Chapter 2, TENSOR ALGEBRA

November 1995

Contents

In this chapter we develop somewhat quickly the basic facts of tensor algebra assuming the reader is familiar with usual linear algebra.

§12 Tensor algebra.

12.1 The language of functions is most suitably generalized into the one 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 notaion of the physicistst, 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 furnishes a coherent model to perform all 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 thus $\sum_{x \in X} f(x)x$ 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 2 variables, it has dimension nm.

Of course, given a function $f(x) \in F^X$ and a function $g(y) \in F^Y$ we can form the 2 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 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 has dimension nm they must be a basis.

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

Definition. 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.

As in the previous case we easily see that, if U, V, W are finite dimensional:

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).

12.3

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

Property iv) insures that two different tensor product maps are canonically isomorphic, in this sense we will speak of the tensor product of two vector spaces denoted by $U \otimes V$ and by $u \otimes v$ the image of the pair (u,v) in the bilinear (tensor product) map.

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

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

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

We have a bilinear map

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

given by $F(\varphi,\psi)(u,v) :=< \varphi|u >< \psi|v >$; in other and more concrete words, the product of two linear functions 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

$$f(\sum_{i=1}^{n}\alpha_i u_i, \sum_{j=1}^{m}\beta_j v_j) = \sum_{i=1}^{n}\sum_{j=1}^{m}\alpha_i\beta_j f(u_i,v_j). \tag{12.3.1}$$

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

$$e^{hk}(\sum_{i=1}^{n}\alpha_i u_i, \sum_{j=1}^{m}\beta_j v_j) = \alpha_h\beta_k, \tag{12.3.2}$$

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

$$f = \sum_{i=1}^{n}\sum_{j=1}^{m} f(u_i,v_j)e^{ij}. \tag{12.3.3}$$

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

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 just 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[1]

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

12.4

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

If $f : U_1 \to V_1$ and $g : U_2 \to V_2$ are 2 linear maps, the map $U_1 \times V_1 \to U_2 \otimes V_2$ given by $(u,v) \to f(u)\otimes g(v)$ is bilinear, hence it factors through a unique linear map denoted by $f\otimes g : U_1\otimes V_1 \to U_2\otimes V_2$.

This is characterized by the property

$$(f\otimes g)(u\otimes v) = f(u)\otimes g(v). \tag{12.4.1}$$

In matrix notations the only difficulty is a notational one, usually it is customary to index basis elements with integral indeces, clearly if we do this for two spaces the tensor product basis is indexed with pairs of indeces and so the corresponding matrices with pairs of pairs of indeces.

[1] 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.

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

$$(f \otimes g)(u_i \otimes v_h) = \sum_{jk} a_{ji} b_{kh} u'_j \otimes v'_k \tag{12.4.2}$$

hence the elements $a_{ji} b_{kh}$ are the entries of the tensor product of the 2 matrices of the 2 operators.

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

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

Using bases and matrix notations we have thus a map

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

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

Finally we have the obviuos associativity conditions. Given

$$\begin{array}{ccccc} U_1 & \xrightarrow{f} & V_1 & \xrightarrow{h} & W_1 \\ U_2 & \xrightarrow{g} & V & \xrightarrow{k} & W_2 \end{array}$$

we have $(h \otimes k)(f \otimes g) = hf \otimes kg$. In particular if we consider endomorphisms 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 2 associative algebras A, B over F. The vector space $A \otimes B$ obtains 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, 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 $B \supset F$ is a field extension then $A \otimes_F G$ can be thought as a G algebra.

Remark Given an algebra A and 2 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 this is induced by $g \to g \otimes g$, $g \in G$).

12.5 We analyze some special cases.

First of all we identify any vector space U with $hom(F, U)$. To a map $f \in hom(F, U)$ associate the vector $f(1)$.

We have also seen that $F \otimes U = U$.

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 to $hom(U, V)$.

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

It is useful to explicit this identification and express the action of a decomposable element $v \otimes \varphi$ on a vector u, and the composition law $hom(V, W) \times hom(U, V) \to hom(U, W)$.

With the obvious notations we easily find:

$$(12.5.1) \qquad (v \otimes \varphi)(u) = v < \varphi|u >, \ w \otimes \psi \circ v \otimes \varphi = w \otimes < \psi|v > \varphi.$$

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

$$(12.5.2) \qquad Tr(u \otimes \varphi) :=< \varphi|u > .$$

Definition. The mapping $Tr : End(U) \to F$ 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 $Tr(A) = \sum_{ij} a_{ij} < e^j|e_i >= \sum_i a_{ii}$.

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

$$(12.5.3) \qquad Tr(A \otimes B) = Tr(A)Tr(B)$$

as one verifies immediately.

12.6 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 the elements a_{ij} has rank ≤ 1.

Another important case is the sequence of identifications:

$$(12.6.1) \quad U^* \otimes V^* = hom(U, F) \otimes hom(V, F) = hom(U \otimes V, F \otimes F) = hom(U \otimes V, F)$$

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 explicitely the duality pairing at the level of decomposable tensors:

(12.6.2) $$< \varphi \otimes \psi | u \otimes v > = < \varphi | u > < \psi | v > .$$

The interpretation of $U \otimes V$ as bilinear functions on $U^* \times V^*$ is completely embedded in this basic pairing.

Summarizing we have:

The intrinsic notion of $U \otimes V$ as solution of a universal problem, its appearence as bilinear functions on $U^* \times V^*$, or finally as the dual of $U^* \otimes V^*$.

12.7 The tensor product construction can be clearly 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:

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

given on the decomposable tensors by

(12.7.1) $$< \varphi_1 \otimes \varphi_2 \otimes \cdots \otimes \varphi_m | u_1 \otimes u_2 \otimes \cdots \otimes u_m > = \prod_{i=1}^{m} < \varphi_i | u_i > .$$

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

Similarly an identification of

$$Hom(U_1 \otimes U_2 \otimes \ldots U_m, V_1 \otimes V_2 \otimes \ldots V_m) \cong Hom(U_1, V_1) \otimes Hom(U_2, V_2) \otimes \ldots \otimes Hom(U_m, V_m).$$

Let us notice the self duality pairing on $End(U)$ given by $Tr(AB)$ in terms of decomposable tensors, if $A = v \otimes \psi$ and $B = u \otimes \varphi$ we have:

(12.7.2) $$Tr(AB) = Tr(v \otimes \psi \circ u \otimes \varphi) = Tr(< \psi | u > v \otimes \varphi) = < \varphi | v > < \psi | u >$$

we recover the simple fact that $Tr(AB) = Tr(BA)$.

Remark also that this is a non degenerate pairing and $End(U) = U \otimes U^*$ is identified by it to the dual:

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

Since $Tr([A, B]) = 0$, the operators with trace 0 form a Lie algebra, 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 $Tr(1) = n$.

If we are in characteristic 0 (or prime with n) we can decompose each matrix as $A = \frac{Tr(A)}{n} 1 + A_0$ where A_0 has zero trace.

Thus the Lie algebra $gl(U)$ decomposes as direct sum $gl(U) = F \oplus sl(U)$ (F being the *scalars* i.e. the multiples of the identity).

It will have a special interest for 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 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 denote by

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

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 $j(u_1)j(u_2)\ldots j(u_k)$ 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. $\square$

12.8 In particular, given a linear automorphism g of U, this extends to an automorphism of the tensor algebra which, on the tensors $U^{\otimes m}$ is $g^{\otimes m} := g \otimes g \otimes g \cdots \otimes g$. 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 indexed by the letters of an *alphabet* A with n letters.[2]

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 as *the non commutative 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 $F < A >$ instead of $T(U)$. In this language the universal property is the one of polynomials, i.e. we can evaluate a polynomial in any algebra once we give the *values* for the variables.

[2] 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

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, in the corresponding product of the values. Thus this map is really the evaluation of a polynomial.

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

We can reconstruct the commutative picture 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 the commutators $[a, b] := ab - ba$.

It is enough to consider the ideal generated by the commutators of a set of generators for the algebra since if an algebra is generated by pairwise commuting elements 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.

In the presentation as free algebra, to make $F < A >$ 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 polynomial language we find again the action 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 $\wedge U$.

The multiplication of elements in $\wedge U$ is also indicated by $a \wedge b$. Again we have an action of $GL(U)$ on $\wedge U = \oplus_k \wedge^k U$ as automorphisms of graded algebra and the algebra satisfies the universal property with respect to linear maps

Given a linear map $j : U \to R$ into an algebra R restricted by the condition $j(u)^2 = 0,\ \forall u \in U$ it extends to a unique homomorphism $\wedge U \to R$.

In the language of an alphabet it appears as follows. The variables in A satisfy the rules:

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

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

We order the letters in A and consider the monomials in which the letters appear in strict increasing order, 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).

Theorem. *The strict monomials are a basis of* $\wedge(U)$.

Proof. We hint 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 $\wedge(U)$.

For a different proof see 12.11 in which we generalize this theorem to Clifford algebras.

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

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

We also suggest to the reader the

Exercise The multiplication map $\wedge^k U \times \wedge^{n-k} U \to \wedge^n U$ is a non degenerate pairing.

spostare

The Grassmann algebra is strictly tied with the theory of determinants. First of all one can define the determinant of a linear map $A : U \to U$ of an $n-$dimensional vector space U as the linear map $\wedge^n A$.

In order to define the determinant as a scalar one has to choose an identification of $\wedge^n U$ with the base field. This is done 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.

More generally, given a linear map $A : U \to V$ the composed map $j : U \to V \to \wedge V$ satisfies the universal property and thus we have a homomorphism, denoted by $\wedge A : \wedge U \to \wedge V$. In this way $\wedge U$ is a *functor* from vector spaces to graded algebras (a similar fact holds for the tensor and symmetric algebras).

