5

Tensor Algebra

Summary. In this chapter we develop somewhat quickly the basic facts of tensor algebra, assuming the reader is familiar with linear algebra. Tensor algebra should be thought of as a natural development of the theory of functions in several vector variables. To some extent it is equivalent, at least in our setting, to this theory.

1 Tensor Algebra

1.1 Functions of Two Variables

The language of functions is most suitably generalized into the language of tensor algebra. The idea is simple but powerful: the dual V^* of a vector space V is a space of functions on V, and V itself can be viewed as functions on V^*.

A way to stress this symmetry is to use the bra-ket $\langle \mid \rangle$ notation of the physicists:[27] given a linear form $\phi \in V^*$ and a vector $v \in V$, we denote by $\langle \phi | v \rangle := \phi(v)$ the value of ϕ on v (or of v on ϕ !).

From linear functions one can construct polynomials in one or several variables. Tensor algebra provides a coherent model to perform these constructions in an intrinsic way.

Let us start with some elementary remarks. Given a set X (with n elements) and a field F, we can form the n-dimensional vector space F^X of functions on X with values in F.

This space comes equipped with a canonical basis: the characteristic functions of the elements of X. It is convenient to identify X with this basis and write $\sum_{x \in X} f(x)x$ for the vector corresponding to a function f.

From two sets X, Y (with n, m elements, respectively) we can construct F^X, F^Y, and also $F^{X \times Y}$. This last space is the space of functions in two variables. It has dimension nm.

[27] This was introduced in quantum mechanics by Dirac.

Of course, given a function $f(x) \in F^X$ and a function $g(y) \in F^Y$, we can form the two variable function $F(x, y) := f(x)g(y)$; the product of the given basis elements is just $xy = (x, y)$. A simple but useful remark is the following:

Proposition. *Given two bases $u_1, \dots, u_n$ of F^X and $v_1, \dots, v_m$ of X^Y the nm elements $u_i v_j$ are a basis of $F^{X \times Y}$.*

Proof. The elements xy are a basis of $F^{X \times Y}$. We express x as a linear combination of the $u_1, \dots, u_n$ and y as one of the $v_1, \dots, v_m$.

We then see, by distributing the products, that the nm elements $u_i v_j$ span the vector space $F^{X \times Y}$. Since this space has dimension nm the $u_i v_j$ must be a basis. □

1.2 Tensor Products

We perform the same type of construction with a tensor product of two spaces, without making any reference to a basis. Thus we define:

Definition 1. Given 3 vector spaces U, V, W a map $f(u, v) : U \times V \to W$ is *bilinear* if it is linear in each of the variables u, v separately.

If U, V, W are finite dimensional we easily see that:

Proposition. *The following conditions on a bilinear map $f : U \times V \to W$ are equivalent:*

(i) There exist bases $u_1, \dots, u_n$ of U and $v_1, \dots, v_m$ of V such that the nm elements $f(u_i, v_j)$ are a basis of W.
(ii) For all bases $u_1, \dots, u_n$ of U and $v_1, \dots, v_m$ of V, the nm elements $f(u_i, v_j)$ are a basis of W.
(iii) $\dim(W) = nm$, and the elements $f(u, v)$ span W.
(iv) Given any vector space Z and a bilinear map $g(u, v) : U \times V \to Z$ there exists a unique linear map $G : W \to Z$ such that $g(u, v) = G(f(u, v))$ (universal property).

Definition 2. A bilinear map is called a *tensor product* if it satisfies the equivalent conditions of the previous proposition.

Property (iv) ensures that two different tensor product maps are canonically isomorphic. In this sense we will speak of W as the *tensor product of two vector spaces* which we will denote by $U \otimes V$. We will denote by $u \otimes v$ the image of the pair (u, v) in the bilinear (tensor product) map.

Definition 3. The elements $u \otimes v$ are called *decomposable tensors.*

Example. The bilinear product $F \times U \to U$ given by $(\alpha, u) \mapsto \alpha u$ is a tensor product.

1.3 Bilinear Functions

To go back to functions, we can again concretely treat our constructions as follows. Consider the space $\mathrm{Bil}(U \times V, F)$ of bilinear functions with values in the field F. We have a bilinear map

$$F : U^* \times V^* \to \mathrm{Bil}(U \times V, F)$$

given by $F(\varphi, \psi)(u, v) := \langle\varphi|u\rangle\langle\psi|v\rangle$. In other and more concrete words, the product of two linear functions, in separate variables, is bilinear.

In given bases $u_1, \dots, u_n$ of U and $v_1, \dots, v_m$ of V we have for a bilinear function

$$f\Big(\sum_{i=1}^{n} \alpha_i u_i, \sum_{j=1}^{m} \beta_j v_j\Big) = \sum_{i=1}^{n}\sum_{j=1}^{m} \alpha_i \beta_j f(u_i, v_j). \tag{1.3.1}$$

Let e^{hk} be the bilinear function defined by the property

$$e^{hk}\Big(\sum_{i=1}^{n} \alpha_i u_i, \sum_{j=1}^{m} \beta_j v_j\Big) = \alpha_h \beta_k. \tag{1.3.2}$$

We easily see that these bilinear functions form a basis of $\mathrm{Bil}(U \times V, F)$, and a general bilinear function f is expressed in this basis as

$$f(u, v) = \sum_{i=1}^{n}\sum_{j=1}^{m} f(u_i, v_j) e^{ij}(u, v), \quad f = \sum_{i=1}^{n}\sum_{j=1}^{m} f(u_i, v_j) e^{ij}. \tag{1.3.3}$$

Moreover let u^i and v^j be the dual bases of the two given bases. We see immediately that $e^{hk}(u, v) = u^h(u)v^k(v)$. Thus we are exactly in the situation of a tensor product, and we may say that $\mathrm{Bil}(U \times V, F) = U^* \otimes V^*$.

In the more familiar language of polynomials, we can think of n variables x_i and m variables y_j. The space of bilinear functions is the span of the bilinear monomials $x_i y_j$.

Since a finite-dimensional vector space U can be identified with its double dual it is clear how to construct a tensor product. We may set[28]

$$U \otimes V := \mathrm{Bil}(U^* \times V^*, F).$$

1.4 Tensor Product of Operators

The most important point for us is that one can also perform the tensor product of operators using the universal property.

[28] Nevertheless, the tensor product construction holds for much more general situations than the one we are treating now. We refer to N. Bourbaki for a more detailed discussion [B].

If $f : U_1 \to V_1$ and $g : U_2 \to V_2$ are two linear maps, the map $U_1 \times U_2 \to V_1 \otimes V_2$ given by $(u, v) \to f(u) \otimes g(v)$ is bilinear. Hence it factors through a unique linear map denoted by $f \otimes g : U_1 \otimes U_2 \to V_1 \otimes V_2$.

This is characterized by the property

$$(1.4.1) \qquad (f \otimes g)(u \otimes v) = f(u) \otimes g(v).$$

In matrix notation the only difficulty is a notational one. Usually it is customary to index basis elements with integral indices. Clearly if we do this for two spaces, the tensor product basis is indexed with pairs of indices and so the corresponding matrices are indexed with pairs of pairs of indices.

Concretely, if $f(u_i) = \sum_j a_{ji} u'_j$ and $g(v_h) = \sum_k b_{kh} v'_k$ we have

$$(1.4.2) \qquad (f \otimes g)(u_i \otimes v_h) = \sum_{jk} a_{ji} b_{kh} u'_j \otimes v'_k.$$

Hence the elements $a_{ji} b_{kh}$ are the entries of the tensor product of the two matrices.

An easy exercise shows that the tensor product of maps is again bilinear and thus defines a map

$$\hom(U_1, V_1) \otimes \hom(U_2, V_2) \to \hom(U_1 \otimes U_2, V_1 \otimes V_2).$$

Using bases and matrix notations (and denoting by $M_{m,n}$ the space of $m \times n$ matrices), we have thus a map

$$M_{m,n} \otimes M_{p,q} \to M_{mp,nq}.$$

We leave it to the reader to verify that the tensor product of the elementary matrices gives the elementary matrices, and hence that this mapping is an isomorphism.

Finally we have the obvious associativity conditions. Given

$$\begin{array}{c} U_1 \xrightarrow{\quad f \quad} V_1 \xrightarrow{\quad h \quad} W_1 \\ U_2 \xrightarrow{\quad g \quad} V_2 \xrightarrow{\quad k \quad} W_2 \end{array}$$

we have $(h \otimes k)(f \otimes g) = hf \otimes kg$. In particular, consider two spaces U, V and endomorphisms $f : U \to U$, $g : V \to V$. We see that:

Proposition. *The mapping $(f, g) \to f \otimes g$ is a representation of $GL(U) \times GL(V)$ in $GL(U \otimes V)$.*

There is an abstraction of this notion. Suppose we are given two associative algebras A, B over F. The vector space $A \otimes B$ has an associative algebra structure, by the universal property, which on decomposable tensors is

$$(a \otimes b)(c \otimes d) = ac \otimes bd.$$

Given two modules M, N on A, and B respectively, $M \otimes N$ becomes an $A \otimes B$-module by

$$(a \otimes b)(m \otimes n) = am \otimes bn.$$

Remark. (1) Given two maps $i, j : A, B \to C$ of algebras such that the images commute, we have an induced map $A \otimes B \to C$ given by $a \otimes b \to i(a)j(b)$.

This is a characterization of the tensor product by universal maps.

(2) If A is an algebra over F and $G \supset F$ is a field extension, then $A \otimes_F G$ can be thought of as a G algebra.

Remark. Given an algebra A and two modules M, N, in general $M \otimes N$ does not carry any natural A-module structure. This is the case for group representations or more generally for Hopf algebras, in which one assumes, among other things, to have a homomorphism $\Delta : A \to A \otimes A$ (for the group algebra of G it is induced by $g \mapsto g \otimes g$), cf. Chapter 8, §7.

1.5 Special Isomorphisms

We analyze some special cases.

First, we can identify any vector space U with $\hom(F, U)$ associating to a map $f \in \hom(F, U)$ the vector $f(1)$. We have also seen that $F \otimes U = U$, and $a \otimes u = au$.

We thus follow the identifications:

$$V \otimes U^* = \hom(F, V) \otimes \hom(U, F) = \hom(F \otimes U, V \otimes F).$$

This last space is identified with $\hom(U, V)$.

Proposition. *There is a canonical isomorphism* $V \otimes U^* = \hom(U, V)$.

It is useful to make this identification explicit and express the action of a decomposable element $v \otimes \varphi$ on a vector u, as well as the composition law of morphisms. Consider the tensor product map

$$\hom(V, W) \times \hom(U, V) \to \hom(U, W).$$

With the obvious notations we easily find:

$$(1.5.1) \qquad (v \otimes \varphi)(u) = v\langle \varphi | u \rangle, \quad w \otimes \psi \circ v \otimes \varphi = w \otimes \langle \psi | v \rangle \varphi.$$

In the case of $\mathrm{End}(U) := \hom(U, U)$, we have the identification $\mathrm{End}(U) = U \otimes U^*$; in this case we can consider the linear map $\mathrm{Tr} : U \otimes U^* \to F$ induced by the bilinear pairing given by duality

$$(1.5.2) \qquad \mathrm{Tr}(u \otimes \varphi) := \langle \varphi | u \rangle.$$

Definition. The mapping $\mathrm{Tr} : \mathrm{End}(U) \to F$ is called the *trace*. In matrix notations, if e_i is a basis of U and e^i the dual basis, given a matrix $A = \sum_{ij} a_{ij} e_i \otimes e^j$ we have $\mathrm{Tr}(A) = \sum_{ij} a_{ij} \langle e^j | e_i \rangle = \sum_i a_{ii}$.

For the tensor product of two endomorphisms of two vector spaces one has

$$(1.5.3) \qquad \mathrm{Tr}(A \otimes B) = \mathrm{Tr}(A)\,\mathrm{Tr}(B),$$

as verified immediately.

Finally, given linear maps $X : W \to U,\ Y : V \to Z,\ v \otimes \phi : U \to V$ we have

$$(1.5.4) \qquad Y \circ v \otimes \phi \circ X = Yv \otimes X^t \phi.$$

1.6 Decomposable Tensors

An immediate consequence of the previous analysis is:

Proposition. *The decomposable tensors in* $V \otimes U^* = \hom(U, V)$ *are the maps of rank* 1.

In particular this shows that most tensors are not decomposable. In fact, quite generally:

Exercise. In a tensor product (with the notations of Section 1) a tensor $\sum a_{ij} u_i \otimes v_j$ is decomposable if and only if the $n \times m$ matrix with entries a_{ij} has rank ≤ 1.

Another important case is the sequence of identifications:

$$(1.6.1) \quad U^* \otimes V^* = \hom(U, F) \otimes \hom(V, F) \\ = \hom(U \otimes V, F \otimes F) = \hom(U \otimes V, F),$$

i.e., the tensor product of the duals is identified with the dual of the tensor product. In symbols

$$(U \otimes V)^* = U^* \otimes V^*.$$

It is useful to write explicitly the duality pairing at the level of decomposable tensors:

$$(1.6.2) \qquad \langle \varphi \otimes \psi | u \otimes v \rangle = \langle \varphi | u \rangle \langle \psi | v \rangle.$$

In other words, if we think of $U^* \otimes V^*$ as the space of bilinear functions $f(u, v)$, $u \in U$, $v \in V$ the tensor $a \otimes b$ is identified, as a linear function on this space and as the evaluation $f \to f(a, b)$. The interpretation of $U \otimes V$ as bilinear functions on $U^* \times V^*$ is completely embedded in this basic pairing.

Summarizing, we have seen the intrinsic notion of $U \otimes V$ as the solution of a universal problem, as bilinear functions on $U^* \times V^*$, and finally as the dual of $U^* \otimes V^*$.[29]

1.7 Multiple Tensor Product

The tensor product construction can clearly be iterated. The multiple tensor product map

$$U_1 \times U_2 \times \cdots \times U_m \to U_1 \otimes U_2 \otimes \cdots \otimes U_m$$

is the universal multilinear map, and we have in general the dual pairing:

$$U_1^* \otimes U_2^* \otimes \cdots \otimes U_m^* \times U_1 \otimes U_2 \otimes \cdots \otimes U_m \to F$$

[29] Of these three definitions, the solution of the universal problem is the one which admits the widest generalizations.

given on the decomposable tensors by

$$(1.7.1)\qquad \langle \varphi_1 \otimes \varphi_2 \otimes \cdots \otimes \varphi_m | u_1 \otimes u_2 \otimes \cdots \otimes u_m \rangle = \prod_{i=1}^{m} \langle \varphi_i | u_i \rangle.$$

This defines a canonical identification of $(U_1 \otimes U_2 \otimes \cdots U_m)^*$ with $U_1^* \otimes U_2^* \otimes \cdots U_m^*$.

Similarly we have an identification

$$\hom(U_1 \otimes U_2 \otimes \cdots U_m, V_1 \otimes V_2 \otimes \cdots V_m)$$
$$\cong \hom(U_1, V_1) \otimes \hom(U_2, V_2) \otimes \cdots \otimes \hom(U_m, V_m).$$

Let us consider the self-dual pairing on $\mathrm{End}(U)$ given by $\mathrm{Tr}(AB)$ in terms of decomposable tensors. If $A = v \otimes \psi$ and $B = u \otimes \varphi$ we have

$$(1.7.2)\qquad \mathrm{Tr}(AB) = \mathrm{Tr}(v \otimes \psi \circ u \otimes \varphi) = \mathrm{Tr}(\langle \psi | u \rangle v \otimes \varphi) = \langle \varphi | v \rangle \langle \psi | u \rangle.$$

We recover the simple fact that $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.

