) \neq 0$.

For instance, if $V = \mathbb{C}^n$ over $\mathbb{K} = \mathbb{C}$, with the standard bilinear form $B(z, w) = \sum_{i=1}^n z_i w_i$, then $v = (1, i, 0, \ldots, 0)$ is an isotropic vector. If $V = \mathbb{R}^2$ over $\mathbb{K} = \mathbb{R}$, with bilinear form $B(x, y) = x_1 y_1 - x_2 y_2$, then the set of isotropic vectors $x = (x_1, x_2)$ is given by the 'light cone' $x_1 = \pm x_2$.

The *orthogonal group* $\mathrm{O}(V)$ is the group

$$\mathrm{O}(V) = \{A \in \mathrm{GL}(V) \mid\ B(Av, Aw) = B(v, w) \text{ for all } v, w \in V\}. \tag{2}$$

The subgroup of orthogonal transformations of determinant 1 is denoted $\mathrm{SO}(V)$, and is called the *special orthogonal group*.

For any subspace $F \subset V$, the *orthogonal* or *perpendicular* subspace is defined as

$$F^\perp = \{v \in V | \ B(v, v_1) = 0 \text{ for all } v_1 \in F\}.$$

The image of $B^\flat(F^\perp) \subset V^*$ is the annihilator of F. From this one deduces the dimension formula

$$\dim F + \dim F^\perp = \dim V \tag{3}$$

and the identities

$$(F^\perp)^\perp = F, \quad (F_1 \cap F_2)^\perp = F_1^\perp + F_2^\perp, \quad (F_1 + F_2)^\perp = F_1^\perp \cap F_2^\perp$$

for all $F, F_1, F_2 \subset V$. For any subspace $F \subset V$ the restriction of B to F has kernel $\ker(B|_{F\times F}) = F \cap F^\perp$.

Definition 1.2. A subspace $F \subset V$ is called a *quadratic subspace* if the restriction of B to F is non-degenerate, that is $F \cap F^\perp = 0$.

Using $(F^\perp)^\perp = F$ we see that F is quadratic $\Leftrightarrow$ $F^\perp$ is quadratic $\Leftrightarrow$ $F \oplus F^\perp = V$.

As a simple application, one finds that any non-degenerate symmetric bilinear form B on V can be 'diagonalized'. Let us call a basis $\epsilon_1, \ldots, \epsilon_n$ of V an *orthogonal basis* if $B(\epsilon_i, \epsilon_j) = 0$ for all $i \neq j$.

Proposition 1.3. *Any quadratic vector space (V, B) admits an orthogonal basis $\epsilon_1, \ldots, \epsilon_n$. If $\mathbb{K} = \mathbb{C}$ one can arrange that $B(\epsilon_i, \epsilon_i) = 1$ for all i. If $\mathbb{K} = \mathbb{R}$ or $K = \mathbb{Q}$, one can arrange that $B(\epsilon_i, \epsilon_i) = \pm 1$ for all i.*

Proof. The proof is by induction on $n = \dim V$, the case $\dim V = 1$ being obvious. If $n > 1$ choose any non-isotropic vector $\epsilon_1 \in V$. The span of ϵ_1 is a quadratic subspace, hence so is $\mathrm{span}(\epsilon_1)^\perp$. By induction, there is an orthogonal basis $\epsilon_2, \ldots, \epsilon_n$ of $\mathrm{span}(\epsilon_1)^\perp$. If $\mathbb{K} = \mathbb{C}$ (resp. $\mathbb{K} = \mathbb{R}, \mathbb{Q}$), one can rescale the ϵ_i such that $B(\epsilon_i, \epsilon_i) = 1$ (resp. $B(\epsilon_i, \epsilon_i) = \pm 1$). $\square$

We will denote by $\mathbb{K}^{n,m}$ the vector space $\mathbb{K}^{n+m}$ with bilinear form given by $B(\epsilon_i, \epsilon_j) = \pm \delta_{ij}$, with a $+$ sign for $i = 1, \ldots, n$ and a $-$ sign for $i = n+1, \ldots, n+m$. If $m = 0$ we simple write $\mathbb{K}^n = \mathbb{K}^{n,0}$, and refer to the bilinear form as *standard*. The Proposition above shows that for $\mathbb{K} = \mathbb{C}$, and quadratic vector space (V, B) is isomorphic to $\mathbb{C}^n$ with the standard bilinear form, while for $\mathbb{K} = \mathbb{R}$ it is isomorphic to some $\mathbb{R}^{n,m}$. (Here n, m are uniquely determined, although it is not entirely obvious at this point.)

2. Isotropic subspaces

Let (V, B) be a quadratic vector space.

Definition 2.1. A subspace $F \subset V$ is called *isotropic*[1] if $B|_{F\times F} = 0$, that is $F \subset F^\perp$.

[1]In some of the literature (e.g. C. Chevalley [?] or L. Grove [?]), a subspace is called isotropic if it contains at least one non-zero isotropic vector, and totally isotropic if all of its vectors are isotropic.

The polarization identity (1) shows that a subspace $F \subset V$ is isotropic if and only if all of its vectors are isotropic. If $F \subset V$ is isotropic, then

$$\dim F \leq \dim V/2 \tag{4}$$

since $\dim V = \dim F + \dim F^{\perp} \geq 2 \dim F$.

PROPOSITION 2.2. *For isotropic subspaces F, F' the following three conditions*

(a) *$F + F'$ is quadratic,*
(b) *$V = F \oplus (F')^{\perp}$,*
(c) *$V = F' \oplus F^{\perp}$*

are equivalent, and imply that $\dim F = \dim F'$. *Given an isotropic subspace $F \subset V$ one can always find an isotropic subspace F' satisfying these conditions.*

PROOF. We have

$$\begin{aligned}(F + F') \cap (F + F')^{\perp} &= (F + F') \cap F^{\perp} \cap (F')^{\perp} \\ &= (F + (F' \cap F^{\perp})) \cap (F')^{\perp} \\ &= (F \cap (F')^{\perp}) + (F' \cap F^{\perp}).\end{aligned}$$

Thus

$$\begin{aligned}(F + F') \cap (F + F')^{\perp} = 0 &\Leftrightarrow F \cap (F')^{\perp} = 0 \text{ and } \ F' \cap F^{\perp} = 0 \\ &\Leftrightarrow F \cap (F')^{\perp} = 0, \text{ and } \ F + (F')^{\perp} = V.\end{aligned} \tag{5}$$

This shows (a)$\Leftrightarrow$(b), and similarly (a)$\Leftrightarrow$(c). Property (b) shows $\dim V = \dim F + (\dim F')^{\perp} = \dim F + \dim V - \dim F'$, hence $\dim F = \dim F'$. Given an isotropic subspace F, we find an isotropic subspace F' satisfying (c) as follows. Choose any complement W to $F^{\perp}$, so that

$$V = F^{\perp} \oplus W.$$

Thus $V = F^{\perp} + W$ and $0 = F^{\perp} \cap W$. Taking orthogonals, this is equivalent to $0 = F \cap W^{\perp}$ and $V = F + W^{\perp}$, that is

$$V = F \oplus W^{\perp}.$$

Let $S \colon W \to F \subset F^{\perp}$ be the projection along $W^{\perp}$. Then $w - S(w) \in W^{\perp}$ for all $w \in W$. The subspace

$$F' = \{w - \tfrac{1}{2}S(w)|\ w \in W\}.$$

(being the graph of a map $W \to F^{\perp}$) is again a complement to $F^{\perp}$, and since for all $w \in$

$$B(w - \tfrac{1}{2}S(w), w - \tfrac{1}{2}S(w)) = B(w, w - S(w)) + \tfrac{1}{4}B(S(w), S(w)) = 0$$

(the first term vanishes since $w - S(w) \in W^{\perp}$, the second term vanishes since $S(w) \in F$ is isotropic) it follows that F' is isotropic. $\square$

An isotropic subspace is called *maximal isotropic* if it is not properly contained in another isotropic subspace. Put differently, an isotropic subspace F is maximal isotropic if and only if it contains all $v \in F^{\perp}$ with $B(v, v) = 0$.

