

Clifford Algebras and Spinors

Notes based on PMATH945: Topics in Algebra at University of Waterloo
Taught by Spiro Karigiannis in Fall 2020
Transcribed and typeset by Michael Astwood

March 14, 2023

Contents

1	Tensors and k-Forms	5
1.1	Lecture 1: Linear Functionals and the Dual Space	5
1.2	Lecture 2: Tensor Product and Dual Map	6
1.2.1	Bilinear Forms and Tensor Products of Forms	6
1.2.2	Tensor Product of Arbitrary Vector Spaces	8
1.2.3	The Tensor Algebra	12
1.2.4	Linear Maps as Tensors, and the Dual Map	14
1.3	Lecture 3: The Exterior Algebra	16
1.3.1	Antisymmetric Forms	16
1.3.2	The Exterior Algebra	20
1.4	Lecture 4: Dual Maps, Annihilators, and Non-degenerate Bilinear Forms	22
1.4.1	Annihilators	22
1.4.2	Bilinear Forms on a Vector Space V	23
2	Classification of Bilinear Forms	27
2.1	Lecture 5: Restrictions of Bilinear Forms	27
2.1.1	Restriction to Subspace	27
2.1.2	Restriction to Quotient	29
2.1.3	Classification of Bilinear Forms: Symmetric Bilinear Forms Part 1	30
2.2	Lecture 6: Definiteness, Isometries and Adjoints	32
2.2.1	Classification of Bilinear Forms: Symmetric Bilinear Forms Part 2	32
2.2.2	Isometries	34
2.2.3	Adjoints	35
2.3	Lecture 7: Reflections, Hyperbolic Planes, and Cancellation	36
2.3.1	Adjoint Basis	36
2.3.2	Classification of Bilinear Forms: Symmetric Bilinear Forms Part 3	36
2.3.3	Hyperbolic Planes	39
2.3.4	Witt's Cancellation Theorem	42
2.4	Lecture 8: Witt Extension Theorem, Maximally Isotropic Subspaces, and the Cartan-Dieudonné Theorem	43
2.4.1	Witt Extension Theorem	43
2.4.2	Maximally Isotropic Subspaces	44
2.4.3	Classification of Bilinear Forms: Symmetric Bilinear Forms Part 4, Cartan-Dieudonné theorem	45
2.5	Lecture 9: Skew-Symmetric Bilinear Forms and the Pfaffian	48
2.5.1	Classification of Bilinear Forms: Skew-Symmetric Bilinear Forms	48
2.5.2	The Pfaffian	49
3	Digression: The Hodge Dual	52
3.1	Lecture 10: Poincaré Duality	52
3.1.1	Orientations	52

3.1.2	Volume Forms	53
3.1.3	Poincaré Duality and the Hodge Star	55
3.1.4	Hodge Star on Orthogonal Direct Sums	56
3.1.5	Square of the Hodge Star	57
4	Relations Between Real and Complex Vector Spaces	60
4.1	Lecture 11: Relations between Complex and Real Vector Spaces. Part I	60
4.1.1	Antilinear Maps and Conjugation	60
4.1.2	Constructing Real Vector Spaces from Complex Vector Spaces	61
4.2	Lecture 12: Relations between Complex and Real Vector Spaces. Part II	63
4.2.1	Constructing Complex Vector Spaces from Real Vector Spaces	63
4.2.2	Non-existence of Field Structure on \mathbb{R}^3	64
5	Quaternions	67
5.1	Lecture 13: The Quaternions	67
5.1.1	The Quaternions	67
5.1.2	Conjugation by Quaternions	71
5.2	Lecture 14: Quaternions, Rotations, and Double Covers	75
5.2.1	Quaternions as Rotations	75
5.2.2	Exact Sequences and Double Covers	76
6	Clifford Algebras	78
6.1	Lecture 15: Clifford Algebras	78
6.1.1	Interior Product Revisited	78
6.1.2	Quotient Algebras	81
6.1.3	The Clifford Algebra	82
6.1.4	Universal Properties	83
6.1.5	Clifford Involutions	88
6.2	Lecture 16: Symmetries and Automorphisms of Clifford Algebras	89
6.2.1	Clifford Involutions (Continued)	89
6.2.2	Clifford Algebras as Subspaces of Clifford Algebras	91
6.2.3	Center and Twisted Center of Clifford Algebras	92
6.3	Lecture 17: The Spin Group and Pin Group	93
6.3.1	Self-Dual and Anti-Self-Dual Forms	93
6.3.2	Examples of Clifford Algebras	95
6.3.3	Pin Groups and Spin Groups	96
7	Composition Algebras	97
7.1	Lecture 18: Composition Algebras	97
7.1.1	Real and Imaginary Parts of Composition Algebras	97
7.1.2	Conjugation in Composition Algebras	98

7.1.3	Commutators and Associators	100
7.1.4	Cayley-Dickson Construction	100
7.2	Lecture 19: Hurwitz Theorem	105
7.2.1	Classification of Composition Algebras	105
7.2.2	Hurwitz Theorem	106
7.3	Lecture 20: Cross Products and Isometries of Composition Algebras	107
7.3.1	Cross Products	107
7.3.2	Automorphisms of Composition Algebras	111
7.4	Lecture 21: The Exceptional Groups G_2 and \tilde{G}_2	112
7.4.1	Automorphisms of \mathbb{H} and $\tilde{\mathbb{H}}$, Continued	112
7.4.2	Automorphisms of \mathbb{O} and $\tilde{\mathbb{O}}$	113
8	Spinors	116
8.1	Lecture 22: Classification of Clifford Algebras	116
8.1.1	Classification of Clifford Algebras	116
8.1.2	Self Duality	119
8.2	Lecture 23: Pinor and Spinor Representations	120
8.2.1	Classification of Pinor and Spinor Representations	120
8.2.2	Duality and Triality	123
8.3	Lecture 24: $\text{Spin}(8)$ and Triality	126
8.3.1	Properties of $\text{Cl}(8, 0)$	126
9	Appendix I: Groups	132
9.1	Sets and Fields	132
9.2	Groups	134
9.2.1	Basic Definitions	134
9.2.2	Permutations	136
9.2.3	Group Homomorphisms	138
9.2.4	Quotient Groups and the First Isomorphism Theorem	140
9.2.5	Group Actions and Cayley's Theorem	141
9.3	Rings	142
10	Appendix II: Linear Algebra	144
10.1	Vector Spaces	144
10.1.1	Basic Definitions	144
10.1.2	Maps Between Vector Spaces	147
10.1.3	Inner Products and Eigenvectors	152
10.1.4	Quotient Spaces	155
10.1.5	Change of Coordinates	157
	References	160

Introduction

These are lecture notes based on a course I took in Fall 2020 at the University of Waterloo titled Clifford Algebras, Spinors, and Calibrations. On the first day, before beginning the main content of the course, Spiro gave us a short overview of the content of the course. I will review this here.

The first major theme of the course is to build up to the notion of spinors, which are elements of certain special vector spaces which arise as representations of Clifford Algebras. This is all motivated by the study of special structures on vector spaces, and so this course will introduce a number of new structures you may not be familiar with.

As mentioned, spinors arise as representations of a certain group. This is the group of even elements of the Clifford Algebra, which is often called the Spin group. A characteristic feature of the spin group is that there is an exact sequence of groups,

$$\{e\} \rightarrow \mathbb{Z}/2\mathbb{Z} \rightarrow \text{Spin}(n) \rightarrow \text{SO}(n) \rightarrow \{e\}$$

Meaning a collection of maps, f_1, f_2, f_3, f_4 which represent the arrows in the above diagram, so that $\ker f_2 = \text{im } f_1$, $\ker f_3 = \text{im } f_2$, and $\ker f_4 = \text{im } f_3$. From the existence of such a sequence, we will find that the spin group appears as a double covering of the special orthogonal group. This makes it extremely useful in physics applications, where certain groups of transformations must be replaced with their universal cover upon quantization. Additionally, in this course we will see that for small values of n , the group $\text{Spin}(n)$ happens to be intimately related to the so-called composition division algebras (these are the real numbers, the complex numbers, the quaternions, and the octonions).

A secondary theme of this course was that of calibrations. Unfortunately we did not have time to talk about calibrations in lectures. However, a few assignment questions dealt with them. Calibrations provide a way to generalize the Cauchy-Schwartz inequality, which we state as $-1 \leq a \cdot b \leq 1$ for all unit vectors a, b . The reformulation roughly goes as follows. We first define a map from \mathbb{R}^n to \mathbb{R} , given a unit vector a , by the formula $L_a(v) = a \cdot v$. This can be viewed as a linear functional on the space of oriented lines through the origin in \mathbb{R}^n . A natural generalization of this then considers a map α which takes oriented k -dimensional subspaces V of \mathbb{R}^n , and returns a number between -1 and 1 . When \mathbb{R}^n is provided with certain special structures, we can find a rich class of maps like this, which are called calibrations. One question of interest is the following: which k -dimensional subspaces of \mathbb{R}^n give us $\alpha(V) = \pm 1$?

These calibrations are related to the composition division algebras mentioned above, through the cross product. Also, it turns out that calibrations can even be related to $\text{Spin}(n)$, albeit in a nontrivial way.

Much of this course can be found in various parts of the following textbooks. However, none of the content of the course will come exactly from one book, except for the sections on bilinear forms which are mostly based on Basic Algebra I by Jacobson.

- [5] Chevalley, The algebraic theory of spinors, Columbia University Press, 1954.
- [6] Conway and Smith, On quaternions and octonions: their geometry, arithmetic, and symmetry, A K Peters, 2003.
- [11] Harvey, Spinors and calibrations (Perspectives in Mathematics), Academic Press, 1990.
- [9] Lawson and Michelson, Spin geometry (Princeton Mathematical Series), Princeton University Press, 1989.

- [10] Porteous, Clifford algebras and the classical groups (Cambridge Studies in Advanced Mathematics), Cambridge University Press, 1995.
- [8] Jacobson, Basic Algebra I, Dover Publications, 1974.

1 Tensors and k -Forms

1.1 Lecture 1: Linear Functionals and the Dual Space

Let us begin with the first new definition of the course:

Definition 1.1 (Dual Space). Let V be a vector space. Consider the set V^* consisting of all linear maps from V to \mathbb{F} . This set is a vector space, as whenever $T, S \in V^*$ we have $(T + S)(v) = T(v) + S(v)$. We call the vector space V^* the **dual space** of V . The elements of the dual space are called **linear functionals**, **1-forms**, or **covectors**.

Theorem 1.1. If V is a finite dimensional vector space, then $\dim V^* = \dim V$.

For a proof of this, it is easiest to use a basis. Therefore we will delay the proof until the next lemma. Note that in the infinite dimensional case, the most we can get is $\dim V \leq \dim V^*$. Also, just because V and V^* have the same dimension, there does not have to be an obvious isomorphism between them. In order to construct a nice isomorphism $f : V \rightarrow V^*$, we typically need a basis.

Theorem 1.2. Let V be a vector space with basis $B = \{e_1, \dots, e_n\}$. We will define $B^* = \{e^1, \dots, e^n\}$ as the (unique up to reordering) collection of n linear maps so that $e^i(e_j) = 0$ if $i \neq j$, and $e^i(e_i) = 1$ for all i . The set B^* is a basis for V^* .

Proof. Consider any arbitrary element of V^* . Let us call it α . Then $\alpha(v) = \alpha(v^i e_i) = v^i \alpha(e_i)$ since α is linear. We then consider the map $\tilde{\alpha} = \alpha(e_i) e^i$. Observe that for all v , we have

$$\begin{aligned} \tilde{\alpha}(v) &= v^i \tilde{\alpha}(e_i) \\ &= v^i \alpha(e_j) e^j(e_i) \\ &= v^i \alpha(e_i) \end{aligned}$$

Therefore we can see that $\tilde{\alpha}(v) = \alpha(v)$ for all v , so $\alpha = \alpha(e_i) e^i$. The numbers $\alpha_i = \alpha(e_i)$, $i = 1, \dots, n$ are the coordinates of α in the basis B^* . □

In particular, this proof shows us that a linear functional is completely determined by its' value when evaluated on each basis vector e_i .

Definition 1.2 (Canonical). Let V and W be vector spaces. Then we say V and W are **canonically isomorphic** if it is possible to construct an isomorphism between V and W without having to use a basis for V or W .

Remark. V and V^* are **not** canonically isomorphic unless we have extra information, such as a metric.

Now, since V^* is a perfectly well defined vector space on its' own, we could construct the **double dual** V^{**} as $(V^*)^*$. Since whenever V is finite dimensional we have $\dim V = \dim V^* = \dim V^{**}$, it is natural to ask whether there is a canonical isomorphism between V and V^{**} . As it turns out, there is.

Definition 1.3 (Canonical Injection). The **canonical injection** is the injective linear map $\xi : V \rightarrow V^{**}$ defined by the formula $\xi(v)(\alpha) = \alpha(v)$ for all $\alpha \in V^*, v \in V$.

If V is finite dimensional, this map is an isomorphism between V and V^{**} , and in particular it does not depend on the choice of basis for V (this is what we usually mean by a canonical linear isomorphism). If V is infinite dimensional, this map is only guaranteed to be injective, and so we sometimes call it the canonical injection (this terminology is often used in functional analysis).

Theorem 1.3. If V is a finite dimensional vector space, the canonical injection $\xi : V \rightarrow V^{**}$ is an isomorphism.

Proof. We first show that ξ is linear. Observe:

$$\begin{aligned}\xi(\lambda v + w)(\alpha) &= \alpha(\lambda v + w) \\ &= \lambda \alpha(v) + \alpha(w)\end{aligned}$$

as required. Next we show that the kernel is trivial. Suppose $\xi(v)(\alpha) = 0$ for all α , then $\alpha(v) = 0$ for all α , which means $v = 0$. Therefore $\ker \xi = \{0\}$. Since $\dim V = \dim V^{**}$, we conclude that ξ is an isomorphism. \square

It is important to note that this proof did not use any basis for V or V^* . This is why we call ξ "canonical".

1.2 Lecture 2: Tensor Product and Dual Map

1.2.1 Bilinear Forms and Tensor Products of Forms

In this lecture we will begin with a very useful tool in linear algebra, the tensor product. First we need to review some definitions.

Definition 1.4 (Bilinear Map). Let V and W be vector spaces over \mathbb{F} . Then a map $G : V \times W \rightarrow \mathbb{F}$ is called bilinear if for all $a, b \in \mathbb{F}$,

$$G(av_1 + bv_2, w) = aG(v_1, w) + bG(v_2, w) \text{ for all } v_1, v_2 \in V, w \in W$$

and

$$G(v, aw_1 + bw_2) = aG(v, w_1) + bG(v, w_2) \text{ for all } v \in V, w_1, w_2 \in W$$

The set of bilinear maps defines a vector space $B(V, W, \mathbb{F})$.

Definition 1.5 (Tensor Product of Forms). Consider two linear forms $\alpha \in V^*$ and $\beta \in W^*$. We define a bilinear map from α and β , called the tensor product of α and β , so that the following identity holds for all $v \in V, w \in W$:

$$\alpha \otimes \beta(v, w) = \alpha(v)\beta(w)$$

Observe that,

$$\begin{aligned} \alpha \otimes \beta(av_1 + bv_2, w) &= \alpha(av_1 + bv_2)\beta(w) \\ &= (a\alpha(v_1) + b\alpha(v_2))\beta(w) \\ &= a\alpha(v_1)\beta(w) + b\alpha(v_2)\beta(w) \\ &= a(\alpha \otimes \beta)(v_1, w) + b(\alpha \otimes \beta)(v_2, w) \end{aligned}$$

So $\alpha \otimes \beta$ is linear in the first input. A similar proof holds for the second input. So we have verified that $\alpha \otimes \beta$ is indeed a bilinear mapping. If we consider all possible finite linear combinations of all these bilinear maps we get the tensor product space.

Definition 1.6 (Tensor Product of Dual Vector Spaces). Define $V^* \otimes W^* = \text{span}\{\alpha \otimes \beta : \alpha \in V^*, \beta \in W^*\}$. This is called the tensor product of the two vector spaces V^* and W^* .

Note that not all elements of $V^* \otimes W^*$ are of the form $\alpha \otimes \beta$. In general, an element is of the form $\omega = \sum_{i=1}^N \alpha_i \otimes \beta_i$. By linearity, every element of $V^* \otimes W^*$ is bilinear. Therefore $V^* \otimes W^* \subseteq B(V, W, \mathbb{F})$. It remains to be shown that $B(V, W, \mathbb{F})$ is a subset of $V^* \otimes W^*$.

Theorem 1.4. Let V, W be finite dimensional vector spaces. Then $B(V, W, \mathbb{F}) = V^* \otimes W^*$.

Proof. Suppose that $\gamma \in V^* \otimes W^*$. Then γ defines a bilinear map as shown above. Now, on the other hand, suppose that κ is any bilinear map taking inputs in $V \times W$. Let $\{e_1, \dots, e_n\} = B$ be a basis for V and let $\{f_1, \dots, f_m\} = C$ be a basis for W . Let $\{e^i\} = B^*$ and $\{f^j\} = C^*$ be the corresponding dual bases. Then let $\kappa_{ij} = \kappa(e_i, f_j)$. We can see that,

$$\begin{aligned} \kappa(v, w) &= v^i w^j \kappa(e_i, f_j) \\ &= e^i(v) f^j(w) \kappa_{ij} \\ &= \kappa_{ij} e^i \otimes f^j(v, w) \end{aligned}$$

Since this holds for all $v, w \in V, W$, we can see that $\kappa = \sum_{i,j} \kappa_{ij} e^i \otimes f^j$. So κ is indeed a linear combination of elements of $V^* \otimes W^*$. Additionally, from this we see that $\{e^i \otimes f^j : 1 \leq i \leq n, 1 \leq j \leq m\}$ is a basis for $V^* \otimes W^*$. \square

Corollary 1.1. Let V, W be finite dimensional vector spaces. Then

$$\dim(V^* \otimes W^*) = (\dim V)(\dim W)$$

Proof. In the proof of the previous theorem we constructed a basis $\{e^i \otimes f^j : 1 \leq i \leq n, 1 \leq j \leq m\}$ for $V^* \otimes W^*$. There are $nm = (\dim V)(\dim W)$ elements in this basis. \square

1.2.2 Tensor Product of Arbitrary Vector Spaces

Now we should define the tensor product in complete generality. For it to make sense as an explicit construction of a vector space takes some work.

One way to define the tensor product of any general vector spaces is as a quotient of $V \times W$. This is the most explicit construction which does not depend on a choice of basis. It also makes it very easy to prove the various properties of the tensor product. We will also have to show that this definition agrees with the above definition of $V^* \otimes W^*$.

Definition 1.7 (Tensor Product of Vector Spaces). Let V and W be vector spaces. Recall that $V \times W$ is a vector space. Then define the following subspace of $V \times W$.

$$R = \text{span} \left\{ \begin{aligned} &(v, w) \in V \times W, \text{ where } (v, w) = (v_1 + v_2, w) - (v_1, w) - (v_2, w) \\ &\text{or } (v, w) = (v, w_1 + w_2) - (v, w_1) - (v, w_2) \\ &\text{or } (v, w) = s(v_1, w) - (sv_1, w) \\ &\text{or } (v, w) = s(v, w_1) - (v, sw_1) \end{aligned} \right\}$$

Then the tensor product $V \otimes W$ is defined by,

$$V \otimes W = \frac{V \times W}{R}$$

Where we also define the notation,

$$v \otimes w = [(v, w)]$$

and in general, any element of $V \otimes W$ is of the form $\sum_{i=1}^N v_i \otimes w_i$. Remember that the quotient of two vector spaces is formed by considering equivalence classes $[v] = \{w \in V : w - v \text{ is equivalent to } 0\}$.

Theorem 1.5. The above definition ensures that the following equations hold:

- i. $(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w.$
- ii. $v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2.$
- iii. $s(v \otimes w) = (sv) \otimes w = v \otimes (sw).$

Proof. We will show, for example, that $(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w$. Observe that,

$$\begin{aligned} & (v_1 + v_2) \otimes w - v_1 \otimes w - v_2 \otimes w \\ &= [(v_1 + v_2, w)] - [(v_1, w)] - [(v_2, w)] \\ &= [(v_1 + v_2, w) - (v_1, w) - (v_2, w)] \end{aligned}$$

Since $(v_1 + v_2, w) - (v_1, w) - (v_2, w)$ is equivalent to 0 for all v_1, v_2, w , we therefore see that

$$[(v_1 + v_2, w) - (v_1, w) - (v_2, w)] = [(0, 0)]$$

In other words, $(v_1 + v_2) \otimes w - v_1 \otimes w - v_2 \otimes w = 0$. The proofs for the other identities are identical. \square

Lemma 1.1. Let V be a vector space over \mathbb{F} . Then $\mathbb{F} \otimes V$ is canonically isomorphic to V .

Proof. Let $a \otimes v \in \mathbb{F} \otimes V$. Observe that $a \otimes v = 1 \otimes (av)$. Let $i : \mathbb{F} \otimes V \rightarrow V$ be defined by $i(a \otimes v) = av$. Then i is invertible, since we can simply set $i^{-1}(v) = 1 \otimes v$. To see that this is valid, observe that $i^{-1}(i(a \otimes v)) = i^{-1}(av) = 1 \otimes av = a \otimes v$. So i is an isomorphism as required. \square

Theorem 1.6. Let V and W be vector spaces over \mathbb{F} . For all $\omega \in V^* \otimes W^*$, there is a **unique** linear map $\tilde{\omega} : V \otimes W \rightarrow \mathbb{F}$ which has the property that $\tilde{\omega}(v \otimes w) = \omega(v, w)$ for all $v \in V, w \in W$.

Proof. Recall that any element $T \in V \otimes W$ can be written in the form $T = \sum_{i=1}^N v_i \otimes w_i$, with each $v_i \in V$ and $w_i \in W$ for $i = 1, \dots, N$. Define $\tilde{\omega}(T) = \sum_{i=1}^N \omega(v_i, w_i)$ for any such $T \in V \otimes W$. We must verify that this is linear and then that it is unique.

To show that it is linear, let $S \in V \otimes W$ and set $S = \sum_{j=1}^M a_j \otimes b_j$, with $a_j \in V$ and $b_j \in W$ for each $j = 1, \dots, M$. Also let $\lambda \in \mathbb{F}$.

Observe,

$$\begin{aligned}
 \tilde{\omega}(T + \lambda S) &= \tilde{\omega} \left(\sum_{i=1}^N v_i \otimes w_i + \lambda \sum_{j=1}^M a_j \otimes w_j \right) \\
 &= \tilde{\omega} \left(\sum_{i=1}^N v_i \otimes w_i + \sum_{j=1}^M (\lambda a_j) \otimes w_j \right) \\
 &= \tilde{\omega} \left(\sum_{k=1}^{N+M} e_k \otimes f_k \right)
 \end{aligned}$$

Where $e_k = v_k$ for $k = 1, \dots, N$ and $e_k = \lambda a_k$ for $k = N + 1, \dots, N + M$, and similarly $f_k = w_k$ for $k = 1, \dots, N$ and $f_k = b_k$ for $k = N + 1, \dots, N + M$. So,

$$\begin{aligned}
 \tilde{\omega}(T + \lambda S) &= \sum_{k=1}^{N+M} \omega(e_k, f_k) \\
 &= \sum_{i=1}^N \omega(v_i, w_i) + \sum_{j=1}^M \omega(\lambda a_j, b_j) \\
 &= \sum_{i=1}^N \omega(v_i, w_i) + \lambda \sum_{j=1}^M \omega(a_j, b_j) \\
 &= \tilde{\omega}(T) + \lambda \tilde{\omega}(S)
 \end{aligned}$$

Therefore $\tilde{\omega}$ is linear as required. Finally we will show that it is unique.

Suppose some other map $\hat{\omega}$ existed so that $\tilde{\omega}(v \otimes w) = \hat{\omega}(v \otimes w)$ for all v, w . Then

$$\begin{aligned}
 \tilde{\omega}(T) &= \sum_{i=1}^N \omega(v_i, w_i) \\
 &= \sum_{i=1}^N \hat{\omega}(v_i \otimes w_i) \\
 &= \hat{\omega} \left(\sum_{i=1}^N v_i \otimes w_i \right)
 \end{aligned}$$

Therefore $\tilde{\omega}(T) = \hat{\omega}(T)$ for all $T \in V \otimes W$. So $\hat{\omega} = \tilde{\omega}$. Therefore this map is indeed unique. □

Theorem 1.7. $\dim(V \otimes W) = (\dim V)(\dim W)$

Corollary 1.2. Let $\{e_i : 1 \leq i \leq n\}$ be a basis for V and let $\{f_j : 1 \leq j \leq m\}$ be a basis for W . Then

$$\{e_i \otimes f_j : 1 \leq i \leq n, 1 \leq j \leq m\}$$

is a basis for $V \otimes W$.

Proof. (of Theorem 1.7 and Corollary 1.2)

Since there are $\dim V \dim W$ elements in the set $\{e_i \otimes f_j\}$ it suffices to show that this set is linearly independent. Let,

$$\omega = \omega^{ij} e_i \otimes f_j$$

Then $\omega = 0$ iff $\tilde{\omega} = 0$. We check:

$$\tilde{\omega}(\alpha \otimes \beta) = \omega^{ij} e_i(\alpha) f_j(\beta) = \omega^{ij} \alpha_i \beta_j$$

This is zero for all $\alpha \in V^*, \beta \in W^*$ iff $\omega^{ij} = 0$ for all i, j . Thus $e_i \otimes f_j$ is a linearly independent set. Finally, it is simple to show that any $\omega \in V \otimes W$ can be written as $\omega^{ij} e_i \otimes f_j$. So $\dim(V \otimes W) = \dim V \dim W$ as required, and this set is a basis. \square

Corollary 1.3. Let V and W be finite dimensional vector spaces. Then $(V \otimes W)^*$ is isomorphic to $V^* \otimes W^*$.

Proof. Consider the map $i : V^* \otimes W^* \rightarrow (V \otimes W)^*$ given by $i(\omega) = \tilde{\omega}$. The previous theorem shows that this map is injective. It suffices to show that $\dim(V^* \otimes W^*) = \dim((V \otimes W)^*)$. We have,

$$\begin{aligned} \dim((V \otimes W)^*) &= \dim(V \otimes W) \\ &= (\dim V)(\dim W) \\ &= \dim(V^* \otimes W^*). \end{aligned}$$

\square

Theorem 1.8. Let V and W be finite dimensional vector spaces. Then $L(V, W)$ is isomorphic to $V^* \otimes W$.

Proof. Recall that $\dim L(V, W) = \dim V \dim W = \dim V^* \otimes W$.

Let us define the map $f : V^* \otimes W \rightarrow L(V, W)$ according to the formula

$$f(\alpha \otimes w)(v) = \alpha(v)w$$

for all $\alpha \in V^*, w \in W, v \in V$.

Then f is linear, since α is linear. Therefore, we can extend f by linearity to all elements of $V^* \otimes W$.

Let $\{e_1, \dots, e_n\}$ be a basis for V and let $\{f_1, \dots, f_m\}$ be a basis for W . Then given any tensor $T \in V^* \otimes W$ we can write $T = T_i^j e^i \otimes f_j$.

Suppose $f(T) = 0$. Then $T_i^j e^i(v) f_j = 0$. But this can only be true if $v = 0$ or if $T_i^j = 0$ for all i, j .

Since the statement holds for any v , not just $v = 0$, we must have the second case, that $T = 0$. Therefore $\ker f = \{0\}$. Hence f is injective, and is therefore an isomorphism.

Observe that we only needed a basis to show that $\ker f = \{0\}$, but the definition of f itself did not depend on any choice of

basis. So we say f is a canonical isomorphism. From here on out, we will therefore simply refer to $f(T)$ as T . □

From here, we can inductively construct higher rank tensors as elements of iterated tensor product spaces.

1.2.3 The Tensor Algebra

Definition 1.8. Let V be a vector space over \mathbb{F} . The vector space of rank- (k, ℓ) tensors over V is defined as follows

$$\mathcal{T}_\ell^k(V) = \underbrace{(V \otimes \dots \otimes V)}_{k \text{ copies}} \otimes \underbrace{(V^* \otimes \dots \otimes V^*)}_{\ell \text{ copies}} \quad (1.1)$$

Definition 1.9. We often write,

$$V^{\otimes n} = \underbrace{V \otimes \dots \otimes V}_{n \text{ times}} \quad (1.2)$$

This is called the n 'th **tensor power** of V .

Theorem 1.9. Let $B = \{e_i : 0 \leq i \leq n\}$ be a basis for V . A basis for $\mathcal{T}_\ell^k(V)$ is given by the set,

$$\mathcal{B} = \left\{ e_{i_1} \otimes \dots \otimes e_{i_k} \otimes e^{j_1} \otimes \dots \otimes e^{j_\ell} : 0 \leq i_p, j_q \leq n \forall 0 \leq p \leq k, 0 \leq q \leq \ell \right\} \quad (1.3)$$

There are exactly $N = n^{k+\ell}$ basis vectors.

In this basis, any tensor $T \in \mathcal{T}_\ell^k$ can be written as a linear combination,

$$T = T^{i_1 \dots i_k}_{j_1 \dots j_\ell} e_{i_1} \otimes \dots \otimes e_{i_k} \otimes e^{j_1} \otimes \dots \otimes e^{j_\ell}$$

Where repeated indices are summed over. The coefficients $T^{i_1 \dots i_k}_{j_1 \dots j_\ell}$ are called the components (or coordinate representations) of T in the basis. It is common practice to define a multi-index to make this easier to write down.

Definition 1.10 (Multi-Index). Suppose we have an expression of the form

$$T = T^{i_1 \dots i_k}_{j_1 \dots j_\ell} e_{i_1} \otimes \dots \otimes e_{i_k} \otimes e^{j_1} \otimes \dots \otimes e^{j_\ell}$$

Then we can bundle up the indices into ordered lists of indices

$$I = \{i_1, \dots, i_k : i_n = 1, \dots, \dim V \forall n\}$$

and

$$J = \{j_1, \dots, j_\ell : j_n = 1, \dots, \dim V \forall n\}$$

so that the above expression is written

$$T = T_J^I e_I \otimes e^J$$

Where $e_I = e_{i_1} \otimes \dots \otimes e_{i_k}$ and $e^J = e^{j_1} \otimes \dots \otimes e^{j_\ell}$. The sets I and J are called multi-indices.

Remark. Let I be a multi index of length k over the set $\{1, \dots, n\}$. This means that I consists of a choice of k elements from $\{1, \dots, n\}$. So for example, if $k = 3, n = 5$ we could have $I = \{1, 2, 3\}$ or $\{1, 2, 5\}$ or $\{2, 4, 5\}$ and so on.

Definition 1.11. Let I be a multi index of length k over the set $\{1, \dots, n\}$. The **complementary multi-index** to I is the set $I^c = \{1, \dots, n\} \setminus I$. That is, I and I^c are completely disjoint, and $I \cup I^c = \{1, \dots, n\}$. Observe that $|I| + |I^c| = n$.

Example 1.1. Let $n = 5$. Then if $I = \{1, 2, 4\}$ we have $I^c = \{3, 5\}$.

By convention, we define $\mathcal{T}_0^0(V) = \mathbb{F}$, since $\mathbb{F} \otimes V$ is canonically isomorphic to V . Similarly, one will note that $\mathcal{T}_0^1(V) = V$ and $\mathcal{T}_1^0(V) = V^*$.

Definition 1.12 (Decomposable Tensor). A **decomposable** tensor $T \in \mathcal{T}_\ell^k(V)$ is a tensor which can be written in the form $T = v_1 \otimes \dots \otimes v_k \otimes \alpha^1 \otimes \dots \otimes \alpha^\ell$ for some $v_1, \dots, v_k, \alpha^1, \dots, \alpha^\ell$. Not all tensors are decomposable.

Example 1.2. The tensor $T = e_1 \otimes e_2$ is decomposable, but not the tensor $e_1 \otimes e_2 - e_2 \otimes e_1$.

Definition 1.13 (Tensor Algebra). Let V be a vector space over \mathbb{F} . The **tensor algebra** of V is the vector space

$$\bigotimes V = \bigoplus_{k=0}^{\infty} V^{\otimes k} \quad (1.4)$$

Where each element $Y \in \bigotimes$ is of the form $Y = \sum_{k=0}^N Y_k$ with each $Y_k \in V^{\otimes k}$, $0 \leq k \leq N$. Note that each element is defined as a finite linear combination of tensors.

Lemma 1.2. The tensor algebra is an associative algebra, where the "multiplication" is the tensor product.

Proof. The tensor product is associative and distributive. □

Additionally, notice that we have defined $1 \otimes v = v \otimes 1 = v$ since $\mathbb{F} \otimes V$ is isomorphic to V . This means that the algebra is **unital**, as in, it has a multiplicative identity.

Definition 1.14 (Unital Algebra). An algebra A is said to be **unital** if there exists some $e \in A$ so that $ev = ve = v$ for all $v \in A$.

1.2.4 Linear Maps as Tensors, and the Dual Map

Now we will shift gears a little bit to talk about something called the dual map associated with a linear map.

Let V, W be finite dimensional vector spaces and let $M : V \rightarrow W$ be a linear map. Let B be a basis of V and let C be a basis for W . Recall that we defined the coordinate matrix of M to be the matrix $[M_i^j]_{C,B}$ where $M_i^j = f^j(M(e_i))$ for each $e_i \in B, f^j \in C^*$. The C, B subscript in the above notation is there to remind the reader that a vector written in the basis B can be inserted on the right hand side, and when matrix multiplication is carried out, a vector written in the basis C comes out the other side. Suppose we were to perform a change of basis. That is, let \tilde{B} be another basis for V and let \tilde{C} be another basis for W . Then there exist unique linear maps $P : \mathbb{F}^n \rightarrow \mathbb{F}^n$ and $Q : \mathbb{F}^m \rightarrow \mathbb{F}^m$ so that $P([v]_{\tilde{B}}) = [v]_B$ for all $v \in V$ and so that $Q([w]_{\tilde{C}}) = [w]_C$ for all $w \in W$. These are called the change of basis transformations. Given these maps, we can convert the matrix $[M_i^j]_{C,B}$ which is written in terms of C, B into a new matrix written in terms of \tilde{C}, \tilde{B} . Observe that since P, Q are invertible, we have,

$$[M]_{C,B}[v]_B = [M(v)]_C = Q[M(v)]_{\tilde{C}} = Q[M]_{\tilde{C},\tilde{B}}[v]_{\tilde{B}} = Q[M]_{\tilde{C},\tilde{B}}P^{-1}[v]_B$$

So the matrix representation of M with respect to the new bases is given by,

$$[M]_{\tilde{C},\tilde{B}} = Q^{-1}[M]_{C,B}P$$

Theorem 1.10. Let $L(V, W)$ denote the set of all linear maps from V to W . Then $L(V, W)$ is canonically isomorphic to $W \otimes V^*$.

Proof. Let $M \in L(V, W)$.

Let $S \in W \otimes V^*$ be defined by $S_M(\alpha, v) = \alpha(M(v))$ for all $v \in V, \alpha \in W^*$.

So $S_M \in W \otimes V^*$ by construction.

Observe that for all $c_1, c_2 \in \mathbb{F}$ and $M_1, M_2 \in L(V, W)$ we have,

$$\begin{aligned} S_{c_1 M_1 + c_2 M_2}(v) &= \alpha(c_1 M_1 + c_2 M_2)(v) \\ &= c_1 \alpha(M_1)(v) + c_2 \alpha(M_2)(v) \\ &= c_1 S_{M_1}(\alpha, v) + c_2 S_{M_2}(\alpha, v) \end{aligned}$$

That is,

$$S_{c_1 M_1 + c_2 M_2} = c_1 S_{M_1} + c_2 S_{M_2}$$

So S defines a linear mapping between elements of $L(V, W)$ and elements of $W \otimes V^*$. We must show that it is injective and surjective.

Suppose that $S_M(\alpha, v) = 0$ for all v, α . Then $\alpha(M(v)) = 0$ for all α, v . So $M = 0$. Therefore $\ker S = \{0\}$, meaning S is injective.

Finally, observe that if $\dim V = n, \dim W = m$, then

$$\dim L(V, W) = \dim M_{nm}(\mathbb{F}) = nm$$

and

$$\dim W \otimes V^* = \dim W \dim V = nm$$

So $\dim L(V, W) = \dim W \otimes V^*$. So S has got to be surjective and therefore an isomorphism.

Since S did not depend on any particular choice of basis, we can see that it is indeed a canonical isomorphism. \square

Definition 1.15. Let $M : V \rightarrow W$ be a linear map. The **dual map** $M^* : W^* \rightarrow V^*$ (also known as the **pullback** of M) is defined so that,

$$M^*(\alpha)(v) = \alpha(M(v)) \quad \text{for all } \alpha \in W^*, v \in V \quad (1.5)$$

Let $\mathbb{1}_V : V \rightarrow V$ be the identity map on V and let $\mathbb{1}_{V^*} : V^* \rightarrow V^*$ be the identity map on V^* . Then $\alpha(\mathbb{1}_V(v)) = \alpha(v)$ for all $v \in V$ and $(\mathbb{1}_V)^*(\alpha)(v) = \alpha(\mathbb{1}_V(v)) = \alpha(v)$ for all $v \in V$. Therefore $\mathbb{1}_V^* = \mathbb{1}_{V^*}$.

Let B be a basis for V and let C be a basis for W . Let $\alpha \in W^*$, $v \in V$. Then $M^*(\alpha) \in V^*$. We write,

$$M^*(\alpha)(v) = [M^*(\alpha)]_{B^*}[v]_B$$

And,

$$\alpha(M(v)) = [\alpha]_{C^*}[M(v)]_C = [\alpha]_{C^*}[M]_{CB}[v]_B$$

So $[M^*(\alpha)]_{B^*} = [\alpha]_{C^*}[M]_{CB}$. Plugging in basis vectors $B = \{e_1, \dots, e_n\}$, $C = \{f_1, \dots, f_m\}$ gives the following result.

Theorem 1.11. Let $M : V \rightarrow W$ be a linear map. Let B, C be bases for V, W . The matrix representation of M^* is then

$$[M^*]_{B^*C^*} = [M]_{CB}^T$$

That is, the matrix of the dual map M^* is the transpose of the matrix of M .

Proof. By direct computation we have,

$$([M^*]_{B^*C^*})_j^i = M^*(f^j)(e_i) = f^j(M(e_i)) = ([M]_{CB})_i^j = ([M]_{CB}^T)_j^i$$

So $[M^*]_{B^*C^*} = [M]_{CB}^T$. \square

Corollary 1.4. $\det T^* = \det T$

1.3 Lecture 3: The Exterior Algebra

1.3.1 Antisymmetric Forms

First, remember that in this course we are working over a field \mathbb{F} of characteristic zero. This means that whenever $x \in \mathbb{F}$, then $x^n = 0$ if and only if $x = 0$. This includes $\mathbb{F} = \mathbb{R}, \mathbb{C}, \mathbb{Q}$ among others. The first way we will define the exterior algebra is as a vector subspace of the tensor algebra.

Definition 1.16 (Totally Antisymmetric Tensor). Let $\alpha \in \mathcal{T}_k^0(V)$. We say α is called totally antisymmetric if whenever two inputs are exchanged, the result picks up a minus sign. That is,

$$\alpha(v_1, \dots, v_i, \dots, v_j, \dots, v_k) = -\alpha(v_1, \dots, v_j, \dots, v_i, \dots, v_k)$$

With $0 \leq i < j \leq k$. In general, if $\sigma : \{1, \dots, k\} \rightarrow \{1, \dots, k\}$ is a permutation operator we have,

$$\alpha(v_{\sigma(1)}, \dots, v_{\sigma(k)}) = \text{sgn}(\sigma) \alpha(v_1, \dots, v_k)$$

Totally anti-symmetric tensors are often referred to as **exterior k -forms**, **alternating forms**, or **skew-symmetric forms**.

Definition 1.17 (Exterior Power of V^*). Let V be a finite dimensional vector space over \mathbb{F} . The set of totally anti-symmetric $(k, 0)$ -tensors over V is called the k 'th **exterior power** of V^* , and is written,

$$\Lambda^k V^* = \{\alpha \in \mathcal{T}_k^0(V) : \alpha \text{ is totally anti-symmetric}\} \quad (1.6)$$

By convention $\Lambda^0 V^* = \mathbb{F}$. Elements of $\Lambda^k V^*$ are called **k -forms**.

Since a linear combination of any two totally anti-symmetric k -forms is also totally anti-symmetric, this set forms a vector subspace of $\mathcal{T}_k^0(V)$. We can construct these totally anti-symmetric forms in a similar way to the tensor product.

Definition 1.18 (Exterior Product). Let $\alpha_1, \dots, \alpha_k \in V^*$ be some one-forms. The exterior product (also known as the wedge product) of $\alpha_1, \dots, \alpha_k$ is an alternating k -form defined by the formula,

$$\alpha_1 \wedge \dots \wedge \alpha_k(v_1, \dots, v_k) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) \alpha_{\sigma(1)} \otimes \dots \otimes \alpha_{\sigma(k)}(v_1, \dots, v_k) \quad (1.7)$$

Remark. In some sources, the right hand side of (1.7) is multiplied by $1/k!$. This is a matter of convention, but be careful with this!

Physics Remark. In physics we would write this formula using the Levi-Civita symbol as,

$$\alpha_1 \wedge \dots \wedge \alpha_k = \varepsilon^{i_1 \dots i_k} \alpha_{i_1} \otimes \dots \otimes \alpha_{i_k}$$

The Levi-Civita symbol makes some proofs easier, and some proofs harder. In this course we will avoid it.

We should verify first that this actually does define a k -form.

Theorem 1.12. The tensor $\alpha_1 \wedge \dots \wedge \alpha_k$ is indeed totally antisymmetric.

Proof. Observe that $\alpha_{\sigma(1)} \otimes \dots \otimes \alpha_{\sigma(k)}(v_1, \dots, v_k) = \alpha_{\sigma(1)}(v_1) \dots \alpha_{\sigma(k)}(v_k)$. Upon exchanging the i 'th entry of this product for the j 'th, we essentially compose the permutation σ with the map τ so that $\tau(p) = p$ for all $p \neq i, j$, but $\tau(i) = j, \tau(j) = i$. This map τ is also a permutation, and is acyclic. So we have,

$$\alpha_1 \wedge \dots \wedge \alpha_k(v_1, \dots, v_j, \dots, v_i, \dots, v_k) = \sum_{\sigma \in S_k} \text{sgn}(\sigma) \alpha_{\sigma \circ \tau(1)} \otimes \dots \otimes \alpha_{\sigma \circ \tau(k)}(v_1, \dots, v_k)$$

Now, observe that this is a sum over all permutations. Since σ is arbitrary, we can replace σ everywhere by $\sigma \circ \tau^{-1}$ and the result will still be the same sum over all permutations.

$$= \sum_{\sigma \in S_k} \text{sgn}(\sigma \circ \tau^{-1}) \alpha_{\sigma(1)} \otimes \dots \otimes \alpha_{\sigma(k)}(v_1, \dots, v_k)$$

Finally, since the sign of a composition of permutations is the product of the signs, we have

$$\begin{aligned} &= \text{sgn}(\tau^{-1}) \sum_{\sigma \in S_k} \text{sgn}(\sigma) \alpha_{\sigma(1)} \otimes \dots \otimes \alpha_{\sigma(k)}(v_1, \dots, v_k) \\ &= -\alpha_1 \wedge \dots \wedge \alpha_k(v_1, \dots, v_i, \dots, v_j, \dots, v_k) \end{aligned}$$

As required. □

Alternate Proof. Another way to prove that this is antisymmetric is by computing the components directly. Let $A = [\alpha^i(v_j)]$ be the matrix formed by computing $\alpha^i(v_j)$ for each $0 \leq i, j \leq k$. We can use the Laplace expansion formula for the determinant,

$$\begin{aligned} \alpha_1 \wedge \dots \wedge \alpha_k(v_1, \dots, v_k) &= \sum_{\sigma \in S_k} \text{sgn}(\sigma) \alpha_{\sigma(1)} \otimes \dots \otimes \alpha_{\sigma(k)}(v_1, \dots, v_k) \\ &= \sum_{\sigma \in S_k} \text{sgn}(\sigma) \prod_{i=1}^k \alpha_{\sigma(i)}(v_i) \\ &= \det A \end{aligned}$$

Since the determinant function acquires a minus sign whenever a row or column is exchanged, we can see that exchanging either α_i for α_j or v_i for v_j introduces a minus sign. Therefore the wedge product indeed results in an alternating form. □

Example 1.3. Consider $k = 2$. We have $\alpha \wedge \beta = \alpha \otimes \beta - \beta \otimes \alpha$.

Example 1.4. Consider $k = 3$. We have

$$\alpha \wedge \beta \wedge \gamma = \alpha \otimes \beta \otimes \gamma + \beta \otimes \gamma \otimes \alpha + \gamma \otimes \alpha \otimes \beta - \alpha \otimes \gamma \otimes \beta - \gamma \otimes \beta \otimes \alpha - \beta \otimes \alpha \otimes \gamma$$

Lemma 1.3. Let V be a finite dimensional vector space with basis $B = \{e_i\}$. Then a basis for $\Lambda^k(V)$ is $\Lambda^k B = \{e_{i_1} \wedge \dots \wedge e_{i_k} : 0 \leq i_1, \dots, i_k \leq n\}$.

Proof. We write $\omega = \omega_{i_1 \dots i_k} e^{i_1} \otimes \dots \otimes e^{i_k}$. Note that if σ is a permutation, then

$$\omega_{i_1 \dots i_k} e^{i_1} \otimes \dots \otimes e^{i_k} = \omega_{i_{\sigma(1)} \dots i_{\sigma(k)}} e^{i_{\sigma(1)}} \otimes \dots \otimes e^{i_{\sigma(k)}}$$

This follows from the fact that relabeling the indices doesn't do anything, and has nothing to do with ω being a k -form. However, the fact that ω is totally antisymmetric comes in the next step. Since there are $k!$ possible permutations, we can write

$$\omega_{i_1 \dots i_k} e^{i_1} \otimes \dots \otimes e^{i_k} = \frac{1}{k!} \sum_{\sigma \in S_k} \omega_{i_{\sigma(1)} \dots i_{\sigma(k)}} e^{i_{\sigma(1)}} \otimes \dots \otimes e^{i_{\sigma(k)}}$$

We can then use the antisymmetry of ω to remove the permutation from the indices on $\omega_{i_1 \dots i_k}$ and insert a $\text{sgn}(\sigma)$ as so,

$$\begin{aligned} \omega &= \omega_{i_1 \dots i_k} e^{i_1} \otimes \dots \otimes e^{i_k} \\ &= \frac{1}{k!} \sum_{\sigma \in S_k} \omega_{i_{\sigma(1)} \dots i_{\sigma(k)}} e^{i_{\sigma(1)}} \otimes \dots \otimes e^{i_{\sigma(k)}} \\ &= \frac{1}{k!} \sum_{\sigma \in S_k} \text{sgn}(\sigma) \omega_{i_1 \dots i_k} e^{i_{\sigma(1)}} \otimes \dots \otimes e^{i_{\sigma(k)}} \\ &= \omega_{i_1 \dots i_k} \frac{1}{k!} \sum_{\sigma \in S_k} \text{sgn}(\sigma) e^{i_{\sigma(1)}} \otimes \dots \otimes e^{i_{\sigma(k)}} \\ &= \frac{1}{k!} \omega_{i_1 \dots i_k} e^{i_1} \wedge \dots \wedge e^{i_k} \end{aligned}$$

As required. □

Remark. Observe that this is where the $k!$ would have come in handy. If we included it in the definition of the wedge product we could cancel it here. However, in this course we are using the convention that does not include the $k!$, so we have to be a bit careful with the components in the tensor basis and the components k -form basis.

Observe that since

$$e_{i_1} \wedge \dots \wedge e_{i_p} \wedge \dots \wedge e_{i_p} \wedge \dots \wedge e_{i_k} = -e_{i_1} \wedge \dots \wedge e_{i_p} \wedge \dots \wedge e_{i_p} \wedge \dots \wedge e_{i_k}$$

we have $2e_{i_1} \wedge \dots \wedge e_{i_p} \wedge \dots \wedge e_{i_p} \wedge \dots \wedge e_{i_k} = 0$. So any time two entries repeat in a wedge product the result is zero.

Corollary 1.5. Let V be a vector space and let $\dim V = n$. Then $\dim \Lambda^k(V) = \binom{n}{k} = \frac{n!}{(n-k)!k!}$.

Corollary 1.6. We can always write $\alpha = \sum_{i=1}^N \alpha_{i,1} \wedge \dots \wedge \alpha_{i,k}$ for some finite list of 1-forms $\alpha_{i,k}$. Tensors of the form $\alpha_{i,1} \wedge \dots \wedge \alpha_{i,k}$ are called **wedge-decomposable**.

Corollary 1.7. Let $\alpha_1, \dots, \alpha_k \in V^*$. Then $\alpha_1 \wedge \dots \wedge \alpha_k = 0$ if and only if $\alpha_1, \dots, \alpha_k$ are not linearly independent.

Proof. Suppose they are not linearly independent. Then $\alpha_1 = \sum_{i=2}^k c_i \alpha_i$. So $\alpha_1 \wedge \dots \wedge \alpha_k = \sum_{i=2}^k c_i \alpha_i \wedge \alpha_2 \wedge \dots \wedge \alpha_k = 0$ since every term in this sum has repeated entries. Now for the other direction of the "if and only if" statement: suppose that $\alpha_1 \wedge \dots \wedge \alpha_k = 0$. Then $\sum_{\sigma \in S_k} \text{sgn}(\sigma) \alpha_{\sigma(1)} \otimes \dots \otimes \alpha_{\sigma(k)} = 0$. Let $\{e_i\}$ be any basis for V . Then,

$$\sum_{\sigma \in S_k} \text{sgn}(\sigma) \alpha_{\sigma(1)} \otimes \dots \otimes \alpha_{\sigma(k)}(e_1, \dots, e_k) = \sum_{\sigma \in S_k} \text{sgn}(\sigma) \alpha_{\sigma(1)}(e_1) \dots \alpha_{\sigma(k)}(e_k) = 0$$

From this, if we group first the terms containing $\alpha_1(e_1), \dots, \alpha_k(e_1)$, and then the terms containing $\alpha_1(e_2), \dots, \alpha_k(e_2)$, and so on, we find that each grouping of terms will contain the same constants C_1, \dots, C_k so that

$$0 = C_1 \alpha_1(e_i) + \dots + C_k \alpha_k(e_i)$$

for all $0 \leq i \leq k$. In other words, $\alpha_1, \dots, \alpha_k$ are linearly dependent. □

Corollary 1.8. If $k > n$ then $\Lambda^k V^* = \{0\}$.

Definition 1.19 (Wedge product of totally antisymmetric tensors). Let $\alpha \in \Lambda^k V^*$ and $\beta \in \Lambda^\ell V^*$. Recall that we can always write α and β as sums of wedge-decomposable tensors,

$$\alpha = \sum_{i=1}^N \alpha_{i,1} \wedge \dots \wedge \alpha_{i,k}, \quad \beta = \sum_{j=1}^M \beta_{j,1} \wedge \dots \wedge \beta_{j,\ell}$$

for some finite list of 1-forms $\alpha_{i,k}, \beta_{j,k}$. We therefore define $\alpha \wedge \beta \in \Lambda^{k+\ell} V^*$ as,

$$\alpha \wedge \beta = \sum_{i=1}^N \sum_{j=1}^M \alpha_{i,1} \wedge \dots \wedge \alpha_{i,k} \wedge \beta_{j,1} \wedge \dots \wedge \beta_{j,\ell}$$

Lemma 1.4. If $\alpha \in \Lambda^k V^*$ and $\beta \in \Lambda^\ell V^*$, then $\alpha \wedge \beta = (-1)^{k\ell} \beta \wedge \alpha$.

Proof. Recall that

$$\alpha \wedge \beta = \sum_{i=1}^N \sum_{j=1}^M \alpha_{i,1} \wedge \dots \wedge \alpha_{i,k} \wedge \beta_{j,1} \wedge \dots \wedge \beta_{j,\ell}$$

We must then move β_1 to the far left, which requires k swaps. Since there are ℓ factors in β , we must repeat this ℓ times. Therefore, we get a total of $k\ell$ swaps, which means a factor of $(-1)^{k\ell}$. \square

1.3.2 The Exterior Algebra

Proof. Let $B = \{e_1, \dots, e_n\}$ be a basis for V . Then suppose $\alpha \in \Lambda^k V^*$ with $k > n$. This would mean that $\alpha = \sum \alpha_I e^{i_1} \wedge \dots \wedge e^{i_k}$. But a set of $k > n$ elements of B is necessarily linearly dependent. So $\alpha = 0$ as required. \square

Definition 1.20 (Exterior Algebra). Let V be a vector space of dimension n . We define the exterior algebra of V to be,

$$\Lambda^\bullet V^* = \bigoplus_{k=0}^n \Lambda^k V^* \quad (1.8)$$

Where the product associated to this algebra is the wedge product defined above. It is associative.

Theorem 1.13. The exterior algebra has dimension $\dim \Lambda^\bullet V^* = 2^n$

Proof. By the binomial theorem,

$$\dim \Lambda^\bullet V^* = \sum_{k=0}^n \binom{n}{k} = \sum_{k=0}^n \binom{n}{k} 1^k 1^{n-k} = (1+1)^n = 2^n$$

\square

We say if $\alpha \in \Lambda^k V^*$, then α is of degree k . We sometimes write $|\alpha| = k$ for the degree of α .

Definition 1.21 (Graded Algebra). Let A be an associative algebra. Then A is said to be **graded** if we can write

$$A = \bigoplus_{i=0}^{\infty} A_i$$

Where the vector spaces $A_i, i \in \mathbb{N}$ have the property that if $v \in A_i$ and $w \in A_j$ then $vw \in A_{i+j}$.

Elements of $A_{2i}, i \in \mathbb{N}$ are called elements of **even degree**. Elements of $A_{2i+1}, i \in \mathbb{N}$ are called elements of **odd degree**.

Theorem 1.14. The exterior algebra $\Lambda^\bullet V^*$ is a graded algebra.

Proof. We have $A_0 = \mathbb{F}$, $A_k = \Lambda^k V^*$ for $k = 1, \dots, n$, and $A_k = \{0\}$ for all $k > n$.

Suppose α is of degree k and β is of degree ℓ . Then if $1 \leq k + \ell \leq n$ it follows that $\alpha \wedge \beta \in \Lambda^{k+\ell} V^*$ by definition. If α is of degree zero, then α is just a scalar, which we have defined to act as $\alpha \wedge \beta = \alpha\beta \in \Lambda^\ell V^* = \Lambda^{0+\ell} V^*$ as required. Otherwise, if $k + \ell > n$ we have $\alpha \wedge \beta = 0 \in A_{k+\ell}$ since α and β will necessarily contain linearly dependent components. So we have verified everything we needed to. \square

Also note that since $\Lambda^\bullet V^*$ does not have the same "product" as $\mathcal{T}_k^0(V)$, it is not a vector subalgebra even though it is a vector subspace. These notions are different, and should be pointed out.

Now we will talk about how a linear map $T : V \rightarrow W$ induces a linear map $\mathcal{T}_\ell^k(T) : \mathcal{T}_\ell^k(V) \rightarrow \mathcal{T}_\ell^k(W)$. Recall that any tensor can be written as the sum of decomposable tensors. We can therefore define the induced map as acting on decomposable tensors, and then extend the definition to all tensors by requiring that the induced map is linear (ie, using the extension by linearity theorem).

Definition 1.22 (Tensor Power of Linear Map). Let $T : V \rightarrow W$ be a linear map. Then let $\omega = v_1 \otimes \dots \otimes v_k \otimes \alpha_1 \otimes \dots \otimes \alpha_\ell$ be a decomposable tensor. We define,

$$\mathcal{T}_\ell^k(T)(\omega) = T(v_1) \otimes \dots \otimes T(v_k) \otimes T^*(\alpha_1) \otimes \dots \otimes T^*(\alpha_\ell)$$

and extend the definition by linearity to all of $\mathcal{T}_\ell^k(V)$.

Definition 1.23 (Exterior Power of Linear Map). Let $T : V \rightarrow W$ be a linear map. We often write $\Lambda^k T^*$ to mean the map with the property that

$$\Lambda^k T^*(\alpha_1 \wedge \dots \wedge \alpha_k) = T^*(\alpha_1) \wedge \dots \wedge T^*(\alpha_k)$$

Where we extend by linearity to the rest of $\Lambda^k V^*$.

We refer to $\Lambda^k T^*$ as the k 'th exterior power of T .

Theorem 1.15. Let $\dim V = n$. and let $T : V \rightarrow V$ be a linear map. Then $\Lambda^n T^* = (\det T) \mathbb{1}$.

Proof. Let $B = \{e_1, \dots, e_n\}$ be a basis for V . Then $\Lambda^n V^* = \text{span}\{e^1 \wedge \dots \wedge e^n\}$. It follows that $\Lambda^n T^*(e^1 \wedge \dots \wedge e^n) = C e^1 \wedge \dots \wedge e^n$

for some constant C . So C is the only coefficient of $\Lambda^n T$. We therefore directly compute the coefficient:

$$\begin{aligned}
C &= (\Lambda^n T^*)_{1, \dots, n} \\
&= \Lambda^n T^* (e^1 \wedge \dots \wedge e^n)(e_1, \dots, e_n) \\
&= \sum_{\sigma \in S_n} \text{sgn}(\sigma) e^{\sigma(1)} \otimes \dots \otimes e^{\sigma(n)} (T(e_1), \dots, T(e_n)) \\
&= \sum_{\sigma \in S_n} \text{sgn}(\sigma) \prod_{i=1}^n e^{\sigma(i)} (T(e_i)) \\
&= \det T
\end{aligned}$$

As required. □

1.4 Lecture 4: Dual Maps, Annihilators, and Non-degenerate Bilinear Forms

1.4.1 Annihilators

Definition 1.24 (Annihilator). Let V be a vector space and let W be a subspace of V . Then

$$W^\circ = \{\alpha \in V^* : \alpha(W) = \{0\}\} \quad (1.9)$$

is called the **annihilator** of W . Be careful to note that W° is a subspace of V^* .