We remark also that this is a nondegenerate pairing, and $\mathrm{End}(U) = U \otimes U^*$ is identified by this pairing with its dual:

$$\mathrm{End}(U)^* = (U \otimes U^*)^* = U^* \otimes (U^*)^* = U^* \otimes U \cong U \otimes U^*.$$

We identify an operator A with the linear function $X \mapsto \mathrm{Tr}(AX) = \mathrm{Tr}(XA)$.

Since $\mathrm{Tr}([A, B]) = 0$, the operators with trace 0 form a Lie algebra (of the group $SL(U)$), called $sl(U)$ (and an ideal in the Lie algebra of all linear operators).

For the identity operator in an n-dimensional vector space we have $\mathrm{Tr}(1) = n$.

If we are in characteristic 0 (or prime with n) we can decompose each matrix as $A = \frac{\mathrm{Tr}(A)}{n} 1 + A_0$ where A_0 has zero trace. Thus the Lie algebra $gl(U)$ decomposes as the direct sum $gl(U) = F \oplus sl(U)$, F being identified with the *scalar matrices*, i.e., with the multiples of the identity matrix.

It will be of special interest to us to consider the tensor product of several copies of the same space U, i.e., the *tensor power* of U, denoted by $U^{\otimes m}$. It is convenient to form the direct sum of all of these powers since this space has a natural algebra structure defined on the decomposable tensors by the formula

$$(u_1 \otimes u_2 \otimes \cdots \otimes u_h)(v_1 \otimes v_2 \otimes \cdots \otimes v_k)$$
$$:= u_1 \otimes u_2 \otimes \cdots \otimes u_h \otimes v_1 \otimes v_2 \otimes \cdots \otimes v_k.$$

We usually use the notation

$$T(U) := \bigoplus_{k=0}^{\infty} U^{\otimes k}.$$

This is clearly a graded algebra generated by the elements of degree 1.

Definition. $T(U)$ is called the *tensor algebra* of U.

This algebra is characterized by the following *universal property*:

Proposition. *Any linear mapping $j : U \to R$ into an associative algebra R extends uniquely to a homomorphism $\overline{j} : T(U) \to R$.*

Proof. The mapping $U \times U \times \cdots \times U \to R$ given by $j(u_1)j(u_2)\dots j(u_k)$ is multilinear and so defines a linear map $U^{\otimes k} \to R$.

The required map is the sum of all these maps and is clearly a homomorphism extending j; it is also the unique possible extension since U generates $T(U)$ as an algebra. □

1.8 Actions on Tensors

In particular, a linear automorphism g of U extends to an automorphism of the tensor algebra which acts on the tensors $U^{\otimes m}$ as $g^{\otimes m} := g \otimes g \otimes g \cdots \otimes g$.

Thus we have:

Proposition. *$GL(U)$ acts naturally on $T(U)$ as algebra automorphisms (preserving the degree and extending the standard action on U).*

It is quite suggestive to think of the tensor algebra in a more concrete way. Let us fix a basis of U which we think of as indexed by the letters of an *alphabet* A with n letters.[30]

If we write the tensor product omitting the symbol $\otimes$ we see that a basis of $U^{\otimes m}$ is given by all the n^m *words* of length m in the given alphabet.

The multiplication of two words is just the *juxtaposition* of the words (i.e., write one after the other as a unique word). In this language we see that the tensor algebra can be thought of as *the noncommutative polynomial ring in the variables A*, or *the free algebra on A*, or the *monoid algebra* of the free monoid.

When we think in these terms we adopt the notation $F\langle A\rangle$ instead of $T(U)$.

In this language the universal property is that of polynomials, i.e., we can evaluate a polynomial in any algebra once we give the *values* for the variables.

In fact, since A is a basis of U, a linear map $j : U \to R$ is determined by assigning arbitrarily the *values for the variables A*. The resulting map sends a *word*, i.e., a product of variables, to the corresponding product of the values. Thus this map is really the evaluation of a polynomial.

The action of a linear map on U is a special substitution of variables, a *linear substitution.*

Notice that we are working in the category of all associative algebras and thus we have to use *noncommutative* polynomials, i.e., elements of the free algebra. Otherwise the evaluation map is either not defined or not a homomorphism.

Remark. As already mentioned the notion of tensor product is much more general than the one we have given. We will use at least one case of the more general definition. If A is an algebra over a field k, M a right A-module, N a left A-module, we

[30] Of course we use the usual alphabet, and so in our examples this restricts n artificially, but there is no theoretical obstruction to think of a possibly infinite alphabet.

define $M \otimes_A N$ to be the quotient of the vector space $M \otimes N$ modulo the elements $ma \otimes n - m \otimes an$. This construction typically is used when dealing with *induced representations*. We will use some simple properties of this construction which the reader should be able to verify.

2 Symmetric and Exterior Algebras

2.1 Symmetric and Exterior Algebras

We can reconstruct the commutative picture by passing to a quotient.

Given an algebra R, there is a unique minimal ideal I of R such that R/I is commutative. It is the ideal generated by all of the commutators $[a, b] := ab - ba$.

In fact, I is even generated by the commutators of a set of generators for the algebra since if an algebra is generated by pairwise commuting elements, then it is commutative.

Consider this ideal in the case of the tensor algebra. It is generated by the commutators of the elements of degree 1, hence it is a homogeneous ideal, and so the resulting quotient is a graded algebra, called the *symmetric algebra* on U.

It is usually denoted by $S(U)$ and its homogeneous component of degree m is called the m^{th} *symmetric power* of the space U and denoted $S^m(U)$.

In the presentation as a free algebra, to make $F\langle A\rangle$ commutative means to impose the commutative law on the variables A. This gives rise to the polynomial algebra $F[A]$ in the variables A. Thus $S(U)$ is isomorphic to $F[A]$.

The canonical action of $GL(U)$ on $T(U)$ clearly leaves invariant the commutator ideal and so induces an action as algebra automorphisms on $S(U)$. In the language of polynomials we again find that the action may be realized by changes of variables.

There is another important algebra, the *Grassmann* or *exterior algebra*. It is defined as $T(U)/J$ where J is the ideal generated by all the elements $u^{\otimes 2}$ for $u \in U$. It is usually denoted by $\bigwedge U$.

The multiplication of elements in $\bigwedge U$ is indicated by $a \wedge b$. Again we have an action of $GL(U)$ on $\bigwedge U = \oplus_k \bigwedge^k U$ as automorphisms of a graded algebra, and the algebra satisfies a universal property with respect to linear maps. Given a linear map $j : U \to R$ into an algebra R, restricted by the condition $j(u)^2 = 0, \forall u \in U$, we have that j extends to a unique homomorphism $\bigwedge U \to R$.

In the language of alphabets we have the following description. The variables in A satisfy the rules:

$$a \wedge b = -b \wedge a, \; a \wedge a = 0.$$

We order the letters in A. A monomial M is 0 if it contains a repeated letter. Otherwise reorder it in alphabetical order, introducing a negative sign if the permutation used to reorder is odd; let us denote by $a(M)$ this value.

Consider the monomials in which the letters appear in strict increasing order, and we call these the *strict monomials* (if A has n elements we have $\binom{n}{k}$ strict monomials of degree k for a total of 2^n monomials). For example, from a basis $e_1, \dots, e_n$ of U we deduce a basis

$$e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_k}, \quad 1 \le i_1 < i_2 < \cdots < i_k \le n$$

of $\bigwedge U$.

Theorem. *The strict monomials are a basis of* $\bigwedge U$.

Proof. We hint at a combinatorial proof. We construct a vector space with basis the strict monomials. We then define a product by $M \wedge N := a(MN)$. A little combinatorics shows that we have an associative algebra R, and the map of A into R determines an isomorphism of R with $\bigwedge U$. □

For a different proof see Section 4.1 in which we generalize this theorem to Clifford algebras.

In particular, we have the following dimension computations. If $\dim U = n$,

$$\dim \bigwedge^k U = \binom{n}{k}, \ \dim \bigwedge U = 2^n, \ \dim \bigwedge^n U = 1.$$

Let $\dim U = n$. The bilinear pairing $\bigwedge^k U \times \bigwedge^{n-k} U \to \bigwedge^n U$ induces a linear map

$$j : \bigwedge^k U \to \hom\left(\bigwedge^{n-k} U, \bigwedge^n U\right) = \bigwedge^n U \otimes \left(\bigwedge^{n-k} U\right)^*,$$
$$j(u_1 \wedge \cdots \wedge u_k)(v_1 \wedge \cdots v_{n-k}) := u_1 \wedge \cdots \wedge u_k \wedge v_1 \wedge \cdots \wedge v_{n-k}.$$

In a given basis $e_1, \dots, e_n$ we have $j(e_{i_1} \wedge \cdots \wedge e_{i_k})(e_{j_1} \wedge \cdots \wedge e_{j_{n-k}}) = 0$ if the elements $i_1, \dots, i_k, j_1, \dots, j_{n-k}$ are not a permutation of $1, 2, \dots, n$. Otherwise, reordering, the value we obtain is $\epsilon_\sigma e_1 \wedge e_2 \dots \wedge e_n$, where σ is the (unique) permutation that brings the elements $i_1, \dots, i_k, j_1, \dots, j_{n-k}$ into increasing order and ϵ_σ denotes its sign.

In particular we obtain

Proposition. *The map* $j : \bigwedge^k U \to \bigwedge^n U \otimes (\bigwedge^{n-k} U)^*$ *is an isomorphism.*

This statement is a duality statement in the exterior algebra; it is part of a long series of ideas connected with duality. It is also related to the Laplace expansion of a determinant and the expression of the inverse of a given matrix. We leave to the reader to make these facts explicit (see the next section).

2.2 Determinants

Given a linear map $A : U \to V$, the composed map $j : U \to V \to \bigwedge V$ satisfies the universal property and thus induces a homomorphism of algebras, denoted by $\bigwedge A : \bigwedge U \to \bigwedge V$. For every k the map $\bigwedge A$ induces a linear map $\bigwedge^k A : \bigwedge^k U \to \bigwedge^k V$ with

$$\bigwedge^k A(u_1 \wedge u_2 \wedge \cdots \wedge u_k) = Au_1 \wedge Au_2 \wedge \cdots \wedge Au_k.$$

In this way $(\bigwedge U, \bigwedge A)$ is a *functor* from vector spaces to graded algebras (a similar fact holds for the tensor and symmetric algebras).

In particular, for every k the space $\bigwedge^k U$ is a linear representation of the group $GL(U)$.

This is the proper setting for the theory of determinants. One can define the determinant of a linear map $A : U \to U$ of an n-dimensional vector space U as the linear map $\bigwedge^n A$.

Since $\dim \bigwedge^n U = 1$ the linear map $\bigwedge^n A$ is a scalar. One can identify $\bigwedge^n U$ with the base field by choosing a basis of U; any other basis of U gives the same identification if and only if the matrix of the base change is of determinant 1.

Definition. Given an n-dimensional space V, the special linear group $SL(V)$ is the group of transformations A of determinant 1, or $\bigwedge^n A = 1$.

Sometimes one refers to a matrix of determinant 1 as *unimodular*.

More generally, given bases $u_1, \dots, u_m$ for U and $v_1, \dots, v_n$ for V, we have the induced bases on the Grassmann algebra, and we can compute the matrix of $\bigwedge A$ starting from the matrix a_j^i of A. We have $Au_j = \sum_i a_j^i v_i$ and

$$\begin{aligned}\bigwedge^k A(u_{j_1} \wedge \cdots \wedge u_{j_k}) &= Au_{j_1} \wedge \cdots \wedge Au_{j_k} \\ &= \Big(\sum_{i_1} a_{j_1}^{i_1} v_{i_1}\Big) \wedge \Big(\sum_{i_2} a_{j_2}^{i_2} v_{i_2}\Big) \wedge \dots \wedge \Big(\sum_{i_k} a_{j_k}^{i_k} v_{i_k}\Big) \\ &= \sum_{i_1,\dots,i_k} A(i_1, \dots, i_k | j_1, \dots, j_k) v_{i_1} \wedge \cdots \wedge v_{i_k}.\end{aligned}$$

Proposition. *The coefficient $A(i_1, \dots, i_k | j_1, \dots, j_k)$ is the determinant of the minor of the matrix extracted from the matrix of A from the rows of indices $i_1, \dots, i_k$ and the columns of indices $j_1, \dots, j_k$.*

Proof. By expanding the product and collecting terms. □

Given two matrices A, B with product BA, the multiplication formula of the two matrices associated to two exterior powers, $\bigwedge^k (BA) = \bigwedge^k B \circ \bigwedge^k A$, is called *Binet's formula.*

2.3 Symmetry on Tensors

The theory developed is tied with the concepts of symmetry. We have a canonical action of the symmetric group S_n on $U^{\otimes n}$, induced by the permutation action on $U \times U \times \cdots \times U$. Explicitly,

$$\sigma(u_1 \otimes u_2 \otimes \cdots \otimes u_n) = u_{\sigma^{-1}1} \otimes u_{\sigma^{-1}2} \otimes \cdots \otimes u_{\sigma^{-1}n}.$$

We will refer to this action as the *symmetry action* on tensors.[31]

[31] It will be studied intensively in Chapter 9.

Definition 1. The spaces $\Sigma_n(U)$, $A_n(U)$ of symmetric and antisymmetric tensors are defined by

$$\Sigma_n(U) = \{u \in U^{\otimes n} | \sigma(u) = u\}, \ A_n(U) = \{u \in U^{\otimes n} | \sigma(u) = \epsilon(\sigma)u, \ \forall \sigma \in S_n\},$$

($\epsilon(\sigma)$ indicates the sign of σ).

In other words, the space of symmetric tensors is the sum of copies of the trivial representation while the space of antisymmetric tensors is the sum of copies of the sign representation of the symmetric group.

One can explicitly describe bases for these spaces along the lines of §1.

Fix a basis of U which we think of as an ordered alphabet, and take for a basis of $U^{\otimes n}$ the words of length n in this alphabet. The symmetric group permutes these words by reordering the letters, and $U^{\otimes n}$ is thus a *permutation representation.*

Each word is equivalent to a unique word in which all the letters appear in increasing order. If the letters appear with multiplicity $h_1, h_2, \ldots, h_k$, the stabilizer of this word is the product of the symmetric groups $S_{h_1} \times \cdots \times S_{h_k}$. The number of elements in its orbit is $\binom{n}{h_1\, h_2 \ldots h_k}$.

The sum of the elements of such an orbit is a symmetric tensor denoted by $e_{h_1}, \ldots, e_{h_k}$ and these tensors are a basis of $\Sigma_n(U)$. For skew-symmetric tensors we can only use words without multiplicity, since otherwise a transposition fixes such a word but by antisymmetry must change sign to the tensor. The sum of the elements of such an orbit taken with the sign of the permutation is an antisymmetric tensor, and these tensors are a basis of $A_n(U)$.

Theorem. *If the characteristic of the base field F is 0, the projections of $T(U)$ on the symmetric and on the Grassmann algebra are linear isomorphisms when restricted to the symmetric, respectively, the antisymmetric, tensors.*

Proof. Take a symmetric tensor sum of $\binom{n}{h_1\, h_2 \ldots h_k}$ elements of an orbit. The image of all the elements in the same orbit in the symmetric algebra is always the same monomial.

Thus the image of this basis element in the symmetric algebra is the corresponding commutative monomial times the order of the orbit, e.g.,

$$aabb + abab + abba + baab + baba + bbaa \to 6a^2b^2.$$

In char $= 0$, this establishes the isomorphism since it sends a basis to a basis. In order to be more explicit, let us denote by $e_1, e_2, \ldots, e_m$ a basis of U. The element

$$\frac{1}{n!} \sum_{\sigma \in S_n} e_{i_{\sigma(1)}} \otimes e_{i_{\sigma(2)}} \otimes \cdots e_{i_{\sigma(n)}}$$

is a symmetric tensor which, by abuse of notation, we identify with the monomial

$$e_{i_1} e_{i_2} \ldots e_{i_n}$$

to which it corresponds in the symmetric algebra.