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 $\wedge A$ starting from the matrix a^i_j of A. we have $Au_j = \sum_i a^i_j v_i$ and

$$\wedge^k A(u_{j_1} \wedge \cdots \wedge u_{j_k}) = Au_{j_1} \wedge \cdots \wedge Au_{j_k} =$$
$$(\sum_{i_1} a^{i_1}_{j_1} v_{i_1}) \wedge (\sum_{i_2} a^{i_2}_{j_2} v_{i_2}) \wedge \cdots \wedge (\sum_{i_k} a^{i_k}_{j_k} v_{i_k}) =$$
$$\sum_{i_1,\dots,i_k} A(i_1, \dots, i_k | j_1, \dots, j_k) v_{i_1} \wedge \cdots \wedge v_{i_k}$$

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

Given two matrices A, B which compose into BA the explicit multiplication formula of the two matrices associated to two exterior powers $\wedge^k A, \wedge^k B, \wedge^k(BA) = \wedge^k B \circ \wedge^k A$ is called *Binet's formula*.

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 $U \times U \times U \ldots \times U$. Explicitely:

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

We will refer to this action as to the *symmetry action* on tensors.[3]

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

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

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

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

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.*

For this we consider again a basis of U which we think of as an ordered alphabet, and take for basis of $U^{\otimes n}$ the words of length n in this alphabet.

The action of 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, the image of these elements in the symmetric algebra is always the same monomial. 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 $S_{h_1} \times \ldots \times S_{h_k}$, the number of elements in a given orbit is $\binom{n}{h_1\ h_2 \ldots h_k}$.

The sum of the elements of a given orbit is a symmetric tensor and these tensors are a basis of $\Sigma_n(U)$. Thus the image of this basis element in the symmetric algebra is the corresponding commutative monomial times the order of the orbit, e.g.:

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

This establishes the isomorphism since it sends a basis to a basis.

In order to be more explicit let us indicate by $e_1, e_2, \ldots, e_m$ a basis of U. The element

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

is a symmetric tensor which in the symmetric algebra corresponds to the monomial

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

[3] It will be studied intensively in Chap.3.

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

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

This tensor maps in the Grassmann algebra to

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

It is often customary to identify $e_{i_1} e_{i_2} \dots e_{i_n}$, resp. $a_1 \wedge a_2 \wedge \dots \wedge a_n$ to the corresponding symmetric or antisymmetric tensor.

Let us now notice one more fact. Given a vector $u \in U$ the tensor $u^n = u \otimes u \otimes u \dots \otimes u$ is clearly symmetric. If $u = \sum_k \alpha_k e_k$ we have:

$$u^n = \sum_{h_1 + h_2 + \dots + h_m = n} \alpha_1^{h_1} \alpha_2^{h_2} \dots \alpha_m^{h_m} \binom{n}{h_1\, h_2 \dots h_m} e_1^{h_1} e_2^{h_2} \dots e_m^{h_m}.$$

We notice a formal implication of this fact. A homogeneous plynomial map f on U of degree n factors through $u \to u^n$ and a uniquely determined linear map on $\Sigma_n(U)$. In other words $P_n[U]$ is canonically isomorphic to the dual of $\Sigma_n(U)$.[4]

12.9 It is important to start to introduce now 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 $< u|v >$.[5]

Choosing bases for the 2 vector spaces, the pairing determines a matrix A with entries $a_{ij} =< u_i|v_j >$.

If we change the two bases with matrices B, C the corresponding matrix becomes BAC^t.

If we fix our attention on one of the 2 variables we can equivalently think of the pairing as a linear map $j : U \to hom(V, F)$ given by $< j(u)|v >=< u|v >$.

We have used the braket 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^*.

With this idea of bilinear pairing there is asociated that of orthogonality. Given a subspace M of U its orthogonal is the subspace

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

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

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

[4] We shall repeat this argument in §17.

[5] this has been introduced in Quantum Mechanics by Dirac.

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

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, the one of symmetry:

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

One can easily see that the symmetry condition can be translated in terms of the associated map $j : U \to U^*$.

We exploit the identification $U = U^{**}$ and so the transpose map $j^* : U^{**} = U \to U^*$. Then the form is symmetric if and only if $j = j^*$ and it is antisymmetric if and only if $j = -j^*$.

Sometimes it is convenient to uniformize the 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:

(12.9.1) $$< u_1|u_2 >= \epsilon < u_2|u_1 > .$$

Example 1) The space $End(U)$ with the form $Tr(AB)$ is an example of a non degenerate symmetric bilinear form. The form is non degenerate 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. 12.7).

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

$$< (v_1, \varphi_1)|(v_2, \varphi_2) >:=< \varphi_1|v_2 > \pm < \varphi_2|v_1 > .12.9.2$$

In 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.

For non degenerate symmetric forms we have also the important notion of adjonction for operators on U. For $T \in End(U)$ one sets T^* by:

(12.9.3) $$(Tu, v) := (u, T^*v).$$

We can use the form to identify U with U^*, by identifying u with the linear form $< j(u)|v >= (u, v)$.

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

(12.9.4) $$(u \otimes v)w = u(v, w),\ (a \otimes b)(c \otimes d) = a \otimes (b, c)d,\ (a \otimes b)^* = \varepsilon b \otimes a.$$

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

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

Equivalently $T^*T = TT^* = 1$ (if the form is non degenerate).

For a non degenerate symetric form the corresponding group of orthogonal transformations is called the **orthogonal group**, for a non degenerate skew-symetric form the corresponding group of orthogonal transformations is called the **symplectic group**.

For the 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 correspond to non equivalent symmetric forms. One chooses a symmetric matrix A and the corresponding form, using as vectors column vectors, in matrix notations is

$$(u, v) := u^t A v.$$

Thus the condition 12.9.3 becomes $u^t T^t A v = u^t A T^* v$, or:

$$T^* = A^{-1} T^t A.$$

If A is the identity matrix we get the usual relation $T^* = T^t$.

Consider the case of the *standard hyperbolic form 12.9.2* 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 $A = \begin{vmatrix} 0 & 1_m \\ 1_m & 0 \end{vmatrix}$.

Similarly for the standard symplectic form we have $A = \begin{vmatrix} 0 & 1_m \\ -1_m & 0 \end{vmatrix}$.

This matrix description is connected to the standard hyperbolic form on the space $V \oplus V^*$ by Then we write a matrix T in block form $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$, and see that

(12.9.6)
$$T^* = \begin{vmatrix} d^t & b^t \\ c^t & a^t \end{vmatrix} \text{ hyperbolic adjoint}$$

(12.9.7)
$$T^* = \begin{vmatrix} d^t & -b^t \\ -c^t & a^t \end{vmatrix} \text{ symplectic adjoint}$$

One could easily write the condition for a block matrix to belong to the corresponding orthogonal or symplectic group.

Rather we deduce the Lie algebras of these groups. We have that $(e^{tX})^* = e^{tX^*},\ (e^{tX})^{-1} = e^{-tX}$ hence;

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

From the previous formulas 12.9.6,7 we get immediately an explicit description of the space of these matrices, which are denoted by $so(2n), sp(2n)$.

(12.9.8)
$$so(2n) := \{\begin{vmatrix} a & b \\ c & -a^t \end{vmatrix};\ b, c \text{ skew symmetric}\}$$

(12.9.9)
$$sp(2n) := \{\begin{vmatrix} a & b \\ c & -a^t \end{vmatrix};\ b, c \text{ symmetric}\}$$

We have thus for their dimensions:

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

We leave to the reder to describe $so(2n+1)$.

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[6] given by $\omega_A := \sum_{i<j} a_{i,j} e^i \wedge e^j$.

Definition.

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

Theorem. *i) For any invertible matrix B we have*

(12.9.11) $$Pf(BAB^t) = det(B) Pf(A).$$

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

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

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, 2×2 blocks $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ bordered by 0.

For this matrix $\Omega_{J_n} = \sum_{j=1}^n e^{2j-1} \wedge e^{2j}$, $\Omega_{J_n}^n = n! e^1 \wedge e^2 \wedge \cdots \wedge e^{2n}$, the identity is verified directly. $\square$

Exercise

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

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

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

---

[6] one refers to an elemento of $\wedge^k V^*$ as a $k-$covector.

12.10 Suppose we have a symmetric form on U, we define its associated *quadratic form* by $Q(u) :=< u|u >$. We see that $Q(u)$ is a homogeneous polynomial of degree 2. We have $Q(u+v) =< u+v|u+v >= Q(u)+Q(v)+2 < u|v >$ by the bilinearity and symmetry properties, thus:

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

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

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:

$$< u_1 \otimes v_1 | u_2 \otimes v_2 >=< u_1|u_2 >< v_1|v_2 >,$$

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

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

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

Proposition. *The tensor product of 2 non degenerate forms is non degenerate.*

Iterating the construction we have a bilinear funtion 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.

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

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

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

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

so if $u = ax+by$, $v = cx+dy$ we get

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

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

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

12.11 Given a quadratic form on a space U we can consider the ideal J of $T(U)$ generated by the elements $u^{\otimes 2} - Q(u)$, the quotient algebra $T(U)/J$ is called

The Clifford algebra of the quadratic form.

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

There are several efficient ways to study the Clifford algebra. One goes through the theory of superalgebras.

Start remarking that, although the relations defining the Clifford algebra are not homogeneous they are of even degree. This suggests the following:

Definition. A superalgebra is an algebra A decomposed as $A_0\oplus A_1$ with $A_iA_j\subset A_{i+j}$ where the indeces are taken modulo 2.