Lemma 1.5. $\dim W^\circ = \dim V - \dim W$

Proof. Let $B = \{e_1, \dots, e_k, f_1, \dots, f_{n-k}\}$ be a basis for V , where $B|_W = \{e_1, \dots, e_k\}$ is a basis for W . $f^i(e_j) = 0$ for all $i = 1, \dots, n-k$ and $j = 1, \dots, k$. Observe that if $\alpha \in W^\circ$, we can write $\alpha = \alpha_i e^i + \tilde{\alpha}_i f^i$. So then for any $w \in W$, we have

$$\begin{aligned}
0 &= \alpha(w) = \alpha(w^i e_i) \\
&= \alpha_j w^j e^k(e_i) + \tilde{\alpha}_p w^i f^p(e_i) \\
&= \alpha_j e^k(e_i) \quad \text{for all } i = 1, \dots, k
\end{aligned}$$

So $\alpha_j = 0$ for all $j = 1, \dots, k$. So $\alpha = \tilde{\alpha}_i f^i, i = 1, \dots, n-k$. Therefore f^1, \dots, f^{n-k} is a basis for W° . □

Let us prove some theorems about how the fundamental subspaces of a linear map and its dual are related to annihilators of subspaces.

Remark. Let $T : V \rightarrow W$ be a linear map. Its dual is $T^* : W^* \rightarrow V^*$. So,

1. $\ker T \subseteq V$
2. $\text{image } T \subseteq W$

$$3. \ker T^* \subseteq W^*$$

$$4. \text{image } T^* \subseteq V^*$$

Theorem 1.16. Let $T : V \rightarrow W$ be a linear map. Then,

$$\ker T^* = (\text{image } T)^\circ, \quad \text{image } T^* = (\ker T)^\circ \quad (1.10)$$

Furthermore,

$$\dim \text{image } T = \dim \text{image } T^* \quad (1.11)$$

Proof. Let $\beta \in W^*$. Then

$$\begin{aligned} \beta \in (\text{image } T)^\circ &\iff \beta(T(v)) = 0 \quad \forall v \\ &\iff T^* \beta(v) = 0 \quad \forall v \\ &\iff T^* \beta = 0 \\ &\iff \beta \in \ker T^* \end{aligned}$$

By rank-nullity theorem,

$$\begin{aligned} \dim \text{image } T^* &= \dim W^* - \dim \ker T^* \\ &= \dim W - \dim (\text{image } T)^\circ \\ &= \dim W - (\dim W - \dim \text{image } T) \\ &= \dim \text{image } T \end{aligned}$$

Now let $\alpha \in \text{image } T^*$. Then $\alpha = T^* \beta$ implies that $\alpha(v) = \beta(T(v))$. So if $v \in \ker T$ then $\alpha(v) = 0$ implies that $\alpha \in (\ker T)^\circ$ as required. So $\text{image } T^* \subseteq (\ker T)^\circ$. But by rank-nullity theorem, $\dim \text{image } T^* = \dim V - \dim \ker T = \dim (\ker T)^\circ$. Therefore $\text{image } T^* = (\ker T)^\circ$. \square

1.4.2 Bilinear Forms on a Vector Space V

Recall that if $T : V \rightarrow V$ is a linear operator, then $T \in V^* \otimes V$, and $\det T$ is well defined because the determinant does not depend on the choice of basis. We will show that for a bilinear form $G \in V^* \otimes V^*$ this is not the case. The determinant of a bilinear form can not be defined. In place of the determinant, we therefore have to turn to other invariants in order to learn about the properties of bilinear forms. One such invariant will be the signature.

Definition 1.25 (Matrix of a Bilinear Form). Let $G \in V^* \otimes V^*$ be a bilinear form. Let $B = \{e_1, \dots, e_n\}$ be a basis for V . Then the matrix $[G]_B$ of G is defined to be the $n \times n$ matrix with entries $G_{ij} = G(e_i, e_j)$. Observe that $G = G_{ij} e^i \otimes e^j$.

Theorem 1.17. Let $G \in V^* \otimes V^*$ be a bilinear form. Let $B = \{e_1, \dots, e_n\}$ and $\tilde{B} = \{\tilde{e}_1, \dots, \tilde{e}_n\}$ be two bases for V . Let $P_{\tilde{B}B} \in M_{nn}(\mathbb{F})$ be the change of basis matrix. Then the matrix $[G]_B$ is related to the matrix $[G]_{\tilde{B}}$ by the formula

$$P_{\tilde{B}B}^T [G]_B P_{\tilde{B}B} = [G]_{\tilde{B}} \quad (1.12)$$

Notice that instead of using the inverse of P we use the transpose! Therefore we have $\det[G]_{\tilde{B}} = \det(P_{\tilde{B}B}^T P_{\tilde{B}B}) \det[G]_B$. So the determinants of the two matrices are not the same! So the determinant is **not** an invariant of G . However, we can say something about G from the rank.

Lemma 1.6. The rank of G is an invariant.

Proof. This can be seen from the fact that $P_{\tilde{B}B}^T$ is the dual map of $P_{\tilde{B}B}$, and as we just showed in the previous section, $\text{rank } T = \text{rank } T^*$. Therefore, the rank of $[G]_B$ is the same as the rank of $[G]_{\tilde{B}}$ because of the fact that $\text{rank } P_{\tilde{B}B} = \text{rank } P_{\tilde{B}B}^T = \dim V$. \square

Definition 1.26 (Non-Degenerate Bilinear Form). Let $G \in V^* \otimes V^*$. Then G is called **non-degenerate** if $[G]_B$ is invertible for all bases B . That is, if $\text{rank } G = \dim V$.

Definition 1.27 (Orthogonal). Let $G \in V^* \otimes V^*$ and let $v, w \in V^*$. Then v is said to be **orthogonal** to w if $G(v, w) = 0$.

Note that the statement " v is orthogonal to w " is only equivalent to " w is orthogonal to v " if it is true that $G(v, w) = 0 \iff G(w, v) = 0$.

Definition 1.28 (Symmetric/Skew-Symmetric Bilinear Form). G is said to be **symmetric** if $G(v, w) = G(w, v)$ for all $v, w \in V$. G is said to be **skew-symmetric** if $G(v, w) = -G(w, v)$ for all $v, w \in V$.

Definition 1.29. The **symmetric part** of G is defined as

$$\text{Sym}(G)(v, w) = \frac{G(v, w) + G(w, v)}{2} \quad (1.13)$$

It is a symmetric bilinear form.

The **skew-symmetric** part of G is defined as

$$\text{Alt}(G)(v, w) = \frac{G(v, w) - G(w, v)}{2} \quad (1.14)$$

It is a skew-symmetric bilinear form.

Observe that $G = \text{Sym } G + \text{Alt } G$, so any bilinear form can be written as the sum of a symmetric form and a skew-symmetric form. This very useful fact allows us to decompose $V^* \otimes V^*$ into two subspaces.

Theorem 1.18. Let $G \in V^* \otimes V^*$. Suppose that $G(v, w) = 0$ if and only if $G(w, v) = 0$. Then either G is symmetric or G is skew-symmetric.

Proof. Let $u, v, w \in V$. Let $y = G(u, v)w - G(u, w)v$. By bilinearity, $G(u, y) = G(u, v)G(u, w) - G(u, w)G(u, v) = 0$. Therefore u is orthogonal to y . Since $G(v, w) = 0$ if and only if $G(w, v) = 0$, we see that this means

$$G(u, v)G(w, u) = G(u, w)G(v, u)$$

for all u, v, w . Call this equation (1)

Now, consider the above formula in the case $u = v$. We get $G(v, v)G(w, v) = G(v, w)G(v, v)$, so

$$G(v, v)(G(v, w) - G(w, v)) = 0.$$

Call this equation (2). This implies that either $G(v, v) = 0$ or $G(v, w) = G(w, v)$. The case $G(v, v) = 0$ implies that G is skew-symmetric, while the case $G(v, w) = G(w, v)$ means that G is symmetric. Therefore G must be symmetric, skew-symmetric, or both.

Suppose that G were both. Then there must be some nonzero $x, y, z \in V$ so that $G(x, y) \neq G(y, x)$ and so that $G(z, z) \neq 0$. Call this fact (3). But then by (1), we get $G(x, x)(G(y, z) - G(z, y)) = 0$ and $G(y, y)(G(x, z) - G(z, x)) = 0$, so

$$G(x, x) = 0, \quad G(y, y) = 0, \quad G(x, z) = G(z, x), \quad G(y, z) = G(z, y)$$

We also have $G(x, z) = 0$ and $G(y, z) = 0$. Call this (4).

By (3) and (4), $G(z, y) = G(y, z) = 0$ and $G(x, z) = G(z, x) = 0$. Call this (5).

From (5) and (3), we see that $G(x, y + z) = G(x, y) + G(x, z) = G(x, y)$. But

$$G(y, x) + G(z, x) = G(y + z, x) \neq G(x, y)$$

So $G(x, y + z) \neq G(y + z, x)$. By (2) we see that $G(y + z, y + z) = G(z, z) = 0$. But $G(z, z) \neq 0$ by (3), so we have reached a contradiction. Therefore G can not be both symmetric and skew-symmetric. \square

Definition 1.30 (Perp Space). Let V be a vector space and let $W \subseteq V$ be a subspace of V . Let $G \in V^* \otimes V^*$. We define

$$W^\perp = \{v \in V : G(v, w) = 0 \ \forall w \in W\} \tag{1.15}$$

These are the elements of V which are orthogonal to every element of W . Since G is bilinear, the set W^\perp forms a subspace of V . This subspace is called the perpendicular subspace to W (or, for short, the **perp space**).

Example 1.5. If G is an inner product, then the set W^\perp can be interpreted literally as the hyperplane which lies perpendicular to W . In general, G does not define a notion of perpendicular unless it is positive definite and non-degenerate.

Lemma 1.7. Let G be a symmetric or skew-symmetric bilinear form. Then $W \subseteq (W^\perp)^\perp$.

Proof. Let $w \in W$ and $v \in W^\perp$. Then $G(w, v) = \pm G(v, w) = 0$, so w is orthogonal to every element of W^\perp . In other words, $w \in (W^\perp)^\perp$. \square

Lemma 1.8. Let W_1 and W_2 be subspaces of V and suppose $W_1 \subseteq W_2$. Then $(W_2)^\perp \subseteq (W_1)^\perp$.

Proof. Let $v \in (W_2)^\perp$. Then $G(v, u) = 0$ for all $u \in W_2$. Since any element u of W_1 is an element of W_2 this means $G(v, u) = 0$ for all $u \in W_1$. So $v \in (W_1)^\perp$. \square

Definition 1.31 (Radical). Let V be a vector space. Note that V is a subspace of itself. Let G be a bilinear form on V . We define the **radical** of G to be the vector subspace $\text{rad}(G) = V^\perp$. This is the set of vectors in V which are orthogonal to every other vector in V .

Definition 1.32. Let V be a vector space and let $G \in V^* \otimes V^*$. We define the **sharp map**, $\sharp : V \rightarrow V^*$, by

$$\sharp(v)(w) = G(v, w) \quad (1.16)$$

Observe that since G is bilinear, \sharp is linear.

Observe that $\ker \sharp = \{v : G(v, w) = 0 \ \forall w\} = V^\perp = \text{rad}(G)$. So \sharp is an isomorphism iff $\text{rad}(G) = \{0\}$.

Remark. Note that $\sharp^* : V^* \rightarrow V^{**} \cong V$. Therefore we define $\sharp^*(v)(w) = \sharp^*(\xi(v))(w) = \xi(v)(\sharp(w)) = \sharp(w)(v) = G(w, v)$. So if G is symmetric, $\sharp(v)(w) = \sharp^*(w)(v)$ and if G is antisymmetric, $\sharp(v)(w) = -\sharp^*(w)(v)$.

Theorem 1.19. Let G be a symmetric or skew-symmetric bilinear form. Then G is non-degenerate if and only if $\text{rad}(G) = \{0\}$.

Proof. Let $\{e_1, \dots, e_n\}$ be a basis for V . Then $G(e_i, v) = 0$ for all i if and only if $G(u, v) = 0$ for all u . Therefore $v \in \text{rad}(G)$ if and only if $G(e_i, v) = 0$ for all $i = 1, \dots, n$. Let us expand v in this basis as $v = v^j e_j$. Then if $G(e_i, v) = 0$ for all i we see that $v^j G_{ij} = 0$ for all i , which is exactly the statement that $[v]_B \in \ker[G]_B$ if we consider $[G]_B$ as a linear map.

Therefore, $V^\perp = \{0\}$ iff $\ker[G]_B = \{0\}$. This is true if and only if $\text{rank } G = \dim V$, if and only if G is non-degenerate. This completes the proof. \square

Corollary 1.9. The map \sharp is an isomorphism if and only if G is non-degenerate. That is, if G is non-degenerate then it defines a canonical isomorphism between V and V^* .

2 Classification of Bilinear Forms

2.1 Lecture 5: Restrictions of Bilinear Forms

2.1.1 Restriction to Subspace

Theorem 2.1 (Riesz Representation Theorem (Finite Dimensional)). Let V be a finite dimensional vector space and let G be a bilinear form. Suppose G is non-degenerate. Then for all $\alpha \in V^*$ there is a unique $v \in V$ so that $\alpha(w) = G(v, w)$ for all $w \in V$.

Proof. Set $\alpha = \sharp(v)$. □

Definition 2.1. Let W be a linear subspace of V . The map $\iota : W \rightarrow V$ given by $\iota(w) = w$ is called the inclusion map.

Remark. Observe that $\iota^* : V^* \rightarrow W^*$. So if $\alpha \in V^*$ we have $(\iota^*\alpha)(w) = \alpha(\iota(w)) = \alpha(w)$ for any $w \in W$.

This definition may seem pointless, but the following lemmas are useful, and easier to prove with use of this inclusion map.

Lemma 2.1. $\ker(\sharp \circ \iota) = W \cap V^\perp$ and $\ker(\iota^* \circ \sharp) = W^\perp$

Proof. Let $w \in W$. Then $\sharp \circ \iota(w)(v) = \sharp w(v) = G(w, v)$. So $\ker(\sharp \circ \iota) = \{w \in W : G(w, v) = 0 \forall v \in V\} = W \cap V^\perp$. Similarly, $\iota^* \circ \sharp(v)(w) = G(v, w)$ for all $w \in W$. Therefore $\ker(\iota^* \circ \sharp) = \{v \in V : G(v, w) = 0 \forall w \in W\} = W^\perp$. □

Theorem 2.2. Let W be a subspace of V . Then $\dim W + \dim W^\perp = \dim V + \dim(W \cap V^\perp)$.

Proof. We begin by noting that,

$$\dim W + \dim \ker \iota^* \circ \sharp = \dim W + \dim V - \dim \text{image } \iota^* \circ \sharp$$

So,

$$\begin{aligned} \dim W &= \dim \ker \sharp \circ \iota + \dim \text{image } \sharp \circ \iota \\ &= \dim W \cap V^\perp + \dim(\ker \iota^* \circ \sharp)^\circ \\ &= \dim W \cap V^\perp + \dim V - \dim W^\perp \end{aligned}$$

Which can be rearranged to the desired equation. □

Corollary 2.1. Let G be non-degenerate. Then $\dim W + \dim W^\perp = \dim V$

Corollary 2.2. Let G be non-degenerate. Then $(W^\perp)^\perp = W$

Proof. First recall that $W \subseteq (W^\perp)^\perp$ as shown earlier. Next, observe that

$$\dim(W^\perp)^\perp = n - \dim W^\perp = n - (n - \dim W) = \dim W$$

This completes the proof. □

Corollary 2.3. Let G be non-degenerate. Let $W \subseteq V$ be a subspace of V , and let $\alpha \in W^*$. Then there exists $v \in V$ so that $\alpha(w) = G(v, w)$ for all $w \in W$. (Note that there could be multiple choices for V if $V \neq W$)

Proof. This is the same as saying $\iota^* \circ \sharp$ is surjective. Observe:

$$\begin{aligned} \dim \text{image } \iota^* \circ \sharp &= \dim \text{image } \sharp \circ \iota \\ &= \dim V - \dim \ker \sharp \circ \iota \\ &= \dim W - \dim W \cap V^\perp \\ &= \dim W \end{aligned}$$

□

Corollary 2.4. Let G be non-degenerate. Then $G|_W \in W^* \otimes W^*$ is nondegenerate iff $V = W \oplus W^\perp$.

Proof. Observe that $\text{rad}(G|_W) = W \cap W^\perp$.

Suppose $W \cap W^\perp = \{0\}$. Then $\dim W + \dim W^\perp \geq \dim V$. But

$$\dim W \oplus W^\perp = \dim W + \dim W^\perp - \dim W \cap W^\perp \geq \dim V$$

That is, $\dim W \oplus W^\perp = \dim V$. Since $W \oplus W^\perp$ is a subspace of V , we see that $W \oplus W^\perp = V$.

Now for the other direction. Suppose $V = W \oplus W^\perp$. Then

$$W \cap W^\perp = (W \cap W^\perp)^\perp = V^\perp = \{0\}$$

So $\text{rad}(G|_W) = 0$ as required. □

2.1.2 Restriction to Quotient

Definition 2.2 (Bilinear Form induced on Quotient). Let V be a vector space and let G be a symmetric or skew-symmetric bilinear form. Suppose $W \subseteq V^\perp$. Then we can define $\overline{G} \in (V/W)^* \otimes (V/W)^*$ by the formula,

$$\overline{G}([v_1], [v_2]) = G(v_1, v_2)$$

Lemma 2.2. The bilinear form \overline{G} is well defined.

Proof. Recall that for any $w_1, w_2 \in W$ we have $v_1 + w_1 \in [v_1]$ and $v_2 + w_2 \in [v_2]$, so we must verify that no matter what choice of w_1, w_2 we use we get the same result.

$$\begin{aligned} \overline{G}([v_1], [v_2]) &= G(v_1 + w_1, v_2 + w_2) \\ &= G(v_1, v_2) + \cancel{G(v_1, w_2)} + \cancel{G(w_1, v_2)} + \cancel{G(w_1, w_2)} \\ &= G(v_1, v_2) \text{ for all } w_1, w_2 \in W \end{aligned}$$

□

Theorem 2.3. $\overline{G} \in (V/W)^* \otimes (V/W)^*$ is non-degenerate if and only if $W = V^\perp$.

Proof.

$$\begin{aligned} \text{rad } \overline{G} &= \{[v] : \overline{G}([v], [u]) = 0 \ \forall [u] \in V/W\} \\ &= \{x \in V : G(x, y) = 0 \ \forall y \in V\} / W \\ &= V^\perp / W = \{0\} \text{ if and only if } V^\perp = W \end{aligned}$$

□

Remark. If G is non-degenerate, then W would have to be $\{0\}$ for the quotient bilinear form to be non-degenerate.

Physics Remark. This theorem is very important for symplectic reduction, which is a useful tool in classical mechanics as well as quantization theory.

2.1.3 Classification of Bilinear Forms: Symmetric Bilinear Forms Part 1

A useful kind of theorem we like in algebra is a classification theorem. This tells us all of the different kinds of a certain object, up to isomorphism. In this section we will be interested in classifying all symmetric bilinear forms.

Definition 2.3 (Homogenous Function). A function $f : V \rightarrow \mathbb{F}$ is said to be **homogenous** of degree n if $f(av) = a^n f(v)$.

Remark. Let $T : V \rightarrow V$ be a linear map and let B be a basis for V . Then the function $\det : L(V) \rightarrow \mathbb{F}$ is homogenous of degree $\dim V$. That is, $\det(aT) = a^{\dim V} \det(T)$.

Definition 2.4. Let $G : V \times V \rightarrow \mathbb{F}$ be a symmetric bilinear form. Let $Q : V \rightarrow \mathbb{F}$ be the function defined by $Q(v) = G(v, v)$. Then Q is called the **quadratic form** of G .

Remark. Note that $Q(av) = a^2 Q(v)$, so Q is homogenous of degree 2.

Theorem 2.4 (Polarization Identity). Given a quadratic form $Q(v)$ we can recover the symmetric bilinear form $G(v, w)$ by the formula,

$$G(v, w) = \frac{1}{2}(Q(v + w) - Q(v) - Q(w)) \quad (2.1)$$

Proof. Note that,

$$Q(v + w) = G(v + w, v + w) = G(v, v) + 2G(v, w) + G(w, w) \quad (2.2)$$

Rearranging this gives the desired result. □

Remark. Let $B = \{e_1, \dots, e_n\}$ is a basis for V and consider $v = x^i e_i$. Then $Q(v) = G_{ij} x^i x^j$. This is a homogenous symmetric polynomial of degree 2 in x^1, \dots, x^n . Any such polynomial defines a symmetric bilinear form.

Definition 2.5. Suppose $u \in V$ and that $G(u, u) = 0$. Then we say u is an **isotropic** vector.

Theorem 2.5. There exists an orthogonal basis $B = \{e_1, \dots, e_r, g_1, \dots, g_{n-r}\}$ of V so that $[G]_B = \text{diag}(s_1, \dots, s_r, 0, \dots, 0)$ where $r = \text{rank}(G)$ and s_1, \dots, s_r are nonzero.

Proof. Suppose $G = 0$. Then G is already in this form in every basis. So we will assume $G \neq 0$. Suppose $Q = 0$. Then $G(u, v) = \frac{1}{2}(Q(u + v) - Q(u) - Q(v)) = 0$ for all u, v . Therefore, we will assume $Q \neq 0$ as well. In other words, there exists some non-isotropic vector u . We will use this as our first basis vector. Let $e_1 = u$ and define $s_1 = G(e_1, e_1)$.

We will do the rest of the proof by induction. Suppose that for some $k \in \mathbb{N}$ there exists some linearly independent set $B_k = \{e_1, \dots, e_k\}$ so that $G(e_i, e_j) = \delta_{ij} s_i$ for $1 \leq i, j \leq k$. To begin, notice that we just proved this for $k = 1$. We therefore just have to prove that if the statement holds for some k , it also holds for $k + 1$.

Let $V_k = \text{span}\{e_1, \dots, e_k\}$. Then $[G|_{V_k}]_{B_k} = \text{diag}(s_1, \dots, s_k)$. So $\det[G|_{V_k}]_{B_k} \neq 0$. So $G|_{V_k}$ is non-degenerate. By Corollary 2.4 we have $V = V_k \oplus V_k^\perp$.

If $G|_{V_k^\perp} = 0$ then we are done, and $r = k$. Otherwise, if $G|_{V_k^\perp} \neq 0$ then there exists some $e_{k+1} \in V_k^\perp$ so that $G(e_{k+1}, e_{k+1}) = s_{k+1} \neq 0$. This completes the proof by induction. \square

Remark. The set $\{g_1, \dots, g_{n-r}\}$ forms a basis of $V^\perp = \text{rad}(G)$.

Remark. This theorem shows that given any symmetric $n \times n$ matrix $A \in M_{nn}(\mathbb{F})$ there exists an invertible P so that $P^T A P = \text{diag}(s_1, \dots, s_n, 0, \dots, 0)$. That is, we have just shown that every symmetric matrix is orthogonally diagonalizable.

Remark. Note that this is **not** a statement similar to the spectral theorem. The numbers s_1, \dots, s_r are not eigenvalues of $[G]$, and in fact the eigenvalues of $[G]$ depend on the choice of basis. The only thing meaningful about the numbers s_1, \dots, s_r is their *signs*.

Definition 2.6 (Orthonormal Basis). Suppose that we are working over an algebraically closed field such as $\mathbb{F} = \mathbb{C}$. Then every element of \mathbb{F} has a square root, so we can replace each e_i with $e_i/\sqrt{G(e_i, e_i)}$. This means we can always put $[G]_B$ in the form $\text{diag}(1, \dots, 1, 0, \dots, 0)$. This basis is called an **orthonormal** basis for G .

Corollary 2.5. Suppose G is a symmetric nondegenerate bilinear form over an algebraically closed field. Then there exists some basis $B = \{e_1, \dots, e_n\}$ for V so that $[G]_B = \mathbb{1}$.

Remark. Therefore, up to orthogonal transformations there is only one symmetric nondegenerate bilinear form.

Remark. The symmetric nondegenerate bilinear form over \mathbb{C} is **not** an inner product. It does **not** satisfy $G(u, v) = \overline{G(v, u)}$.

However, it is useful for defining complexified Clifford algebras.

Definition 2.7 (Metric). We often call a nondegenerate symmetric bilinear form a **metric**.

Definition 2.8 (Signature). Suppose we are working over $\mathbb{F} = \mathbb{R}$. Then we can always reorder the vectors $\{e_1, \dots, e_r\}$ so that s_1, \dots, s_p are positive and s_{p+1}, \dots, s_r are negative. Let $\tilde{e}_i = e_i/\sqrt{|s_i|}$ and set $\tilde{B} = \{\tilde{e}_1, \dots, \tilde{e}_r\}$. Then

$$[G]_{\tilde{B}} = \text{diag}(\underbrace{1, \dots, 1}_p, \underbrace{-1, \dots, -1}_{r-p}, \underbrace{0, \dots, 0}_{n-r}) \quad (2.3)$$

We define the **signature** of $[G]_{\tilde{B}}$ to be the pair of numbers describing how many positive and negative entries there are on the diagonal. That is,

$$\text{signature}(G) = (p, r - p) \quad (2.4)$$

Theorem 2.6 (Sylvester's Law of Inertia). Let G and \tilde{G} be symmetric bilinear forms and suppose there exists some bases B, \tilde{B} and a matrix P so that $P^T[G]_B P = [\tilde{G}]_{\tilde{B}}$. Then $\text{signature}([G]_B) = \text{signature}([\tilde{G}]_{\tilde{B}})$. In other words, signature is an **invariant** of a bilinear form.

Proof. Let $\{e_1, \dots, e_p, e_{p+1}, \dots, e_r, g_1, \dots, g_{n-r}\} = B$ be a basis for V and let $\{\tilde{e}_1, \dots, \tilde{e}_{\tilde{p}}, \tilde{e}_{\tilde{p}+1}, \dots, \tilde{e}_{\tilde{r}}, \tilde{g}_1, \dots, \tilde{g}_{n-\tilde{r}}\} = \tilde{B}$ be some other basis. Suppose that $[G]$ is diagonal in both bases. We want to show that the number of positive, negative, and zero entries on the diagonal of $[G]$ does not depend on which basis we use to diagonalize it.

First let $v = a^1 e_1 + \dots + a^p e_p$ and observe that $G(v, v) = \sum_{i=1}^p a_i^2 > 0$ if $v \neq 0$.

Similarly let $w = b^{\tilde{p}+1} \tilde{e}_{\tilde{p}+1} + \dots + b^{\tilde{r}} \tilde{e}_{\tilde{r}} + c^1 \tilde{g}_1 + \dots + c^{n-\tilde{r}} \tilde{g}_{n-\tilde{r}}$. Observe that for any vector of this form, $G(w, w) \leq 0$.

Then consider the following subspaces of V ,

$$U = \text{span}\{a^1 e_1 + \dots + a^p e_p\}$$

and

$$\tilde{U} = \text{span}\{b^{\tilde{p}+1} \tilde{e}_{\tilde{p}+1} + \dots + b^{\tilde{r}} \tilde{e}_{\tilde{r}} + c^1 \tilde{g}_1 + \dots + c^{n-\tilde{r}} \tilde{g}_{n-\tilde{r}}\}$$

Where we consider all possible values for the coefficients a^i, b^i, c^i . Since $G(v, v) > 0$ for all $0 \neq v \in U$ and $G(v, v) \leq 0$ for all $v \in \tilde{U}$, it is clear that U and \tilde{U} share no vectors except the zero vector. Therefore $\dim(U \cap \tilde{U}) = 0$.

This tells us that $n \geq \dim(U \oplus \tilde{U}) = \dim U + \dim \tilde{U} - \dim(U \cap \tilde{U})$. So $n \geq p + n - \tilde{p}$. So $p \leq \tilde{p}$.

Finally, if we repeat the above proof and swap all the variables in B with their counterpart in \tilde{B} and vice versa, we can show in the exact same way that $\tilde{p} \leq p$. So $p = \tilde{p}$ as required. \square

2.2 Lecture 6: Definiteness, Isometries and Adjoints

2.2.1 Classification of Bilinear Forms: Symmetric Bilinear Forms Part 2

Definition 2.9. Let V be a finite dimensional vector space over \mathbb{R} . Let G be a symmetric bilinear form and let $Q(v) = G(v, v)$.

- G is positive definite if $Q(v) > 0$ for all $v \in V$.
- G is negative definite if $Q(v) < 0$ for all $v \in V$.
- G is non-negative (positive semidefinite) if $Q(v) \geq 0$ for all $v \in V$.
- G is non-positive (negative semidefinite) if $Q(v) \leq 0$ for all $v \in V$.
- G is indefinite if there exists some $v, w \in V$ so that $Q(v) > 0, Q(w) < 0$.

By Sylvester's theorem, there is a basis $B = \{e_1, \dots, e_p, f_1, \dots, f_q, g_1, \dots, g_{n-p-q}\}$ where $[G]_B = \text{diag}(1, \dots, 1, -1, \dots, -1, 0, \dots, 0)$, with p ones, q minus ones, and $n - p - q$ zeros.

Remark. $G|_{\text{span}\{e_1, \dots, e_p\}}$ is positive definite and $G|_{\text{span}\{f_1, \dots, f_q\}}$ is negative definite.

Remark. G is positive definite if and only if $p = n$, negative definite if and only if $q = n$, and indefinite if and only if $q > 0$ and

$p > 0$. If $n - p - q > 0$ then G is degenerate. G is nonpositive iff $p = 0$, and G is nonnegative iff $q = 0$. If G is nondegenerate then the signature (p, q) is well defined.

Remark. The space $\text{span}\{e_1, \dots, e_p\}$ might depend on the choice of basis (ie. different bases can give different hyperplanes, but G is still positive definite on these hyperplanes). Only the dimension is invariant of the basis.

Example 2.1. Consider $V = \mathbb{R}^2$ and let $S = \{e_1, e_2\}$ be the standard basis. Let $v = ae_1 + be_2$. That is, $[v]_S = (a, b)$. But in the standard basis we usually just write $v = (a, b)$ since it is understood what basis we mean.

Set,

$$[G]_S = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Then,

$$Q(v) = [v]_S^T [G]_S [v]_S = \begin{bmatrix} a \\ b \end{bmatrix}^T \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = a^2 - b^2$$

That is, $Q(ae_1 + be_2) = a^2 - b^2$.

Now consider the vectors $u_1 = (1, 1)$ and $u_2 = (1, -1)$. Let $U_1 = \text{span}\{u_1\}$ and $U_2 = \text{span}\{u_2\}$. Let $\lambda \in \mathbb{R}$ be any real number. We can see that $Q(\lambda u_1) = \lambda^2(1 - 1) = 0$, and $Q(\lambda u_2) = \lambda^2(1 - 1) = 0$. Therefore $G|_{U_1} = 0$ and $G|_{U_2} = 0$.

Similarly, for any $v_1 = (1, \epsilon)$ with $-1 < \epsilon < 1$, we have $Q(\lambda v_1) > 0$. For $v_2 = (\epsilon, 1)$ with $-1 < \epsilon < 1$ we have $Q(\lambda v_2) < 0$. This allows us to draw a diagram of the regions where G is positive and negative definite.

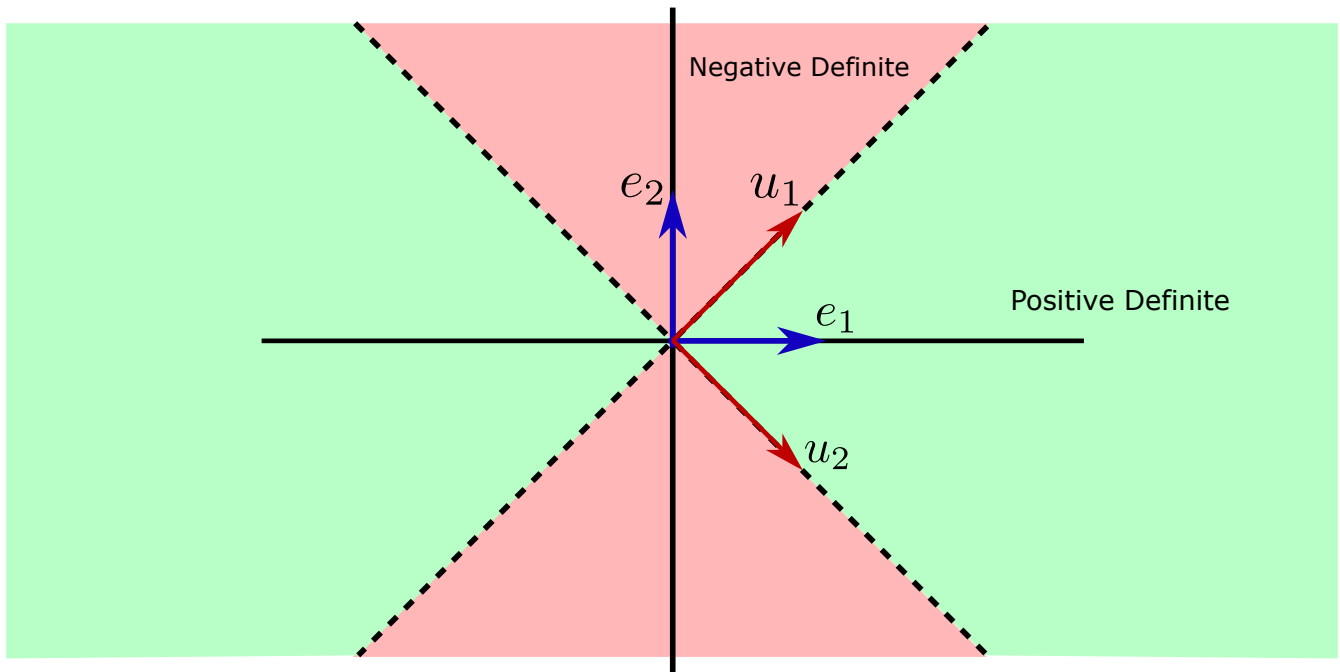


Figure 1: Diagram showing the regions where Q from example 2.1 is positive and negative definite, as well as the two lines where it is zero.

Remark. In example 2.1, we saw that the region where G is positive or negative definite, are not lines through \mathbb{R}^2 , and instead they form a kind of flat cone shaped region. Each region consists of infinitely many lines, on which G can be restricted in order

to get a positive or negative definite form. In summary, there is no unique subspace V_0 of \mathbb{R}^2 where $G|_{V_0}$ is positive definite, and the same can be said for negative definite.

2.2.2 Isometries

A nice way to classify bilinear forms is through their set of isometries. Isometries are transformations on V which preserve the bilinear form, and isometry of vector spaces forms an equivalence relation.

Definition 2.10 (Pullback of Bilinear Form). Let V be an n dimensional vector space over a field \mathbb{F} of characteristic zero. Let G be a nondegenerate symmetric bilinear form on V . Let W be some other vector space over \mathbb{F} and consider a linear map $P : W \rightarrow V$. Then we define the **pullback of G with respect to P** to be a bilinear form P^*G on W defined by

$$P^*G(w_1, w_2) = G(P(w_1), P(w_2)) \quad (2.5)$$

Definition 2.11 (Isometry). Let V_1, V_2 be two vector spaces, and let G_1, G_2 be nondegenerate symmetric bilinear forms on V_1 and V_2 respectively. Then a map $P : V_1 \rightarrow V_2$ is called an **isometry** if $P^*G_2 = G_1$.

Lemma 2.3. Let $\dim V_1 = \dim V_2$. Then any isometry between V_1 and V_2 is an isomorphism.

Proof. Let $v_1 \in \ker P$. Then $G_1(v_1, v_2) = G_2(P(v_1), P(v_2)) = 0$ for all v_2 . But since G_1 is nondegenerate, this is true for all v_2 iff $v_1 = 0$. So $v_1 = 0$ for all $v_1 \in \ker P$. Thus $\ker P = \{0\}$ and so we are done. \square

Definition 2.12. Let V_1, G_1 and V_2, G_2 be vector spaces with nondegenerate symmetric bilinear forms. We say V_1 is **isometric** to V_2 if there exists an isometry between them.

Remark. Isometry is an equivalence relation.

Definition 2.13 (Orthogonal Group). Let V be a vector space and let G be a metric on V . The set of all isometries $P : V \rightarrow V$ with $P^*G = G$ is a subgroup of $L(V)$ called the **orthogonal group** and we denote it $O(V, G)$.

$$O(V, G) = \{P \in L(V) : P^*G = G\} \quad (2.6)$$

Remark. If G is of signature (p, q) we often write this as $O(p, q, V)$.

If the vector space is $V = \mathbb{R}^n$, and G is of signature $(n, 0)$ we just write $O(n, \mathbb{R})$.

Definition 2.14. The **special orthogonal group** $SO(V, G)$ is the set of all orthogonal transformations $T \in O(V, G)$ with $\det T = 1$.

Remark. An isometry of V, G is sometimes called an orthogonal transformation.

Lemma 2.4. Let B be a basis of V . Then P is an isometry of G iff $[P]_B^T [G]_B [P]_B = [G]_B$.

Proof. Recall that $G(v, w) = [v]_B^T [G]_B [w]_B$. Then

$$\begin{aligned} G(P(v), P(w)) &= ([P]_B [v]_B)^T [G]_B [P]_B [w]_B \\ &= [v]_B [P]_B^T [G]_B [P]_B [w]_B \end{aligned}$$

So we need $[P]_B^T [G]_B [P]_B = [G]_B$. □

2.2.3 Adjoint

Lemma 2.5. Let V be a finite dimensional vector space over a field \mathbb{F} of characteristic 0. Let G be a symmetric or skew nondegenerate bilinear form on V . Let B be any other bilinear form on V . Then there exists a unique linear map $T_B : V \rightarrow V$ so that $B(v, w) = G(T_B(v), w)$ for all v, w .

Proof. Recall the flat map, $\flat_G : V \rightarrow V^*$ defined by $\flat_G(v)(w) = G(v, w)$. We know \flat_G is an isomorphism since G is nondegenerate. Let $\flat_B(v)(w) = B(v, w)$ as well. So $B(v, w) = G(T_B(v), w)$ iff $\flat_B(v)(w) = \flat_G(T_B(v))(w)$ for all w , if and only if $\flat_B = \flat_G \circ T_B$. So we just have to set $T_B = \flat_G^{-1} \circ \flat_B$. Since \flat_G is an isomorphism and hence invertible, this definition is perfectly good. □

Remark. So given a fixed nondegenerate bilinear form G on V we have an isomorphism ι between $V^* \otimes V^*$ and $V \otimes V^*$, given by $\iota(B) = T_B$ as above.

Definition 2.15 (Adjoint). Let G be a nondegenerate bilinear form on V (symmetric or skew). For any map $T : V \rightarrow V$ we define a linear map $T^\dagger : V \rightarrow V$ by the formula $G(v, T(w)) = G(T^\dagger(v), w)$ for all $v, w \in V$. T^\dagger is called the **adjoint** of T .

Remark. T^\dagger is not the same as the dual map T^* . But they are related by a formula.

Lemma 2.6. The map T^\dagger is the only map satisfying $G(v, T(w)) = G(T^\dagger(v), w)$ for all v, w .

Proof. Let $B \in V^* \otimes V^*$ where $B(v, w) = G(v, Tw)$. Then there is a unique map $T_B : V \rightarrow V$ given by $B(v, w) = G(T_B(v), w)$. Therefore we have $G(T_B(v), w) = G(v, T(w))$ so $T_B = T^\dagger$ as required. □

Remark. Since G is either skew or symmetric, $G(v, Tw) = G(T^\dagger v, w)$ implies that $\pm G(T(w), v) = \pm G(w, T^\dagger V)$. So $G(T(v), w) = G(v, T^\dagger(w))$. Therefore T satisfies the same definition as $(T^\dagger)^\dagger$. Since adjoints are unique, $T = (T^\dagger)^\dagger$.

It also follows that $(T \circ S)^\dagger = S^\dagger \circ T^\dagger$.

Remark. The dual map and the adjoint are related. Observe, $G(v, T(w)) = G(T^\dagger(v), w)$, so $\flat_G(v)(T(w)) = \flat_G(T^\dagger(v))(w)$. But $\flat_G(v)(T(w)) = T^*\flat_G(v)(w)$, so $T^*(\flat_G) = \flat_G \circ T^\dagger$. Hence,

$$T^\dagger = \flat_G^{-1} \circ T^*(\flat_G) = \sharp_G \circ T^* \circ \flat_G \quad (2.7)$$

Definition 2.16 (Self/Skew-Adjoint). A map $T : V \rightarrow V$ is called self-adjoint if $T^\dagger = T$. It is called skew-adjoint if $T = -T^\dagger$.

Remark. Let $T \in L(V)$. Then we can write $T = \frac{1}{2}(T + T^\dagger) + \frac{1}{2}(T - T^\dagger)$. Observe that $\frac{1}{2}(T + T^\dagger)$ is self adjoint and $\frac{1}{2}(T - T^\dagger)$ is skew-adjoint!

Lemma 2.7. B is symmetric iff T_B is self-adjoint, and B is skew iff T_B is skew-adjoint.

Proof. Let $B \in V^* \otimes V^*$ and let $T_B \in \text{End}(V)$. We have $B(v, w) = G(T_B(v), w) = G(v, T_B^\dagger(w)) = \pm G(T^\dagger(w), v)$. Suppose G is symmetric. Then $B(v, w) = G(T_B^\dagger(w), v) = G(T_B(w), v)$ for all v, w . Since G is nondegenerate, we have $T_B^\dagger(w) = T_B(w)$ for all w . This means T_B is self adjoint.

For the other direction, we see that if T_B is self adjoint then $B(v, w) = G(T_B(v), w) = G(w, T_B(v)) = B(w, v)$.

If we did this proof again for skew G , then we would have a minus sign everywhere. Therefore the proof for skew-adjoint T_B is identical. \square

2.3 Lecture 7: Reflections, Hyperbolic Planes, and Cancellation

2.3.1 Adjoint Basis

For all of this lecture, V is a finite dimensional vector space over a field \mathbb{F} of characteristic zero, and G is a symmetric bilinear form.

Recall that $\flat : V \rightarrow V^*$ and $\flat(v)(w) = G(v, w)$. Let $B = \{e_1, \dots, e_n\}$ and $B^* = \{e^1, \dots, e^n\}$. Let $\flat(e_i) = A_{ij}e^j$. Observe that $\flat(e_i)(e_j) = G_{ij} = A_{ij}$. So $[\flat]_{B^*B} = [G]_B$.

Definition 2.17 (G -adjoint Basis). The set $\{\flat(e_i) : i = 1, \dots, n\}$ is a basis for V^* called the G -adjoint basis. We denote it by B^\dagger .

Remark. If B is an orthonormal basis then $B^* = B^\dagger$. In this basis $[T]_B^\dagger = [T]_B^T$.

2.3.2 Classification of Bilinear Forms: Symmetric Bilinear Forms Part 3

The problem is to classify all symmetric nondegenerate bilinear forms (metrics) on V . Given a vector space V , can we classify all possible metrics up to isometry (that is, we only say two metrics are different if we can not transform one into another via an isometry). This classification will be based off the proof in the book Basic Algebra I by Jacobson [8].

Definition 2.18 (Reflection). Let v be a vector which is not isotropic, so $Q(v) \neq 0$. Then we define a linear map $R_v : V \rightarrow V$ by $R_v(u) = u - 2\frac{G(u,v)}{Q(v)}v$. This map is called a **reflection generated by v**

Remark. $R_{tv} = R_v$ for all $t \in \mathbb{R}$

Lemma 2.8. The map R_v is an isometry.

Proof. By the polarization identity, the map R_v is an isometry iff $Q(R_v(u)) = Q(u)$ for all $u \in V$. We have

$$\begin{aligned} Q(R_v(u)) &= Q\left(u - 2\frac{G(u,v)}{Q(v)}v\right) \\ &= G\left(u - 2\frac{G(u,v)}{Q(v)}v, u - 2\frac{G(u,v)}{Q(v)}v\right) \\ &= Q(u) - 4\frac{G(u,v)G(v,u)}{Q(v)} + 4\frac{G(u,v)^2G(v,v)}{Q(v)} \\ &= Q(u) \end{aligned}$$

Where we have used the fact that G is symmetric to cancel out the last two terms. □

Remark. Let $u \in \text{span}\{v\}^\perp$. Then $G(u,v) = 0$, so $R_v(u) = u$.

Remark. Suppose $u = v$. Then $R_v(u) = v - 2v = -v$.

Remark. Based on the above two facts, we observe that R_v is a reflection over the hyperplane $\text{span}\{v\}^\perp$.

Example 2.2. Let $V = \mathbb{R}^3$ and let G be the standard inner product, $G(v,w) = v^1w^1 + v^2w^2 + v^3w^3$. We have the following picture.

Remark. Let $e_1 = v$ and let e_2, \dots, e_n be a basis for $\text{span}\{e_1\}^\perp$. Then $B = \{e_1, \dots, e_n\}$ is a basis for V and $[R]_B = \text{diag}(-1, 1, \dots, 1)$. So $\det R_v = -1$.

Definition 2.19. Let $T : V \rightarrow V$ be an isometry. Recall that since T is an isomorphism, $\det T \neq 0$. We say that T is **proper** if $\det T > 0$ and **improper** if $\det T < 0$.

Remark. Reflections are improper isometries.

Lemma 2.9. Let P be an isometry. Then PR_vP^{-1} is a reflection, and in fact $PR_vP^{-1} = R_{Pv}$.

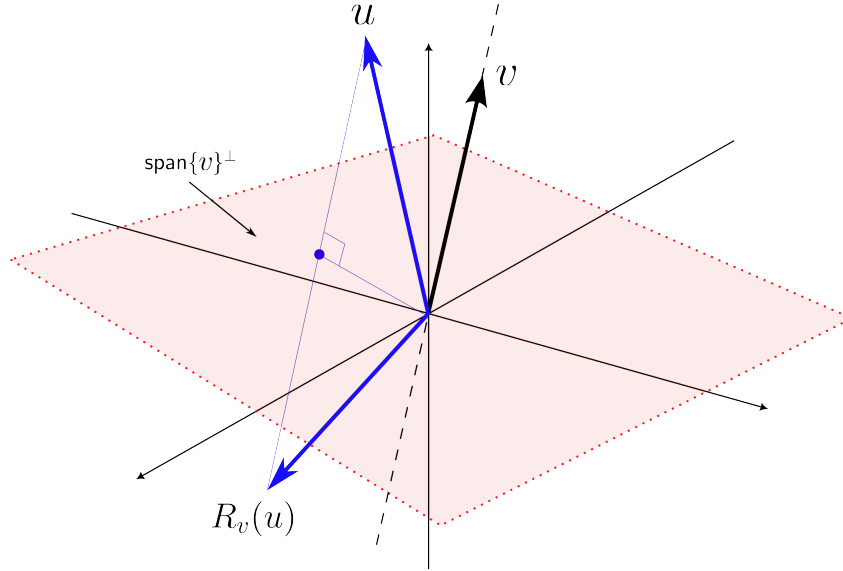


Figure 2: Diagram of a reflection in \mathbb{R}^3 using the standard inner product.

Proof.

$$\begin{aligned}
 PR_vP^{-1}(u) &= P\left(P^{-1}(u) - \frac{2G(P^{-1}(u), v)v}{Q(v)}\right) \\
 &= u - \frac{2G(P^{-1}(u), P^{-1}(P(v)))}{Q(P(v))}P(v) \\
 &= u - \frac{2G(u, P(v))}{Q(P(v))}P(v) \\
 &= R_{P(v)}(u)
 \end{aligned}$$

Where we have used the fact that P is an isometry to say $Q(P(v)) = Q(v)$ and $G(P^{-1}(x), P^{-1}(y)) = G(x, y)$. \square

Definition 2.20 (T -invariant Subspace). Let $T : V \rightarrow V$ be linear and let U be a subspace of V . We say U is T -invariant if $T(U) \subseteq U$.

Lemma 2.10. Let U be a T -invariant subspace. Then U^\perp is a T^\dagger -invariant subspace.

Proof. Let $v \in U^\perp$. This means that $G(u, v) = 0$ for all $u \in U$. Therefore $G(u, T^\dagger(v)) = G(T(u), v) = 0$ for all v since, because U is T -invariant, we know $T(u) \in U$. In other words, we have $G(u, T^\dagger(v)) = 0$ for all $u \in U$, which means by definition that $T^\dagger(v) \in U^\perp$. \square

Corollary 2.6. Let $T \in O(V, G)$. If $T(U) \subseteq U$, then $T(U^\perp) = U^\perp$.

Proof. We know $T(U) \subseteq U$. Since T is injective, $\dim T(U) = \dim U$. Therefore $T(U) = U$ and $U = T^{-1}(U)$. Therefore $T^\dagger(U) = U$. Applying the lemma, we have $(T^\dagger)^\dagger(U^\perp) \subseteq U^\perp$, so $T(U^\perp) \subseteq U^\perp$, and because T is injective, we furthermore have $T(U^\perp) = U^\perp$. \square

Remark. Note that we have not said anything about whether $G|_U$ is nondegenerate or not.

Definition 2.21. We say V is an **orthogonal direct sum** of the subspaces $U_1, \dots, U_k \subseteq V$ if $V = \bigoplus_{i=1}^k U_i$ and $U_i \perp U_j$ for all $i \neq j$. Notationally we write $V = U_1 \oplus U_2 \oplus \dots \oplus U_k = \bigoplus_{i=1}^k U_i$.

Remark. The above definition is about decomposing a vector space into a direct sum of smaller, mutually orthogonal vector spaces.

Remark. Notice that for each j , we have $U_j^\perp = \bigoplus_{i=1, i \neq j}^k U_i$.

Remark. Suppose that $V = \bigoplus_{i=1}^k U_i$, which means that any vector v can be written as $v = \sum_{i=1}^k u_i$, where $G(u_i, u_j) = 0$ if $i \neq j$. Then $Q(v) = \sum_{i=1}^k Q(u_i)$.

Lemma 2.11. Suppose that $V = \bigoplus_{i=1}^k \tilde{U}_i$, and $V = \bigoplus_{i=1}^k U_i$, and let $v = \sum_{i=1}^k u_i$ and $v = \sum_{i=1}^k \tilde{u}_i$. Suppose for each $i = 1, \dots, k$ we have an isometry $P_i : U_i \rightarrow \tilde{U}_i$. Define $P : V \rightarrow V$ by $P(\sum_{i=1}^k u_i) = \sum_{i=1}^k P_i u_i$. Then P is an isometry.

Remark.

$$\begin{aligned} Q(P(v)) &= Q\left(\sum_{i=1}^k P_i(v_i)\right) \\ &= \sum_{i=1}^k Q(P_i(v_i)) \\ &= \sum_{i=1}^k Q(v_i) \\ &= Q(v) \end{aligned}$$

So we can build isometries of V out of isometries on the orthogonal components of V .

2.3.3 Hyperbolic Planes

We will take a detour here to talk about hyperbolic planes.

Definition 2.22. Let U be a vector subspace of V . Then U is said to be isotropic if $Q|_U = 0$, or equivalently $U \subseteq U^\perp$.

Remark. An isotropic vector generates an isotropic vector subspace. That is, if v is isotropic then $\text{span}\{v\}$ is an isotropic subspace.

If $\mathbb{F} = \mathbb{R}$ and G is positive definite or negative definite then there are no isotropic subspaces.

Definition 2.23. Let G be a metric on V . We say that V is a **hyperbolic plane** if $\dim V = 2$ and V contains an isotropic vector.

Theorem 2.7. Let $\dim V = 2$, and let G be a metric. The following are equivalent.

- i) V is a hyperbolic plane.
- ii) There exists a basis $B = \{u, v\}$ of V so that $Q(u) = 0$, $Q(v) = 0$, and $G(u, v) = G(v, u) = 1$. That is,

$$[G]_B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Such a basis is called a **hyperbolic basis**

- iii) For all bases B of V , $\det[G]_B = -x^2$ for some $x \in \mathbb{F}$.

Proof. First we show that i implies ii. First, there exists u so that $Q(u) = 0$. Then by nondegeneracy there is some v so that $G(u, v) \neq 0$. Then since $G(u, u) = 0$, $v \neq u$, so $\{u, v\}$ is a linearly independent set and hence a basis. Without loss of generality we may set $G(u, v) = 1$ by rescaling v . Then let $a \in \mathbb{F}$, and set $\tilde{v} = v + au$. Then $Q(\tilde{v}) = Q(v) + 2a$. Choose $a = -\frac{1}{2}Q(v)$. So $Q(\tilde{v}) = 0$ and $\{u, \tilde{v}\}$ is linearly independent, and therefore is a basis satisfying the requirements of ii.

Now we show that ii implies iii. Let B be the aforementioned basis. Then

$$[G]_B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

which has determinant -1 as required. If \tilde{B} is another basis then $[G]_{\tilde{B}} = -\det(P_{B\tilde{B}})^2$ which is of the form $-x^2$ as required. Finally we show that iii implies i. Suppose $\det[G]_B = -x^2, x \neq 0$ for all bases B . This means G is nondegenerate, and furthermore there exists a basis $B = \{e_1, e_2\}$ so that $[G]_B = \text{diag}(b_1, b_2)$. This means $b_1 b_2 = -x^2$ for some x . Now let $w = c_1 e_1 + c_2 e_2$. Then

$$Q(w) = c_1^2 b_1 + c_2^2 b_2$$

Then let $c_1 = a$ and $c_2 = b_2$, so then

$$Q(w) = a^2 b_1 + b_1^2 b_2 = b_1(a^2 + b_1 b_2) = b_1(a^2 - a^2) = 0$$

So w is an isotropic vector, which is what we needed to find. This completes the proof. □

Corollary 2.7. Any two hyperbolic planes are isometric.

Proof. Let V, G and \tilde{V}, \tilde{G} be two hyperbolic planes. Then there are bases $B = \{e_1, e_2\}, \tilde{B} = \{\tilde{e}_1, \tilde{e}_2\}$ so that

$$[G]_B = [\tilde{G}]_{\tilde{B}} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Let $P : V \rightarrow \tilde{V}$ so that $P(e_1) = \tilde{e}_1, P(e_2) = \tilde{e}_2$, and extend the definition by linearity. Then $P^T[G]_B P = [\tilde{G}]_{\tilde{B}}$ which means P is an isometry as required. \square

Theorem 2.8. Any hyperbolic plane V contains exactly two one-dimensional isotropic subspaces.

Proof. Let $\{u, v\} = B$ be a hyperbolic basis for V . Then let $w \in V$. Then $w = au + bv$ and $Q(w) = 2ab$. This is zero iff $a = 0$ or $b = 0$. Therefore $w = au$ or $w = bv$. Thus w is isotropic iff $w \in \text{span}\{u\}$ or $w \in \text{span}\{v\}$. These are the two isotropic subspaces. \square

Remark. The figure presented earlier, figure 1, shows the isotropic subspaces for the choice of metric,

$$[G]_S = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Any hyperbolic plane metric can be put in this form, if we use the basis $e_1 = \frac{u+v}{4G(u,v)}$ and $e_2 = \frac{u-v}{4G(u,v)}$. Therefore, figure 1 provides a decent visual of what the isotropic subspaces look like. In general, we get the same picture but warped by some shear/rotation.

Definition 2.24. Let \mathbb{F} be a field. The **group of units** of \mathbb{F} is the set of all invertible elements of \mathbb{F} . That is,

$$\mathbb{F}^* = \{a \in \mathbb{F} : \exists b \in \mathbb{F}, ba = ab = 1\}. \quad (2.8)$$

Theorem 2.9. Let V, G be a hyperbolic plane. Then,

- a) Any improper isometry is a reflection.
- b) The special orthogonal group of G is isomorphic to the group of units of \mathbb{F} . That is, $\text{SO}(V, G) = \mathbb{F}^*$

Proof. Let $P \in \text{O}(V, G)$ be any orthogonal transformation and let $\{u, v\}$ be a hyperbolic basis. Then P maps isotropic subspaces to isotropic subspaces since $Q(P(v)) = Q(v)$. So either P takes $\text{span}\{u\}$ to $\text{span}\{u\}$ and $\text{span}\{v\}$ to $\text{span}\{v\}$ or it exchanges them. In the first case, $P(u) = au$ and $P(v) = bv$ for some $a, b \in \mathbb{F}$. Therefore $[P]_B = \text{diag}(a, b)$. So

$1 = G(u, v) = G(P(u), P(v)) = G(u, v)ab = ab$. Therefore $b = a^{-1}$, meaning $[P]_B = \text{diag}(a, a^{-1})$ and $\det[P]_B = 1$. So $P \in \text{SO}(V, G)$. Therefore every isometry which leaves the two isotropic vector spaces where they are is an element of $\text{SO}(V, G)$. In the other case, we have $P(u) = av$ and $P(v) = bu$. So $G(u, v) = ab$, and

$$[P]_B = \begin{bmatrix} 0 & a^{-1} \\ a & 0 \end{bmatrix}$$

so $\det[P]_B = -1$. This is not an element of $\text{SO}(V, G)$, but it is a reflection. Observe that $P(u + av) = av + aa^{-1}u = u + av$ and $P(u - av) = av - u = -(u - av)$. Since $G(u + av, u - av) = 0 + a - a + 0 = 0$, we finally have $P = R_{u-av}$. This proves part a).

The previous two points imply that $\text{SO}(V, G)$ consists entirely of isometries which leave the two isotropic subspaces where they are.