Now for the antisymmetric tensors, an orbit gives rise to an antisymmetric tensor if and only if the stabilizer is 1, i.e., if all the $h_i = 1$. Then the antisymmetric tensor corresponding to a word $a_1 a_2 \dots a_n$ is

$$\frac{1}{n!} \sum_{\sigma \in S_n} \epsilon_\sigma a_{\sigma(1)} a_{\sigma(2)} \dots a_{\sigma(n)}.$$

This tensor maps in the Grassmann algebra to

$$a_1 \wedge a_2 \wedge \dots \wedge a_n.$$

It is often customary to identify $e_{i_1} e_{i_2} \dots e_{i_n}$ or $a_1 \wedge a_2 \wedge \dots \wedge a_n$, with the corresponding symmetric or antisymmetric tensor, but of course one loses the algebra structure. □

Let us now notice one more fact. Given a vector $u \in U$ the tensor $u^{\otimes n} = u \otimes u \otimes u \otimes \cdots \otimes u$ is clearly symmetric and identified with $u^n \in S_n(U)$. If $u = \sum_k \alpha_k e_k$ we have

$$u^{\otimes n} = \sum_{h_1 + h_2 + \cdots + h_m = n} \alpha_1^{h_1} \alpha_2^{h_2} \cdots \alpha_m^{h_m} e_{h_1, \dots, h_k}$$

$$u^n \equiv u^{\otimes n}, \quad e_{h_1, \dots, h_k} \equiv \binom{n}{h_1\, h_2 \cdots h_m} e_1^{h_1} e_2^{h_2} \cdots e_m^{h_m}.$$

We notice a formal fact. A homogeneous polynomial function f on U of degree n factors through $u \mapsto u^{\otimes n}$ and a uniquely determined linear map on $\Sigma_n(U)$. In other words, with $P_n(U)$ the space of homogeneous polynomials of degree n on U, we have:

Proposition. *$P_n(U)$ is canonically isomorphic to the dual of $\Sigma_n(U)$.*[32]

The identification is through the factorization

$$U \xrightarrow{u^{\otimes n}} \Sigma_n(U) \xrightarrow{f} F, \quad f(u^{\otimes n}) = \sum_{h_1 + h_2 + \cdots + h_m = n} \alpha_1^{h_1} \alpha_2^{h_2} \cdots \alpha_m^{h_m} f(e_{h_1, \dots, h_k}).$$

One in fact can more generally define:

Definition 2. A *polynomial map* $F : U \to V$ between two vector spaces is a map which in coordinates is given by polynomials.

In particular, one can define homogeneous polynomial maps. We thus have that the map $U \to \Sigma_n(U)$ given by $u \mapsto u^{\otimes n}$ is a polynomial map, homogeneous of degree n and *universal*, in the sense that:

Corollary. *Every homogeneous polynomial map of degree n from U to a vector space V factors through the map $u \to u^{\otimes n}$, with a linear map $\Sigma_n(U) \to V$.*

In characteristic 0 we use instead $u \to u^n$ as a universal map.

[32] This will be useful in the next chapter.

3 Bilinear Forms

3.1 Bilinear Forms

At this point it is important to start introducing the language of bilinear forms in a more systematic way.

We have already discussed the notion of a bilinear mapping $U \times V \to F$. Let us denote the value of such a mapping with the *bra-ket* notation $\langle u|v\rangle$.

Choosing bases for the two vector spaces, the pairing determines a matrix A with entries $a_{ij} = \langle u_i|v_j\rangle$. Using column notation for vectors, the form is given by the formula

$$(u, v) := u^t A v. \tag{3.1.1}$$

If we change the two bases with matrices B, C, and $u = Bu'$, $v = Cv'$, the corresponding matrix of the pairing becomes $B^t AC$.

If we fix our attention on one of the two variables we can equivalently think of the pairing as a linear map $j : U \to \hom(V, F)$ given by $\langle j(u)|v\rangle = \langle u|v\rangle$ or $j(u) : v \mapsto \langle u|v\rangle$.

We have used the bracket notation for our given pairing as well as the duality pairing, thus we can think of a pairing as a linear map from U to V^*.

Definition. We say that a pairing is *nondegenerate* if it induces an isomorphism between U and V^*.

Associated to this idea of bilinear pairing is the notion of *orthogonality*. Given a subspace M of U, its orthogonal is the subspace

$$M^\perp := \{v \in V|\langle u|v\rangle = 0,\ \forall u \in M\}.$$

Remark. The pairing is an isomorphism if and only if its associated (square) matrix is nonsingular. In the case of nondegenerate pairings we have:

(a) $\dim(U) = \dim(V)$.
(b) $\dim(M) + \dim(M^\perp) = \dim(U)$; $(M^\perp)^\perp = M$ for all the subspaces.

3.2 Symmetry in Forms

In particular consider the case $U = V$. In this case we speak of a *bilinear form* on U. For such forms we have a further important notion, that of symmetry:

Definition. We say that a form is *symmetric*, respectively *antisymmetric* or *symplectic*, if $\langle u_1|u_2\rangle = \langle u_2|u_1\rangle$ or, respectively, $\langle u_1|u_2\rangle = -\langle u_2|u_1\rangle$, for all $u_1, u_2 \in U$.

One can easily see that the symmetry condition can be written in terms of the associated map $j : U \to U^* : \langle j(u)|v\rangle = \langle u|v\rangle$.

We take advantage of the identification $U = U^{**}$ and so we have the transpose map $j^* : U^{**} = U \to U^*$.

Lemma. *The form is symmetric if and only if $j = j^*$. It is antisymmetric if and only if $j = -j^*$.*

Sometimes it is convenient to give a uniform treatment of the two cases and use the following language. Let ϵ be 1 or -1. We say that the form is *ϵ-symmetric* if

$$\langle u_1|u_2\rangle = \epsilon\langle u_2|u_1\rangle. \tag{3.2.1}$$

Example 1. The space $\mathrm{End}(U)$ with the form $\mathrm{Tr}(AB)$ is an example of a nondegenerate symmetric bilinear form. The form is nondegenerate since it induces the isomorphism between $U^* \otimes U$ and its dual $U \otimes U^*$ given by exchanging the two factors of the tensor product (cf. 1.7.2).

Example 2. Given a vector space V we can equip $V \oplus V^*$ with a canonical symmetric form, and a canonical antisymmetric form, by the formula

$$\langle (v_1, \varphi_1)|(v_2, \varphi_2)\rangle := \langle \varphi_1|v_2\rangle \pm \langle \varphi_2|v_1\rangle. \tag{3.2.2}$$

On the right-hand side we have used the dual pairing to define the form. We will sometimes refer to these forms as *standard hyperbolic* (resp. *symplectic*) *form*. One should remark that the group $GL(V)$ acts naturally on $V \oplus V^*$ preserving the given forms.

The previous forms are nondegenerate. For an ϵ-symmetric form $(\ ,\ )$ on V we have

$$\{v \in V \mid (v, w) = 0,\ \forall w \in V\} = \{v \in V \mid (w, v) = 0,\ \forall w \in V\}.$$

This subspace is called the *kernel* of the form. The form is nondegenerate if and only if its kernel is 0.

3.3 Isotropic Spaces

For bilinear forms we have the important notion of an *isotropic* and a *totally isotropic subspace*.

Definition. A subspace $V \subset U$ is isotropic if the restriction of the form to V is degenerate and totally isotropic if the restricted form is identically 0.

From the formulas of the previous section it follows:

Proposition. *For a nondegenerate bilinear form on a space U, a totally isotropic subspace V has dimension* $\dim V \leq \dim U/2$.

In particular if $\dim U = 2m$, a maximal totally isotropic subspace has at most dimension m.

Exercise.

(1) Prove that a nondegenerate symmetric bilinear form on a space U of dimension $2m$ has a maximal totally isotropic subspace of dimension m if and only if it is isomorphic to the standard hyperbolic form.
(2) Prove that a nondegenerate antisymmetric bilinear form on a space U exists only if U is of even dimension $2m$. In this case, it is isomorphic to the standard symplectic form.

The previous exercise shows that for a given even dimension there is only one symplectic form up to isomorphism. This is not true for symmetric forms, at least if the field F is not algebraically closed. Let us recall the theory for real numbers. Given a symmetric bilinear form on a vector space over the real number $\mathbb{R}$ there is a basis in which its matrix is diagonal with entries $+1, -1, 0$. The number of 0 is the dimension of the kernel of the form. The fact that the number of $+1$'s (or of -1's) is independent of the basis in which the form is diagonal is Sylvester's law of inertia. The form is *positive* (resp. *negative*) *definite* if the matrix is $+1$ (resp. -1). Since the positive definite form is the usual Euclidean norm, one refers to such space as Euclidean space. In general the number of $+1$'s minus the number of -1's is an invariant of the form called its *signature*.

3.4 Adjunction

For a nondegenerate ϵ-symmetric form we have also the important notion of *adjunction* for operators on U. For $T \in \mathrm{End}(U)$ one defines T^*, the adjoint of T, by

$$(u, T^*v) := (Tu, v). \tag{3.4.1}$$

Using the matrix notation $(u, v) = u^t A v$ we have

(3.4.2)
$$(Tu, v) = (Tu)^t Av = u^t\, T^t Av = u^t AA^{-1}T^t Av = u^t AT^*v \implies T^* = A^{-1}T^t A.$$

Adjunction defines an involution on the algebra of linear operators. Let us recall the definition:

Definition. An involution of an F-algebra R is a linear map $r \mapsto r^*$ satisfying:

$$(rs)^* = s^*r^*, \quad (r^*)^* = r. \tag{3.4.3}$$

In other words $r \mapsto r^*$ is an isomorphism between R and its opposite R^o and it is of order 2. Sometimes it is also convenient to denote an involution by $r \to \bar{r}$.

Let us use the form to identify U with U^* as in 3.1, by identifying u with the linear form $\langle j(u)|v\rangle = (u, v)$.

This identifies $\mathrm{End}(U) = U \otimes U^* = U \otimes U$. With these identifications we have:

$$\begin{gathered}(u \otimes v)w = u(v, w),\ (a \otimes b)(c \otimes d) = a \otimes (b, c)d,\\ (a \otimes b)^* = \varepsilon b \otimes a, \mathrm{tr}(a \otimes b) = (b, a).\end{gathered} \tag{3.4.4}$$

These formulas will be used systematically in Chapter 11.

3.5 Orthogonal and Symplectic Groups

Another important notion is that of the symmetry group of a form. We define *an orthogonal transformation* T for a form to be one for which

$$(3.5.1) \qquad (u, v) = (Tu, Tv), \text{ for all } u, v \in U.$$

Equivalently $T^*T = TT^* = 1$ (if the form is nondegenerate), in matrix notations. From 3.4.2 we have $A^{-1}T^tAT = TA^{-1}T^tA = 1$ or $T^tAT = A, \ TA^{-1}T^t = A^{-1}$.

One checks immediately that if the form is nondegenerate, the orthogonal transformations form a group. For a nondegenerate symmetric form the corresponding group of orthogonal transformations is called the *orthogonal group*. For a nondegenerate skew-symmetric form the corresponding group of orthogonal transformations is called the *symplectic group*.

We will denote by $O(V)$, $Sp(V)$ the orthogonal or symplectic group when there is no ambiguity with respect to the form.

For explicit computations it is useful to have a matrix representation of these groups. For the orthogonal group there are several possible choices, which for a non-algebraically closed field may correspond to non-equivalent symmetric forms and non-isomorphic orthogonal groups.

If A is the identity matrix we get the usual relation $T^* = T^t$. In this case the orthogonal group is

$$O(n, F) := \{X \in GL(n, F) \mid XX^t = 1\}.$$

It is immediate then that for $X \in O(n, F)$ we have $\det(X) = \pm 1$. Together with the orthogonal group it is useful to consider the *special orthogonal group*:

$$SO(n, F) = \{X \in O(n, F) \mid \det X = 1\}.$$

Often one refers to elements in $SO(n, F)$ as *proper orthogonal transformations* while the elements of determinant -1 are called *improper*.

Consider the case of the *standard hyperbolic form 3.2.2* where $U = V \oplus V^*$, $\dim(U) = 2m$ is even.

Choose a basis v_i in V and correspondingly the dual basis v^i in V^*. We see that the matrix of the standard hyperbolic form is $A = \begin{vmatrix} 0 & 1_m \\ 1_m & 0 \end{vmatrix}$ (note that $A = A^{-1}$).

It is useful to consider the orthogonal group for this form, which for non-algebraically closed fields is usually different from the standard form and is called the *split form* of the orthogonal group.

Similarly for the standard symplectic form we have $A = \begin{vmatrix} 0 & 1_m \\ -1_m & 0 \end{vmatrix}$. Notice that $A^{-1} = -A = A^t$. The standard matrix form of the symplectic group is

$$(3.5.2) \qquad Sp(2m, F) := \{X \in GL(m, F) \mid X^tAX = A \text{ or } XAX^t = A\}.$$

In the previous cases, writing a matrix T in block form $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$, we see that

(3.5.3)
$$T^* = \begin{vmatrix} d^t & b^t \\ c^t & a^t \end{vmatrix} \qquad \text{(hyperbolic adjoint)}$$

(3.5.4)
$$T^* = \begin{vmatrix} d^t & -b^t \\ -c^t & a^t \end{vmatrix} \qquad \text{(symplectic adjoint)}$$

One could easily write the condition for a block matrix to belong to the corresponding orthogonal or symplectic group. Rather we work on the real or complex numbers and deduce the Lie algebras of these groups. We have that $(e^{tX})^* = e^{tX^*}$, $(e^{tX})^{-1} = e^{-tX}$, hence:

Proposition. *The Lie algebra of the orthogonal group of a form is the space of matrices with $X^* = -X$.*

From the formulas 3.5.3 and 3.5.4 we get immediately an explicit description of these spaces of matrices, which are denoted by $so(2n, F)$, $sp(2n, F)$.

(3.5.5)
$$so(2n, F) := \left\{ \begin{vmatrix} a & b \\ c & -a^t \end{vmatrix} ;\ b, c \text{ skew-symmetric} \right\}$$

(3.5.6)
$$sp(2n, F) := \left\{ \begin{vmatrix} a & b \\ c & -a^t \end{vmatrix} ;\ b, c \text{ symmetric} \right\}.$$

Thus we have for their dimensions:

$$\dim(so(2n, F)) = 2n^2 - n, \ \dim(sp(2n, F)) = 2n^2 + n.$$

We leave it to the reader to describe $so(2n + 1, F)$. The study of these Lie algebras will be taken up in Chapter 10 in a more systematic way.

3.6 Pfaffian

We want to complete this treatment recalling the properties and definitions of the *Pfaffian* of a skew matrix.

Let V be a vector space of dimension $2n$ with basis e_i. A skew-symmetric form ω_A on V corresponds to a $2n \times 2n$ skew-symmetric matrix A defined by $a_{i,j} := \omega_A(e_i, e_j)$.

According to the theory of exterior algebras we can think of ω_A as the 2-covector[33] given by $\omega_A := \sum_{i<j} a_{i,j} e^i \wedge e^j = 1/2 \sum_{i,j} a_{i,j} e^i \wedge e^j$.

Definition. We define $Pf(A)$ through the formula

(3.6.1)
$$\omega_A^n = n!\, Pf(A) e^1 \wedge e^2 \wedge \cdots \wedge e^{2n}.$$

[33] One refers to an element of $\bigwedge^k V^*$ as a k-covector.

Theorem.

(i) For any invertible matrix B we have

$$Pf(BAB^t) = \det(B)Pf(A). \tag{3.6.2}$$

(ii) $\det(A) = Pf(A)^2$.