A superalgebra is thus graded modulo 2. For a homogeneous element we set $d(a)$ its degree (modulo 2). We have the obvious notion of (graded) homomorphism of superalgebras.

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 super-tensorproduct of associative superalgebras.

Given 2 superalgebras A, B we define a superalgebra:
(12.11.1)
$A\hat{\otimes}B := (A_0 \otimes B_0 \oplus A_1 \otimes B_1) \oplus (A_0 \otimes B_1 \oplus A_1 \otimes B_0),\ (a \otimes b)(c \otimes d) := (-1)^{d(b)d(c)} ac \otimes bd.$

It is left to the reader to show that this defines an associative superalgebra. In this vein of definitions we have the notion of **supercommutator** which on homogeneous elements is:

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

Accordingly we say that a superalgebra is supercommutative if $\{a, b\} = 0$ for all the elements.

The connection between supertensor product and supercommutativity is in the following (cf. 12.4):

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** ($D(ab) = D(a)b + (-1)^{d(a)} aD(b)$). Of supermodule and super tensorproduct of such supermodules.

We can now formulate the main:

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

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

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 the definitions given, that this map satisfies the universal property for the Clifford algebra and so it induces a map $\overline{j} : Cl(U) \to Cl(U_1)\hat{\otimes}Cl(U_2)$.

Now consider the 2 inclusions of U_1, U_2 in U which define two maps of $Cl(U_1) \to Cl(U)$, $Cl(U_2) \to Cl(U)$.

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

On the generating subspaces U, $U_1 \otimes 1 \oplus 1 \otimes U_2$ the maps $\overline{j}, \overline{i}$ are isomorphisms inverse of each other, hence the claim.

For a 1-dimesional 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, if we fix an orthogonal basis $u_1, \dots, u_n$ the 2^n elements $u_{i_1} u_{i_2} \dots u_{i_k},\ i_1 < i_2 < \dots < i_k$ are a basis of $Cl(U)$.

In particular if we have an orthonormal basis e_i we have the defining commuting relations $e_i^2 = 1,\ e_i e_j = -e_j e_i,\ i \neq j$.

It is useful also to present the Clifford algebra in a hyperbolic basis, i.e. the Clifford algebra of the standard quadratic form on $V \oplus V^*$.

The most efficient way to treat it is to exhibit the exterior algebra $\wedge V$ as an irreducible module over $Cl(V \oplus V^*)$. So that $Cl(V \oplus V^*) = End(\wedge V)$.

This is usually called the *spin formalism.*

For this let us define two linear maps i, j from V, V^* to $End(\wedge V)$.

(12.11.4) $$i(v)(u) := v\wedge u,\ j(\varphi)(v_1\wedge v_2 \dots \wedge v_k) := \sum_{t=1}^{k}(-1)^{t-1} < \varphi|v_t > v_1\wedge v_2 \dots \check{v}_t \dots \wedge v_k,$$

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 $\wedge V$ while $j(\varphi)$ is the superderivation induced by the contraction by φ on V.

One immediately verifies:

(12.11.5) $$i(v)^2 = j(\varphi)^2 = 0,\ i(v)j(\varphi) + j(\varphi)i(v) =< \varphi|v > .$$

We thus have that the map $i + j : V \oplus V^* \to End(\wedge V)$ 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 $End(\wedge V)$ one has several options.

One is to show directly that $\wedge V$ is an irreducible module under the Clifford algebra and then remark that, if $k = \dim(V)$ then $\dim Cl(V \oplus V^*) = 2^{2n} = \dim End(\wedge V)$. Otherwise we can view the exterior algebra as tensor product of the exterior algebras on 1 dimensional spaces. Each is a graded irreducible module over the corresponding Cliffors algebra and we get the identity by taking supertensorproducts.

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 the Clifford algebra C_{2n+1}.

Lemma. *The element $c := (-1)^n e_1e_2 \dots e_{2n+1}$ is a central idempotent of the Clifford algebra.*

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

$$e_ic = (-1)^{i-1}c = (-1)^{n-i}c = ce_i,\ c^2 = (-1)^{\sum_1^{2n} i}(-1)^n c = c.$$

Take now the Clifford algebra C_{2n} on the first $2n$ basis elements. We easily see that $C_{2n+1} = C_{2n} + C_{2n}c$. It is then clear that C_{2n+1} as an algebra is isomorphic to $C_{2n} \oplus C_{2n}$.

12.12 We collect now a few odd items. First of all when one works over the complex numbers there are several notions associated to complex conjugation[7]

Civen 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:

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

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

A linear map from $A : \overline{U} \to V$ to another vector space V is the same as a *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 Hermitian form and Hilbert space structure on a vector space U.

From the algebraic point of view 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:

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

A pre-Hilbert structure is an Hermitian form with $\|u\| := (u, u) > 0$ for all $u \neq 0$.

The Hilbert space condition is the completeness of U under the metric induced by the Hilbert norm.

In a finite dimensional space the completeness is always insured, such a Hilbert space has always an orthonormal basis u_i with $(u_i, u_j) = \delta_{ij}$.

The group of linear transformations preserving a given Hilbert structure is called the Unitary group, in an orthonormal basis it is formed by the matrices A such that $A\overline{A}^t = 1$.

The matrix $A\overline{A}^t$ 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 an orthonormal basis the matrix of the adjoint of an operator is the adjint matrix.

Given two or more Hilbert spaces one can form the tensor product of the Hilbert structures by the obvious formula $(u \otimes v, w \otimes x) := (u, w)(v, x)$.

The real and imaginary part of an 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 posiive quadratic form while $A(u, v)$ is a non degenerate alternating form.

An orthonormal basis $u_1, \dots, u_n$ for U defines a basis 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 non degenerate.

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

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

We finish with some remarks on the exterior algebra.

Let $\dim U = n$, the bilinear pairing $\wedge^k U \times \wedge^{n-k} U \to \wedge^n U$ induces a linear map

$$j : \wedge^k U \to \hom(\wedge^{n-k} U, \wedge^n U) = \wedge^n U \otimes (\wedge^{n-k} U)^*,$$

$$j(u_1 \wedge \ldots \wedge u_k)(v_1 \wedge \ldots v_{n-k}) := u_1 \wedge \ldots \wedge u_k \wedge v_1 \wedge \ldots v_{n-k}$$

In a given basis $e_1, \dots, e_n$ we have $j(e_{i_1} \wedge \ldots \wedge e_{i_k})(e_{j_1} \wedge \ldots e_{j_{n-k}}) := 0$ if the elements $i_1, \dots, i_k, j_1, \dots, j_{n-k}$ are not a permutation of $1, 2, \dots, n$ otherwise the value is $\epsilon_\sigma e_1 \wedge e_2 \ldots \wedge e_n$ where σ is the permutation that brings the elements $i_1, \dots, i_k, j_1, \dots, j_{n-k}$ in increasing order.

In particular we obtain

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

§13 Basic constructions on representations.

13.1 Having some of the formalism of tensor algebra we can go back to representation theory.

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.

Let then be given 2 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 of all for a group. We already have (cf 1.2.2) 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

$$\begin{array}{ll} \text{for } hom(V,W) & (gf)(v) := g(f(g^{-1})) \\ \text{for } V^* & < g\phi|v > := < \phi|g^{-1}v > \\ \text{for } V \otimes W & g(v \otimes w) := gv \otimes gw \end{array}$$

We can now verify:

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

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

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

First of all let us assume that G is a Lie group with Lie algebra $Lie(G)$ and let us consider a 1 parameter subgroup $exp(tA)$ generated by an element $A \in Lie(G)$.

Given a representation ϱ of G we have the induced representation $d\varrho$ of $Lie(G)$ 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 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 $< exp(tA)\varphi|v >=< \varphi|exp(-tA)v >$ we see that the Lie algebra action on V^* is given by

$$< A\varphi|v >=< \varphi| - Av >,$$

in matrix notations the contragredient action of a Lie algebra is given by minus the transpose of a given matrix.

Similarly we have the formulas:

$$A(v \otimes w) = Av \otimes w + v \otimes Aw, \ (Af)(v) = A(f(v)) - f(A(v)). \tag{13.1.1}$$

for the action on tensor product or on homomorphisms.

On the various algebras $T(U), S[U], \wedge U$ 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:

$$A(u_1 \wedge u_2 \cdots \wedge u_k) = Au_1 \wedge u_2 \cdots \wedge u_k + u_1 \wedge Au_2 \cdots \wedge u_k + \ldots u_1 \wedge u_2 \cdots \wedge Au_k. \tag{13.1.2}$$

On the Clifford algebra we have an action as derivations only of the Lie algebra of the ortogonal group of the quadratic form, since only this group preserves the defining ideal. As a consequence we have:

Proposition. *If G is a connected Lie group*

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

Proof. Same as 1.4.7.

13.2 Let analyze now, as a first elementary example, the orbit structure of some basic representations.

We start with $hom(V, W)$ thought of as a representation of $GL(V) \times GL(W)$.

It is convenient to introduce bases and use the usual matrix notation.

Let n, m be the dimensions of V, W and, using bases we identify $hom(V, W)$ to the space $M_{m\,n}$ of rectangular matrices, the group $GL(V) \times GL(W)$ is also identified to $GL(n) \times GL(m)$ and the action is $(A, B)C = BCA^{-1}$.

The notion of rank of an operator is an invariant notion, furthermore we have:

Proposition. *Two elements of* $hom(V, W)$ *are in the same* $GL(V) \times GL(W)$ *orbit if and only if they have the same rank.*

Proof. This is an elementary fact, one can give an abstract proof as follows.