Finally, let $\varphi : \text{SO}(V, G) \rightarrow \mathbb{F}^*$ be defined by $\varphi(P) = \varphi(\text{diag}(a, a^{-1})) = a$. This is well defined because we just showed that the only elements of $\text{SO}(V, G)$ are those which can be put in the form $\text{diag}(a, a^{-1})$. This is a group isomorphism since $\varphi^{-1}(a) = \text{diag}(a, a^{-1})$ is defined and

$$\begin{aligned} \varphi(AB) &= \varphi(\text{diag}(a, a^{-1}) \text{diag}(b, b^{-1})) \\ &= \varphi(\text{diag}(ab, (ab)^{-1})) = ab \\ &= \varphi(A)\varphi(B) \end{aligned}$$

□

2.3.4 Witt's Cancellation Theorem

Definition 2.25. Let U be a subspace of V , and let G be nondegenerate. If $G|_U$ is nondegenerate we say U is a **nondegenerate subspace**.

Theorem 2.10 (Witt Cancellation Theorem). Let V, G be a vector space over a field \mathbb{F} of characteristic zero where G is a metric. Suppose U_1, U_2 are nondegenerate subspaces of V . If U_1 is isometric to U_2 then U_1^\perp is isometric to U_2^\perp .

Proof. If U_1, U_2 are nondegenerate subspaces then $V = U_1 \oplus U_1^\perp$ and $V = U_2 \oplus U_2^\perp$. So what we are trying to show is that if U_1 and U_2 are isometric, we can "cancel" them on both sides of the equation $U_1 \oplus U_1^\perp = U_2 \oplus U_2^\perp$.

First observe that $G|_{U_1}$ and $G|_{U_2}$ are both nondegenerate.

Now, let U_1 and U_2 be isometric. Then first of all, $\dim U_1 = \dim U_2 = k$. We will prove this by using mathematical induction on k . Let us begin with $k = 1$ as the base case.

In the base case, $U_1 = \text{span}\{u_1\}$, $U_2 = \text{span}\{u_2\}$, and $Q(u_1) \neq 0, Q(u_2) \neq 0$. By scaling u_2 appropriately, let $Q(u_1) = Q(u_2)$. Then

$$Q(u_1 \pm u_2) = Q(u_1) \pm 2G(u_1, u_2) + Q(u_2) = 2Q(u_1) \pm 2G(u_1, u_2)$$

We claim that either $Q(u_1 + u_2) \neq 0$ or $Q(u_1 - u_2) \neq 0$ (but not both!)

If they are both zero, then we could expand then add them together and get $Q(u_1) = Q(u_2) = 0$.

If we know that $Q(u_1 + u_2) \neq 0$, then we also know that $G(u_1 + u_2, u_1 - u_2) = Q(u_1) - Q(u_2) = 0$, so $u_1 + u_2 \perp u_1 - u_2$.

Furthermore, $R_{u_1+u_2}(u_1 - u_2) = u_1 - u_2$, and so

$$R_{u_1+u_2}(u_1 - u_2) + R_{u_1+u_2}(u_1 + u_2) = R_{u_1+u_2}(2u_1) = -2u_2$$

This means that $R_{u_1+u_2}$ takes $\text{span}\{u_1\}$ to $\text{span}\{u_2\}$ and vice versa, and also that $R_{u_1+u_2}$ takes $\text{span}\{u_1\}^\perp$ to $\text{span}\{u_2\}^\perp$.

Therefore $R_{u_1+u_2}$ is an isometry between U_1^\perp and U_2^\perp as required.

Now suppose instead that we know $Q(u_1 - u_2) \neq 0$. Then we could take $R_{u_1-u_2}$ and find $R_{u_1-u_2}(u_1) = u_2$, and then get the same result. This finishes the base case.

See Jacobson chapter 6 for the inductive step [8]. □

Remark. Witt's Cancellation theorem, along with lemma 2.11 hints at the fact that if we have an isometry between nondegenerate subspaces U_1 and U_2 , then we can extend it to an isometry from V to V . This will be proven at the beginning of next lecture.

2.4 Lecture 8: Witt Extension Theorem, Maximally Isotropic Subspaces, and the Cartan-Dieudonné Theorem

2.4.1 Witt Extension Theorem

Lemma 2.12. Let V, G be a vector space with metric, and let $U \subseteq V$ be a degenerate subspace. That is, a subspace with $\text{rad}(G|_U) \neq \{0\}$. Let U' be any subspace with the property that $V = U' \oplus \text{rad}(G|_U)$. Let $r = \dim \text{rad}(G|_U)$. Then there exist hyperbolic planes H_1, \dots, H_r and a subspace $W \subseteq V$ so that $U \oplus W = U' \oplus H_1 \oplus \dots \oplus H_r$.

Remark. This lemma says that if we have a degenerate subspace U , there is always some other subspace W we can add it to and get a nondegenerate subspace. Furthermore, the sum of these two subspaces decomposes into a nice nondegenerate space U' along with a bunch of hyperbolic planes H_i , which are easy to work with.

Proof. Let $\{u_1, \dots, u_r\}$ be a basis of $\text{rad}(G|_U)$. Let $\alpha \in U^*$ be defined so that $\alpha(u_1) = 1$ and $\alpha(u_i) = 0$ for $i = 2, \dots, r$ and $\alpha(U') = 0$.

Then since G is nondegenerate on V there exists some $v_1 \in V$ so that $\alpha(u) = G(u, v_1)$ for all u . So, there is some v_1 with $G(u_1, v_1) = 1$, $G(u_i, v_1) = 0$ for $i = 2, \dots, r$, and $G(u, v_1) = 0$ for all $u \in U'$.

Note that v_1 is not an element of U . Then replace v by $\tilde{v}_1 = v_1 + \sigma u_1$, for some $\sigma \in \mathbb{F}$.

Then $Q(\tilde{v}_1) = Q(v_1) + 2\sigma$. So we set $\sigma = -Q(v_1)/2$, to get $\{u_1, \tilde{v}_1\}$ to span a hyperbolic plane. So we set $H_1 = \text{span}\{u_1, \tilde{v}_1\}$.

Furthermore, by definition of \perp we have $V = H_1 \oplus H_1^\perp$. Now let $\tilde{U} = U' \oplus \text{span}\{u_2, \dots, u_r\} \subseteq \text{span}\{u_1, v_1\}^\perp$.

Note that $\text{rad}(G|_{\tilde{U}}) = \text{span}\{u_2, \dots, u_r\}$. If $r = 1$ we have $U' = \tilde{U}$ and we can let $W = \text{span}\{v_1\}$ so that $U \oplus W = U' \oplus H_1$.

Otherwise, set $U, V = \tilde{U}, \tilde{V}$ in the above argument and repeat. Inductively, we will get r hyperbolic planes H_1, \dots, H_r until finally we have $U \oplus W = U' \oplus H_1 \oplus \dots \oplus H_r$ as required. We also verify that $U \oplus W$ is nondegenerate, since everything in the

big orthogonal direct sum is. □

Remark. This proof is quite inscrutable. The idea is that, since U is degenerate, we can find an isotropic vector u_1 in U . This vector gives us the first basis vector for the hyperbolic plane. Then we find another vector, \tilde{v}_1 which is also isotropic, and the span of the two isotropic vectors gives us the first hyperbolic plane H_1 . Then, we check whether the complement of H_1 in U is degenerate. If not, we are done and W is just the span of v_1 . Otherwise, since the complement is degenerate we can find another isotropic vector and repeat the process.

The main idea is to find all of the linearly independent isotropic vectors in U , and pair each of them up with an isotropic vector in the larger space V so that their span is non degenerate. To do this, we have to make U into a bigger subspace, which is $U \oplus W$.

Theorem 2.11 (Witt Extension Theorem). Let U_1, U_2 be subspaces of V and let $P : U_1 \rightarrow U_2$ be an isometry. Then there exists an isometry $P : V \rightarrow V$ so that $P|_{U_1} = \tilde{P}$.

Proof. If U_1 is nondegenerate, this is a direct consequence of Theorem 2.10 and Lemma 2.11.

If U_1 is degenerate, then there exists some W_1 so that $U_1 \oplus W_1 = U'_1 \oplus H_1 \oplus \dots \oplus H_r$, where U'_1 has the property that $U_1 = \text{rad}(G|_{U_1}) \oplus U'_1$. Let $\tilde{P} = P|_{U_1}$. Then \tilde{P} maps U_1 to U_2 , and $\tilde{P}|_{\text{rad}(G|_{U_1})}$ is an isomorphism between $\text{rad}(G|_{U_1})$ and $\text{rad}(G|_{U_2})$. We then set $U'_2 = P(U'_1)$ which means there exists W_2 so that $U_2 \oplus W_2 = P(U'_1) \oplus H_1 \oplus \dots \oplus H_r = U'_2 \oplus \tilde{H}_1 \oplus \dots \oplus \tilde{H}_r$. Here, $H_i = \text{span}\{v_i, u_i\}$ and $\tilde{H}_i = \text{span}\{\tilde{v}_i, P(u_i)\}$.

Now we define $L : U_1 \oplus W_1 \rightarrow U_2 \oplus W_2$ by $L|_{U_1} = P$ and $L|_{W_1}(v_i) = \tilde{v}_i$ for $i = 1, \dots, r$. Since $v_i \in W_1$, $\tilde{v}_i \in W_2$, and $U_1 \oplus W_1$ is nondegenerate, this is an isometry by construction. Since $U_1 \oplus W_1$ is nondegenerate, we then see that by Theorem 2.10 and Lemma 2.11, L can be extended to all of V . Since L is in turn an extension of P , we are done. □

Remark. A useful application of this theorem is the Witt index.

2.4.2 Maximally Isotropic Subspaces

Remark. Recall that U_1 is an isotropic subspace if $G|_{U_1} = 0$. If U_2 is another isotropic with $\dim U_1 = \dim U_2$ then any isomorphism between U_1 and U_2 is an isometry, and can be extended to V to get an isometry.

Definition 2.26 (Maximally Isotropic Subspace). An isotropic subspace $U \subseteq V$ is said to be **maximal** if whenever we have another isotropic space W with $U \subseteq W \subseteq V$, then $U = W$.

Lemma 2.13. If U is maximally isotropic and \tilde{U} is isotropic then there is an isometry $P : V \rightarrow V$ with $P(\tilde{U}) \subseteq U$.

Proof. Suppose $\dim \tilde{U} < \dim U$. Then there is a subspace of U of dimension $\dim \tilde{U}$, and therefore an isomorphism between \tilde{U} and a subspace of U . Since any isomorphism between isotropic subspaces is an isometry, this case is done.

If $\dim \tilde{U} \geq \dim U$, then there is an isometry P with $P(U) \subseteq \tilde{U}$, which means $P^{-1}(\tilde{U})$ is isotropic. This means $U \subseteq P^{-1}(\tilde{U}) \subseteq V$. Since U is maximal, $P^{-1}(\tilde{U}) = U$, so $\tilde{U} = P(U)$. This completes the proof for all possible cases. \square

Corollary 2.8. All maximally isotropic subspaces of V are the same dimension.

Proof. If U and \tilde{U} are maximally isotropic, then $\dim U \leq \dim \tilde{U}$ by the previous lemma, and vice versa, so $\dim U = \dim \tilde{U}$. \square

Definition 2.27 (Witt Index). Let U be a maximally isotropic subspace. Then we define $\nu(V) = \dim U$ to be the **Witt index** of V . Notice that it does not depend on the choice of maximally isotropic subspace.

Lemma 2.14. Let $\dim V = n$. Then $\nu(V) \leq \lfloor \frac{n}{2} \rfloor$.

Proof. Let U be a maximally isotropic subspace of V . Then U can be extended into $U \oplus W = \bigoplus_{i=1}^r \perp H_i$, where $U' = \{0\}$ since U is maximally isotropic. Since $\dim H_i = 2$, we have $\dim U \oplus W = \nu(V) + \dim W = 2r \leq n$. So $r \leq \frac{n}{2}$. Since r is an integer, we furthermore have $r \leq \lfloor \frac{n}{2} \rfloor$. Since U is isotropic, $\text{rad}(G|_U) = U$, so $r = \dim U$. This completes the proof. \square

Definition 2.28 (Anisotropic Kernel). Let U be maximally isotropic. Then there is some W so that $U \oplus W$ is nondegenerate and $V = (U \oplus W) \oplus (U \oplus W)^\perp$, where $(U \oplus W)^\perp$ is anisotropic. The space $(U \oplus W)^\perp$ is called an **anisotropic kernel** of G .

Remark. Let $\mu \leq \nu$. Suppose we had the following decompositions of V , where X and \tilde{X} are anisotropic subspaces,

$$V = H_1 \oplus \dots \oplus H_\nu \oplus X$$

$$V = \tilde{H}_1 \oplus \dots \oplus \tilde{H}_\mu \oplus \tilde{X}$$

Let $\{\tilde{u}_i, \tilde{v}_i\}$ be a hyperbolic basis for \tilde{H}_i for each $i = 1, \dots, \mu$. Then $\text{span}\{\tilde{u}_1, \dots, \tilde{u}_\mu\}$ is isotropic, and $\mu \leq \nu$, so there is an isometry between $H_1 \oplus \dots \oplus H_\mu$ and $\tilde{H}_1 \oplus \dots \oplus \tilde{H}_\mu$. By the extension theorem we can extend this to an isometry $H_{\mu+1} \oplus \dots \oplus H_\nu \oplus X \rightarrow \tilde{X}$. But \tilde{X} is anisotropic, so $\mu = \nu$ and $\tilde{X} = X$. So the study of metrics reduces up to isometry to the study of anisotropic metric vector spaces. For general \mathbb{F} this becomes complicated.

2.4.3 Classification of Bilinear Forms: Symmetric Bilinear Forms Part 4, Cartan-Dieudonné theorem

Theorem 2.12 (Cartan-Dieudonné Theorem). Let \mathbb{F} be a field of characteristic not equal to 2. Let G be a metric on V , and let $\dim V = n$. Let P be an isometry of V . Then P is the product of $k \leq n$ reflections.

Proof. We prove this by induction on n .

First suppose $n = 1$ for the base case. Let $\{v\}$ be a basis of V . Then $Pv = \lambda v$ for some λ , so $Q(P(v)) = \lambda^2 Q(v)$. So $\lambda = \pm 1$. In the positive case $P = \mathbb{1}$ and in the negative case $P = R_{-v}$. So P is either a product of zero or one reflections as required. The inductive step splits into three cases.

1. Let $V_1 = \{v \in V : Pv = v\}$. Suppose V_1 is anisotropic. Then there exists $u \neq 0$ so that $Q(u) \neq 0$ and $P(u) = u$. So P maps $\text{span}\{u\}^\perp$ to $\text{span}\{u\}^\perp$, and so $P|_{\text{span}\{u\}} = \mathbb{1}$, so $P = \mathbb{1}_{1 \times 1} \oplus P|_{\text{span}\{u\}^\perp}$. So the problem reduces to $\text{span}\{u\}^\perp$ which is $n - 1$ dimensional. If it is anisotropic, we repeat the procedure. Otherwise we go to another one of the cases, and try to show that $P|_{\text{span}\{v\}^\perp}$ and hence P is the product of at most $n - 1$ reflections.
2. Suppose there exists $u \neq 0$ so that $Q(v) \neq 0$ and $Q(v - P(v)) \neq 0$. Since P is an isometry, $Q(P(v)) = Q(u) \neq 0$. So

$$\begin{aligned} R_{v-P(v)}(v + P(v)) &= v + P(v) - \frac{G(v + P(v), v - P(v))(v - P(v))}{Q(v - P(v))} \\ &= v + P(v) - \frac{Q(v) - Q(P(v))}{Q(v - P(v))}(v - P(v)) \\ &= v + P(v) \end{aligned}$$

Similarly, $R_{v-P(v)}(v - P(v)) = -(v - P(v))$. From this, we see that

$$\begin{aligned} R_{v-P(v)}(P(v)) &= \frac{1}{2}R_{v-P(v)}(P(v) - v + P(v) + v) \\ &= \frac{1}{2}(v - P(v)) + \frac{1}{2}(v + P(v)) \\ &= v \end{aligned}$$

This means $R_{v-P(v)} \circ P|_{\text{span}\{v\}} = \mathbb{1}$. So $P = \mathbb{1}_{1 \times 1} \oplus P|_{\text{span}\{v\}^\perp}$. By induction, we see if we can find v satisfying the original condition but on $\text{span}\{v\}^\perp$. In general the problem reduces to $\text{span}\{v\}^\perp$ which is $n - 1$ dimensional, so we try to show that $P|_{\text{span}\{v\}^\perp}$ and hence P is the product of at most $n - 1$ reflections

3. If V is anisotropic then suppose there exists a vector $v \in V_1$ as above. Then case 1 holds. Suppose V is anisotropic and V_1 is empty, then for any vector $v \neq 0$ we have $v - P(v) \neq 0$, so $Q(v) \neq 0$ and $Q(v - P(v)) \neq 0$. Then case 2 holds. Finally, we come to the last case. Suppose V is not anisotropic, and there exists $u \neq 0$ so that $Q(u) = 0$. Then we have two sub-cases, $n = 2$ and $n > 2$.
 - i) $n = 2$. If $n = 2$ and V has an isotropic vector, then V must be a hyperbolic plane. Let $\{u, v\}$ be a hyperbolic basis. Then P is an isometry, we have $P : u \mapsto au$ and $P : v \mapsto bv$ or $P : u \mapsto av$ and $P : v \mapsto bu$. Consider the first one, where P is diagonal. If $a = 1$ then P is just the identity. Otherwise, assume $a \neq 1$ and let $w = u + v$. Then $w - P(w) = u + v - au - a^{-1}v \neq 0$, so $Q(w) = 2$ and $Q(w - P(w)) = 2(1 - a)(1 - a^{-1})$. So case 2 holds. Now suppose P swaps u and v . Then $P = R_{u-av}$. Therefore we are done.
 - ii) Assume $n \geq 3$. Assume $V_1 = \ker(\mathbb{1} - P)$ is not anisotropic, and if $Q(v) \neq 0$ then $Q(v - P(v)) = 0$. The claim is that $Q(u - P(u)) = 0$ for all $u \in V$. Since we have assumed this is true for all anisotropic vectors, it suffices to prove it for all isotropic vectors. First of all, we have $Q(w - P(w)) = Q(P(w)) - 2G(w, P(w))$.

Note that $\text{span}\{w\}^\perp$ is $n - 1$ dimensional, and since $n \geq 3$ we have $n - 1 \geq \lfloor n/2 \rfloor$. So $\text{span}\{w\}^\perp$ is not isotropic because if it were, it would violate Lemma 2.14. So there must be some $u \in \text{span}\{w\}^\perp$ with $Q(u) \neq 0$ and $u \neq 0$. Then $G(w \pm u, w) = \pm G(u, w) = 0$. So $w \pm u \in \text{span}\{w\}^\perp$. So $Q(w \pm u) = Q(u) \neq 0$. Let $L = \mathbb{1} - P$. Then $Q(L(u)) = 0$ for all $u \in V$ as we stated at the beginning, so $Q(L(w \pm u)) = 0$. But then $Q(L(w) + L(u)) = Q(L(w)) + 2G(L(w), L(u))$ and $Q(L(w) - L(u)) = Q(L(w)) - 2G(L(w), L(u))$. So $Q(L(w) + L(u)) + Q(L(w) - L(u)) = 2Q(L(w)) = 0$. So $Q(w - P(w)) = 0$ as required.

Since $Q(u - P(u)) = 0$ for all u , $\text{image}(L)$ is isotropic. From assignment 2 we also have that $\ker(\mathbb{1} - P) = \text{im}(\mathbb{1} - P)^\perp$, so $\text{image}(\mathbb{1} - P)^\perp$ and $\text{image}(\mathbb{1} - P)$ are isotropic. So $V_1 \subseteq V_1^\perp$. Since V_1^\perp is isotropic, $V_1^\perp \subseteq V_1$. So $V_1 = V_1^\perp$. Hence $\text{image}(\mathbb{1} - P) = \ker(\mathbb{1} - P)$. So $(\mathbb{1} - P)^2 = 0$. So $\mathbb{1} - P$ is nilpotent. From the Jordan canonical form of $\mathbb{1} - P$ we see that the matrix of $\mathbb{1} - P = L$ in the Jordan basis J will look like all zeros, except for the first upper diagonal line which will be all ones. It follows that $P = \mathbb{1} - L$ will have all 1's on the diagonal.

$$[P]_J = \begin{bmatrix} 1 & -1 & & & & \\ & 1 & -1 & & & \\ & & 1 & -1 & & \\ & & & \ddots & \ddots & \\ & & & & 1 & -1 \\ & & & & & 1 \end{bmatrix} \quad (2.9)$$

Where there are zeros everywhere except shown. So $\det P = 1$, which means P is a proper isometry. In this case we also have $\dim V_1 = \dim V_1^\perp$, so $\dim V = 2 \dim V_1$, which means $\dim V$ is even.

This means that any improper isometry meets the criteria of one of the previous cases! So any improper isometry is the product of at most $n - 1$ reflections.

Now let R_w be a reflection and let $\tilde{P} = R_w P$. Since P is proper, \tilde{P} is improper. So \tilde{P} is the product of at most $n - 1$ reflections. Since $R_w^{-1} = R_w$, we see that $P = R_w \tilde{P}$, so P is the product of at most n reflections as required.

□

Theorem 2.13 (Classification of Symmetric Bilinear Forms). The classification of symmetric bilinear forms refers to the following theorems

- Sylvester's Law of Inertia
- Witt Extension Theorem
- Cartan-Dieudonné Theorem

Together, these three theorems allow one to determine a massive amount of information about a given metric.

2.5 Lecture 9: Skew-Symmetric Bilinear Forms and the Pfaffian

2.5.1 Classification of Bilinear Forms: Skew-Symmetric Bilinear Forms

The skew symmetric case is actually a lot easier than the symmetric case.

Theorem 2.14 (Classification of Skew-Symmetric Bilinear Forms). Let $\text{char } \mathbb{F} \neq 2$ and let V be a vector space over \mathbb{F} . Let G be a skew-symmetric bilinear form over V . Then there exists a basis $B = \{e_1, f_1, \dots, e_r, f_r, g_1, \dots, g_{n-2r}\}$ so that

1. $G(e_i, e_j) = 0$ for all i, j , $G(f_i, f_j) = 0$ for all i, j , $G(g_i, g_j) = 0$ for all i, j .
2. $G(e_i, f_j) = -G(f_j, e_i) = \delta_{ij}$.
3. $G(v, g_\ell) = 0$ for all $\ell = 1, \dots, n - 2r$.

In this basis we have,

$$[G]_B = \begin{bmatrix} 0 & 1 & & & & \\ -1 & 0 & & & & \\ & & 0 & 1 & & \\ & & -1 & 0 & & \\ & & & & \ddots & \\ & & & & & 0 & 1 \\ & & & & & -1 & 0 \\ & & & & & & & 0 \end{bmatrix} \quad (2.10)$$

Where there are zeros everywhere nothing is shown. In the bottom right corner, note that there is only a zero if $\dim V$ is odd.

Remark. If $\dim V$ is odd, then $\det[G]_B = 0$. So $\det[G] = 0$ in every basis. So G is degenerate.

Corollary 2.9. If G is nondegenerate then $\dim V$ is even.

Proof. If $G = 0$ then $r = 0$ and we are done. So suppose $G \neq 0$. Then there exist some e_1, \tilde{f}_1 so that $G(e_1, \tilde{f}_1) = a \neq 0$. Set $f_1 = \tilde{f}_1/a$. So $G(f_1, e_1) = -1$ and, because G is skew symmetric, this means $B_1 = \{e_1, f_1\}$ is linearly independent. Let $V_1 = \text{span}\{e_1, f_1\}$. Then

$$[G|_{V_1}]_{B_1} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

and $|\det[G|_{V_1}]_{B_1}| = 1$, so V_1 is nondegenerate. Set $V = V_1 \oplus V_1^\perp$, then pick any $e_2, \tilde{f}_2 \in V_1^\perp$ satisfying $G(e_2, \tilde{f}_2) = a \in \mathbb{F}$. Since they are elements of V_1^\perp we also satisfy the other required identities. We keep doing this, and by induction we can find some number r of pairs $\{e_i, f_i\}$ which satisfy $G(e_i, f_i) \neq 0$ and the other identities. Finally, we observe that $V = V_1 \oplus \dots \oplus V_r \oplus V^\perp$. So we finally let $\{g_1, \dots, g_{n-2r}\}$ be a basis for V^\perp , and the theorem is complete. \square

Remark. Recall that a skew symmetric bilinear form on V is also an element of $\Lambda^2 V^*$. So we have just completely characterized all possible elements of $\Lambda^2 V^*$.

Definition 2.29. Let $\dim V = 2k$ be even. A nondegenerate skew-symmetric bilinear form on V is called a **symplectic form**. It is usually denoted ω .

Definition 2.30. Let V and W be even dimensional vector spaces, and let G and H be symplectic forms on V and W respectively. Then an isomorphism $\phi : V \rightarrow W$ with $\phi^*H = G$ is called a **symplectic isomorphism**, or symplectomorphism. It is the equivalent of an isometry for skew-symmetric bilinear forms.

2.5.2 The Pfaffian

We've just shown that given any skew-symmetric bilinear form on V there is a basis where either $\det[G]_B = 0$ or $\det[G]_B = 1$. Suppose we choose a different basis C . Then let P_{CB} be the change of basis map. We have

$$\det[G]_C = \det(P_{CB})^2 \det[G]_B = \begin{cases} 0 & n = 2k + 1 \\ 1 & n = 2k \end{cases}$$

So for any basis C , the value of $\det[G]_C$ is a perfect square in \mathbb{F} . We can use this to construct an invariant of G (that is, a scalar function of G which is invariant under isometries of G).

Corollary 2.10. Any two skew-symmetric matrices A and B are related by a change of basis transformation $P^T B P = A$ if and only if $\text{rank } A = \text{rank } B$.

Proof. Suppose $\text{rank } A = \text{rank } B$. Then the dimension of $V^{\perp A}$ is the same as the dimension of $V^{\perp B}$ by definition. Therefore both matrices can be put into the same canonical form, which by transitivity means they are related by a change of basis transformation.

The converse holds due to the fact that rank is already known to be an invariant of bilinear forms, so $\text{rank } A = \text{rank } P^T B P = \text{rank } B$. □

Example 2.3. Some examples.

i) Let $n = 3$. Then

$$A = \begin{bmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{bmatrix}, \quad \det A = 0$$

ii) Let $n = 2$. Then

$$A = \begin{bmatrix} 0 & a \\ -a & 0 \end{bmatrix}, \quad \det A = a^2$$

Let $\{x_{ij} : i, j = 1, \dots, n\}$ be some variables (it does not matter what set they take values in, we are just using them

symbolically).

Definition 2.31. Consider the set of all polynomials, whose variables are $\{x_{ij}\}$ and whose coefficients are in \mathbb{Z} . We will call this set $\mathbb{Z}[x_{ij}]$.

Remark. Some remarks are as follows.

1. The only elements $f \in \mathbb{Z}[x_{ij}]$ with $f^{-1} \in \mathbb{Z}[x_{ij}]$ are $f = \pm 1$.
2. The set $\mathbb{Z}[x_{ij}]$ forms a **ring** with respect to multiplication and addition.
3. If $f \neq 0, g \neq 0$ then $fg \neq 0$. So the ring $\mathbb{Z}[x_{ij}]$ is an **integral domain**.
4. The ring $\mathbb{Z}[x_{ij}]$ is a **unique factorization domain**. For any $f \in \mathbb{Z}[x_{ij}]$ there is a unique set of polynomials $g_1, \dots, g_r \in \mathbb{Z}[x_{ij}]$ so that $f = \pm g_1 \dots g_r$.

Definition 2.32. Let $\mathbb{Q}(x_{ij})$ denote the **field of fractions** of $\mathbb{Z}[x_{ij}]$. That is, all rational functions f/g where f and g are polynomials with integer coefficients.

Remark. We are about to use the fact that $x_{ij} \in \mathbb{Q}(x_{ij})$ and $-x_{ij} \in \mathbb{Q}(x_{ij})$.

Lemma 2.15. Let X be the matrix formed by setting $X_{ii} = 0$ for all $i = 1, \dots, n$, $X_{ij} = x_{ij}$ for all $i < j$, and $X_{ij} = -x_{ij}$ for all $i > j$. Then there exist integer polynomials f, g so that $\det X = f^2/g^2$.

Proof. Since X is a skew symmetric matrix and $\mathbb{F} = \mathbb{Q}(x_{ij})$ is a field, we know that $\det X$ is a perfect square in \mathbb{F} . □

Remark. Without loss of generality, f^2 and g^2 share no common factors in \mathbb{F} . But since \mathbb{F} is a field, $f^2 \in \mathbb{F}$ is automatically divisible by $g^2 \in \mathbb{F}$. Therefore it is sufficient to take $g = \pm 1$, so $g^2 = 1$. Thus we have $\det X = f^2$ for some $f \in \mathbb{Z}[x_{ij}]$. This means that f is determined up to a sign by the formula $\det X = f^2$. We can use this to define the Pfaffian for any commutative ring R with $1 + 1 \neq 0$.

Remark. Given a ring R , there is a unique ring homomorphism from $\mathbb{Z}[x_{ij}]$ to R , which takes each monomial x_{ij} to some element of R . That is, there exists $\phi : \mathbb{Z}[x_{ij}] \rightarrow R$ with $\phi(x_{ij}) = A_{ij}$ for each i, j and so that $\phi(h(x_{11}, \dots)) = h(\phi(x_{11}), \dots)$. This is similar to extending a definition by linearity, except we are extending it as a ring homomorphism instead.

Remark. Let A_0 denote the standard matrix,

$$A_0 = \begin{bmatrix} 0 & 1 & & & \\ -1 & 0 & & & \\ & & 0 & 1 & \\ & & -1 & 0 & \\ & & & & \ddots \\ & & & & & 0 & 1 \\ & & & & & -1 & 0 \end{bmatrix}$$

Then $f(A_0) = \pm 1$, since $\det A_0 = 1$.

Definition 2.33 (Pfaffian). Let R be a commutative ring with $1 + 1 \neq 0$ and let $\text{Skew}(M_{nn}(R))$. Then we define the Pfaffian to be the unique function $\text{Pf} : \text{Skew}(M_{nn}(R)) \rightarrow R$ with $\text{Pf}(A_0) = 1$ and $\text{Pf}(A)^2 = \det A$ for all skew symmetric A .

Example 2.4. Some examples.

1. When $n = 2$, we have

$$\text{Pf} \begin{bmatrix} 0 & a \\ -a & 0 \end{bmatrix} = a$$

2. When $n = 4$ we have,

$$\text{Pf} \begin{bmatrix} 0 & A_{12} & A_{13} & A_{14} \\ -A_{12} & 0 & A_{23} & A_{24} \\ -A_{13} & -A_{23} & 0 & A_{34} \\ -A_{14} & -A_{24} & -A_{34} & 0 \end{bmatrix} = A_{12}A_{34} - A_{13}A_{24} + A_{14}A_{23}$$

Theorem 2.15. Let ω be a skew-symmetric bilinear form. Then $\text{Pf}(\omega)$ is well defined, and is invariant with respect to isometries.

Proof. If $\dim V$ is odd then $\text{Pf}(\omega) = 0$ for all ω . So assume $\dim V = 2k$ is even.

Let $\omega \in \Lambda^2 V^*$ be a skew-symmetric bilinear form. Then $\omega^k = \omega \wedge \dots \wedge \omega$ (k times) is an element of $\Lambda^{2k} V^*$. Let $\{e_1, \dots, e_{2k}\}$ be a basis for V . Then set $\omega = \frac{1}{2} A_{ij} e^i \wedge e^j$. We have,

$$\frac{1}{k!} \omega^k = \frac{1}{2^{2k} k!} A_{i_1 j_1} \dots A_{i_k j_k} e^{i_1} \wedge e^{j_1} \wedge \dots \wedge e^{i_k} \wedge e^{j_k}$$

Exercise: check that this simplifies to,

$$\frac{1}{k!} \omega^k = \text{Pf}(A) e^1 \wedge \dots \wedge e^{2k}$$

Now let P be an isometry of ω . So $P^*\omega = \omega$. Therefore,

$$\begin{aligned} (\Lambda^{2k} P^*) \frac{1}{k!} \omega^k &= \frac{1}{k!} P^* \omega \wedge \dots \wedge P^* \omega \\ &= \det(P)^2 \frac{1}{k!} \omega^k \\ &= \frac{1}{k!} \omega^k \end{aligned}$$

But it is also known that,

$$\frac{1}{k!} (P^* \omega)^k = \text{Pf}([P]^T A [P]) e^1 \wedge \dots \wedge e^{2k}$$

Therefore, since these are equal, we have

$$\text{Pf}([P]^T A [P]) = \text{Pf}(A)$$

So $\text{Pf}(\omega) = \text{Pf}(A)$ is unique up to isometry. □

3 Digression: The Hodge Dual

3.1 Lecture 10: Poincaré Duality

3.1.1 Orientations

Definition 3.1 (Oriented Basis). Let V be a vector space. Then an **oriented basis** for V is a basis B , where **the order of the basis matters**.

Definition 3.2 (Orientation). Let V be a vector space of dimension n . Then an **orientation** on V is an element of $\Lambda^n V$.

Definition 3.3 (Equivalence of Orientations). Let μ_1 and μ_2 be orientations on V . Note that $\dim \Lambda^n V = 1$, so $\mu_1 = a\mu_2$ with $a \neq 0$.

We define an equivalence relation by the statement " μ_1 and μ_2 are equivalent if and only if $a > 0$ ".

The equivalence classes of orientations are written using the notation $[\mu]$.

Remark. An oriented basis B determines an orientation μ_B by the formula $\mu_B = e_1 \wedge \dots \wedge e_n$.

Definition 3.4 (Equivalence of Oriented Bases). Let $B = \{e_1, \dots, e_n\}$ and $C = \{f_1, \dots, f_n\}$ be oriented bases. Then $\mu_B = e_1 \wedge \dots \wedge e_n \neq 0$ and $\mu_C = f_1 \wedge \dots \wedge f_n \neq 0$ are orientations on V . We say B and C determine equivalent orientations if $[\mu_C] = [\mu_B]$.

Remark. There are exactly two unique orientations on any vector space up to equivalence. They are often called "right handed" and "left handed". Which one is right and which one is left is a matter of convention.

Example 3.1. Consider $V = \mathbb{R}^3$ and let $e_1 = \hat{x}, e_2 = \hat{y}, e_3 = \hat{z}$ be the standard basis vectors. Then $\{e_1, e_2, e_3\}$ is an oriented basis with orientation $e^1 \wedge e^2 \wedge e^3$. By convention we call this basis the standard right handed basis. Another oriented basis is $\{e_1, e_3, e_2\}$. But $e^1 \wedge e^3 \wedge e^2 = -e^1 \wedge e^2 \wedge e^3$, so this has the opposite orientation. This is an example of a left handed basis.

3.1.2 Volume Forms

Let V be a vector space over \mathbb{R} with dimension n and let G be a nondegenerate symmetric or skew-symmetric bilinear form (in the symmetric case, with signature (p, q)). If G is skew symmetric let $n = 2m$.

Definition 3.5 (Dual of Bilinear Form). Recall that given nondegenerate G we have an isomorphism $\sharp : V^* \rightarrow V, \flat = \sharp^{-1}$. We can define a nondegenerate bilinear form G^* by the formula,

$$G^*(\alpha, \beta) = G(\sharp\alpha, \sharp\beta) \quad (3.1)$$

That is, $G^* = \sharp^*G$. Also observe that $G = \flat^*G^*$.

Remark. By definition, \flat is an isometry between (V, G) and (V^*, G^*) if G is symmetric, and if G is skew-symmetric then \flat is a symplectic isomorphism.

Definition 3.6 (Induced Metric on Exterior Product). Let $v_1, \dots, v_k, w_1, \dots, w_k \in V$. Then we define a metric $\Lambda^k G : \Lambda^k V \times \Lambda^k V \rightarrow \mathbb{R}$ by the formula,

$$\Lambda^k G(v_1 \wedge \dots \wedge v_k, w_1 \wedge \dots \wedge w_k) = \det \begin{bmatrix} G(v_1, w_1) & G(v_1, w_2) & \dots & G(v_1, w_k) \\ G(v_2, w_1) & G(v_2, w_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ G(v_k, w_1) & \dots & \dots & G(v_k, w_k) \end{bmatrix} \quad (3.2)$$

Since this is linear in each v_i, w_j we then extend this definition by linearity to all $\omega \in \Lambda^k V$.

We often simply write $\Lambda^k G = G$ to make the notation a bit easier.

Remark. If $v_i = w_i, i = 1, \dots, k$ then the matrix on the right hand side of (3.2) is called the **Gram matrix** or the **overlap matrix**.

Remark. If G is symmetric then $\Lambda^k G$ is symmetric. If G is skew-symmetric then $\Lambda^k G(\omega_1, \omega_2) = (-1)^k \Lambda^k G(\omega_2, \omega_1)$.

Remark. If $k = 0$ then $\Lambda^k V = \mathbb{R}$, so we define $\Lambda^0 G(a, b) = ab$.

Definition 3.7 (Volume Form). Let G be a metric of signature (p, q) and let μ be an orientation on V . Then there exists an orthonormal basis $B = \{e_1, \dots, e_p, f_1, \dots, f_q\}$ so that $[G]_B = \text{diag}(1, \dots, 1, -1, \dots, -1)$. Let $\lambda = e_1 \wedge \dots \wedge e_p \wedge f_1 \wedge \dots \wedge f_q$.

Since this is an orthonormal basis, all the entries of the Gram matrix of λ will be 1, so $\Lambda^n G(\lambda, \lambda) = \pm 1$. We therefore define,

$$\mu_G = \pm \lambda \quad (3.3)$$

Where the sign is chosen so that $[\mu_G] = [\mu]$. The orientation μ_G is called the **volume form** of G with respect to μ .

Remark. μ_G is the unique multiple of μ so that $\Lambda^n G(\mu_G, \mu_G) = (-1)^q$.

Remark. It follows that $\Lambda^n G$ is positive definite if q is even and negative definite if q is odd.

Now we will cover the skew-symmetric case. In the skew symmetric case we have $\dim V = n = 2m$.

Lemma 3.1 (Assignment 3 Question 4). Let ω be a symplectic form on V . Then $\omega^m = \omega \wedge \dots \wedge \omega \neq 0$. In other words, ω^m is an orientation on V^* .

Proof. By the characterization of skew symmetric bilinear functionals we can choose a basis $\{e_1, f_1, \dots, e_m, f_m\}$ for V such that $\omega(e_i, f_j) = \delta_{ij}$ and $\omega(e_i, e_j) = 0 = \omega(f_i, f_j)$.

Notice that because $\dim \Lambda^n V = 1$, it follows that ω^k is some multiple of $e^1 \wedge f^1 \wedge \dots \wedge e^m \wedge f^m$. We must show that it is not zero. We can do this by induction.

First, we have $(e^1 \wedge f^1 + e^2 \wedge f^2)^2 = e^1 \wedge f^1 \wedge e^2 \wedge f^2 + e^2 \wedge f^2 \wedge e^1 \wedge f^1$, which is equal to $e^1 \wedge f^1 \wedge e^2 \wedge f^2 + e^1 \wedge e^2 \wedge f^2 \wedge f^1$, which becomes $2e^1 \wedge f^1 \wedge e^2 \wedge f^2$.

Our inductive hypothesis is that for all $j \geq 1$, if we let $\omega_0 = \sum_{i=1}^{m-j} e^i \wedge f^i$, then we have $\omega_0^{k-j} = (m-j)! e^1 \wedge f^1 \wedge \dots \wedge e^{m-j} \wedge f^{m-j}$.

Since this is a wedge product of $2m$ forms, as long as we collect only 2-forms when exchanging terms we can treat the wedge product as commutative, so we have $(\omega_0 + e^m \wedge f^m)^m = \sum_{i=1}^m \binom{m}{i} \omega_0^{m-i} \wedge (e^m \wedge f^m)^i$. Since $\omega_0^m = 0$ the only term here which survives is where $i = 1$. Then $(\omega_0 + e^m \wedge f^m)^m = \omega^m = m\omega_0^{m-1} \wedge e^m \wedge f^m = m(m-1)! e^1 \wedge f^1 \wedge \dots \wedge e^m \wedge f^m$ as desired.

So we can see that $\omega^k = m! e^1 \wedge f^1 \wedge \dots \wedge e^m \wedge f^m \neq 0$ which is all we needed to show. \square

Theorem 3.1. There exists a unique element μ_ω of $\Lambda^n V$ so that $\frac{1}{m!} \omega^m(\mu_\omega) = 1$.

Proof. Let $\mu \in \Lambda^n V$. Then $\mu = C e_1 \wedge f_1 \wedge \dots \wedge e_m \wedge f_m$. Then,

$$\begin{aligned} \frac{1}{m!} \omega^m(\mu) &= \frac{C}{m!} m! e^1 \wedge f^1 \wedge \dots \wedge e^m \wedge f^m (e_1 \wedge f_1 \wedge \dots \wedge e_m \wedge f_m) \\ &= \frac{C}{m!} m! \\ &= C \end{aligned}$$

We therefore set $C = 1$, which gives us the unique solution. \square

Definition 3.8 (Symplectic Volume Element). The n -form μ_ω defined in theorem 3.1 is called the **symplectic volume element** ω .

Remark. Notice that if G is a metric, we need an orientation μ in order to define μ_G . But if G is symplectic, then μ_G does not depend on any choice of orientation.

3.1.3 Poincaré Duality and the Hodge Star

Remark. Let G be a metric or symplectic form and let μ_G be the associated volume form. Then there exists an isomorphism $M : \mathbb{R} \rightarrow \Lambda^n V$ defined by $M(1) = \mu_G$.

Definition 3.9 (Poincaré Dual Isomorphism). Let M be the canonical isomorphism $M : \mathbb{R} \rightarrow \Lambda^n V$ with $M(1) = \mu_G$. Then define a map $D : \Lambda^k V \times \Lambda^{n-k} V \rightarrow \mathbb{R}$ by the formula,

$$D(a, b) = M^{-1}(a \wedge b)$$

Then using D we can create a map $D_0 : \Lambda^{n-k} V \rightarrow (\Lambda^k V)^*$ according to the formula

$$P_1(b)(a) = D(a, b) \tag{3.4}$$

That is, b gets mapped to the object $D(\cdot, b)$.

Theorem 3.2 (Poincaré Duality). The map P_1 defined in (3.4) is an isomorphism. We refer to the isomorphism between $\Lambda^{n-k} V$ and $(\Lambda^k V)^*$ as **Poincaré duality**.

Proof. First notice that $\dim(\Lambda^k V^*) = \binom{n}{k} = \binom{n}{n-k} = \dim(\Lambda^{n-k} V)$. Therefore we just need to show that P_1 is injective. First let e_1, \dots, e_n be a basis for V and recall multi-index notation, where we let $e_I = e_{i_1} \wedge \dots \wedge e_{i_k}$. Then $\{e_I\}$ is a basis for $\Lambda^k V$. We set $b = b^I e_I$. Let J be a complementary multi-index to I . That is, if I contains some k numbers between 1 and n then J contains the $n - k$ numbers which are not in I . In this case $e_J \wedge b = 0$ for all J if and only if $b = 0$. So, since M is an isomorphism, $P_1(b)(a) = D(a, b) = M^{-1}(a \wedge b) = 0$ for all a if and only if $b = 0$. Therefore P_1 is injective. Furthermore, by unravelling the definitions we have $a \wedge b = P_1(b)(a) \mu_G$. □

Lemma 3.2. There exists an isometry $P_2 : \Lambda^k V \rightarrow (\Lambda^k V)^*$ given by $P_2(b)(a) = G(a, b)$.

Proof. P_2 is basically just the exterior power of the flat map, $(\Lambda^k b)(b)(a) = \Lambda^k G(a, b)$ by definition. □

Definition 3.10 (Hodge Star Operator). The map $\star : \Lambda^k V \rightarrow \Lambda^{n-k} V$ given by $\star = P_1^{-1} \circ P_2$ is called the **Hodge star** operator.

Theorem 3.3. The Hodge star operator is a linear isomorphism. Furthermore, it is the unique map satisfying

$$a \wedge \star(b) = G(a, b) \mu_G \quad (3.5)$$

for all a, b .

Proof. Since P_1 and P_2 are both linear and bijective it is clearly an isomorphism. For the other part, we directly compute:

$$\begin{aligned} a \wedge \star(b) &= a \wedge (P_1^{-1} \circ P_2(b)) \\ &= (P_1(P_1^{-1}(P_2(b))))(a) \mu_G \\ &= P_2(b)(a) \mu_G \\ &= G(a, b) \mu_G \end{aligned}$$

As required. □

Remark. If we replace μ_G by $-\mu_G$, then the sign of \star also changes.

Example 3.2. Let $V = \mathbb{R}^3$ and let G be the standard inner product. Let $\mu = e_1 \wedge e_2 \wedge e_3$ and let $[G]_B = \mathbb{1}$ in this basis. Let $a, b, c \in \mathbb{R}$ so that $\star(e_1) = ae_1 \wedge e_2 + be_2 \wedge e_3 + ce_1 \wedge e_3$ (such a, b, c exist by definition!). Then,

- $e_1 \wedge \star(e_1) = be_1 \wedge e_2 \wedge e_3$, so $b = 1$.
- $e_2 \wedge \star(e_1) = -ce_1 \wedge e_2 \wedge e_3 = G(e_1, e_2)(e_1 \wedge e_2 \wedge e_3) = 0$ so $c = 0$
- $e_3 \wedge \star(e_1) = 0$ as well

Overall we see that $\star(e_1) = e_2 \wedge e_3$. Applying the same argument to the other basis vectors gives us the result

$$\star(e_1) = e_2 \wedge e_3 \quad (3.6)$$

$$\star(e_2) = e_3 \wedge e_1 \quad (3.7)$$

$$\star(e_3) = e_1 \wedge e_2 \quad (3.8)$$

Remark. If we remove the stars on the left side, this example bears striking resemblance to the cross product. We will explore this connection later.

3.1.4 Hodge Star on Orthogonal Direct Sums

Suppose $V = V_1 \oplus V_2$ and G is nondegenerate. We already showed that $G|_{V_1} = G_1$ and $G|_{V_2} = G_2$ are both nondegenerate.

Remark. If G is symplectic, then $\mu_G = \mu_{G_1} \wedge \mu_{G_2}$.

If G is symmetric, then we can always choose an orientation μ so that $\mu_G = \mu_{G_1} \wedge \mu_{G_2}$ as well.

Let \star_1 be the Hodge star of G_1 and similarly define \star_2 .

Remark. Note that $\Lambda^\bullet V = \Lambda^\bullet V_1 \otimes \Lambda^\bullet V_2$. So any element $\zeta \in \Lambda^k V$ can be written in the form,

$$\zeta = \sum_{j=1}^n a_j \otimes b_j$$

with $a_j \in \Lambda^j V_1$ and $b_j \in \Lambda^{n-j} V_2$ for each $j = 1, \dots, n$.

This means that we can determine \star on all of V if we just know $\star(a_1 \wedge a_2)$ with $a_1 \in \Lambda^\bullet V_1$ and $a_2 \in \Lambda^\bullet V_2$.

Theorem 3.4. For all $b_1 \in \Lambda^r V_1, b_2 \in \Lambda^s V_2$ we have

$$\star(b_1 \wedge b_2) = (-1)^{(\dim V_1 - r)s} (\star_1 b_1) \wedge (\star_2 b_2) \quad (3.9)$$

Proof. Note that since $V_1 \perp V_2$, we have $G(a_1 \wedge a_2, b_1 \wedge b_2) = G(a_1, b_1)G(a_2, b_2)$. Thus we have,

$$\begin{aligned} a_1 \wedge a_2 \wedge \star(b_1 \wedge b_2) &= G(a_1 \wedge a_2, b_1 \wedge b_2) \mu_G \\ &= G(a_1, b_1) G(a_2, b_2) \mu_{G_1} \wedge \mu_{G_2} \\ &= (G(a_1, b_1) \mu_{G_1}) \wedge (G(a_2, b_2) \mu_{G_2}) \\ &= (a_1 \wedge \star_1(b_1)) \wedge (a_2 \wedge \star_2(b_2)) \\ &= (-1)^{(\dim V_1 - 1)j} a_1 \wedge a_2 \wedge \star(b_1) \wedge \star(b_2) \end{aligned}$$

Since this holds for all a_1, a_2 we are done. □

3.1.5 Square of the Hodge Star

Remark. Let \star be the Hodge star operator. Then $\star^2 : \Lambda^k V \rightarrow \Lambda^k V$ is an isomorphism between $\Lambda^k V$ and itself and only depends on G ! It does not depend on any other choices.

Theorem 3.5. If G is a metric with signature (p, q) , and $\star : \Lambda^k V \rightarrow \Lambda^{n-k} V$ is the Hodge star, then $\star^2 = (-1)^{k(n-k)+q} \mathbb{1}$.

Proof. Let $\{e_1, \dots, e_p, f_1, \dots, f_q\}$ be an oriented orthonormal basis of V with $\mu_G = e_1 \wedge \dots \wedge e_p \wedge f_1 \wedge \dots \wedge f_q$. A basis for $\Lambda^k V$ is given by $\{e_I \wedge f_J\}$ where I and J are multi-indices, and $|I| + |J| = k$. Notice that $e_I \perp f_J$ if $I \neq J$. That is, whenever I', J' are multi-indices with $|I'| + |J'| = n - k$, we have

$$G(e_I \wedge f_J, e_{I'} \wedge f_{J'}) = \pm \delta_{II'} \delta_{JJ'} = \pm \delta_{i_1 i'_1} \dots \delta_{i_p i'_p} \delta_{j_1 j'_1} \dots \delta_{j_q j'_q}$$

So then

$$(e_{I'} \wedge f_{J'}) \wedge \star(e_I \wedge f_J) = G(e_{I'} \wedge f_{J'}, e_I \wedge f_J) \mu_G = \pm \delta_{II'} \delta_{JJ'} \mu_G$$

Let us express the components of $\star(e_I \wedge f_J)$ as $\star(e_I \wedge f_J) = C_{IJ}^{I'J'} e_{I'} \wedge f_{J'}$. Then from the above computation we see that $C_{IJ}^{I'J'} = 0$ unless I', J' are complementary multi-indices to I and J respectively. Therefore, only one of the coefficients $C_{IJ}^{I'J'}$ is nonzero, and so we have

$$\star(e_I \wedge f_J) = C_{IJ} e_{I'} \wedge f_{J'}$$

Where I' and J' are complementary to I and J , and $C_{IJ} = \pm 1$. So let I^c and J^c be the complementary indices to I and J . Then we write

$$\begin{aligned} C_{IJ} e_I \wedge f_J \wedge e_{I^c} \wedge f_{J^c} &= e_I \wedge f_J \wedge \star(e_I \wedge f_J) \\ &= G(e_I \wedge f_J, e_I \wedge f_J) \mu_G \\ &= (-1)^{|J|} \mu_G \end{aligned}$$

Next, observe that $(I^c)^c = I$ and similarly for J . Therefore, the same equation holds for the complements,

$$\begin{aligned} \star(e_{I^c} \wedge f_{J^c}) \wedge e_I \wedge f_J &= C_{I^c J^c} e_{I^c} \wedge f_{J^c} \wedge e_I \wedge f_J \\ &= G(e_{I^c} \wedge f_{J^c}, e_I \wedge f_J) \\ &= (-1)^{|J^c|} \mu_G \end{aligned}$$

We then rearrange this to get,

$$\begin{aligned} \mu_G &= (-1)^{|J^c|} C_{I^c J^c} e_{I^c} \wedge f_{J^c} \wedge e_I \wedge f_J \\ &= (-1)^{k(n-k)+|J^c|} C_{I^c J^c} e_I \wedge f_J \wedge e_{I^c} \wedge f_{J^c} \end{aligned}$$

So we have

$$C_{IJ} (-1)^{|J|} = (-1)^{k(n-k)+|J'|} C_{I^c J^c}$$

Which implies that,

$$\star^2(e_I \wedge f_J) = C_{IJ} C_{I^c J^c} e_I \wedge f_J = (-1)^{k(n-k)+|J|+|J^c|} e_I \wedge f_J$$

Since $|J| + |J^c| = q$ we get,

$$\star^2 = (-1)^{k(n-k)+q} \mathbb{1}$$

□

Theorem 3.6. Suppose G is symplectic and \star is the Hodge star, then $\star^2 = \mathbb{1}$.

Proof. Let $\{e_1, \dots, e_m, f_1, \dots, f_m\}$ be a symplectic basis of V . That is, $G(e_i, f_j) = \delta_{ij}$, $G(e_i, e_j) = 0$, $G(f_i, f_j) = 0$, and $G(f_i, e_j) = -\delta_{ij}$.

Note that $\mu_G = e_1 \wedge f_1 \wedge \dots \wedge e_m \wedge f_m$. We will use induction on m . First the base cases.

First if $m = 0$. Then we only have $\Lambda^0 V = \mathbb{R}$ so $\mu_G = 1$, and $\star^2 1 = \mu_G = 1$. By linearity $\star^2 = 1$.

Next, suppose $m = 1$. Then we have $\mu_G = e_1 \wedge f_1$. So $\star 1 = \mu_G$, and $\star \mu_G = s \in \mathbb{R}$. Notice that $\mu_G \wedge \star(\mu_G) = G(\mu_G, \mu_G) \mu_G = \mu_G$. So $s = 1$ which implies $\star \mu_G = 1$ and $\star^2|_{\Lambda^0 V} = 1$. Next, we show that $\star^2|_{\Lambda^1 V} = 1$. We have $\star e_1 = a e_1 + b f_1$. So $e_1 \wedge \star e_1 = G(e_1, e_1) \mu_G = 0$, so $b = c$. Similarly, $f_1 \wedge \star e_1 = G(f_1, e_1) \mu_G = -\mu_G$, and $f_1 \wedge \star e_1 = a f_1 \wedge e_1 = -a \mu_G$. So $a = 1$. Therefore $\star e_1 = e_1$ and $\star f_1 = f_1$. We also know $\star^2 \mu_G = \mu_G$ so $\star^2|_{\Lambda^2 V} = 0$. This completes the case where $m = 1$.

Now suppose $m \geq 2$. The inductive hypothesis is that for all symplectic vector spaces of dimension $2(m-1)$, $\star^2 = 1$.

First, we write $V = V_1 \oplus V_2$ where $V_1 = \text{span}\{e_1, f_1\}$ and $V_2 = \text{span}\{e_2, f_2, \dots, e_m, f_m\}$. We have $\star_1^2 = 1$ and $\star_2^2 = 1$ by the inductive hypothesis. Additionally,

$$\begin{aligned} \star(b_1 \wedge b_2) &= (-1)^{(n_1-r)s} \star_1 b_1 \wedge \star_2 b_2 \\ &= (-1)^{rs} \star_1 b_1 \wedge \star_2 b_2 \\ \star^2(b_1 \wedge b_2) &= (-1)^{rs} (-1)^{(2-r)(2(m-1)-s)} b_1 \wedge b_2 \\ &= (-1)^{2rs} b_1 \wedge b_2 \\ &= b_1 \wedge b_2 \end{aligned}$$

So $\star^2 = 1$ on $\Lambda^k V$ for all k . □

Definition 3.11 (Anti-Isometry). Let G be a metric. An **anti-isometry** is a map $P : V \rightarrow V$ so that $P^* G = -G$.

Corollary 3.1. Let G be a metric on a vector space V . Let (p, q) be the signature of G . If G is symmetric then \star is an isometry if q is even and an anti-isometry if q is odd.

Proof. Let $a, b \in \Lambda^k V$, $\star a, \star b \in \Lambda^{n-k} V$. Then,

$$\begin{aligned} G(\star a, \star b) \mu_G &= \star a \wedge \star^2 b \\ &= t_k \star a \wedge b \\ &= (-1)^{k(n-k)} t_k G(b, a) \mu_G \end{aligned}$$

In the symmetric case, $G(b, a) = G(a, b)$ so $t_k = (-1)^{k(n-k)+q}$. So $G(\star a, \star b) = (-1)^q G(a, b)$. Therefore G is either an anti-isometry or an isometry. □

Corollary 3.2. Let V be a symplectic vector space and let G be the symplectic form. Then $\Lambda^k \star : \Lambda^k V \rightarrow \Lambda^{2m-k} V$ is a symplectomorphism when k is odd and an isometry when k is even.

Proof. Let $a, b \in \Lambda^k V$, $\star a, \star b \in \Lambda^{n-k} V$. Then,

$$\begin{aligned} G(\star a, \star b) \mu_G &= \star a \wedge \star^2 b \\ &= t_k \star a \wedge b \\ &= (-1)^{k(n-k)} t_k G(b, a) \mu_G \end{aligned}$$

We have $G(\star a, \star b) = (-1)^{k^2+k} G(a, b) = G(a, b)$. This completes the proof. \square

4 Relations Between Real and Complex Vector Spaces

4.1 Lecture 11: Relations between Complex and Real Vector Spaces. Part I

4.1.1 Antilinear Maps and Conjugation

Definition 4.1 (Antilinear). Let V and W be vector spaces over \mathbb{C} . A map $T : V \rightarrow W$ is called **antilinear** if $T(v + w) = T(v) + T(w)$ and $T(\lambda v) = \lambda^* v$ for all $v, w \in V, \lambda \in \mathbb{C}$.

Remark. In words, anti-linear map conjugates scalars.

Remark. Observe that for any $\lambda \in \mathbb{R}$, we have $\lambda^* = \lambda$, so $T(\lambda v) = \lambda T(v)$. Therefore we say that anti-linear maps are \mathbb{R} -linear but not \mathbb{C} -linear.

Example 4.1. Let $V = W = \mathbb{C}^m$. Define $T : \mathbb{C}^m \rightarrow \mathbb{C}^m$, where $T(v) = \bar{v}$ (which means we just conjugate each element of the tuple $v = (v_1, \dots, v_m)$). Then T is antilinear and a bijection.