PROPOSITION 2.3. *Suppose F, F' are maximal isotropic. Then*

(a) *the kernel of the restriction of B to $F + F'$ equals $F \cap F'$. (In particular, $F + F'$ is quadratic if and only if $F \cap F' = 0$.)*
(b) *The images of F, F' in the quadratic vector space $(F + F')/(F \cap F')$ are maximal isotropic.*
(c) $\dim F = \dim F'$.

PROOF. Since F is maximal isotropic, it contains all isotropic vectors of $F^{\perp}$, and in particular it contains $F^{\perp} \cap F'$. Thus

$$F^{\perp} \cap F' = F \cap F'$$

Similarly $F \cap (F')^{\perp} = F \cap F'$ since F' is maximal isotropic. The calculation (5) hence shows

$$(F + F') \cap (F + F')^{\perp} = F \cap F',$$

proving (a). Let $W = (F + F')/(F \cap F')$ with the bilinear form B_W induced from B, and $\pi \colon F + F' \to W$ the quotient map. Clearly, B_W is non-degenerate, and $\pi(F), \pi(F')$ are isotropic. Hence the sum $W = \pi(F) + \pi(F')$ is a direct sum, and the two subspaces are maximal isotropic of dimension $\frac{1}{2} \dim W$. It follows that $\dim F = \dim \pi(F) + \dim(F \cap F') = \dim \pi(F') + \dim(F \cap F') = \dim F'$. □

DEFINITION 2.4. The *Witt index* of a non-degenerate symmetric bilinear form B is the dimension of a maximal isotropic subspace.

By (4), the maximal Witt index is $\frac{1}{2} \dim V$ if $\dim V$ is even, and $\frac{1}{2}(\dim V - 1)$ if $\dim V$ is odd.

3. Split bilinear forms

DEFINITION 3.1. The non-degenerate symmetric bilinear form B on an even-dimensional vector space V is called *split* if its Witt index is $\frac{1}{2} \dim V$. In this case, maximal isotropic subspaces are also called *Lagrangian subspaces*.

Equivalently, the Lagrangian subspaces are characterized by the property

$$F = F^{\perp}.$$

Split bilinear forms are easily classified:

PROPOSITION 3.2. *Let (V, B) be a quadratic vector space with a split bilinear form. Then there exists a basis $e_1, \ldots, e_k, f_1, \ldots, f_k$ of V in which the bilinear form is given as follows:*

$$B(e_i, e_j) = 0, \; B(e_i, f_j) = \delta_{ij}, \; B(f_i, f_j) = 0. \tag{6}$$

PROOF. Choose a pair of complementary Lagrangian subspaces, F, F'. Since B defines a non-degenerate pairing between F and F', it defines an isomorphism, $F' \cong F^*$. Choose a basis $e_1, \ldots, e_k$, and let $f_1, \ldots, f_k$ be the dual basis of F' under this identification. Then $B(e_i, f_j) = \delta_{ij}$ by definition of dual basis, and $B(e_i, e_j) = B(f_i, f_j) = 0$ since F, F' are Lagrangian. □

Our basis $e_1, \ldots, e_k, f_1, \ldots, f_k$ for a quadratic vector space (V, B) with split bilinear form is not orthogonal. However, it may be replaced by an orthogonal basis

$$\epsilon_i = e_i + \tfrac{1}{2} f_i, \quad \tilde{\epsilon}_i = e_i - \tfrac{1}{2} f_i.$$

In the new basis, the bilinear form reads,

$$B(\epsilon_i, \epsilon_j) = \delta_{ij}, \; B(\epsilon_i, \tilde{\epsilon}_j) = 0, \; B(\tilde{\epsilon}_i, \tilde{\epsilon}_j) = -\delta_{ij}. \tag{7}$$

Put differently, Proposition 3.2 (and its proof) say that any quadratic vector space with split bilinear form is isometric to a vector space

$$V = F^* \oplus F,$$

where the bilinear form is given by the pairing:

$$B((\mu, v), (\mu', v')) = \langle \mu', v \rangle + \langle \mu, v' \rangle.$$

The corresponding orthogonal group will be discussed in Section **??** below. At this point we will only need the following fact:

LEMMA 3.3. *Let $V = F^* \oplus F$, with the split bilinear form B given by the pairing. Then the subgroup $\mathrm{O}(V)_F \subset \mathrm{O}(V)$ fixing F pointwise consists of transformations of the form*

$$A_D \colon (\mu, v) \mapsto (\mu, v + D\mu)$$

where $D \colon F^* \to F$ is skew-adjoint: $D^* = -D$. In particular, $\mathrm{O}(V)_F \subset \mathrm{SO}(V)$.

PROOF. A linear transformation $A \in \mathrm{GL}(V)$ fixes F pointwise if and only if it is of the form

$$A(\mu, v) = (S\mu, v + D\mu)$$

for some linear maps $D \colon F^* \to F$ and $S \colon F^* \to F^*$. Suppose A is orthogonal. Then

$$0 = B(A(\mu, 0), A(0, v)) - B((\mu, 0), (0, v)) = \langle S\mu - \mu, v \rangle$$

for all $v \in F, \; \mu \in F^*$; hence $S = I$. Furthermore

$$0 = B(A(\mu, 0), A(\mu', 0)) - B((\mu, 0), (\mu', 0)) = \langle \mu, D\mu' \rangle + \langle \mu', D\mu \rangle,$$

so that $D = -D^*$. Conversely, it is straightforward to check that transformations of the form $A = A_D$ are orthogonal. □

4. E.Cartan-Dieudonné's Theorem

Throughout this Section, we assume that (V, B) is a quadratic vector space. The following simple result will be frequently used.

LEMMA 4.1. *For any $A \in \mathrm{O}(V)$, the orthogonal of the space of A-fixed vectors equals the range of $A - I$:*

$$\operatorname{ran}(A - I) = \ker(A - I)^{\perp}.$$

PROOF. For any $L \in \mathrm{End}(V)$, the transpose $L^{\top}$ relative to B satisfies $\operatorname{ran}(L) = \ker(L^{\top})^{\perp}$. We apply this to $L = A - I$, and observe that $\ker(A^{\top} - I) = \ker(A - I)$ since a vector is fixed under A if and only if it is fixed under $A^{\top} = A^{-1}$. □

DEFINITION 4.2. An orthogonal transformation $R \in \mathrm{O}(V)$ is called a *reflection* if its fixed point set $\ker(R - I)$ has codimension 1.

Equivalently, $\operatorname{ran}(R - I) = \ker(R - I)^{\perp}$ is 1-dimensional. If $v \in V$ is a non-isotropic vector, then the formula

$$R_v(w) = w - 2\frac{B(v, w)}{B(v, v)}v,$$

defines a reflection, since $\operatorname{ran}(R_v - I) = \operatorname{span}(v)$ is 1-dimensional.

PROPOSITION 4.3. *Any reflection R is of the form R_v, where the non-isotropic vector v is unique up to a non-zero scalar.*

PROOF. Suppose R is a reflection, and consider the 1-dimensional subspace $F = \operatorname{ran}(R - I)$. We claim that F is a quadratic subspace of V. Once this is established, we obtain $R = R_v$ for any non-zero $v \in F$, since R_v then acts as -1 on F and as $+1$ on $F^{\perp}$. To prove the claim, suppose on the contrary that F is not quadratic. Since $\dim F = 1$ it is then isotropic. Let F' be an isotropic subspace such that $F + F'$ is quadratic. Since R fixes $(F + F')^{\perp} \subset F^{\perp} = \ker(R - I)$, it may be regarded as a reflection of $F + F'$. This reduces the problem to the case $\dim V = 2$, with $F \subset V$ maximal isotropic and $R \in \mathrm{O}(V)_F$. As we had seen, $\mathrm{O}(V)_F$ is identified with the group of skew-symmetric maps $F^* \to F$, but for $\dim F = 1$ this group is trivial. Hence R is the identity, contradicting $\dim \operatorname{ran}(R - I) = 1$. □

Some easy properties of reflections are,

(1) $\det(R) = -1$,
(2) $R^2 = I$,
(3) if v is non-isotropic, $A R_v A^{-1} = R_{Av}$ for all $A \in \mathrm{O}(V)$,
(4) distinct reflections $R_1 \neq R_2$ commute if and only if the lines $\operatorname{ran}(R_1 - I)$ and $\operatorname{ran}(R_2 - I)$ are orthogonal.