Given a matrix C of rank k we choose a basis of V such that the last $n - k$ vectors are a basis of its kernel. Then the image of the first k vectors are linearly independent and we can conplete them to a basis of W. In these bases the operator has matrix (in block form):

$$\begin{pmatrix} 1_k & 0 \\ 0 & 0 \end{pmatrix}.$$

1_k is the identity matrix of size k. This matrix is otained from C by the action of the group and so it is in the same orbit, we have found a canonical representative for matrices of rank k.

In practice this abstract proof can be made into an effective algorithm, for instance Gaussian elimination on rows and columns.

As a consequence we also have:

Consider $V \otimes W$ as a representation of $GL(V) \times GL(W)$, then there are exactly $min(n, m) + 1$ orbits, formed by the tensors which can be expressed as sum of k decomposable tensors (and not less), $k = 0, \dots, min(m, n)$.

This is left to the reader using the identification $V \otimes W = hom(V^*, W)$.

Remark that these results are quite general and make no particular assumptions on the field F.

We suggest to the reader a harder exercise which is in fact quite interesting and has far reaching generalizations. Consider again the space of $m \times n$ matrices but restrict the action to $B^+(m) \times B^-(n)$ where $B^+(m)$ is the group of upper triangular matrices ($B^-(n)$ the group of lower triangular matrices). As an example show:

The orbits of $B^+(n) \times B^-(n)$ acting on $GL(n)$ by $(A, B)X := AXB^{-1}$ are in 1-1 correspondence with the symmetric group S_n.

Let us now consider the action on bilinear forms, for the moment we restrict to the simpler case of ϵ symmetric forms. Representing them as matrices the action is XAX^t.

For antisymmetric forms on U the only invariant is again the rank. For symmetric forms the same theorem holds if F is algebraically closed otherwise there are deeper arithmetical invariants, in the special case of the real numbers the signature is a sufficient invariant.

The proof is by induction, if an antisymmetric form is non zero we can find a pair of vectors on which the matrix of the form is:

$$\begin{vmatrix} 0 & 1 \\ -1 & 0 \end{vmatrix},$$

if V is the span of these vectors, the space U decomposes in the orthogonal sum $V \oplus V^\perp$. Then one procedes by induction on $V^\perp$.

For a symmetric form instead one chooses a vector of norm 1 and procedes with the orthogonal decomposition.

13.3 We want to discuss an important point of representation theory of groups. Let U be a representation of a group G. Let $\phi \in U^*, u \in U$ be given, then:

Definition. *The function* $< \phi|gu >$ *on the group* G *is called a matrix coefficient.*

Consider the bilinear map

$$i_U : U^* \times U \to F[G],\ i_U(\phi, u)(g) =< \phi|gu >$$

Proposition. *The induced linear map:*

$$i_U : U^* \otimes U \to F[G],\ i_U(\phi \otimes u)(g) =< \phi|gu >$$

is $G \times G$ *equivariant, where on* $U^* \otimes U$ *we have the actions on the two factors and on the functions we consider the left and right action.*

Proof. We have $(h \times k)(\phi \otimes u) = h\phi \otimes ku$ thus:

$$< h\phi|g(ku) >=< \phi|h^{-1}gku >= i_U(\phi, u)(h^{-1}gk)$$

as desired.

13.4 We want to complete this general discussion stressing some properties of 1-dimensional representations of a group G.

First of all, clearly, a 1-dimensional representation is just a homomorphism of G in the multiplicative group F^* of the base field. Such a homomorphism is also called a *multiplicative character*(cf. §16).

The tensor product of two 1-dimensional spaces is clearly 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, so is the inverse. The multiplicative characters of a group* G *form a group, called the character group of* G.

Notice in particular that, if V is 1-dimensional, $V \otimes V^*$ is canonically isomorphic to the trivial representation by mapping $v \otimes \phi \to< \phi|v >$ (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. *Given a 1-dimensional representation* L *and another representation* U *of a group* G *we have that* 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.

§14 Semisimple algebras.

14.1 One of the main themes of our theory will be related to completely reducible representations, it is thus important to establish these notions in full detail and generality.

Definition.

i) A set S of operators on a vector space U is *irreducible* or *simple* if the only subspaces of U which are stable under S are 0 and U.

ii) A set S of operators on a vector space U is *completely reducible* or *semisimple* if U decomposes as a direct sum of stable irreducible subspaces.

iii) A set S of operators on a vector space U is *indecomposable* if the space U cannot be decomposed in the direct sum of two non trivial stable subspaces.

Of course a space is irreducible if and only if it is completely reducible and indecomposable.

A typical example of completely reducible sets of operators is the following. Let $U = \mathbb{C}^n$ and S a set of matrices. For a matrix A denote by $A^* = \overline{A}^t$ its adjoint, it is characterized by the fact that$(Au, v) = (u, A^*v)$ for the standard Hermitian product $\sum z_i \overline{w}_i$.

Lemma. *If a subspace M of $\mathbb{C}^n$ is stable under A then $M^\perp$ (the orthogonal under the Hermitian product) is stable under A^*.*

Proof. If $m \in M, u \in M^\perp$ we have $(m, A^*u) = (Am, u) = 0$ since M is A stable. Thus $A^*u \in M^\perp$.

Proposition. *If $S = S^*$ then $\mathbb{C}^n$ is the orthogonal sum of irreducible submodules, in particular it is semisimple.*

Proof. Take an S stable subspace M of $\mathbb{C}^n$ of minimal dimension, it is then necessarily irreducible.

Consider its orthogonal complement $M^\perp$, by adjonction and the previous lemma we get that $M^\perp$ is S stable and $\mathbb{C}^n = M \oplus M^\perp$.

We then procede in the same way on $M^\perp$.

A special case is when S is a group of unitary operators.

More generally we say that S is *unitarizable* if there is an Hermitian product for which the operators of S are unitary. If we consider a matrix mapping the standard basis of $\mathbb{C}^n$ to a basis orthonormal for some given Hermitian product we see that

Lemma. *A set of matrices is unitarizable if and only if it is conjugate to a set of unitary matrices.*

These ideas have an important consequence.

Theorem. *A finite group G of linear operators on a finite dimensional complex space U is unitarizable and hence the module is semisimple.*

Proof. We fix an arbirary positive Hermitian product (u, v) on U. Define a new Hermitian product as:

$$< u, v >:= \sum_{g \in G} (gu, gv). \tag{14.1.1}$$

Then $< hu, hv >= \sum_{g \in G}(ghu, ghv) = \sum_{g \in G}(gu, gv) =< u, v >$ and G is unitary for this new product.

The previous theorem has a far reaching generalization, by replacing the average given by the sum with an integral. Recall some definitions.

Definition. *A topological group is a group with a topology such that the multiplication and the inverse are continuous maps.*

Given a locally compact Hausdorff topological space X, let $\mathcal{C}^0(X)$ denote the space of continuous functions with compact support.

Definition. *An integral is a linear map $I : \mathcal{C}^0(X) \to \mathbb{R}$ such that $f \geq 0$ implies $I(f) \geq 0$.*

Recall that, for a function on a group we have the right action $f^h(g) := f(gh)$.

Definition. *An integral on a locally compact group G is right invariant if $I(f) = I(f^h)$ for every continuous function with compact support and every $h \in G$. (Similarly for a left invariant integral).*

Theorem. *On a locally compact group there exists a right invariant integral which is unique up to scale ($I' = \alpha I$).*

This is the theorem of existence of Haar measure. In general a right invariant measure is not left invariant, it is so for compact groups (and for several other interesting groups).

If G is compact we can integrate all continuous functions and so as in the previous theorem we construct $\int_G (gu, gv)dg =< u, v >$.

Theorem. *A compact group G of linear operators on a finite dimensional complex space U is unitarizable and hence the module is semisimple.*

We might remark that the unitary group $U(n, \mathbb{C}) := \{A | AA^* = 1\}$ is a bounded and closed set in $M_n(\mathbb{C})$ hence it is compact. Thus from all the previous remarks we have.

Proposition. *$U(n, \mathbb{C})$ is a maximal compact subgroup of $GL(n, \mathbb{C})$ any other maximal compact subgroup of $GL(n, \mathbb{C})$ is conjugate to $U(n, \mathbb{C})$.*

Proof. Let K be a compct linear group, since it is unitarizable there exists a matrix g such that $K \subset gU(n, \mathbb{C})g^{-1}$. If K is maximal this inclusion is an equality.

For non compact groups there is an important class that we have already introduced for which similar results are valid.

These are the self adjoint subgroups of $GL(n, \mathbb{C})$.

For a self adjoint group on a given (finite dimensional) Hilbert space U the orthogonal of every invariant subspace is also invariant, thus any subspace or quotient module of U is completely reducible.

Take a self adjoint group G and consider its induced action on the tensor algebra. The tensor powers of $U = \mathbb{C}^n$ have an induced canonical Hermitian form for which:

$$< u_1 \otimes u_2 \otimes \ldots \otimes u_n | v_1 \otimes v_2 \otimes \ldots \otimes v_n > = < u_1 | v_1 >< u_2 | v_2 > \ldots < u_n | v_n >$$

It is clear that this is a positive Hermitian form for which the tensor power of an orthonormal basis is also an orthonormal basis.

The map $g \to g^{\otimes n}$ is compatible with adjonction, i.e. $(g^*)^{\otimes n} = (g^{\otimes n})^*$:

$$(g^{\otimes m}(v_1 \otimes v_2 \otimes \ldots \otimes v_m) | w_1 \otimes w_2 \otimes \ldots \otimes w_m) := \prod_{i=1}^{m} (g v_i | w_i) =$$

$$\prod_{i=1}^{m} (v_i | g^* w_i) = (v_1 \otimes v_2 \otimes \ldots \otimes v_m | (g^*)^{\otimes m}(w_1 \otimes w_2 \otimes \ldots \otimes w_m))$$

Thus:

Proposition. *The action of G on $T(U)$ is self adjoint, hence all tensor powers of U are completely reducible under G.*

Corollary. *The action of G on the polynomial ring $\mathcal{P}[U]$ is completely reducible.*

Proof. The action of G on U^* is selfadjoint, so is on $T[U^*]$ and $\mathcal{P}[U]$ is a quotient of $T[U^*]$ (for each degree).