Proof. (i) One quickly verifies that $\omega_{BAB^t} = (\wedge^2 B)(\omega_A)$. The identity (i) follows since the linear group acts as algebra automorphisms of the exterior algebra.

The identity (ii) follows from (i) since every skew matrix is of the form BJ_kB^t, where J_k is the standard skew matrix of rank $2k$ given by the direct sum of k blocks of size 2×2 $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. The corresponding covector is $\Omega_{J_k} = \sum_{j=1}^k e^{2j-1} \wedge e^{2j}$,

$$\Omega_{J_k}^n = 0, \text{ if } k < n, \quad \Omega_{J_n}^n = n!\, e^1 \wedge e^2 \wedge \cdots \wedge e^{2n}, \quad Pf(J_n) = 1.$$

The identity is verified directly. □

From formula 3.6.2 it follows that *the determinant of a symplectic matrix is* 1.

Exercise. Let $x_{ij} = -x_{ji}$, $i, j = 1 \dots, 2n$, be antisymmetric variables and X the *generic antisymmetric matrix* with entries x_{ij}. Consider the symmetric group S_{2n} acting on the matrix indices and on the monomial $x_{1\,2}x_{3\,4} \dots x_{2n-1\,2n}$. Up to sign this monomial is stabilized by a subgroup H isomorphic to the semidirect product $S_n \ltimes \mathbb{Z}/(2)^n$. Prove that

$$Pf(X) = \sum_{\sigma \in S_{2n}/H} \epsilon_\sigma x_{\sigma(1)\,\sigma(2)} x_{\sigma(3)\,\sigma(4)} \dots x_{\sigma(2n-1)\,\sigma(2n)}.$$

Prove that the polynomial $Pf(X)$ (in the variables which are the coordinates of a skew matrix) is irreducible.

There is another interesting formula to point out. We will return to this in Chapter 13.

Let us introduce the following notation. Given X as before, $k \le n$, and indices $1 \le i_1 < i_2 < \cdots < i_{2k} \le 2n$ we define the symbol $[i_1, i_2, \dots, i_{2k}]$ to be the Pfaffian of the principal minor of X extracted from the rows and the columns of indices $i_1 < i_2 < \cdots < i_{2k}$. If $\omega_X := \sum_{i<j} x_{i,j} e_i \wedge e_j$, we have

$$\exp(\omega_X) = \sum_k \sum_{i_1 < i_2 < \cdots < i_{2k}} [i_1, i_2, \dots, i_{2k}] e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_{2k-1}} \wedge e_{i_{2k}}. \tag{3.6.3}$$

3.7 Quadratic Forms

Given a symmetric bilinear form on U, we define its associated *quadratic form* by $Q(u) := \langle u|u \rangle$. We see that $Q(u)$ is a homogeneous polynomial of degree 2. We have $Q(u + v) = \langle u + v|u + v \rangle = Q(u) + Q(v) + 2\langle u|v \rangle$ by the bilinearity and symmetry properties. Thus (if 2 is invertible):

$$\frac{1}{2}(Q(u+v) - Q(u) - Q(v)) = \langle u|v\rangle.$$

Notice that this is a very special case of the theory of polarization and restitution, thus quadratic forms or symmetric bilinear forms are equivalent notions (if 2 is invertible).[34]

Suppose we are now given two bilinear forms on two vector spaces U, V. We can then construct a bilinear form on $U \otimes V$ which, on the decomposable tensors, is

$$\langle u_1 \otimes v_1 | u_2 \otimes v_2\rangle = \langle u_1|u_2\rangle\langle v_1|v_2\rangle.$$

We see immediately that if the forms are ϵ_1, ϵ_2-symmetric, then the tensor product is $\epsilon_1\epsilon_2$-symmetric.

One easily verifies that if the two forms are associated to the maps

$$j : U \to U^*,\ k : V \to V^*,$$

then the tensor product form corresponds to the tensor product of the two maps. As a consequence we have:

Proposition. *The tensor product of two nondegenerate forms is nondegenerate.*

Iterating the construction we have a bilinear function on $U^{\otimes m}$ induced by a bilinear form on U.

If the form is symmetric on U, then it is symmetric on all the tensor powers, but if it is antisymmetric, then it will be symmetric on the even and antisymmetric on the odd tensor powers.

Example. We consider the classical example of binary forms ([Hilb]).

We start from a 2-dimensional vector space V with basis e_1, e_2. The element $e_1 \wedge e_2$ can be viewed as a skew-symmetric form on the dual space.

The symplectic group in this case is just the group $SL(2, \mathbb{C})$ of 2×2 matrices with determinant 1.

The dual space of V is identified with the space of linear forms in two variables x, y where x, y represent the dual basis of $e_1,\ e_2$.

A typical element is thus a linear form $ax + by$. The skew form on this space is

$$[ax + by, cx + dy] := ad - bc = \det\begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

This skew form determines corresponding forms on the tensor powers of V^*. We restrict such a form to the symmetric tensors which are identified with the space of binary forms of degree n. We obtain on the space of binary forms of even degree, a nondegenerate symmetric form and, on the ones of odd degree a nondegenerate skew-symmetric form.

The group $SL(2, \mathbb{C})$ acts correspondingly on these spaces by orthogonal or symplectic transformations.

[34] The theory in characteristic 2 can be developed but it is rather more complicated.

One can explicitly evaluate these forms on the special symmetric tensors given by taking the power of a linear form

$$[u \otimes u \otimes \dots u, v \otimes v \otimes \dots \otimes v] = [u, v]^n.$$

If $u = ax + by,\ v = cx + dy$, we get

$$\left[\sum_{i=0}^{n} \binom{n}{i} a^{n-i} b^i x^{n-i} y^i, \sum_{j=0}^{n} \binom{n}{j} c^{n-j} d^j x^{n-j} y^j\right] = (ad - bc)^n.$$

Setting $u_{ij} := [x^{n-i} y^i, x^{n-j} y^j]$ we get

$$\sum_{i=0}^{n} \binom{n}{i} \sum_{j=0}^{n} \binom{n}{j} u_{ij} a^{n-i} b^i c^{n-j} d^j = \sum_{k=0}^{n} \binom{n}{k} (-1)^k (ad)^k (bc)^{n-k}.$$

Comparing the coefficients of the monomials we finally have

$$u_{ij} = 0, \text{ if } i + j \neq n, \ u_{i,n-i} = (-1)^{n-i} \binom{n}{i}^{-1}.$$

3.8 Hermitian Forms

When one works over the complex numbers there are several notions associated with complex conjugation.[35]

Given a vector space U over $\mathbb{C}$ one defines the *conjugate space* $\overline{U}$ to be the group U with the new scalar multiplication $\circ$ defined by

$$\alpha \circ u := \overline{\alpha} u.$$

A linear map from $A : \overline{U} \to V$ to another vector space V, is the same as an *antilinear* map from U to V, i.e., a map A respecting the sum and for which $A(\alpha u) = \overline{\alpha} A(u)$.

The most important concept associated to antilinearity is perhaps that of a Hermitian form and Hilbert space structure on a vector space U.

From the algebraic point of view:

Definition. An Hermitian form is a bilinear map $U \times \overline{U} \to \mathbb{C}$ denoted by (u, v) with the property that (besides the linearity in u and the antilinearity in v) one has

$$(v, u) = \overline{(u, v)}, \ \forall u, v \in U.$$

An Hermitian form is positive if $\|u\|^2 := (u, u) > 0$ for all $u \neq 0$.

A positive Hermitian form is also called a pre-Hilbert structure on U.

[35] One could extend several of these notions to automorphisms of a field, or automorphisms of order 2.

Remark. The Hilbert space condition is not algebraic, but is the completeness of U under the metric $\|u\|$ induced by the Hilbert norm.

In a finite-dimensional space, completeness is always ensured. A Hilbert space always has an orthonormal basis u_i with $(u_i, u_j) = \delta_{ij}$. In the infinite-dimensional case this basis will be infinite and it has to be understood in a topological way (cf. Chapter 8). The most interesting example is the *separable* case in which any orthonormal basis is countable.

A pre-Hilbert space can always be completed to a Hilbert space by the standard method of Cauchy sequences modulo sequences converging to 0.

The group of linear transformations preserving a given Hilbert structure is called the unitary group. In the finite-dimensional case and in an orthonormal basis it is formed by the matrices A such that $A\overline{A}^t = 1$.

The matrix $\overline{A}^t$ is denoted by A^* and called the *adjoint* of A. It is connected with the notion of adjoint of an operator T which is given by the formula $(Tu, v) = (u, T^*v)$. In the infinite-dimensional case and in an orthonormal basis the matrix of the adjoint of an operator is the adjoint matrix.

Given two Hilbert spaces H_1, H_2 one can form the tensor product of the Hilbert structures by the obvious formula $(u \otimes v, w \otimes x) := (u, w)(v, x)$. This gives a pre-Hilbert space $H_1 \otimes H_2$; if we complete it we have a *complete tensor product* which we denote by $H_1 \hat{\otimes} H_2$.

Exercise. If u_i is an orthonormal basis of H_1 and v_j an orthonormal basis of H_2, then $u_i \otimes v_j$ is an orthonormal basis of $H_1 \hat{\otimes} H_2$.

The real and imaginary parts of a positive Hermitian form $(u, v) := S(u, v) + iA(u, v)$ are immediately seen to be bilinear forms on U as a real vector space. $S(u, u)$ is a positive quadratic form while $A(u, v)$ is a nondegenerate alternating form.

An orthonormal basis $u_1, \dots, u_n$ for U (as Hilbert space) defines a basis for U as real vector space given by $u_1, \dots, u_n, iu_1, \dots, iu_n$ which is an orthonormal basis for S and a standard symplectic basis for A, which is thus nondegenerate.

The connection between S, A and the complex structure on U is given by the formula

$$A(u, v) = S(u, iv), \quad S(u, v) = -A(u, iv).$$

3.9 Reflections

Consider a nondegenerate quadratic form Q on a vector space V over a field F of characteristic $\neq 2$. Write $\|v\|^2 = Q(v)$, $(v, w) = \frac{1}{2}(Q(v + w) - Q(v) - Q(w))$.

If $v \in V$ and $Q(v) \neq 0$, we may construct the map $s_v : w \to w - \frac{2(v,w)}{Q(v)} v$. Clearly $s_v(w) = w$ if w is orthogonal to v, while $s_v(v) = -v$.

The map s_v is called the *orthogonal reflection* relative to the hyperplane orthogonal to v. It is an improper orthogonal transformation (of determinant -1) of order two ($s_v^2 = 1$).

Example. Consider the hyperbolic space V of dimension 2. In a hyperbolic basis the matrices $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$ of the orthogonal transformations satisfy

$$\begin{vmatrix} a & c \\ b & d \end{vmatrix}\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} \implies ac = bd = 0,\ cb + ad = 1.$$

From the above formulas one determines the proper transformations $\begin{vmatrix} a & 0 \\ 0 & a^{-1} \end{vmatrix}$, and improper transformations $\begin{vmatrix} 0 & a \\ a^{-1} & 0 \end{vmatrix}$ which are the reflections $s_{(-a,1)}$.

Theorem (Cartan–Dieudonné). *If* $\dim V = m$*, every orthogonal transformation of* V *is the product of at most* m *reflections.*

Proof. Let $T : V \to V$ be an orthogonal transformation. If T fixes a non-isotropic vector v, then T induces an orthogonal transformation in the orthogonal subspace $v^\perp$, and we can apply induction.

The next case is when there is a non-isotropic vector v such that $u := v - T(v)$ is non-isotropic. Then $(v, u) = (v, v) - (v, T(v)), (u, u) = (v, v) - (v, T(v)) - (T(v), v) + (T(v), T(v)) = 2((v, v) - (v, T(v)))$ so that $\frac{2(u,v)}{(u,u)} = 1$ and $s_u(v) = T(v)$. Now $s_u T$ fixes v and we can again apply induction.

This already proves the theorem in the case of a definite form, for instance for the Euclidean space.

The remaining case is that in which every fixed point is isotropic and for every non-isotropic vector v we have $u := v - T(v)$ is isotropic. We claim that if $\dim V \geq 3$, then:

(1) $v - T(v)$ is always isotropic.
(2) V has even dimension $2m$.
(3) T is a proper orthogonal transformation.

Let $s := 1 - T$, and let $V_1 = \ker s$ be the subspace of vectors fixed by T.

Let v be isotropic and consider $v^\perp$ which is a space of dimension $n - 1 > n/2$. Thus $v^\perp$ contains a non-isotropic vector w, and also $\lambda v - w$ is non-isotropic for all λ. Thus by hypothesis

$$0 = Q(s(w)) = Q(s(v - w)) = Q(s(-v - w)).$$

From these equalities follows

$$0 = Q(s(v)) + Q(s(w)) - 2(s(v), s(w)) = Q(s(v)) - 2(s(v), s(w)),$$
$$0 = Q(s(v)) + Q(s(w)) + 2(s(v), s(w)) = Q(s(v)) + 2(s(v), s(w)).$$

Hence $Q(s(v)) = 0$. From $(v - T(v), v - T(v)) = 0$ for all v follows $(v - T(v), w - T(w)) = 0$ for all v, w. From the orthogonality of T follows that

$$(v, 2w) + (v, -T(w)) + (v, -T^{-1}w) = 0$$

for all v, w or $2 - T - T^{-1} = 0$. Hence $2T - T^2 - 1 = 0$ or $s^2 = (1 - T)^2 = 0$.

We have, by hypothesis, that $V_1 = \ker s$ is a totally isotropic subspace, so $2 \dim V_1 \leq \dim V$. Since $s^2 = 0$ we have that $s(V) \subset V_1$, thus $s(V)$ is made of isotropic vectors. Since $\dim V = \dim s(V) + \dim(\ker s)$, it follows that $V_1 = s(V)$ is a maximal totally isotropic subspace and V is of even dimension $2m$. We have that $T = 1 + s$ has only 1 as an eigenvalue and so it has determinant 1, and is thus a proper orthogonal transformation. If S_w is any reflection, we have that $S_w T$ cannot satisfy the same conditions as T, otherwise it would be of determinant 1. Thus we can apply induction and write it as a product of $\leq 2m$ reflections, but this number must be odd, since $S_w T$ has determinant -1. So $S_w T$ is a product of $< 2m$ reflections hence $T = S_w(S_w T)$ is the product of $\leq 2m$ reflections.

In the case $\dim V = 2$, we may assume that the space has isotropic vectors. Hence it is hyperbolic, and we have the formulas of the example. The elements $\begin{vmatrix} 0 & a \\ a^{-1} & 0 \end{vmatrix}$ are reflections, and clearly the proper transformations are products of two reflections. □

3.10 Topology of Classical Groups

Over the complex or real numbers, the groups we have studied are Lie groups. In particular it is useful to understand some of their topology. For the orthogonal groups we have just one group $O(n, \mathbb{C})$ over the complex numbers. Over the real numbers, the group depends on the signature of the form, and we denote by $O(p, q; \mathbb{R})$ the orthogonal group of a form with p entries $+1$ and q entries -1 in the diagonal matrix representation. Let us study the connectedness properties of these groups. Since the special orthogonal group has index 2 in the orthogonal group, we always have for any form and field $O(V) = SO(V) \cup SO(V)\eta$ where η is any given improper transformation. Topologically this is a disjoint union of closed and open subsets, so for the study of the topology we may reduce to the special orthogonal groups.

Let us remark that if T_1, T_2 can be joined by a path to the identity in a topological group, then so can $T_1 T_2$. In fact if $\phi_i(t)$ is a path with $\phi_i(0) = 1$, $\phi_i(1) = T_i$, we take the path $\phi_1(t)\phi_2(t)$.