Lemma 4.1. Let V_1, V_2, V_3 be complex vector spaces. Let $T_1 : V_1 \rightarrow V_2, T_2 : V_2 \rightarrow V_3$, and note that we have $T_2 \circ T_1 : V_1 \rightarrow V_3$. If T_1 and T_2 are both antilinear then $T_2 \circ T_1$ is \mathbb{C} -linear. If T_1 or T_2 is \mathbb{C} -linear but the other is antilinear, then $T_2 \circ T_1$ is antilinear.

Proof. Remember that $(z^*)^* = z$, $(\lambda \mu)^* = \lambda^* \mu^*$, and $(\lambda + \mu)^* = \lambda^* + \mu^*$. The proof follows. \square

Definition 4.2. Let V be a complex vector space. Then we define the **complex conjugate vector space** \bar{V} to be the vector space constructed over the same underlying set V except where the scalar multiplication operation $f : \mathbb{C} \times V \rightarrow V$ is defined by $f(\lambda, v) = \lambda^* v$ instead of by $f(\lambda, v) = \lambda v$.

Lemma 4.2. Let V be a complex vector space. The map $C_V : V \rightarrow \bar{V}$ given by $C_V(u) = \bar{u}$ is antilinear and a bijection. Note that this implies that $\dim \bar{V} = \dim V$.

Proof. $C_V(\lambda u) = \overline{\lambda u} = \lambda^* \overline{u}$. □

Remark. $\overline{\overline{V}} = V$ so $C_{\overline{V}} = C_V^{-1}$. That is, $C_V \circ C_{\overline{V}} = C_{\overline{V}} \circ C_V = \mathbb{1}$.

Remark. This also means \dim and \overline{V} are isomorphic (but not canonically, since C_V is not linear, so it is technically not an isomorphism).

Lemma 4.3. Let $P : V \rightarrow W$ be a linear map between complex vector spaces. Then P is antilinear if and only if $P \circ C_V$ is linear, if and only if $C_W \circ P$ is linear.

Proof. C_V is antilinear. □

Corollary 4.1. Let $P : V \rightarrow W$. Then P is an antilinear bijection if and only if $P \circ C_V$ is an isomorphism, if and only if $C_W \circ P$ is an isomorphism.

So any antilinear bijection between V and W is an isomorphism $V \cong \overline{W}$ and $W \cong \overline{V}$.

Corollary 4.2. Let $T : V \rightarrow W$ be complex linear. Then $\overline{T} = C_W \circ T \circ C_V$ is also complex linear. We have $\overline{T}(\overline{v}) = \overline{T(v)}$.

4.1.2 Constructing Real Vector Spaces from Complex Vector Spaces

Let V be a complex n -dimensional vector space.

Definition 4.3. Let $\{v_1, \dots, v_k\}$ be a set of k linearly independent vectors. Then the **span over \mathbb{C}** is defined by $\text{span}_{\mathbb{C}}\{v_1, \dots, v_n\} = \{\sum a^i v_i : a^i \in \mathbb{C}\}$. We also define the **span over \mathbb{R}** by $\text{span}_{\mathbb{R}}\{v_1, \dots, v_n\} = \{\sum a^i v_i : a^i \in \mathbb{R}\}$.

Definition 4.4. Suppose we take V as a set, and we define the vector space $V_{\mathbb{R}} = \text{span}_{\mathbb{R}} V$ to be the real vector space where scalar multiplication $\lambda v, v \in V$ is only defined for real numbers $\lambda \in \mathbb{R}$.

The vector space $V_{\mathbb{R}}$ is called the **underlying real vector space** of V .

Example 4.2. $\mathbb{C}_{\mathbb{R}}^n = \mathbb{R}^{2n}$.

To see this, let e_1, \dots, e_n be a basis for \mathbb{C}^n . Then as elements in $V_{\mathbb{R}}$ we can only get $\text{span}_{\mathbb{R}}\{e_1, \dots, e_n\} = \{\sum a^i e_i : a^i \in \mathbb{R}\}$, but we can't possibly get something like $ie_4 \in \text{span}_{\mathbb{R}}\{e_1, \dots, e_n\}$. That is to say, the vectors iv and v are linearly dependent in V , but they are linearly independent in $V_{\mathbb{R}}$. Therefore a basis for $\mathbb{C}_{\mathbb{R}}^n$ is actually $\{e_1, ie_1, \dots, e_n, ie_n\}$. So then $\text{span}_{\mathbb{R}}\{e_1, ie_1, \dots, e_n, ie_n\} = \mathbb{C}^n$.

Remark. Keep in mind the statement from the above example! The vectors iv and v are linearly dependent in V , but they are linearly independent in $V_{\mathbb{R}}$.

Remark. Suppose V is complex. Then $V_{\mathbb{R}} = \overline{V}_{\mathbb{R}}$ since $V = \overline{V}$ as sets and since scalar multiplication by real numbers satisfies $\lambda^* = \lambda$ for $\lambda \in \mathbb{R}$.

Definition 4.5. Let V be a vector space over \mathbb{C} . Then let B be a complex basis for V , so that $\text{span}_{\mathbb{C}} B = V$, then the complex dimension of V is $\dim_{\mathbb{C}} V = |B|$. Let \tilde{B} be a real basis for V , so that $\text{span}_{\mathbb{R}} \tilde{B} = V_{\mathbb{R}} = V$ as sets. Then the real dimension of V is $\dim_{\mathbb{R}} V = \dim V_{\mathbb{R}}$.

Lemma 4.4. $\dim V_{\mathbb{R}} = 2 \dim_{\mathbb{C}} V$.

Proof. Let $\{e_1, \dots, e_n\}$ be a basis for V . Then $\{e_1, ie_1, \dots, e_n, ie_n\}$ is a basis for $V_{\mathbb{R}}$. □

Question: Given a real vector space U , when can we find a complex vector space V so that $V_{\mathbb{R}} = U$? First of all, we would need $\dim_{\mathbb{R}} U = 2m$. Now, to characterize such a construction let us consider the map $J : U \oplus iU \rightarrow U \oplus iU$ defined by $J(v) = \sqrt{-1}v$. Since U is a real vector space, iU is linearly independent from U . We also have that J is linear over \mathbb{R} and that $J^2 = -\mathbb{1}$. This leads to the following definition.

Definition 4.6. Let U be an n -dimensional vector space over \mathbb{R} . A **complex structure** on U is an \mathbb{R} -linear map $J : U \rightarrow U$ such that $J^2 = -\mathbb{1}$.

If $U = V_{\mathbb{R}}$, then J definitely exists!

Remark. Suppose V_1, J_1 and V_2, J_2 are real vector spaces with complex structures. Then an \mathbb{R} -linear map $T : V_1 \rightarrow V_2$ is also \mathbb{C} -linear iff $T \circ J_1 = J_2 \circ T$.

Theorem 4.1. Any $2m$ dimensional real vector space then U admits a complex structure, and furthermore there exists a complex vector space V so that $V_{\mathbb{R}} = U$.

Additionally, any real vector space admitting a complex structure must be even dimensional, and furthermore there exists a complex vector space V so that $V_{\mathbb{R}} = U$.

Proof. Suppose U is a real vector space. Suppose there exists a complex structure J on U . Let $e_1 \neq 0$ be any element of V . Then since $J^2 = -\mathbb{1}$, we have $a_1 e_1 + a_2 J(e_1) = a_1 J(e_1) - a_2 e_1$. Therefore $a_1 e_1 + a_2 J(e_1) = 0$ if and only if $a_1 + a_2 = a_1 - a_2 = 0$, if and only if $a_1 = a_2 = 0$. So e_1 and $J(e_1)$ are linearly independent. Furthermore, this means that if $\{e_1, \dots, e_n\}$ are linearly independent then $\{e_1, J(e_1), \dots, e_n, J(e_n)\}$ is. Therefore $\dim U = 2m$. That is, the existence of a complex structure implies that the vector space is even dimensional.

Conversely, suppose $\dim U = 2m$. Let $\{e_1, \dots, e_m, f_1, \dots, f_m\}$ be a basis for U . Define $J : U \rightarrow U$ by $J(e_i) = f_i$ and $J(f_i) = -e_i$ for all $i = 1, \dots, m$. Then $J^2 = -\mathbb{1}$. So U admits a complex structure.

Now, define a complex vector space V so that as sets $V = U$, but so that the scalar multiplication is defined by $f(\lambda, v) = \text{Re}(\lambda)v + \text{Im}(\lambda)J(v)$ for all $v \in U$. Then V is a complex vector space of dimension m as required. □

4.2 Lecture 12: Relations between Complex and Real Vector Spaces. Part II

4.2.1 Constructing Complex Vector Spaces from Real Vector Spaces

Definition 4.7. Let U be an n dimensional real vector space. Define the map $J : U \oplus U \rightarrow U \oplus U$ by $J(u, 0) = (0, u)$ for all $u \in U$, and $J(0, u) = (-u, 0)$, and then linearly extending. The **complexification** of U is defined by making $U_{\mathbb{C}} = U \oplus U$ as sets and defining scalar multiplication by $f(i, v) = J(v)$.

Remark. We write $iU = \{0\} \oplus U$ and use the notation $iv = J(v)$ for all $v \in U_{\mathbb{C}}$.

Lemma 4.5. $\dim_{\mathbb{R}} U = \dim_{\mathbb{C}} U_{\mathbb{C}}$

Proof. Let $\{e_1, \dots, e_n\}$ be a basis for U . That is, $\text{span}_{\mathbb{R}}\{e_1, \dots, e_n\}$. Let $\tilde{e}_k = (e_k, 0)$. Then $U_{\mathbb{C}} = \text{span}_{\mathbb{C}}\{\tilde{e}_1, \dots, \tilde{e}_k\}$. \square

Remark. Observe that $U_{\mathbb{C}}$ is canonically isomorphic to $U \otimes \mathbb{C}$. The complex scalar multiplication is given by $\lambda(u \otimes z) = u \otimes (\lambda z)$. That is, we just associate $U \otimes \text{span}_{\mathbb{R}}\{i\} \cong \{0\} \oplus U$ and $U \otimes \text{span}_{\mathbb{R}}\{1\} \cong U \oplus \{0\}$.

Remark. This gives us an alternative proof, using the fact that $\dim_{\mathbb{R}} V \otimes W = (\dim_{\mathbb{R}} V)(\dim_{\mathbb{R}} W)$, and $\dim_{\mathbb{R}} \mathbb{C} = 2$. We have $\dim_{\mathbb{C}}(U \otimes \mathbb{C}) = \frac{1}{2} \dim_{\mathbb{R}}(U \otimes \mathbb{C}) = \frac{2}{2} \dim_{\mathbb{R}}(U) = \dim_{\mathbb{R}}(U)$.

Definition 4.8. Suppose $V = U_{\mathbb{C}}$ for some real vector space U . We can define an \mathbb{R} -linear map $K : U_{\mathbb{C}} \rightarrow U_{\mathbb{C}}$ by $K(u_1, u_2) = (u_1, -u_2)$. Then $K = \mathbb{1}|_U \oplus (-\mathbb{1})|_{iU}$. So $K^2 = \mathbb{1}$. The map K is called a **real structure** for V .

Definition 4.9. Let $V = U_{\mathbb{C}}$ and let K be a real structure for V . We then define,

$$V^+ = \{v \in V : K(v) = v\}, \quad V^- = \{v \in V : K(v) = -v\} \quad (4.1)$$

Remark. Observe that $V = V^+ \oplus V^-$ and $V = V^+ \cap V^- = \{0\}$.

Remark. We have $v = \frac{1}{2}(K(v) + v) + \frac{1}{2}(K(v) - v)$.

Remark. If $v = (u_1, u_2) = u_1 + iu_2$ then $K(iv) = (-u_2, -u_1)$, but $iK(v) = (u_2, u_1)$. So K is antilinear.

Definition 4.10. Let J be the complex structure for $U_{\mathbb{C}}$. That is, the map $J : U_{\mathbb{C}} \rightarrow U_{\mathbb{C}}$ so that $J(u_1, u_2) = (u_2, -u_1)$ for all $u \in U$. Define

$$U^{(1,0)} = \{v \in U_{\mathbb{C}} : J(v) = iv\}, \quad U^{(0,1)} = \{v \in U_{\mathbb{C}} : J(v) = -iv\} \quad (4.2)$$

Lemma 4.6. There exists a complex linear isomorphism $\overline{U^{(1,0)}} \cong U^{(0,1)}$.

Proof. Let $K : U_{\mathbb{C}} \rightarrow U_{\mathbb{C}}$ be the canonical real structure defined by $K(v_1 + iv_2) = v_1 - iv_2$. Then K maps $U^{(1,0)}$ to $U^{(0,1)}$. We then set $P = K \circ C_{U_{\mathbb{C}}}$. So P is a linear isomorphism. \square

Corollary 4.3. We have $\dim_{\mathbb{C}} U^{(1,0)} = \dim_{\mathbb{C}} U^{(0,1)} = \frac{1}{2} \dim_{\mathbb{C}} U$.

Theorem 4.2. $U_{\mathbb{C}} = U^{(1,0)} \oplus U^{(0,1)}$.

Proof. Let $v \in U_{\mathbb{C}}$. Then $v = (v - iJ(v))/2 + (v + iJ(v))/2$. We can see that $U^{(1,0)} \cap U^{(0,1)} = \{0\}$ since $iv \neq -iv$ for all $v \neq 0$. We then define $T : U \rightarrow U^{(1,0)}$ by $T(v) = v - iJ(v)$. Then T is \mathbb{R} -linear and injective, and $\dim U = \dim U^{(1,0)}$, so it is an isomorphism. Finally, we have

$$T(J(v)) = J(v) - iJ^2(v) = J(v) - iv = i(v - iJ(v))$$

So $T \circ J = iT$, which means it is \mathbb{C} -linear. This completes the proof. \square

4.2.2 Non-existence of Field Structure on \mathbb{R}^3

In this section we will change topics a little bit. Up until now we have covered the theory of bilinear forms, and now complex vector spaces. The purpose of this was to build up our tool-set for studying the geometry of finite vector spaces, so that we can construct various useful algebras.

A key question is whether or not we can endow \mathbb{R}^3 with a vector multiplication which is invertible. That is, can we turn \mathbb{R}^3 into a field? As we will see in this subsection, the answer is no.

Definition 4.11 (Division Algebra). Let A be a vector algebra over \mathbb{F} . Then we say A is a division algebra if there exists an element e so that for all $0 \neq x \in A$ there exists $x^{-1} \in A$ with $xx^{-1} = x^{-1}x = e$. The element e is called the identity.

Remark. Let A be a division algebra. Then if A is associative and commutative, it is also a field.

Any field is a division algebra over itself.

Example 4.3 (Real Numbers). Consider the field \mathbb{R} . We know this can be interpreted as a one-dimensional vector space. It can also be interpreted as a one-dimensional associative algebra. If we have two real numbers a and b , then we can also interpret them as vectors, and so their product ab is basically an algebra structure. In fact, we know that since they form a field, they are a division algebra.

We should also remark that $|ab| = |a||b|$ for any $a, b \in \mathbb{R}$.

Definition 4.12 (Norm). Let V be a vector space over \mathbb{R} . Then a **norm** on V is a function $v \mapsto \|v\|$ which satisfies the three following properties,

$$\|v + w\| \leq \|v\| + \|w\| \quad (4.3)$$

$$\|v\| \geq 0 \quad (4.4)$$

$$\|av\| = |a| \|v\| \quad (4.5)$$

$$\|v\| = 0 \implies v = 0 \quad (4.6)$$

Where $v, w \in V, a \in \mathbb{R}$.

Definition 4.13. Let A be a real vector algebra. Then we say the algebra structure is compatible with a norm $\|v\|$ if $\|vw\| = \|v\| \|w\|$ for all $v, w \in V$. We say that the norm is **multiplicative**, and we say that A is a **composition algebra**.

Remark. In the actual course, we called these algebras "normed algebras". However, on wikipedia one will find that a "normed algebra" means something else. Therefore we will use the other terminology so that things are easier to look up.

Example 4.4. The complex numbers \mathbb{C} form a two dimensional vector space over \mathbb{R} . In fact, they also form a two-dimensional real associative division algebra.

Furthermore, $|zw| = |z||w|$ for all $z, w \in \mathbb{C}$. So the complex numbers form a **compositional associative division algebra**.

Remark. In the weaker case, where $|ab| \leq C|a||b|$ for some C , we say A is a **normed algebra**. All composition algebras are normed.

Physics Remark. Normed algebras arise in quantum physics and computational PDEs quite often. Infinite dimensional normed algebras are often called Banach algebras.

Remark. Recall that a unital algebra is one with an identity, $1_V \in V$. That is, $1_V v = v 1_V = v$ for all $v \in V$. Any associative unital commutative division algebra is a field.

Now we ask a question. Is there a way to make \mathbb{R}^3 into a field by defining a suitable vector multiplication?

Theorem 4.3. The vector space \mathbb{R}^3 can not be made into a unital division algebra. So, as a corollary, it can not be made into a field.

Proof. This proof is by contradiction.

Suppose $V = \mathbb{R}^3$ was a field. Let $1_V \neq 0$ be the identity. Then let $x \in \mathbb{R}^3$ be any vector.

We know that the set $\{1_V, x, x^2, x^3\}$ is linearly dependent because $\dim V = 3$.

Therefore there exist constants c_0, c_1, c_2, c_3 so that

$$c_0 1_V + c_1 x + c_2 x^2 + c_3 x^3 = 0$$

Suppose that x is not a multiple of 1_V . Then c_2 and c_3 must not be both zero.

If $c_2 = c_3 = 0$ we would get a linear equation, which would contradict the fact that x is not a multiple of 1_V .

Suppose that $c_3 \neq 0$. Then we can divide the equation by c_3 to get

$$x^3 + b_2 x^2 + b_1 x + b_0 1_V = 0$$

Where we have defined $b_2 = c_2/c_3, b_1 = c_1/c_3, b_0 = c_0/c_3$.

Consider the analogous polynomial with real variables, rather than elements of \mathbb{R}^3 . Recall that real cubics always have at least one real root.

Therefore, we can definitely factor the above equation into

$$(x - r 1_V)(x^2 + \mu x + \nu 1_V) = 0$$

However, since x is not a multiple of 1_V we can't possibly have $x - r 1_V = 0$. Therefore, while $(x - r 1_V)$ could be a factor of the polynomial, it must not be zero. This is the difference between the case for real polynomials and polynomials with elements in the algebra.

Therefore, for the equation to be satisfied we must instead have $x^2 + \mu x + \nu 1_V = 0$, for some μ, ν . We also remark that the same thing would hold if $c_3 = 0$ and $c_2 \neq 0$. This accounts for all cases.

Now, notice that we can't factor the quadratic into linear terms, since otherwise we would again have x be a multiple of 1_V . In order to avoid this, we must have that $\mu^2 - 4\nu < 0$ so that there are no roots.

We then complete the square in the above equation, to get

$$(x + \mu 1_V/2)^2 + (\mu^2/4 - \nu) 1_V = 0$$

If we set

$$y = \frac{x + \mu 1_V/2}{\sqrt{\nu - \mu^2/4}}$$

then the above equation tells us that $y^2 = -1_V$. Additionally, y is not a multiple of 1_V .

Essentially, what we have just proved is that given any vector \tilde{v} which is not a multiple of 1_V , we can construct a vector $v = a\tilde{v} + b 1_V$ which satisfies $\tilde{v}^2 = -1_V$ and which is not a multiple of 1_V .

Now, let $\{e_0, \tilde{e}_1, \tilde{e}_2\}$ be a basis for \mathbb{R}^3 . Set $e_0 = 1_V$. Then we can construct e_1, e_2 , so that $\{e_0, e_1, e_2\}$ is a basis for \mathbb{R}^3 , and so that $e_1^2 = e_2^2 = -1_V$.

Furthermore, $e_1 e_2 = s_0 1_V + s_1 e_1 + s_2 e_2$. Multiplying on the right by e_2 gives $e_1 = s_0 e_2 + s_1 e_1 e_2 - s_2 e_0$, so $s_1 \neq 0$ or else e_1 is not linearly independent of e_0, e_2 .

Therefore, $e_1 e_2 = (s_2/s_1)e_0 - e_1/s_1 - (s_0/s_1)e_2$. Comparing this to the original equation, we have $s_1 = -1/s_1$, so $s_1^2 = -1$.

But $s_1 \in \mathbb{R}$, so this is not possible.

We have now run into a contradiction. Therefore there is no multiplication structure on \mathbb{R}^3 which makes it into a unital division algebra. □

5 Quaternions

5.1 Lecture 13: The Quaternions

5.1.1 The Quaternions

Definition 5.1 (Quaternions). Consider the vector space \mathbb{R}^4 . Recall the standard inner product, $\langle v, w \rangle = G(v, w) = \sum_{i=0}^3 v^i w^i$. The standard basis $\{e_0, e_1, e_2, e_3\}$ is orthonormal for G .

We set $\mathbb{R}^4 = \mathbb{R} \oplus \mathbb{R}^3$, with $\mathbb{R} = \text{span}\{e_0\}$ and $\mathbb{R}^3 = \text{span}\{e_1, e_2, e_3\}$.

We then define an associative algebra structure on \mathbb{R}^4 by first requiring that,

$$e_0 q = q e_0 \quad \forall q \in \mathbb{R}^4, \quad (5.1)$$

$$e_1^2 = e_2^2 = e_3^2 = -e_0, \quad (5.2)$$

$$e_1 e_2 = -e_2 e_1 = e_3, \quad (5.3)$$

$$e_2 e_3 = -e_3 e_2 = e_1, \quad (5.4)$$

$$e_3 e_1 = -e_1 e_3 = e_2 \quad (5.5)$$

and then extending by linearity. This algebra is not commutative, but it does have an identity, e_0 .

This vector algebra is called the **quaternion algebra**, and it is referred to as \mathbb{H} .

Remark. Since $\mathbb{H} = \mathbb{R}^4 = \mathbb{R} \oplus \mathbb{R}^3$, we often write $e_0 = 1, e_1 = i, e_2 = j, e_3 = k$. Then we have $ij = k, jk = i, ki = j$, and $i^2 = j^2 = k^2 = ijk = -1$. This is how we most often see the quaternions written.

In general, let $p = (p^0, p^1, p^2, p^3), q = (q^0, q^1, q^2, q^3) \in \mathbb{R}^4$. Then,

$$\begin{aligned} pq &= (p^0 q^0 - p^1 q^1 - p^2 q^2 - p^3 q^3) e_0 \\ &\quad + (p^0 q^1 + q^0 p^1 + p^2 q^3 - p^3 q^2) e_1 \\ &\quad + (p^0 q^2 + q^0 p^2 + p^3 q^1 - p^1 q^3) e_2 \\ &\quad + (p^0 q^3 + q^0 p^3 + p^1 q^2 - p^2 q^1) e_3 \end{aligned}$$

Suppose we write $p = p^0 + \mathbf{p}$, with p^0 representing $(p^0, 0, 0, 0) \in \mathbb{R}$ and $\mathbf{p} = (0, p^1, p^2, p^3) \in \mathbb{R}^3$. We refer to p^0 as the **real part** of p and \mathbf{p} as the **imaginary part**. That is, we are utilizing the isomorphism $\mathbb{R}^4 = \mathbb{R} \oplus \mathbb{R}^3$. If we do this, we will then observe that,

$$pq = (p^0 q^0 - \langle \mathbf{p}, \mathbf{q} \rangle) + (p^0 \mathbf{q} + q^0 \mathbf{p} + \mathbf{p} \times \mathbf{q}) \quad (5.6)$$

So the real part is $\text{Re}(pq) = p^0 q^0 - \langle \mathbf{p}, \mathbf{q} \rangle$, and the imaginary part is $\text{Im}(pq) = p^0 \mathbf{q} + q^0 \mathbf{p} + \mathbf{p} \times \mathbf{q}$

Physics Remark. Physicists will note that the real part is the Minkowski metric applied to p, q . That is, $\text{Re}(pq) = \eta(p, q)$.

Remark. If $\text{Im}(p) = \text{Im}(q) = 0$, then $pq = p^0 q^0$, so \mathbb{R} is a subalgebra, since the product of two real numbers is real.

Remark. Set $e_1 = i$. Then $i^2 = -e_0 = -1$. So we see that \mathbb{C} is also a subalgebra of \mathbb{H} .

Definition 5.2 (Center of an Algebra). Let A be an algebra. Then we define the center of A to be the set $\text{cent}(A) = \{a \in A : ab = ba \forall b \in A\}$. It is the set of all elements which commute with everything in A .

Lemma 5.1. $\text{cent}(\mathbb{H}) = \mathbb{R}$.

Proof. Write $p = p^0 + \mathbf{p}$, $q = q^0 + \mathbf{q}$. Then $pq - qp = \mathbf{p} \times \mathbf{q} - \mathbf{q} \times \mathbf{p} = 2\mathbf{p} \times \mathbf{q}$. And $\mathbf{p} \times \mathbf{q} = 0$ for all \mathbf{q} only if $\mathbf{p} = 0$. Therefore $\text{cent}(\mathbb{H}) = \mathbb{R}$. \square

Definition 5.3 (Conjugate of Quaternion). Let $p = p^0 + \mathbf{p}$. We define $\bar{p} = p^0 - \mathbf{p}$. This is the real linear map defined by the matrix $\text{diag}(1, -1, -1, -1) = [K]_S$ in the standard basis, with $K^2 = \mathbb{1}$. In particular, this is an isomorphism, and also an isometry with respect to the standard inner product on \mathbb{R}^4 . We can also write $K = -R_{e_0}$ - it is a reflection along e_0 .

Remark. $\langle \bar{p}, \bar{q} \rangle = \langle p, q \rangle$, so K is an improper isometry.

Remark. We have $\text{Re}(p) = p^0 = \frac{1}{2}(p + \bar{p})$ and $\text{Im}(p) = \mathbf{p} = \frac{1}{2}(p - \bar{p})$.

Remark. Quaternions with $\text{Re}(p) = 0$ are called imaginary quaternions. The set of imaginary quaternions is $\text{Im}(\mathbb{H}) = \mathbb{R}^3$.

Remark. Unit vectors in \mathbb{R}^4 are often called "unit quaternions" when interpreted as quaternions.

Remark. We have $p\bar{p} = \langle \mathbf{p}, \mathbf{p} \rangle + (p^0)^2 = |p|^2$.

Suppose $|p| = 1$ and $\text{Re}(p) = 0$, ie, p is an imaginary unit quaternion. Then $p^2 = -p\bar{p} = -1$.

The set of all imaginary unit quaternions is the 3 dimensional unit sphere, S^2 .

Since $p^2 = -1$ for any imaginary unit quaternion, any choice of imaginary unit quaternion provides a complex structure for \mathbb{R}^4 ! So the set of all complex structures at least contains a copy of the unit sphere.

Remark. The unit sphere in n dimensions forms a $n - 1$ -dimensional hypersurface. We therefore refer to the unit sphere in n dimensions as S^{n-1} . For example, the circle is S^1 and the usual unit sphere in \mathbb{R}^3 (as in the previous remark) is S^2 .

Remark. Let p, q be any quaternions. Recall that by equation (5.6) we can write,

$$p\bar{q} = (p^0 q^0 + \langle \mathbf{p}, \mathbf{q} \rangle) + \text{Im}(p\bar{q})$$

The first term (the real part) on the right hand side is equal to $\langle p, q \rangle$. In other words we have the identity

$$\langle p, q \rangle = \text{Im}(p\bar{q}) \tag{5.7}$$

Remark. Let v, w be imaginary quaternions. Then $-2\langle v, w \rangle = -2\operatorname{Re}(v\bar{w}) = 2\operatorname{Re}(vw) = vw + \overline{vw} = vw + wv$. Therefore,

$$\forall v, w \in \operatorname{Im}(\mathbb{H}), \quad vw + wv = -2\langle v, w \rangle \quad (5.8)$$

Physics Remark. Equation (5.8) is an anti-commutation relation as seen in quantum mechanics.

Remark. Let $p \in \mathbb{H}, p \neq 0$, from $p\bar{p} = |p|^2$ we see that $p^{-1} = \bar{p}/|p|^2$ is the inverse of p .

Definition 5.4 (Algebra Homomorphism). Let A, B be two vector algebras. Let $L : A \rightarrow B$ be a linear map and suppose $L(ab) = L(a)L(b)$ for all $a \in A, b \in B$ we say that L is an algebra homomorphism.

Remark. We always use the word "homomorphism" to mean a map which "preserves" some structure of the space we are interested in. In this case, it preserves the multiplicative structure. The same holds for group homomorphisms.

Definition 5.5 (Algebra Isomorphism). If L is a linear isomorphism and an algebra homomorphism, we say it is an algebra isomorphism.

Remark. If two algebras have the same vector space structure, but different multiplication structures, then they could be isomorphic as vector spaces but not as algebras.

Definition 5.6 (Automorphism). If L is an algebra isomorphism from A to itself, we say it is an automorphism. The set of automorphisms of A is denoted $\operatorname{Aut}(A)$.

Remark. The automorphisms of an algebra are like the "symmetries" of the algebra.

Definition 5.7 (Anti-Automorphism). If L is an antilinear map, a bijection, and an algebra homomorphism from A to itself, we say L is an anti-automorphism.

Remark. We view anti-automorphisms as another kind of symmetry.

Lemma 5.2. Conjugation is an anti-automorphism. In other words, $\overline{pq} = \bar{p}\bar{q}$.

Proof.

$$\begin{aligned} \overline{pq} &= (p^0 q^0 + \langle -\mathbf{p}, -\mathbf{q} \rangle) + (-p^0 \mathbf{q} - q^0 \mathbf{p} + \mathbf{p} \times \mathbf{q}) \\ &= (p^0 q^0 + \langle \mathbf{p}, \mathbf{q} \rangle) - (-p^0 \mathbf{q} - q^0 \mathbf{p} - \mathbf{q} \times \mathbf{p}) \\ &= \bar{p}\bar{q} \end{aligned}$$

□

Lemma 5.3. The quaternions form a composition algebra. In other words, $|pq| = |p||q|$.

Proof. Recall that $|p|^2 = p\bar{p}$. So,

$$\begin{aligned} |pq|^2 &= (pq)(\overline{pq}) \\ &= pq\bar{q}\bar{p} \\ &= p|q|^2\bar{p} \\ &= p\bar{p}|q|^2 \\ &= |p|^2|q|^2 \end{aligned}$$

Since $|p| \geq 0$, $|q| \geq 0$, we have $|pq| = |p||q|$ as required. □

Corollary 5.1. The quaternions form a compositional division algebra.

Definition 5.8 (Group of Unit Quaternions). Observe that the set $S^3 = \{p \in \mathbb{H} : |p| = 1\}$ forms a group with respect to quaternion multiplication. That is, if $p, q \in S^3$ then $p^{-1} \in S^3$ and $pq \in S^3$. We therefore refer to S^3 as the **group of unit quaternions**

Remark. The set $S^3 \subseteq \mathbb{H}$ forms a group. However, the set of imaginary unit quaternions $S^2 \subseteq \text{Im}(\mathbb{H})$ does **not** form a group.

Remark. If you have two groups G and H , we can form a new group out of $G \times H$ where the multiplication is just $(g_1, h_1)(g_2, h_2) = (g_1g_2, h_1h_2)$. This is called the direct product of the two groups.

Definition 5.9. Let $p, q \in S^3$ and let $h \in \mathbb{H}$. Define $K_{p,q} : \mathbb{H} \rightarrow \mathbb{H}$ by the formula,

$$K_{p,q}(h) = ph\bar{q}$$

Note that since $q \in S^3$, $\bar{q} = q^{-1}$. Also note that $K_{p,q}$ is linear with respect to h and invertible, so that $K_{p,q}^{-1} = K_{\bar{p},\bar{q}}$.

Lemma 5.4. The linear map $K_{p,q}$ is an isometry. In other words, $K_{p,q} \in O(4, \mathbb{R})$.

Proof. We have

$$\begin{aligned} |K_{p,q}(h)| &= |ph\bar{q}| \\ &= |p||h||\bar{q}| \\ &= |h| \end{aligned}$$

By the polarization identity we therefore have $\langle K_{p,q}a, K_{p,q}b \rangle = \langle a, b \rangle$ for all $a, b \in \mathbb{H}$. □

Lemma 5.5. Let $K : S^3 \times S^3 \rightarrow O(4, \mathbb{R})$ be the map defined by $K(p, q) = K_{p,q}$ for all $p, q \in S^3$. Then K is a group homomorphism.

Proof. We have,

$$\begin{aligned} K_{p_1p_2, q_1q_2}(x) &= p_1p_2x\overline{q_1q_2} \\ &= p_1(p_2x\bar{q}_2)\bar{q}_1 \\ &= K_{p_1, q_1}(K_{p_2, q_2}(x)) \end{aligned}$$

So $K(p_1p_2, q_1q_2) = K(p_1, q_1)K(p_2, q_2)$. Thus K is a homomorphism. □

Remark. K is **almost** an isomorphism but not quite. Let us figure out its' kernel.

Lemma 5.6. $\ker K = \{(1, 1), (-1, -1)\}$

Proof. Suppose $(p, q) \in \ker K$. Then $K(p, q) = \mathbb{1}$. Therefore $px\bar{q} = x$ for all $x \in \mathbb{H}$. If we substitute in $x = 1$ we find that $p\bar{q} = 1$. Therefore $p = q$ and furthermore we have $px\bar{p} = x$, so $p \in \text{cent}(\mathbb{H})$, which means $p \in \mathbb{R}$. Since $p\bar{p} = 1$ we therefore get $p = \pm 1$ and $q = p$. □

Remark. To interpret this, we pick an element $P \in \text{image}(K)$. Then there are two elements, (p, q) and $(-p, -q)$ which have the property that $K(p, q) = K(-p, -q) = P$. Is one of these a preferred choice? The answer will turn out to be no.

Remark. Later, we will prove that K is surjective.

5.1.2 Conjugation by Quaternions

In this section we will show how rotation matrices in \mathbb{R}^3 correspond to unit quaternions. For any rotation matrix, there are two unit quaternions which represent it.

Definition 5.10 (Diagonal Subgroup). Let G be a group, recall that $G \times G$ is also a group. The **diagonal subgroup** is the subgroup G_Δ of $G \times G$ defined by,

$$G_\Delta = \{(g, g) : g \in G\}.$$

Remark. The map $p \mapsto (p, p) \in G \times G$ is an isomorphism between G and G_Δ .

Remark. If S_Δ^3 is the diagonal subgroup of $S^3 \times S^3$, then the restriction $K|_{S_\Delta^3}$ is a group homomorphism. We denote $K_{p,p} = C_p$, so that $C_p(x) = px\bar{p}$. In other words, the map $p \mapsto C_p$ is a homomorphism from S^3 to $O(4, \mathbb{R})$

Remark. Observe that $C_p(1) = p\bar{p} = 1$. So C_p preserves \mathbb{R} . Since $K_{p,q}$ is an isometry, C_p is also an isometry, and so C_p preserves \mathbb{R}^3 . That is, C_p takes imaginary quaternions to imaginary quaternions.

Remark. Since $C_p|_{\mathbb{R}}$ is the identity, we see that $C_p = \mathbb{1} \oplus \tilde{C}_p$, where \tilde{C} is now a homomorphism from S^3 to $O(3, \mathbb{R})$. In other words, the outputs of C are just elements of $O(3, \mathbb{R})$ embedded in the larger group $O(4, \mathbb{R})$. So we usually write $C : S^3 \rightarrow O(3, \mathbb{R})$ rather than $C : S^3 \rightarrow O(4, \mathbb{R})$ since it is understood that acting it on the real part of \mathbb{H} has no effect.

Remark. Let $x \in \text{Im}(\mathbb{H})$. We have $x = -\bar{x}$. So $\overline{C_p x} = \overline{px\bar{p}} = p\bar{x}\bar{p} = -C_p x$.

Remark. Suppose $p \in \ker C$. Then $px\bar{p} = x$ for all x , so $px = xp$. Therefore $p \in \mathbb{R}$. We also have $|p| = 1$, so $p = \pm 1$. Therefore $\ker C = \{1, -1\}$. In other words, any $P \in \text{image}(C)$ is the image of exactly two points $p, -p \in S^3$.

Lemma 5.7. $SO(3, \mathbb{R}) \subseteq \text{image } C$.

Proof. Let $u \in S^2 = S^3 \cap \text{Im}(\mathbb{H})$. That is, $u^2 = -1$ and $\bar{u} = -u$.

If $x \in \text{Im}(\mathbb{H})$ then $2\langle u, x \rangle = -(ux + xu)$. So $ux = -2\langle u, x \rangle - xu$.

We also have $C_u(x) = ux\bar{u} = -uxu$. So,

$$\begin{aligned} C_u(x) &= -(-2\langle u, x \rangle - xu)u \\ &= 2\langle u, x \rangle u - xu^2 \\ &= -(x - 2\langle u, x \rangle u) \\ &= -R_u(x) \end{aligned}$$

Therefore, if u is an imaginary unit quaternion then C_u is a reflection.

Now, suppose $P \in SO(3, \mathbb{R})$. Recall that any proper isometry is the product of an even number of reflections, by Cartan-Dieudonné theorem. So there exist $u_1, \dots, u_{2m} \in S^2$ with,

$$P = (-R_{u_1}) \dots (R_{u_{2m}}) = C_{u_1} \dots C_{u_{2m}} = C_{u_1 \dots u_{2m}}$$

Since $u_i \in S^3$ for each i , we see that $u_1 \dots u_{2m} \in S^3$. Therefore $P \in \text{image}(C)$. Thus $SO(3, \mathbb{R}) \subseteq \text{image}(C)$. \square

Remark. This means that rotations in \mathbb{R}^3 can be represented by unit quaternions. This is a remarkably useful fact in computer graphics.

Physics Remark. The fact that any rotation can be written as conjugation by a unit quaternion is deeply connected to the fact that the Pauli spin matrices are generators of infinitesimal rotations of the Bloch sphere. We will understand this more deeply when we come to the sections about spinors.

Corollary 5.2. For all $u, v \in S^3$ there exists $r \in S^3$ so that $ru\bar{r} = v$.

Proof. Note that there exist $P, Q \in \text{SO}(3, \mathbb{R})$ so that $P(e_1) = u$ and $Q(e_1) = v$. Recall that P, Q are matrices, since they are linear maps from \mathbb{R}^3 to \mathbb{R}^3 .

To construct P and Q , take $u \in S^2 \subseteq \mathbb{R}^3 \subseteq \mathbb{R}^4$ and extend $\{u\}$ to an orthonormal basis for \mathbb{R}^3 , and similarly for v . Then let u be the first column of P and let v be the first column of Q . So $P(e_1) = u$ and $Q(e_1) = v$ by construction.

Due to the fact that $\text{SO}(3, \mathbb{R}) \subseteq \text{image } C$, there must exist $p, q \in S^3$ so that $P(e_1) = pe_1\bar{p} = u$ and $Q(e_1) = qe_1\bar{q} = v$. We get $pe_1 = up$ and $qe_1 = vq$. So $q\bar{p}u\bar{q}\bar{p} = v$. Let $r = q\bar{p} \in S^3$. Then $ru\bar{r} = v$.

Therefore, for all $u, v \in S^3$ there exists $r \in S^3$ so that $ru\bar{r} = v$. □

Corollary 5.3. Any unit quaternion $p \in S^3$ can be written as the product of two imaginary unit quaternions.

Proof. Let $p = \lambda 1 + \mu v_1$ where $\mu v = \text{Im}(p)$, $v_1 \in S^2$, and $\lambda 1 = \text{Re}(p)$. So we must have $|p|^2 = \lambda^2 + \mu^2 = 1$. By the previous corollary there exists $r \in S^3$ so that $re_1\bar{r} = v_1$. Define $v_2 = re_2\bar{r}$ and $v_3 = re_3\bar{r}$. Then,

$$\begin{aligned} \bar{r}(-e_2)r(\bar{r}(\lambda e_2 - \mu e_3)r) &= -\bar{r}e_2(\lambda e_2 - \mu e_3)r \\ &= -\bar{r}(\lambda(-1) - \mu e_1)r \\ &= \bar{r}(\lambda + \mu e_1)r \\ &= \lambda + \mu v_1 \\ &= p \end{aligned}$$

A direct calculation shows that $\bar{r}(-e_2)r \in S^2$ and $\bar{r}(\lambda e_2 - \mu e_3)r \in S^2$. Therefore we see that p is the product of two imaginary unit quaternions as required. □

We will now prove that in fact, $\text{SO}(3, \mathbb{R})$ is **equal** to the image of C , not just a subset of it. One way to show this is to show that any improper isometry can't be in the image of C . We will give three different proofs.

Theorem 5.1 (Unit Quaternions represent Rotations I). $\text{SO}(3, \mathbb{R}) = \text{image } C$

Proof using Topology. Topology is not required as a prerequisite for this course, so you may skip this proof if you like.

First, notice that $C : S^3 \rightarrow O(3, \mathbb{R})$ is a continuous map, since $C_p(x)$ is just a polynomial in the entries of p and x .

Since S^3 is a connected topological space, image C must be connected since C is continuous. So image C must be equal to one of the connected components of $O(3, \mathbb{R})$.

$O(3, \mathbb{R})$ can be split into two connected components, $SO(3, \mathbb{R})$ and $O^-(3, 0, \mathbb{R})$, the latter of which is composed of improper isometries.

Since $SO(3, \mathbb{R})$ contains the identity matrix, which is an element of image C , we see that image $C = SO(3, \mathbb{R})$ as required. \square

Proof by Contradiction. Suppose P is an improper isometry. Then by Cartan-Dieudonné theorem, P can be written as $P = R_1 \circ \dots \circ R_{2n+1}$, where each R_i is a reflection.

Observe that $R_1 \circ \dots \circ R_{2n}$ is already known to be an element of image C . Therefore, since image C is a subgroup of $O(3, \mathbb{R})$, we know that $P = R_1 \circ \dots \circ R_{2n+1} \in \text{image } C$ if and only if $R_{2n+1} \in \text{image } C$.

Therefore, what we must show is that the image of C does not contain any reflections.

Suppose this was not the case, and that $R_v \in \text{image } C$ for some v . Then $R_v = C_p$ for some $p \in S^3$. So for all $x \in \text{Im } \mathbb{H}$ we have

$$R_v(x) = x - 2\langle x, v \rangle v = px\bar{p}$$

Set $v_1 = v$ and let $\{v_1, v_2, v_3\}$ be an orthonormal basis of \mathbb{R}^3 . Earlier we saw how $vw + wv = -2\langle v, w \rangle$. So $v_i v_j + v_j v_i = -2\delta_{ij}$.

Set $x = v_k$. Then $v_k - 2\langle v_k, v_1 \rangle v_1 = pv_k \bar{p}$. For $k = 1$ we get $-v_1 = pv_1 \bar{p}$, so $v_1 p = -pv_1$. For $k = 2, 3$ we get $pv_k = v_k p$.

So p must commute with v_2, v_3 but anti-commute with v_1 .

Suppose $p = (a + bv_1 + cv_2 + dv_3)$. Then

$$(a + bv_1 + cv_2 + dv_3)v_2 = v_2(a + bv_1 + cv_2 + dv_3)$$

So

$$av_2 + bv_1 v_2 - c + dv_3 v_2 = av_2 + bv_2 v_1 - c + dv_2 v_3$$

Then

$$b(v_1 v_2 - v_2 v_1) + d(v_3 v_2 - v_2 v_3) = 0$$

We have $(bv_1 + dv_3)v_2 = 0$. So unless $b = d = 0$, we get a contradiction since v_1 and v_3 are linearly independent.

So $p = a + bv_2$. The same argument, except starting with,

$$(a + bv_1 + cv_2 + dv_3)v_3 = v_3(a + bv_1 + cv_2 + dv_3)$$

leads us to find that $b = c = 0$. So $p \in \mathbb{R}$. But this contradicts the fact that p anticommutes with v_1 .

Therefore, we have reached a contradiction. \square

Physics Remark. The equation $v_i v_j + v_j v_i = -2\delta_{ij}$ is quite similar to the canonical anti-commutation relation. In fact, with an appropriate choice of metric/basis this relationship can be made precise.

Physics Remark. Observe that the exterior algebra satisfies $v_i \wedge v_j + v_j \wedge v_i = 0$, while the imaginary quaternions satisfy $v_i v_j + v_j v_i = -2\delta_{ij}$. These are very similar, and the reason is that the imaginary quaternions are elements of a Clifford algebra, which is like a quantization of the exterior algebra. We will explore Clifford algebras very soon.

5.2 Lecture 14: Quaternions, Rotations, and Double Covers

5.2.1 Quaternions as Rotations

Theorem 5.2 (Unit Quaternions represent Rotations II). Let $\mathbf{x} \in \mathbb{R}^3, p \in S^3$ and let $\mathbf{u} = \mathbf{p}/\sqrt{1 - (p^0)^2}$. Observe that $|\mathbf{u}| = 1$. Suppose that $\mathbf{u} \perp \mathbf{x}$. Then

$$C_p(\mathbf{x}) = \cos(2\theta)\mathbf{x} + \sin(2\theta)\mathbf{u} \times \mathbf{x} \quad (5.9)$$

Where $\cos(2\theta) = p^0$. Similarly, suppose that $\mathbf{x} \in \text{span}\{\mathbf{u}\}$. Then

$$C_p(\mathbf{x}) = \mathbf{x} \quad (5.10)$$

This means that C_p is a rotation of the plane $\text{span}\{\mathbf{u}\}^\perp$ by an angle 2θ in the direction from \mathbf{x} to $\mathbf{u} \times \mathbf{x}$. The rotation axis is the line spanned by the vector \mathbf{u} .

Proofs of Theorems 5.1 and 5.2. In this proof we will determine the rotation defined by C_p , which will also allow us to show that $\text{image } C = \text{SO}(3, \mathbb{R})$.

Let $p = p^0 + \mathbf{p}$. Then we know $(p^0)^2 + |\mathbf{p}|^2 = 1$.

Since $-1 \leq p^0 \leq 1$ we can set $p^0 = \cos \theta$ for some θ and $\mathbf{p} = \sin \theta \mathbf{u}$ for some $\mathbf{u} \in S^3$.

Now suppose $x = \mathbf{x} \in \mathbb{R}^3$ is an imaginary quaternion.

We then get $uxu = x - 2\langle x, \mathbf{u} \rangle \mathbf{u}$. So

$$\begin{aligned} C_p(\mathbf{x}) &= px\bar{p} = (\cos \theta + \sin \theta \mathbf{u}) \times (\cos \theta - \sin \theta \mathbf{u}) \\ &= (\cos^2 \theta - \sin^2 \theta)\mathbf{x} + 2 \sin \theta \cos \theta \mathbf{u} \times \mathbf{x} + 2 \sin^2 \theta \langle \mathbf{x}, \mathbf{u} \rangle \mathbf{u} \end{aligned}$$

Observe that if $x = u$ then this implies $C_p u = u$. So u is fixed by C_p . Meanwhile, if x is perpendicular to u , then by the double angle identity we get

$$C_p(\mathbf{x}) = \cos(2\theta)\mathbf{x} + \sin(2\theta)\mathbf{u} \times \mathbf{x}$$

Since $p^0 = \cos \theta$, this also tells us that any value of p^0 between -1 and 1 gives us exactly two choices of θ , which are $\theta = \cos^{-1}(p^0)$ and $\theta = \pi - \cos^{-1}(p^0)$.

Furthermore, this acts as another proof of the fact that $\text{image } C = \text{SO}(3, \mathbb{R})$, since we have just shown that $C_p \in \text{SO}(3, \mathbb{R})$. \square

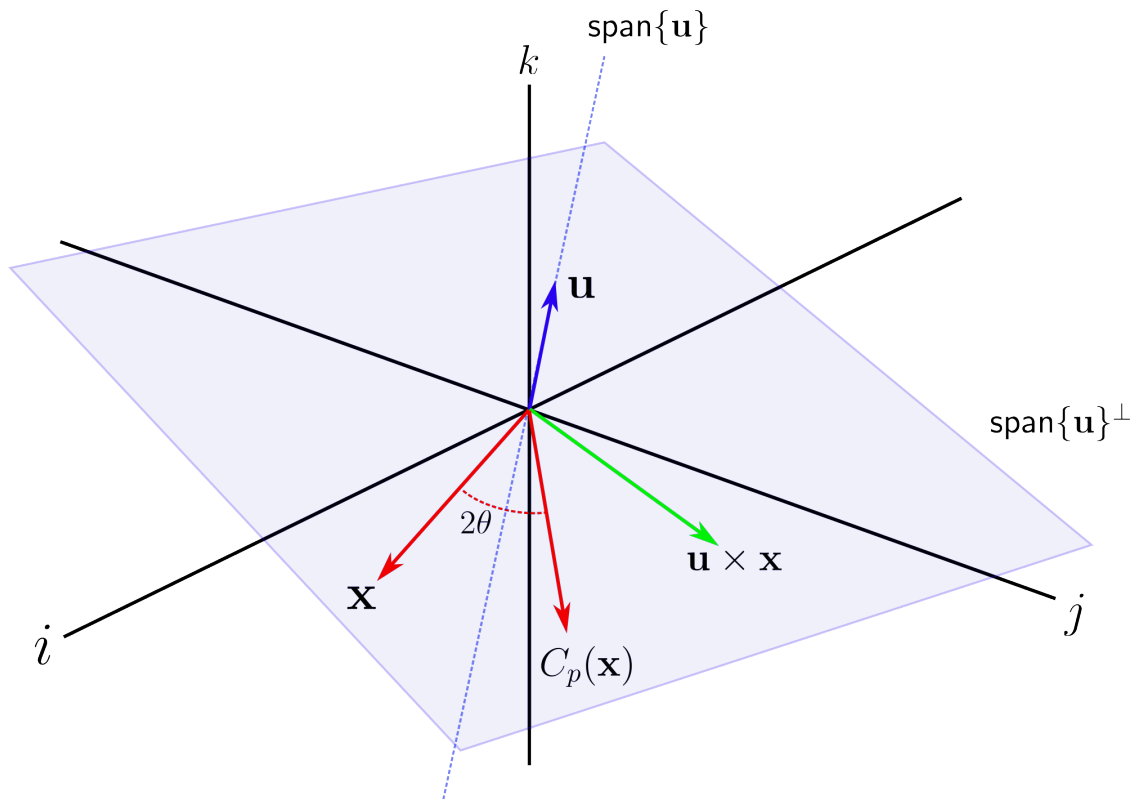


Figure 3: Diagram of rotation including the vectors $\mathbf{u}, \mathbf{x}, \mathbf{u} \times \mathbf{x}$.

5.2.2 Exact Sequences and Double Covers

Definition 5.11 (Exact Sequence). Let G_1, \dots, G_n be a list of groups, and suppose there are maps $f_0 : \{1\} \rightarrow G_1, f_1 : G_1 \rightarrow G_2, \dots, f_{n-1} : G_{n-1} \rightarrow G_n, f_n : G_n \rightarrow \{1\}$, where $\{1\}$ is the trivial group. That is, we have a diagram

$$1 \rightarrow G_1 \rightarrow \dots \rightarrow G_n \rightarrow 1$$

Suppose these maps satisfy $\ker f_i = \text{image } f_{i-1}$ for all $i = 1, \dots, n+1$. Then we say the sequence G_1, \dots, G_n forms an **exact sequence**.

Example 5.1. Observe that the set $\mathbb{Z}/2 = \{-1, 1\}$ is a group with respect to multiplication. Define

$$f_0 : 1 \rightarrow \mathbb{Z}/2$$

by $f_0(1) = 1$. Define

$$f_1 : \mathbb{Z}/2 \rightarrow S^3$$

by $f_1(1) = 1, f_1(-1) = -1$. Define

$$f_2 : S^3 \rightarrow \text{SO}(3, \mathbb{R})$$

by $f_2(p) = C_p$, and define

$$f_3 : \mathrm{SO}(3, \mathbb{R}) \rightarrow 1$$

by $f_3(R) = 1$.

This gives us an exact sequence of maps,

$$1 \rightarrow \mathbb{Z}/2 \rightarrow S^3 \rightarrow \mathrm{SO}(3, \mathbb{R}) \rightarrow 1 \quad (5.11)$$

Definition 5.12. Let A, B, C be groups. Let $\iota : A \rightarrow A \times C$ be given by $\iota(a) = (a, 1)$, and let $p : A \times C \rightarrow C$ be given by $p(a, c) = c$. Suppose that we have an exact sequence,

$$1 \rightarrow A \xrightarrow{a} B \xrightarrow{b} C \rightarrow 1$$

Then we say this sequence **splits** if there exists an isomorphism $f : B \rightarrow A \oplus C$ so that $f \circ a = \iota$ and $p \circ f = b$.

Theorem 5.3. The exact sequence (5.11) does not split. In other words, there does not exist any map $\varphi : \mathrm{SO}(3, \mathbb{R}) \rightarrow S^3$ so that $\varphi \circ C = \mathbb{1}$.

Proof. Suppose φ existed. Then $[S^3 : \text{image } \varphi] = 2$, so $\text{image } \varphi$ is a normal subgroup of S^3 .

By the first isomorphism theorem, this would mean $S^3 / \text{image } \varphi$ is isomorphic to $\mathbb{Z}/2$. Let $\rho : S^3 \rightarrow S^3 / \text{image } \varphi$ be the quotient map.

Recall that if $u, v \in S^2$, there exists $r \in S^3$ so that $rvr^{-1} = u$. Since $\mathbb{Z}/2$ is abelian, we have $\rho(u) = \rho(rvr^{-1}) = \rho(v)$ for all $u, v \in S^2$.

Now recall that a basis for \mathbb{H} is given by $\{e_0, e_1, e_2, e_3\}$. We have $\rho(e_3) = \rho(e_1)\rho(e_2)$, and since $u, e_1 \in S^2$ we have $\rho(u) = \rho(e_1) = \rho(e_2)^{-1}\rho(e_3) = \rho(e_2^{-1}e_2) = 1$. Therefore, $\rho(u) = 1$ for all $u \in S^2$.

Now, if $p \in S^3$, then recall that there exist $u, v \in S^2$ with $p = uv$. So $\rho(p) = 1$ as well.

But this means that ρ is not surjective, since $\text{image } \rho = \{1\} \neq S^3 / \text{image } \varphi$. This is a contradiction, so the proof is complete. \square

Remark. We say that since this sequence does not split, and since $\ker C$ has two elements, the map $C : S^3 \rightarrow \mathrm{SO}(3, \mathbb{R})$ is a **nontrivial double cover**.

Remark. As we will soon see, this is really a statement about the **spin group**, $\mathrm{Spin}(3) = S^3$.

Now we will return to the study of the map $K_{p,q}$.

Lemma 5.8. The sequence,

$$1 \rightarrow \mathbb{Z}/2 \rightarrow S^3 \times S^3 \xrightarrow{K} \mathrm{SO}(4, \mathbb{R}) \rightarrow 1$$

is exact.

Proof. All we need is for K to be surjective.

Topologically, we can see that since $\text{image } K$ contains the identity, we must have $\text{image } K \subseteq \text{SO}(4, \mathbb{R})$.

Furthermore, by Cartan-Dieudonné theorem any P in $\text{SO}(4, \mathbb{R})$ can be written as $P = R_{u_{2m}} \circ \dots \circ R_{u_1}$, with each $u_i \in S^3$.

Therefore since K is a homomorphism, if we can show that $R_{u_2} \circ R_{u_1} \in \text{image } K$ for all u_1, u_2 then it follows that $\text{SO}(4, \mathbb{R}) \subseteq \text{image } K$.

Suppose $u_1 = \pm u_2$. Then $R_{u_2} \circ R_{u_1} = \mathbb{1}$. So this case is done.

Now suppose u_1 and u_2 are linearly independent. Let $u_1 = v_1$. Then there exists v_2 with $u_2 = \cos \theta v_1 + \sin \theta v_2$, where $\theta = \arccos(\langle u_1, u_2 \rangle)$ is the angle between the two unit quaternions. Now let $x \in \mathbb{R}^4$. We have,

$$R_{u_2}(R_{u_1}(x)) = x - 2\langle x, u_1 \rangle u_1 - 2(\langle x - 2\langle x, u_1 \rangle u_1, u_2 \rangle) u_2$$

Exercise: check that if $x \perp \text{span}\{u_1, u_2\}$ then $R_{u_2}(R_{u_1}(x)) = x$, and if $x \in \text{span}\{u_1, u_2\}$ then $R_{u_2}(R_{u_1}(x)) = \cos(2\theta)x + \sin(2\theta)v_3$ for some v_3 .

Finally we just need to show that any rotation of this form is in the image of K . Let $L = \text{span}\{u_1, u_2\}^\perp$. Then $\mathbb{H} = L \oplus L^\perp$, where P fixes L and rotates L^\perp . This is done in two cases.

First, if $1 \in L$ then P must be in $\text{SO}(3, \mathbb{R})$ and $P = K_{p,p}$ for some P .

The other case: MISSING

□

Remark. This exact sequence does not split. That is, $S^3 \times S^3 \rightarrow \text{SO}(4, \mathbb{R})$ is a nontrivial double cover. We will soon see that this is a fact about spin groups, $S^3 \times S^3 = \text{Spin}(4)$.

6 Clifford Algebras

6.1 Lecture 15: Clifford Algebras

6.1.1 Interior Product Revisited

Recall that the interior product is defined by,

$$\lrcorner : V \times \Lambda^k V^* \rightarrow \Lambda^{k-1} V^*$$

where,

$$v \lrcorner \omega(u_1, \dots, u_{k-1}) = \omega(v, u_1, \dots, u_{k-1})$$

Given a metric G and $a \in \Lambda^k V$, we can also set,

$$v \lrcorner a(\alpha_2, \dots, \alpha_{k-1}) = a(v, \alpha_2, \dots, \alpha_{k-1})$$

Remark. Given this, we see that $v \lrcorner v = G(v, v)$ for any $v \in V$. Therefore the interior product is a generalization of the inner product.