14.2 It is usually more convenient to use the language of modules since the irreducibility or complete reducibility of a space U under a set S of operators is clearly equivalent to the same property under the subalgebra of operators generated by S.

Let us reacall in a slightly greater generality, the results of 3.1. Given a group G or even a semigroup, one can form its *group algebra* denoted $F[G]$.

Definition. $F[G]$ is the set of formal linear combinations of elements of G where the multiplication extends bilinearly the one of G.

The following statement can be immediately verified.

Proposition. *Every linear representation of G extends by linearity to an $F[G]$ module and conversely. Moreover a map between $F[G]$ modules is a (module) homomorphism if and only if it is G equivariant.*

Thus from the point of view of representation theory it is equivalent to study the category of G representations or that of $F[G]$ modules.

14.3 We consider thus a ring R and its modules, using the same definitions for reducible, irreducible modules. We define $R^\vee$ to be the set of (isomorphism classes) of irreducible modules of R, the *spectrum* of R.

Given an irreducible module N we will say that it is of *type* α if α indicates its isomorphism class.

Given a set S of operators on U we set $S' := \{A \in End(U) | As = sA,\ \forall s \in S\}$, S' is called the *centralizer* of S, equivalently it should be thought as the set of all S linear endomorphisms. One immediately verifies:

Proposition.
i) S' *is an algebra.*
ii) $S \subset S''$
iii) $S' = S'''$

The centralizer of the operators induced by R in a module M is also usually indicated by $End_R(M, M)$ or $End_R(M)$ and called the *endomorphism ring.*

Any ring R can be considered as a module on itself by left multiplication (and as a module on the opposite of R by right multiplication), this module is usually called *the regular representation*; of course in this case a submodule is the same as a left ideal, an irreducible submodule is also referred to as a *minimal left ideal.*

A trivial but useful fact on the regular representation is:

Proposition. *The ring of endomorphisms of the regular representation is the opposite of R acting by right multiplications.*

Proof. Let $f \in End_R(R)$ we have $f(a) = f(a1) = af(1)$ by linearity, thus f is the right multiplication by $f(1)$.

Given two homomorphisms f, g we have $fg(1) = f(g(1)) = g(1)f(1)$ and so the mapping $f \to f(1)$ is an isomorphism between $End_R(R)$ and R^0.

One can generalize the previous considerations as follows: Let R be a ring.

Definiton. *A cyclic module is a module generated by a single element.*

A cyclic module should be thought as the linear analogue of a single orbit.

The structure of cyclic modules is quite simple, if M is generated by an element m we have the map $\varphi : R \to M$ given by $\varphi(r) = rm$ (analogue of the orbit map).

By hypothesis φ is surjective, its kernel is a left ideal J and so M is identified to R/J.

Thus a module is cyclic if and only if it is a quotient of the regular representation.

Given 2 cyclic modules R/J, R/I we can compute $Hom_R(R/J, R/I)$ as follows (cf. Exercise).

If f is a homomorphism it can be lifted to a R mapping of R to R, thus a map of the form $r \to rx$ for some $x \in R$ and this map must send J into I.

Thus we define the set

$$(I : J) := \{x \in R | Jx \subset I\},$$

(in particular for $J = I$ we have the *idealizer* $\mathcal{I}(J)$ of J , $\mathcal{I}(J) := \{x \in R | Ix \subset I\}$, i.e. the idealizer is the maximal subring of R in which J is a two sided ideal).

We then see that $Hom_R(R/J, R/I) = (I : J)/I$. In particular the ring $End_R(R/J) = \mathcal{I}(J)/J$ (cf. Chap 1 inserire)

Example Consider for A the full ring of $n \times n$ matrices over a field F. As a module we take F^n and in it the basis element e_1.

Its annihilator is the left ideal I_1 of matrices with the first column 0. In this case though we have a more precise picture.

Let J_1 denote the left ideal of matrices having 0 in all columns except the first. Then $M_n(F) = J_1 \oplus I_1$ and the map $a \to ae_1$ restricted to J_1 is an isomorphism.

In fact we can define in the same way J_i (the matrices with 0 outside the i^{th} column).

Then $M_n(F) = \oplus_{i=1}^n J_i$ is a direct sum of the algebra $M_n(F)$ into irreducible left ideals isomorphic, as modules, to the representation F^n.

Remark The same proof, with small variations, applies with a division algebra D in place of F.

Lemma. *The module D^n is irreducible, we will call it the standard module of $M_n(D)$.*

Proof. Let us consider a column vector u with its i^{th} coordinate u_i non zero. Acting on u with a diagonal matrix which has u_i^{-1} in the i^{th} position we transform u into a vector with i^{th} coordinate 1. Acting with elementary matrices we can make all the other coordinates 0, finally acting with a permutation matrix we can bring 1 in the first position. This shows that any submodule contains the vector of coordinates $(1, 0, 0, \dots, 0)$. This vector, in turn, generates the entire space again acting on it by elementary matrices.

Theorem. *The regular representation of $M_n(D)$ is the direct sum of n copies of the standard module.*

Remark In order to understand $M_m(D)$ as module endomorphisms we have to take D^n as a *right* vector space over D or as a left vector space over D^0.

Another action of some interest is the action of $R \otimes R^0$ on R given by $a \otimes b(c) := acb$; in this case the submodules are the 2 sided ideals and the centralizer is easily seen to be the *center* of R.

The example of matrices suggests the following:

Definition. We say that a ring R is *semisimple* if it is semisimple as a left module on itself.

This definition is a priori not symmetric although it will be proved to be so from the structure theorem of semisimple rings.

Remark Let us decompose a semisimple ring R as direct sum of irreducible left ideals, since 1 generates R and it must be in a finite sum of the given sum we see:

Proposition. *A semisimple ring is a direct sum of finitely many minimal left ideals.*

Corollary. *If D is a division ring then $M_m(D)$ is semisimple.*

14.4 We wish to collect some examples of semisimple rings.

First of alll from the results in 14.1 and 14.2 we deduce:

Theorem. *The group algebra $\mathbb{C}[G]$ of a finite group is semisimple.*

Remark. In fact it is not difficult to generalize to an arbitrary field. The general statement is:

The group algebra $F[G]$ of a finite group over a field F is semisimple if and only if the characteristic of F does not divide the order of G.

Next we have the obvious fact:

Proposition. *The direct sum of two semisimple rings is semisimple.*

In fact we let the following simple exercise to the reader.

EXERCISE To decompose a ring A in a direct sum of two rings is equivalent to give an element $e \in A$ such that:

$$e^2 = e, \ ea = ae, \ \forall a \in A. \ e \text{ is called a central idempotent.}$$

Having a central idempotent e, every A module M decomposes canonically as

$$eM \oplus (1 - e)M.$$

Thus if $A = A_1 \oplus A_2$ the module theory of A reduces to the one of A_1, A_2.

From the previous paragraph and these remarks we deduce:

Theorem. *A ring $A := \oplus_i M_{n_i}(D_i)$, with the D_i division algebras is semisimple.*

14.5 Our next task will be to show that also the converse to theorem 14.4 is true, i.e. that every semisimple ring is a finite direct sum of rings of type $M_m(D)$, D division ring.

For the moment we collect one further remark, let us recall that:

Definition. A ring R is called *simple* if it does not possess any non trivial two sided ideals.

Equivalently it is irreducible as a module over $R \otimes R^0$ under the left and right action $(a \otimes b)r := arb$.

This definition is slightly confusing since a simple ring is by no means semisimple, unless it satisfies further properties (the d.c.c. on left ideals). A classical example is the algebra of differential operators $F < x_i, \dfrac{\partial}{\partial x_i} >$ (F a field of characteristic 0).

Nevertheless we have:

Proposition. *If D is a division ring $M_m(D)$ is simple.*

Proof. Let I be a non trivial two sided ideal, $a \in I$ a non zero element. We write a as a linear combination of elementary matrices $a = \sum a_{ij} e_{ij}$, thus $e_{ii} a e_{jj} = a_{ij} e_{ij}$ and at least one of these elements must be non zero. Multiplying it by a scalar matrix we can obtain an element e_{ij} in the ideal I, then we have $e_{hk} = e_{hi} e_{ij} e_{jk}$ and we see that the ideal coincides with the full ring of matrices.

Exercise The same argument shows more generally that for any ring A the ideals of the ring $M_m(A)$ are all of the form $M_m(I)$ for I an ideal of A.

14.6 We start now with the general theory and with the following basic facts:

Theorem. *(Schur's lemma) The centralizer $\Delta := End_R(M, M)$ of an irreducible module M is a division algebra.*

Proof. Let $a : U \to U$ be a non zero R linear endomorphism. Its kernel and image are submodules of M. Since M is irreducible and $a \neq 0$ we must have $Ker(a) = 0$, $Im(a) = M$ hence a is an isomorphism and so it is invertible. This means that every non 0 element in Δ is invertible, this is the definition of a division algebra.