The last Property may be seen as follows: suppose $R_1 R_2 = R_2 R_1$ and apply to $v_1 \in \operatorname{ran}(R_1 - I)$. Then $R_1(R_2 v_1) = -R_2 v_1$, which implies that $R_2 v_1$ is a multiple of v_1; in fact $R_2 v_1 = \pm v_1$ since R_2 is orthogonal. Since $R_2 v_1 = -v_1$ would imply that $R_1 = R_2$, we must have $R_2 v_1 = v_1$, or $v_1 \in \ker(R_2 - I)$.

For any $A \in \mathrm{O}(V)$, let $l(A)$ denote the smallest number l such that $A = R_1 \cdots R_l$ where $R_i \in \mathrm{O}(V)$ are reflections. We put $l(I) = 0$, and for the time being we put $l(A) = \infty$ if A cannot be written as such a product. (The Cartan-Dieudonne theorem below states that $l(A) < \infty$ always.) The following properties are easily obtained from the definition, for all $A, g, A_1, A_2 \in \mathrm{O}(V)$,

$$\begin{aligned} l(A^{-1}) &= l(A), \\ l(gAg^{-1}) &= l(A), \\ |l(A_1) - l(A_2)| \le l(A_1A_2) &\le l(A_1) + l(A_2), \\ \det(A) &= (-1)^{l(A)} \end{aligned}$$

A little less obvious is the following estimate.

PROPOSITION 4.4. *There is a lower bound*

$$\dim(\mathrm{ran}(A - I)) \le l(A)$$

for any $A \in \mathrm{O}(V)$.

PROOF. Let $n(A) = \dim(\mathrm{ran}(A - I))$. If $A_1, A_2 \in \mathrm{O}(V)$, we have

$$\ker(A_1A_2 - I) \supseteq \ker(A_1A_2 - I) \cap \ker(A_1 - I) = \ker(A_2 - I) \cap \ker(A_1 - I)$$

Taking orthogonals,

$$\mathrm{ran}(A_1A_2 - I) \subseteq \mathrm{ran}(A_2 - I) + \mathrm{ran}(A_1 - I)$$

which shows

$$n(A_1A_2) \le n(A_1) + n(A_2).$$

Thus, if $A = R_1 \cdots R_l$ is a product of $l = l(A)$ reflections, we have

$$n(A) \le n(R_1) + \ldots + n(R_l) = l(A).$$ □

The following upper bound for $l(A)$ is much more tricky:

THEOREM 4.5 (E.Cartan-Dieudonné). *Any orthogonal transformation $A \in \mathrm{O}(V)$ can be written as a product of $l(A) \le \dim V$ reflections.*

PROOF. By induction, we may assume that the Theorem is true for quadratic vector spaces of dimension $\le \dim V - 1$. We will consider three cases.

Case 1: $\ker(A - I)$ is non-isotropic. Choose any non-isotropic vector $v \in \ker(A - I)$. Then A fixes the span of v and restricts to an orthogonal transformation A_1 of $V_1 = \mathrm{span}(v)^\perp$. Using the induction hypothesis, we obtain

$$l(A) = l(A_1) \le \dim V - 1. \tag{8}$$

Case 2: $\mathrm{ran}(A - I)$ is non-isotropic. We claim:

(C) There exists a non-isotropic element $w \in V$ such that $v = (A - I)w$ is non-isotropic.

Given v, w as in (C), we may argue as follows. Since $v = (A-I)w$, and hence $(A+I)w \in \operatorname{span}(v)^\perp$, we have

$$R_v(A-I)w = -(A-I)w, \quad R_v(A+I)w = (A+I)w.$$

Adding and dividing by 2 we find $R_v A w = w$. Since w is non-isotropic, this shows that the kernel of $R_v A - I$ is non-isotropic. Equation (8) applied to the orthogonal transformation $R_v A$ shows $l(R_v A) \le \dim V - 1$. Hence $l(A) \le \dim V$. It remains to prove the claim (C). Suppose it is false, so that we have:

$(\neg C)$ The transformation $A - I$ takes the set of non-isotropic elements into the set of isotropic elements.

Let $v = (A-I)w$ be a non-isotropic element in $\operatorname{ran}(A-I)$. By $(\neg C)$ the element w is isotropic. The orthogonal space $\operatorname{span}(w)^\perp$ is non-isotropic for dimensional reasons, hence there exists a non-isotropic element w_1 with $B(w, w_1) = 0$. Then $w_1,\ w + w_1,\ w - w_1$ are all non-isotropic, and by $(\neg C)$ their images

$$v_1 = (A-I)w_1,\ v + v_1 = (A-I)(w + w_1),\ v - v_1 = (A-I)(w - w_1)$$

are isotropic. But then the polarization identity

$$Q_B(v) = \tfrac{1}{2}(Q_B(v+v_1) + Q_B(v-v_1)) - Q_B(v_1) = 0$$

shows that v is isotropic, a contradiction. This proves (C).

Case 3: Both $\ker(A-I)$ and $\operatorname{ran}(A-I)$ are isotropic. Since these two subspaces are orthogonal, it follows that they are equal, and are both Lagrangian. This reduces the problem to the case $V = F^* \oplus F$, where $F = \ker(A-I)$, that is $A \in \mathrm{O}(V)_F$. In particular $\det(A) = 1$. Let R_v be *any* reflection, then $A_1 = R_v A \in \mathrm{O}(V)$ has $\det(A_1) = -1$. Hence $\ker(A_1 - I)$ and $\operatorname{ran}(A_1 - I)$ cannot be both isotropic, and by the first two cases $l(A_1) \le \dim V = 2\dim F$. But since $\det(A_1) = -1$, $l(A_1)$ must be *odd*, hence $l(A_1) < \dim V$ and therefore $l(A) \le \dim V$. □

Remark 4.6. Our proof of Cartan-Dieudonne's theorem is a small modification of Artin's proof in [**?**]. If $\operatorname{char}(\mathbb{K}) = 2$, the statement of the Cartan-Dieudonne theorem is still true, except in some very special cases. See Chevalley [**?**, page 83].

Example 4.7. Let $\dim F = 2$, and $V = F^* \oplus F$ with bilinear form given by the pairing. Suppose $A \in \mathrm{O}(V)_F$, so that $A(\mu, v) = (\mu, v + D\mu)$ where $D\colon F^* \to F$ is skew-adjoint: $D^* = -D$. Assuming $D \ne 0$ we will show how to write A as a product of four reflections. Choose a basis e_1, e_2 of F, with dual basis f_1, f_2 of F^*, such that D has the normal form $Df_1 = e_2,\ Df_2 = -e_1$. Let $Q \in \mathrm{GL}(F)$ be the diagonal transformation,

$$Q(e_1) = 2e_1, Q(e_2) = e_2$$

and put

$$g = \begin{pmatrix} (Q^*)^{-1} & 0 \\ 0 & Q \end{pmatrix}, \quad .$$

Then g is a product of two reflections, for example $g = R'R$ where

$$R = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}, \; R' = gR.$$

On the other hand, using $QD(Q^*)^{-1} = 2D$ we see

$$gAg^{-1} = \begin{pmatrix} I & 0 \\ QD(Q^*)^{-1} & I \end{pmatrix} = \begin{pmatrix} I & 0 \\ 2D & I \end{pmatrix} = A^2,$$

or $A = gAg^{-1}A^{-1}$. Since $g = R'R$ we obtain the desired presentation of A as a product of 4 reflections:

$$A = R'R(ARA^{-1})(AR'A^{-1}).$$

5. Witt's Theorem

The following result is of fundamental importance in the theory of quadratic forms.