Remark. An explicit formula is given as follows. Let $\alpha = \alpha_1 \wedge \dots \wedge \alpha_k$. Then,

$$\begin{aligned}
v \lrcorner \alpha(v_2, \dots, v_k) &= \alpha(v, v_1, \dots, v_{k-1}) \\
&= \sum_{\sigma \in S_k} \text{sgn}(\sigma) \alpha_{\sigma(1)}(v) \alpha_{\sigma(2)}(v_2) \dots \alpha_{\sigma(k)}(v_k) \\
&= \sum_{j=1}^k \sum_{\sigma \in S_k, \sigma(1)=j} \text{sgn}(\sigma) \alpha_{\sigma(1)}(v_1) (\alpha_{\sigma(2)}(v_2) \dots \alpha_{\sigma(k)}(v_k)) \\
&= \sum_{j=1}^k (-1)^{j+1} \alpha_j(v_1) \sum_{\sigma \in S_{k-1}} \text{sgn}(\sigma) \alpha_{i_{\sigma(1)}}(v_2) \dots \alpha_{i_{\sigma(k-1)}}(v_k)
\end{aligned}$$

Where $i_1 = 2, \dots, i_{j-1} = j-1, i_j = j+1, \dots, i_{k-1} = k$.

$$= \sum_{j=1}^k (-1)^{j+1} \alpha_j(v) (\alpha_1 \wedge \dots \wedge \alpha_{j-1} \wedge \alpha_{j+1} \wedge \dots \wedge \alpha_k(v_2, \dots, v_k))$$

In other words,

$$v \lrcorner \alpha = \sum_{j=1}^k (-1)^{j+1} \alpha_j(v) \alpha_1 \wedge \dots \wedge \alpha_{j-1} \wedge \alpha_{j+1} \wedge \dots \wedge \alpha_k \quad (6.1)$$

Here we have factored out the $\alpha_{\sigma(1)}(v)$ and set $j = \sigma(1)$. Then we can sum over $\sigma(1)$ separately from the other outputs, which effectively means we are summing over the permutations of $2, \dots, k$. Therefore we get the wedge product of all but $\alpha_{\sigma(1)}$ in the end.

Lemma 6.1. Let $v \in V$, $\alpha \in \Lambda^p V^*$, $\beta \in \Lambda^q V^*$. Then,

$$v \lrcorner (\alpha \wedge \beta) = (v \lrcorner \alpha) \wedge \beta + (-1)^p \alpha \wedge (v \lrcorner \beta) \quad (6.2)$$

In other words, the interior product is an **(anti)-derivation**. It obeys a "graded" version of the Leibniz rule.

Proof. By linearity it suffices to show this for wedge-decomposable forms. Let $\alpha = \alpha_1 \wedge \dots \wedge \alpha_p$, $\beta = \beta_1 \wedge \dots \wedge \beta_q$.

To make the proof easier, we will introduce the notation $\gamma_i = \alpha_i, i = 1, \dots, p$ and $\gamma_{i+k} = \beta_i, i = 1, \dots, q$. This way, $\alpha = \gamma_1 \wedge \dots \wedge \gamma_p$ and $\beta = \gamma_{p+1} \wedge \dots \wedge \gamma_{p+q}$. Then,

$$\begin{aligned}
(v \lrcorner \alpha) \wedge \beta &= \sum_{i=1}^p (-1)^{i+1} \gamma_i(v) \gamma_1 \wedge \dots \wedge \gamma_{i-1} \wedge \gamma_{i+1} \wedge \dots \wedge \gamma_p \wedge \gamma_{p+1} \wedge \dots \wedge \gamma_{p+q} \\
(-1)^p \alpha \wedge (v \lrcorner \beta) &= (-1)^p \sum_{i=p+1}^{p+q} (-1)^{i+1} \gamma_i(v) \gamma_1 \wedge \dots \wedge \gamma_p \wedge \dots \wedge \gamma_{i-1} \wedge \gamma_{i+1} \wedge \dots \wedge \gamma_{p+q} \\
&= \sum_{i=p+1}^{p+q} (-1)^{i+1} \gamma_i(v) \gamma_1 \wedge \dots \wedge \gamma_p \wedge \dots \wedge \gamma_{i-1} \wedge \gamma_{i+1} \wedge \dots \wedge \gamma_{p+q}
\end{aligned}$$

Adding these together, we get,

$$(v \lrcorner \alpha) \wedge \beta + (-1)^p \alpha \wedge (v \lrcorner \beta) = \sum_{i=1}^{p+q} (-1)^{i+1} \gamma_i(v) \gamma_1 \wedge \dots \wedge \gamma_{i-1} \wedge \gamma_{i+1} \wedge \dots \wedge \gamma_{p+q}$$

The right hand side is exactly $v \lrcorner (\alpha \wedge \beta)$ as required. \square

Theorem 6.1. Let G be a metric on V . Let $v \in V, \alpha \in \Lambda^{k-1}V, \beta \in \Lambda^k V$. Then,

$$G(v \wedge \alpha, \beta) = G(\alpha, v \lrcorner \beta) \quad (6.3)$$

In other words, the interior product is the **adjoint** of the exterior product.

Proof. By linearity it is enough to show that this is true for wedge-decomposable forms. Let $\alpha = \alpha_2 \wedge \dots \wedge \alpha_{k-1}, \beta = \beta_1 \wedge \dots \wedge \beta_k$. Then we have,

$$\begin{aligned} G(v \wedge \alpha, \beta) &= G(v \wedge \alpha_2 \wedge \dots \wedge \alpha_{k-1}, \beta_1 \wedge \dots \wedge \beta_k) \\ &= \det \begin{bmatrix} G(v, \beta_1) & \dots & G(v, \beta_k) \\ \vdots & \ddots & \vdots \\ G(\alpha_k, \beta_1) & \dots & G(\alpha_k, \beta_k) \end{bmatrix} \\ &= \sum_{j=1}^k (-1)^{j+1} G(v, \beta_j) \det[G(\alpha_i, \beta_j)]_{i \neq j} \\ &= \sum_{j=1}^k (-1)^{j+1} G(v, \beta_j) \beta_1 \wedge \dots \wedge \beta_{j-1} \wedge \beta_{j+1} \wedge \dots \wedge \beta_k (\alpha_2, \dots, \alpha_{k-1}) \\ &= v \lrcorner \beta (\alpha_2 \wedge \dots \wedge \alpha_{k-1}) \\ &= G(\alpha, v \lrcorner \beta) \end{aligned}$$

Where in line 3 we have used the cofactor expansion to sum over the first row of the overlap matrix, and from line 4 to 5 we have used the formula for the interior product derived above. \square

Lemma 6.2. Let G be a metric on V . Let $w \in V$ and $\alpha \in \Lambda^k V$. Then,

$$w \lrcorner (w \wedge \alpha) = |w|^2 \alpha \quad (6.4)$$

Proof. By the product rule for the interior product we have $w \lrcorner (w \wedge \alpha) = (w \lrcorner w) \wedge \alpha - w \wedge (w \lrcorner \alpha)$. We know $w \lrcorner w = w(b(w)) = G(w, w)$. So $(w \lrcorner w) \wedge \alpha = |w|^2 \alpha$.

Therefore we just need to show that the second term vanishes. We have,

$$\begin{aligned} w \wedge (w \lrcorner \alpha)(v_1, \dots, v_k) &= \left[w \wedge \sum_{j=1}^k (-1)^{j+1} \alpha_j(w) \alpha_1 \wedge \dots \wedge \alpha_{j-1} \wedge \alpha_{j+1} \wedge \dots \wedge \alpha_k \right] (v_1, \dots, v_k) \\ &= \left[\sum_{j=1}^k \alpha_j(w) \alpha_1 \wedge \dots \wedge \alpha_{j-1} \wedge w \wedge \alpha_{j+1} \wedge \dots \wedge \alpha_k \right] (v_1, \dots, v_k) \end{aligned}$$

□

Definition 6.1. We define the following notation.

Let $I_v : \Lambda^k V \rightarrow \Lambda^{k-1} V$ be given by $I_v(\alpha) = v \lrcorner \alpha$, and let $E_v : \Lambda^k V \rightarrow \Lambda^{k-1} V$ be given by $E_v(\alpha) = v \wedge \alpha$.

Remark. Note that $E_v^2 = I_v^2 = 0$.

Corollary 6.1. In the proof of lemma 6.2 we showed that $I_v \circ E_v + E_v \circ I_v = \langle v, v \rangle \mathbb{1}$.

6.1.2 Quotient Algebras

Here we will first go over the definition of a quotient algebra. The definition is almost the same as a quotient vector space, except in order for the quotient to form an algebra we will require that the denominator is more than just a vector space. We will require that it is an ideal.

Definition 6.2 (Ideal of an Algebra). Let A be an algebra. Then an **ideal** is a vector subspace B of A with the additional property that if $v \in A$ and $w \in B$ then $vw \in B$ and $wv \in B$.

In other words, an ideal is a vector subspace which *absorbs vector multiplication from the left and right*.

Definition 6.3. Let A be an algebra and let B be an ideal of A . Then the **quotient algebra** is the set of equivalence classes $[v] = \{v + b : b \in B\}$, and is denoted,

$$A/B = \{[v] : v \in A\}$$

Sometimes we use the notation $[v] = v + B$.

Remark. Observe that since B is an ideal, we have $wB = Bw = B$ for all $w \in A$, and $B^2 = \{bB : b \in B\} = B$, by the absorption property. Therefore,

$$[v][w] = (v + B)(w + B) = vw + vB + wB + BB = vw + B = [vw]$$

This means that the quotient of an algebra by an ideal is indeed a well defined algebra, and the multiplication rule from the original algebra still makes sense.

6.1.3 The Clifford Algebra

Recall definition 1.13 of the **tensor algebra**. It is given by

$$\bigotimes V = \bigoplus_{k=0}^{\infty} V^{\otimes k}$$

Each element of $\bigotimes V$ is of the form $\sum_{k=0}^N Y_k$ with each $Y_k \in V^{\otimes k}$.

Example 6.1. Let $V = \mathbb{R}^3$. Then, for example,

$$1 + e_1 + 3e_1 \otimes e_2 \otimes e_3 \in \bigotimes V$$

Remark. The set $\bigotimes V$ is a real algebra, where multiplication is given by the tensor product. The subset $\mathbb{R} \subseteq \bigotimes V$ forms a subalgebra.

Definition 6.4. Let V be a vector space with a metric $G(v, w)$ which we denote by $G(v, w) = \langle v, w \rangle$. The **Clifford ideal** of V is the set

$$I(V) = \left\{ \sum_{k=1}^N \alpha_k \otimes (v_k \otimes v_k + \langle v_k, v_k \rangle 1) \otimes \beta_k : \alpha_k, \beta_k \in \bigotimes V, v_k \in V \right\}$$

That is, $I(V)$ is the **ideal** generated by elements of the form $v \otimes v + \langle v, v \rangle 1$.

Definition 6.5 (Clifford Algebra). Let V be a vector space and let G be a metric on V . Then the **Clifford algebra** of V is the quotient,

$$\text{Cl}(V, G) = \frac{\bigotimes V}{I(V)} \quad (6.5)$$

It is associative.

Remark. The multiplication is denoted $\alpha \cdot \beta = [\alpha \otimes \beta] = \alpha \otimes \beta + I(V)$.

Theorem 6.2. Let $v, w \in \text{Cl}(V, G)$. Then,

$$v \cdot w + w \cdot v = -2G(v, w) = -2\langle v, w \rangle \quad (6.6)$$

And in particular,

$$v \cdot v = -Q(v) = -|v|^2 \quad (6.7)$$

For example, if A, B, C, D are taken to be groups, and all the functions in the diagram are group homomorphisms, then the above universal property says that $C = A \times B$ if and only if for all groups D and group homomorphisms $f_1 : D \rightarrow A, f_2 : D \rightarrow B$, there exists a group isomorphism $\phi : D \rightarrow C$ satisfying the equations $f_1 = \pi_A \circ \phi$ and $f_2 = \pi_B \circ \phi$. Similarly we can construct the equivalent of the product for vector spaces, algebras and so on.

Theorem 6.3 (Universal Property of the Clifford Algebra). Let V be a vector space over \mathbb{R} and let A be a unital algebra over \mathbb{R} . Suppose that $\phi : V \rightarrow A$ is a linear map, and suppose $\phi(x)\phi(x) = -G(x, x)1_A$ for all $x \in V$.

Suppose B is an algebra, and that there exists an injective linear map $\iota : V \rightarrow B$, and a unique algebra homomorphism $\tilde{\phi} : B \rightarrow A$ so that $\tilde{\phi} \circ \iota = \phi$. Then there is an algebra isomorphism between B and the Clifford algebra of V .

That is, $B \cong \text{Cl}(V, G)$ iff for all unital associative algebras A and all linear maps $\phi : V \rightarrow A$ with $\phi(x)\phi(x) = -G(x, x)1_A$, there exists a unique diagram

$$\begin{array}{ccc} V & \xrightarrow{\phi} & A \\ \downarrow \iota & \nearrow \tilde{\phi} & \\ B & & \end{array}$$

Where ι is injective, $\tilde{\phi} = \phi \circ \iota$, and $\tilde{\phi}$ is a homomorphism.

Proof. We must first show that $\bigotimes V/I(V)$ satisfies this property, and then show that if B satisfies the same property then $B \cong \bigotimes V/I(V)$.

Let $\phi : V \rightarrow A$ be a linear map.

First note that any algebra homomorphism $g : \bigotimes V \rightarrow A$ must satisfy the formula, $g(v_1 \otimes \dots \otimes v_k) = g(v_1) \dots g(v_k)$.

We then define an algebra homomorphism $\hat{\phi} : \bigotimes V \rightarrow A$ by the formula $\hat{\phi}(v_1 \otimes \dots \otimes v_k) = \phi(v_1) \dots \phi(v_k)$ for all v_1, \dots, v_k , and then extend it by linearity.

We can see that $\hat{\phi}(v) = \phi(v)$, so $\hat{\phi}|_V = \phi \circ \iota_0$, where $\iota_0 : V \rightarrow \bigotimes V$ is the inclusion map.

Suppose that we had any other algebra homomorphism $g : \bigotimes V \rightarrow A$ which satisfied $g|_V = \phi \circ \iota_0$. Then we would have $g(v) = \phi(v)$, so

$$\begin{aligned} \hat{\phi}(v_1 \otimes \dots \otimes v_k) &= \phi(v_1) \dots \phi(v_k) \\ &= g(v_1) \dots g(v_k) \\ &= g(v_1 \otimes \dots \otimes v_k) \end{aligned}$$

By linearity, we see that $\hat{\phi} = g$. Therefore $\hat{\phi}$ is unique.

Since $\phi(x)\phi(x) = -|x|^2 1$, we have $\phi(x \otimes x + |x|^2 1) = 0$. Therefore $I(V) \subseteq \ker \phi$. Let $q : \bigotimes V \rightarrow \bigotimes V/I(V)$ be the quotient map, and let $\tilde{\phi} : \bigotimes V/I(V) \rightarrow A$ be given by $\tilde{\phi}(v_1 \dots v_k) = \phi(v_1) \dots \phi(v_k)$. Then since $\ker \phi$ contains $I(V)$ we have $\tilde{\phi} = q \circ \hat{\phi}$. Since $\hat{\phi}$ is unique, $\tilde{\phi}$ is the unique extension of ϕ to $\text{Cl}(V)$. That is, we have shown that $\bigotimes V/I(V)$ satisfies the universal property of $\text{Cl}(V)$. This completes the proof of the first half.

Now we must prove that if A is any other algebra which satisfies the universal property of the Clifford algebra, then A is isomorphic to $\bigotimes V/I(V)$.

First, we have that any linear map $\varphi : V \rightarrow A$ with $\varphi(x)\varphi(x) = -|x|^2$ has a unique extension to an algebra homomorphism $\varphi : \text{Cl}(V) \rightarrow A$. \square

Corollary 6.2. Let $\{e_1, \dots, e_n\}$ be an orthonormal basis for V . That is, $\langle e_i, e_j \rangle = \pm \delta_{ij}$. Suppose $\varphi : V \rightarrow A$ satisfies $\varphi(e_j^2) = -|e_j|^2$ and $\varphi(e_i)\varphi(e_j) + \varphi(e_j)\varphi(e_i) = 0$. Then φ has a unique extension to $\varphi : \text{Cl}(V) \rightarrow A$.

Proof. We have $\varphi(e_i \otimes e_i + |e_i|^2 1) = 0$ by definition. Using the polarization identity and linearity, one can expand each term in $x \otimes y + y \otimes x + 2\langle x, y \rangle 1$ in terms of a basis, before applying φ to show that $\varphi(x)\varphi(y) + \varphi(y)\varphi(x) = -2\langle x, y \rangle 1$. So $\ker \varphi$ contains $I(V)$. This completes the proof. \square

Theorem 6.4. The vector space $\Lambda^\bullet V$ is canonically isomorphic as a vector space to $\text{Cl}(V)$.

Furthermore, if $v, w \in \text{Cl}(V)$, then $vw = v \wedge w - v \lrcorner w$.

We will delay the proof until after some remarks.

Remark. Let $\phi : \Lambda^\bullet V \rightarrow \text{Cl}(V)$ be the canonical isomorphism. Let $\iota : V \rightarrow \text{Cl}(V)$ be the inclusion map. Then $\iota \circ \phi$ is injective.

Remark. It is not isomorphic as an algebra to $\text{Cl}(V)$ because they have different multiplications.

Definition 6.7 (Clifford Product). When we interpret elements v, w of $\text{Cl}(V)$ as elements of $\Lambda^\bullet V$, the product vw is referred to as the **Clifford product**.

Remark. The Clifford algebra is not a graded algebra (recall Definition 1.21) since the Clifford product of two k -forms is not a $k + \ell$ -form.

Remark. Some books define $\text{Cl}(V)$ to be generated by the ideal $I(V) = \langle v \otimes v - |v|^2 1 \rangle$ instead, in which case the Clifford product is $vw = v \wedge w + v \lrcorner w$. This is equivalent to replacing the metric G with $-G$, or replacing the signature (p, q) with (q, p) .

Now we will get to the proof of the theorem.

Proof. We first show that the restriction of the quotient map $q : \otimes V \rightarrow \text{Cl}(V)$ to the subspace $\Lambda^\bullet V \subset \otimes V$ is surjective. That is, every equivalence class can be represented by a skew-symmetric tensor.

For a tensor of degree 0 or 1 this is trivial. Let $\beta \in \otimes V$ be decomposable. Then

$$\beta = \frac{1}{2}(x \otimes y + y \otimes x) + \frac{1}{2}(x \otimes y - y \otimes x)$$

Adding and subtracting $G(x, y)1$, we can get this to be in the form

$$\beta = \frac{1}{2}(x \otimes y + y \otimes x + 2G(x, y)1) - G(x, y)1$$

Where the first term is an element of $I(V)$ and the second term is an element of $\Lambda^0 V$.

Inductively we can show that any decomposable k -tensor can be written as something in $I(V)$ plus something in $\Lambda^0 V$.

We have a surjection from $\Lambda^\bullet V$ to $\text{Cl}(V)$. Now we need to construct a surjection from $\text{Cl}(V)$ to $\Lambda^\bullet V$.

We will use the universal property to construct a map $\varphi : V \rightarrow \text{End}(\Lambda^\bullet V)$. First we remark that $\text{End}(\Lambda^\bullet V)$ is an algebra over \mathbb{R} with identity. Then set $\varphi(x) = [E_x - I_x]$ for $x \in V$. Then $\varphi(v)(w) = v \wedge w - w \wedge v$ and so $\varphi(v) \circ \varphi(w) = [E_x - I_x]^2 = -|x|^2 \mathbb{1}$. So φ extends uniquely to an algebra homomorphism $\varphi : \text{Cl}(V) \rightarrow \text{End}(\Lambda^\bullet V)$.

Now let us define $\psi : \text{End}(\Lambda^\bullet V) \rightarrow \Lambda^\bullet V$ by $\psi(A) = A(1)$. Observe that this map is linear. Let us define $L = \frac{1}{k!} \psi \circ \varphi : \text{Cl}(V) \rightarrow \Lambda^\bullet V$, which is linear.

Let $\alpha = v_1 \wedge \dots \wedge v_k \in \Lambda^\bullet V$. We have

$$\begin{aligned} L \circ q(\alpha) &= L([\alpha]) \\ &= L([v_1 \wedge \dots \wedge v_k]) \\ &= L\left(\left[\sum_{\sigma \in S_k} \text{sgn}(\sigma) v_{\sigma(1)} \otimes \dots \otimes v_{\sigma(k)}\right]\right) \\ &= L\left(\sum_{\sigma \in S_k} \text{sgn}(\sigma) v_{\sigma(1)} \dots v_{\sigma(k)}\right) \\ &= \frac{1}{k!} \psi \circ \varphi \left(\sum_{\sigma \in S_k} \text{sgn}(\sigma) v_{\sigma(1)} \dots v_{\sigma(k)}\right) \\ &= \frac{1}{k!} \psi \left(\sum_{\sigma \in S_k} \text{sgn}(\sigma) (E_{v_{\sigma(1)}} - I_{v_{\sigma(1)}}) \dots (E_{v_{\sigma(k)}} - I_{v_{\sigma(k)}})\right) \\ &= \frac{1}{k!} \sum_{\sigma \in S_k} \text{sgn}(\sigma) (E_{v_{\sigma(1)}} - I_{v_{\sigma(1)}}) \dots (E_{v_{\sigma(k)}} - I_{v_{\sigma(k)}})(1) \\ &= \frac{1}{k!} \sum_{\sigma \in S_k} \text{sgn}(\sigma) (E_{v_{\sigma(1)}} - I_{v_{\sigma(1)}}) \dots (E_{v_{\sigma(k-1)}} - I_{v_{\sigma(k-1)}})(v_{\sigma(k)}) \end{aligned}$$

Note that $(E_{v_{\sigma(k-1)}} - I_{v_{\sigma(k-1)}})(v_{\sigma(k)}) = v_{\sigma(k-1)} \wedge v_{\sigma(k)} - G(v_{\sigma(k-1)}, v_{\sigma(k)})$. In the above sum, the term $(v_{\sigma(k-1)}, v_{\sigma(k)})$ will always appear twice, with opposite signs. Since G is symmetric, this means all the scalar terms vanish.

$$\begin{aligned} &= \frac{1}{k!} \sum_{\sigma \in S_k} \text{sgn}(\sigma) v_{\sigma(1)} \wedge \dots \wedge v_{\sigma(k)} \\ &= v_1 \wedge \dots \wedge v_k \end{aligned}$$

Therefore, $L \circ q$ is an isomorphism. Since q is surjective, this means L has to be injective. Therefore, we have an injection $L : \Lambda^\bullet V \rightarrow \text{Cl}(V)$.

Finally, we have $\varphi(x\alpha) = \varphi(x)\varphi(\alpha)$, so $\psi(\varphi(x\alpha)) = \varphi(x\alpha)1 = \varphi(x)\psi(\varphi(\alpha))$. So $L(x\alpha) = x \wedge L(\alpha) - x \lrcorner L(\alpha)$. Therefore, for

any $x \in V, \alpha \in \Lambda^\bullet V$ we know that L maps $x\alpha$ to $x \wedge \alpha - x \lrcorner \alpha$. We also know

$$\begin{aligned} q \circ L(v_1 \dots v_k) &= q(v_1 \wedge \dots \wedge v_k) \\ &= v_1 \dots v_k \end{aligned}$$

So $q \circ L = \mathbb{1}$ and $L \circ q = \mathbb{1}$. This means $q = L^{-1}$. In other words, L is an isomorphism. \square

Remark. We can therefore just think of $Cl(V)$ as the vector space $\Lambda^\bullet V$ along with the new multiplication rule $v\alpha = v \wedge \alpha - v \lrcorner \alpha$ for $v \in V$ and $\alpha \in \Lambda^\bullet V$.

Example 6.3. If $\alpha = 1$ then $v\alpha = v$.

If $v, w \in V$ then $vw = v \wedge w - \langle v, w \rangle 1$.

If $v, w, u \in V$ then

$$\begin{aligned} u(vw) &= u(v \wedge w - \langle v, w \rangle 1) \\ &= u \wedge v \wedge w - \langle v, w \rangle u - u \lrcorner v \wedge w + u \lrcorner \langle v, w \rangle 1 \\ &= u \wedge v \wedge w - \langle v, w \rangle u - \langle u, v \rangle w + \langle u, w \rangle v \end{aligned}$$

The Clifford algebra has a lot of extra structure, such as involutions, automorphisms, and anti-automorphisms. Also, a special case of the universal property is when A is also a Clifford algebra.

Theorem 6.5. Let V_1, G_1 and V_2, G_2 be two vector spaces with metrics. Then for each isometry $P : V_1 \rightarrow V_2$ we have a unique extension to an algebra homomorphism $Cl(V_1) \rightarrow Cl(V_2)$.

Proof. We have $P(x)P(x) = -|P(x)|^2 1 = -|x|^2 1$. So P satisfies the universal property. \square

Remark. Each linear map $T : V_1 \rightarrow V_2$ has a unique extension to an algebra homomorphism $T : \Lambda^\bullet V_1 \rightarrow \Lambda^\bullet V_2$ (given by the exterior power). This extension is the same as the extension to a map $Cl(V_1) \rightarrow Cl(V_2)$ as long as T preserves the metric.

Remark. If $Q : V_2 \rightarrow V_3$ is another isometry then $Q \circ P$ is an isometry and has a unique extension to a map $Q \circ P : Cl(V_1) \rightarrow Cl(V_3)$. It is equal to $\tilde{Q} \circ \tilde{P}$.

Definition 6.8. Let $\text{Aut}_{\text{alg}}(Cl(V))$ denote the set of all algebra automorphisms of $Cl(V)$.

The set of **Clifford automorphisms** is the set of all automorphisms which map V into V , and is denoted $\text{Aut}(Cl(V))$.

$$\text{Aut}(Cl(V)) = \{P \in \text{Aut}_{\text{alg}}(Cl(V)) | P(V) \subseteq V\} \quad (6.8)$$

Theorem 6.6. The set of Clifford automorphisms of V is equal to the orthogonal group of V . In other words $\text{Aut}(Cl(V)) = O(V)$.

Proof. Suppose $P \in O(V)$. Then P maps V to V and preserves the inner product. Therefore, P extends to a map $Cl(V) \rightarrow Cl(V)$, and since P is invertible the extension is invertible. So P is a Clifford automorphism.

Now suppose $P \in \text{Aut}(Cl(V))$. Then

$$\begin{aligned}\langle P(x), P(x) \rangle &= -P(x)P(x) \\ &= -P(x^2) \\ &= P(-x^2) \\ &= P(\langle x, x \rangle 1) \\ &= \langle x, x \rangle\end{aligned}$$

So $P \in O(V)$.

Therefore every Clifford automorphism is an extension of an isometry of V . □

6.1.5 Clifford Involutions

We will now see that $Cl(V)$ comes equipped with a special Clifford automorphism and two anti-automorphisms.

Definition 6.9 (Canonical Clifford Automorphism). The isometry $P(v) = -v$ extends to an automorphism of $Cl(V)$. We will denote the extension by the tilde symbol, so $\tilde{P}(\alpha) = \tilde{\alpha}$. It is called the canonical Clifford automorphism. Notice that $-(-v) = v$, so $\tilde{\tilde{v}} = v$. So the canonical automorphism is an involution.

Definition 6.10 (Even/Odd parts of the Clifford Algebra). The even/odd parts of the Clifford algebra are defined by,

$$Cl^{\text{even}}(V) = \{\alpha \in Cl(V) : \tilde{\alpha} = \alpha\} \tag{6.9}$$

$$Cl^{\text{odd}}(V) = \{\alpha \in Cl(V) : \tilde{\alpha} = -\alpha\} \tag{6.10}$$

Lemma 6.3. Since $\alpha = \frac{1}{2}(\alpha + \tilde{\alpha}) + \frac{1}{2}(\alpha - \tilde{\alpha})$ we have

$$Cl(V) = Cl^{\text{even}}(V) \oplus Cl^{\text{odd}}(V)$$

Recall that $\Lambda^\bullet V$ is a graded algebra, and we have $\Lambda^\bullet V = \Lambda^{\text{even}} V \oplus \Lambda^{\text{odd}} V$, where $\Lambda^{\text{even}} V = \bigoplus \Lambda^{2k} V$ and $\Lambda^{\text{odd}} V = \bigoplus \Lambda^{2k+1} V$.

Lemma 6.4. There exists vector space isomorphisms $\Lambda^{\text{even}} V \cong Cl^{\text{even}}(V)$ and $\Lambda^{\text{odd}} V \cong Cl^{\text{odd}}(V)$.

Proof. Let $\{e_1, \dots, e_n\}$ be an orthonormal basis. Then $e_{i_1} \dots e_{i_k} = e_{i_1} \wedge \dots \wedge e_{i_k}$ since $e_{i_j} \lrcorner e_{i_\ell} = 0$ for all j, ℓ . This means $\widetilde{e_{i_1} \dots e_{i_k}} = \tilde{e}_{i_1} \dots \tilde{e}_{i_k} = (-1)^k e_{i_1} \dots e_{i_k}$. So the parity of the form $\alpha = e_{i_1} \dots e_{i_k}$ with respect to the canonical automorphism is the same as the parity of α with respect to the grading on $\Lambda^k V$. Extending by linearity completes the proof. \square

Definition 6.11 (Canonical Anti-Automorphism). Consider the map $T : \otimes V \rightarrow \otimes V$ defined by,

$$T(v_1 \otimes \dots \otimes v_k) = v_k \otimes v_{k-1} \otimes \dots \otimes v_2 \otimes v_1$$

That is, it reverses the order of all components. This map is clearly an involution, and we can also see that it is an anti-automorphism with respect to the tensor product since $T(\alpha \otimes \beta) = T(\beta) \otimes T(\alpha)$. We write $T(\alpha) = \check{\alpha}$.

6.2 Lecture 16: Symmetries and Automorphisms of Clifford Algebras

6.2.1 Clifford Involutions (Continued)

Lemma 6.5. Let $T(\alpha) = \check{\alpha}$ and $P(\alpha) = \tilde{\alpha}$. Then $T \circ P \circ T$ is an automorphism. Furthermore, since $(T \circ P \circ T)|_V = P|_V$, by uniqueness we see that $T \circ P \circ T = P$.

Definition 6.12 (Hat Anti-Automorphism). Let $T(\alpha) = \check{\alpha}$ and $P(\alpha) = \tilde{\alpha}$. Then we define $\hat{\alpha} = T(P(\alpha))$.

Remark. By the above lemma, $T(P(\alpha)) = P(T(\alpha))$. That is, $\hat{\alpha} = \check{\tilde{\alpha}} = \tilde{\check{\alpha}}$.

Theorem 6.7. Let $a \in Cl(V)$ be of degree k . Then $\tilde{a} = \pm a$, $\check{a} = \pm a$, $\hat{a} = \pm a$, where the signs are determined by k according to the following table,

$k \bmod 4$	0	1	2	3
\sim	+	-	+	-
\vee	+	+	-	-
\wedge	+	-	-	+

Proof. Write down an orthonormal basis, and check directly. \square

These involutions we have discussed have some important properties with respect to the metric induced on $Cl(V)$ by the metric on V .

Remark. Recall that the metric G on V induces a metric on $\Lambda^\bullet V$ (see equation 3.2). Since $\Lambda^\bullet V \cong Cl(V)$ as vector spaces, we therefore also have a metric on $Cl(V)$.

Remark. If $\{e_1, \dots, e_n\}$ is an orthonormal basis for V then $\{e_{i_1} \cdot \dots \cdot e_{i_k} | 1 \leq i_1, \dots, i_k \leq n\}$ is an orthonormal basis for $Cl(V)$.

Theorem 6.8. Let $a, b \in Cl(V)$. Then,

$$G(a, b) = G(1, \widehat{ab})$$

$$G(a, b) = G(1, \widehat{ba})$$

Proof. Check directly using the aforementioned orthonormal basis. □

Corollary 6.3. The map $L_a : Cl(V) \rightarrow Cl(V)$ defined by $L_a(b) = ab$ satisfies $L_a^\dagger = L_{\widehat{a}}$. In other words, left-multiplication by a is adjoint to left-multiplication by \widehat{a} .

Proof.

$$G(ab, c) = G(1, \widehat{abc})$$

$$= G(1, \widehat{b\widehat{a}c})$$

$$= G(b, \widehat{ac})$$

□

Corollary 6.4. Let $a = v_1 \dots v_k$. Then $|a|^2 = G(a, a) = |v_1|^2 \dots |v_k|^2$.

Proof. Recall that Clifford multiplication satisfies $-v^2 = |v|^2 1$. Therefore we have,

$$\widehat{a}a = \widehat{v_1 \dots v_k} v_1 \dots v_k$$

$$= (-v_k) \dots (-v_1) v_1 \dots v_k$$

$$= |v_1|^2 \dots |v_k|^2$$

□

Theorem 6.9 (Hadamard Identity). Let v_1, \dots, v_k be linearly independent. Then,

$$|v_1|^2 \dots |v_k|^2 = |v_1 \wedge \dots \wedge v_k|^2 + \sum_{j=1}^{k-1} |\pi_j(v_1 \dots v_k)|^2 \quad (6.11)$$

Where $\pi_j(v_1 \dots v_k)$ is the projection of $v_1 \dots v_k$ onto $\Lambda^j V \subset \Lambda^\bullet V = Cl(V)$.

Corollary 6.5 (Hadamard Inequality). Suppose G is a positive definite metric. The above equation then implies that

$$|v_1 \wedge \dots \wedge v_k| \leq |v_1| \dots |v_k| \quad (6.12)$$

This is often referred to as the Hadamard inequality.

Proof. Note that by definition,

$$\sum_{j=1}^k \pi_j(v_1 \dots v_k) = v_1 \dots v_k$$

Since each $\Lambda^\bullet V$ is an orthogonal direct sum over j , we therefore have

$$\left| \sum_{j=1}^k \pi_j(v_1 \dots v_k) \right|^2 = |v_1 \wedge \dots \wedge v_k|^2 + \sum_{j=1}^{k-1} |\pi_j(v_1 \dots v_k)|^2$$

as required. □

6.2.2 Clifford Algebras as Subspaces of Clifford Algebras

Definition 6.13. We define the following notation,

$$\mathbb{R}^{p,q} = (\mathbb{R}^{p+q}, G)$$

Where G is any metric of signature (p, q) .

Definition 6.14. We define the following notation,

$$Cl(p, q) = Cl(\mathbb{R}^{p,q})$$

Remark. Up to isometry it does not matter what specific metric G we choose.

Physics Remark. One of the most commonly used vector spaces in physics is $\mathbb{R}^{3,1}$. This is commonly referred to as **Minkowski space**.

Let $e_0, e_1, \dots, e_p, e_{p+1}, \dots, e_{p+q}$ be the standard basis of $\mathbb{R}^{p+1,q}$, where e_{p+1}, \dots, e_q are the negative definite components and e_0, \dots, e_p are the positive definite components. Let us define a map $\phi : \mathbb{R}^{p,q} \rightarrow Cl(p+1, q)$ by the formula $\phi(e_j) = e_0 e_j$, where $j = 1, \dots, n$. Then,

$$\phi(e_i)\phi(e_i) = e_0 e_i e_0 e_i = -e_0^2 e_i^2 = e_i^2 = -|e_i|^2 1$$

$$\phi(e_i)\phi(e_j) = e_i e_j = -\phi(e_j)\phi(e_i)$$

By the universal property, ϕ extends to a unique algebra homomorphism $\phi : Cl(p, q) \rightarrow Cl(p+1, q)$.

Similarly, suppose $e_0, \dots, e_p, e_{p+1}, \dots, e_{p+q}$ is the standard basis of $\mathbb{R}^{p,q+1}$ where now e_0 is part of the negative-definite subspace. Suppose we define a map $\psi : \mathbb{R}^{p,q} \rightarrow \text{Cl}(p, q+1)$ by $\psi(e_j) = e_0 e_j$. Then ψ extends to a unique algebra homomorphism $\phi : \text{Cl}(p, q) \rightarrow \text{Cl}(q, p+1)$.

Theorem 6.10. We have,

1. The map ϕ defined in the above paragraph is an isomorphism from $\text{Cl}(p, q)$ to $\text{Cl}^{\text{even}}(p+1, q)$.
2. The map ψ is an isomorphism from $\text{Cl}(p, q)$ to $\text{Cl}^{\text{even}}(q, p+1)$.

Proof. Since $\phi(e_j) = e_0 e_j$, it is clear that ϕ takes elements of V to even elements of the Clifford algebra.

Next, notice that $e_i e_j = \pm e_0 e_i e_0 e_j = \pm \phi(e_i e_j)$, so ϕ is surjective.

Furthermore, since ϕ simply inserts copies of e_0 into a given input, we have that $\hat{\phi}(\alpha) = \phi(\hat{\alpha})$ since the order doesn't matter, and the sign pulls out.

Finally, we note that ϕ sends even elements to even elements, since it would insert an even number of e_0 's into a given input which all cancel. Additionally, when we act ϕ on an odd element (of degree k for example), we get an odd number of e_0 's, so all but one cancel and the result is an element of degree $k+1$. So the output of ϕ always has even degree.

Therefore we see that since ϕ is surjective, its' image is exactly $\text{Cl}^{\text{even}}(p+1, q)$, which completes the proof.

The proof for the other map is identical, except we see that since e_0 is part of the negative definite part of $\mathbb{R}^{p,q+1}$ the definiteness of the output of ϕ always gets swapped. This is why the output is $\text{Cl}^{\text{even}}(q, p+1)$. \square

6.2.3 Center and Twisted Center of Clifford Algebras

Definition 6.15 (Twisted Center). Recall that the center of $\text{Cl}(V)$ is the set of all elements of $\text{Cl}(V)$ which commute with every other element of $\text{Cl}(V)$,

$$\text{cent}(\text{Cl}(V)) = \{a \in \text{Cl}(V) : ab = ba \ \forall b \in \text{Cl}(V)\}$$

We similarly define the **twisted center** of $\text{Cl}(V)$ to be all the elements which anti-commute with everything else,

$$\text{twcent}(\text{Cl}(V)) = \{a \in \text{Cl}(V) : ab = -ba \ \forall b \in \text{Cl}(V)\}$$

Theorem 6.11. We have,

1. If $\dim V = n$ is even, then $\text{cent}(\text{Cl}(V)) = \Lambda^0 V$ and $\text{twcent}(\text{Cl}(V)) = \Lambda^n V$.
2. If $\dim V = n$ is odd, then $\text{cent}(\text{Cl}(V)) = \Lambda^0 V \oplus \Lambda^n V$ and $\text{twcent}(\text{Cl}(V)) = \{0\}$.

Proof. Let $v \in V$ be non-null. We say an element $x \in \text{Cl}(V)$ does not involve v if there is no way to write $x = \sum x_i$ where $x_i = x_{i1} \dots v \dots x_{ik}$ for some i, k . That is, x does not involve v if the value of x does not depend at all on the value of v .

Let $a \in \text{Cl}(V)$ and set $a = b + vc$ where b and c don't involve v . Then

$$\begin{aligned} av &= (b + vc)v = bv + vcv \\ va &= vb + v^2c = vb - |v|^2c \end{aligned}$$

Suppose a and b are even while c is odd. Then

$$\begin{aligned} av &= vb + |v|^2c \\ va &= vb - |v|^2c \end{aligned}$$

From this we see that $va = av$ if and only if $c = 0$ (in other words, a does not involve v). Similarly $va = -av$ iff $b = 0$ (which would mean a involves v).

Now suppose a and b are odd, while c is even. Then

$$\begin{aligned} av &= -vb - |v|^2c \\ va &= vb - |v|^2c \end{aligned}$$

So $av = va$ iff $b = 0$, and $av = -va$ iff $c = 0$.

Overall, if a is even and $a \in \text{cent}(\text{Cl}(V))$ then $av = va$ for all v , which means that a can't possibly involve v . So there is no way to decompose a in terms of vectors, which means a must be a scalar. In other words, $a \in \Lambda^0 V$ as required.

Similarly, suppose a is odd and $a \in \text{cent}(\text{Cl}(V))$. Then $av = va$ for all v , which means that a must involve v for all v . In other words, a can be decomposed into a product of all linearly independent vectors. So $a \in \Lambda^n V$. Furthermore, since a is odd, this means a can only be nonzero if n is odd.

Therefore, we have shown that $\text{cent}(\text{Cl}(V)) = \Lambda^0 V$ if n is even and $\text{cent}(\text{Cl}(V)) = \Lambda^0 V \oplus \Lambda^n V$ if n is odd.

Finally, by the exact same argument we see that if a is even and $a \in \text{twcent}(\text{Cl}(V))$ then a must involve all vectors v , which means $a \in \Lambda^n V$ and so n must be even for a to be nonzero. Similarly if a is odd then $av = -va$ iff a does not involve any vectors, which means that a is either a scalar or zero. Since scalars have even degree, a must be zero. This completes the proof. □

6.3 Lecture 17: The Spin Group and Pin Group

6.3.1 Self-Dual and Anti-Self-Dual Forms

Remark. Let μ be a metric-compatible orientation (i.e. a volume form) on $\mathbb{R}^{p,q}$. Recall that this means $G(\mu, \mu) = |\mu|^2 = (-1)^q$. Then by Theorem 6.7 we know that $\hat{\mu} = \pm\mu$ depends only on $p + q \bmod 4$. So $\mu\mu = \pm\hat{\mu}\mu = \pm|\mu| = \pm 1$, where the sign depends on n, p , and q .

Lemma 6.6. Let μ be a volume form on $\mathbb{R}^{p,q}$. Then,

1. $\mu^2 = 1$ if $p - q = 0 \bmod 4$ or $p - q = 3 \bmod 4$.

2. $\mu^2 = -1$ if $p - q = 1 \pmod{4}$ or $p - q = 2 \pmod{4}$.

Proof. We have $\mu = e_1 \dots e_n$ for some orthonormal basis $\{e_1, \dots, e_n\}$. This means $\check{\mu}\mu = (-1)^p$.

We also know that $\check{\mu} = \mu$ if $n = 0$ or $1 \pmod{4}$ and $\check{\mu} = -\mu$ if $n = 2$ or $3 \pmod{4}$. So if $n = 2k$, then $\check{\mu} = (-1)^k \mu$ and if $n = 2k + 1$ then $\check{\mu} = (-1)^k \mu$ as well.

Therefore, $\mu^2 = \mu\check{\mu}(-1)^k = (-1)^{p-k} 1$.

When n is even, this becomes $(-1)^{(p-q)/2}$ and when n is odd it becomes $(-1)^{(p-q+1)/2}$. One can now check that this gives the desired result by substituting in each of $p - q = 1, 2, 3, 4$. □

Remark. When $p - q = 1$ or $2 \pmod{4}$, then μ is a **complex structure** on $\mathcal{Cl}(p, q)$.

Definition 6.16. Let μ be an orientation satisfying $\mu^2 = 1$ and let $a \in \mathcal{Cl}(p, q)$

1. If $\mu a = a$ we say a is **self-dual**.
2. If $\mu a = -a$ we say a is **anti-self-dual**.

The vector subspace of all self-dual forms is denoted $\mathcal{Cl}^+(p, q)$ and the subspace of all anti-self-dual forms is denoted $\mathcal{Cl}^-(p, q)$.

Remark. The sets $\mathcal{Cl}^+(p, q)$ and $\mathcal{Cl}^-(p, q)$ are vector subspaces, but not necessarily subalgebras.

Remark. Since $\mu^2 = 1$ we require that $p - q = 0$ or $3 \pmod{4}$ in order to define these subspaces.

Remark. By setting $a = \frac{1}{2}(a + \mu a) + \frac{1}{2}(a - \mu a)$ we see that $\mathcal{Cl}(p, q) = \mathcal{Cl}^+(p, q) \oplus \mathcal{Cl}^-(p, q)$.

Definition 6.17 (Left/Right Ideals). Let A be an algebra and let I be an ideal. We say I is a two-sided ideal since $aI = I$ and $Ia = I$ for all $a \in A$.

A weaker definition is that of left and right ideals. I is said to be a **right ideal** if $Ia = I$ for all $a \in A$. It is said to be a **left ideal** if $aI = I$ for all $a \in A$.

Lemma 6.7. 1. When $p - q = 3 \pmod{4}$, the subspaces $\mathcal{Cl}^\pm(p, q)$ are two-sided ideals.

2. When $p - q = 0 \pmod{4}$, the subspaces $\mathcal{Cl}^\pm(p, q)$ are right ideals.

Proof. Let $\alpha \in \mathcal{Cl}^\pm(p, q)$ and let $b \in \mathcal{Cl}(p, q)$. Then $\mu\alpha = \pm\alpha$, so

$$\mu(\alpha b) = (\mu\alpha)b = \pm\alpha b$$

So $\alpha b \in \mathcal{Cl}^\pm(p, q)$. Therefore $\mathcal{Cl}^\pm(p, q)$ is always a right ideal.

Now we show that if $p - q = 3 \bmod 4$ then it is also a left ideal. Note that when this is the case, $\mu b = b\mu$, so

$$\mu(b\alpha) = b(\mu\alpha) = b(\pm\alpha) = \pm b\alpha$$

This completes the proof. □

Remark. Why isn't $Cl^\pm(p, q)$ two-sided when $p - q = 0 \bmod 4$? The answer is that in general, if $p - q = 0 \bmod 4$ then $\mu(b\alpha) = \pm b\alpha$, which is not in general equal to $\pm b\alpha$.

Remark. In fact, the algebra $Cl(p, q)$ **only** has non-trivial two-sided ideals when $p - q = 3 \bmod 4$, and they are always given by $Cl^\pm(p, q)$ for some orientation μ .

6.3.2 Examples of Clifford Algebras

Remark. Recall that $\dim Cl(V) = 2^n$.

Example 6.4. Suppose $n = 0$. Then $V = \{0\}$. We have $Cl(V) = \mathbb{R}$, where it is given the positive definite inner product $\langle a, b \rangle = ab$.

Example 6.5. Let $n = 1$. So $V = \text{span}\{e_1\}$. There are two possible signatures, $(1, 0)$ and $(0, 1)$. In the first case, $\langle e_1, e_1 \rangle = 1$ and in the second, $\langle e_1, e_1 \rangle = -1$.

(i) The first case is the signature $(1, 0)$.

In this case, $e_1 e_1 = |e_1|^2 = -1$, so $Cl(1, 0) = \{a1 + be_1 | e_1^2 = -1\}$. This is exactly the same as the definition of the complex numbers, so we see that

$$\mathbb{C} = Cl(1, 0)$$

(ii) The second case is the signature $(0, 1)$.

In this case, $e_1 e_1 = |e_1|^2 = 1$. So we have $Cl(0, 1) = \{a1 + be_1 | e_1^2 = 1\}$. This is called the set of **Lorentz numbers** (also known as the split-complex numbers or the hyperbolic numbers). We denote it,

$$\mathbb{L} = Cl(0, 1)$$

Remark. The Lorentz numbers **do not** form a field since $(1 + e_1)(1 - e_1) = 1 - e_1^2 = 0$.

Example 6.6. We will soon see that $Cl(2, 0) = \mathbb{H}$ and that $Cl(0, 2)$ and $Cl(1, 1)$ are both isomorphic to $M_{2 \times 2}(\mathbb{R})$ (although the isomorphisms are different).

Remark. Using the $k \bmod 4$ which appears in many places will allow us to classify all Clifford algebras just by knowing the first few cases. This is a phenomenon called Bott periodicity.

6.3.3 Pin Groups and Spin Groups

For the past few lectures we have worked out many properties of Clifford algebras. Now we will examine some important subsets of Clifford algebras which form groups.

Remark. Let $u \in \text{Cl}(V)$. Then if $|u|^2 \neq 0$ then $u \frac{\bar{u}}{|u|^2} = 1$, so u is invertible.

Definition 6.18 (Group of Invertible Elements of Clifford Algebra). The set of all invertible elements of $\text{Cl}(V)$ is denoted $\text{Cl}^*(V)$. Observe that it forms a group with respect to Clifford multiplication.

Remark. Try not to confuse this with the dual space, $(\text{Cl}(V))^*$.

Definition 6.19 (Pin Group). The Pin group is defined to be the subgroup of $\text{Cl}^*(V)$ generated by all unit vectors in V ,

$$\text{Pin}(V) = \{u_1 \dots u_k : |u_i|^2 = 1 \ \forall i = 1, \dots, k\} \quad (6.13)$$

Remark. By definition, every element $v \in \text{Pin}(V)$ is decomposable. So every element of $\text{Pin}(V)$ is either even or odd.

Definition 6.20 (Spin Group). We define the Spin group to be all the even elements of the Pin group,

$$\text{Spin}(V) = \text{Cl}^{\text{even}}(V) \cap \text{Pin}(V) \quad (6.14)$$

Lemma 6.8. Let $u \in V$ be a non-null vector. Then the action of a reflection R_u on a vector $v \in V$ is given by $R_u(v) = -uvu^{-1}$.

Proof. We have $uv = u \wedge v - \langle u, v \rangle$. So $uv = vu$ iff u and v are co-linear, since the wedge product part vanishes and we are left with the metric term which is symmetric.

So $uvu^{-1} = v$ if $v = \lambda u$ for some λ . Furthermore, $uvu^{-1} = -v$ if u and v are perpendicular. These are the two properties a reflection needs to satisfy, so we are done. \square

The above lemma allows us to relate the Pin and Spin groups to the group of isometries of a vector space. To make this precise we will introduce some ideas from representation theory. The representation theory of the Spin group is extremely important in physics.

Definition 6.21 (Representation). Let V be a vector space and let G be a group. Let $\text{GL}(V)$ be the set of all invertible linear maps from V to V . Then a **representation of G** is a group homomorphism $\rho : G \rightarrow \text{GL}(V)$.

The vector space V is called a representation space of G , and $\dim V$ is referred to as the dimension of the representation.

Remark. For an element of $\text{GL}(V)$ in the image of ρ , we use the notation $\rho_g = \rho(g)$.

Remark. We call such a map a representation because it provides us with a way to represent group elements as linear maps,

and group multiplication as composition of linear maps. When V is finite dimensional, we can even represent group elements using invertible matrices. As you might imagine, this is extremely powerful because linear algebra makes everything easier.

Definition 6.22 (Adjoint Representation). The **adjoint representation** of $\text{Cl}^*(V)$ is the map $\text{Ad} : \text{Cl}^*(V) \rightarrow \text{GL}(\text{Cl}(V))$ given by

$$\text{Ad}_a(b) = aba^{-1} \quad (6.15)$$

Definition 6.23 (Twisted Adjoint Representation). The **twisted adjoint representation** of $\text{Cl}^*(V)$ is the map $\widetilde{\text{Ad}} : \text{Cl}^*(V) \rightarrow \text{GL}(\text{Cl}(V))$ given by

$$\widetilde{\text{Ad}}_a(b) = \tilde{a}ba^{-1} \quad (6.16)$$

Theorem 6.12. There exist exact sequences,

$$1 \rightarrow \{1, -1\} \rightarrow \text{Pin}(V) \xrightarrow{\widetilde{\text{Ad}}} \text{O}(V) \rightarrow 1 \quad (6.17)$$

and

$$1 \rightarrow \{1, -1\} \rightarrow \text{Spin}(V) \xrightarrow{\widetilde{\text{Ad}}} \text{SO}(V) \rightarrow 1 \quad (6.18)$$

Proof. Let $u \in V$, $|u| \neq 0$. Then $R_u = \widetilde{\text{Ad}}_u \in \text{O}(V)$ as we saw earlier. So by Cartan-Dieudonné theorem, $\widetilde{\text{Ad}}$ surjects onto $\text{O}(V)$. Now suppose $\widetilde{\text{Ad}}_a$ equals the identity for some $a \in \text{Pin}(V)$. This means $\tilde{a}va^{-1} = v$ for all $v \in V$.

If a is odd we have $\tilde{a} = -a$, so $av = -va$, or in other words $a \in \text{twcent}(\text{Cl}(V))$ which is impossible since $\text{twcent}(\text{Cl}(V))$ never contains odd elements.

If a is even, we have $ava^{-1} = v$, so $a \in \text{cent}(\text{Cl}(V))$. Since a is even, we have $a \in \mathbb{R}$. So $a = \pm 1$.

Therefore, the only way to have $\widetilde{\text{Ad}}_a = \mathbb{1}$ is to have $a = \pm 1$, so we see that $\ker \widetilde{\text{Ad}}|_{\text{Pin}(V)} = \{1, -1\}$. Therefore we have shown that the first sequence is exact.

Now recall that any element of $\text{SO}(V)$ can be written as an even number of reflections. Since $\widetilde{\text{Ad}}(v)$ is a reflection for any $v \in V$, and since elements of $\text{Spin}(V)$ are the product of an even number of vectors, we see that $\widetilde{\text{Ad}}(\alpha)$ is the product of an even number of reflections as required. \square

Remark. The above exact sequences do not split. The proof of this is on assignment 5.

7 Composition Algebras

7.1 Lecture 18: Composition Algebras

7.1.1 Real and Imaginary Parts of Composition Algebras

We have seen so far that a few different composition algebras (algebras with $|ab| = |a||b|$) can be constructed using Clifford algebras. In this section we will consider these kinds of algebras in more generality, and later we will relate this back to Clifford

algebras. We will see in this section that any composition algebra can be broken down into a "real" and "imaginary" part (although more work will be required to show that this has anything to do with the complex numbers or quaternions).

Lemma 7.1. Let A be a unital algebra (not necessarily associative) over \mathbb{R} with metric $G = \langle \cdot, \cdot \rangle$. The following are equivalent conditions for A to be compositional.

- (i) $|ab|^2 = |a|^2|b|^2$
- (ii) $\langle ac, bc \rangle = \langle a, b \rangle |c|^2$
- (iii) $\langle ca, cb \rangle = |c|^2 \langle a, b \rangle$
- (iv) $\langle ac, bd \rangle + \langle bc, ad \rangle = 2\langle a, b \rangle \langle c, d \rangle$

Proof. First notice that (iv) implies (ii) and (iii), since we can set $c = d$ or $a = b$ in the equation.

Next, we see that (i) implies (ii) by the polarization identity, $|\langle a, b \rangle c|^2 = |a + b|^2 |c|^2$

The same argument shows that (i) implies (iii).

Again we can use the polarization identity to show that (ii) implies (iv) and that (iii) implies (iv).

Finally, we can set $a = b$ in (ii) or (iii) to show that (ii) implies (i) and that (iii) implies (i). □

Definition 7.1. We will define a right-multiplication map by $R_a(b) = ba$, and a left-multiplication map by $L_a(b) = ab$.

Remark. The condition (ii) in the previous lemma becomes $\langle R_c a, R_c b \rangle = |c|^2 \langle a, b \rangle$, and similar statements hold for (iii) and (iv).

Given a unital composition algebra A with a metric of signature (p, q) we will then define $\text{Re}(A) = \text{span}\{1\}$. Then $|a1|^2 = |a|^2|1|^2$, so $|1|^2 = 1$ as long as $p \geq 1$. This means that $\text{Re}(A)$ is a nondegenerate subspace of A , and in particular we have $A = \text{Re}(A) \oplus_{\perp} (\text{Re}(A))^{\perp}$. We then define $\text{Im}(A) = \text{Re}(A)^{\perp}$. We therefore get $a = \text{Re}(a) + \text{Im}(a)$ for any $a \in A$.

Observe that $\text{Re}(A)$ is a one-dimensional subalgebra, but $\text{Im}(A)$ is not necessarily a subalgebra and may just be a vector subspace.

We then define a map $a \mapsto \bar{a}$ by the formula $\bar{a} = \text{Re}(a) - \text{Im}(a)$. This is conjugation, which we will note defines an element of $O(V)$. Next time we will show that $R_a^{\dagger} = R_{\bar{a}}$ and $L_a^{\dagger} = L_{\bar{a}}$ for all $a \in A$, and that $\overline{ab} = \bar{b}\bar{a}$.

7.1.2 Conjugation in Composition Algebras

Before continuing our discussion of composition algebras we will prove the claim made at the end of the previous lecture. In this section A is a unital composition algebra with metric of signature (p, q) .

Lemma 7.2. Let $c \in A$. Then $R_c^{\dagger} = R_{\bar{c}}$ and $L_c^{\dagger} = L_{\bar{c}}$.

Proof. Suppose $c \in \text{Re}(A)$. Then $\bar{c} = c$, and furthermore, by the definition of $\text{Re}(A)$ we have that $\langle ca, b \rangle = \langle ta, b \rangle$ for some $t \in \mathbb{R}$. By linearity we can then move t to the other side to get $\langle ca, b \rangle = \langle a, tb \rangle = \langle a, cb \rangle$ as required.

Now suppose $c \in \text{Im}(A)$. Then $\bar{c} = -c$. By property (iv), we also have $\langle a, R_c b \rangle + \langle R_c a, b \rangle = 0$. So

$$\begin{aligned}\langle a, R_c b \rangle &= \langle -R_c a, b \rangle \\ &= \langle R_{\bar{c}} a, b \rangle\end{aligned}$$

By linearity the proof is complete. □

Lemma 7.3. We have,

1. $\bar{\bar{c}} = c$
2. $\langle a, b \rangle = \text{Re}(a\bar{b}) = \text{Re}(\bar{a}b) = \text{Re}(\bar{b}\bar{a}) = \text{Re}(\bar{\bar{b}}\bar{\bar{a}})$.
3. $\overline{ab} = \bar{b}\bar{a}$
4. $\bar{a}a = a\bar{a} = |a|^2$

Proof. We have

$$\begin{aligned}\langle a, b \rangle &= \langle a, R_b 1 \rangle \\ &= \langle R_{\bar{b}} a, 1 \rangle \\ &= \langle a\bar{b}, 1 \rangle \\ &= \text{Re}(a\bar{b})\end{aligned}$$

and,

$$\begin{aligned}\langle \overline{ab}, c \rangle &= \langle ab, \bar{c} \rangle \\ &= \langle b, \bar{a}\bar{c} \rangle \\ &= \langle bc, \bar{a} \rangle \\ &= \langle c, \bar{b}\bar{a} \rangle \\ &= \langle \bar{b}\bar{a}, c \rangle\end{aligned}$$

We also have $\overline{\bar{a}\bar{a}} = \bar{\bar{a}}\bar{\bar{a}} = a\bar{a}$, so $a\bar{a}$ is real. Therefore $\langle a\bar{a}, 1 \rangle = \langle a, a \rangle = |a|^2$. □

7.1.3 Commutators and Associators

Definition 7.2. Let A be an algebra. We define the **commutator** of two elements $a, b \in A$ as $[a, b] = ab - ba \in A$.