This lemma has several variations, the same proof shows that:

Corollary. *If $a : M \to N$ ia a homomorphism between two irreducible modules then either $a = 0$ or a is an isomorphism.*

14.7 A particularly important case is when U is a finite dimensional vector space over $\mathbb{C}$, in this case since the only division algebras over $\mathbb{C}$ is $\mathbb{C}$ itself we have that:

Theorem. *Given an irreducible set S of operators on a finite dimensional space over $\mathbb{C}$ then its centralizer S' is formed by the scalars $\mathbb{C}$.*

Proof. Rather than applying the structure theorem of finite dimensional dvision algebras one can argue that, given an element $x \in S'$ and an eigenvalue α of x the space of eigenvectors of x for this eigenvalue is stable under S and so, by irreducibility, it is the whole space hence $x = \alpha$.

Remarks. 1. If the base field is the field of real numbers we have (according to the theorem of Frobenius) 3 possibilities for Δ: $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$ the algebra of quaternions.

2. It is not necessary to assume that U is finite dimensional, it is enough to assume that it is of countable dimension.

In fact U is also a vector space over Δ and so Δ, being isomorphic to a $\mathbb{C}$ (or $\mathbb{R}$) subspace of U is also countably dimensional.

This implies that every element of Δ is algebraic over $\mathbb{R}$. Otherwise Δ would contain a field isomorphic to the rational function field $\mathbb{R}(t)$ which is impossible, since this field contains the uncountably many linearly independent elements $\frac{1}{t-r}$, $r \in \mathbb{R}$.

Now one can prove that a division algebras over $\mathbb{R}$ in which every element is algebraic is necessarily finite dimensional[8] and thus the theorem of Frobenius applies.

In order to understand semisimple algebras from this point of view we make a general remark about matrices.

Let $M = M_1 \oplus M_2 \oplus M_3 \oplus \ldots \oplus M_k$ be an R module decomposed in a direct sum.

For each i, j consider $A(j,i) := hom_R(M_i, M_j)$. For 3 indeces we have the composition map $A(k,j) \times A(j,i) \to A(k,i)$.

The groups $A(j,i)$ together with the composition maps allow us to recover the full endomorphism algebra of M as **block matrices**:

$$A = (a_{ji}), \ a_{ji} \in A(j,i).$$

(One can give a formal abstract construction starting from the associativity properties).

In more concrete form let $e_i \in End(M)$ be the projection on the summand M_i with kernel $\oplus_{j \neq i} M_j$. The elements e_i are *orthogonal idempotents in* $End(M)$ i.e. they satisfy the properties

$$e_i^2 = e_i, \ e_i e_j = e_j e_i = 0, \ i \neq j, \text{ and } \sum_{i=1}^{k} e_i = 1.$$

When we have in a ring S such a set of idempotents we decompose S as

$$S = (\sum_{i=1}^{k} e_i) S (\sum_{i=1}^{k} e_i) = \oplus_{i,j} e_i S e_j.$$

This sum is direct by the remarks on idempotents.

We have $e_i S e_j e_j S e_k \subset e_i S e_k, \ e_i S e_j e_h S e_k = 0, \ j \neq h$. In our case $S = End_R(M)$ and $e_i S e_j = hom_R(M_j, M_i)$.

In particular assume that the M_i are all isomorphic to a module N and let $A := End_R(N)$ then:

$$End_R(N^{\oplus k}) = M_k(A).$$

Assume now we have two modules N, P such that $hom_R(N,P) = hom_R(P,N) = 0$, let $A := End_R(N)$, $B := End_R(P)$ then:

$$End_R(N^{\oplus k} \oplus P^{\oplus h}) = M_k(A) \oplus M_h(B).$$

Clearly we have a similar statement for several modules.

We can add together all these remarks in the case in which a module is a finite direct sum of irreducibles.

---

[8]this depends on the fact that every element algebraic over   satisfies a quadratic polynomial.

Assume $N_1, N_2, \dots, N_k$ are the distinct irreducible which appear with multiplicities $h_1, h_2, \dots, h_k$, let $D_i = End_R(N_i)$ (a division ring) then:

$$End_R(\oplus_{i=1}^k N_i^{h_i}) = \oplus_{i=1}^k M_{h_i}(D_i). \quad (14.7.1)$$

14.8 We are now ready to characterize semisimple rings. If R is semisimple we have that $R = \oplus_{i=1}^k N_i^{m_i}$ as in the previous section as a left R module, then:

$$R^0 = End_R(R) = \oplus_{i\in I} N_i^{m_i} = \oplus_{i\in I} M_{m_i}(\Delta_i).$$

We deduce that $R = R^{00} = \oplus_{i\in I} M_{m_i}(\Delta_i)^0$.

The opposite of the matrix ring over a ring A is the matrices over the opposite ring (use transposition) and so we deduce finally:

Theorem. *A semisimple ring is isomorphic to the direct sum of matrix rings R_i over division algebras.*

Some comments are in order.

1. We have seen that the various blocks of this sum are simple rings, they are thus distinct irreducible representations of the ring $R \otimes R^0$ acting by the left and right action.

We deduce that the matrix blocks are minimal 2 sided ideals. From the theory of isotypic components which we will discuss it follows that the only ideals of R are direct sums of these minimal ideals.

2. We have now the left right symmetry, if R is semisimple so is also R^0.

Since any irreducible module N is cyclic there is a surjective map $R \to N$. This map restricted to one of the N_i must be non zero hence:

Corollary. *Each irreducible R module is isomorphic to one of the N_i (appearing in the regular representation).*

14.9 We will complete the theory with some general remarks:

Lemma. *Given a module M and two submodules A, B such that A is irreducible, either $A \subset B$ or $A \cap B = 0$.*

Proof. Trivial since $A \cap B$ is a submodule of A and A is irreducible.

Lemma. *Given a module M a submodule N and an element $m \notin N$ there exists a maximal submodule $N_0 \supset N$ such that $m \notin N_0$. M/N_0 is indecomposable.*

Proof. Consider the set of all submodules containing N and which do not contain m, this has a maximal element since it satisfies the hypotheses of Zorn's lemma. If we could decompose M/N then the class of m cannot be contained in both summands and we could find a larger submodule not containing m.

The basic fact on semisimple modules is the following:

Theorem. *For a module M the following conditions are equivalent:*

i) M is a sum or irreducible submodules.

ii) Every submodule N of M admits a complement, i.e. a submodule P such that $M = N \oplus P$.

iii) M is completely reducible.

Proof. This is a rather abstract theorem and the proof is correspondingly abstract.

iii implies i clearly, so we prove i) implies ii) implies iii).

i) implies ii). Assume i) holds and write $M = \sum_{i \in I} N_i$ where I is some set of indeces.

For a subset A of I set $N_A := \sum_{i \in A} N_i$. Let N be a given submodule and consider all subsets A such that $N \cap N_A = 0$, it is clear that these subsets satisfy the conditions of Zorn's lemma and so we can find a maximal set among them, let this be A_0. We claim that $M = N \oplus N_{A_0}$.

For every $i \in I$ consider $(N + N_{A_0}) \cap N_i$. If $(N + N_{A_0}) \cap N_i = 0$ we have that $i \notin A_0$. We can add i to A_0 getting a contradiction to the maximality.

Hence by the first lemma $N_i \subset (N + N_{A_0})$ and since i is arbitrary $M = \sum_{i \in I} N_i \subset (N + N_{A_0})$ as desired.

ii) implies iii). Assume ii) and consider the set J of all irreducible submodules of M (at this point we do not even know if it is not empty!).

Consider all the subsets A of J for which the modules in A form a direct sum. This clearly satisfies the hypotheses of Zorn's lemma and so we can find a maximal set adding to a submodule N.

We must prove that $N = M$. Otherwise we can find an element $m \notin N$ and a maximal submodule $N_0 \supset N$ such that $m \notin N$. By hypothesis there is a direct summand P of N_0.

We claim that P is irreducible, in fact otherwise let T be a non trivial submodule of P and consider a direct summand Q of $N_0 \oplus T$, we have thus that M/N_0 is isomorphic to $T \oplus Q$ and so is decomposable, against the conclusions of lemma 2.

P irreducible is also a contradiction since P and N form a direct sum and this contradicts the maximal choice of N as direct sum of irreducibles.

Comment If the reader is confused by the transfinite induction he should easily realize that all these inductions, in the case in which M is a finite dimensional vector space, can be replaced with ordinary inductions on the dimensions of the various submodules constructed.

Corollary. *Given a semisimple module M as a sum $\sum_{i \in I} N_i$ of irreducible submodules, we can extract from this sum a direct sum.*

Proof. We consider a maximal direct sum out of the given one, then any other irreducible module N_i must be in the sum and so this sum gives M.

Corollary. *Given a semisimple module M which is given as a direct sum $\oplus_{i\in I}N_i$ of irreducible submodules and an irreducible submodule N of M then the projection to one of the N_i must be an isomorphism.*

14.10

Corollary.

i) Submodules and quotients of a semisimple module are semisimple, as well as direct sums of semisimple modules.

ii) R is semisimple if and only if every R module is semisimple, in this case its spectrum is finite and consists of the irreducible modules appearing in the regular representation.

iii) If R has a faithful semisimple module M then R is semisimple.

Proof. i) Since the quotient of a sum of irreducible modules is again a sum of irreducibles the statement is clear for quotients. But every submodule has a complement and so it is isomorphic to a quotient. For direct sums the statement is clear.

ii) Every R module is a quotient of a free module so this statement follows immediately from i), remark 14.3 and a variation of corollary 14.6. In fact we consider $R = \oplus_{i\in I}N_i$, we take any irreducible module N a non 0 element $n \in N$. The map $R \to N$ given by $r \to rn$ is not 0 and so induces an isomorphism between one of the summands N_i and N.

iii) Take the direct sum of as many copies of M as necessary (for instance one copy $M(m)$ for each element $m \in M$ and take $\oplus_{m\in M}M(m)$) which is then a semisimple module.