THEOREM 5.1 (Witt's Theorem). *Suppose $F, \tilde{F}$ are subspaces of a quadratic vector space (V, B), such that there exists an isometric isomorphism $\phi \colon F \to \tilde{F}$, i.e. $B(\phi(v), \phi(w)) = B(v, w)$ for all $v, w \in F$. Then ϕ extends to an orthogonal transformation $A \in \mathrm{O}(V)$.*

PROOF. By induction, we may assume that the Theorem is true for quadratic vector spaces of dimension $\leq \dim V - 1$. We will consider two cases.

Case 1: F is non-isotropic. Let $v \in F$ be a non-isotropic vector, and let $\tilde{v} = \phi(v)$. Then $Q_B(v) = Q_B(\tilde{v}) \neq 0$, and $v + \tilde{v}$ and $v - \tilde{v}$ are orthogonal. The polarization identity $Q_B(v) + Q_B(\tilde{V}) = \frac{1}{2}(Q_B(v + \tilde{v}) + Q_B(v - \tilde{v}))$ show that are not both isotropic; say $w = v + \tilde{v}$ is non-isotropic. The reflection R_w satisfies

$$R_w(v + \tilde{v}) = -(v + \tilde{v}), \quad R_w(v - \tilde{v}) = v - \tilde{v}.$$

Adding, and dividing by 2 we find that $R_w(v) = -\tilde{v}$. Let $Q = R_w R_v$. Then Q is an orthogonal transformation with $Q(v) = \tilde{v} = \phi(v)$.

Replacing F with $F' = Q(F)$, v with $v' = Q(v)$ and ϕ with $\phi' = \phi \circ Q^{-1}$, we may thus assume that $F \cap \tilde{F}$ contains a non-isotropic vector v such that $\phi(v) = v$. Let

$$V_1 = \mathrm{span}(v)^\perp, \quad F_1 = F \cap V_1, \quad \tilde{F}_1 = \tilde{F} \cap V_1$$

and $\phi_1 \colon F_1 \to \tilde{F}_1$ the restriction of ϕ. By induction, there exists an orthogonal transformation $A_1 \in \mathrm{O}(V_1)$ extending ϕ_1. Let $A \in \mathrm{O}(V)$ with $A(v) = v$ and $A|_{V_1} = A_1$; then A extends ϕ.

Case 2: F is isotropic. Let F' be an isotropic complement to $F^\perp$, and let $\tilde{F}'$ be an isotropic complement to $\tilde{F}^\perp$. The pairing given by B identifies

$F' \cong F^*$ and $\tilde{F}' \cong \tilde{F}^*$. The isomorphism $\phi \colon F \to \tilde{F}$ extends to an isometry $\psi \colon F \oplus F' \to \tilde{F} \oplus \tilde{F}'$, given by $(\phi^{-1})^*$ on $F' \cong F^*$. By Case 1 above, ψ extends further to an orthogonal transformation of V. $\square$

Some direct consequences are:

(1) $\mathrm{O}(V)$ acts transitively on the set of isotropic subspaces of any given dimension.
(2) If $F, \tilde{F}$ are isometric, then so are $F^\perp, \tilde{F}^\perp$. Indeed, any orthogonal extension of an isometry $\phi \colon F \to \tilde{F}$ restricts to an isometry of their orthogonals.
(3) Suppose $F \subset V$ is a subspace isometric to $\mathbb{K}^n$, with standard bilinear form $B(\epsilon_i, \epsilon_j) = \delta_{ij}$, and F is maximal relative to this property. If $F' \subset V$ is isometric to $\mathbb{K}^{n'}$, then there exists an orthogonal transformation $A \in \mathrm{O}(V)$ with $F' \subset A(F)$. In particular, the dimension of such a subspace F is an invariant of (V, B).

A subspace $W \subset V$ of a quadratic vector space is called *anisotropic* if it does not contain isotropic vectors other than 0. In particular, W is a quadratic subspace.

Proposition 5.2 (Witt decomposition). *Any quadratic vector space (V, B) admits a decomposition $V = F \oplus F' \oplus W$ where F, F' are maximal isotropic, W is anisotropic, and $W^\perp = F \oplus F'$. If $V = F_1 \oplus F_1' \oplus W_1$ is another such decomposition, then there exists $A \in \mathrm{O}(V)$ with $A(F) = F_1,\ A(F') = F_1',\ A(W) = W_1$.*

Proof. To construct such a decomposition, let F be a maximal isotropic subspace, and F' an isotropic complement to $F^\perp$. Then $F \oplus F'$ is quadratic, hence so is $W = (F \oplus F')^\perp$. Since F is maximal isotropic, the subspace W cannot contain isotropic vectors other than 0. Hence W is anisotropic. Given another such decomposition $V = F_1 \oplus F_1' \oplus W_1$, choose an isomorphism $F \cong F_1$. As we had seen (e.g. in the proof of Witt's Theorem), this extends canonically to an isometry $\phi \colon F \oplus F' \to F_1 \oplus F_1'$. Witt's Theorem gives an extension of ϕ to an orthogonal transformation $A \in \mathrm{O}(V)$. It is automatic that A takes $W = (F \oplus F')^\perp$ to $W = (F_1 \oplus F_1')^\perp$. $\square$

Example 5.3. If $\mathbb{K} = \mathbb{R}$, the bilinear form on the anisotropic part of the Witt decomposition is either positive definite (i.e. $Q_B(v) > 0$ for non-zero $v \in W$) or negative definite (i.e. $Q_B(v) < 0$ for non-zero $v \in W$). By Proposition 1.3, any quadratic vector space (V, B) over $\mathbb{R}$ is isometric to $\mathbb{R}^{n,m}$ for some n, m. The Witt decomposition shows that n, m are uniquely determined by B. Indeed $\min(n, m)$ is the Witt index of B, while the sign of $n - m$ is given by the sign of Q_B on the anisotropic part.

6. Orthogonal groups for $\mathbb{K} = \mathbb{R}, \mathbb{C}$

In this Section we discuss the structure of the orthogonal group $\mathrm{O}(V)$ for quadratic vector spaces over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.

Being a closed subgroup of $\mathrm{GL}(V)$, the orthogonal group $\mathrm{O}(V)$ is a Lie group. (If $\mathbb{K} = \mathbb{C}$ it is an algebraic Lie group since the defining equations are polynomial.) Recall that for a Lie subgroup $G \subset \mathrm{GL}(V)$, the corresponding Lie algebra $\mathfrak{g}$ is the subspace of all $\xi \in \mathrm{End}(V)$ with the property $\exp(t\xi) \in G$ for all $t \in \mathbb{K}$ (using the exponential map of matrices). We have:

Proposition 6.1. *The Lie algebra of* $\mathrm{O}(V)$ *is given by*

$$\mathfrak{o}(V) = \{A \in \mathrm{End}(V) | \quad B(Av, w) + B(v, Aw) = 0 \text{ for all } v, w \in V\},$$

with bracket given by commutator.

Proof. Suppose $A \in \mathfrak{o}(V)$, so that $\exp(tA) \in \mathrm{O}(V)$ for all t. Taking the t-derivative of $B(\exp(tA)v, \exp(tA)w) = B(v, w)$ we obtain $B(Av, w) + B(v, Aw) = 0$ for all $v, w \in V$. Conversely, given $A \in \mathfrak{gl}(V)$ with $B(Av, w) + B(v, Aw) = 0$ for all $v, w \in V$ we have

$$\begin{aligned}
B(\exp(tA)v, \exp(tA)w) &= \sum_{k,l=0}^{\infty} \frac{t^{k+l}}{k!l!} B(A^k v, A^l w) \\
&= \sum_{k=0}^{\infty} \sum_{i=0}^{k} \frac{t^k}{i!(k-i)!} B(A^i v, A^{k-i} w) \\
&= \sum_{k=0}^{\infty} \frac{t^k}{k!} \sum_{i=0}^{k} \binom{k}{i} B(A^i v, A^{k-i} w) \\
&= \sum_{k=0}^{\infty} \frac{t^k}{k!} B(v, A^k w) \sum_{i=0}^{k} (-1)^i \binom{k}{i} \\
&= B(v, w)
\end{aligned}$$

since $\sum_{i=0}^{k} (-1)^i \binom{k}{i} = \delta_{k,0}$. ∎

Thus $A \in \mathfrak{o}(V)$ if and only if $B^\flat \circ A \colon V \to V^*$ is a skew-adjoint map. In particular

$$\dim_{\mathbb{K}} \mathfrak{o}(V) = N(N-1)/2$$

where $N = \dim V$.