Remark. The map $[\cdot, \cdot] : A \times A \rightarrow A$ is bilinear and vanishes iff a, b commute. We also have $[a, b] = -[b, a]$.

Definition 7.3. Let A be an algebra. We define the **associator** to be the trilinear map $[\cdot, \cdot, \cdot] : A^3 \rightarrow A$ given by $[a, b, c] = (ab)c - a(bc)$.

Remark. The algebra A is associative if and only if the associator is always zero.

Lemma 7.4. Let A be a composition algebra. Then the associator is totally skew-symmetric.

Proof. By trilinearity it is enough to show that $[a, a, b] = [a, b, a] = [b, a, a] = 0$.

It is clear that if a or b are real the associator vanishes. So assume $a, b \in \text{Im}(A)$.

Then,

$$\begin{aligned} -\langle [a, b, b], c \rangle &= \langle [a, b, \bar{b}], c \rangle \\ &= \langle (ab)\bar{b} - a(b\bar{b}), c \rangle \\ &= \langle (ab)\bar{b}, c \rangle - \langle a(b\bar{b}), c \rangle \\ &= \langle ab, cb \rangle - |b|^2 \langle a, c \rangle \\ &= 0 \end{aligned}$$

Since the metric is nondegenerate, $[a, b, b] = 0$ for all a, b . We can show in exactly the same way that $[a, a, b] = 0$ and $[a, b, a] = 0$. \square

Definition 7.4 (Alternative Algebra). An algebra A over \mathbb{R} is said to be **alternative** if the associator is totally skew-symmetric.

Remark. We just showed that any composition algebra is alternative.

7.1.4 Cayley-Dickson Construction

Now we will show how any real unital composition algebra can be constructed out of \mathbb{R} by using the direct product and appropriately defining $\text{Re}(A)$ and $\text{Im}(A)$. First we need some lemmas.

Lemma 7.5. Let A be a unital composition algebra. Then,

$$(bc)\bar{d} + (bd)\bar{c} = 2\langle c, d \rangle b \quad (7.1)$$

and

$$\bar{a}(bc) + \bar{b}(ac) = 2\langle a, b \rangle c \quad (7.2)$$

Proof. We can rewrite the property that

$$\langle R_d a, R_c b \rangle + \langle R_c a, R_d b \rangle = 2\langle a, b \rangle \langle c, d \rangle$$

as,

$$\langle a, R_{\bar{d}} R_c b + R_{\bar{c}} R_d b \rangle = \langle 2\langle c, d \rangle b, a \rangle$$

and so that property is equivalent to,

$$(bc)\bar{d} + (bd)\bar{c} = 2\langle c, d \rangle b$$

Similarly we can write

$$\bar{a}(bc) + \bar{b}(ac) = 2\langle a, b \rangle c$$

□

Corollary 7.1. Let A be a unital composition algebra. If x and y are orthogonal, then

$$x(\bar{y}w) = -y(\bar{x}w)$$

$$(w\bar{y})x = -(w\bar{x})y$$

$$x\bar{y} = -y\bar{x}$$

Proof. Simply apply orthogonality to get rid of the metric terms in the above equations. □

Lemma 7.6. Let B be a unital composition algebra and let $A \subseteq B$ be a subalgebra which is also unital and compositional. Suppose $e \in A^\perp$ and $|e|^2 = \pm 1$. Then $A \oplus (Ae)$ is also a unital and compositional subalgebra of B .

Proof. Since A is unital we have $1 \in A$, so $x \in A$ implies that $x - 2\text{Re}(x)1 = \bar{x} \in A$, and vice versa. So $x \in A$ iff $\bar{x} \in A$.

Let $a, b \in A$. Then $\langle a, be \rangle = \langle \bar{b}a, e \rangle = 0$ since $e \in A^\perp$. So $Ae = \{ae | a \in A\}$ is orthogonal to A .

Since $e \in A^\perp$ we also have that $\langle e, 1 \rangle = 0$. So $\bar{e} = -e$. Furthermore, $e^2 = e(-\bar{e}) = -e\bar{e} = -|e|^2 = \pm 1$.

Now let us try some calculations. Suppose $a, b, c, d \in A$.

We have, $(a + be)(c + de) = ac + (be)c + a(de) + (be)(de)$. What we want to do is show that this is an element of $A \oplus (Ae)$.

We also have $(be)c = -(b\bar{e})c = (b\bar{e})e$ using the properties we proved earlier.

Similarly, using the fact that $ad \in A$, we have

$$\begin{aligned}
 a(de) &= -a(d\bar{e}) \\
 &= a(e\bar{d}) \\
 &= -a(\bar{e}\bar{d}) \\
 &= e(\bar{a}\bar{d}) \\
 &= e(\bar{d}\bar{a}) \\
 &= -(da)\bar{e} \\
 &= (da)e
 \end{aligned}$$

We also have,

$$\begin{aligned}
 (be)(de) &= -(\bar{b}\bar{e})(de) \\
 &= \bar{d}((be)e) \\
 &= -\bar{d}((b\bar{e})e) \\
 &= -(\bar{d}b)|e|^2 \\
 &= \mp(\bar{d}b)
 \end{aligned}$$

Overall we find that,

$$(a + be)(c + de) = (ac \mp \bar{d}b) + (da + b\bar{c})e$$

So $A \oplus (Ae)$ is indeed closed with respect to multiplication, so it is a subalgebra. □

The above proof gives us a hint. If we want to construct bigger unital composition algebras, we can start with a known unital composition algebra A and embed it in a larger algebra, and then construct $A \oplus (Ae)$ which is a new and larger composition algebra. This is called the Cayley-Dickson doubling process.

Definition 7.5 (Cayley-Dickson Process). Let A be a real unital composition algebra. Then we define the following algebras,

$$\begin{aligned}
 A(+) &= A \oplus A, & (a, b)(c, d) &= (ac - \bar{d}b, da + b\bar{c}) \\
 A(-) &= A \oplus A, & (a, b)(c, d) &= (ac + \bar{d}b, da + b\bar{a})
 \end{aligned}$$

Both $A(+)$ and $A(-)$ are real vector algebras, and the element $(1, 0) \in A \oplus A$ is the identity, so $A(+)$ and $A(-)$ are unital. It remains to show that these are composition algebras.

Let $e = (0, 1) \in A \oplus A$. Then $A(\pm) = A \oplus_{\pm} (Ae)$ since $e^2 = \mp 1$.

So A satisfies exactly the same conditions as above. So we just need to show that the metric makes sense, and soon we will

be able to show that these algebras are indeed compositional. Let (p, q) be the signature of the metric on A . We define,

$$|(a, b)|^2 = |a|^2 \pm |b|^2$$

Where the sign is chosen so that $e^2 = \mp 1$. Then the polarization identity gives us a metric with signature $(2p, 2q)$ for $A(+)$, and a metric with signature $(p + q, p + q)$ for $A(-)$.

Definition 7.6. Let $x = (a, b) \in A(\pm)$. Then we define $\bar{x} = (\bar{a}, -b)$. The map taking x to \bar{x} is an isometry, since conjugation is known to be an isometry on A and the map $b \mapsto -b$ is also trivially an isometry.

Lemma 7.7. Let A be a unital composition algebra. Let $x = a + pe$, $y = b + qe$, $z = c + re$ all be elements of $A(\pm)$. Then,

$$\begin{aligned}\overline{xy} &= \bar{y}\bar{x} \\ \overline{xx} &= x\bar{x} = |x|^2 \\ \frac{1}{2}(x\bar{y} + y\bar{x}) &= \text{Re}(x\bar{y}) = \langle x, y \rangle\end{aligned}$$

Proof. The following calculations are sufficient.

$$\begin{aligned}xy &= (a + pe)(b + qe) \\ &= (ab \mp \bar{q}p) + (qa + p\bar{b})e \\ \overline{xy} &= (\bar{b}\bar{a} \mp \bar{p}q) - (qa + p\bar{b})e \\ \bar{y}\bar{x} &= (\bar{b}\bar{a} \mp \bar{p}q) + (-p\bar{b} - qa)e \\ x\bar{x} &= (a\bar{a} \mp (-\bar{p})p) + (-pa + pa)e \\ &= |a|^2 \pm |p|^2 = |x|^2 \\ \overline{xx} &= (\bar{a}a \mp p(-\bar{p})) = |x|^2 \\ \text{Re}(x\bar{y}) &= \frac{1}{2}(x\bar{y} + \overline{(x\bar{y})}) \\ &= \frac{1}{2}(x\bar{y} + y\bar{x}) \\ &= \frac{1}{2}((a\bar{b} \pm \bar{q}p) + (-qa + p\bar{b})e + (-p\bar{b} + qa)e + (\bar{b}a \pm \bar{p}q)) \\ &= \text{Re}(a\bar{b}) + \text{Re}(\bar{p}q) \\ &= \langle a, b \rangle \pm \langle p, q \rangle \\ &= \langle x, y \rangle\end{aligned}$$

□

Lemma 7.8. Let $x = a + pe$, $y = b + qe$, $z = c + re$. We have the following identities,

$$(i) \quad \frac{1}{2}[x, y] = \frac{1}{2}[a, b] \pm \text{Im}(\bar{p}q) + (q\text{Im}(a) - p\text{Im}(b))e \quad (7.3)$$

(ii) If A is associative,

$$[x, y, z] = \pm[a, \bar{r}q] \pm [\bar{b}, \bar{r}p] \pm [c, \bar{q}p] + (p[\bar{b}, \bar{c}] + q[a, \bar{c}] + r[a, b])e \pm (p\bar{q}r - r\bar{q}p)e \quad (7.4)$$

(iii)

$$[x, \bar{x}, y] = \pm[a, \bar{q}, p] + [p, \bar{b}, a]e \quad (7.5)$$

(iv)

$$|x|^2|y|^2 = |xy|^2 + 2\langle a, [\bar{q}, p, \bar{b}] \rangle \quad (7.6)$$

In particular, if we want $A(\pm)$ to be a composition algebra we require that $2\langle a, [\bar{q}, p, \bar{b}] \rangle = 0$.

Theorem 7.1. We have,

- (i) $A(\pm)$ is commutative if and only if $A = \mathbb{R}$
- (ii) $A(\pm)$ is associative if and only if A is commutative and associative.
- (iii) $A(\pm)$ is alternative if and only if $A(\pm)$ is compositional, if and only if A is associative.

Proof. Let $x = a + pe$, $y = b + qe$, $z = c + re$

(i) Suppose $A = \mathbb{R}$. Then $\text{Im}(a) = 0$ and $[a, b] = 0$. Conversely, suppose $[a, b] = 0$. Then

$$0 = \mp \text{Im}(\bar{p}q) + (q\text{Im}(a) + p\text{Im}(b))e$$

Set $a = b = 0$, $p = 1$, and $q \in A$. Then we find that $\text{Im}(q) = 0$ for all $q \in A$. In other words, $A = \mathbb{R}$.

(ii) Suppose A is commutative and associative. Then both the associator and commutator are zero. So, from equation 7.4 we see that $[x, y, z] = p\bar{q}r - r\bar{q}p$, which is zero.

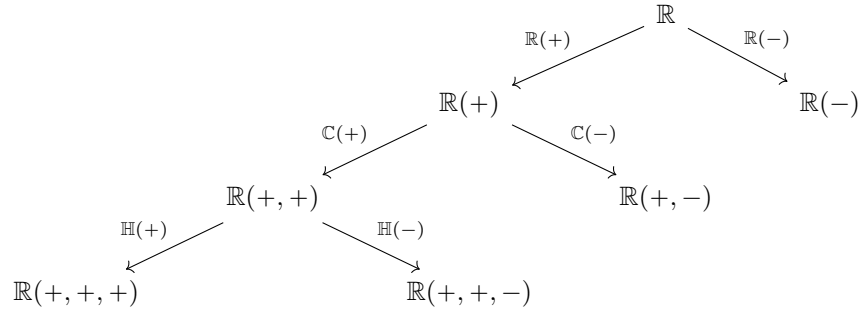
Conversely, suppose $A(\pm)$ is associative. Then A is associative because it is a subalgebra. Now let $b = 0$, $c = 0$, $p = 0$, $r = 1$. Then we get $[a, q] = 0$ for all $a, q \in A$. So A is commutative.

(iii) Suppose $A(\pm)$ is alternative. Then $[x, \bar{x}, y] = 0$, so $\pm[a, \bar{q}, p] + [p, \bar{b}, a]e = 0$. Let $b = 0$. Then we see that $[a, \bar{q}, p] = 0$ for all a, \bar{q}, p , so $A(\pm)$ is associative, and hence A is as well.

Now suppose A is associative. Then by 7.6, $|xy|^2 = |x|^2|y|^2$. So $A(\pm)$ is compositional.

Finally, suppose $A(\pm)$ is compositional. Then it is alternative as shown earlier. This completes the proof.

This process gives us the following sequence of constructions.

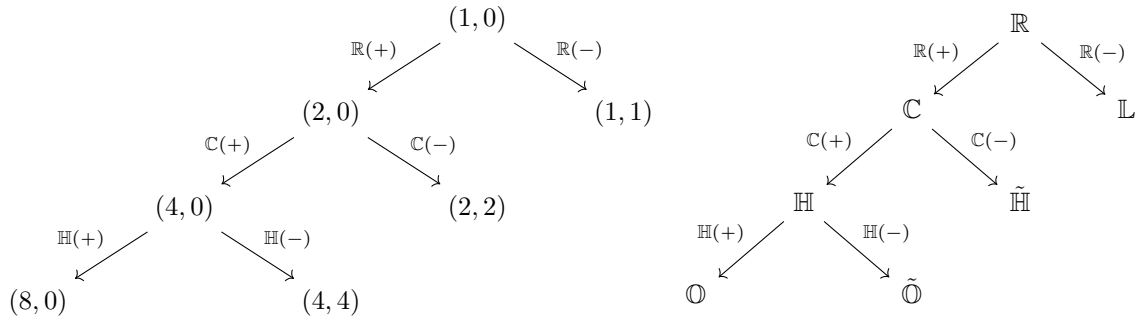


Soon we will show that only four of these are division algebras, and so there are exactly four real compositional division algebras up to isometry.

7.2 Lecture 19: Hurwitz Theorem

7.2.1 Classification of Composition Algebras

Recall the classification at the end of the previous section. It can also be phrased in terms of the signature of the metric on each algebra.



Corollary 7.2 (Corollary of Theorem 7.1). All 2-dimensional algebras generated by the Cayley Dickson process are commutative, associative, and compositional.

All the 4-dimensional algebras are associative and compositional.

All the 8-dimensional algebras generated are alternative and compositional.

Lemma 7.9. We have $\mathbb{R}(+) = \mathbb{C}, \mathbb{R}(-) = \mathbb{L}, \mathbb{R}(+, +) = \mathbb{H}, \tilde{\mathbb{H}} = \mathbb{R}(+, -) \cong M_{2 \times 2}(\mathbb{R})$.

Proof. We can see that $\mathbb{R}(+) = \mathbb{C}$ and $\mathbb{R}(-) = \mathbb{L}$, since $\mathbb{R}(+) = \mathbb{R} \oplus i\mathbb{R}$ with $i^2 = -1$ and $\mathbb{R}(-) = \mathbb{R} \oplus e\mathbb{R}$ with $e^2 = 1$.

We have $\mathbb{R}(+, +) = \mathbb{C} \oplus_{e_2} \mathbb{C}$, with $e_2^2 = -1$. So $i, j = e_2, k = ie_2$ all square to minus 1, and we get $i^2 = j^2 = k^2 = ijk = -1$.

So $\mathbb{R}(+, +) = \mathbb{H}$.

We also have $\mathbb{R}(+, -) \cong M_{2 \times 2}(\mathbb{R})$. The isomorphism is as follows. Consider the norm on $M_{2 \times 2}(\mathbb{R})$ given by $\|A\|^2 = \det A$. An

orthonormal basis is given by,

$$B = \left\{ \mathbb{1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, e_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, e_3 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, e_4 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \right\}$$

We can see that $e_2^2 = -\mathbb{1}$, so $\text{span}\{\mathbb{1}, e_1\}$ is isometric to \mathbb{C} , and then we see that $e_2 e_3 = -e_4$, so overall $M_{2 \times 2}$ is isometric to $\mathbb{C} \oplus \mathbb{C} e_3 = \mathbb{C}(-)$ as required. \square

7.2.2 Hurwitz Theorem

Many proofs of Hurwitz' theorem are done using topology. However, Hurwitz' theorem was proven in 1898, and the original proof only involved linear algebra. We will give a similar proof here.

Theorem 7.2 (Hurwitz' Theorem). Up to isomorphism there are only seven real composition algebras, $\mathbb{R}, \mathbb{C}, \mathbb{L}, \mathbb{H}, M_{2 \times 2}(\mathbb{R}), \tilde{\mathbb{O}}, \mathbb{O}$. Of these, the only real composition algebras which are division algebras are $\mathbb{R}, \mathbb{C}, \mathbb{H}, \mathbb{O}$.

Proof. Let B be a composition algebra. Let $A_1 = \text{Re}(B) = \mathbb{R}$. If $A_1 = B$ we are done. Otherwise set $e_1 \in A_1^\perp$ be a unit vector. Let $A_2 = A_1 \oplus A_1 e_1 = A_1(\pm)$ according to $e_1^2 = \mp 1$. We have $A_2 = \mathbb{C}$ or \mathbb{L} .

If $B = A_2$ we're done. Otherwise set $e_2 \in A_2^\perp$ with $|e_2|^2 = 1$. Let $A_3 = A_2(\pm)$. There are two cases,

1. If $A_2 = \mathbb{C}$ then $A_2(\pm)$ is either \mathbb{H} or $M_{2 \times 2}(\mathbb{R})$.
2. If $A_2 = \mathbb{L}$ then $e_1^2 = 1$, $e_2^2 = \pm 1$, and $e_1 \perp e_2$. Recall that $x\bar{y} + y\bar{x} = 0$ if $x \perp y$. So if $x \perp y$ and x, y are imaginary, we have $xy = -yx$. So $e_1 e_2 = -e_2 e_1$. Let $e_3 = e_1 e_2$. Then $e_3^2 = \mp 1$. So one of e_2, e_3 square to 1 while the other squares to -1 . Without loss of generality set $e_2^2 = e_3^2 = -1$. Notice that $e_1 e_2 = e_3$ implies that $e_3 e_1 = \pm e_2$, $e_2 e_3 = \pm e_1$. So $A_3 = (A_1 \oplus A_1 e_3) \oplus ((A_1 \oplus A_1 e_3)e_2)$. Therefore $A_3 = \mathbb{C}(-) \cong M_{2 \times 2}(\mathbb{R})$.

Now if $B = A_3$ we are done. Otherwise there exists $e_4 \in A_3^\perp$ with $e_4^2 = \pm 1$, and we set $A_4 = A_3 \oplus A_3 e_4$. If $A_3 = \mathbb{H}$ then clearly $A_4 = \mathbb{O}$ when $e_4^2 = -1$. If $e_4^2 = 1$ then we get $\tilde{\mathbb{O}}$ which is defined to be $\mathbb{H}(-)$.

If $A_3 = \mathbb{C}(-)$, then instead we get $A_3 = (A_1 \oplus A_1 e_1) \oplus (A_1 e_2 \oplus A_1 e_3)$, which means

$$A_3 e_4 = A_1 e_4 + A_1 e_1 e_4 + A_1 e_2 e_4 + A_1 (e_1 e_2) e_4$$

This means we can write A_4 as,

$$A_4 = A_3 \oplus A_3 e_4 = (A_1 \oplus A_1 e_1 \oplus A_1 e_2 \oplus A_1 e_1 e_2 \oplus A_1 e_4 \oplus A_1 e_1 e_4 \oplus A_1 e_2 e_4 \oplus A_1 (e_1 e_2) e_4)$$

We want to show that this reduces to $\mathbb{H} \oplus \mathbb{H} e$ where $e^2 = 1$. Note that in this case, the resulting algebra is alternative. Furthermore, we can show that $(e_1 e_4)^2 = -e_1 e_4 e_4 e_1 = e_4^2 = \pm 1$. Similarly $(e_2 e_4)^2 = (e_3 e_4)^2 = -1$. Finally, if $x \perp y$ we get

$(w\bar{y})x + (w\bar{x})y = 0$, so since $e_1 \perp e_2e_4$ we get,

$$(e_2e_4)e_1 = \pm e_3e_4$$

$$(e_2e_4)e_2 = \pm e_4$$

$$((e_2e_4)e_1)e_2 = \pm e_1e_4$$

This allows us to finally reduce A_4 to $\mathbb{H}(-)$, so $A_4 \cong \tilde{\mathbb{O}}$.

Finally, since any higher dimensional algebra produced by this construction is not compositional, we are done. \square

Remark. One can also show that if A is a real finite dimensional division algebra then $\dim A$ is 1, 2, 4 or 8 using topology. These include the composition algebras, $\mathbb{R}, \mathbb{C}, \mathbb{H}, \mathbb{O}$, but also more which are not composition algebras.

7.3 Lecture 20: Cross Products and Isometries of Composition Algebras

7.3.1 Cross Products

Definition 7.7 (Cross Product on Composition Algebra). Let A be a composition algebra and define a map $\times : \text{Im}(A) \times \text{Im}(A) \rightarrow \text{Im}(A)$ by the formula,

$$a \times b = \text{Im}(ab) = \frac{ab - \bar{a}b}{2} = \frac{1}{2}[a, b] \quad (7.7)$$

Notice that $a \times b = -b \times a$ and that this product is \mathbb{R} -linear in both inputs.

Remark.

$$\begin{aligned} ab &= \text{Re}(ab) + \text{Im}(ab) \\ &= -\text{Re}(a\bar{b}) + a \times b \\ &= a \times b - \langle a, b \rangle 1 \end{aligned}$$

Lemma 7.10.

$$\langle a \times b, a \rangle = 0 \quad (7.8)$$

$$\langle a \times b, b \rangle = 0 \quad (7.9)$$

Proof. The second follows from the first by antisymmetry. Therefore we just prove one,

$$\begin{aligned}
 \langle a \times b, a \rangle &= \langle ab + \langle a, b \rangle, a \rangle \\
 &= \langle a, b \rangle \langle 1, a \rangle + \langle ab, a \rangle \\
 &= \langle b, |a|^2 \rangle \\
 &= 0
 \end{aligned}$$

□

Lemma 7.11.

$$|a \times b|^2 = |a|^2 |b|^2 - \langle a, b \rangle^2 \quad (7.10)$$

Proof.

$$\begin{aligned}
 |a \times b|^2 &= \frac{1}{4} |ab - ba|^2 \\
 &= \frac{1}{4} (|ab|^2 - 2\langle ab, ba \rangle + |ba|^2) \\
 &= \frac{1}{2} (|a|^2 |b|^2 - \langle ab, ba \rangle) \\
 &= \frac{1}{2} (|a|^2 |b|^2 - \langle a, b \rangle^2 + |a \times b|^2)
 \end{aligned}$$

Rearranging this gives the desired result.

□

Definition 7.8. Let $a \in \text{Im}(A)$. Then define $L_a^\times : \text{Im}(A) \rightarrow \text{Im}(A)$ by $L_a^\times(b) = a \times b$.

Lemma 7.12. $(L_a^\times)^2(b) = -|a|^2 b + \langle a, b \rangle a$.

Proof.

$$\begin{aligned}
 (L_a^\times)^2(b) &= a \times (a \times b) \\
 &= a(a \times b) + \langle a, a \times b \rangle \\
 &= a(ab + \langle a, b \rangle) \\
 &= \langle a, b \rangle a + a(ab) \\
 &= \langle a, b \rangle a - \bar{a}(ab) \\
 &= \langle a, b \rangle a - |a|^2 b
 \end{aligned}$$

□

Remark. Fix $|a| = 1$ and let $V = \text{span}\{a\}^\perp$. Then $(L_a^\times)^2|_V = -\mathbb{1}$, so $(L_a^\times)^2$ is a complex structure on V .

Definition 7.9 (Associative 3-Form). The **associative 3-form** of an algebra with a cross product and a metric is the 3-form $\varphi : (\text{Im}(A))^3 \rightarrow \mathbb{R}$, $\varphi \in \Lambda^3(\text{Im}(A)^*)$ defined by the formula,

$$\varphi(a, b, c) = \langle a \times b, c \rangle = \langle ab, c \rangle \quad (7.11)$$

Lemma 7.13 (Triple Product Identity). Let A be a composition algebra with a cross product. Then,

$$a \times (b \times c) = -\langle a, b \rangle c + \langle a, c \rangle b - \frac{1}{2}[a, b, c] \quad (7.12)$$

Proof. We first show that,

$$\begin{aligned} a(bc) &= -a(\bar{bc}) \\ &= b(\bar{ac}) - 2\langle a, b \rangle c \\ &= b(\bar{ca} - 2\langle a, c \rangle) - 2\langle a, b \rangle c \\ &= -b(\bar{ca}) + 2\langle a, c \rangle b - 2\langle a, b \rangle c \\ &= c(\bar{ba}) - 2\langle b, c \rangle a + 2\langle a, c \rangle b - 2\langle a, b \rangle c \\ &= -c(\bar{ab}) - 2\langle b, c \rangle a + 2\langle a, c \rangle b - 2\langle a, b \rangle c \\ &= (ab)\bar{c} - 2\varphi(a, b, c) - 2\langle b, c \rangle a + 2\langle a, c \rangle b - 2\langle a, b \rangle c \end{aligned}$$

We then use this to determine the triple product identity in the general cross product case,

$$\begin{aligned} a \times (b \times c) &= \frac{1}{2}(a \times (b \times c)) + \frac{1}{2}(a \times (b \times c)) \\ &= \frac{1}{2}\varphi(a, b, c) + \frac{1}{2}\langle b, c \rangle a + \frac{1}{2}a(bc) \\ &\quad + \frac{1}{2}\varphi(a, b, c) + \frac{1}{2}\langle b, c \rangle a + \frac{1}{2}a(bc) \\ &= -\langle a, b \rangle c + \langle a, c \rangle b - \frac{1}{2}[a, b, c] \end{aligned}$$

□

We have defined a cross product for a composition algebra, which extends the usual definition from \mathbb{R}^3 . Now we will show that these are the **only** algebras with a cross product.

Definition 7.10 (Cross Product Algebra). Let V be an n -dimensional vector space with a metric. We say V is a cross-product algebra if there is a product $\times : V^2 \rightarrow V$ which satisfies,

1. $a \times b = -b \times a$
2. $a \times b \perp a$ and $a \times b \perp b$
3. $|a \times b|^2 = |a \wedge b|^2$

Theorem 7.3. If V is a cross-product algebra, then V is isomorphic as an algebra to $\text{Im}(A)$ for some composition algebra A , where $a \times b = \text{Im}(ab)$.

Corollary 7.3. Any cross-product algebra is either 0, 1, 3, or 7 dimensional.

Proof. Define $A = \mathbb{R} \oplus V$ and define a product on A by,

$$(t, a)(s, b) = (ts - \langle a, b \rangle, tb + sa + a \times b)$$

Where $t, s \in \mathbb{R}$ and $a, b \in V$. Note that this product is bilinear and that $(1, 0)$ is the identity. Define a norm on A by $|(t, s)|^2 = t^2 + |a|^2$. We just need to show that A is a composition algebra. Let $a = (a_0, \mathbf{a}), b = (b_0, \mathbf{b})$. We have,

$$\begin{aligned} |ab|^2 &= (a_0b_0 - \langle \mathbf{a}, \mathbf{b} \rangle)^2 + |a_0\mathbf{b} + b_0\mathbf{a} + \mathbf{a} \times \mathbf{b}|^2 \\ &= a_0^2b_0^2 - 2\langle \mathbf{a}, \mathbf{b} \rangle a_0b_0 + \langle \mathbf{a}, \mathbf{b} \rangle^2 + a_0^2|\mathbf{b}|^2 + b_0^2|\mathbf{a}|^2 + |\mathbf{a} \times \mathbf{b}|^2 + 2a_0b_0\langle \mathbf{a}, \mathbf{b} \rangle \\ &= a_0^2b_0^2 + a_0^2|\mathbf{b}|^2 + b_0^2|\mathbf{a}|^2 \\ &= |a|^2|b|^2 \end{aligned}$$

As required. □

This gives us a summary of all cross product algebras, related to the composition algebras.

$\dim V$	0	1	3	7
$\dim A$	1	2	4	8
$\text{Im}(A)$	$\{0\}$	\mathbb{R}	\mathbb{R}^3	\mathbb{R}^7
A	\mathbb{R}	\mathbb{C}	\mathbb{H} or $\tilde{\mathbb{H}}$	\mathbb{O} or $\tilde{\mathbb{O}}$
\times	Trivial	Trivial	Nontrivial	Nontrivial

Table 1: Summary of Cross Product Algebras and their relation to Composition Algebras.

Example 7.1. The cross product on $\{0\}$ is clearly meaningless, and just amounts to multiplying zero by zero.

The cross product on \mathbb{R} is just multiplication, $ab = a \times b$ if $a, b \in \mathbb{R}$.

In the case \mathbb{R}^3 we get the usual cross product that is well-known from linear algebra.

In the case \mathbb{R}^7 we discover a modified cross product, related to the multiplication of octonions.

7.3.2 Automorphisms of Composition Algebras

The cross product defined on the imaginary part of a composition algebra allows us to understand the automorphisms. Recall that the automorphisms of an algebra A are the linear isomorphisms from A to itself which satisfy $T(ab) = T(a)T(b)$.

Lemma 7.14. $\text{Aut}(A) \subseteq \text{O}(\text{Im}(A))$

Proof. Let $T \in \text{Aut}(A)$. Then $T(b) = T(1b) = T(1)T(b)$. This means $T(1) = 1$. So T fixes $\text{Re}(A)$. If $a \in A$, then we have $\overline{\text{Im}(a)} = -\text{Im}(a)$, so

$$\begin{aligned} a^2 &= \text{Re}(a)^2 + \text{Im}(a)^2 + 2\text{Re}(a)\text{Im}(a) \\ &= (\text{Re}(a)^2 - |\text{Im}(a)|^2) + 2\text{Re}(a)\text{Im}(a) \end{aligned}$$

Let $a \in \text{Im}(A)$. Then we see that $a^2 \in \text{Re}(A)$, so $g(a^2) = g(a)^2$, so $g(a)$ must be purely imaginary or purely real. Since g fixes \mathbb{R} , we can't possibly have $g(a) \in \text{Re}(A)$. So $g(\text{Im}(A)) \subseteq \text{Im}(A)$, meaning $g \in \text{GL}(\text{Im}(A))$.

Also, $\overline{g(a)} = g(\bar{a})$, so $|g(a)|^2 = |a|^2$, so $g \in \text{O}(\text{Im}(A))$.

Therefore we have shown that $\text{Aut}(A) \subseteq \text{O}(\text{Im}(A))$. □

In fact, we can go further.

Lemma 7.15. Let $A = \mathbb{H}$ or $\tilde{\mathbb{H}}$, and let $T \in \text{Aut}(A)$. Let φ be the associative 3-form of $\text{Im}(A)$. Then $T^*\varphi = \varphi$.

Proof. Since $\dim \text{Im}(A) = 3$, the associative 3-form φ is a top-degree form. For all $a, b, c \in \text{Im}(A)$ we have

$$\begin{aligned} T^*\varphi(a, b, c) &= \langle g(a)g(b), g(c) \rangle \\ &= \langle g(ab), g(c) \rangle \\ &= \langle ab, c \rangle \\ &= \varphi(a, b, c) \end{aligned}$$

Since g is an isometry. □

Lemma 7.16. Let $A = \mathbb{H}$ or $\tilde{\mathbb{H}}$, then $\text{Aut}(A) \subseteq \text{SO}(\text{Im}(A))$

Proof. Let e_1, e_2, e_3 be an orthonormal basis of $\text{Im}(A)$. Then let $\mu = e_1 \wedge e_2 \wedge e_3$. Since φ is a top form we have $\varphi = \lambda\mu$ for some $\lambda \in \mathbb{R}$. Therefore, if $T \in \text{Aut}(A)$ we have

$$\begin{aligned} T^*\mu &= T^*(e_1 \wedge e_2 \wedge e_3) \\ &= \lambda^{-1}T^*\varphi \\ &= \lambda^{-1}\varphi \\ &= \mu \end{aligned}$$

We also have $T^*\mu = \det(T)\mu$. So $\det T = 1$, which means $T \in \text{SO}(\text{Im}(A))$ as required. \square

A summary of the automorphism groups of every composition algebra is shown in the following table. We just looked at the 4-dimensional case, and the smaller cases are trivial. In the next lecture we will study the 8 dimensional case and define the exceptional groups G_2 and \tilde{G}_2 .

A	$\text{Im}(A)$	$\text{O}(\text{Im}(A))$	$\text{Aut}(A)$
\mathbb{R}	$\{0\}$	$\{1\}$	$\{1\}$
\mathbb{C}	\mathbb{R}	$\mathbb{Z}/2\mathbb{Z}$	$\mathbb{Z}/2\mathbb{Z}$
\mathbb{H}	\mathbb{R}^3	$\text{O}(3)$	$\text{SO}(3)$
\mathbb{O}	\mathbb{R}^7	$\text{O}(7)$	G_2
$\tilde{\mathbb{L}}$	\mathbb{R}	$\mathbb{Z}/2\mathbb{Z}$	$\mathbb{Z}/2\mathbb{Z}$
$\tilde{\mathbb{H}}$	\mathbb{R}^3	$\text{O}(1, 2)$	$\text{SO}(1, 2)$
$\tilde{\mathbb{O}}$	\mathbb{R}^7	$\text{O}(3, 4)$	\tilde{G}_2

Table 2: Table of Automorphism Groups of the Real Composition Algebras

7.4 Lecture 21: The Exceptional Groups G_2 and \tilde{G}_2 .

7.4.1 Automorphisms of \mathbb{H} and $\tilde{\mathbb{H}}$, Continued

We will first prove the following theorem, and then move on to the 8 dimensional case.

Theorem 7.4. Let A be a composition algebra. If $\dim A = 4$ then $\text{Aut}(A) = \text{SO}(\text{Im}(A))$.

Proof. We have shown that $\text{Aut}(A) \subseteq \text{SO}(\text{Im}(A))$ in the previous lecture.

Let $g \in \text{SO}(\text{Im}(A))$. We know $\langle a, b \rangle = \langle g(a), g(b) \rangle$. Then,

$$\begin{aligned} g^*\varphi(a, b, c) &= \langle g(a \times b), g(c) \rangle \\ &= \langle g(ab + \langle a, b \rangle), c \rangle \\ &= \langle g(a)g(b) + \langle a, b \rangle, c \rangle \\ &= \langle g(a)g(b), c \rangle + \langle g(a), g(b) \rangle \langle 1, c \rangle \\ &= \langle g(a)g(b) + \langle g(a), g(b) \rangle, c \rangle \end{aligned}$$

So $g(a \times b) = g(a) \times g(b)$. So,

$$\begin{aligned}
 g(a \times b) &= g(ab + \langle a, b \rangle) \\
 &= g(ab) + \langle a, b \rangle g(1) \\
 &= g(ab) + \langle a, b \rangle \\
 g(a) \times g(b) &= g(a)g(b) + \langle g(a), g(b) \rangle \\
 &= g(ab) + \langle a, b \rangle
 \end{aligned}$$

Therefore $g(ab) = g(a)g(b)$, so $g \in \text{Aut}(A)$. □

7.4.2 Automorphisms of \mathbb{O} and $\tilde{\mathbb{O}}$

Definition 7.11 (Exceptional Groups G_2 and \tilde{G}_2). We define the group $G_2 = \text{Aut}(\mathbb{O})$ and the group $\tilde{G}_2 = \text{Aut}(\tilde{\mathbb{O}})$. These are the smallest real **exceptional lie groups**.

Physics Remark. The groups $O(n)$ and $SO(n)$ are the symmetry groups of an object with rotational symmetry.

The question is then, what physical system is G_2 the symmetry group of? Élie Cartan showed in 1893 that one option is to interpret the group G_2 as the symmetry group of a ball which rolls without slipping or twisting [4].

Theorem 7.5. $G_2 \subseteq SO(7)$ and $\tilde{G}_2 \subseteq SO(3, 4)$.

Proof. We already showed that $\text{Aut}(A) \subseteq O(\text{Im}(A))$. Let us now show that G_2 and \tilde{G}_2 preserve orientations on \mathbb{R}^7 .

First, recall that in the four dimensional case we used the associative 3-form in order to show that the automorphisms were isometries of determinant 1. In this case we will use the associative 3-form as well as its' Hodge dual to come up with a similar argument. We define a four form $\psi \in \Lambda^4(\text{Im}(A)^*)$ by the formula,

$$\psi(a, b, c, d) = \frac{1}{2} \langle a, [b, c, d] \rangle$$

Observe that since A is compositional, it is also alternative, so ψ is antisymmetric in b, c, d . To show that ψ is totally antisymmetric it suffices to show that $\psi(a, a, b, c) = 0$. We have,

$$\begin{aligned}
 2\psi(a, a, b, c) &= \langle a, (ac)d - a(cd) \rangle \\
 &= -\langle a\bar{d}, ac \rangle - |a|^2 \langle 1, cd \rangle \\
 &= |a|^2 \langle \bar{d}, c \rangle - |a|^2 \langle 1, cd \rangle \\
 &= 0
 \end{aligned}$$

So ψ is indeed a four-form. Clearly if $\dim A$ is not 8, then ψ vanishes because A would be associative. So this construction only works in this case.

We also see that if $g \in \text{Aut}(A)$ then

$$g^*\psi(a, b, c, d) = \langle g(a), [g(b), g(c), g(d)] \rangle = \langle g(a), g([b, c, d]) \rangle$$

Since $g \in \text{O}(\text{Im}(A))$ it follows that

$$2g^*\psi(a, b, c, d) = \langle g(a), g([b, c, d]) \rangle = \langle a, [b, c, d] \rangle = 2\psi$$

So $g^*\psi = \psi$.

Now choose a basis e_1, \dots, e_7 for $\text{Im}(A)$. Consider the case $A = \mathbb{O}$, where $|e_i|^2 = 1$. Then we have,

$$\varphi = e_1 \wedge e_2 \wedge e_3 - (e_1 \wedge e_5 \wedge e_7 + e_5 \wedge e_2 \wedge e_7 + e_5 \wedge e_6 \wedge e_3) - e_4 \wedge (e_1 \wedge e_5 + e_2 \wedge e_6 + e_3 \wedge e_7)$$

Similarly, we can expand ψ in a basis and show that $\star\varphi = \psi$.

Therefore $\varphi \wedge \psi = \langle \varphi, \varphi \rangle e_1 \wedge \dots \wedge e_7 = 7\mu_G$, where μ_G is the volume form.

So if $g \in \text{Aut}(A)$,

$$7g^*\mu = g^*(\varphi \wedge \psi) = g^*\varphi \wedge g^*\psi = \varphi \wedge \psi = 7\mu$$

So $g \in \text{SO}(7)$.

Similarly, we can show that if $A = \tilde{\mathbb{O}}$ then $g \in \text{SO}(3, 4)$. This completes the proof. □

Remark. The four-form ψ is related to the associator. Similarly, $\varphi(a, b, c) = \langle a \times b, c \rangle = \frac{1}{2}\langle [a, b], c \rangle$, so the three-form φ is related to the commutator.

Definition 7.12 (Co-associative 4-Form). Let A be a composition algebra of dimension 8. The four-form $\psi \in \Lambda^4(\text{Im}(A)^*)$ defined by $\psi(a, b, c, d) = \frac{1}{2}\langle a, [b, c, d] \rangle$ is called the **co-associative 4-form** (also known as the **Cayley form**). It satisfies $\star\varphi = \psi$.

Theorem 7.6 (Bryant, 1982 [3]).

$$G_2 = \{g \in \text{GL}(\text{Im}(\mathbb{O})) : g^*\varphi = \varphi\} \tag{7.13}$$

$$\tilde{G}_2 = \{g \in \text{GL}(\text{Im}(\tilde{\mathbb{O}})) : g^*\varphi = \varphi\} \tag{7.14}$$

Proof. We will give the proof for G_2 . First, by plugging into a basis one can show through tedious computation that,

$$(a \lrcorner \varphi) \wedge (b \lrcorner \varphi) \wedge \varphi = -6\langle a, b \rangle \mu_G$$

Suppose $g \in \text{GL}(\text{Im}(\mathbb{O}))$ and $g^*\varphi = \varphi$. Then $(g^{-1}(a) \lrcorner \varphi) \wedge (g^{-1}(b) \lrcorner \varphi) \wedge \varphi = -6 \det G \langle a, b \rangle \mu_G = \langle g^{-1}(a), g^{-1}(b) \rangle \mu_G$.

So $\det g \langle a, b \rangle = \langle g^{-1}(a), g^{-1}(b) \rangle$ for all a, b . Then by inserting any normalized vector, we get $(\det g)^7 = 1$. So $\det g = 1$, meaning $g \in \text{SO}(7)$. Furthermore, we saw earlier how this implies $g(a \times b) = g(a) \times g(b)$, which implies $g(ab) = g(a)g(b)$. So $g \in G_2$. This completes the proof for the G_2 case.

The \tilde{G}_2 case differs only by a few signs when plugging in basis vectors, which end up canceling anyway. \square

Remark. In an 8-dimensional composition algebra, a map g is an automorphism iff it preserves the associative 3-form. So everything is completely determined by φ .

Remark. This is somewhat similar to the symmetries of complex vector spaces, in which all the automorphisms, $\text{U}(n)$, are determined by preserving two out of three special structures. However, with G_2 we end up with a much stronger requirement. Let us make this more precise.

The vector space \mathbb{C}^n has a complex structure J defined by $J(v) = iv$, a metric induced by the standard inner product on \mathbb{R}^{2n} , and a two-form ω defined by $\omega(u, v) = \langle J(u), v \rangle$.

We call the tuple $(J, \omega, \langle \cdot, \cdot \rangle)$ a **compatible triple**. The equivalent structure on a manifold is called a **Kähler structure**, where the metric, complex structure, and 2-form are defined on the tangent bundle and have to satisfy some additional conditions.

The group which preserves J is the set of \mathbb{C} -linear invertible transformations, $\text{GL}(n, \mathbb{C})$. The group preserving the inner product is the orthogonal group $\text{O}(2n, \mathbb{R})$, and the group preserving ω is the symplectic group $\text{Sp}(2n, \mathbb{R})$.

As we can see, if a transformation g preserves two out of three of these structures, it must preserve the third. It turns out that the intersection of two of these groups is always $\text{U}(n)$.

$$\text{Sp}(2n, \mathbb{R}) \cap \text{GL}(n, \mathbb{C}) = \text{U}(n)$$

$$\text{Sp}(2n, \mathbb{R}) \cap \text{O}(2n, \mathbb{R}) = \text{U}(n)$$

$$\text{O}(2n, \mathbb{R}) \cap \text{GL}(n, \mathbb{C}) = \text{U}(n)$$

In this octonionic case, we only need to preserve one thing, which is φ . This means that the conditions on automorphisms are stricter.

Remark. A Kähler manifold is a manifold with a compatible triple $(J, \omega, \langle \cdot, \cdot \rangle)$. Equivalently, it is a Riemannian manifold whose holonomy group is a subgroup of $\text{U}(n)$. So the holonomy must preserve two out of three of the objects in the Kähler structure.

Similarly, a G_2 -manifold is a Riemannian manifold whose holonomy group is a subgroup of G_2 . The condition of being a G_2 -manifold is therefore a bit stricter, since the holonomy must preserve φ .

Lemma 7.17. If $A \in G_2$, then we can write the matrix of A as,

$$[A] = [v_1 | v_2 | v_1 \times v_2 | v_4 | v_1 \times v_4 | v_2 \times v_4 | (v_1 \times v_2) \times v_4]$$

Where v_1, v_2, v_4 are orthonormal. This means the columns of A are an oriented orthonormal basis of \mathbb{R}^7 . Furthermore, we see that $\dim G_2 = 14$.

Proof. Proof omitted for time. □

8 Spinors

8.1 Lecture 22: Classification of Clifford Algebras

8.1.1 Classification of Clifford Algebras

In this section we will use what we have just learned about composition algebras in order to classify all Clifford algebras.

Remark. Recall some facts from the earlier sections. We have

$$Cl(p, q) = \bigotimes \mathbb{R}^{p,q} / I$$

Where $I = \langle v \otimes v + |v|^2 1 \rangle$. We have $Cl(p, q) \cong \Lambda^\bullet \mathbb{R}^{p,q}$ as vector spaces but not as algebras, where if $v \in \mathbb{R}^{p,q}$ and $\alpha \in \Lambda^\bullet \mathbb{R}^{p,q}$ we have $v\alpha = v \lrcorner \alpha - v \wedge \alpha$.

Example 8.1. We see the beginning of a pattern by looking at $Cl(1, 0), Cl(0, 1), Cl(1, 1)$. As shown earlier in the notes, we have,

$$\begin{aligned} Cl(1, 0) &\cong \mathbb{C}, \\ Cl(0, 1) &\cong \mathbb{L}, \\ Cl(2, 0) &\cong \mathbb{H}, \\ Cl(0, 2) &\cong Cl(1, 1) \cong \tilde{\mathbb{H}} \cong M_{2 \times 2}(\mathbb{R}) \end{aligned}$$

The pattern ends here since every Clifford algebra is associative but \mathbb{O} is non-associative. However, we can then show that every other Clifford algebra decomposes into a direct sum of these simple algebras.

Example 8.2. The algebra $Cl(1, 1)$ has a basis consisting of $1, e_1, e_2, e_1 e_2$. This gets mapped to $M_{2 \times 2}(\mathbb{R})$ by,

$$1 \mapsto \mathbb{I}, \quad e_1 \mapsto \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad e_2 \mapsto \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad e_1 e_2 \mapsto \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Example 8.3. We have $Cl^{\text{even}}(1, 1) \cong \mathbb{L}$.

We also have,

$$Cl^{\text{odd}}(1, 1) \cong \left\{ \begin{bmatrix} a & -b \\ b & -a \end{bmatrix} \mid a, b \in \mathbb{R} \right\}$$

Where matrices act on each other by $A \cdot B = ABA^{-1}$ rather than the usual formula.

Example 8.4. The algebra $Cl(0, 2)$ has basis $\{1, e_1, e_2, e_1e_2\}$, and we again use the map

$$1 \mapsto \mathbb{1}, \quad e_1 \mapsto \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad e_2 \mapsto \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad e_1e_2 \mapsto \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Where now $Cl^{\text{even}}(0, 2) \cong \mathbb{C}$ instead of \mathbb{L} .

Lemma 8.1 (Clifford Reduction Lemma). The following reduction formulas allow us to classify Clifford algebras,

$$Cl(p+1, q+1) \cong Cl(p, q) \otimes Cl(1, 1) \cong Cl(p, q) \otimes M_{2 \times 2}(\mathbb{R}) \quad (8.1)$$

$$Cl(q, p+2) \cong Cl(p, q) \otimes Cl(0, 2) \cong Cl(p, q) \otimes M_{2 \times 2}(\mathbb{R}) \quad (8.2)$$

$$Cl(q+2, p) \cong Cl(p, q) \otimes Cl(2, 0) \cong Cl(p, q) \otimes \mathbb{H} \quad (8.3)$$

Proof. We start with $Cl(p, q) \otimes Cl(V)$ and show that it is equal to the Clifford algebra of some vector space W with the right signature.

Let e_1, \dots, e_n be a basis of $\mathbb{R}^{p,q}$. Let V be a real two-dimensional vector space with metric and orthonormal basis $\{f_1, f_2\}$.

Also let $\mu = e_1 \dots e_n \in Cl(p, q)$ be the volume form, and let $\nu = f_1 f_2 \in Cl(V)$.

Then $\nu^2 = -1$ if the signature of the metric on V is $(2, 0)$ or $(0, 2)$, and $\nu^2 = 1$ if the signature is $(1, 1)$.

We then define a metric on $Cl(p, q) \otimes Cl(V)$ by $\langle a_1 \otimes b_1, a_2 \otimes b_2 \rangle = \langle a_1, a_2 \rangle \langle b_1, b_2 \rangle$.

We then define a subspace,

$$W = \text{span}\{e_1 \otimes \nu, \dots, e_n \otimes \nu, 1 \otimes f_1, 1 \otimes f_2\}$$

The inner product on W is then

$$|e_i \otimes \nu|^2 = \nu^2 |e_i|^2, \quad |1 \otimes f_i|^2 = |f_i|^2$$

We also have that $\nu \in \text{twcent}(Cl(V))$. So,

$$(e_i \otimes \nu)(e_j \otimes \nu) = -(e_j \otimes \nu)(e_i \otimes \nu)$$

$$(1 \otimes f_1)(1 \otimes f_2) = -(1 \otimes f_2)(1 \otimes f_1)$$

$$(e_i \otimes \nu)(1 \otimes f_j) = -(1 \otimes f_j)(e_i \otimes \nu)$$

Furthermore, this means

$$(e_i \otimes \nu)^2 = -|e_i \otimes \nu|^2 (1 \otimes 1)$$

$$(1 \otimes f_i)^2 = -|1 \otimes f_i|^2 (1 \otimes 1)$$

We can therefore see that the inclusion map $\phi : W \rightarrow Cl(p, q) \otimes Cl(V)$ extends to an algebra homomorphism $\phi : Cl(W) \rightarrow Cl(p, q) \otimes Cl(V)$.

We also have $(\nu^2)(e_i \otimes 1) = (e_i \otimes \nu)(1 \otimes f_i)(1 \otimes f_2)$, and $\nu^2 = \pm 1$, which means that $e_i \otimes 1 = \pm \phi(e_i \otimes \nu)\phi(1 \otimes f_i)\phi(1 \otimes f_2)$, so $e_i \otimes 1$ and $1 \otimes f_i$ are both in the image of ϕ , and together they span all of $\mathcal{C}\ell(p, q) \otimes \mathcal{C}\ell(V)$. So ϕ is an isomorphism. Finally, if the signature of V is $(1, 1)$, the signature of W is $(p, q) + (1, 1) = (p + 1, q + 1)$, and otherwise we get $(q, p) + (2, 0)$ or $(q, p) + (0, 2)$ where p and q get swapped because $\nu^2 = -1$. This completes the proof. \square

Lemma 8.2. Let us use the notation $M_{k \times k}(R) = M_k(R)$, where $R = \mathbb{R}, \mathbb{C}$, or \mathbb{H} (note here we are considering matrices whose entries are quaternions, but quaternions are not a field, hence we use the letter R).

$$(1) \quad M_k(R) \otimes M_\ell(R) \cong M_{k\ell}(R)$$

$$(2) \quad \mathbb{H} \otimes \mathbb{H} = M_4(\mathbb{R})$$

$$(3) \quad \mathbb{C} \otimes \mathbb{H} \cong M_2(\mathbb{C})$$

$$(4) \quad \mathbb{C} \otimes \mathbb{C} \cong \mathbb{C} \oplus \mathbb{C}$$

Proof. (1) This follows from the fact that $L(V \otimes W) = L(V) \otimes L(W)$.

(2) We will construct an isomorphism directly. Recall that $\mathbb{H} = \mathbb{R}^4$ as vector spaces, so multiplying a vector $v \in \mathbb{R}^4$ by a quaternion can be defined by simply interpreting v as an element of \mathbb{H} . Let $T : \mathbb{H} \otimes \mathbb{H} \rightarrow M_4(\mathbb{R})$ be defined by,

$$T(p \otimes q)(v) = pv\bar{q}$$

Clearly T is injective, since $|pv\bar{q}| = |v|^2|p|^2|q|^2 = 0$ iff $p = 0$ or $\bar{q} = 0$.

Finally, $\dim \mathbb{H} \otimes \mathbb{H} = 16 = \dim M_4(\mathbb{R})$. So T is an isomorphism.

Furthermore, if we define $|A|^2 = |\det A|^2$ we see that $|T(p \otimes q)|^2 = 4|p|^2|q|^2$, so T is an isometry.

(3) Recall that \mathbb{C} is a subalgebra of \mathbb{H} . So we can restrict T to $\mathbb{C} \otimes \mathbb{H}$.

Then $T(i \otimes 1)^2 = -1$, so $J = T(i \otimes 1)$ is a complex structure on \mathbb{H} and we have $\mathbb{H} = \mathbb{C}^2$ as vector spaces.

Similarly, $T(1 \otimes q)$ is right multiplication by \bar{q} .

Therefore $T(1 \otimes q)$ acts as a linear map from \mathbb{C}^2 to \mathbb{C}^2 , and since T was shown to be surjective on the previous part, we are done.

(4) We know $\mathbb{C} \otimes \mathbb{C} = \mathbb{H}$.

We also know that $T(1 \otimes i)$ right-multiplies by i . We have,

$$T(1 \otimes i)(a, b) = (a + T(i \otimes 1)b)i = ai + T(i \otimes 1)bi = ai + T((i \otimes 1)(1 \otimes i))b = ai - iT(i \otimes 1)b$$

So, in matrix form we can write

$$[T(1 \otimes i)] = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$$

Which means we have,

$$\mathbb{C} \otimes \mathbb{C} = \left\{ t \mathbb{1} + s \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} \mid t, s \in \mathbb{R} \right\} = \mathbb{C} \oplus \mathbb{C}$$

□

Theorem 8.1 (Classification of Clifford Algebras). Let p, q be positive integers and set $n = p + q$. Then $Cl(p, q)$ is given by the following table according to the value of $p - q \bmod 8$.

$p - q \bmod 8$	μ^2	$Cl(p, q)$
0,6	± 1	$M_n(\mathbb{R})$
2,4	± 1	$M_n(\mathbb{H})$
1,5	-1	$M_n(\mathbb{C})$
3	1	$M_n(\mathbb{H}) \oplus M_n(\mathbb{H})$
7	1	$M_n(\mathbb{R}) \oplus M_n(\mathbb{R})$

Proof. Apply the previous lemmas.

□

Remark. The fact that this pattern repeats with period 8 is called the **Bott periodicity** phenomenon. It can also be proven using algebraic topology, and can be seen to relate to the homotopy groups of spheres.

8.1.2 Self Duality

Lemma 8.3. Let A be an algebra of matrices. Then $\text{cent}(A)$ consists of matrices which are constant multiples of the identity.

Proof. Let $\mu_0 \in \text{cent}(A)$. Then let m_{ij} denote the matrix consisting only of a 1 in the i, j position.

Since μ_0 commutes with everything, $\mu_0 m_{ii} = m_{ii} \mu_0$ for all i . The matrix $\mu_0 m_{ii}$ consists just of the i 'th column of μ_0 , with zeros everywhere else, and similarly the matrix $m_{ii} \mu_0$ consists of the i 'th row of μ_0 and zeros everywhere else. The only way these two matrices can possibly be equal is if every element except those on the diagonal is zero. So μ_0 is diagonal.

Furthermore, $\mu_0 m_{ij} = m_{ij} \mu_0$. The matrix $\mu_0 m_{ij}$ consists of just the j 'th column of μ_0 , shifted to the i 'th spot, and similarly $m_{ij} \mu_0$ consists of just the i 'th row of μ_0 shifted to the j 'th spot. That is, we have $(\mu_0)_{ii} = (\mu_0)_{jj}$ for all i, j . So $\mu_0 = c \mathbb{1}$ for some constant c .

□

Recall that the volume form is $\mu = e_1 \dots e_n$, and $\mu^2 = 1$ if $p - q = 0$ or $3 \bmod 4$, and $\mu^2 = -1$ otherwise.

In the case where $\mu^2 = -1$, we see that μ defines a complex structure on $Cl(p, q)$.

In the case where $\mu^2 = 1$, we see that μ defines a real structure, and hence splits $Cl(p, q)$ into self-dual and anti-self-dual

parts, $\text{Cl}(p, q) = \text{Cl}^+(p, q) \oplus \text{Cl}^-(p, q)$, where these are two-sided ideals iff $p - q = 3 \pmod 4$.

Furthermore, when n is even, $\mu \in \text{cent}(\text{Cl}^{\text{even}}(p, q))$ and when n is odd, $\mu \in \text{cent}(\text{Cl}(p, q))$. So we have the following lemma.

Lemma 8.4. Let $\mu \in \text{Cl}(p, q)$. Let $A = \text{Cl}(p, q)$ when n is odd or $\text{Cl}^{\text{even}}(p, q)$ when n is even.

- (i) If $\mu^2 = 1$ and $\mu \in \text{cent}(A)$ then A decomposes into $A^+ \oplus A^-$, and with respect to this decomposition we have $\mu = \pm(1, -1)$
- (ii) If $\mu^2 = -1$ and $\mu \in \text{cent}(A)$, then $A = B_{\mathbb{R}}$, where B is some complex algebra and $\mu = i$ when regarded as an element of B .

Proof. By the classification theorem, $\text{Cl}(p, q)$ is a matrix algebra or a direct sum of matrix algebras.

Suppose first that $\text{Cl}(p, q)$ can be written as $M_n(\mathbb{F})$. We then know that either $\text{cent}(\text{Cl}(p, q)) = \text{span}\{1, \mu\}$ (if n is odd) or $\text{cent}(\text{Cl}(p, q)) = \text{span}\{1\}$ (if n is even). So the center of $\text{Cl}(p, q)$ contains at most all of $\{1, -1, \mu, -\mu\}$.

If n is odd, then $\mu^2 = 1$ implies that $\mu = \pm 1$. Therefore $\text{cent}(\text{Cl}(p, q)) = \text{span}\{1\}$.