Map $R \to \oplus_{m\in M}M(m)$ by $r \to (rm)_{m\in M}$. This map is clearly injective so R, as a submodule of a semisimple module, is semisimple.

14.11 An essential notion in the theory of semisimple modules is that of *isotypic component.*

Definition. Given an isomorphism class of irreducible representations, i.e. a point of the spectrum $\alpha \in R^\vee$ and a module M we set M^α to be the sum of all the irreducible submodules of M of type α. This submodule is called *the isotypic component of type α.*

Let us also use the notation M_α to be the sum of all the irreducible submodules of M which are *not* of type α.

Theorem. *The isotypic components of M decompose M in a direct sum.*

Proof. We must only prove that, given an isomorphism class α, $M^\alpha \cap M_\alpha = 0$.

M^α can be presented as a direct sum of irreducibles of type α, while M_α can be presented as a direct sum of irreducibles of type different from α.

Thus every irreducible submodule in their intersection must be 0, otherwise by Corollary 14.9 2 it is at the same time of type α and of type different from α. From the third corollary any submodule is semisimple and so this implies that the intersection is 0.

Proposition. *Given any homomorphism $f : M \to N$ between semisimple modules it induces a morphism $f_\alpha : M^\alpha \to N^\alpha$ for every α and f is the direct sum of the f_α. Conversely:*

$$hom_R(M, N) = \oplus_\alpha hom(M^\alpha, N^\alpha).$$

Proof. The image under a homomorphism of an irreducible module of type α is either 0 or of the same type, since the isotypic component is the sum of all the submodules of a given type and since each module is the direct sum of its isotypic components the claim follows.

Now that we have the canonical decomposition $M = \oplus_{\alpha \in R^\vee} M^\alpha$ we can consider the projection $\pi^\alpha : M \to M^\alpha$ with kernel M_α. We have:

14.12

Proposition. *Given any homomorphism $f : M \to N$ between semisimple modules we have a commutative diagram for each $\alpha \in R^\vee$:*

$$\begin{array}{ccc} M & \xrightarrow{f} & N \\ {\scriptstyle \pi^\alpha}\downarrow & & {\scriptstyle \pi^\alpha}\downarrow \\ M^\alpha & \xrightarrow{f} & N^\alpha \end{array}$$

Proof. This is an immediate consequence of the proposition 14.7.

14.13 This rather formal analysis has an important implication. Let us assume that we have a group G acting as automorphisms on an algebra A, let us furthermore assume that A is semisimple as a G module. Thus the subalgebra of invariants A^G is the isotypic component of the trivial representation.

Let us denote by A_G the sum of all the other irreducible representations so that $A = A^G \oplus A_G$.

Definition. *The canonical projection $\pi^G : A \to A^G$ is usually indicated by the symbol R and called the Reynolds operator.*

R satisfies the general properties of 14.8. Let us now consider an element $a \in A^G$; since, by hypothesis, G acts as algebra automorphisms, both left and right multiplication by a are G equivariant. We thus can apply the general proposition 14.8 and deduce the so called *Reynolds identities.*

Proposition.

$$R(ab) = aR(b), \;\; R(ba) = R(b)a, \;\; \forall b \in A, \; a \in A^G.$$

We will see that these identities are the main tool to develop the theory of Hilbert on Invariants of forms (and its generalizations).

14.14 Although the theory could be pursued in the generality of Artinian rings let us revert to finite dimensional representations.

Let $R = \oplus_{i \in I} R_i = \oplus_{i \in I} M_{m_i}(\Delta_i)$ be a finite dimensional semisimple algebra over a field F. In particular now all the division algebras Δ_i will be finite dimensional over F.

In the case of the complex numbers (or of an algeraically closed field) they will coincide with F, for the real numbers we have the 3 possibilities already discussed.

If we consider any finite dimensional module M over R we have seen from 14.10 that M is isomorphic to a finite sum

$$M = \oplus_i M_i = \oplus_{i \in I} N_i^{p_i}$$

where M_i is the isotypic component relative to the block R_i and $N_i = \Delta_i^{m_i}$.

We have also from 14.7.1:

$$S := End_R(M) = \oplus_{i \in I} End_R(M_i) = \oplus_{i \in I} End_R(N_i^{p_i}) = \oplus_{i \in I} M_{p_i}({\Delta_i}^0) := \oplus_{i \in I} S_i.$$

The block S_i acts on $N_i^{p_i}$ as m_i copies of its standard representation, and as zero on the other isotypic components.

In fact by definition $S_i = End_R(N_i^{p_i})$ acts as 0 on all isotypic components different from the i^{th} one.

As for the action on this component we may identify $N_i := \Delta_i^{m_i}$ and thus the space $N_i^{p_i}$ as the set $M_{m_i, p_i}(\Delta_i)$ of $m_i \times p_i$ matrices. Multiplication on the right by a $p_i \times p_i$ matrix with entries in Δ_i induces a typical endomorphism in S_i. The algebra of such endomorphisms is isomorphic to $M_{p_i}(\Delta_i^0)$ and the m_i subspaces of $M_{m_i, p_i}(\Delta_i)$ formed by the rows decompose this space into irreducible representations of S_i isomorphic to the standard representation $(\Delta_i^0)^{p_i}$.

Summarizing we have:

Theorem.

i) Given a finite dimensional semisimple R module M the centralizer S of R is semisimple.

ii) The isotypic components of R and S coincide.

iii) The multiplicities and the dimensions (relative to the corresponding division ring) of the irreducibles appearing in an isotypic component are exchanged, passing from R to S.

iv) If for a given i, $M_i \neq 0$ then the centralizer of S on M_i is R_i (or rather the ring of operators induced by R_i on M_i), in particular if R acts faithfully on M we have $R = S' = R''$ (double centralizer theorem).

All the statements are implicit in our previous analysis.

We wish to restate the result in case $F = \mathbb{C}$ for a semisimple algebra of operators as follows:

Given two sequences of positive integers $m_1, m_2, \ldots, m_k$ and $p_1, p_2, \ldots, p_k$ we form the two semisimple algebras $A = \oplus_{i=1}^k M_{m_i}(\mathbb{C})$ and $B = \oplus_{i=1}^k M_{p_i}(\mathbb{C})$.

We form also the vector space $W = \oplus_{i=1}^{k} \mathbb{C}^{m_i} \otimes \mathbb{C}^{p_i}$ and consider A, B as commuting algebras of operators on W in the obvious way, i.e.

$(a_1, a_2, \ldots, a_k) \sum u_i \otimes v_i = \sum a_i u_i \otimes v_i$ for A and $(b_1, b_2, \ldots, b_k) \sum u_i \otimes v_i = \sum u_i \otimes b_i v_i$ for B

then:

Corollary. *Given a semisimple algebra R of operators on a finite dimensional vector space M over $\mathbb{C}$ and calling $S = R'$ its centralizer then there exists two sequences of integers $m_1, m_2, \ldots, m_k$ and $p_1, p_2, \ldots, p_k$ and an isomorphism of M with $W = \oplus_{i=1}^{k} \mathbb{C}^{m_i} \otimes \mathbb{C}^{p_i}$ under which the algebras R, S are identified to the algebras A, B.*

This corollary gives a very precise information on the nature of the two algebras since it claims that:

On each isotypic component we can find a basis indexed by pairs of indeces such that, if we order the pairs by setting first all the pairs which *end* with 1 then all that *end* with 2 and so on, the matrices of R appear as diagonal block matrices.

Similarly for the matrices of S if we order the indeces by setting first all the pairs which *begin* with 1 then all that *begin* with 2 and so on.

Let us continue a moment with the same hypotheses as in the previous section. Choose a semisimple algebra $A = \oplus_{i=1}^{k} M_{m_i}(\mathbb{C})$ and two representations:

$$W_1 = \oplus_{i=1}^{k} \mathbb{C}^{m_i} \otimes \mathbb{C}^{p_i}, \quad \text{and } W_2 = \oplus_{i=1}^{k} \mathbb{C}^{m_i} \otimes \mathbb{C}^{q_i}$$

which we have presented as decomposed into isotypic components, according to 14.7.1 we can compute thus:

$$Hom_A(W_1, W_2) = \oplus_{i=1}^{k} Hom_A(\mathbb{C}^{m_i} \otimes \mathbb{C}^{p_i}, \mathbb{C}^{m_i} \otimes \mathbb{C}^{q_i}) = \oplus_{i=1}^{k} Hom_{\mathbb{C}}(\mathbb{C}^{p_i}, \mathbb{C}^{q_i}).$$

We will need this computation for the theory of invariants.

14.15 We want to deduce an important application. Let H, K be two groups (not necessarily finite) and let us give two finite dimensional irreducible representations U, V of these two groups over $\mathbb{C}$.

Proposition. *$U \otimes V$ is an irreducible representation of $H \times K$, any finite dimensional irreducible representation of $H \times K$ is of this form.*

Proof. The maps $\mathbb{C}[H] \to End(U)$, $\mathbb{C}[K] \to End(V)$ are surjective hence the map $\mathbb{C}[H \times K] \to End(U) \otimes End(V) = End(U \otimes V)$ is also surjective and so $U \otimes V$ is irreducible.

Conversely assume that we are given an irreducible representation W of $H \times K$ so that the image of the algebra $\mathbb{C}[H \times K]$ is the whole algebra $End(W)$.

Let W' be the sum of all irreducible H submodules of a given type appearing in W, since K commutes with H we have that W' is K stable, since W is irreducible we have $W = W'$ so W is semisimple and consists of a unique isotypic component under H.