Let us now first discuss the case $\mathbb{K} = \mathbb{R}$. We have shown that any quadratic vector space (V, B) over $\mathbb{R}$ is isometric to $\mathbb{R}^{n,m}$, for unique n, m. The corresponding orthogonal group will be denoted $\mathrm{O}(n, m)$, the special orthogonal group $\mathrm{SO}(n, m)$, and its identity component $\mathrm{SO}_0(n, m)$. The dimension of $\mathrm{O}(n, m)$ coincides with the dimension of its Lie algebra $\mathfrak{o}(n, m)$, $N(N-1)/2$ where $N = n + m$. If $m = 0$ we will write $\mathrm{O}(n) = \mathrm{O}(n, 0)$ and $\mathrm{SO}(n) = \mathrm{SO}(n, 0)$. These groups are compact, since they are closed subsets of the unit ball in $\mathrm{Mat}(n, \mathbb{R})$.

Lemma 6.2. *The groups* $\mathrm{SO}(n)$ *are connected for all* $n \geq 1$*, and have fundamental group* $\pi_1(\mathrm{SO}(n)) = \mathbb{Z}_2$ *for* $n \geq 3$.

PROOF. The defining action of $\mathrm{SO}(n)$ on $\mathbb{R}^n$ restricts to a transitive action on the unit sphere S^{n-1}, with stabilizer at $(0,\ldots,0,1)$ equal to $\mathrm{SO}(n-1)$. Hence, for $n \geq 2$ the Lie group $\mathrm{SO}(n)$ is the total space of a principal fiber bundle over S^{n-1}, with fiber $\mathrm{SO}(n-1)$. This shows by induction that $\mathrm{SO}(n)$ is connected. The long exact sequence of homotopy groups

$$\cdots \to \pi_2(S^{n-1}) \to \pi_1(\mathrm{SO}(n-1)) \to \pi_1(\mathrm{SO}(n)) \to \pi_1(S^{n-1})$$

shows furthermore that the map $\pi_1(\mathrm{SO}(n-1)) \to \pi_1(\mathrm{SO}(n))$ is an isomorphism for $n > 3$ (since $\pi_2(S^{n-1}) = 0$ in that case). But $\pi_1(\mathrm{SO}(3)) = \mathbb{Z}_2$, since $\mathrm{SO}(3)$ is diffeomorphic to $\mathbb{R}P(3) = S^3/\mathbb{Z}_2$ (see below). □

The groups $\mathrm{SO}(3)$ and $\mathrm{SO}(4)$ have a well-known relation with the group $\mathrm{SU}(2)$ of complex 2×2-matrices X satisfying $X^\dagger = X^{-1}$ and $\det X = 1$. Recall that the center of $\mathrm{SU}(2)$ is $\mathbb{Z}_2 = \{+I, -I\}$.

PROPOSITION 6.3. *There are isomorphisms of Lie groups,*

$$\mathrm{SO}(3) = \mathrm{SU}(2)/\mathbb{Z}_2, \quad \mathrm{SO}(4) = (\mathrm{SU}(2) \times \mathrm{SU}(2))/\mathbb{Z}_2$$

where in the second equality the quotient is by the diagonal subgroup $\mathbb{Z}_2 \subset \mathbb{Z}_2 \times \mathbb{Z}_2$.

PROOF. Consider the algebra of quaternions $\mathbb{H} \cong \mathbb{C}^2 \cong \mathbb{R}^4$,

$$\mathbb{H} = \left\{ X = \begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix},\ z, w \in \mathbb{C} \right\}.$$

For any $X \in \mathbb{H}$ let $||X|| = (|z|^2 + |w|^2)^{\frac{1}{2}}$. Note that $X^\dagger X = X X^\dagger = ||X||^2\, I$ for all $X \in \mathbb{H}$. Define a symmetric $\mathbb{R}$-bilinear form on $\mathbb{H}$ by

$$B(X_1, X_2) = \tfrac{1}{2}\,\mathrm{tr}(X_1^\dagger X_2).$$

The identification $\mathbb{H} \cong \mathbb{R}^4$ takes this to the standard bilinear form on $\mathbb{R}^4$ since $B(X,X) = \frac{1}{2}||X||^2\,\mathrm{tr}(I) = ||X||^2$. The unit sphere $S^3 \subset \mathbb{H}$, characterized by $||X||^2 = 1$ is the group $\mathrm{SU}(2) = \{X|\ X^\dagger = X^{-1},\ \det(X) = 1\}$. Define an action of $\mathrm{SU}(2) \times \mathrm{SU}(2)$ on $\mathbb{H}$ by

$$(X_1, X_2) \cdot X = X_1 X X_2^{-1}.$$

This action preserves the bilinear form on $\mathbb{H} \cong \mathbb{R}^4$, and hence defines a homomorphism $\mathrm{SU}(2) \times \mathrm{SU}(2) \to \mathrm{SO}(4)$. The kernel of this homomorphism is the finite subgroup $\{\pm(I,I)\} \cong \mathbb{Z}_2$. (Indeed, $X_1 X X_2^{-1} = X$ for all X implies in particular $X_1 = X X_2 X^{-1}$ for all invertible X. But this is only possible if $X_1 = X_2 = \pm I$.) Since $\dim \mathrm{SO}(4) = 6 = 2 \dim \mathrm{SU}(2)$, and since $\mathrm{SO}(4)$ is connected, this homomorphism must be onto. Thus $\mathrm{SO}(4) = (\mathrm{SU}(2) \times \mathrm{SU}(2))/\{\pm(I,I)\}$.

Similarly, identify $\mathbb{R}^3 \cong \{X \in \mathbb{H}|\ \mathrm{tr}(X) = 0\} = \mathrm{span}(I)^\perp$. The conjugation action of $\mathrm{SU}(2)$ on $\mathbb{H}$ preserves this subspace; hence we obtain a group homomorphism $\mathrm{SU}(2) \to \mathrm{SO}(3)$. The kernel of this homomorphism is $\mathbb{Z}_2 \cong \{\pm I\} \subset \mathrm{SU}(2)$. Since $\mathrm{SO}(3)$ is connected and $\dim \mathrm{SO}(3) = 3 = \dim \mathrm{SU}(2)$, it follows that $\mathrm{SO}(3) = \mathrm{SU}(2)/\{\pm I\}$. □

To study the more general groups $\mathrm{SO}(n,m)$ and $\mathrm{O}(n,m)$, we recall the polar decomposition of matrices. Let

$$\mathrm{Sym}(k) = \{A|\ A^\top = A\} \subset \mathfrak{gl}(k,\mathbb{R})$$

be the space of real symmetric $k \times k$-matrices, and $\mathrm{Sym}^+(k)$ its subspace of positive definite matrices. As is well-known, the exponential map for matrices restricts to a diffeomorphism,

$$\exp\colon \mathrm{Sym}(k) \to \mathrm{Sym}^+(k),$$

with inverse $\log\colon \mathrm{Sym}^+(k) \to \mathrm{Sym}(k)$. Furthermore, the map

$$\mathrm{O}(k) \times \mathrm{Sym}(k) \to \mathrm{GL}(k,\mathbb{R}),\ (O,X) \mapsto Oe^X$$

is a diffeomorphism. The inverse map

$$\mathrm{GL}(k,\mathbb{R}) \to \mathrm{O}(k) \times \mathrm{Sym}(k),\ \mapsto (A|A|^{-1}, \log|A|),$$

where $|A| = (A^\top A)^{1/2}$, is called the *polar decomposition* for $\mathrm{GL}(k,\mathbb{R})$. We will need the following simple observation:

LEMMA 6.4. *Suppose $X \in \mathrm{Sym}(k)$ is non-zero. Then the closed subgroup of $\mathrm{GL}(k,\mathbb{R})$ generated by e^X is non-compact.*