Proposition. *The groups $SO(n, \mathbb{C})$, $SO(n, \mathbb{R})$ are connected.*

Proof. It is enough to show that any element T can be joined by a path to the identity. If we write T as a product of an even number of reflections, it is enough by the previous remark to treat a product of two reflections. In this case the fixed vectors have codimension 2, and the transformation is essentially a rotation in 2-dimensional space. Then for the complex groups these rotations can be identified (in a hyperbolic basis) with the invertible elements of $\mathbb{C}$ which is a connected set. $SO(2, \mathbb{R})$ is the *circle* group of rotations of the plane, which is clearly connected. □

Things are different for the groups $SO(p, q; \mathbb{R})$. For instance $SO(1, 1; \mathbb{R}) = \mathbb{R}^*$ has two components. In this case we claim that if $p, q > 0$, then $SO(p, q; \mathbb{R})$ has two components. Let us give the main ideas of the proof, leaving the details as exercise.

Suppose $p = 1$. The quadratic form can be written as $Q(x, y) := x^2 - \sum_{i=1}^{q} y_i^2$. The set of elements with $x^2 - \sum_{i=1}^{q} y_i^2 = 1$ has two connected components. The group $SO(1, q; \mathbb{R})$ acts transitively on these vectors, and the stabilizer of a given vector is $SO(q; \mathbb{R})$ which is connected.

If $p > 1$, the set of elements $\sum_{j-1}^{p} x_j^2 = 1 + \sum_{i=1}^{q} y_i^2$ is connected. The stabilizer of a vector is $SO(p-1, q; \mathbb{R})$ so we can apply induction.

Exercise. Prove that the groups $GL(n, \mathbb{C})$, $SL(n, \mathbb{C})$, $SL(n, \mathbb{R})$ are connected while $GL(n, \mathbb{R})$ has two connected components.

More interesting is the question of which groups are simply connected. We have seen in Chapter 4, §3.7 that $SL(n, \mathbb{C})$ is simply connected. Let us analyze $SO(n, \mathbb{C})$ and $Sp(2n, \mathbb{C})$. We use the same method of fibrations developed in Chapter 4, §3.7. Let us treat $Sp(2n, \mathbb{C})$ first. As for $SL(n, \mathbb{C})$, we have that $Sp(2n, \mathbb{C})$ acts transitively on the set W of pairs of vectors u, v with $[u, v] = 1$. The stabilizer of e_1, f_1 is $Sp(2(n-1), \mathbb{C})$. Now let us understand the topology of W. Consider the projection $(u, v) \mapsto u$ which is a surjective map to $\mathbb{C}^{2n} - \{0\}$ with fiber at e_1 the set $f_1 + ae_1 + \sum_{i>2} a_i e_i + b_i f_i$, a contractible space.

This is a special case of Chapter 4, §3.7.1 for the groups $Sp(2(n-1), \mathbb{C}) \subset H \subset Sp(2n, \mathbb{C})$ where H is the stabilizer, in $Sp(2n, \mathbb{C})$, of the vector e_1.

Thus $\pi_1(Sp(2n, \mathbb{C})) = \pi_1(Sp(2(n-1), \mathbb{C}))$. By induction we have:

Theorem. *$Sp(2n, \mathbb{C})$ is simply connected.*

As for $SO(n, \mathbb{C})$ we make two remarks. In Chapter 8, §6.2 we will see a fact, which can be easily verified directly, implying that $SO(n, \mathbb{C})$ and $SO(n, \mathbb{R})$ are homotopically equivalent. We discuss $SO(n, \mathbb{R})$ at the end of §5, proving that $\pi_1(SO(n, \mathbb{R})) = \mathbb{Z}/(2)$.

At this point, if we restrict our attention only to the infinitesimal point of view (that is, the Lie algebras), we could stop our search for classical groups.

This, on the other hand, misses a very interesting point. The fact that the special orthogonal group is not simply connected implies that, even at the infinitesimal level, not all the representations of its Lie algebra arise from representation of this group.

In fact we miss the rather interesting *spin representations*. In order to discover them we have to construct the *spin group*. This will be done, as is customary, through the analysis of Clifford algebras, to which the next section is devoted.

4 Clifford Algebras

4.1 Clifford Algebras

Given a quadratic form on a space U we can consider the ideal J of $T(U)$ generated by the elements $u^{\otimes 2} - Q(u)$.

Definition 1. The quotient algebra $T(U)/J$ is called the *Clifford algebra of the quadratic form.*

Notice that the Clifford algebra is a generalization of the Grassmann algebra, which is obtained when $Q = 0$. We will denote it by $Cl_Q(U)$ or by $Cl(U)$ when there is no ambiguity for the quadratic form.

By definition the Clifford algebra is the universal solution for the construction of an algebra R and a linear map $j : U \to R$ with the property that $j(u)^2 = Q(u)$. Let us denote by $(u, v) := 1/2(Q(u+v) - Q(u) - Q(v))$ the bilinear form associated to Q. We have in the Clifford algebra:

$$v, w \in U, \implies vw + wv = (v+w)^2 - v^2 - w^2 = 2(v, w). \tag{4.1.1}$$

In particular if v, w are orthogonal they anticommute in the Clifford algebra $vw = -wv$.

If $G \supset F$ is a field extension, the given quadratic form Q on U extends to a quadratic form Q_G on $U_G := U \otimes_F G$. By the universal property it is easy to verify that

$$Cl_{Q_G}(U_G) = Cl_Q(U) \otimes_F G.$$

There are several efficient ways to study the Clifford algebra. We will go through the theory of superalgebras ([ABS]).

One starts by remarking that, although the relations defining the Clifford algebra are not homogeneous, they are of even degree. In other words, decompose the tensor algebra $T(U) = T_0(U) \oplus T_1(U)$, where $T_0(U) = \oplus_k \otimes^{2k} U$, $T_1(U) = \oplus_k \otimes^{2k+1} U$, as a direct sum of its even and odd parts. The ideal I defining the Clifford algebra decomposes also as $I = I_0 \oplus I_1$ and $Cl(U) = T_0(U)/I_0 \oplus T_1(U)/I_1$. This suggests the following:

Definition 2. A superalgebra is an algebra A decomposed as $A_0 \oplus A_1$, with $A_i A_j \subset A_{i+j}$, where the indices are taken modulo 2.[36]

A superalgebra is thus graded modulo 2. For a homogeneous element a, we set $d(a)$ to be its degree (modulo 2). We have the obvious notion of (graded) homomorphism of superalgebras. Often we will write $A_0 = A^+$, $A_1 = A^-$.

Given a superalgebra A, a *superideal* is an ideal $I = I_0 \oplus I_1$, and the quotient is again a superalgebra.

More important is the notion of a *super-tensor product* of associative superalgebras.

Given two superalgebras A, B we define a superalgebra:

$$A \hat{\otimes} B := (A_0 \otimes B_0 \oplus A_1 \otimes B_1) \oplus (A_0 \otimes B_1 \oplus A_1 \otimes B_0),$$

$$(a \otimes b)(c \otimes d) := (-1)^{d(b)d(c)} ac \otimes bd. \tag{4.1.2}$$

It is left to the reader to show that this defines an associative superalgebra.

[36] Superalgebras have been extensively used by physicists in the context of the theory of elementary particles. In fact several basic particles like electrons are *Fermions*, i.e., they obey a special statistics which is suitably translated with the spin formalism. Some further proposed theories, such as supersymmetry, require the systematic use of superalgebras of operators.

Exercise. A superspace is just a $\mathbb{Z}/(2)$ graded vector space $U = U_0 \oplus U_1$, we can grade the endomorphism ring $\mathrm{End}(U)$ in an obvious way:

$$\mathrm{End}(U)_0 = \mathrm{End}(U_0) \oplus \mathrm{End}(U_1), \quad \mathrm{End}(U)_1 = \hom(U_0, U_1) \oplus \hom(U_1, U_0).$$

Prove that, given two superspaces U and V, we have a natural structure of superspace on $U \otimes V$ such that $\mathrm{End}(U \otimes V)$ is isomorphic to $\mathrm{End}(U) \hat{\otimes} \mathrm{End}(V)$ as superalgebra.

In this vein of definitions we have the notion of a *supercommutator*, which on homogeneous elements is

$$\{a, b\} := ab - (-1)^{d(a)d(b)} ba. \tag{4.1.3}$$

Definitions 4.1.2,3 are then extended to all elements by bilinearity.

Accordingly we say that a superalgebra is *supercommutative* if $\{a, b\} = 0$ for all the elements. For instance, the Grassmann algebra is supercommutative.[37]

The connection between supertensor product and supercommutativity is in the following:

Exercise. Given two graded maps $i, j : A, B \to C$ of superalgebras such that the images supercommute we have an induced map $A \hat{\otimes} B \to C$ given by $a \otimes b \to i(a)j(b)$.

Exercise. Discuss the notions of *superderivation* ($D(ab) = D(a)b + (-1)^{d(a)} aD(b)$), supermodule, and super-tensor product of such supermodules.

We can now formulate the main theorem:

Theorem 1. *Given a vector space U with a quadratic form and an orthogonal decomposition $U = U_1 \oplus U_2$, we have a canonical isomorphism*

$$Cl(U) = Cl(U_1) \hat{\otimes} Cl(U_2). \tag{4.1.4}$$

Proof. First consider the linear map $j : U \to Cl(U_1) \hat{\otimes} Cl(U_2)$ which on U_1 is $j(u_1) := u_1 \otimes 1$ and on U_2 is $j(u_2) = 1 \otimes u_2$.

It is easy to see, by all of the definitions given, that this map satisfies the universal property for the Clifford algebra and so it induces a map $\bar{j} : Cl(U) \to Cl(U_1) \hat{\otimes} Cl(U_2)$.

Now consider the two inclusions of U_1, U_2 in U which define two maps $Cl(U_1) \to Cl(U)$, $Cl(U_2) \to Cl(U)$.

It is again easy to see (since the two subspaces are orthogonal) that the images supercommute. Hence we have a map $\bar{i} : Cl(U_1) \hat{\otimes} Cl(U_2) \to Cl(U)$.

On the generating subspaces $U, U_1 \otimes 1 \oplus 1 \otimes U_2$, the maps $\bar{j}, \bar{i}$ are isomorphisms inverse to each other. The claim follows. $\square$

---

[37] Sometimes in the literature just the term commutative, instead of supercommutative, is used for superalgebras.

For a 1-dimensional space with basis u and $Q(u) = \alpha$, the Clifford algebra has basis $1, u$ with $u^2 = \alpha$.

Thus we see by induction that:

Lemma 1. *If we fix an orthogonal basis $e_1, \dots, e_n$ of the vector space U, then the 2^n elements $e_{i_1} e_{i_2} \dots e_{i_k}$, $i_1 < i_2 < \dots < i_k$ give a basis of $Cl(U)$.*

For an orthogonal basis e_i we have the defining commuting relations $e_i^2 = Q(e_i)$, $e_i e_j = -e_j e_i$, $i \neq j$. If the basis is orthonormal we have also $e_i^2 = 1$.

It is also useful to present the Clifford algebra in a hyperbolic basis, i.e., the Clifford algebra of the standard quadratic form on $V \oplus V^*$ which is convenient to renormalize by dividing by 2, so that $Q((v, \phi)) = \langle \phi \mid v \rangle$. If $V = \mathbb{C}^n$ we denote this Clifford algebra by C_{2n}.

The most efficient way to treat C_{2n} is to exhibit the exterior algebra $\bigwedge V$ as an irreducible module over $Cl(V \oplus V^*)$, so that $Cl(V \oplus V^*) = \operatorname{End}(\bigwedge V)$. This is usually called the *spin formalism.*

For this let us define two linear maps i, j from V, V^* to $\operatorname{End}(\bigwedge V)$:

$$i(v)(u) := v \wedge u,$$

$$j(\varphi)(v_1 \wedge v_2 \dots \wedge v_k) := \sum_{t=1}^{k} (-1)^{t-1} \langle \varphi | v_t \rangle v_1 \wedge v_2 \dots \check{v}_t \dots \wedge v_k, \tag{4.1.5}$$

where $\check{v}_t$ means that this term has been omitted.

Notice that the action of $i(v)$ is just the left action of the algebra $\bigwedge V$ while $j(\varphi)$ is the superderivation induced by the contraction by φ on V.

One immediately verifies

$$i(v)^2 = j(\varphi)^2 = 0, \quad i(v)j(\varphi) + j(\varphi)i(v) = \langle \varphi | v \rangle. \tag{4.1.6}$$

Theorem 2. *The map $i + j : V \oplus V^* \to \operatorname{End}(\bigwedge V)$ induces an isomorphism between the algebras $Cl(V \oplus V^*)$ and $\operatorname{End}(\bigwedge V)$ (as superalgebras).*

Proof. From 4.1.6 we have that $i + j$ satisfies the universal condition defining the Clifford algebra for $1/2$ of the standard form.

To prove that the resulting map is an isomorphism between $Cl(V \oplus V^*)$ and $\operatorname{End}(\bigwedge V)$ one has several options. One option is to show directly that $\bigwedge V$ is an irreducible module under the Clifford algebra and then remark that, if $n = \dim(V)$ then $\dim Cl(V \oplus V^*) = 2^{2n} = \dim \operatorname{End}(\bigwedge V)$. The second option is to analyze first the very simple case of $\dim V = 1$ where we verify the statement by direct inspection. Next we decompose the exterior algebra as the tensor product of the exterior algebras on 1-dimensional spaces. Each such space is a graded irreducible 2-dimensional module over the corresponding Clifford algebra, and we get the identity by taking super-tensor products. $\square$

The Clifford algebra in the odd-dimensional case is different. Let us discuss the case of a standard orthonormal basis, $e_1, e_2, \dots, e_{2n+1}$. Call this Clifford algebra C_{2n+1}.

Lemma 2. *The element $c := e_1 e_2 \dots e_{2n+1}$ is central and $c^2 = (-1)^n$.*

Proof. From the defining commutation relations we immediately see that.

$$e_i c = e_i e_1 e_2 \dots e_{i-1} e_i \dots e_{2n+1} = (-1)^{i-1} e_1 e_2 \dots e_{i-1} e_i^2 \dots e_{2n+1};$$
$$c e_i = e_1 e_2 \dots e_{i-1} e_i \dots e_{2n+1} e_i = (-1)^{2n+1-i} e_1 e_2 \dots e_{i-1} e_i^2 \dots e_{2n+1}.$$

As for c^2, we have again by commutation relations:

$$\begin{aligned} e_1 e_2 \dots e_{2n+1} e_1 e_2 \dots e_{2n+1} &= e_1^2 e_2 \dots e_{2n+1} e_2 \dots e_{2n+1} \\ &= e_2 \dots e_{2n+1} e_2 \dots e_{2n+1} = -e_2^2 e_3 \dots e_{2n+1} e_3 \dots e_{2n+1} \\ &= -e_3 \dots e_{2n+1} e_3 \dots e_{2n+1} = \dots = (-1)^n. \end{aligned}$$ □

Take the Clifford algebra C_{2n} on the first $2n$ basis elements. Since $e_{2n+1} = e_{2n} e_{2n-1} \dots e_1 c$, we see that $C_{2n+1} = C_{2n} + C_{2n} c$. We have proved that:

Theorem 3. *$C_{2n+1} = C_{2n} \otimes_F F[c]$. If $F[c] = F \oplus F$, which happens if $(-1)^n$ has a square root in F, then C_{2n+1} is isomorphic to $C_{2n} \oplus C_{2n}$.*

4.2 Center

One may apply the previous results to have a first study of Clifford algebras as follows. Let U be a quadratic space of dimension n over a field F and let $G = \overline{F}$ be an algebraic closure. Then $U \otimes_F \overline{F}$ is hyperbolic, so we have that $Cl_Q(U) \otimes_F G$ is the Clifford algebra of a hyperbolic form. Thus if $n = 2k$ is even, $Cl_Q(U) \otimes_F G = M_{2^k}(G)$ is the algebra of matrices over G, while if $n = 2k + 1$ is odd, we have $Cl_Q(U) \otimes_F G = M_{2^k}(G) \oplus M_{2^k}(G)$. We can draw some consequences from this statement using the following simple lemma:

Lemma. *Let R be an algebra over a field F with center Z and let G be a field extension of F. Then the center of $R \otimes_F G$ is $Z \otimes_F G$.*

Proof. Let u_i be a basis of G over F. Consider an element $s := \sum_i r_i \otimes u_i \in R \otimes_F G$. To say that it is in the center implies that for $r \in R$ we have $0 = rs - sr = \sum_i (r r_i - r_i r) \otimes u_i$. Hence $r r_i - r_i r = 0$ for all i and $r_i \in Z$ for all i. The converse is also obvious. □

As a consequence we have that:

Proposition. *The center of $Cl_Q(U)$ is F if n is even. If $n = 2k+1$, then the center is $F + cF$, where $c := u_1 u_2 \dots u_{2k+1}$ for any orthogonal basis $u_1, u_2, \dots, u_{2k+1}$ of U.*

Proof. First, one uses the fact that the center of the algebra of matrices over a field is the field itself. Second, up to multiplying c by a nonzero scalar, we may assume that the basis is orthonormal. Finally we are reduced to the theory developed in the previous paragraph. □

Remark. In the case of odd dimension the center may either be isomorphic to $F \oplus F$ or to a quadratic extension field of F. This depends on whether the element $c^2 \in F^*$ is a square or not (in F^*).