Similarly if $\text{Cl}^{\text{even}}(p, q)$ can be written as a matrix algebra, $\text{cent}(\text{Cl}^{\text{even}}(p, q)) = \text{span}\{1\}$.

So the cases where $A \subseteq M_n(R)$ are trivial, since $A^+ = A$ and $A^- = \{0\}$.

Now we can prove the lemma in the case where $A \subseteq M_n(R) \oplus M_n(R)$.

Suppose $\mu^2 = 1$ and $\mu \in \text{cent}(A)$

Then, with respect to the decomposition $M_n(R) \oplus M_n(R)$ we have $\mu = (a, b)$, where a and b both square to the identity.

Furthermore μ being in the center implies that $a = \pm 1$ and $b = \pm 1$.

Finally, note that when n is odd (even) the isomorphism $\phi : \text{Cl}(p, q) \rightarrow A$ ($\phi : \text{Cl}^{\text{even}}(p, q) \rightarrow A$) takes $\phi(1) = (1, 1)$ and $\phi(-1) = (-1, -1)$. Since $\mu \neq \pm 1$, we must have $\mu = \pm(1, -1)$ as required.

The case $\mu^2 = -1$ is trivial from the fact that it is a complex structure. □

8.2 Lecture 23: Pinor and Spinor Representations

8.2.1 Classification of Pinor and Spinor Representations

The irreducible representations of the pin group (or, more generally, $\text{Cl}(p, q)$) are called pinors, and the irreducible representations of the spin group (more generally $\text{Cl}^{\text{even}}(p, q)$) are called spinors.

Recall that a representation is a mapping from a group or algebra A to the general linear group (note that this forms a matrix algebra, which is why the tools of representation theory works for both groups and algebras).

In particular, an algebra representation is an algebra homomorphism.

Definition 8.1 (Irreducible Representation I). Let $\rho : A \rightarrow \text{GL}(V)$ be an algebra representation. Then ρ is said to be **irreducible** if it is impossible to split V into a direct sum $V = V_1 \oplus V_2$, so that ρ also splits into $\rho = \rho_1 \oplus \rho_2$, where $\rho_1 : A \rightarrow \text{GL}(V_1)$ and $\rho_2 : A \rightarrow \text{GL}(V_2)$ are both representations.

Definition 8.2 (Invariant Subspace). Let $\rho : A \rightarrow \text{GL}(V)$ be a representation. Then a subspace $W \subseteq V$ is said to be A -invariant if $\rho(a)V \subseteq V$ for all $a \in A$.

Remark. The trivial invariant subspaces are $\{0\}$ and $V \subseteq V$.

Definition 8.3 (Irreducible Representation II). Let $\rho : A \rightarrow \text{GL}(V)$ be a representation. Then ρ is said to be **irreducible** if there does not exist any non-trivial A -invariant subspace of V .

Lemma 8.5. Definitions I and II of irreducible representations are equivalent.

Proof. Suppose V had a non-trivial A -invariant subspace W . It suffices to show that there exists an A -invariant subspace U so that $W \cap U = \{0\}$ and so that $V = W \oplus U$.

The reverse direction is much simpler. □

Example 8.5. The algebra $A = M_n(\mathbb{F})$, where \mathbb{F} is not \mathbb{C} , has only one irreducible representation, which is the standard one on \mathbb{F}^n where $\rho(a)$ just acts on a vector by left-multiplication.

Example 8.6. If $\mathbb{F} = \mathbb{C}$ then there are two irreducible representations, which are on \mathbb{C}^n and $\overline{\mathbb{C}}^n$.

That is, there are two maps $\rho_1 : M_n(\mathbb{C}) \rightarrow \text{GL}(\mathbb{C}^n)$ and $\rho_2 : M_n(\mathbb{C}) \rightarrow \text{GL}(\overline{\mathbb{C}}^n)$, where $\rho_1(a)v = av$ and $\rho_2(a)v = \bar{a}v$.

A consequence of the previous two examples is the following.

Theorem 8.2 (Pinor Representations). The irreducible representations of $\text{Cl}(p, q)$ (equivalently $\text{Pin}(p, q)$) are classified in the following table. We denote the representation space by \mathbb{P} .

$p - q \bmod 8$	$\text{Cl}(p, q)$	\mathbb{P}	ρ
0,6	$M_n(\mathbb{R})$	\mathbb{R}^n	$\rho(a)v = av$
2,4	$M_n(\mathbb{H})$	\mathbb{H}^n	$\rho(a)v = av$
1,5	$M_n(\mathbb{C})$	\mathbb{C}^n and $\overline{\mathbb{C}}^n$	$\rho(a)v = av$ and $\rho(a)v = \bar{a}v$
3	$M_n(\mathbb{H}) \oplus M_n(\mathbb{H})$	$\mathbb{P}^+ = \mathbb{H}^n \oplus \{0\}$ and $\mathbb{P}^- = \{0\} \oplus \mathbb{H}^n$	$\rho(a)(v, w) = (a_1v, w)$ and $\rho(a)(v, w) = (v, a_2w)$
7	$M_n(\mathbb{R}) \oplus M_n(\mathbb{R})$	$\mathbb{P}^+ = \mathbb{R}^n \oplus \{0\}$ and $\mathbb{P}^- = \{0\} \oplus \mathbb{R}^n$	$\rho(a)(v, w) = (a_1v, w)$ and $\rho(a)(v, w) = (v, a_2w)$

Note that overall, in the last two cases we have a representation on $\mathbb{P}^+ \oplus \mathbb{P}^-$, which splits into two irreducible representations.

Remark. The representations of $\text{Cl}(p, q)$ for $p - q = 3$ or $7 \bmod 8$ are inequivalent. We can see that this is the case because of the fact that $\mu \in \text{Cl}(p, q)$ acts differently on either factor, since $\mu = (1, -1)$ or $(-1, 1)$. That is, in the first case we find $\rho(\mu)(v, w) = (v, -w)$ and in the second case we have $\rho(\mu)(v, w) = (v, w)$. Therefore, a choice of orientation (ie $\mu = (1, -1)$ or $\mu = (-1, 1)$) determines the representation.

Remark. The same table classifies representations of the group $\text{Pin}(p, q)$ because we can just restrict the domain of ρ , and it turns out that the representations are still irreducible.

Theorem 8.3 (Spinor Representations). The irreducible representations of $Cl^{\text{even}}(p, q)$ (equivalently $\text{Spin}(p, q)$) are classified in the following table, and are called the **spinor representations** of $\text{Spin}(p, q)$. We denote the representation space by \mathbb{P} .

$p - q \bmod 8$	$Cl^{\text{even}}(p, q)$	\mathbb{P}	ρ
0	$M_n(\mathbb{R}) \oplus M_n(\mathbb{R})$	$\mathbb{P}^+ = \mathbb{R}^n \oplus \{0\}$ and $\mathbb{P}^- = \{0\} \oplus \mathbb{R}^n$	$\rho(a)(v, w) = (av, w)$ and $\rho(a)(v, w) = (v, aw)$
1, 7	$M_n(\mathbb{R})$	\mathbb{R}^n	$\rho(a)v = av$
3, 5	$M_n(\mathbb{H})$	\mathbb{H}^n	$\rho(a)v = av$
2, 6	$M_n(\mathbb{C})$	\mathbb{C}^n and $\overline{\mathbb{C}}^n$	$\rho(a)v = av$ and $\rho(a)v = \bar{a}v$
4	$M_n(\mathbb{H}) \oplus M_n(\mathbb{H})$	$\mathbb{P}^+ = \mathbb{H}^n \oplus \{0\}$ and $\mathbb{P}^- = \{0\} \oplus \mathbb{H}^n$	$\rho(a)(v, w) = (av, w)$ and $\rho(a)(v, w) = (v, aw)$

Example 8.7. The representation $\text{Spin}(1) \rightarrow \text{GL}(\mathbb{R})$ is just the trivial representation. We just have $\rho(a)s = as$, where $s \in \mathbb{R}$ and $a = \pm 1$.

There are two obvious representations of $\text{Spin}(2, 0)$ on \mathbb{C} . Recall that $\text{Spin}(2, 0)$ is isomorphic to $\text{SO}(2) \cong \text{U}(1)$, which means that $\text{Spin}(2, 0)$ acts on \mathbb{C} by multiplying on the left by $e^{i\theta}$, and on $\overline{\mathbb{C}}$ by multiplication by $e^{-i\theta}$.

For $\text{Spin}(4, 0)$, the irreducible representations are $\rho^+ : \text{Spin}(4, 0) \rightarrow M_n(\mathbb{H}) \oplus \{0\}$ and $\rho^- : \text{Spin}(4, 0) \rightarrow \{0\} \oplus M_n(\mathbb{H})$. One can show that $\text{Spin}(4, 0) \cong \text{SU}(2) \times \text{SU}(2)$, which means each representation simply acts on $\mathbb{H} \oplus \{0\}$ or $\{0\} \oplus \mathbb{H}$ by multiplication on the left by an imaginary unit quaternion.

Physics Remark. Finally, we can discuss the spin representations which appear in physics.

In physics, we essentially take the wave function of a particle to be a function which takes values in some vector space \mathbb{P} . This vector space is in general a representation space of some spin group.

Recall that the metric used in physics has signature $(3, 1)$. Therefore $p - q = 2 \bmod 8$, so the representation spaces of $\text{Spin}(3, 1)$ are all going to be of the form \mathbb{C}^n or $\overline{\mathbb{C}}^n$.

In the case $n = 2$ we get two representations. One on \mathbb{C}^2 and one on $\overline{\mathbb{C}}^2$. When we interpret these as representation spaces, the elements of \mathbb{C}^2 and $\overline{\mathbb{C}}^2$ are called **uncharged Weyl spinors**.

Similarly, the **uncharged Dirac spinors** are elements of the reducible representation space $\mathbb{C}^2 \oplus \overline{\mathbb{C}}^2$. As we can see, a Dirac spinor consists of a pair of Weyl spinors. In general, other reducible representations of $\text{Spin}(3, 1)$ for $n = 2$ give rise to other types of uncharged spin-1/2 particles.

For $n = 3$, we get a representation on \mathbb{C}^3 and a representation on $\overline{\mathbb{C}}^3$. Wave functions of uncharged spin-1 particles therefore give values in \mathbb{C}^3 .

In general, a non-electrically-charged spin $k/2$ particle can be described by one of the $n = 2k + 1$ dimensional representations of $\text{Spin}(3, 1)$.

Physics Remark. To take this a bit further, each irreducible representation ρ of $\text{Spin}(p, q)$ on some representation space \mathbb{P} also induces a representation of $\text{Spin}(p, q)$ on $Cl(\mathbb{P})$. This allows us to describe the action of an element of $\text{Spin}(p, q)$ on a system of particles. One can find that when we are dealing with spin- $n/2$ representations with n odd, we find that the system of particles satisfies a kind of antisymmetry, meaning we get **fermions**. Similarly, for spin- n particles we get symmetry, which implies these

are **bosons**. The relationship between spin and particle statistics is called the **spin-statistics theorem**.

Physics Remark. Charged particles are described by representations of the **complexified** Clifford algebra $Cl^{\mathbb{C}}(p, q) = Cl(p, q) \otimes \mathbb{C}$. In this case we get a complexified Spin group as the even unit elements, which can be related to the usual spin group by the formula $Spin^{\mathbb{C}}(p, q) = Spin(p, q) \times S^1 / \sim$, where we are quotienting by the relation $(a, \theta) \sim (-a, -\theta)$. The **electric charge** of a particle is the conserved quantity associated to the S^1 symmetry, and the **spin** of the particle is the conserved quantity associated to the $Spin(3, 1)$ symmetry.

Physics Remark. Describing particles which have spin on the domain $\mathbb{R}^{3,1}$ is fairly straightforward since we can just describe them as functions $\psi : \mathbb{R}^{3,1} \rightarrow \mathbb{C}^n$. However, the situation for other domains is more difficult. In general, for a pseudo-Riemannian manifold M of signature (p, q) we construct a principal $Spin(p, q)$ bundle $S \rightarrow M$, and then the spinors are sections of the associated representation bundle \tilde{S} , which is defined so that the representation is of the form $\rho : (S \rightarrow M) \rightarrow (\mathfrak{gl}(\tilde{S}) \rightarrow M)$. The bundle \tilde{S} is called the **Spinor bundle**. The study of manifolds with a principal $Spin(p, q)$ bundle is called **spin geometry**.

8.2.2 Duality and Triality

Recall that there is a map $\widetilde{\text{Ad}} : Spin(p, q) \rightarrow SO(p, q)$, and that $SO(p, q) \subseteq GL(\mathbb{R}^{p,q})$. Therefore, we get a representation of $Spin(p, q)$ on $\mathbb{R}^{p,q}$, given by the twisted adjoint representation. When viewing this as a representation of the spin group we call this the vector representation.

Definition 8.4 (Vector Representation). The **vector representation** of $Spin(p, q)$ is the map $\widetilde{\text{Ad}} : Spin(p, q) \rightarrow SO(p, q)$.

Remark. For $Spin(8)$ we therefore have the two spinor representations $\rho^+ : Spin(8) \rightarrow GL(\mathbb{R}^8 \oplus \{0\})$, $\rho^- : Spin(8) \rightarrow GL(\{0\} \oplus \mathbb{R}^8)$, and now a third representation $\widetilde{\text{Ad}} : Spin(8) \rightarrow GL(\mathbb{R}^8)$. These are all **inequivalent** representations! They are all related in a mysterious way called **triality**.

Definition 8.5 (Duality). Let V and W be real vector spaces. A **duality** between V and W is a bilinear map $F : V \times W \rightarrow \mathbb{R}$, where if $v, w \neq 0$ we have $F(\cdot, w) \neq 0 \in V^*$ and $F(v, \cdot) \neq 0 \in W^*$.

Remark. The map $L : V \rightarrow W^*$ given by $L(v)(w) = F(v, w)$ is injective since $L(v) \neq 0$ by definition, as is $R : W \rightarrow V^*$ where $R(w)(v) = F(v, w)$. Therefore if two vector spaces are dual, we have $\dim V = \dim W = \dim V^* = \dim W^*$ and L and R are isomorphisms.

Definition 8.6 (Norm-Duality). Let V and W be vector spaces with positive definite inner products. Then a **norm-duality** is a bilinear map F so that $|F(v, w)| \leq \|v\|\|w\|$ and so that for all $v \neq 0$ there exists $w \neq 0$ with $|F(v, w)| = \|v\|\|w\|$ (and vice versa for all $w \neq 0$)

Lemma 8.6. Any norm-duality is a duality.

Proof. Suppose $v \neq 0$. Then there is $w \neq 0$ so that $|F(v, w)| = |v||w| \neq 0$, so $F(v, w) \neq 0$ and vice versa. So $L(v) \neq 0$ and $L(w) \neq 0$ □

Example 8.8. Let $V = W$ with the same inner product. Let $F(v, w) = \langle v, w \rangle$. Then clearly by Cauchy-Schwartz, $|\langle v, w \rangle| \leq |v||w|$. Additionally, if $v \neq 0$ then $|\langle v, v \rangle| = |v||v|$. So a positive definite inner product defines a norm-duality from a vector space to itself.

Remark. Since a norm-duality is a duality, we get a map $L : V \rightarrow W^*$, and we already have a map $\sharp : W^* \rightarrow W$.

Lemma 8.7. The map $T = \sharp \circ L : V \rightarrow W$ is an isometry.

Proof. Let $w = T(v)$. Then $|T(v)|^2 = |F(v, T(v))|^2 \leq |v||T(v)|$. So $|T(v)| \leq |v|$.
Conversely, for all $v \neq 0$ there exists w so that $|v||w| = |F(v, w)|$. We then have $|F(v, w)| = |\langle T(v), w \rangle| \leq |T(v)||w|$. So $|v||w| \leq |T(v)||w|$, which means $|v| \leq |T(v)|$.
So $|T(v)| = |v|$ as required. □

Remark. The key thing here is that given a norm-duality between V and W , we get an isometry between V and W .

Definition 8.7 (Triality). Let U, V, W be real vector spaces. A **triality** is a trilinear map $T : U \times V \times W \rightarrow \mathbb{R}$ so that if any two arguments are nonzero, then the induced linear functional on the third argument is nonzero.

That is, $T(a, b, \cdot) \neq 0$, $T(\cdot, b, c) \neq 0$, and $T(a, \cdot, c) \neq 0$ whenever $a, b, c \neq 0$.

Remark. Suppose we have a triality between U, V, W . Then if we fix $c \neq 0$ we get a duality between U and V by the map $m = T(\cdot, \cdot, c) : U \times V \rightarrow \mathbb{R}$. Without loss of generality, we therefore have $\dim U = \dim V = \dim W$.

Remark. Given $a \in U$ and $a \neq 0$, we get $L_a = m(a, \cdot) : V \rightarrow W^*$. Similarly for $b \in V$ with $b \neq 0$ we get $R_b = m(\cdot, b)$. These are isomorphisms $V \cong W^*$ and $U \cong W^*$. So we get an isomorphism between U and V which depends on the choice of a, b .

In general, by fixing two nonzero elements from two of U, V or W we get an isomorphism between the two vector spaces.

Lemma 8.8. Let $e \in W$ be defined by $e = \sharp(T(a, b, \cdot))$ and define $M : W \times W \rightarrow W$ by

$$M(w_1, w_2) = \sharp(T(R_b^{-1}(w_1), L_a^{-1}(w_2), \cdot))$$

Then e is an identity for M . That is $M(e, v) = M(v, e) = v$ for all v .

Proof. This is done directly. Let $w \in W$

$$\begin{aligned}
 M(e, w) &= \sharp(T(R_b^{-1}(e), L_a^{-1}(w), \cdot)) \\
 &= \sharp(T(R_b^{-1}(R_b(a)), L_a^{-1}(w), \cdot)) \\
 &= \sharp(T(a, L_a^{-1}(w), \cdot)) \\
 &= L_a(L_a^{-1}(w)) \\
 &= w
 \end{aligned}$$

The other case is identical. □

Remark. Since we can construct a multiplication $M : W \rightarrow W \rightarrow W$, with an identity e , we see that a triality along with a nonzero element from U and V gives W the structure of a real unital algebra.

Remark. If we use M to define multiplication of elements $w_1, w_2 \in W$, we can see that left or right-multiplication gives an isomorphism. That is, if $w_1 \neq 0$ then $R_b^{-1}(w_1) \neq 0$, and if $w_2 \neq 0$ we have $L_a^{-1}(w_2) \neq 0$, so $T(R_b^{-1}(w_1), L_a^{-1}(w_2), \cdot) \neq 0$. Hence $w_1 w_2 = 0$ implies that either $w_1 = 0$ or $w_2 = 0$.

Therefore, we in fact have constructed a real division algebra out of W .

Corollary 8.1. If T is a triality of U, V, W , then each of U, V, W can be made into a real unital division algebra.

Remark. We saw in an earlier section that the only real unital division algebras are in dimensions 1, 2, 4, and 8. Therefore, there are not many possible trialities!

Definition 8.8 (Norm-Triality). Let U, V, W be equipped with positive definite inner products. A **norm-triality** is a trilinear map $T : U \times V \times W \rightarrow \mathbb{R}$ with $T(u, v, w) \leq |u||v||w|$ and so that whenever two elements are nonzero there is a third which satisfies $T(u, v, w) = |u||v||w|$.

Remark. In this case, if we fix two elements a and b with $|a| = |b| = 1$, then we see that we get **isometries** between U, V and W .

Corollary 8.2. A normed triality between U, V , and W gives rise to a compositional division algebra structure on each of U, V , and W .

Proof. We have,

$$\begin{aligned}
 \langle M(a, b), c \rangle &= T(a, b, c) \\
 \implies |M(a, b), M(a, b)| &= |M(a, b)|^2 \leq |a||b||M(a, b)| \\
 \implies |M(a, b)| &\leq |a||b|
 \end{aligned}$$

and similarly $|a||b| \leq |M(a, b)|$. □

In summary, a normed triality gives rise to a composition algebra and vice versa.

8.3 Lecture 24: $\text{Spin}(8)$ and Triality

8.3.1 Properties of $\text{Cl}(8, 0)$

Recall that the octonions \mathbb{O} are isometric to \mathbb{R}^8 with the standard inner product. We can therefore describe elements of $\text{GL}(\mathbb{O} \oplus \mathbb{O})$ as elements of $M_{16}(\mathbb{R})$.

Definition 8.9. Define a map $\beta : \mathbb{O} \rightarrow M_{16}(\mathbb{O})$ by the formula,

$$\beta(v) = \begin{bmatrix} 0 & R_v \\ -R_{\bar{v}} & 0 \end{bmatrix}$$

Where R_v and $R_{\bar{v}}$ represent right-multiplication by v and \bar{v} respectively.

Remark. The map β is \mathbb{R} -linear.

Lemma 8.9. The map β is injective.

Proof. Suppose $v \in \ker(\beta)$. Then for all $\alpha, \beta \in \mathbb{R}^8$ we have

$$\beta(v) \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} R_v b \\ -R_{\bar{v}} a \end{bmatrix}$$

Therefore we need $R_v b = R_{\bar{v}} a = 0$. So $R_v b = bv = 0$ for all b , which means $v = 0$. So $\ker(\beta) = \{0\}$ as required. □

Definition 8.10. Define the vector space $V = \text{image}(\beta)$. Then we define an inner product on V by the formula,

$$\langle \beta(v), \beta(u) \rangle = \langle u, v \rangle$$

Remark. Therefore, β is an isometry by definition.

Remark. We have

$$\beta(u)\beta(v) = \begin{bmatrix} -R_u R_{\bar{v}} & 0 \\ 0 & -R_{\bar{v}} R_u \end{bmatrix}$$

Furthermore, this implies that,

$$(\beta(u)\beta(v) + \beta(v)\beta(u)) \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -(a\bar{v})u - (a\bar{u})v \\ -(bv)\bar{u} - (bu)\bar{v} \end{bmatrix} = -2\langle u, v \rangle \begin{bmatrix} a \\ b \end{bmatrix}$$

Where we have used equation (7.2) in the last step.

Corollary 8.3. By the universal property, β extends to a map $\beta : Cl(8, 0) \rightarrow M_{16}(\mathbb{R})$. Furthermore, since $M_{16}(\mathbb{R})$ has no non-trivial two-sided ideals, the extension of β is injective. Finally, since $\dim Cl(8, 0) = \dim GL(\mathbb{O} \oplus \mathbb{O})$, we see that β is an isomorphism between $Cl(8, 0)$ and $GL(\mathbb{O} \oplus \mathbb{O})$.

Lemma 8.10. Let $\{e_0, \dots, e_7\}$ be an oriented orthonormal basis for \mathbb{R}^8 . Recall that the volume form is $\mu = e_0 \dots e_7$. Then

$$\beta(\mu) = \begin{bmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{bmatrix}$$

Proof. One directly computes $\beta(\mu) = \beta(e_0) \dots \beta(e_7)$ by tedious calculation, using the fact that $R_{e_i} R_{e_j} u = (ue_j)e_i$ for all i, j as well as the other octonion identities. \square

Remark. Let $\mathbb{P}^+ = \mathbb{O} \oplus \{0\}$ and $\mathbb{P}^- = \{0\} \oplus \mathbb{O}$. Then the above says $Cl(8, 0) \cong GL(\mathbb{P}^+ \oplus \mathbb{P}^-)$.

Corollary 8.4. We have,

$$Cl^{\text{even}}(8, 0) \cong \left\{ \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} : A, B \in GL(\mathbb{O}) \right\} = GL(\mathbb{P}^+) \oplus GL(\mathbb{P}^-)$$

Which means $Cl^{\text{even}}(8, 0)$ splits up into

$$Cl_+^{\text{even}}(8, 0) \cong GL(\mathbb{P}^+), \quad Cl_-^{\text{even}}(8, 0) \cong GL(\mathbb{P}^-)$$

Remark. Define projection maps p_+ and p_- by $p_+(a, b) = (a, 0)$ and $p_-(a, b) = (0, b)$. Then the above corollary tells us that $p_+ \circ \beta$ is the positive spinor representation of $Spin(8)$ and that $p_- \circ \beta$ is the negative spinor representation.

Definition 8.11. We define an inner product on $\mathbb{P}^+ \oplus \mathbb{P}^-$ by the formula

$$\left\langle \begin{bmatrix} u \\ v \end{bmatrix}, \begin{bmatrix} u' \\ v' \end{bmatrix} \right\rangle = \langle u, u' \rangle + \langle v, v' \rangle$$

Lemma 8.11. With respect to the inner product defined above, $\beta(\alpha)$ is skew-adjoint for all $\alpha \in Cl(8, 0)$.

Proof. We calculate,

$$\left\langle \beta(\alpha) \begin{bmatrix} u \\ v \end{bmatrix}, \begin{bmatrix} u' \\ v' \end{bmatrix} \right\rangle = \langle v\alpha, u' \rangle - \langle u\bar{\alpha}, v' \rangle$$

Then note that $R_\alpha^\dagger = R_{\bar{\alpha}}$, so

$$\begin{aligned} \langle v\alpha, u' \rangle - \langle u\bar{\alpha}, v' \rangle &= \langle R_\alpha(v), u' \rangle - \langle R_\alpha^\dagger(u), v' \rangle \\ &= \langle v, u'\bar{\alpha} \rangle - \langle u, v'\alpha \rangle \\ &= -(\langle u, v'\alpha \rangle - \langle v, u'\bar{\alpha} \rangle) \\ &= \left\langle \begin{bmatrix} u \\ v \end{bmatrix}, -\beta(\alpha) \begin{bmatrix} u' \\ v' \end{bmatrix} \right\rangle \end{aligned}$$

As required. □

Corollary 8.5. If $\alpha \in \mathcal{Cl}(8, 0)$, then $\beta(\hat{\alpha}) = \beta(\alpha)^\dagger$.

Proof. Let A, B be the matrices defined so that $\beta(\alpha) = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$. Since $\beta(\alpha)$ is skew adjoint, A and B must also be. Now suppose we decompose α into $\alpha = \sum_{i \in \mathbb{N}, k \leq 8} \alpha_{i1} \dots \alpha_{ik}$. Then $\hat{\alpha} = \sum_{i \in \mathbb{N}, k \leq 8} (-1)^k \alpha_{ik} \dots \alpha_{i1}$. In matrix form we get,

$$\beta(\hat{\alpha}) = \sum_{i \in \mathbb{N}, k \leq 8} (-1)^k \begin{bmatrix} A_{ik} \dots A_{i1} & 0 \\ 0 & B_{ik} \dots B_{i1} \end{bmatrix}$$

Since each A_{ij} is skew-adjoint, we have $(A_1 \dots A_k)^\dagger = A_k^\dagger \dots A_1^\dagger = (-1)^k A_k \dots A_1$. So we are done. □

Corollary 8.6. $\text{Spin}(8) \subseteq \text{SO}(\mathbb{P}^+) \oplus \text{SO}(\mathbb{P}^-)$.

Proof. Consider a decomposable even element $\alpha = v_1 \dots v_{2r} \in \text{Spin}(8)$. Then,

$$\alpha \hat{\alpha} = (-1)^{2r} v_1 \dots v_{2r} v_{2r} \dots v_1 = |v_1|^2 \dots |v_{2r}|^2 = 1$$

So

$$\beta(\alpha \hat{\alpha}) = \beta(\alpha) \beta(\hat{\alpha}) = \beta(\alpha) \beta(\alpha)^\dagger = \mathbb{1}$$

In other words, $AA^\dagger = \mathbb{1}$ and $BB^\dagger = \mathbb{1}$, so $\beta(\alpha) \in \text{O}(\mathbb{P}^+) \times \text{O}(\mathbb{P}^-)$. Since $\text{Spin}(8)$ is connected, $\mathbb{1} \in \text{image}(\beta)$, and β is a continuous mapping, we therefore see that $\beta(\alpha) \subseteq \text{SO}(\mathbb{P}^+) \times \text{SO}(\mathbb{P}^-)$ as required. □

Remark. So we have two representations $\rho_+ = p_+ \circ \beta|_{\text{Spin}(8)}$ and $\rho_- = p_- \circ \beta|_{\text{Spin}(8)}$ which are the positive and negative spinor representations.

We also have a third representation $\rho_0 = \widetilde{\text{Ad}}$ defined so that $\widetilde{\text{Ad}}_\alpha(v) = \tilde{\alpha}v\alpha^{-1}$. Since every element of $\text{Spin}(8)$ is even, we have $\tilde{\alpha} = \alpha$ for all $\alpha \in \text{Spin}(8)$. Therefore $\widetilde{\text{Ad}} = \text{Ad}$ when we restrict to $\text{Spin}(8)$.

Remark. One can show that for any vector space V with a metric of signature p, q , the group $\text{Spin}(V)$ can be written as,

$$\text{Spin}(V) = \{\alpha \in \text{Cl}^{\text{even}}(V)^* : \widetilde{\text{Ad}}_\alpha(V) \subseteq V, |\alpha|^2 = \pm 1\}$$

Where we have used the notation that A^* denotes the group of invertible elements of an algebra A .

Remark. Let us write

$$g = \begin{bmatrix} g_+ & 0 \\ 0 & g_- \end{bmatrix} \in \text{Cl}^{\text{even}}(8, 0) = \text{Cl}_+^{\text{even}}(8, 0) \oplus \text{Cl}_-^{\text{even}}(8, 0)$$

Then if $g \in \text{Spin}(8)$, we have $g_+, g_- \in \text{SO}(\mathbb{O})$. Furthermore, we can abuse notation and write $\rho_+(g) = g_+$ and $\rho_-(g) = g_-$. Then how do we write $\rho_0(g)$?

Theorem 8.4 ($\text{Spin}(8)$ Triality Theorem). Let (g_0, g_+, g_-) be an ordered triple of elements of $\text{O}(\mathbb{O})$. Then the following statements are equivalent.

1. $g = \text{diag}(g_+, g_-) \in \text{Spin}(8)$, and $g_0 = \text{Ad}_g$.
2. $g_+(xy) = g_-(x)g_0(y)$ for all $x, y \in \mathbb{O}$.

Remark. Since having g_+ and g_- determines g_0 completely, it is useful to denote $\text{Spin}(8)$ as consisting of all ordered triples (g_0, g_+, g_-) satisfying the above criteria.

Proof. The goal of the proof is to show that an equivalent condition for $\text{Ad}_\rho(v) \in \mathbb{O}$ is that $g_-(x)g_0(v) = g_+(xv)$ for all $x, v \in \mathbb{O}$. Recall that \mathbb{O} is isometric to $V = \beta(\mathbb{O})$, where we have defined $\beta(v)$ by the formula,

$$\beta(v) = \begin{bmatrix} 0 & R_v \\ -R_{\bar{v}} & 0 \end{bmatrix}$$

Also recall that,

$$\text{Spin}(V) = \{\alpha \in \text{Cl}^{\text{even}}(V)^* : \widetilde{\text{Ad}}_\alpha(V) \subseteq V, |\alpha|^2 = \pm 1\}$$

Now let

$$\rho = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

Then by the above definition of $\text{Spin}(V)$ we have that $\rho \in \text{Spin}(8)$ iff it is true that $A, B \in \text{O}(\mathbb{O}) \cong \text{O}(V)$ and that $\beta(\text{Ad}_\rho(v)) \in V$. Let us then write out $\text{Ad}_\rho(v)$ directly. We have,

$$\text{Ad}_\rho(v) = \tilde{\rho}v\rho^{-1}$$

Applying the isometry, this is equivalent to writing,

$$\beta(\text{Ad}_\rho(v)) = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} 0 & R_v \\ -R_{\bar{v}} & 0 \end{bmatrix} \begin{bmatrix} A^\dagger & 0 \\ 0 & B^\dagger \end{bmatrix} = \begin{bmatrix} 0 & AR_v B^\dagger \\ -BR_{\bar{v}} A^\dagger & 0 \end{bmatrix}$$

So $\text{Ad}_\rho(v) \in \mathbb{O}$ if and only if $AR_v B^\dagger = R_u$ and $BR_{\bar{v}} A^\dagger = R_{\bar{u}}$ for some $u \in V$. Note that since these are adjoint to each other, we only have to prove one of these. We can take $AR_v B^\dagger = R_u$ and apply both sides to the element $1 \in \mathbb{O}$, which gives us,

$$\begin{aligned} u &= R_u(1) \\ &= AR_v B^\dagger(1) \\ &= A(B^\dagger(1)v) \end{aligned}$$

So we have $g_0(v) = u$ defined by the above equation. This means that $\text{Ad}_g \in \text{Spin}(8)$ if and only if $g_+(g_-^{-1}(1)v) = g_0(v)$ for all v . In general, we can apply the same equation to $w \in \mathbb{O}$ to get

$$\begin{aligned} wu &= R_u(w) \\ &= AR_v B^\dagger(w) \\ &= A(B^\dagger(w)v) \end{aligned}$$

Set $x = g_-^{-1}(w)$. Then the above implies that $g_-(x)u = g_+(xv)$, so $g_-(x)g_0(v) = g_+(xv)$ as required. □

Remark. If $\text{diag}(g_+, g_-) \in \text{Spin}(8)$, then $g_+, g_- \in \text{SO}(\mathbb{O})$, and so the above theorem implies that $g_0 \in \text{SO}(\mathbb{O})$ as well.

Definition 8.12. Let C denote conjugation of octonions. Then we define $\rho'_\pm = C \circ \rho_\pm \circ C$.

Remark. Observe that $CR_v C = L_{\bar{v}}$.

Lemma 8.12. If the ordered triple (g_0, g_+, g_-) satisfies the triality theorem, then the triple (g'_-, g'_+, g'_0) also satisfies the triality theorem.

Proof. Direct calculation gives,

$$\begin{aligned} g'_+(xy) &= \overline{g_+(\overline{xy})} \\ &= \overline{g_+(\overline{yx})} \\ &= \overline{g_-(\overline{y})g_0(\overline{x})} \\ &= \overline{g_0(\overline{x})g_-(\overline{y})} \\ &= g'_0(x)g'_-(y) \end{aligned}$$

As we can see, this is the same condition as in the theorem except reordered as required. □

Definition 8.13 (Inner/Outer Automorphism). Let G be a group. An automorphism $a : G \rightarrow G$ is said to be **inner** if there is some $h \in G$ so that $a(g) = hgh^{-1}$. If a is not inner, we say it is **outer**.

Remark. If a is an inner automorphism, then since $hgh^{-1} = g$ for all $g \in \text{cent}(G)$, it follows that $a \text{cent}(G) = G$.

Definition 8.14 (Triality Automorphism I). The first outer automorphism of $\text{Spin}(8)$ is the map $\alpha : \text{Spin}(8) \rightarrow \text{Spin}(8)$ defined so that

$$\alpha(g_0, g_+, g_-) = (g'_-, g'_+, g'_0)$$

Remark. Clearly $\alpha \circ \alpha = \mathbb{1}$ since conjugation C is an involution. Additionally, one can check that α is a group homomorphism using the formula $g_+(xy) = g_-(x)g_0(y)$. Finally, since it is bijective it is a group automorphism.

Remark. Recall that $g_+(xy) = g_-(x)g_0(y)$ is equivalent to the equations $R_v = AR_vB^\dagger$ and $R_{\bar{v}} = BR_{\bar{v}}A^\dagger$. We have $g_0(v) = A(B^\dagger(1)v)$, so $g'_0(v) = \overline{A(B^\dagger(1)\bar{v})} = B(A^\dagger(1)v)$.

We also have $g_+(v) = Av$ and $g_-(v) = Bv$. Therefore

$$g'_0(v) = g_-(g_+^{-1}(1)v)$$

This means that (g'_0, g_-, g_+) satisfies the triality theorem.

Definition 8.15 (Triality Automorphism II). Let $\gamma : \text{Spin}(8) \rightarrow \text{Spin}(8)$ be the map defined by $\gamma(g_0, g_+, g_-) = (g'_0, g_-, g_+)$. Then γ is also a group automorphism.

Theorem 8.5 (Outer Automorphisms of $\text{Spin}(8)$). Let α and γ be the automorphisms defined above. Then α, γ , and $\alpha \circ \gamma$ are **outer automorphisms**. Furthermore, they are the **only** outer automorphisms of $\text{Spin}(8)$.

Proof. To show these are outer automorphisms it suffices to show that they do not fix $\text{cent}(\text{Spin}(8))$. Recall that $1, \mu, -1$, and $-\mu$ are elements of $\text{cent}(\text{Spin}(8))$.

Consider the isomorphism taking $a \in \text{Spin}(8)$ to $g = (\rho_0(a), \rho_+(a), \rho_-(a))$. Then we have,

$$1 \mapsto (1, 1, 1)$$

$$\mu \mapsto (-1, 1, -1)$$

$$-1 \mapsto (-1, -1, -1)$$

$$-\mu \mapsto (-1, -1, 1)$$

From this we immediately see that,

$$\alpha(-1) = -\mu \neq -1$$

$$\gamma(\mu) = -\mu \neq \mu$$

$$\alpha \circ \gamma(\mu) = -1 \neq \mu$$

So none of them are inner automorphisms.

Proving that these are the **only** outer automorphisms takes more work. □

9 Appendix I: Groups

The core content of the course begins with a review of linear algebra. In the actual course, it was assumed that the student is already familiar with the notion of a vector space, as well as the notions of linear maps, eigenvalues and eigenvectors, bases, isomorphisms, and inner product spaces.

In this chapter we will provide a summary of many of the definitions we will use in the course. See standard books for more detail [1, 8, 7, 12, 2]. I have attempted to make this chapter thorough enough that the reader does not need to constantly look up definitions and theorems from linear algebra and abstract algebra.

The section on sets is required knowledge, although the definition of the characteristic of a field is not too important. In this course the main fields we will see are \mathbb{R} and \mathbb{C} , so don't worry about knowing abstract algebraic field theory.

Much of the material on group theory I have included is also extraneous, and can be skipped. The main important thing to note is the definition of a **permutation**. Permutations are very useful tools inside and outside of group theory, and we will use them extensively. I have also included some proofs of important theorems in group theory, such as Cayley's theorem, which I believe lend some intuition to more abstract groups. However, we will not use these theorems in the course.

I have also included a short section on rings, although we will not use very many ring theoretic concepts in this course. I only included this section because when defining the Pfaffian of an antisymmetric bilinear form, we required the fact that $\mathbb{Z}[x]$ is a unique factorization domain. In general, feel free to skip the section on rings entirely.

Finally, I would like to emphasize the sections on linear algebra. Many readers will have taken a first year linear algebra course, but will not have seen some abstract definitions. If the reader wants to understand the proofs in later sections, it will be important to understand isomorphisms, linear maps, and bases.

9.1 Sets and Fields

Sets allow us to essentially describe the "setting"(s) of a mathematical problem. The settings in linear algebra are vector spaces, but there are many more special kinds of sets.

We often model new problems by combining sets in various ways. One of the simplest ways to combine two sets is called the cartesian product. The cartesian product is the collection of ordered pairs of elements from two sets. If we were to visually lay out the two sets on lines, then their Cartesian product would look like a rectangle, cut into grid squares for each pair of elements.

Definition 9.1 (Cartesian Product). Let S_1 and S_2 be any sets. Then the cartesian product $S_1 \times S_2$ of S_1 and S_2 is defined as the collection of ordered pairs of elements from S_1 and S_2 . That is,

$$S_1 \times S_2 = \{(a, b) : a \in S_1 \text{ and } b \in S_2\}$$

If S is a set and n is a positive integer, we often write,

$$S^n = \underbrace{S \times \dots \times S}_{n \text{ times}}$$

Example 9.1. If $S_1 = \{1, 2\}$ and $S_2 = \{3, 4\}$ then $S_1 \times S_2 = \{(1, 3), (1, 4), (2, 3), (2, 4)\}$.

Definition 9.2 (Set Union/Intersection). Let A and B be sets. Then we define the **union** of A and B by

$$A \cup B = \{x : x \in A \text{ or } x \in B\}$$

and the **intersection** of A and B by

$$A \cap B = \{x : x \in A \text{ and } x \in B\}$$

Definition 9.3 (Set Complement). Let A be a set and let B be a subset of A . Then the **complement** of B in A is denoted B^c and is given by

$$B^c = \{x \in A : x \text{ is not in } B\}$$

Definition 9.4 (Set Subtraction). Let A be a set and let B, C be subsets of A . Then we define the notation:

$$B \setminus C = B \cap C^c$$

That is, it is the set you get by removing all the common elements of B and C from B .

Definition 9.5 (Injective, Surjective, Bijective). Let $f : A \rightarrow B$ be any function between sets A and B .

1. Suppose that for all $b \in B$, there exists $a \in A$ so that $f(a) = b$. Then f is called **surjective**, or "onto"
2. Suppose that for all $a_1, a_2 \in A$, that $f(a_1) = f(a_2)$ if and only if $a_1 = a_2$. Then f is called **injective**, or "one-to-one".
3. If f is both injective and surjective, then we say f is **bijective**.

Definition 9.6 (Inverse Image/Fiber). Let $f : A \rightarrow B$ be any function between sets A and B . Let $C \subseteq B$ be any subset of B . Then we define,

$$f^{-1}(C) = \{a \in A : f(a) \in C\} \quad (9.1)$$

This is called the **inverse image** of C under f . If $C = \{b\}$, then we say $f^{-1}(\{b\})$ is the **fiber** of f over b .

The fiber of f over b is the set of all elements of A which get mapped to b by f .

A special kind of set that we often use to construct vector spaces is called a field. A field is a generalization of the idea of "rational numbers". A field is a set, and given some elements of this set you can perform operations which obey all the usual rules of arithmetic. That is, given two elements of a field, you can multiply, add, subtract, or divide them (unless the denominator is zero!)

Definition 9.7 (Field). A field \mathbb{F} is a set equipped with a notion of addition and multiplication. That is, for any two elements $x, y \in \mathbb{F}$ we can add them to produce a third element $x + y \in \mathbb{F}$ or we can multiply them to produce an element $xy \in \mathbb{F}$. These operations must follow the field axioms, which are listed as follows.

1. Commutativity: $x + y = y + x$, $xy = yx$
2. Associativity: $x + (y + z) = (x + y) + z$, $x(yz) = (xy)z$
3. Distributivity: $x(y + z) = xy + xz$

One example of a field we know and love is the rational numbers \mathbb{Q} . The rational numbers are any numbers which can be formed as the ratio of two integers. Another example is the real numbers \mathbb{R} . The real numbers are all the numbers which can be expressed as the limit of some convergent sequence of rational numbers. We will not worry too much about how the real numbers are defined rigorously. All that matters is that you can multiply, divide, add, and subtract them. We also care about the property of the real numbers, that adding 1 to a positive number never results in zero. This property is encapsulated in the following definition.

Definition 9.8 (Characteristic). Let \mathbb{F} be a field. Then the characteristic $\text{char}(\mathbb{F})$ is the smallest natural number n so that $1 + \dots + 1 = \sum_{i=1}^n (1) = 0$. If no such natural number exists, we say $\text{char}(\mathbb{F}) = 0$. In this course, we will primarily deal with the fields \mathbb{R} and \mathbb{C} which have characteristic zero.

9.2 Groups

9.2.1 Basic Definitions

In this course we will make some use of group theory. We are especially interested in groups of matrices, but we will also make a lot of use of permutations. This content can all be found in standard books [1, 7, 8].

Definition 9.9 (Group). Let S be a set and suppose that there is some function $f : S \times S \rightarrow S$ which takes two elements a and b of S and produces a third element, $f(a, b)$. We say that S is a group, with operation f , if the function f satisfies the following properties:

1. Associativity: for all $a, b, c \in S$ we have $f(a, f(b, c)) = f(f(a, b), c)$
2. Identity: There exists some element $e \in S$ so that $f(e, a) = f(a, e) = a$ for all $a \in S$.
3. Invertibility: For all $a \in S$ there is some element $a^{-1} \in S$ so that $f(a^{-1}, a) = f(a, a^{-1}) = e$.

Definition 9.10 (Subgroup). Let G be a group with operation f and let $H \subseteq G$. Then H is a subgroup of G if H is a group with respect to the same operation f as G .

Definition 9.11 (Order). Let G be a group. The order of G , denoted $|G|$, is the number of elements of G .

Example 9.2 (Trivial Group). The only group of order 1 is the trivial group, $\{1\}$. This is clearly a group since $1 \cdot 1 = 1$, and $1 = 1^{-1}$.

Example 9.3 (Integers). The integers \mathbb{Z} form a group, where the operation is $f(a, b) = a + b$. Observe that this indeed satisfies the axioms:

1. Associativity: $a + (b + c) = (a + b) + c$
2. The identity is $e = 0$. We have, $a + 0 = 0 + a = a$
3. The inverse element of a is $-a$. We have, $a + (-a) = 0$

The integers have infinite order.

Example 9.4. The set of $n \times n$ invertible matrices with entries in a field \mathbb{F} forms a group. We often refer to this group as the **general linear group**. The symbol for it is $\text{GL}_n(\mathbb{F})$.

Example 9.5. The set of unitary matrices (i.e. matrices with the property that $U^\dagger U = \mathbb{1}$) forms a group, $U(n)$, called the **unitary group**. We will define this properly later.

Example 9.6. The set of all orthogonal matrices forms a group called the orthogonal group, $O(n)$. We will study it in detail much later.

Example 9.7. The set of unitary matrices with determinant 1 forms the special unitary group, $SU(n)$.
The set of orthogonal matrices with determinant 1 forms the special orthogonal group, $SO(n)$.

Definition 9.12 (Generating Set). Let G be a group and let S be a subset of G . We define the **group generated by S** to be the unique subgroup $\langle S \rangle$ of G so that $\langle S \rangle$ contains S , and if K is any other subgroup of G containing S we have $\langle S \rangle \subseteq K$. That is, $\langle S \rangle$ is the **smallest subgroup containing S** . We call S a **generating set** for $\langle S \rangle$. If $\langle S \rangle = G$ then we say S **generates G** . If $S = \{g_1, \dots, g_n\}$ we usually write $\langle S \rangle = \langle g_1, \dots, g_n \rangle$.

Remark. The generators of a Lie group are **not** the same concept. It is unlucky terminology.

Definition 9.13 (Cyclic Group). Let G be a group. Then we say G is **cyclic** if there is an element $g \in G$ so that $\langle g \rangle = G$. In other words, if G is generated by one element.

Remark. If $G = \langle g \rangle$ is cyclic then any element k of G can be written as $k = g^n$, with $n \in \mathbb{Z}$. **Reminder:** Here by “exponentiation” we really mean repeating the group operation n times! For an addition group, we actually have $k = g + \dots + g = ng$.

Lemma 9.1. The integers are cyclic. We have $\mathbb{Z} = \langle 1 \rangle$.

Proof. The proof is simply the statement that $n = 1n$ for any $n \in \mathbb{Z}$. □

9.2.2 Permutations

Example 9.8 (Permutation Group). Let $\mathbb{Z}_n = \{1, 2, \dots, n\}$ denote the positive integers up to n . We define the **permutation group** (also known as the **symmetric group**) to be the following set.

$$S_n = \{\sigma : \mathbb{Z}_n \rightarrow \mathbb{Z}_n : \sigma \text{ is bijective}\} \quad (9.2)$$

The symmetric group has order $|S_n| = n!$. The elements of S_n are functions which simply rearrange the order of the first n positive integers without removing any. The group operation is given by

$$f(\sigma, \tau) = \sigma \circ \tau, \quad \sigma \in S_n, \tau \in S_n$$

Which is just the composition of the two permutations. This satisfies the group axioms:

1. Associative: $\sigma \circ (\tau \circ \chi)(1, \dots, n) = \sigma(\tau(\chi(1, \dots, n))) = (\sigma \circ \tau) \circ \chi(1, \dots, n)$
2. Identity: Let $e(1, \dots, n) = (1, \dots, n)$. Then $\sigma \circ e = e \circ \sigma = \sigma$.
3. Invertibility: Since σ is bijective, we have $\sigma \circ \sigma^{-1} = \sigma^{-1} \circ \sigma = e$.

A permutation map is primarily defined as a function from \mathbb{Z}_n to \mathbb{Z}_n . But we also often use them to act on an ordered list of

symbols (v_1, \dots, v_n) to get a new list $\sigma(v_1, \dots, v_n) = (v_{\sigma(1)}, \dots, v_{\sigma(n)})$ which is the same list, just reordered.

Example 9.9 (Transposition). Consider the permutation which swaps the order of the first two inputs:

$$\sigma_{12}(v_1, v_2, v_3, \dots, v_n) = (v_2, v_1, v_3, \dots, v_n)$$

A permutation which does nothing but swap two symbols is called a **transposition**. Observe that if we applied σ_{12} twice, we would get,

$$\sigma_{12}^2(v_1, v_2, v_3, \dots, v_n) = \sigma_{12}(v_2, v_1, v_3, \dots, v_n) = (v_1, v_2, v_3, \dots, v_n)$$

So $\sigma_{12}^{-1} = \sigma_{12}$. In general, any transposition is its own inverse.

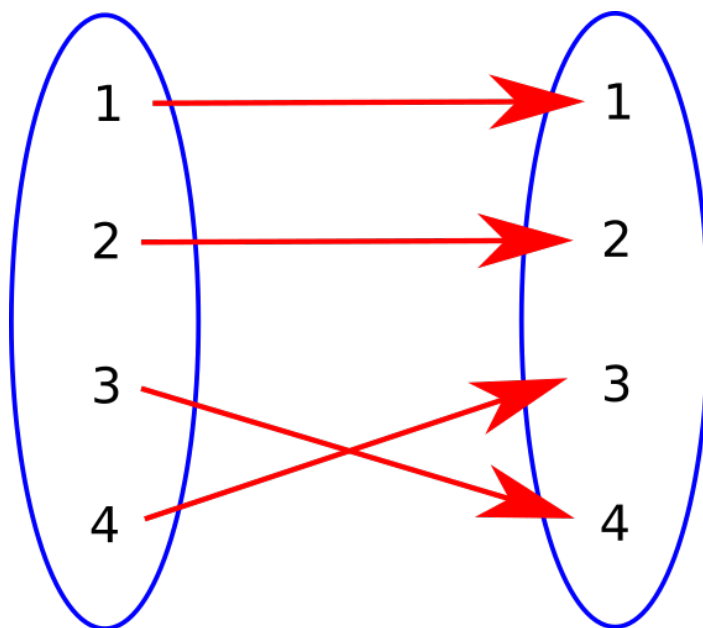


Figure 4: Diagram of an example transposition.

Definition 9.14. Let S_n denote the permutation group on $\{1, \dots, n\}$. We say that $\sigma \in S_n$ is an **even permutation** if it can be reproduced by an even number of transpositions. It is called an **odd permutation** if it can be reproduced by an odd number of transpositions.

The **parity** $\text{sgn}(\sigma)$ of a permutation is defined as follows.

$$\text{sgn}(\sigma) = \begin{cases} 1 & \text{if } \sigma \text{ is even} \\ -1 & \text{if } \sigma \text{ is odd} \end{cases}$$

Example 9.10. Any transposition is an odd permutation by definition.

Example 9.11. Let $n = 3$. The permutation taking $(1, 2, 3)$ to $(2, 3, 1)$ can be achieved by swapping 1 and 3, then swapping 2 and 3. It is therefore an even permutation.

Theorem 9.1 (Decomposition of Permutations). Let $\sigma \in S_n$ be a permutation. Then there exists a finite list $\{\tau_1, \dots, \tau_N\}$ of transpositions so that $\sigma = \tau_1 \circ \dots \circ \tau_N$.

In other words, if T_n is the set of all transpositions in S_n then $S_n = \langle T_n \rangle$.

Theorem 9.2. Parity satisfies the property that $\text{sgn}(\sigma \circ \tau) = \text{sgn}(\sigma) \text{sgn}(\tau)$.

Corollary 9.1. The parity of σ can also be written as,

$$\text{sgn}(\sigma) = \begin{cases} 1 & \text{if } \sigma \text{ is the product of an even number of transpositions} \\ -1 & \text{if } \sigma \text{ is the product of an odd number of transpositions} \end{cases}$$

9.2.3 Group Homomorphisms

Definition 9.15 (Group Homomorphism). Let G and H be groups. Let $\phi : G \rightarrow H$ be a function.

Denote the group operation of G by $f_G(g_1, g_2)$ and the group operation of H by $f_H(h_1, h_2)$. We say ϕ is a **group homomorphism** (or just a **homomorphism**) if the following equation holds,

$$\phi(f_G(g_1, g_2)) = f_H(\phi(g_1), \phi(g_2))$$

Definition 9.16 (Kernel and Image). Let $\phi : G \rightarrow H$ be a group homomorphism.

The image of ϕ is defined by $\text{image}(\phi) = \{\phi(x) : x \in G\}$. It is a subgroup of H .

The kernel of ϕ is defined by $\ker(\phi) = \{x \in G : \phi(x) = e\}$. It is a subgroup of G .

Remark. The kernel consists of anything getting mapped to the identity.

The image is the set of all outputs.

Theorem 9.3. Consider the group $G = \{-1, 1\}$, where the group operation is $f_G(a, b) = ab$. The map $\text{sgn} : S_n \rightarrow G$ is a group homomorphism.

Proof. As we saw earlier, $\text{sgn}(\sigma \circ \tau) = \text{sgn}(\sigma) \text{sgn}(\tau)$. □

Definition 9.17 (Alternating Group). The kernel of the map $\text{sgn} : S_n \rightarrow G$ is denoted $\ker(\text{sgn}) = A_n$, and is called the alternating group. It consists of all even permutations in S_n . It has order $|A_n| = n!/2$.

Definition 9.18 (Group Isomorphism). Let G and H be groups, and let $\phi : G \rightarrow H$ be a group homomorphism. If ϕ is bijective, we say it is a **group isomorphism**. Group isomorphism is an equivalence relation.

Lemma 9.2. Suppose $\phi : G \rightarrow H$ is an injective group homomorphism. Then ϕ is an isomorphism between G and $\text{image}(\phi)$.

Proof. Let $y \in \text{image}(\phi)$. Then by definition, there is an element x of G with $\phi(x) = y$. Therefore ϕ is surjective onto $\text{image}(\phi)$. This completes the proof. \square

Lemma 9.3. Any two cyclic groups G, H of order $|G| = |H|$ are isomorphic.

Proof. Recall that $G = \langle g \rangle$ and $H = \langle h \rangle$.

First we show that there is a unique homomorphism $\phi : G \rightarrow H$ so that $\phi(g) = h$. Suppose ϕ_1 and ϕ_2 are homomorphisms and that $\phi_1(g) = \phi_2(g) = h$. Then $\phi_1(g)^n = \phi_1(g^n)$ and $\phi_2(g)^n = \phi_2(g^n)$. So $\phi_1(g^n) = h^n = \phi_2(g^n)$ for all n . Since every element of G is of the form g^n , we see that $\phi_1(x) = \phi_2(x)$ for all $x \in G$. Thus ϕ is unique.

Now we show that it is surjective. Suppose $k \in H$. Then for some $|n| \leq |G| = |H|$, we have $k = h^n = \phi(g)^n = \phi(g^n)$.

Finally, we show it is injective. Suppose $\phi(x) = e$. Then $\phi(x)k = k$ for all $k \in H$. So for some n we have $\phi(x)h^n = h^n$, so $\phi(xg^n) = h^n$. Since $x \in G$ we have $x = g^q$ for some q , so $\phi(g^{q+n}) = h^n$. But $\phi(g^{q+n}) = h^{q+n}$. So we must have $q = a|G|$ for some $a \in \mathbb{Z}$. But $|G| = |H|$, so $g^{n+a|G|} = g^n$ and $h^{n+a|G|} = h^n$. This completes the proof. \square

Lemma 9.4. Let $G = \langle g \rangle$ be cyclic and finite. Then $g^{|G|} = e$. Furthermore, $g^{a+|G|k} = g^a$ for all $k \in \mathbb{Z}$.

Remark. This is why we call it a cyclic group.

Proof. Let ℓ be the smallest integer so that $g^\ell = e$. We can always write $k = \ell q + r$ for some remainder r , so $g^{\ell q + r} = g^{\ell q} g^r = e$. But $g^{\ell q} = (g^\ell)^q = e^q = e$. So $r = 0$. Therefore if $g^k = e$ we must have $k = \ell q$ for some q . We also have that for all $0 < k \leq \ell - 1$, $g^{-k} = g^a$ for some a . Since $g^\ell = e$, this means that $g^{\ell-k} = g^{-k}$. Since $0 \leq \ell - k \leq \ell - 1$ we have $a = \ell - k$. So $g^{-k} = g^{\ell-k}$.

Finally, observe that since $g^{-k} = g^{\ell-k}$ for all k , we have the following distinct elements of G : $\{e, g^1, g^2, \dots, g^{\ell-1}\}$. Thus $|G| = \ell$ and $g^{|G|} = e$ as required. \square

9.2.4 Quotient Groups and the First Isomorphism Theorem

Definition 9.19 (Coset of h). Let G be a group and let $h \in G$. Then the set $hG = \{hg : g \in G\}$ is called the right coset of h . The set $Gh = \{gh : g \in G\}$ is called the left coset of h .

Definition 9.20 (Coset of H). Let G be a group and let H be a subgroup of G . Then we define the left coset of H to be $GH = \{gh : g \in G, h \in H\}$ and the right coset to be $HG = \{hg : g \in G, h \in H\}$.

Remark. In general, we write $aHb = \{ahb : h \in H\}$. Similarly, $AHB = \{ahb : h \in H, a \in A, b \in B\}$.

Definition 9.21 (Normal Subgroup). Let G be a group and let H be a subgroup. Then we say H is **normal** if $g^{-1}Hg = H$ for all $g \in G$.

Theorem 9.4. The following are equivalent.

1. H is normal.
2. $g^{-1}Hg = H$ for all $g \in G$.
3. $gH = Hg$ for all $g \in G$.
4. If $x \in aH$ and $y \in bH$ then $xy \in (ab)H$.
5. For all $x, y \in G$ we have $(xH)(yH) = (xy)H$.
6. There exists a group K and a group homomorphism $\phi : G \rightarrow K$ with $\ker(\phi) = H$.

Remark. In other words, normal subgroups are just all the possible kernels of different maps out of G .