PROOF. Replacing X with $-X$ if necessary, we may assume $||e^X|| > 1$. But then $||e^{nX}|| = ||e^X||^n$ goes to ∞ for $n \to \infty$. $\square$

This shows that $\mathrm{O}(k)$ is a maximal compact subgroup of $\mathrm{GL}(k,\mathbb{R})$. The polar decomposition for $\mathrm{GL}(k,\mathbb{R})$ restricts to a polar decomposition for any closed subgroup G that is invariant under the involution $A \mapsto A^\top$. Let

$$K = G \cap \mathrm{O}(k,\mathbb{R}),\ P = G \cap \mathrm{Sym}^+(k),\ \mathfrak{p} = \mathfrak{g} \cap \mathrm{Sym}(k).$$

The diffeomorphism $\exp\colon \mathrm{Sym}(k) \to \mathrm{Sym}^+(k)$ restricts to a diffeomorphism $\exp\colon \mathfrak{p} \to P$, with inverse the restriction of log. Hence the polar decomposition for $\mathrm{GL}(k,\mathbb{R})$ restricts to a diffeomorphism

$$K \times \mathfrak{p} \to G$$

whose inverse is called the polar decomposition of G. (It is a special case of a *Cartan decomposition.*) Using Lemma 6.4, we see that K is a maximal compact subgroup of G. Since $\mathfrak{p}$ is just a vector space, K is a deformation retract of G.

We will now apply these considerations to $G = \mathrm{O}(n,m)$. Let B_0 be the standard bilinear form on $\mathbb{R}^{n+m}$, and define the endomorphism τ by

$$B(v,w) = B_0(\tau v, w).$$

Thus τ acts as the identity on $\mathbb{R}^n \oplus 0$ and as minus the identity $0 \oplus \mathbb{R}^m$, and an endomorphism of $\mathbb{R}^{n+m}$ commutes with τ if and only if it preserves the direct sum decomposition $\mathbb{R}^{n+m} = \mathbb{R}^n \oplus \mathbb{R}^m$. A matrix $A \in \mathrm{Mat}(n+m,\mathbb{R})$ lies in $\mathrm{O}(n,m)$ if and only if $A^\top \tau A = \tau$, where $\top$ denotes as before the usual transpose of matrices, i.e. the transpose relative to B_0 (not relative to B). Similarly $X \in \mathfrak{o}(n,m)$ if and only if $X^\top \tau + \tau X = 0$.

Remark 6.5. In block form we have

$$\tau=\begin{pmatrix} I_n & 0\\ 0 & -I_m\end{pmatrix}$$

For $A\in\mathrm{Mat}(n+m,\mathbb{R})$ in block form

$$A=\begin{pmatrix} a & b\\ c & d\end{pmatrix} \tag{9}$$

we have $A\in\mathrm{O}(n,m)$ if and only if

$$a^\top a=I+c^\top c,\quad d^\top d=I+b^\top b,\quad a^\top b=c^\top d. \tag{10}$$

Similarly, for $X\in\mathrm{Mat}(n+m,\mathbb{R})$, written in block form

$$X=\begin{pmatrix} \alpha & \beta\\ \gamma & \delta\end{pmatrix} \tag{11}$$

we have $X\in\mathfrak{o}(n,m)$ if and only if

$$\alpha^\top=-\alpha,\ \beta^\top=\gamma,\ \delta^\top=-\delta. \tag{12}$$

Since $\mathrm{O}(n,m)$ is invariant under $A\mapsto A^\top$, (and likewise for the special orthogonal group and its identity component) the polar decomposition applies. We find:

Proposition 6.6. *Relative to the polar decomposition of* $\mathrm{GL}(n+m,\mathbb{R})$, *the maximal subgroups of*

$$G=\mathrm{O}(n,m),\ \mathrm{SO}(n,m),\ \mathrm{SO}_0(n,m),$$

are, respectively,

$$K=\mathrm{O}(n)\times\mathrm{O}(m),\ \mathrm{S}(\mathrm{O}(n)\times\mathrm{O}(m)),\ \mathrm{SO}(n)\times\mathrm{SO}(m).$$

(Here $\mathrm{S}(\mathrm{O}(n)\times\mathrm{O}(m))$ *are elements of* $(\mathrm{O}(n)\times\mathrm{O}(m))$ *of determinant* 1.*) In all of these cases, the space* $\mathfrak{p}$ *in the Cartan decomposition is given by matrices of the form*

$$\mathfrak{p}=\left\{\begin{pmatrix} 0 & x\\ x^\top & 0\end{pmatrix}\right\}$$

where x *is an arbitrary* $n\times m$*-matrix.*

Proof. We start with $G=\mathrm{O}(n,m)$. Elements in $K=G\cap\mathrm{O}(n+m)$ are characterized by $A^\top\tau A=\tau$ and $A^\top A=I$. The two conditions give $A\tau=\tau A$, so that A is a block diagonal element of $\mathrm{O}(n+m)$. Hence $A\in\mathrm{O}(n)\times\mathrm{O}(m)\subset\mathrm{O}(n,m)$. This shows $K=\mathrm{O}(n)\times\mathrm{O}(m)$. Elements $X\in\mathfrak{p}=\mathfrak{o}(n,m)\cap\mathrm{Sym}(n+m)$ satisfy $X^\top\tau+\tau X=0$ and $X^\top=X$, hence they are symmetric block off-diagonal matrices. This proves our characterization of $\mathfrak{p}$, and proves the polar decomposition for $\mathrm{O}(n,m)$. The polar decompositions for $\mathrm{SO}(n,m)$ is an immediate consequence, and the polar decomposition for $\mathrm{SO}_0(n,m)$ follows since $\mathrm{SO}(n)\times\mathrm{SO}(m)$ is the identity component of $\mathrm{S}(\mathrm{O}(n)\times\mathrm{O}(m))$. $\square$

COROLLARY 6.7. *Unless $n = 0$ or $m = 0$ the group $\mathrm{O}(n,m)$ has four connected components and $\mathrm{SO}(n,m)$ has two connected components.*

We next describe the space $P = \exp(\mathfrak{p})$.

PROPOSITION 6.8. *The space $P = \exp(\mathfrak{p}) \subset G$ consists of matrices*

$$P = \left\{ \begin{pmatrix} (I + bb^\top)^{1/2} & b \\ b^\top & (I + b^\top b)^{1/2} \end{pmatrix} \right\}$$

where b ranges over all $n \times m$-matrices. In fact,

$$\log \begin{pmatrix} (I + bb^\top)^{1/2} & b \\ b^\top & (I + b^\top b)^{1/2} \end{pmatrix} = \begin{pmatrix} 0 & x \\ x^\top & 0 \end{pmatrix}$$

where x and b are related as follows,

$$b = \frac{\sinh(xx^\top)}{xx^\top} x, \quad x = \frac{\operatorname{arsinh}((bb^\top)^{1/2})}{(bb^\top)^{1/2}} b. \tag{13}$$

Note that $xx^\top$ (resp. $bb^\top$) need not be invertible. The quotient $\frac{\sinh(xx^\top)}{xx^\top}$ is to be interpreted as $f(xx^\top)$ where $f(z)$ is the entire holomorphic function $\frac{\sinh z}{z}$, and $f(xx^\top)$ is given in terms of the spectral theorem or equivalently in terms of the power series expansion of f.

PROOF. Let $X = \begin{pmatrix} 0 & x \\ x^\top & 0 \end{pmatrix}$. By induction on k,

$$X^{2k} = \begin{pmatrix} (xx^\top)^k & 0 \\ 0 & (x^\top x)^k \end{pmatrix}, \ X^{2k+1} = \begin{pmatrix} 0 & (xx^\top)^k x \\ x(x^\top x)^k & 0 \end{pmatrix}.$$

This gives

$$\exp(X) = \begin{pmatrix} \cosh(xx^\top) & \frac{\sinh(xx^\top)}{xx^\top} x \\ x \frac{\sinh(x^\top x)}{x^\top x} & \cosh(x^\top x) \end{pmatrix},$$

which is exactly the form of elements in P with $b = \frac{\sinh(xx^\top)}{xx^\top} x$. The equation $\cosh(xx^\top) = (1 + bb^\top)^{1/2}$ gives $\sinh(xx^\top) = (bb^\top)^{1/2}$. Plugging this into the formula for b, we obtain the second equation in (13). $\square$

For later reference, we mention one more simple fact about the orthogonal and special orthogonal groups. Let $\mathbb{Z}_2$ be the center of $\mathrm{GL}(n+m, \mathbb{R})$ consisting of $\pm I$.