4.3 Structure Theorems

It is also important to study the Clifford algebras $C(n)$, $C'(n)$ over $\mathbb{R}$ for the negative and positive definite forms $-\sum_i^n x_i^2$, $\sum_i^n x_i^2$.

For $n = 1$ we get $C(1) := \mathbb{R}[x]/(x^2+1) = \mathbb{C}$, $C'(1) := \mathbb{R}[x]/(x^2-1) = \mathbb{R}\oplus\mathbb{R}$.

For $n = 2$ we get $C(2) := \mathbb{H}$, the quaternions, since setting $i := e_1$, $j := e_2$, the defining relations are $i^2 = -1$, $j^2 = -1, ij = -ji$.

$C'(2)$ is isomorphic to $M_2(\mathbb{R})$, the 2×2 matrices over $\mathbb{R}$. In fact in this case the defining relations are $i^2 = 1$, $j^2 = 1, ij = -ji$, which are satisfied by the matrices:

$$i := \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix}, \quad j := \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}, \quad -ji = ij = \begin{vmatrix} 0 & 1 \\ -1 & 0 \end{vmatrix}.$$

To study Clifford algebras in general we make a remark. Let Q be a nondegenerate quadratic form on the space U. Decompose U into an orthogonal direct sum $U = U_1 \oplus U_2$, with $\dim U_1 = 2$. Denote by Q_1, Q_2 the induced quadratic forms on U_1, U_2.

Fix an orthogonal basis u, v of U_1 and an orthogonal basis $e_1, \dots, e_k$ of U_2. Let $u^2 = \alpha$, $v^2 = \beta$, $\lambda_i := e_i^2 = Q(e_i)$ and set $\delta := \alpha\beta \neq 0$, since the form is nondegenerate. We have the commutation relations:

$$ue_i = -e_i u, \quad ve_i = -ve_i, \quad uve_i = e_i uv, \quad uv = -vu, \quad (uv)^2 = -\delta.$$

Moreover, $uuv = -uvu$, $vuv = -uvv$. If we set $f_i := uve_i$ we deduce the following commutation relations:

$$f_i f_j = -f_j f_i, \ i \neq j; \quad f_i^2 = -\delta\lambda_i; \quad uf_i = f_i u; \quad vf_i = f_i v.$$

From these commutation relations we deduce that the subalgebra $F(f_1, \dots, f_k)$ of the Clifford algebra generated by the elements f_i is a homomorphic image of the Clifford algebra on the space U_2 but relative to the quadratic form $-\delta Q_2$. Moreover, this subalgebra commutes with the subalgebra $F(u, v) = Cl_{Q_1}(U_1)$. We have thus a homomorphism:

$$Cl_{Q_1}(U_1) \otimes Cl_{-\delta Q_2}(U_2) \xrightarrow{i} Cl_Q(U).$$

Proposition. *The map i is an isomorphism.*

Proof. Since the dimensions of the two algebras are the same, it is enough to show that the map is surjective, i.e., that the elements u, v, f_i generate $C_Q(U)$. This is immediate since $e_i = -\delta^{-1}uvf_i$. □

We can apply this proposition to the Clifford algebras $C(n)$, $C'(n)$ over $\mathbb{R}$ for the negative and positive definite quadratic form. In this case $\delta = 1$, so we get

$$C(n) = \mathbb{H} \otimes C'(n-2), \quad C'(n) = M_2(\mathbb{R}) \otimes C(n-2).$$

Iterating, we get the recursive formulas:

$$C(n) = C(4k) \otimes C(n-4k), \quad C'(n) = C'(4k) \otimes C'(n-4k).$$

In order to complete the computations we need the following simple facts:

Lemma.

(1) If A, B are two F-algebras and $M_h(A)$, $M_k(B)$ denote matrix algebras (over A, B respectively), we have

$$M_h(A) \otimes_F M_k(B) = M_{hk}(A \otimes_F B).$$

(2)
$$\mathbb{H} \otimes_{\mathbb{R}} \mathbb{H} = M_4(\mathbb{R}), \quad \mathbb{H} \otimes_{\mathbb{R}} \mathbb{C} = M_2(\mathbb{C}).$$

Proof. (1) is an easy exercise. (2) can be shown as follows. We have a homomorphism $\psi : \mathbb{H} \otimes_{\mathbb{R}} \mathbb{H} \to \mathrm{End}_{\mathbb{R}} \mathbb{H}$ given by $\psi(a \otimes b)(c) := ac\bar{b}$ which one easily verifies is an isomorphism.

For the second consider $\mathbb{C} \subset \mathbb{H}$ in the usual way, and consider $\mathbb{H}$ as vector space over $\mathbb{C}$ by right multiplication. We have a homomorphism $\phi : \mathbb{H} \otimes_{\mathbb{R}} \mathbb{C} \to \mathrm{End}_{\mathbb{C}} \mathbb{H}$ given by $\phi(a \otimes b)(c) := acb$ which one easily verifies is an isomorphism. □

We deduce the following list for $C(n)$, $n = 0, 1, 2 \ldots, 8$:

$$\mathbb{R},\ \mathbb{C},\ \mathbb{H},\ \mathbb{H} \oplus \mathbb{H},\ M_2(\mathbb{H}),\ M_4(\mathbb{C}),\ M_8(\mathbb{R}),\ M_8(\mathbb{R}) \oplus M_8(\mathbb{R}),\ M_{16}(\mathbb{R})$$

and *periodicity 8:* $C(n) = M_{16}(C(n-8))$.[38]

The list of the same algebras but over the complex numbers

$$C(n) \otimes_{\mathbb{R}} \mathbb{C} := C_{\mathbb{C}}(n) = C'_{\mathbb{C}}(n) = C'(n) \otimes_{\mathbb{R}} \mathbb{C}$$

is deduced by tensoring with $\mathbb{C}$ as

$$\mathbb{C},\ \mathbb{C} \oplus \mathbb{C},\ M_2(\mathbb{C}),\ M_2(\mathbb{C}) \oplus M_2(\mathbb{C}),\ M_4(\mathbb{C})$$

and *periodicity 2:* $C_{\mathbb{C}}(n) = M_2(C_{\mathbb{C}}(n-2))$. Of course over the complex numbers the form is hyperbolic, so we get back the result we already knew by the spin formalism.

4.4 Even Clifford Algebra

It is also important to study the degree 0 part of the Clifford algebra, i.e., $Cl_Q^+(U)$, since it will appear in the definition of the spin groups. This is the subalgebra of $Cl_Q(U)$ spanned by products $u_1 u_2 \ldots u_{2k}$ of an even number of vectors in U. Let $\dim U = s + 1$, and fix an orthogonal basis which we write $u, e_1, \ldots, e_s$ to stress the decomposition $U = Fu \oplus U'$. Let Q' be the restriction of Q to U'. Define $f_i := ue_i$. The f_i are elements of $Cl_Q^+(U)$. Let $\delta := u^2 = Q(u)$.

We have from the commutation relations:

$$i \neq j, \quad f_i f_j = ue_i ue_j = -ue_i e_j u = ue_j e_i u = -ue_j ue_i = -f_j f_i,$$
$$f_i^2 = ue_i ue_i = -u^2 e_i^2.$$

It follows then that the elements f_i satisfy the commutation relations for the Clifford algebra of the space U' equipped with the form $-\delta Q'$. Thus we have a homomorphism $i : Cl_{-\delta Q'}(U') \to Cl_Q^+(U)$.

[38] This is related to the Bott periodicity theorem for homotopy groups and K-theory.

Proposition 1. *The map i is an isomorphism.*

Proof. Since the dimensions of the two algebras are the same it is enough to show that the map is surjective, i.e., that the elements f_i generate $C_Q^+(U)$. Let $\alpha = u^2 \neq 0$. We have that $e_i = \alpha^{-1}uf_i$. Moreover $f_iu = ue_iu = -uf_i$. Take an even monomial in the elements $u, e_1, \dots, e_s$, such that u appears h times and the e_i appear $2k - h$ times. Substitute $e_i = \alpha^{-1}uf_i$. Up to a nonzero scalar, we obtain a monomial in u, f_i in which u appears $2k$ times. Using the commutation relations we can bring u^{2k} in front and then use the fact that this is a scalar to conclude. □

If we apply the previous result to $C(n)$, we obtain $C^+(n) \equiv C(n-1)$.

As preparation for the theory of the spin group let us make an important remark.

Let us take a space V with a nondegenerate symmetric form (a, b) and let $C(V)$ be the Clifford algebra for $1/2(a, b)$.[39] Consider the space $L := \{ab \mid a, b \in V\} \subset C^+(V)$ and the map $a : \bigwedge^2 V \to C^+(V)$, $a(v \wedge w) := [v, w]/2$. Fixing an orthogonal basis e_i for V we have $a(e_i \wedge e_j) = e_ie_j$, $i < j$. From Lemma 1, 4.1 it then follows that a is injective, so we identify $\bigwedge^2 V \subset C^+(V)$.

Proposition 2. $L = F \oplus \bigwedge^2 V$ *is a Lie subalgebra of* $C^+(V)$, $[L, L] = \bigwedge^2 V$. *Under the adjoint action, V is an L-submodule for which $\bigwedge^2 V$ is isomorphic to the Lie algebra of $SO(V)$.*

Proof. $ab + ba = (a, b)$ means that $ab = \frac{[a,b]}{2} + (a, b)/2$ so $ab = a \wedge b + (a, b)/2$, $\forall a, b \in V$. It follows that $L = F \oplus \bigwedge^2 V$ is the span of the products ab, $a, b \in V$. Next, given $a, b, c, d \in V$ we have (applying the relations):

$$cdab = abcd + [(b, d)ac + (a, d)cb - (a, c)db - (b, c)ad].$$

$$\text{Hence}\quad [cd, ab] = [(b, d)ac + (a, d)cb - (a, c)db - (b, c)ad]$$
$$= 1/2\{(b, d)[a, c] + (a, d)[c, b] - (a, c)[d, b] - (b, c)[a, d]\}.$$

So L is a Lie algebra and $[L, L] \subset \bigwedge^2 V$. Furthermore

$$[c \wedge d, a \wedge b] = (b, d)a \wedge c + (a, d)c \wedge b - (a, c)d \wedge b - (b, c)a \wedge d. \tag{4.4.1}$$

$$[ab, c] = (b, c)a - (a, c)b, \quad [a \wedge b, c] = (b, c)a - (a, c)b. \tag{4.4.2}$$

Then

$$([a \wedge b, c], d) = (b, c)(a, d) - (a, c)(b, d)$$

is skew-symmetric as a function of c, d. This shows that L acts as $so(V)$. An element of L is in the kernel of the action if and only if it is a scalar. Of course we get F from the elements a^2.

Since $\bigwedge^2 V$ and $so(V)$ have the same dimension we must have the isomorphism. □

[39] The normalization 1/2 is important to eliminate a lot of factors of 2 and also reappears in the spin formalism.

4.5 Principal Involution

The Clifford algebra has an *involution* (cf. 3.4.3).

Consider the embedding of V in $Cl_Q(V)^o$. This embedding still satisfies the property $i(v)^2 = Q(v)$, and hence it extends to a homomorphism $* : Cl_Q(V) \to Cl_Q(V)^o$. In other words there is an *antihomomorphism* $* : Cl_Q(V) \to Cl_Q(V)$ such that $v^* = v, \forall v \in V$. The homomorphism $r \to (r^*)^*$ is the identity on V and therefore, by the universal property, it must be the identity of $Cl_Q(V)$. This proves the existence of an involution, called the *principal involution* on $Cl_Q(V)$, such that $v^* = v, \forall v \in V$.

Remark. The subalgebra $Cl_Q^+(V)$ is stable under the principal involution.

In fact, for the elements defined in 4.4, we have $f_i^* = (ue_i)^* = e_i^* u^* = e_i u = -f_i$.

This formula could of course be defined in general; one could have defined an involution setting $v^* := -v$. For $C(1) = \mathbb{C}$ this last involution is just complex conjugation. For $C(2) = \mathbb{H}$ it is the standard quaternion involution $q = a + b\underline{i} + c\underline{j} + d\underline{k} \mapsto \overline{q} := a - b\underline{i} - c\underline{j} - d\underline{k}$.

5 The Spin Group

5.1 Spin Groups

The last of the groups we want to introduce is the spin group. Consider again the Clifford algebra of a quadratic form Q on a vector space V over a field F. Write $\|v\|^2 = Q(v)$, $(v, w) = \frac{1}{2}(Q(v + w) - Q(v) - Q(w))$.

Definition 1. The Clifford group $\Gamma(V, Q)$ is the subgroup of invertible elements $x \in Cl_Q(V)^*$ with $xVx^{-1} = V$. The Clifford group $\Gamma^+(V, Q)$ is the intersection $\Gamma^+(V, Q) := \Gamma(V, Q) \cap Cl_Q^+(V)$.

Let $x \in \Gamma(V, Q)$ and $u \in V$. We have $(xux^{-1})^2 = xu^2x^{-1} = Q(u)$. Therefore the map $u \mapsto xux^{-1}$ is an orthogonal transformation of V. We have thus a homomorphism $\pi : \Gamma(V, Q) \to O(V)$. If $v, w \in V$ and $Q(v) \neq 0$, we have that v is invertible and

$$v^{-1} = \frac{v}{Q(v)}, \quad vw + wv = 2(v, w), \quad vwv^{-1} = \frac{2(v, w)}{Q(v)} v - w.$$

The map $w \to w - \frac{2(v,w)}{Q(v)} v$ is the orthogonal reflection r_v relative to the hyperplane orthogonal to v. Thus conjugation by v induces $-r_v$. If $\dim V$ is even, it is an improper orthogonal transformation.