Definition 9.22 (Quotient Group). Let G be a group and let H be a normal subgroup of G . Then we define the quotient of G by H to be the collection of all cosets of H .

$$G/H = \{gH : g \in G\} \tag{9.3}$$

The set G/H of all cosets of H forms a group with the same operation as G , since $(aH)(bH) = (ab)H$.

Theorem 9.5 (First Isomorphism Theorem/Noether's First Theorem). Let $\phi : G \rightarrow K$ be a surjective group homomorphism. Then K is isomorphic to $G/\ker(\phi)$.

Proof. Let $g\ker(\phi) \in G/\ker(\phi)$. Then for any $gh \in g\ker(\phi)$ we have

$$\phi(gh) = \phi(g)\phi(h) = \phi(g)e = \phi(g)$$

Therefore, the map $\tilde{\phi} : G/\ker(\phi) \rightarrow K$, given by $\tilde{\phi}(g\ker(\phi)) = \phi(g)$, is well defined.

As we can see, $\tilde{\phi}(g\ker(\phi)) = e$ if and only if $\phi(g) = e$, which means $\tilde{\phi}$ is injective. This completes the proof \square

Corollary 9.2. $G/\ker(\phi)$ is isomorphic to $\text{image}(\phi)$.

Theorem 9.6 (Lagrange's Theorem). If H is a subgroup of G then $|G|/|H|$ is an integer. In other words, the order of H divides the order of G .

Proof. Consider the collection of left-cosets of H .

Suppose $x = gh_1 \in gH$. Then $x \in kH$ if and only if $x = kh_2$ for some $h_2 \in H$, which means $gh_1 = kh_2$, or in other words, $gh_1h_2^{-1} = k$.

Since $h_1h_2^{-1} \in H$, it follows that $gH = kH$. Thus, two left cosets must be either equal or entirely disjoint.

Additionally, every left coset of H is the same size since $|gH| = |H|$ for all g . Since there are an integer number of left cosets, it follows that $n|H| = |G|$, where n is the number of left cosets of H .

The number n is called the **index** of H in G , and we often write it as $n = [G : H]$ so that $|H|[G : H] = |G|$. \square

9.2.5 Group Actions and Cayley's Theorem

Definition 9.23 (Group Action). Let S be a set and let G be a group. Then a (left) **group action** is a function $A : G \times S \rightarrow S$, with the following properties:

1. Compatibility: $A(gh, s) = A(g, A(h, s))$.
2. Identity: $A(e, s) = s$.

Notationally, we often write $A(g, s) = g \cdot s$ as long as this can't be confused with some other operation.

Definition 9.24 (Transitive Action). A group action $A : G \times S \rightarrow S$ is called transitive, if for all $x, y \in S$ there exists some g with $A(g, x) = y$.

Theorem 9.7 (Cayley's Theorem). Any finite group G is isomorphic to a subgroup H of S_n , for some $n \in \mathbb{N}$.

Proof. Let $k = |G|$ be the order of G . Consider the left group-action of G on itself, $A : G \times G \rightarrow G$ given by $A(g, h) = gh$.

We can see that the action is invertible, since $A(g^{-1}, g) = A(g, g^{-1}) = e$.

We can also see that A is transitive, since if $x, y \in G$ we have $x^{-1}y \in G$, so we can simply take $A(x, x^{-1}y) = y$.

Therefore, for each $g \in G$ we have an invertible function $A_g : G \rightarrow G$ defined by $A_g(h) = A(g, h)$.

Since $|G| = k$, each function A_g simply permutes the k elements of G around. Therefore we have a mapping $\tilde{A} : G \rightarrow S_k$ given

by $\tilde{A}(g) = A_g$.

We also have

$$\tilde{A}(g)\tilde{A}(h)(x) = ghx = \tilde{A}(gh)(x)$$

Therefore \tilde{A} is a homomorphism, and $\text{image}(\tilde{A})$ is a subgroup of S_k .

Since $\tilde{A}(g)(h) = gh = e$ iff $g = h^{-1}$, it also follows that \tilde{A} is injective, and is hence an isomorphism between G and $\text{image}(\tilde{A})$ as required. \square

9.3 Rings

We will only review a few definitions for rings.

Definition 9.25 (Ring). Let R be a commutative group. That is, there exists an operation $+$: $R \times R \rightarrow R$ which is associative, commutative, invertible, and has an identity 0 . Suppose there is another operation, \cdot : $R \times R \rightarrow R$, which satisfies the following properties,

1. Associative: $a \cdot (b \cdot c) = (a \cdot b) \cdot c$
2. Distributive: $a \cdot (b + c) = a \cdot b + a \cdot c$ and $(b + c) \cdot a = b \cdot a + c \cdot a$
3. Identity: There exists $1 \in R$ so that $1 \cdot a = a \cdot 1 = a$ for all a .

Then we say that the data $(R, +, \cdot)$ forms a **ring**. For short, we just call R a ring and we usually write $a \cdot b = ab$.

Definition 9.26 (Ideal). A **ring ideal** is a subset $I \subseteq R$ with the following properties.

1. I is a subgroup of R with respect to addition.
2. For all $a \in R$, $b \in I$, we have $ab \in I$

Definition 9.27 (Generating Set of Ideal). The **ideal generated by** a subset S is defined to be the smallest ideal containing S . We denote it (S) . For $S = \{a_1, \dots, a_n\}$ we let $(S) = (a_1, \dots, a_n)$.

Example 9.12. $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ are all rings.

Example 9.13. Let R be a ring. Let $R[x]$ denote the set of all polynomials whose coefficients are elements of R . Then $R[x]$ is also a ring.

Definition 9.28 (Coset of Ideal). Let I be an ideal in R . Then we define the cosets of I to be the sets of the form $a + I$, $a \in R$. This is exactly the same as the construction for subgroups.

Definition 9.29 (Quotient Ring). Let I be an ideal in R . Then we define R/I to be the set of all cosets of I . This forms a ring.

Example 9.14. Let $R = \mathbb{R}$ and consider the set of all real polynomials $\mathbb{R}[x]$. Observe that $x^2 + 1$ has no roots in \mathbb{R} , which means it can not be factored. Let $(x^2 + 1)$ be the ideal generated by this polynomial.

The quotient, $\mathbb{R}[x]/(x^2 + 1)$, consists of all objects of the form $f(x) + I$, and two polynomials in $\mathbb{R}[x]/(x^2 + 1)$ are said to be equivalent if they differ by a polynomial of the form $h(x)(x^2 + 1)$.

Let $\phi : \mathbb{R}[x]/(x^2 + 1) \rightarrow \mathbb{C}$ be defined as follows. $\phi([f]) = f(i)$. Observe that since $i^2 - 1 = 0$, this does not depend on the choice of representative for $[f] = f + I$. Additionally,

$$\phi([f] + [g]) = f(i) + g(i) = \phi([f]) + \phi([g])$$

and

$$\phi([fg]) = f(i)g(i) = \phi([f])\phi([g])$$

So ϕ is a ring homomorphism.

Suppose f is a degree k polynomial. Then it is of the form $f(x) = a_k x^k + \dots + a_2 x^2 + a_1 x + a_0$. Then we can rearrange this as follows.

$$\begin{aligned} f(x) + I &= a_k x^{k-2}(x^2 - 1 + 1) + \dots + a_2(x^2 - 1 + 1) + a_1 x + a_0 + I \\ &= (a_k x^{k-2} + \dots + a_2)(x^2 + 1) - (a_k x^{k-2} + \dots + a_2) + a_1 x + a_0 + I \\ &= -(a_k x^{k-2} + \dots + a_2) + a_1 x + a_0 + I \\ &= -(a_k x^{k-4}(x^2 - 1 + 1) + \dots + a_4(x^2 - 1 + 1)) + (a_1 - a_3)x + (a_0 - a_2) + I \\ &= \dots \\ &= (a_1 \pm a_3 \pm \dots \pm a_p)x + (a_0 \pm a_2 \pm \dots \pm a_q) + I \end{aligned}$$

Where p is the largest even number less than or equal to k , and q is the largest odd number less than or equal to k . Therefore we see that for any $f + I \in \mathbb{R}[x]/(x^2 + 1)$ we can reduce f to a first order polynomial. Since the coefficients a_i can be chosen however we want, this means that for any $a, b \in \mathbb{R}$ there is a polynomial f so that $\phi([f]) = a + bi$. So ϕ is surjective.

Finally, let $\phi^{-1}(a + bi) = a + bx$. Then since any $[f]$ has a first order representative, ϕ^{-1} is surjective.

Therefore ϕ is a ring isomorphism.

Thus we have shown that \mathbb{C} is isomorphic to $\mathbb{R}[x]/(x^2 + 1)$.

Remark. This is one way of defining the complex numbers. In this course we will study another definition in depth, called the Cayley-Dickson construction, which defines \mathbb{C} as a vector algebra over \mathbb{R} .

Definition 9.30 (Field (alternative definition)). Let R be a ring. Then if for all $0 \neq a \in R$ there exists $a^{-1} \in R$ so that $a \cdot a^{-1} = a^{-1} \cdot a = 1$, we say R is a **field**.

10 Appendix II: Linear Algebra

10.1 Vector Spaces

10.1.1 Basic Definitions

This content can all be found in standard books [2, 12]

Definition 10.1 (Vector Space). A vector space is defined by the data of a set V , along with a field \mathbb{F} called the set of scalars. The set V is equipped with a notion of addition. That is, for any two vectors v and w in V , we can add them to produce a third vector $v + w$.

In addition to this notion of vector addition, we also have a notion of scalar multiplication. That is, the vectors $v \in V$ can be scaled by an element a of \mathbb{F} to produce a new vector $av \in V$.

These two operations must follow the vector space axioms.

1. Commutativity: if $v, w \in V$ then $v + w = w + v$
2. Associativity: if $a, b \in \mathbb{F}$ and $v \in V$, we have $a(bv) = (ab)v$
3. Distributivity: if $a \in \mathbb{F}$ and $v, w \in V$ we have $a(v + w) = av + aw$.

Remark. If we replace the field \mathbb{F} in the above definition with an arbitrary ring R , then we get something called a **module** instead of a vector space.

Definition 10.2 (Linear Subspace). Let V be a vector space. A linear subspace is a subset $U \subseteq V$ which is also a vector space, with the same underlying field, scalar multiplication, and addition operations.

Example 10.1. There are a great many examples of vector spaces. I will list just a few kinds of vector spaces we will encounter.

1. Vector spaces of the form \mathbb{F}^n , such as \mathbb{R}^n or \mathbb{C}^n . These are formed by taking the Cartesian product of a field with itself n times. The addition operation is defined by $(a_1, \dots, a_n) + (b_1, \dots, b_n) = (a_1 + b_1, \dots, a_n + b_n)$.
2. Vector spaces of functions. For example, $C_0(\mathbb{R})$ denotes the set of continuous functions from \mathbb{R} to \mathbb{R} . This set is a vector space since we know that two continuous functions can be added to produce a third continuous function. They can also be scaled by real numbers, and the result will still be continuous.
3. Vector spaces formed by combining two vector spaces in some nice way. For example, if V and W are vector spaces we could form the direct sum of V and W . The direct sum, denoted $V \oplus W$, is the set of all linear combinations of elements of V and W .

Definition 10.3 (Span). Let V be a vector space over \mathbb{F} . Let v_1, \dots, v_n be a collection of vectors in V . We define

$$\text{span}(\{v_1, \dots, v_n\}) = \left\{ \sum_{i=1}^n a_i v_i : a_i \in \mathbb{F} \right\} \quad (10.1)$$

Definition 10.4 (Direct Sum). Let V and W be vector spaces. The direct sum of V and W is the vector space

$$V \oplus W = \{v + w : v \in V, w \in W\} = \text{span}(V \cup W) \quad (10.2)$$

For a list of k vector spaces V_1, \dots, V_k we write,

$$V_1 \oplus \dots \oplus V_k = \bigoplus_{i=1}^k V_i \quad (10.3)$$

Example 10.2. Let V and W be vector spaces. Then their Cartesian product $V \times W$ is a vector space.

Proof. Consider $a \in \mathbb{F}$, $(v_1, w_1) \in V \times W$ and $(v_2, w_2) \in V \times W$. Then we define addition and scalar multiplication according to the rule $(v_1, w_1) + a(v_2, w_2) = (v_1 + av_2, w_1 + aw_2)$. As we can see, this is an element of $V \times W$. We can verify each of the vector space axioms as well:

1. $(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2) = (v_2 + v_1, w_2 + w_1) = (v_2, w_2) + (v_1, w_1)$
2. $a(b(v, w)) = a(bv, bw) = ((ab)v, (ab)w) = (ab)(v, w)$
- 3.

$$\begin{aligned} a((v_1, w_1) + (v_2, w_2)) &= a(v_1 + v_2, w_1 + w_2) \\ &= (av_1 + av_2, aw_1 + aw_2) \\ &= (av_1, aw_1) + (av_2, aw_2) \\ &= a(v_1, w_1) + a(v_2, w_2) \end{aligned}$$

□

An important property of vectors is linear independence.

Definition 10.5 (Linearly Independent). Let V be a vector space over \mathbb{F} . Then a collection of n nonzero vectors v_1, \dots, v_n are called linearly independent if the following property holds.

Whenever $a_1, \dots, a_n \in \mathbb{F}$ then $a_1 v_1 + \dots + a_n v_n = 0$ if and only if $a_1 = \dots = a_n = 0$.

Example 10.3. Consider $V = \mathbb{R}^3$. Then the vectors $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, and $e_3 = (0, 0, 1)$ are linearly independent.

Proof. Suppose $a_1e_1 + a_2e_2 + a_3e_3 = 0$. Observe that $a_1e_1 + a_2e_2 + a_3e_3 = (a_1, a_2, a_3)$. Since $(a_1, a_2, a_3) = (0, 0, 0)$, we must have $a_1 = 0, a_2 = 0, a_3 = 0$. Therefore e_1, e_2, e_3 must be linearly independent. \square

Definition 10.6 (Dimension). Let V be a vector space over \mathbb{F} . The dimension $\dim V$ is the largest possible number n with the property that there exists a set of linearly independent vectors of size n .

Definition 10.7 (Basis). Let V be a vector space and let B be a set of linearly independent vectors such that $V = \text{span}(B)$. We say that B is a basis for V .

Theorem 10.1. Let V be a finite dimensional vector space. Then every basis for V has $\dim V$ elements. Furthermore, if B is a linearly independent subset of V of size $\dim V$, then B is a basis for V .

Theorem 10.2. Let V be a vector space and let W be a subspace of V . Then if $\dim W = \dim V$ we have $W = V$.

Proof. Let $\{e_1, \dots, e_n\}$ be a basis for W . Then if V is n -dimensional, e_1, \dots, e_n must also be a basis for V . Therefore $V = \text{span}\{e_1, \dots, e_n\} = W$. \square

Theorem 10.3. Let V be a vector space and let U, W be subspaces of V . Then their intersection $U \cap W$ is a vector space, and $U \cap W = \{0\}$ if and only if U and W are linearly independent subspaces.

Remark. The union of two vector subspaces is **not necessarily** a vector space.

A very special kind of vector space merits its own definition. Vector algebras are vector spaces with a notion of vector multiplication.

Definition 10.8 (Algebra). An **algebra** is a vector space equipped with a notion of vector multiplication. That is, for any two vectors $v, w \in V$ we can multiply them to produce a third vector, $v \times w$. This operation is not required to be commutative, but must satisfy the following axioms:

1. Distributivity: for all $u, v, w \in V$ we have $u \times (v + w) = u \times v + u \times w$
2. Scaling rule: for all $a \in \mathbb{F}$, and $u, v \in V$, we have $a(v \times w) = (av) \times w = v \times (aw)$

If in addition the algebra satisfies $u \times (v \times w) = (u \times v) \times w$, we say A is an **associative algebra**.

Remark. We call this type of vector space an “algebra” because it has all of the tools required to do algebra: multiplication and addition.

Definition 10.9. Let V be an algebra and let W be a vector subspace of V . Then W is said to be a subalgebra of V if for all $a, b \in W$ we have $ab \in W$ and $ba \in W$.

Example 10.4 (Examples of Algebras). The following are a few examples of algebras.

1. Matrix algebras. The set $M_{nn}(\mathbb{R})$ of n by n matrices forms an algebra, where the “vector multiplication” is really just matrix multiplication. Note that this is not commutative! Many algebras are isomorphic to matrix algebras (this isomorphism is called a faithful representation).
2. The set of continuous functions, $C_0(\mathbb{R})$, has the property that if f and g are continuous, then their composition, $f \circ g$, is continuous. Therefore $C_0(\mathbb{R})$ is a vector algebra, where the “multiplication” is composition. Note that this is not commutative! $f \circ g \neq g \circ f$.
3. The spaces \mathbb{R}^1 , \mathbb{R}^3 , and \mathbb{R}^7 all have a special operation defined on them called the cross product. This is denoted $v \times w$, and it also has the property that $v \times w = -w \times v$. Therefore, \mathbb{R} , \mathbb{R}^3 and \mathbb{R}^7 are algebras when equipped with the cross product.

Example 10.5 (Non-example of an Algebra). The set $GL_n(\mathbb{R})$ of invertible $n \times n$ matrices is **not** an algebra since not every linear combination of invertible matrices is invertible. It is closed under matrix multiplication but not addition. For instance, if $M \in GL_n(\mathbb{R})$ then $M - M = 0 \notin GL_n(\mathbb{R})$ since the zero matrix is not invertible.

10.1.2 Maps Between Vector Spaces

The vast majority of the time we are only going to be interested in maps between vector spaces which don't interfere with the laws of arithmetic. That is, you should be able to scalar multiply or add vectors before applying a map, and get the same thing if you were to apply the map and then add/scale the result.

Maps with this property are called homomorphisms, or “structure preserving maps”. In the case where the operations we would like to preserve are vector space operations, these maps are what we call “linear”.

Definition 10.10 (Linear Map). Let V and W be vector spaces over \mathbb{F} . Then a map $T : V \rightarrow W$ is called linear if for all $a \in \mathbb{F}$ and $u, v \in V$, we have $T(au + v) = aT(u) + T(v)$. If $W = V$ we say T is a **linear operator**.

Definition 10.11 (Space of Linear Maps). The set of all linear maps from V to W is denoted $L(V, W)$. It forms a vector space. For simplicity we often write $L(V, V) = L(V)$ when the domain and range are the same.

Remark. Why are these maps called linear? The answer is that if $L \subseteq V$ is a line through the origin in V (that is, a one-dimensional subspace of V), then if you apply a linear map T to every point in L , the result $T(L)$ is also a line through the

origin in W . Therefore, linear maps are exactly the functions which send lines through the origin to lines through the origin.

Definition 10.12 (Isomorphism). Let $T : V \rightarrow W$ be a linear map. If T is also bijective, then T is an **isomorphism**.

Remark. An isomorphism is a way to make two vector spaces equivalent. If one completes some calculation in a vector space V , and V is isomorphic to some other space W , then the same calculation can easily be translated to W .

Definition 10.13. Let V be a vector space. An isomorphism from V to itself is called a **endomorphism**. The set of endomorphisms is a vector space called $\text{End}(V)$, and it is a subspace of $L(V)$.

Definition 10.14 (Fundamental Subspaces). Let V, W be vector spaces and let $T : V \rightarrow W$ be a linear map. The fundamental vector subspaces associated to T are the following sets.

1. The image of T (also known as the range). Defined as

$$\text{image}(T) = \{T(v) : v \in V\}$$

This is a subspace of W . The **rank** of T is defined by $\text{rank}(T) = \dim(\text{image}(T))$. This is an *invariant* of T .

2. The kernel of T (also known as the null-space). Defined as

$$\ker(T) = \{v \in V : T(v) = 0\}$$

This is a subspace of V . It contains all vectors which are mapped to zero by T . The **nullity** of T is the dimension of $\ker T$.

Theorem 10.4 (Rank-Nullity Theorem). Let $T : V \rightarrow W$ be a linear map. Then

$$\dim \text{image } T = \dim W - \dim \ker T \quad (10.4)$$

Theorem 10.5. Let $T : V \rightarrow W$ be a linear map. Then T is injective if and only if $\ker T = \{0\}$.

Proof. First suppose T is injective. Now let $v_1, v_2 \in \ker(T)$. This means that $T(v_1) = T(v_2) = 0$.

But then since T is injective, $v_1 = v_2$. Therefore every vector in $\ker(T)$ is equal to every other vector in $\ker(T)$.

The only way for this to be possible is for $\ker(T)$ to contain exactly one vector.

Since $T(0) = 0$, this vector must be the zero vector. So $\ker(T) = \{0\}$.

Now suppose $\ker(T) = \{0\}$. Then let $v_1, v_2 \in V$ and suppose $T(v_1) = T(v_2)$.

So $T(v_1 - v_2) = 0$, which means $v_1 - v_2 \in \ker(T)$.

But the only thing in $\ker(T)$ is the zero vector, so $v_1 = v_2$. Therefore T is injective. □

Theorem 10.6. Let $T : V \rightarrow W$ be an injective linear map. Then T is an isomorphism if $\dim W = \dim V$.

Proof. Suppose $\dim W = \dim V$. Let $B = \{v_1, \dots, v_n\}$ be a basis for V . Then let $T(B) = \{T(v_1), \dots, T(v_n)\}$. Observe that since T is injective, $T(v_i) = T(v_j)$ if and only if $v_i = v_j$.

Now, we would like to show that $T(B)$ is a basis for W . We can do this by showing that it is linearly independent.

Consider any set $a_1, \dots, a_n \in \mathbb{F}$. Then $a_1T(v_1) + \dots + a_nT(v_n) = 0$ if and only if $T(a_1v_1 + \dots + a_nv_n) = 0$.

But since T is injective, this means

$$a_1v_1 + \dots + a_nv_n = 0$$

Since v_1, \dots, v_n are linearly independent we conclude that

$$a_1 = \dots = a_n = 0$$

Therefore $T(v_1), \dots, T(v_n)$ are linearly independent. Since $\dim W = \dim V$, we conclude that $T(B)$ is a basis for W .

Finally, since $T(B)$ is a basis for W , that means for any $w \in W$ we have

$$w = \sum_{i=1}^n w_i T(v_i) = T\left(\sum_{i=1}^n w_i v_i\right)$$

So for all $w \in W$ there exists v so that $T(v) = w$. So T is surjective, and hence T is an isomorphism. □

Corollary 10.1. Let $T : V \rightarrow W$ be a linear map and let B be any basis for V . Then if $T(B)$ is a basis for W , it follows that T is surjective.

Proof. Exercise. Hint: See the proof of the previous theorem. □

Theorem 10.7 (Extension by Linearity). Let V, W be two vector spaces and let $B = \{v_1, \dots, v_n\}$ be a basis for V . Then if we specify some desired outputs $w_1, \dots, w_n \in W$, there is a **unique** linear map $T : V \rightarrow W$ which satisfies $T(v_i) = w_i$ for each $i = 1, \dots, n$. Additionally, $\text{image}(T) = \text{span}\{w_1, \dots, w_n\}$.

Proof. If we require that T is linear, and that $T(v_i) = w_i$ for each $i = 1, \dots, n$ then for any $v \in V$ we get

$$\begin{aligned} T(v) &= T\left(\sum_{i=1}^n a_i v_i\right) \\ &= \sum_{i=1}^n a_i T(v_i) \\ &= \sum_{i=1}^n a_i w_i \end{aligned}$$

Therefore we only need to know $T(v_i)$ for each $i = 1, \dots, n$ in order to **completely** specify $T(v)$ for all $v \in V$. □

When we do calculations, we don't always want to have an abstract definition for our linear map. For numerical calculations it is much more convenient to put the linear map in matrix form. To do this, we need a basis.

Definition 10.15. The set of all $n \times n$ matrices with entries in a field \mathbb{F} will be denoted $M_{nn}(\mathbb{F})$. Any matrix is also a linear map from \mathbb{F}^n to \mathbb{F}^n .

Definition 10.16 (Matrix of a Linear Map). Let V, W be finite dimensional vector spaces and let $B = \{e_1, \dots, e_n\}$ be a basis for V and $C = \{f_1, \dots, f_n\}$ be a basis for W . Let $T : V \rightarrow W$ be a linear map. Then the matrix $[T]_{CB}$ is defined by the following formula

$$[T]_{CB}[v]_B = [T(v)]_C \quad \text{for all } v \in V$$

Where $[v]_B = (v_1, \dots, v_n)$ is the $n \times 1$ matrix of coefficients defined by $v = \sum_{i=1}^n v_i e_i$. In other words, we can write

$$[T]_{CB} = \sum_{i=1}^n \sum_{j=1}^n T_{ij} v_i f_j$$

Where T_{ij} are the entries of the matrix $[T]_{CB}$. In an upcoming section, we will employ Einstein notation to make things like this less tedious to read and write.

Theorem 10.8. Let $T : V \rightarrow V$ be a linear map, and let $[T]_{BB} \in M_{nn}(\mathbb{F})$ be the matrix for T in some basis. The following are equivalent.

1. T is invertible
2. The columns of $[T]$ are linearly independent
3. The only solution to the equation $T(v) = 0$ is $v = 0$

A very useful computational tool when dealing with finite dimensional linear maps is the determinant, which allows us to easily determine whether a given matrix is invertible.

Definition 10.17 (Determinant). Let M be an $n \times n$ matrix with entries M_{ij} , $i, j = 1, \dots, n$. The determinant of M is given by the following formula

$$\det M = \sum_{\sigma \in S_n} \left(\text{sgn } \sigma \prod_{i=1}^n M_{\sigma(i)i} \right)$$

This formula is called the **Laplace expansion** of the determinant.

Theorem 10.9. The function $\det : M_{nn}(\mathbb{F}) \rightarrow \mathbb{F}$ is the unique function satisfying the following properties:

1. For all $A, B \in M_{nn}(\mathbb{F})$ we have $\det(AB) = \det A \det B$.

If A is invertible we therefore have $\det(A^{-1}) = (\det A)^{-1}$

2. Let $A = [a_1, a_2, \dots, a_n]$ be the matrix formed from the columns $a_1, \dots, a_n \in \mathbb{F}^n$. Then

$$\det[\dots, a_i, \dots, a_j, \dots] = -\det[\dots, a_j, \dots, a_i, \dots]$$

That is, swapping two columns introduces a minus sign to the result.

3. For any $\lambda \in \mathbb{F}, v \in \mathbb{F}^n, w \in \mathbb{F}^n$ we have

$$\det[\dots, \lambda v + w, \dots] = \lambda^n \det[\dots, v, \dots] + \det[\dots, w, \dots]$$

That is, \det is homogenous of degree n in each column.

Proof. The first point that $\det(AB) = \det A \det B$ is a bit difficult to prove. However, the other two points follow straightforwardly from the definition of the determinant above. We will not show that the determinant is the unique map satisfying these properties. □

Theorem 10.10. $\det A^T = \det A$.

Theorem 10.11. Let \mathbb{F} be a field of characteristic $\text{char}(\mathbb{F}) \neq 2$. A matrix $A \in M_{nn}(\mathbb{F})$ is invertible if and only if $\det A \neq 0$.

Theorem 10.12. Let V be a vector space and let $T : V \rightarrow V$ be a linear map. Let B and C be two bases for V . Then $\det[T]_B = \det[T]_C$. We therefore simply write $\det T$ for the determinant of a linear map, since it does not matter what bases we use to compute it. That is, $\det T$ is an *invariant* of T .

Proof. Recall that there exists an invertible matrix P_{BC} so that $[T]_B = P_{BC}[T]_C P_{BC}^{-1}$. Therefore

$$\begin{aligned} \det[T]_B &= \det(P_{BC}[T]_C P_{BC}^{-1}) \\ &= \det(P_{BC}) \det[T]_C \det(P_{BC}^{-1}) \\ &= \frac{\det(P_{BC})}{\det(P_{BC})} \det[T]_C \\ &= \det[T]_C \end{aligned}$$

as required. □

10.1.3 Inner Products and Eigenvectors

Definition 10.18 (Eigenvector). Let V be a vector space over \mathbb{F} and let $T : V \rightarrow V$ be a linear map. Then a vector $v \in V$ is said to be an **eigenvector** of T if $T(v) = \lambda v$ for some $\lambda \in \mathbb{F}$. The value λ is said to be an **eigenvalue** of T .

Remark. Observe that $T(v) = \lambda v$ iff $T(v) - \lambda v = 0$ iff $(T - \lambda \mathbb{1})v = 0$. This equation only has a nonzero solution if $\det(T - \lambda \mathbb{1}) = 0$. The expression $\chi(\lambda) = \det(T - \lambda \mathbb{1})$ is a polynomial of degree n , and is called the **characteristic polynomial**.

Definition 10.19 (Characteristic Polynomial). The **characteristic polynomial** of T is $\chi(\lambda) = \det(T - \lambda \mathbb{1})$. It is an *invariant* of T .

Suppose we are working over $\mathbb{F} = \mathbb{C}$. Then there are guaranteed to be n roots of $\chi(\lambda)$ up to multiplicity.

Definition 10.20 (Algebraic Multiplicity). Let $\mathbb{F} = \mathbb{C}$, let $T : V \rightarrow V$, $\dim V = n$, and let $\chi(\lambda)$ be the characteristic polynomial of T . Let $\lambda_1, \dots, \lambda_k$ be the roots of χ . Then there are some integers $a_i, i = 1, \dots, k$ so that $\chi(\lambda) = \prod_{i=1}^k (\lambda - \lambda_i)^{a_i}$ and $\sum_{i=1}^k a_i = n$. The integer a_i is called the **algebraic multiplicity** of the eigenvalue λ_i .

For a given eigenvalue λ_i of T , the solutions to $T(v) = \lambda_i v$ form a vector space of dimension $\dim \ker(T - \lambda_i \mathbb{1})$.

Definition 10.21 (Geometric Multiplicity). Let λ_i be an eigenvalue of T . Then $g_i = \dim \ker(T - \lambda_i \mathbb{1})$ is called the **geometric multiplicity** of λ_i .

In practice, we write $T - \lambda_i \mathbb{1}$ in some basis and solve the resulting system of equations. There are some useful theoretical results that help us bound the number of linearly independent sets of solutions.

Theorem 10.13. Let T be a linear map and let $\lambda_i, g_i, i = 1, \dots, k$ be the eigenvalues and their algebraic multiplicities. Then $1 \leq \dim \ker(T - \lambda_i \mathbb{1}) \leq g_i$.

Definition 10.22 (Inner Product). Let V be a complex vector space. Then an **inner product** on V is a bilinear function $g : V \times V \rightarrow \mathbb{C}$ with the following properties

1. $g(v, v) \geq 0$, with $g(v, v) = 0$ iff $v = 0$
2. $g(v, w) = \overline{g(w, v)}$

Definition 10.23 (Matrix of Inner Product). Let $\{e_1, \dots, e_n\}$ be a basis for V . Then set $g_{ij} = g(e_i, e_j)$. The matrix whose entries are g_{ij} is denoted $[g]_B$. We have $\langle v, w \rangle = \overline{[v]_B^T} [g]_B [w]_B$ for any basis B . Additionally, if $P_{B\tilde{B}}$ is a change of basis matrix, then $[g]_{\tilde{B}} = P_{B\tilde{B}}^T [g]_B P_{B\tilde{B}}$.

Definition 10.24 (Orthonormal Basis). Let $B = \{e_1, \dots, e_n\}$ be a basis for V and suppose $g(e_i, e_j) = \delta_{ij}$. Then B is said to be an **orthonormal basis**.

Theorem 10.14 (Gram-Schmidt Procedure). Let g be an inner product on V . Then there exists an orthonormal basis for V .

Proof. Let $B = \{e_1, \dots, e_n\}$ be any basis for V . Then set,

$$\begin{aligned} b_1 &= \frac{e_1}{\sqrt{g(e_1, e_1)}} \\ b_2 &= \frac{e_2}{\sqrt{g(e_2, e_2)}} - g(e_1, e_2)b_1 \\ &\vdots \\ b_k &= \frac{e_k}{\sqrt{g(e_k, e_k)}} - \sum_{i=1}^k g(e_i, e_k)b_i \end{aligned}$$

The vectors b_1, \dots, b_n are linearly independent and orthonormal and are a basis for V . The remainder of the proof is left as an exercise. \square

Definition 10.25 (Hermitian Transpose). Let $T : \mathbb{C}^n \rightarrow \mathbb{C}^n$ be a linear map and let $\langle v, w \rangle$ denote the standard complex inner product. Then the **Hermitian transpose** T^\dagger is the map defined by $\langle T(v), w \rangle = \langle v, T^\dagger(W) \rangle$ for all $v, w \in \mathbb{C}^n$.

Definition 10.26 (Hermitian Adjoint). In general, if V is any complex vector space with a complex inner product, we can define the **Hermitian adjoint** by the formula $\langle T(v), w \rangle = \langle v, T^\dagger(W) \rangle$. Note that $[T^\dagger]_{BB}$ is not always the same as $\overline{[T]_{BB}}^T$.

Definition 10.27 (Normal Operator). Let V be a complex vector space. A **normal operator** $A : V \rightarrow V$ is a linear map with the property that $AA^\dagger = A^\dagger A$.

Definition 10.28 (Symmetric Matrix). A **symmetric matrix** A is a real or complex matrix with the property that $A^T = A$. Symmetric matrices are normal.

Example 10.6. Consider $\mathbb{F} = \mathbb{R}$ with the standard inner product $\langle v, w \rangle = v^T w$. Let $A \in M_{nn}(\mathbb{R})$ and suppose $A^T A = \mathbb{1}$. Then the columns of A are all orthonormal with respect to the standard inner product on \mathbb{R}^n . We therefore say that A is an **orthogonal matrix** if $A^T A = \mathbb{1}$. Orthogonal matrices are normal and invertible.

Definition 10.29 (Unitary Operator). A linear map $T : \mathbb{C}^n \rightarrow \mathbb{C}^n$ is said to be **unitary** if $A^\dagger A = \mathbb{I}$. Unitary operators are normal and invertible.

Definition 10.30 (Hermitian Matrix). An operator $T : \mathbb{C}^n \rightarrow \mathbb{C}^n$ is called **Hermitian** if $T = T^\dagger$. Hermitian operators are normal.

Definition 10.31 (Diagonalizable). Let V be a vector space. A linear map $T : V \rightarrow V$ is said to be diagonalizable if there exists a basis B where $[T]_{BB}$ is diagonal.

Theorem 10.15. Let $T : V \rightarrow V$ be a linear map. Then T is diagonalizable if and only if the sum of all of the geometric multiplicities of T is n . That is, if $\sum_{i=1}^k g_i = n$.

Definition 10.32 (Orthogonally Diagonalizable). A linear map $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is said to be orthogonally diagonalizable if there exists an orthogonal basis B so that $[T]_{BB}$ is diagonal. Equivalently, if for any basis \tilde{B} there exists a real orthogonal matrix P so that $P^T [T]_{\tilde{B}\tilde{B}} P = [T]_{BB}$.

Remark. Later, we will prove that any symmetric matrix is orthogonally diagonalizable (see Theorem 2.5)

Definition 10.33 (Unitarily Diagonalizable). A linear map $T : \mathbb{C}^n \rightarrow \mathbb{C}^n$ is said to be unitarily diagonalizable if $[T]_{BB}$ is diagonal in some basis B , and if for all bases \tilde{B} there exists a unitary matrix U so that $U^\dagger [T]_{\tilde{B}\tilde{B}} U = [T]_{BB}$.

Theorem 10.16 (Spectral Theorem for Normal Matrices). Let V be a vector space over \mathbb{C} and let $\langle v, w \rangle$ be a complex inner product on V . Suppose $T : V \rightarrow V$ is normal with respect to this inner product. Then there exists an orthonormal basis $A = \{a_1, \dots, a_n\}$ for V consisting entirely of eigenvectors for T . Furthermore, T is unitarily diagonalizable, is diagonal in A , and the matrix $[T]_{AA}$ has all the eigenvalues of T on the diagonal.

$$[T]_{AA} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \quad (10.5)$$

Definition 10.34 (Jordan Canonical Form). Let V be a vector space over \mathbb{C} and let $\dim V = n$. Let $T : V \rightarrow V$ be **any** linear operator with eigenvalues $\lambda_1, \dots, \lambda_k$. Let a_i be the algebraic multiplicity of λ_i and let g_i be the geometric multiplicity. Then

there exists a basis B for V in which $[T]_{BB}$ is in the **Jordan normal form**.

$$[T]_{BB} = \begin{bmatrix} J_1 & O & \cdots & O \\ O & J_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ O & \cdots & \cdots & J_p \end{bmatrix} \quad (10.6)$$

Where O is the zero matrix, and for each $q = 1, \dots, p$ the **Jordan block** J_q is given by

$$J_q = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \ddots & 0 \\ \vdots & 0 & \lambda_i & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & \lambda_i \end{bmatrix} \quad (10.7)$$

for some λ_i . The number of times an eigenvalue λ_i appears in $[T]_A$ is given by the algebraic multiplicity. The number of Jordan blocks J_q containing λ_i is given by the geometric multiplicity.

Remark. Diagonalization is a special case of the Jordan canonical form.

10.1.4 Quotient Spaces

A very interesting way to produce new vector spaces is called quotienting. Suppose we have a vector space V and a linear subspace $U \subseteq V$. Then we can define a new vector space V/U as follows.

Definition 10.35 (Quotient Vector Space). Let V be a vector space and let $U \subseteq V$ be a linear subspace. We define an equivalence on V where two vectors v_1 and v_2 are said to be equivalent if and only if $v_1 - v_2 \in U$. Given a vector $v \in V$, we can construct the set

$$[v] = \{w \in V : w \text{ is equivalent to } v\}$$

One can also interpret this like:

$$[v] = \{w \in V : w - v \text{ is equivalent to } 0\} = \{v + w : w \in U\} = v + U$$

This collection is called the **equivalence class represented by** v . The set of all equivalence classes is denoted,

$$V/U = \{[v] : v \in V\}$$

This set forms a vector space. The sum of two equivalence classes, $[v]$ and $[w]$, is simply $[v] + [w] = [v + w]$. Similarly, the scalar multiple of an equivalence class is just $a[v] = [av]$.

Sometimes instead of writing $[v]$ we write $v + U$, so that $(x + U) + (y + U) = (x + y) + U$, and so on. Any factors in x or y

belonging to W can then be absorbed into W . This can be a little easier to use of than the square brackets, depending on the situation.

Remark. Notice that this is the same as the quotient group construction, where the group operation is vector addition.

Definition 10.36 (Quotient Map). Let V be a vector space and let $U \subseteq V$ be a subspace. Then the map $q : V \rightarrow V/U$ defined by $q(v) = [v]$ is called the quotient map.

Theorem 10.17. The quotient map $q : V \rightarrow V/U$ is linear and surjective.

Proof. Suppose $w \in V/U$. Then by definition, $w = [v]$ for some v . Therefore, for all w there exists v so that $q(v) = w$. \square

Example 10.7. Let $V = \mathbb{R}^2$. Recall that the real line $\mathbb{R} = \text{span}(\{(1, 0)\})$ forms a one-dimensional linear subspace of \mathbb{R}^2 . We can quotient out V by \mathbb{R} as follows. Consider two vectors $v = (v_1, v_2) \in \mathbb{R}^2$ and $w = (w_1, w_2) \in \mathbb{R}^2$. Then $v - w \in \mathbb{R}$ if and only if $v_2 = w_2$. Therefore, for any $v = (v_1, v_2) \in \mathbb{R}^2$ we have

$$[v] = \{(x, v_2) : x \in \mathbb{R}\}$$

Therefore, the vector space $\mathbb{R}^2 / \mathbb{R}$ contains equivalence classes of vectors in \mathbb{R}^2 , where two vectors are equivalent if they have the same y coordinates. Each equivalence class can therefore be represented by exactly one number, the y coordinate v_2 . We therefore see that the quotient space $\mathbb{R}^2 / \mathbb{R}$ is one dimensional. In fact, consider the following map

$$i : \mathbb{R}^2 / \mathbb{R} \rightarrow \mathbb{R}, \quad i([(x, v_2)]) = v_2$$

This map is invertible! Observe that $i^{-1} : \mathbb{R} \rightarrow \mathbb{R}^2 / \mathbb{R}$ with $i^{-1}(v_2) = [(x, v_2)]$ works. Both i and i^{-1} are linear maps, and $i \circ i^{-1}$ and $i^{-1} \circ i$ are both the identity map. Therefore, i is an isomorphism, so

$$\mathbb{R}^2 / \mathbb{R} \text{ is isomorphic to } \mathbb{R}$$

Theorem 10.18 (First Isomorphism Theorem). Let U and W be vector spaces and let $V = U \oplus W$. Then V/U is canonically isomorphic to W .

Proof. Since any vector $v \in V$ is of the form $v = u + w$, we can always decompose v into unique u and w . Consider the quotient map $q : V \rightarrow V/U$. Let $q|_W$ denote the restriction of q to the domain W .

Observe that $q(v) = q(u + w) = q(w) + q(u) = q(w) + [0] = q(w)$ for all u . Therefore for all $[v] \in V/U$ there is a unique $w \in W$ so that $q|_W(w) = [v]$. Therefore $q|_W$ is surjective.

Now observe that if $[w_1] = [w_2]$, then $[w_1 - w_2] = [0]$. But $[w_1 - w_2] = \{w_1 - w_2 + u : u \in U\}$. The only way this can be equal to $[0] = \{u : u \in U\}$ is if $w_1 = w_2$. Therefore $q|_W$ is injective. In other words, $q|_W$ is an isomorphism. \square

Corollary 10.2. Let V and U be finite dimensional vector spaces. Then $\dim V/U = \dim V - \dim U$.

Definition 10.37 (Well-Defined). Let V, U be vector spaces and let $W \subseteq V$ be a vector subspace. Let $f : V \rightarrow U$ be any function and let $\bar{f} : V/W \rightarrow U$ be defined by $\bar{f}([v]) = f(v)$ for all v . Then $\bar{f}([v])$ is **well-defined** if and only if $f(v) = f(x)$ for all $x \in [v]$.

Remark. The definition of \bar{f} **only** makes sense if \bar{f} is well defined because then the value of $\bar{f}([v])$ does not depend on the representative we choose for the class $[v]$.

10.1.5 Change of Coordinates

Here we will introduce Einstein notation, which means that wherever we see an expression including a repeated index, so that the index appears in both the superscript and subscript positions, we will ignore the summation symbol. Therefore the previous example would be written $v = v^i e_i$.

Definition 10.38 (Einstein Notation). Consider any sum of the form $\sum_i a_i b_i$. Whenever we see an expression like this, we will remove the summation symbol, and raise the index of one of the variables. That is, we define the following notation:

$$a_i b^i = \sum_i a_i b_i$$

Warning: Sometimes, you can't put a sum into the above form. For example, we will often write expressions of the form $\sum_{i=1}^N e_i \otimes e_i$. We could not write this as $e_i \otimes e^i$, since this means something completely different!

Given a basis B , if we were to construct an alternative basis $C = \{f_1, \dots, f_n\}$ for V , then there is necessarily some isomorphism $P : V \rightarrow V$ so that $P(e_i) = f_i$ for each $i = 1, \dots, n$. In coordinates, we would say there is an $n \times n$ matrix $[P_j^i]$ of coefficients so that $f_j = P_j^i e_i$ for all $j = 1, \dots, n$.

To make this translation between coordinates and abstract vectors a little bit more precise, we will introduce the following framework.

Definition 10.39 (Coordinate Map). Let V be a vector space and let $B = \{e_1, \dots, e_n\}$ be a basis for V . Then there is an isomorphism $\beta : V \rightarrow \mathbb{F}^n$ so that $\beta(v) = [v]_B = (v^1, \dots, v^n)$. The function β is called the **coordinate map** associated to B .

Therefore, in this framework, if C is another basis with associated isomorphism $\gamma : V \rightarrow \mathbb{F}^n$, then the matrix P_j^i can be seen as the coordinates of the linear map $[P]_{CB} = \gamma \circ P \circ \beta^{-1} : \mathbb{F}^n \rightarrow \mathbb{F}^n$.

To calculate the entries of the matrix of P in practice, we take each of the standard basis vectors in \mathbb{F}^n , which are of the form $v_i = (0, \dots, 1, \dots, 0)$, and apply the map β^{-1} to get $\beta^{-1}(v_i) = e_i$. We then apply the map P to this vector, to get $P(e_i)$,

and compute its expansion in terms of the basis vectors of C . So we will have something of the form $P(b_i) = P_i^j c_j$, which is the same as saying $\gamma \circ P \circ \beta^{-1}(v_i) = (P_i^1, \dots, P_i^n)$.

Lemma 10.1. The set of all linear maps, $L(V, W)$, is isomorphic to $M_{nm}(\mathbb{F})$.

Proof. Let B be a basis for V and let C be a basis for W . Let $\beta : V \rightarrow \mathbb{R}^n$ and $\gamma : W \rightarrow \mathbb{R}^m$ be the mappings associated to these bases.

Let $f : L(V, W) \rightarrow M_{nm}(\mathbb{F})$ be defined by $f(A) = \gamma \circ A \circ \beta^{-1} = [A]_{CB}$. Then

$$\begin{aligned} f^{-1}([A]_{CB}) &= \gamma^{-1} \circ [A]_{CB} \circ \beta \\ &= \gamma^{-1} \circ \gamma \circ A \circ \beta^{-1} \circ \beta \\ &= A \end{aligned}$$

So f is invertible.

Finally, since γ and β are both linear, f must be linear. Therefore f is an isomorphism as required. \square

Corollary 10.3. $\dim L(V, W) = \dim V \dim W$.

Let us go through an explicit example.

Example 10.8. Consider $\mathbb{F} = \mathbb{R}$, and let V be defined as follows.

$$V = \{a + bx + cx^2 : a, b, c \in \mathbb{R}\}$$

This is the set of quadratic polynomials with coefficients in \mathbb{R} . One possible basis for this is the following,

$$B = \{1, x, x^2\}$$

An alternative basis could be,

$$C = \{1 - x, 1 + x, x^2\}$$

Now let us consider the map defined by

$$P(a + bx + cx^2) := b + ax + cx^2$$

Let us first compute the matrix $[P]_{BB}$. For the basis B , we have $\beta(a + bx + cx^2) = (a, b, c)$. When we apply P to this, we get

$P \circ \beta^{-1}(a, b, c) = b + ax + cx^2$. Therefore $\beta \circ P \circ \beta^{-1}(a, b, c) = (b, a, c)$. In other words,

$$[P]_{BB} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Now let us compute $[P]_{CB}$. This will take a vector (a, b, c) as input, apply the map P to it, and then compute the result in terms of the C basis. Let us find the map γ associated to the C basis. We have $\gamma(1+x) = (1, 0, 0)$, $\gamma(1-x) = (0, 1, 0)$, and $\gamma(x^2) = (0, 0, 1)$. Since γ is linear we therefore have $\frac{1}{2}(\gamma(1+x) + \gamma(1-x)) = \gamma(1) = \frac{1}{2}(1, 1, 0)$. Similarly, $\gamma(x) = \frac{1}{2}(1, -1, 0)$ and $\gamma(x^2) = (0, 0, 1)$. So now we know how to put any polynomial into the C basis. We then calculate $\gamma \circ P \circ \beta^{-1}(a, b, c) = \gamma(b + ax + cx^2)$, which is

$$\begin{aligned} \gamma \circ P \circ \beta^{-1}(a, b, c) &= \gamma(b + ax + cx^2) \\ &= \frac{1}{2}b(1, 1, 0) + \frac{1}{2}a(1, -1, 0) + c(0, 0, 1) \\ &= \left(\frac{b+a}{2}, \frac{b-a}{2}, c \right) \end{aligned}$$

Therefore,

$$[P]_{CB} = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

References

- [1] Paolo Aluffi. “Algebra: Chapter Zero”. In: *Graduate Studies in Mathematics*. Graduate studies in mathematics. Providence, Rhode Island: American Mathematical Society, July 2009.
- [2] Sheldon Axler. “Linear Algebra Done Right”. In: *Undergraduate Texts in Mathematics*. Cham: Springer International Publishing, 2015.
- [3] Robert L Bryant. “Submanifolds and special structures on the octonians”. In: *Journal of Differential Geometry* 17.2 (1982), pp. 185–232.
- [4] Elie Cartan. “Sur la structure des groupes simples finis et continus”. In: *Comptes rendus de l’Académie des Sciences* 116 (1893), pp. 784–786.
- [5] Claude C Chevalley. *The algebraic theory of spinors and Clifford algebras*. en. Ed. by Pierre Cartier and Catherine Chevalley. Collected works of Claude Chevalley. Berlin, Germany: Springer, Dec. 1995.
- [6] John H Conway and Derek A Smith. *On Quaternions and Octonions*. Natick, MA: A K Peters, July 2001.
- [7] David S Dummit and Richard M Foote. *Abstract Algebra*. en. 3rd ed. Nashville, TN: John Wiley & Sons, June 2003.
- [8] Nathan Jacobson. *Basic Algebra I*. 2nd ed. Dover Books on Mathematics. Mineola, NY: Dover Publications, June 2009.
- [9] Herbert Blaine Lawson and Marie-Louise Michelsohn. *Spin geometry (PMS-38), volume 38*. en. Princeton Mathematical Series. Princeton, NJ: Princeton University Press, Feb. 1990.
- [10] Ian R Porteous. *Clifford Algebras and the Classical Groups*. Cambridge: Cambridge University Press, Oct. 1995.
- [11] Frank Reese Harvey. *Spinors and Calibrations*. San Diego, CA: Academic Press, Apr. 1990.
- [12] Steven Roman. “Advanced Linear Algebra”. In: *Graduate Texts in Mathematics*. New York, NY: Springer New York, 2008.

Index

Algebra, [146](#)

 Alternative, [100](#)

 Center, [68](#)

 Clifford Algebra, [82](#)

 Composition Algebra, [65](#)

 Division, [64](#)

 Exterior Algebra, [20](#)

 Graded, [20](#)

 Ideal, [81](#)

 Left/Right Ideal, [94](#)

 Subalgebra, [147](#)

 Tensor Algebra, [13](#)

 Twisted Center, [92](#)

 Unital, [13](#)

Annihilator, [22](#)

Automorphism, [69](#)

 Anti-Automorphism, [69](#)

 Clifford Automorphism, [87](#)

 Inner/Outer, [131](#)

Basis

 G-adjoint, [36](#)

 Hyperbolic, [40](#)

 Orthonormal, [31](#)

Bijjective, [133](#)

Bilinear Form

 Pullback, [34](#)

Bilinear Map, [6](#)

Bott Periodicity, [119](#)

Canonical Injection, [6](#)

Cayley Form, [114](#)

Characteristic, [134](#)

Characteristic Polynomial, [152](#)

Complex Structure, [62](#)

Complexification, [63](#)

Coordinate Map, [157](#)

Coset, [140](#)

 Ideal, [142](#)

Covector, [5](#)

Determinant, [150](#)

Diagonal Subgroup, [72](#)

Direct

 Sum, [145](#)

Dual

 Bilinear Form, [53](#)

 Map, [15](#)

 Space, [5](#)

Eigenvector, [152](#)

Einstein Notation, [157](#)

Endomorphism, [148](#)

Exceptional Groups, [113](#)

Extension by Linearity, [149](#)

Exterior

k -form, [16](#)

 Algebra, [20](#)

 Power

 Linear Map, [21](#)

 Vector Space, [16](#)

 Product, [16](#)

Fiber, [134](#)

Field, [134](#)

 Definition using rings, [143](#)

 of Fractions of Polynomial Ring, [50](#)

First Isomorphism Theorem, [156](#)

Form

k -form, [16](#)

 1-form, [5](#)

 Cayley, [114](#)

 Co-associative, [114](#)

 Non-Degenerate Bilinear, [24](#)

 Symmetric/Skew-Symmetric Bilinear, [24](#)

 Symplectic, [49](#)

 Volume, [53](#)

Gram-Schmidt Procedure, [153](#)

- Group, [135](#)
 - Action, [141](#)
 - Transitive, [141](#)
 - Alternating, [139](#)
 - Cyclic, [136](#)
 - Direct Product, [70](#)
 - Exceptional, [113](#)
 - Generating Set, [136](#)
 - of Integers, [135](#)
 - of Permutations, [136](#)
 - of Units, [41](#)
 - Order, [135](#)
 - Orthogonal, [34](#)
 - Quotient, [140](#)
 - Special Orthogonal, [34](#)
 - Unit Quaternions, [70](#)
- Hadamard Inequality, [91](#)
- Hodge Star, [56](#)
- Homomorphism
 - Algebra, [69](#)
 - Group, [138](#)
- Hyperbolic Plane, [40](#)
- Ideal
 - Algebra, [81](#)
 - Left/Right Ideal of Algebra, [94](#)
- Image
 - of Group Homomorphism, [138](#)
 - of Linear Map, [148](#)
- Inclusion Map, [27](#)
- Index
 - of Subgroup, [141](#)
- Injective, [133](#)
- Integral Domain, [50](#)
- Invariant
 - Characteristic Polynomial, [152](#)
 - Determinant of Linear Map, [151](#)
 - Pfaffian of Skew-Symmetric Bilinear Form, [51](#)
 - Rank of a Linear Map, [148](#)
- Rank of Bilinear Form, [24](#)
- Signature, [32](#)
- Isometric, [34](#)
- Isometry, [34](#)
 - Anti-Isometry, [59](#)
 - Proper/Improper, [37](#)
- Isomorphism, [148](#)
 - Algebra, [69](#)
 - of Groups, [139](#)
 - Symplectic, [49](#)
- Isotropic
 - Subspace, [39](#)
- Kernel
 - of Group Homomorphism, [138](#)
 - of Linear Map, [148](#)
- Laplace expansion, [150](#)
- Linear
 - \mathbb{C} -Linear, [60](#)
 - \mathbb{R} -Linear, [60](#)
 - Extension, [149](#)
 - Functional, [5](#)
 - Independence, [145](#)
 - Map, [147](#)
 - Operator, [147](#)
 - Subspace, [144](#)
- Matrix
 - Algebra, [147](#)
 - Gram, [53](#)
 - Normal, [153](#)
 - of a Bilinear Form, [23](#)
 - of a Linear Map, [150](#)
 - of Inner Product, [152](#)
 - Orthogonal, [153](#)
 - Overlap, [53](#)
 - Symmetric, [153](#)
- Module, [144](#)
- Multi-Index, [12](#)
- Multiplicity

- Algebraic, [152](#)
- Geometric, [152](#)
- Musical Isomorphism, [26](#)
- Non-Degenerate, [24](#)
- Norm, [65](#)
- Operator
 - Normal, [153](#)
 - Self-Adjoint, [36](#)
 - Skew-Adjoint, [36](#)
- Orthogonal
 - General Definition, [24](#)
 - Group, [34](#)
 - Matrix, [153](#)
- Orthonormal
 - Basis, [31](#)
- Orthonormal Basis, [153](#)
- Parity, [137](#)
- Poincaré Dual Isomorphism, [55](#)
- Product
 - Cartesian, [133](#)
 - Exterior/Wedge, [16](#)
 - Inner, [152](#)
 - Tensor
 - Dual Spaces, [7](#)
 - Linear Forms, [7](#)
 - Vector Spaces, [8](#)
- Pullback, [15](#)
- Quaternion, [67](#)
 - Conjugate, [68](#)
 - Imaginary, [68](#)
 - Unit, [68](#)
- Quotient
 - Bilinear Form, [29](#)
 - Map, [156](#)
 - of Groups, [140](#)
 - of Vector Spaces, [155](#)
- Radical, [26](#)
- Real Structure, [63](#)
- Reflection, [37](#)
- Representation, [96](#)
 - Adjoint, [97](#)
 - Twisted Adjoint, [97](#)
- Ring, [142](#)
 - of Polynomials, [50](#)
- Sharp, [26](#)
- Signature, [31](#)
- Space
 - Perp, [25](#)
- Subalgebra, [147](#)
- Subgroup, [135](#)
 - Normal, [140](#)
- Subspace
 - Maximally Isotropic, [44](#)
 - Nondegenerate, [42](#)
- Surjective, [133](#)
- Symmetric
 - Bilinear Form, [24](#)
 - Negative (Semi)Definite, [32](#)
 - Positive (Semi)Definite, [32](#)
 - Matrix, [153](#)
- Symplectic Form, [49](#)
- Symplectomorphism, [49](#)
- Tensor
 - Algebra, [13](#)
 - Decomposable, [13](#)
 - Power
 - Linear Map, [21](#)
 - Vector Space, [12](#)
 - Product
 - Dual Spaces, [7](#)
 - Linear Forms, [7](#)
 - Vector Spaces, [8](#)
 - Totally Antisymmetric, [16](#)
- Theorem
 - Spin(8) Triality Theorem, [129](#)

- Cartan-Dieudonné Theorem, [45](#)
- Classification of Clifford Algebras, [119](#)
- Classification of Skew-Symmetric Bilinear Forms, [48](#)
- Classification of Symmetric Bilinear Forms, [47](#)
- First Isomorphism Theorem, [156](#)
- First Isomorphism Theorem (for Groups), [140](#)
- Hadamard Inequality, [91](#)
- Lagrange's Theorem, [141](#)
- Rank-Nullity Theorem, [148](#)
- Riesz Representation Theorem, [27](#)
- Spectral Theorem, [154](#)
- Sylvester's Law of Inertia, [32](#)
- Witt Cancellation Theorem, [42](#)
- Witt Extension Theorem, [44](#)
- Transposition, [137](#)
- Triality, [124](#)
- Unique Factorization Domain, [50](#)
- Universal Property
 - Clifford Algebra, [84](#)
 - Definition, [83](#)
 - Products, [83](#)
- Vector
 - Algebra, [146](#)
 - Space, [144](#)
- Vector Space
 - Complex Conjugate Space, [60](#)
 - Dimension, [146](#)
 - Underlying Real Vector Space, [61](#)
- Wedge Product, [16](#)
- Well Defined, [157](#)
- Witt
 - Cancellation Theorem, [42](#)
 - Extension Theorem, [44](#)
 - Index, [45](#)