PROPOSITION 6.9. *For all n, m, the center of the group $\mathrm{O}(n,m)$ is $\mathbb{Z}_2$. Except in the cases $(n,m) = (0,2), (2,0)$, the center of $\mathrm{SO}(n,m)$ is $\mathbb{Z}_2$ if $-I$ lies in $\mathrm{SO}(n,m)$, and is trivial otherwise. The statement for the identity component is similar.*

The proof is left as an exercise. (Note that the elements of the center of G commute in particular with the diagonal elements of G. In the case of hand, one uses this fact to argue that the central elements are themselves diagonal, and finally that they are multiples of the identity.)

The discussion above carries over to $\mathbb{K} = \mathbb{C}$, with only minor modifications. It is enough to consider the case $V = \mathbb{C}^n$, with the standard symmetric bilinear form. Again, our starting point is the polar decomposition, but now for complex matrices. Let $\mathrm{Herm}(n) = \{A|\ A^\dagger = A\}$ be the space of Hermitian $n \times n$ matrices, and $\mathrm{Herm}^+(n)$ the subset of positive definite matrices. The exponential map gives a diffeomorphism

$$\mathrm{Herm}(n) \to \mathrm{Herm}^+(n),\ X \mapsto e^X.$$

This is used to show that the map

$$\mathrm{U}(n) \times \mathrm{Herm}(n) \to \mathrm{GL}(n, \mathbb{C}), \quad (U, X) \mapsto Ue^X$$

is a diffeomorphism; the inverse map takes A to (Ae^{-X}, X) with $X = \frac{1}{2}\log(A^\dagger A)$. The polar decomposition of $\mathrm{GL}(n, \mathbb{C})$ gives rise to polar decompositions of any closed subgroup $G \subset \mathrm{GL}(n, \mathbb{C})$ that is invariant under the involution $\dagger$. In particular, this applies to $\mathrm{O}(n, \mathbb{C})$ and $\mathrm{SO}(n, \mathbb{C})$. Indeed, if $A \in \mathrm{O}(n, \mathbb{C})$, the matrix $A^\dagger A$ lies in $\mathrm{O}(n, \mathbb{C}) \cap \mathrm{Herm}(n)$, and hence its logarithm $X = \frac{1}{2}\log(A^\dagger A)$ lies in $\mathfrak{o}(n, \mathbb{C}) \cap \mathrm{Herm}(n)$. But clearly,

$$\mathrm{O}(n, \mathbb{C}) \cap \mathrm{U}(n) = \mathrm{O}(n, \mathbb{R}),$$
$$\mathrm{SO}(n, \mathbb{C}) \cap \mathrm{U}(n) = \mathrm{SO}(n, \mathbb{R})$$

while

$$\mathfrak{o}(n, \mathbb{C}) \cap \mathrm{Herm}(n) = \sqrt{-1}\mathfrak{o}(n, \mathbb{R}).$$

Hence, the maps $(U, X) \mapsto Ue^X$ restrict to polar decompositions

$$\mathrm{O}(n, \mathbb{R}) \times \sqrt{-1}\mathfrak{o}(n, \mathbb{R}) \to \mathrm{O}(n, \mathbb{C}),$$
$$\mathrm{SO}(n, \mathbb{R}) \times \sqrt{-1}\mathfrak{o}(n, \mathbb{R}) \to \mathrm{SO}(n, \mathbb{C}),$$

which shows that the algebraic topology of the complex orthogonal and special orthogonal group coincides with that of its real counterparts. Arguing as in the real case, the center of $\mathrm{O}(n, \mathbb{C})$ is given by $\{+I, -I\}$ while the center of $\mathrm{SO}(n, \mathbb{C})$ is trivial for n odd and $\{+I, -I\}$ for n even, provided $n \geq 3$.

7. Lagrangian Grassmannians

If (V, B) is a quadratic vector space with split bilinear form, denote by $\mathrm{Lag}(V)$ the set of Lagrangian subspaces. Recall that any such V is isomorphic to $\mathbb{K}^{n,n}$ where $\dim V = 2n$. For $\mathbb{K} = \mathbb{R}$ we have the following result.

THEOREM 7.1. *Let $V = \mathbb{R}^{n,n}$ with the standard basis satisfying* (7)*. Then the maximal compact subgroup $\mathrm{O}(n) \times \mathrm{O}(n)$ of $\mathrm{O}(n,n)$ acts transitively on the space $\mathrm{Lag}(\mathbb{R}^{n,n})$ of Lagrangian subspaces, with stabilizer at*

$$L_0 = \mathrm{span}\{\epsilon_1 + \tilde{\epsilon}_1, \ldots, \epsilon_n + \tilde{\epsilon}_n\} \tag{14}$$

the diagonal subgroup $\mathrm{O}(n)_\Delta$. Thus

$$\mathrm{Lag}(\mathbb{R}^{n,n}) \cong \mathrm{O}(n) \times \mathrm{O}(n) / \mathrm{O}(n)_\Delta \cong \mathrm{O}(n).$$

In particular, it is a compact space with two connected components.

PROOF. Let B_0 be the standard positive definite bilinear form on the vector space $\mathbb{R}^{n,n} = \mathbb{R}^{2n}$, with corresponding orthogonal group $\mathrm{O}(2n)$. Introduce an involution $\tau \in \mathrm{O}(2n)$, by

$$B(v, w) = B_0(\tau v, w).$$

That is $\tau\epsilon_i = \epsilon_i$, $\tau\tilde{\epsilon}_i = -\tilde{\epsilon}_i$. Then the maximal compact subgroup $\mathrm{O}(n) \times \mathrm{O}(n)$ consists of all those transformations $A \in \mathrm{O}(n, n)$ which commute with τ. At the same time, $\mathrm{O}(n) \times \mathrm{O}(n)$ is characterized as the orthogonal transformations in $\mathrm{O}(2n)$ commuting with τ.

The ± 1 eigenspaces $V_\pm$ of τ are both anisotropic, i.e. they do not contain any isotropic vectors. Hence, for any $L \subset \mathbb{R}^{n,n}$ is Lagrangian, then $\tau(L)$ is transverse to L:

$$L \cap \tau(L) = (L \cap V_+) \oplus (L \cap V_-) = 0.$$

For any L, we may choose a basis $v_1, \ldots, v_n$ that is orthonormal relative to B_0. Then $v_1, \ldots, v_n, \tau(v_1), \ldots, \tau(v_n)$ is a B_0-orthonormal basis of $\mathbb{R}^{n,n}$. If L' is another Lagrangian subspace, with B_0-orthonormal basis $v_1', \ldots, v_n'$, then the orthogonal transformation $A \in \mathrm{O}(2n)$ given by

$$Av_i = v_i', \quad A\tau(v_i) = \tau(v_i'), \ i = 1, \ldots, n$$

commutes with τ, hence $A \in \mathrm{O}(n) \times \mathrm{O}(n)$. This shows that $\mathrm{O}(n) \times \mathrm{O}(n)$ acts transitively on $\mathrm{Lag}(\mathbb{R}^{n,n})$. For the Lagrangian subspace (14), with $v_i = \frac{1}{\sqrt{2}}(\epsilon_i + \tilde{\epsilon}_i)$, the stabilizer of L_0 under the action of $\mathrm{O}(n) \times \mathrm{O}(n)$ consists of those transformations $A \in \mathrm{O}(n) \times \mathrm{O}(n)$ for which $v_1', \ldots, v_n'$ is again a B_0-orthonormal basis of L_0. But this is just the diagonal subgroup $\mathrm{O}(n)_\Delta \subset \mathrm{O}(n) \times \mathrm{O}(n)$. Finally, since the multiplication map