We have that $v \in \Gamma(V, Q)$ and that a product $v_1 v_2 \dots v_{2k}$ of an even number of such v induces an even product of reflections, hence a special orthogonal transformation in V. By the Cartan–Dieudonné theorem, any special orthogonal transformation can be so induced. Similarly, a non-special (or improper) orthogonal transformation is the product of an odd number of reflections. We obtain:

Proposition 1. *The image of the homomorphism $\pi : \Gamma(V, Q) \to O(V)$ contains $SO(V)$. If $\dim V$ is even, π is surjective. If $\dim V$ is odd, the image of π is $SO(V)$.*

Proof. Only the last statement has not been checked. If in the odd case π were surjective, there is an element $x \in \Gamma(V, Q)$ with $xvx^{-1} = -v, \forall v \in V$. It follows that $xcx^{-1} = -c$ which is absurd, since c is in the center. □

The kernel of π is composed of elements which commute with the elements of V. Since these elements generate the Clifford algebra, we deduce that $\operatorname{Ker} \pi$ is the set Z^* of invertible elements of the center of $Cl_Q(V)$. We have thus the exact sequence:

(5.1.1) $$1 \to Z^* \to \Gamma(V, Q) \xrightarrow{\pi} O(V).$$

If $n = \dim V$ is even the center is the field F, otherwise it is the set $\alpha + \beta c$, $\alpha, \beta \in F$ and $c := u_1 u_2 \dots u_{2k+1}$ for any given orthogonal basis $u_1, u_2, \dots, u_{2k+1}$ of U (cf. 4.2). Let us consider now $\Gamma^+(V, Q)$, its intersection with the center is clearly F^*. Since every element of $O(V)$ is a product of reflections, we deduce that every element γ of $\Gamma(V, Q)$ is a product $\alpha v_1 v_2 \cdots v_k$, of an element $\alpha \in Z^*$ and elements $v_i \in V$. If $\dim V$ is odd, by Proposition 1 we can assume that this last product is even $(k = 2h)$. If $\gamma \in \Gamma^+(V, Q)$ we deduce $\alpha \in F^*$. If $\dim V$ is even we have $Z^* = F^*$ and so again, if $\gamma \in \Gamma^+(V, Q)$ we deduce that $k = 2h$ is even. The image of γ in $O(V)$ is contained in $SO(V)$, and we have an exact sequence:

(5.1.2) $$1 \to F^* \to \Gamma^+(V, Q) \xrightarrow{\pi} SO(V) \to 1.$$

Let us compute $N(r) := rr^*$ when $r = v_1 v_2 \cdots v_j \in \Gamma(V, Q)$. We have $r^* = v_j v_{j-1} \dots v_1$ and by easy induction we obtain

(5.1.3) $$rr^* = Q(v_1)Q(v_2)\dots Q(v_j) \in F^*.$$

Lemma. *The map $r \to N(r) = rr^*$ restricted to $\Gamma^+(V, Q)$ is a homomorphism to F^*.*

Proof. For two elements $r = v_1 v_2 \dots v_j, s = u_1 u_2 \dots u_h$ we have $N(r) = \prod_i Q(v_i)$, $N(s) = \prod_k Q(u_k)$ and $N(rs) = \prod Q(v_i) \prod_k Q(u_k)$. Every element of $\Gamma^+(V, Q)$ is of the form $\alpha v_1 v_2 \dots v_{2j}$, $\alpha \in F^*$ and the claim follows from 5.1.3. □

Proposition 2. *(a) The Lie algebra of $\Gamma^+(V, Q)$ is the Lie algebra L of 4.4.*
(b) Given $ab \in L$ we have $N(\exp(ab)) = \exp(2(a, b))$.

Proof. (a) First, taking $ab \in L$ we claim that $\exp(t\,ab) \in \Gamma^+(V, Q), \forall t$.

Clearly $\exp(t\,ab) \in Cl^+(V)$, on the other hand, by Proposition 4.4, we have $[ab, V] \subset V$. Hence $\exp(t\,ab) V \exp(-t\,ab) \subset V$, and $\exp(t\,ab)$ is by definition in the Clifford group. To prove that L is the entire Lie algebra L' of $\Gamma^+(V, Q)$, we remark that L and L' induce the same Lie algebra $so(V)$ by acting on V. In both cases the kernel of the Lie algebra action is the scalars.

(b) $\exp(ab)^* = \exp(ba) = \exp(-ab + 2(a, b)) = \exp(-ab)\exp(2(a, b))$. □

Example. For $V = \mathbb{R}^3$ with the negative definite quadratic form, we have seen that $C(3) = \mathbb{H} \oplus \mathbb{H}$ and that $C^+(3) = \mathbb{H}$. From the preceding analysis, we may explicitly identify $C^+(2) = \mathbb{H}$ by setting $\underline{i} := e_1e_2$, $\underline{j} := e_1e_3$ and $\underline{k} := e_1e_2e_1e_3 = e_2e_3$. Let us consider $c = e_1e_2e_3$ which generates the center of $C(3)$. The map $v \mapsto cv$ is a linear map which embeds V into the subspace $\mathbb{H}^0$ of $C^+(2) = \mathbb{H}$ of quaternions q with $q = -\overline{q}$.

We claim that $\Gamma^+(\mathbb{R}^3, Q) = \mathbb{H}^*$, the group of all invertible quaternions.

In fact we have the elements cv with $v \neq 0$ which are the nonzero quaternions in $\mathbb{H}^0$, and we leave to the reader the easy verification that these elements generate the group $\mathbb{H}^*$.

Definition 2. The spin group is the subgroup:

$$\text{Spin}(V) := \{r \in \Gamma^+(V, Q) \mid N(r) = 1\}. \tag{5.1.4}$$

We assume now that F is the field of real numbers and Q is definite, or F is the complex numbers.

Theorem. *(a) The spin group is a double cover of the special orthogonal group. We have an exact sequence* $1 \to \pm 1 \to \text{Spin}(V) \xrightarrow{\pi} SO(V) \to 1$.

(b) The Lie algebra of $\text{Spin}(V)$ *is* $\bigwedge^2 V = [L, L]$ *(notations of 4.4).*

Proof. (a) Consider $r := v_1v_2 \dots v_{2j}$, and compute $N(r) = \prod_{i=1}^{2j} Q(v_i)$. If we are in the case of complex numbers we can fix an $f \in \mathbb{C}$ so that $f^2N(r) = 1$ and $N(fr) = 1$. Similarly, if $F = \mathbb{R}$ and Q is definite, we have that $N(r) > 0$ and we can fix an $f \in \mathbb{R}$ so that $f^2N(r) = 1$. In both cases we see that $\text{Spin}(V) \xrightarrow{\pi} SO(V)$ is surjective. As for the kernel, if $f \in F^*$ we have $N(f) = f^2$. So $f \in \text{Spin}(V)$ if and only if $f = \pm 1$.

(b) Since the spin group is a double cover of $SO(V)$ it has the same Lie algebra, which is $[L, L] = so(V)$ by Proposition 4.4. □

When $V = F^n$ with form $-\sum_{i=1}^n x_i^2$ we denote $\text{Spin}(V) = \text{Spin}(n, F)$.

Example. For $V = \mathbb{R}^3$ with the negative definite quadratic form, we have seen that $\Gamma^+(\mathbb{R}^3, Q) = \mathbb{H}^*$, the group of all invertible quaternions, hence:

$$\text{Spin}(3, \mathbb{R}) = \{q \in \mathbb{H} \mid q\overline{q} = 1\}, \quad q = a + bi + cj + dk,$$
$$N(q) = q\overline{q} = a^2 + b^2 + c^2 + d^2.$$

Therefore, topologically $\text{Spin}(3, \mathbb{R}) = S^3$, the 3-dimensional sphere.

As groups, $\mathbb{H}^* = R^+ \times SU(2, \mathbb{C})$, $\text{Spin}(3, \mathbb{R}) = SU(2, \mathbb{C}) = Sp(1)$. This can be seen using the formalism of 5.2 for $\mathbb{H}$.

$$q = \alpha + j\beta, \quad N(q) = \alpha\overline{\alpha} + \beta\overline{\beta}, \quad (\alpha + j\beta)j = -\overline{\beta} + j\overline{\alpha}.$$

Formula 5.2.1 expresses q as the matrix

$$q := \begin{vmatrix} \alpha & -\overline{\beta} \\ \beta & \overline{\alpha} \end{vmatrix}, \quad N(q) = \det(q) = \alpha\overline{\alpha} + \beta\overline{\beta}.$$

The previous statements follow immediately from this matrix representation.

If we take $F = \mathbb{C}$, the spin group is an algebraic group (see next chapter) so we have the *algebraic form* $\mathrm{Spin}(n, \mathbb{C})$. If we take $F = \mathbb{R}$ and the negative definite quadratic form we have the *compact form* $\mathrm{Spin}(n, \mathbb{R})$ of the spin group.

The main point is that the extension $1 \to \pm 1 \to \mathrm{Spin}(n, \mathbb{R}) \xrightarrow{\pi} SO(n, \mathbb{R}) \to 1$ $(n > 2)$ is not split, or in better words, that $\mathrm{Spin}(n, \mathbb{R})$ is a connected and simply connected group.

Let us sketch the proof using some elementary algebraic topology. First, the map $\mathrm{Spin}(n, \mathbb{R}) \xrightarrow{\pi} SO(n, \mathbb{R})$ is a locally trivial fibration (as for all surjective homomorphisms of Lie groups, cf. Chapter 4, 3.7). Since $SO(n, \mathbb{R})$ is connected it is enough to exhibit a curve connecting $\pm 1 \in \mathrm{Spin}(n, \mathbb{R})$. Since $\mathrm{Spin}(n-1, \mathbb{R}) \subset \mathrm{Spin}(n, \mathbb{R})$ it is enough to look at $\mathrm{Spin}(2, \mathbb{R})$.

In this case the Clifford algebra $C(2)$ is the quaternion algebra. The space V is spanned by $\underline{i}, \underline{j}$ and $C^+(2) = \mathbb{C} = \mathbb{R} + \mathbb{R}\underline{k} \quad (\underline{k} = \underline{i}\,\underline{j})$.

$$\mathrm{Spin}(2, \mathbb{R}) = U(1) = \{\alpha \in \mathbb{C} \mid |\alpha| = 1\} = \{\cos\phi + \sin\phi\, \underline{k} = e^{\phi \underline{k}}\}$$

and we have, from $\underline{k}\,\underline{i} = -\underline{i}\,\underline{k}, \underline{k}\,\underline{j} = -\underline{j}\,\underline{k}$

$$e^{\phi \underline{k}}\underline{i}e^{-\phi \underline{k}} = e^{2\phi \underline{k}}\underline{i}, \quad e^{\phi \underline{k}}\underline{j}e^{-\phi \underline{k}} = e^{2\phi \underline{k}}\underline{j},$$

from which the double covering and the connectedness is clear.

For the simple connectedness of $\mathrm{Spin}(n, \mathbb{R})$, we need some computations in homotopy theory. Basically we need to compute the fundamental group of the special orthogonal group and prove that

$$\pi_1(SO(n, \mathbb{R})) = \mathbb{Z}/(2), \quad \forall n \geq 3.$$

This can be seen by considering the transitive action of $SO(n, \mathbb{R})$ on the $n-1$-dimensional sphere S^{n-1} by rotations. The stabilizer of a given point is $SO(n-1, \mathbb{R})$, and thus $S^{n-1} = SO(n, \mathbb{R})/SO(n-1, \mathbb{R})$. We thus have that $SO(n, \mathbb{R})$ fibers over S^{n-1} with fiber $SO(n-1, \mathbb{R})$. We have therefore an exact sequence of homotopy groups:

$$\pi_2(S^{n-1}) \to \pi_1(SO(n-1, \mathbb{R})) \to \pi_1(SO(n, \mathbb{R})) \to \pi_1(S^{n-1}).$$

If $n > 3$, $\pi_2(S^{n-1}) = \pi_1(S^{n-1}) = 0$. Hence $\pi_1(SO(n-1, \mathbb{R})) = \pi_1(SO(n, \mathbb{R}))$ and we have $\pi_1(SO(n, \mathbb{R})) = \pi_1(SO(3, \mathbb{R}))$, $\forall n \geq 3$. For $n = 3$ we have seen that we have a double covering $1 \to \mathbb{Z}/(2) \to SU(2, \mathbb{C}) \to SO(3, \mathbb{R}) \to 1$. Since $SU(2, \mathbb{C}) = S^3$, it is simply connected. The exact sequence of the fibration gives the isomorphism

$$\pi_1(S^3) = 0 \to \pi_1(SO(3, \mathbb{R})) \to \mathbb{Z}/(2) = \pi_0(\mathbb{Z}/(2)) \to 0 = \pi_0(S^3).$$

For further details we refer to standard texts in algebraic topology (cf. [Sp], [Hat]).

6 Basic Constructions on Representations

6.1 Tensor Product of Representations

Having the formalism of tensor algebra we can go back to representation theory. Here representations are assumed to be finite dimensional.

The distinctive feature of the theory of representations of a group, versus the general theory of modules, lies in the fact that we have several ways to compose representations to construct new ones. This is a feature that groups share with Lie algebras and which, once it is axiomatized, leads to the idea of Hopf algebra (cf. Chapter 8, §7).

Suppose we are given two representations V, W of a group, or of a Lie algebra.

Theorem. *There are canonical actions on V^*, $V \otimes W$ and* $\hom(V, W)$*, such that the natural mapping $V^* \otimes W \to \hom(V, W)$ is equivariant.*

First we consider the case of a group. We already have (cf. Chapter 1, 2.4.2) general definitions for the actions of a group on $\hom(V, W)$: recall that we set $(gf)(v) := g(f(g^{-1}))$. This definition applies in particular when $W = F$ with the trivial action and so defines the action on the dual (the contragredient action).

The action on $V \otimes W$ is suggested by the existence of the tensor product of operators. We set $g(v \otimes w) := gv \otimes gw$. In other words, if we denote by $\varrho_1, \varrho_2, \varrho$ the representation maps of G into: $GL(V)$, $GL(W)$, $GL(V \otimes W)$ we have $\varrho(g) = \varrho_1(g) \otimes \varrho_2(g)$. Summarizing

$$\begin{aligned} &\text{for } \hom(V, W), \quad && (gf)(v) := g(f(g^{-1})), \\ &\text{for } V^*, && \langle g\phi | v\rangle := \langle \phi | g^{-1}v\rangle, \\ &\text{for } V \otimes W, && g(v \otimes w) := gv \otimes gw. \end{aligned}$$

We can now verify:

Proposition 1. *The natural mapping $i : W \otimes V^* \to \hom(V, W)$ is equivariant.*

Proof. Given $g \in G$, $a = w \otimes \varphi \in W \otimes V^*$, we have $ga = gw \otimes g\varphi$ where $\langle g\varphi | v\rangle = \langle \varphi | g^{-1}v\rangle$. Thus, $(ga)(v) = \langle g\varphi | v\rangle gw = \langle \varphi | g^{-1}v\rangle gw = g(a(g^{-1}v))$, which is the required equivariance by definition of the action of G on $\hom(V, W)$. □

Let us now consider the action at the level of Lie algebras.

First, let us assume that G is a Lie group with Lie algebra $\mathrm{Lie}(G)$, and let us consider a one-parameter subgroup $\exp(tA)$ generated by an element $A \in \mathrm{Lie}(G)$.

Given a representation ϱ of G we have the induced representation $d\varrho$ of $\mathrm{Lie}(G)$ such that $\varrho(\exp(tA)) = \exp(td\varrho(A))$. In order to understand the mapping $d\varrho$ it is enough to expand $\varrho(\exp(tA))$ in a power series up to the first term.

We do this for the representations V^*, $V \otimes W$ and $\hom(V, W)$. We denote the actions on V, W simply as gv or Aw both for the group or Lie algebra.

Since $\langle \exp(tA)\varphi | v\rangle = \langle \varphi | \exp(-tA)v\rangle$ the Lie algebra action on V^* is given by

$$\langle A\varphi|v\rangle = \langle\varphi| - Av\rangle.$$

In matrix notation the contragredient action of a Lie algebra is given by minus the transpose of a given matrix.

Similarly we have the formulas:

$$A(v \otimes w) = Av \otimes w + v \otimes Aw, \ (Af)(v) = A(f(v)) - f(A(v)). \tag{6.1.1}$$

for the action on tensor product or on homomorphisms. As a consequence we have:

Proposition 2. *If G is a connected Lie group*

$$\hom_G(V, W) = \{f \in \hom(V, W) | Af = 0, \ A \in \mathrm{Lie}(G)\}.$$

Proof. Same as Chapter 4, Remark 1.4 on fixed points. □

One final remark, which is part of the requirements when axiomatizing Hopf algebras, is that both a group G and a Lie algebra L have a trivial 1-dimensional representation, which behaves as unit element under tensor product.

On the various algebras $T(U)$, $S(U)$, $\bigwedge U$ the group $GL(U)$ acts as automorphisms. Hence the Lie algebra $gl(U)$ acts as derivations induced by the linear action on the space U of generators. For example,

$$\begin{aligned} A(u_1 \wedge u_2 \ldots \wedge u_k) &= Au_1 \wedge u_2 \ldots \wedge u_k + u_1 \wedge Au_2 \ldots \wedge u_k + \cdots \\ &\quad + u_1 \wedge u_2 \ldots \wedge Au_k. \end{aligned} \tag{6.1.2}$$

On the Clifford algebra we have an action as derivations only of the Lie algebra of the orthogonal group of the quadratic form, since only this group preserves the defining ideal. We have seen in 4.4 that these derivations are inner (Chapter 4, §1.1) and induced by the elements of $\bigwedge^2 V$.

6.2 One-dimensional Representations

We complete this part with some properties of 1-dimensional representations.

A 1-dimensional representation is just a homomorphism of G into the multiplicative group F^* of the base field. Such a homomorphism is also called a *multiplicative character*.

The tensor product of two 1-dimensional spaces is 1-dimensional and so is the dual.

Moreover a linear operator on a 1-dimensional space is just a scalar, the tensor product of two scalars is their product and the inverse transpose is the inverse. Thus:

Theorem. *The product of two multiplicative characters is a multiplicative character, and so is the inverse. The multiplicative characters of a group G form a group, called the character group of G (usually denoted by $\hat{G}$).*

Notice in particular that if V is one-dimensional, $V \otimes V^*$ is canonically isomorphic to the trivial representation by the map $v \otimes \phi \to \langle \phi | v \rangle$ (the trace). Sometimes for a 1-dimensional representation it is convenient to use the notation V^{-1} instead of V^*.

Let us show a typical application of this discussion:

Proposition. *If L and U are representations of a group G such that* $\dim L = 1$, *then U is irreducible if and only if $L \otimes U$ is irreducible.*

Proof. If $W \subset U$ is a proper submodule then also $L \otimes W \subset L \otimes U$ is a proper submodule, so we have the implication in one direction. But now

$$U = (L^{-1} \otimes L) \otimes U = L^{-1} \otimes (L \otimes U),$$

and we have also the reverse implication. □

7 Universal Enveloping Algebras

7.1 Universal Enveloping Algebras

There is one further construction we want to briefly discuss since it is the natural development of the theory of Capelli of Chapter 3. Given a Lie algebra L we consider in the tensor algebra $T(L)$ the ideal I_L generated by the quadratic relations $a \otimes b - b \otimes a - [a, b]$.

Definition. The associative algebra $U(L) := T(L)/I_L$ is called the *universal enveloping algebra* of the Lie algebra L.

The meaning of these relations is that the commutator of the elements $a, b \in L \subset T(L)$ performed in the associative algebra $U(L)$ must coincide with the commutator defined in L by the Lie algebra law. In other words we impose the minimal relations which imply that the morphism $L \to T(L)/I_L$ is a Lie homomorphism (where on the associative algebra $T(L)/I_L$ the Lie algebra structure is induced by the usual commutator).

As for other universal constructions this algebra satisfies the universal property of *mapping into associative algebras*. In fact we have that:

Proposition 1. *A Lie homomorphism $i : L \to A$ where A is an associative algebra with induced Lie structure, extends uniquely to a homomorphism of algebras $U(L) \to A$.*

Proof. By the universal property of tensor algebra the linear map i extends to a homomorphism $i : T(L) \to A$. Since i is a Lie homomorphism we have

$$i(a \otimes b - b \otimes a - [a, b]) = i(a)i(b) - i(b)i(a) - i([a, b]) = 0$$

so i factors through I_L to the required homomorphism. □

The first important result on universal enveloping algebras is the Poincaré–Birkhoff–Witt theorem, which states that:

PBW Theorem. *(1) If $u_1, u_2, \dots, u_k \dots$ form a linear basis of L, the ordered monomials $u_1^{h_1} u_2^{h_2} \dots u_k^{h_k} \dots$ give a linear basis of $U(L)$.*

(2) In characteristic 0 we have a direct sum decomposition $T(L) = I_L \oplus \sum_{i=0}^{\infty} S_i(L)$ (where $S_i(L)$ denotes the space of symmetric tensors of degree i).

Proof. It is almost immediate by induction that the monomials $u_1^{h_1} u_2^{h_2} \dots u_k^{h_k} \dots$ are linear generators. In fact if we have in a product a pair $u_i u_j$ in the *wrong order*, $j < i$ we replace it be $u_j u_i - [u_i, u_j]$. Expanding $[u_i, u_j]$ in the given basis we obtain lower degree monomials and proceed by induction. The independence requires a nontrivial argument.

Consider thus the tensors $M = u_{i_1} \otimes \cdots \otimes u_{i_k}$ which as the indices i_t and k vary, give a basis of the tensor algebra.

We look at the sequence of indices $i_1, i_2, \dots, i_k$, for a tensor M, and count the number of *descents*, i.e., the positions j for which $i_j > i_{j+1}$: we call this number the *index*, $i(M)$ of M. When $i(M) = 0$, i.e., when $i_1 \leq i_2 \leq \cdots \leq i_k$ we say that M is *standard*. Let us define T_n to be the span of all tensors of degree $\leq n$ and by $T_n^s \subset T_n$ the span of the standard tensors. We need a basic lemma.

Lemma. *For each n, there is a unique linear map $\sigma : T_n \to T_n^s$ such that:*

(1) σ is the identity on T_n^s.
(2) Given a tensor $A \otimes a \otimes b \otimes B$, $a, b \in L$, $A, B \in T(L)$, we have

$$\sigma(A \otimes a \otimes b \otimes B) = \sigma(A \otimes b \otimes a \otimes B) + \sigma(A \otimes [a, b] \otimes B).$$

Proof. We define σ on the tensors $M = u_{i_1} \otimes \cdots \otimes u_{i_k}$ by induction on the degree k and on the index $i(M)$. When $i(M) = 0$ by definition we must set $\sigma(M) = M$. When $i(M) > 0$ we have an expression $A \otimes u_i \otimes u_j \otimes B$ with $i > j$ and hence $i(A \otimes u_j \otimes u_i \otimes B) = i(M) - 1$. Thus we may set recursively:

$$\sigma(A \otimes u_i \otimes u_j \otimes B) = \sigma(A \otimes u_j \otimes u_i \otimes B) + \sigma(A \otimes [u_i, u_j] \otimes B).$$

If $i(M) = 1$ this definition is well posed; otherwise, when we have at least two descents we have to prove that the definition of $\sigma(M)$ is independent of the descent we choose. We have two cases: (1) The descents are in $A \otimes b \otimes a \otimes B \otimes d \otimes c \otimes C$. (2) We have consecutive descents $A \otimes c \otimes b \otimes a \otimes B$.

In the first case we have by induction, starting from the descent in $b \otimes a$:

$$\begin{aligned}
&\sigma(A \otimes a \otimes b \otimes B \otimes d \otimes c \otimes C) + \sigma(A \otimes [b, a] \otimes B \otimes d \otimes c \otimes C) \\
&\quad = \sigma(A \otimes a \otimes b \otimes B \otimes c \otimes d \otimes C) + \sigma(A \otimes a \otimes b \otimes B \otimes [d, c] \otimes C) \\
&\qquad + \sigma(A \otimes [b, a] \otimes B \otimes c \otimes d \otimes C) + \sigma(A \otimes [b, a] \otimes B \otimes [d, c] \otimes C).
\end{aligned}$$

Clearly when we start from the other descent we obtain the same result.

For the other case, write for convenience $\tau(X) := \sigma(A \otimes X \otimes B)$. We need to compare:

$$\tau(b \otimes c \otimes a + [c, b] \otimes a), \quad \tau(c \otimes a \otimes b + c \otimes [b, a])$$

We iterate the formulas by induction, the two terms are:

$$1: \ \tau(b \otimes a \otimes c + b \otimes [c, a] + [c, b] \otimes a),$$
$$2: \ \tau(a \otimes c \otimes b + [c, a] \otimes b + [b, a] \otimes c + [c, [b, a]])$$

Applying again the same rules (notice that either the index or the degree is diminished so we can apply induction) we have:

$$1: \ \tau(a \otimes b \otimes c + [b, a] \otimes c + [c, a] \otimes b + [b, [c, a]] + [c, b] \otimes a)$$
$$2: \ \tau(a \otimes b \otimes c + [b, a] \otimes c + [c, a] \otimes b + [c, b] \otimes a + [a, [c, b]] + [c, [b, a]]).$$

From the Jacobi identity $[b, [c, a]] = [a, [c, b]] + [c, [b, a]]$ so the claim follows. □

We can now conclude the proof of the PBW theorem.

The linear map σ by definition vanishes on the ideal I_L defining $U(L)$, thus it defines a linear map $U(L) \to T^s$ which, by the previous remarks, maps the images of the standard tensors which span $U(L)$ into themselves, thus it is a linear isomorphism.

The second part follows easily since a basis of symmetric tensors is given by symmetrization of standard tensors. The image under σ of the symmetrization of a standard tensor M is M. □

There is a simple but important corollary. Let us filter $U(L)$ by setting $U(L)_i$ to be the span of all monomials in elements of L of degree $\leq i$. Then we have:

Proposition 2. *(i) The graded algebra $\oplus U(L)_i / U(L)_{i-1}$ is isomorphic to the symmetric algebra $S(L)$.*

(ii) If the characteristic is 0, for every i we have $U_{i+1}(L) = \overline{S}_i(L) \oplus U_i(L)$, where $\overline{S}_i(L)$ is the image in $U_{i+1}(L)$ of the symmetric tensors.

The importance of the second statement is this. The Lie algebra L acts on $T(L)$ by derivations and, over $\mathbb{C}$, on its associated group G by automorphisms. Both I_L and $S_i(L)$ are stable subspaces. Thus the actions factor to actions on $U(L)$. The subspaces $U_i(L)$ are subrepresentations and in characteristic 0, we have that $U_i(L)$ has the invariant complement $\overline{S}_i(L)$ in $U_{i+1}(L)$.

Exercise. If $L \subset M$ are Lie algebras, the PBW theorem for M implies the same theorem for L; we also can prove it for linear Lie algebras from Capelli's theory.[40]

[40] The PBW theorem holds for any Lie algebra, not necessarily finite dimensional. For finite-dimensional algebras there is a deep theorem stating that these algebras are indeed linear (Ado's theorem, Chapter 10). Usually this theorem is proved *after* proving the PBW theorem.

7.2 Theorem of Capelli

We want to use the previous analysis to study the center of $U(L)$, in particular to give the full details of the Theorem of Capelli, sketched in Chapter 3, §5.3 on the center of the algebra of polarizations.

From Proposition 2 of the preceding section, if G is a group of automorphisms of the Lie algebra L, G acts as automorphisms of the tensor algebra $T(L)$ and preserves the ideal I_L. Thus G extends to a group of automorphisms of $U(L)$. Moreover it clearly preserves the spaces $\overline{S}_i(L)$. In particular we can consider as in Chapter 4, §4.1 the *adjoint group* G^0 generated by the one-parameter groups $e^{t\,\mathrm{ad}(a)}$. Notice that $\mathrm{ad}(a)$ extends to the *inner derivation*: $r \to ar - ra$ of $U(L)$, preserving all the terms $U(L)_i$ of the filtration. We have:

Proposition. *The center of $U(L)$ coincides with the invariants under G^0.*

Proof. By definition an element of $U(L)$ is fixed under G^0 if and only if it is fixed by all the one-parameter subgroups $e^{t\,\mathrm{ad}(a)}$. An element is fixed by a one-parameter subgroup if and only if it is in the kernel of the generator, in our case $\mathrm{ad}(a)$, i.e., if it commutes with a. Since $U(L)$ is generated by L it is clear that its center is the set of elements which commute with L. $\square$

Let us apply the theory to $gl(n, \mathbb{C})$, the Lie algebra of $n \times n$ matrices. In this case $gl(n, \mathbb{C})$ is also an associative algebra. Its group G of automorphisms is induced by conjugation by invertible matrices. Given a matrix A, we have that the group of linear transformations $B \to e^{tA}Be^{-tA}$ has as infinitesimal generator $B \mapsto AB - BA = \mathrm{ad}(A)(B)$.

The Lie algebra $gl(n, \mathbb{C})$ is isomorphic to the Lie algebra of polarizations of Chapter 3. The elementary matrix e_{ij} with 1 in the i, j position and 0 otherwise corresponds to the operator Δ_{ij}. The universal enveloping algebra of $gl(n, \mathbb{C})$ is isomorphic to the algebra $\mathcal{U}_n$ generated by polarizations. Formulas 5.3.3 and 5.3.6 of Chapter 3, give elements K_i in the center of $\mathcal{U}_n$, K_i is a polynomial of degree i in the generators.

Let us analyze the development of the determinant 5.3.3 and, in particular, the terms which contribute to $K_i\rho^{m-i}$. In such a term the contribution of a factor $\Delta_{ii} + m - i$ can be split into the part involving Δ_{ii} and the one involving $m - i$. This last one produces terms of strictly lower degree. Therefore in the associated grading the images of the K_i can be computed by dropping the constants on the diagonal and thus are, up to sign, the coefficients σ_i of the characteristic polynomial of a matrix with entries the classes x_{ij} of the $e_{ij} = \Delta_{ij}$. By Chapter 2, Theorem 5.1 we have that these coefficients generate the invariants under conjugation. We then get:

Theorem. *The elements K_i generate the center of U_n which is the polynomial ring in these generators.*

Proof. Let f be in the center, say $f \in (\mathcal{U}_n)_i$. Its symbol is an invariant in the graded algebra. Hence it is a polynomial in the coefficients σ_i. We have thus a polynomial g in the K_i which lies also $(\mathcal{U}_n)_i$ and has the same symbol. Therefore $f - g \in (\mathcal{U}_n)_{i-1}$, and we can finish by induction. $\square$

This theorem has of course a natural generalization to semisimple Lie algebras. One has to replace the argument of Capelli with more general arguments of Chevalley and Harish-Chandra but the final result is quite similar. The center of the universal enveloping algebra of a semisimple Lie algebra is a ring of polynomials in generators which correspond to symmetric functions under the appropriate group, the Weyl group.

7.3 Free Lie Algebras

As usual, given a Lie algebra L and a set $X \subset L$ we say:

Definition. L is free over X if, given any Lie algebra M and any map $f : X \to M$, f extends to a unique homomorphism $f : L \to M$ of Lie algebras.

The PBW Theorem immediately tells us how to construct free Lie algebras. Let $F\langle X\rangle$ be the free associative noncommutative polynomial algebra over X (the tensor algebra on a vector space with basis X). Let L be the Lie subalgebra of $F\langle X\rangle$ generated by X.

Proposition. *L is free over X.*

Proof. Let M and $f : X \to M$ be given. Consider the universal enveloping algebra U_M of M. By PBW we have $M \subset U_M$. Since $F\langle X\rangle$ is the free associative algebra, f extends to a homomorphism $\tilde{f} : F\langle X\rangle \to U_M$. Since L is generated by the elements X as Lie algebra, $\tilde{f}$ restricted to L maps L to M extending the map f on the generators X. The extended map is also uniquely determined since L by construction is generated by X. $\square$

The free Lie algebra is a very interesting object. It has been extensively studied ([Reu]).