$$(\mathrm{O}(n) \times \{1\}) \times \mathrm{O}(n)_\Delta \to \mathrm{O}(n) \times \mathrm{O}(n)$$

is a bijection, the quotient is just $\mathrm{O}(n)$. □

Theorem 7.1 does not, as it stands, hold for other fields $\mathbb{K}$. Indeed, for $V = \mathbb{K}^{n,n}$ the group $\mathrm{O}(n, \mathbb{K}) \times \mathrm{O}(n, \mathbb{K})$ takes (14) to a Lagrangian subspace transverse to $V_+ = \mathbb{K}^n \oplus 0$, $V_- = 0 \oplus \mathbb{K}^n$, and any Lagrangian subspace transverse to V_+, V_- is of this form. However, there may be other Lagrangian subspaces: E.g. if $\mathbb{K} = \mathbb{C}$ and $n = 2$, the span of $\epsilon_1 + \sqrt{-1}\epsilon_2$ and $\tilde{\epsilon}_1 + \sqrt{-1}\tilde{\epsilon}_2$ is a Lagrangian subspace that is not transverse to $V_\pm$. Nonetheless, there is a good description of the space Lag in the complex case $\mathbb{K} = \mathbb{C}$.

THEOREM 7.2. *Let* $V = \mathbb{C}^{2m}$ *with the standard bilinear form. The action of the maximal compact subgroup* $\mathrm{O}(2m) \subset \mathrm{O}(2m, \mathbb{C})$ *on* $\mathrm{Lag}(V)$ *is transitive, with stabilizer at the Lagrangian subspace*

$$L_0 = \mathrm{span}\{\epsilon_1 - \sqrt{-1}\epsilon_{m+1}, \ldots, \epsilon_m - \sqrt{-1}\epsilon_{2m}\}$$

the unitary group $\mathrm{U}(m)$. *That is,* $\mathrm{Lag}(\mathbb{C}^{2m})$ *is a homogeneous space*

$$\mathrm{Lag}(\mathbb{C}^{2m}) \cong \mathrm{O}(2m)/\mathrm{U}(m);$$

In particular, it is a compact and has two connected components.

PROOF. Let $\tau\colon v \mapsto \overline{v}$ be complex conjugation in $V = \mathbb{C}^{2m}$, so that $\langle v, w\rangle = B(\overline{v}, w)$ is the standard Hermitian inner product on $\mathbb{C}^{2m}$. Then $\tau \in \mathrm{O}(4m, \mathbb{R})$ is a real orthogonal transformation of $\mathbb{C}^{2m} \cong \mathbb{R}^{4m}$. Let $V_\pm \subset \mathbb{C}^{2m}$ be the ± 1 eigenspaces of τ; thus $V_+ = \mathbb{R}^{2m}$ (viewed as a real subspace) and $V_- = \sqrt{-1}\mathbb{R}^{2m}$. Note that $V_\pm$ do not contain non-zero isotropic vectors. Hence, for $L \in \mathrm{Lag}(V)$ we have $L \cap \tau(L) \neq 0$, and hence $V = L \oplus \tau(L)$ is a direct sum. Let $v_1, \ldots, v_n$ be a basis of L that is orthonormal for the Hermitian inner product. Then $v_1, \ldots, v_n, \overline{v}_1, \ldots, \overline{v}_n$ is an orthonormal basis of V. Given another Lagrangian subspace L' with orthonormal basis $v_1,', \ldots, v_n'$, the unitary transformation $A \in \mathrm{U}(2m)$ with $Av_i = v_i'$ and $A\overline{v}_i = \overline{v}_i'$ commutes with τ, hence it is contained in $\mathrm{O}(2m) \subset \mathrm{U}(2m)$. This shows that $\mathrm{O}(2m)$ acts transitively. Note that any unitary transformation $U\colon L \to L'$ between Lagrangian subspaces extends uniquely to an element A of the maximal compact subgroup $\mathrm{O}(2m) \subset \mathrm{O}(2m, \mathbb{C})$, where $Av = Uv$ for $v \in L$ and $A\tau v = \tau U v$ for $\tau v \in \tau L$. In particular, the stabilizer in $\mathrm{O}(2m)$ of L_0 is the unitary group $U(L_0) \cong U(n)$. $\square$

REMARK 7.3. The orbit of L_0 under $\mathrm{O}(m, \mathbb{C}) \times \mathrm{O}(m, \mathbb{C})$ is open and dense in $\mathrm{Lag}(\mathbb{C}^{2m})$, and as in the real case is identified with $\mathrm{O}(m, \mathbb{C})$. Thus, $\mathrm{Lag}(\mathbb{C}^{2m})$ is a smooth compactification of the complex Lie group $\mathrm{O}(m, \mathbb{C})$.

Theorem 7.2 has a well-known geometric interpretation. View $\mathbb{C}^{2m}$ as the complexification of $\mathbb{R}^{2m}$. Recall that an *orthogonal complex structure* on $\mathbb{R}^{2m}$ is an automorphism $J \in \mathrm{O}(2m)$ with $J^2 = -I$. We denote by J_0 the standard complex structure.

Let $\mathcal{J}(2m)$ denote the space of all orthogonal complex structures. It carries a transitive action of $\mathrm{O}(2m)$, with stabilizer at J_0 equal to $\mathrm{U}(m)$. Hence the space of orthogonal complex structures is identified with the complex Lagrangian Grassmannian:

$$\mathcal{J}(2m) = \mathrm{O}(2m)/\,\mathrm{U}(m) = \mathrm{Lag}(\mathbb{C}^{2m}).$$

Explicitly, this correspondence takes $J \in \mathcal{J}(2m)$ to its $+\sqrt{-1}$ eigenspace

$$L = \ker(J - \sqrt{-1}I).$$

This has complex dimension m since $\mathbb{C}^{2m} = L \oplus \overline{L}$, and it is isotropic since $v \in L$ implies

$$B(v, v) = B(Jv, Jv) = B(\sqrt{-1}v, \sqrt{-1}v) = -B(v, v).$$

Any Lagrangian subspace L determines J, as follows: Given $w \in \mathbb{R}^{2n}$, we may uniquely write $w = v + \overline{v}$ where $v \in L$. Define a linear map J by $Jw := -2\,\mathrm{Im}(v)$. Then $v = w - \sqrt{-1}Jw$. Since L is Lagrangian, we have

$$\begin{aligned} 0 = B(v, v) &= B(w - \sqrt{-1}Jw, w - \sqrt{-1}Jw) \\ &= B(w, w) - B(Jw, Jw) - 2\sqrt{-1}B(w, Jw), \end{aligned}$$

which shows that $J \in \mathrm{O}(2m)$ and that $B(w, Jw) = 0$ for all w. Multiplying the definition of J by $\sqrt{-1}$, we get

$$\sqrt{-1}v = \sqrt{-1}w + Jw$$

which shows that $J(Jw) = -w$. Hence J is an orthogonal complex structure.

Remark 7.4. There are parallel results in symplectic geometry, for vector spaces V with a non-degenerate *skew*-symmetric linear form ω. If $\mathbb{K} = \mathbb{R}$, any such V is identified with $\mathbb{R}^{2n} = \mathbb{C}^n$ with the standard symplectic form, $L_0 = \mathbb{R}^n \subset \mathbb{C}^n$ is a Lagrangian subspace, and the action of $\mathrm{U}(n) \subset \mathrm{Sp}(V, \omega)$ on L_0 identifies

$$\mathrm{Lag}_\omega(\mathbb{R}^{2n}) \cong \mathrm{U}(n)/\,\mathrm{O}(n)$$

For the space $\mathrm{Lag}(V)$ of complex Lagrangian Grassmannian subspaces of the complex symplectic vector space $\mathbb{C}^{2n} \cong \mathbf{H}^n$ one has

$$\mathrm{Lag}_\omega(\mathbb{C}^{2n}) \cong \mathrm{Sp}(n)/U(n)$$

where $\mathrm{Sp}(n)$ is the *compact symplectic group* (i.e. the quaternionic unitary group). See e.g. [**?**, p.67].