The algebra of operators induced by H is isomorphic to a full matrix algebra $M_n(\mathbb{C})$ its centralizer is isomorphic to $M_m(\mathbb{C})$, W is nm dimensional and $M_n(\mathbb{C}) \otimes M_m(\mathbb{C})$ is isomorphic to $End(W)$.

Finally the image R of $\mathbb{C}[K]$ is contained in the centralizer $M_m(\mathbb{C})$, but since the operators from $H \times K$ span $End(W)$ this algebra must coincide with the centralizer and the theorem follows from the general theory.

This theorem has an important application to matrix coefficients.

Let U be a (finite dimensional) irreducible representation of G then $U^* \otimes U$ is an irreducible representation of $G \times G$ thus the map $i_U : U^* \otimes U \to \mathbb{C}[G]$ is injective.

Theorem. *$i_U(U^* \otimes U)$ equals the isotypic component of type U in $\mathbb{C}[G]$ under the right action and equals the isotypic component of type U^* in $\mathbb{C}[G]$ under the left action.*

Proof. We do it for the right action, the left is similar.

Let us consider a $G-$equivariant embedding $j : U \to \mathbb{C}[G]$ where $\mathbb{C}[G]$ is considered as G module under right action, we must show that its image is in $i_U(U^* \otimes U)$.

Let $\phi \in U^*$ be defined by:

$$< \phi | u > := j(u)(1).$$

Then:

$$j(u)(g) = j(u)(1g) = j(gu)(1) = < \phi | gu > = i_U(\phi \otimes u)(g). \tag{14.15.1}$$

Thus $j(u) = i_U(\phi \otimes u)$.

Remark. The theorem proved is completely general and refers to any group. It will be possible to apply it also to continuous representations of topological groups and to rational representations of algebraic groups.

14.16 We discuss now the Jacobson density theorem, this is a generalization of Wedderburn's theorem which we will discuss presently.

Theorem. *Let N be an irreducible R module, Δ its centralizer,*

$$u_1, u_2, \ldots, u_n \in N$$

elements which are linearly independent relative to Δ and

$$v_1, v_2, \ldots, v_n \in N$$

arbitrary. Then there exists an element $r \in N$ such that $ru_i = v_i, \ \forall i$.

Proof. The theorem states that the free module R^n is generated over R by the element $a := (u_1, u_2, \ldots, u_n)$.

Since R^n is completely reducible, R^n decomposes as $Ra \oplus P$. Let $\pi \in End_R(R^n)$ be the projection to P vanishing on the submodule Ra.

By the previous analysis this operator is given by an $n \times n$ matrix d_{ij} in Δ and so we have $\sum_j d_{ij} u_j = 0, \ \forall i$ since these are the components of $\pi(a)$.

By hypothesis the elements $u_1, u_2, \ldots, u_n \in N$ are linearly independent over Δ, thus the elements d_{ij} must be 0 and so $\pi = 0$ and $P = 0$ as desired.

The name density comes from the fact that one can define a topology (of finite approximations) on the ring $End_\Delta(N)$ so that R is dense in it (cf. [J]).

14.17 Let again N be an irreducible R module Δ its centralizer, assume that N is a finite dimensional vector space over Δ of dimension n.

There are a few formal difficulties in the non commutative case to be discussed.

If we choose a basis $u_1, u_2, \ldots, u_n$ of N we identify N with Δ^n. Given the set of $n-$tuples of elements of a ring A thought as column vectors, we can act on the left with the algebra $M_n(A)$ of $n \times n$ matrices. This action commutes clearly with the multiplication *on the right* by elements of A.

If we want to think of operators as always acting on the left then we have to think of left multiplication for the *opposite* ring A^0.

We thus have dually the general fact that the endomorphism ring of a free module of rank n on a ring A is the ring of $n \times n$ matrices over A^0. We return now to modules.

Theorem. *(Wedderburn) R induces on N the full ring $End_\Delta(N)$ isomorphic to the ring of $m \times m$ matrices $M_m(\Delta)$.*

Proof. This is an immediate consequence of the density theorem, taking a basis $u_1, u_2, \ldots, u_n$ of N.

We end our abstract discussion with another generality on characters.

Let R be an algebra over $\mathbb{C}$ (we make no assumption of finite dimensionality). Let M be a finite dimensional semisimple representation. The homomorphism $\rho_M : R \to End(M)$ allows to define a *character* on R stting $t_N(a) := tr(\rho_M(a))$.

Theorem. *Two finite dimensional semisimple modules M, N are isomorphic if and only if they have the same character.*

Proof. It is clear that if the two modules are isomorphic the traces are the same, we must prove the converse.

Let I_M, I_N be the kernels respectively of ρ_M, ρ_N.

By the theory of semisimple algebras we know that R/I_M is isomorphic to a direct sum $\oplus_i M_{n_i}(\mathbb{C})$ of matrix algebras and similarly for R/I_N.

Assume that M decomposes under $\oplus_{i=1}^k M_{n_i}(\mathbb{C})$ with multiplicity p_i for the i^{th} isotypic component.

Then the trace of an element $(a_1, a_2, \ldots, a_k)$ as operator on M is $\sum_{i=1}^k p_i Tr(a_i)$ where $Tr(a_i)$ is the ordinary trace as an $n_i \times n_i$ matrix.

We deduce that the bilinear form $tr(ab)$ is non degenerate on R/I_M and so I_M is the kernel of the form induced by this trace on R. Similarly for R/I_N.

If the two traces are the same we deduce then that the kernel is also the same and so $I_M = I_N$. Now in order to prove that the representations are the same we have to check that the isotypic components have the same multiplicities. This is clear from the formula $\sum_{i=1}^k p_i Tr(a_i)$. □

§15 Algebraic groups.

15.1

Definition. *An algebraic group G is an algebraic variety with a group action, such that the two maps of multiplication $G \times G \to G$ and inverse $G \to G$ are algebraic.*

If G is an affine variety it is called an affine algebraic group.

For algebraic groups it is important to study algebraic actions, i.e. actions $\pi : G \times V \to V$ where V is an algebraic variety and the map π is algebraic.

Among algebraic actions there are the left and right action of G on itself:

$$(g, h) \to gh, \ (g, h) \to hg^{-1}.$$

Remark. The map $h \to h^{-1}$ is an isomorphism between the left and right action.

For algebraic groups we will usually restrict to algebraic homomorphisms (i.e. regular algebraic). In particular a linear representation $\rho : G \to GL(n)$ is called a **rational** representation if the homomorphism is algebraic.

It is useful to extend the notion to infinite dimensional representations.

Definition. *A linear action of an algebraic group G on a vector space V is called a* **rational** *representation if V is the union of finite dimensional subrepresentations which are algebraic.*

The main example is the general linear group $GL(n)$ of invertible matrices, it is an affine group. A subgroup H of $GL(n)$ is called a linear group, in particular a Zariski closed subgroup H of $GL(n)$ is a linear algebraic group.

Thus linear algebraic groups are affine. There are precise reasons to restrict our study to affine groups (cf.).

Consider an algebraic action $\pi : G \times V \to V$ on an affine algebraic variety V. The action induces an action on functions. The regular functions $F[V]$ on V are then a representation.

Proposition. *$F[V]$ is a rational representation.*

Proof. Let $F[G]$ be the coordinate ring of G so that $F[G] \otimes F[V]$ is the coordinate ring of $G \times V$. The action π induces a map $\pi^* : F[G] \to F[G] \otimes F[V]$, where $\pi^* f(g, v) := f(gv)$.

For a given function $f(v)$ on V we have that $f(gv) = \sum_i a_i(g) b_i(v)$ thus we see that the translated functions f^g lie in the linear span of the functions b_i.

This shows that any finite dimensional subspace U of the space $F[V]$ is contained in a finite dimensional G stable subspace W.

Given a basis u_i of W, we have:

$$u_i(g^{-1}v) = \sum_j a_{ij}(g)u_i(v),$$

with the a_{ij} regular functions on G, thus W is a rational representation and the union of these representations is clearly $F[V]$.

The previous theorem has an important corollary.

Theorem. *i) Given an action of an affine group G on an affine variety V there exists a linear representation W of G and a G equivariant embedding of V in W.*

ii) An affine group is isomorphic to a linear group.

Proof. i) Choose a finite set of generators of the algebra $F[V]$ and then a finite dimensional $G-$stable subspace $W \subset F[V]$ containing this set of generators.

W defines an embedding i of V into W^* by $< i(v)|w >= w(v)$. This embedding is clearly equivariant if on W^* we put the dual of the action on W.

ii) Consider the right action of G on itself. If W is as before and $u_i,\ i = 1, \dots, n$ is a basis of W we have: $u_i(xy) = \sum_j a_{ij}(y)u_j(x)$.

Consider the homomorphism ρ of G to matrices given by the matrix $(a_{ij}(y))$. Since $u_i(y) = \sum_j a_{ij}(y)u_j(1)$ we have that the functions a_{ij} generate the coordinate ring of G and thus ρ is an embedding of G into matrices. Thus the image of ρ is a linear group and ρ is an isomorphism from G to its image.

15.2 We start with:

Lemma. *Any finite dimensional rational representation U of an algebraic group G can be embedded in an equivariant way in the direct sum of finitely many copies of the coordinate ring $F[G]$ under the right (or the left) action.*

Proof. Take a basis $u^i, i = 1, \dots, n$ of the dual of U. The map:

$$u \to (< u^1|gu >, \dots, < u^n|gu >)$$

is clearly equivariant, with respect to the right action on $F[G]$. Computing these functions in $g = 1$ we see that this map is an imbedding.

If we start from an irreducible representation U, then U can be embedded in a single factor i.e. in $F[G]$. We are thus in the situation studied in 14.5.

We come to the main class of algebraic groups of our interst.

Proposition. *For an affine group G the following are equivalent:*

i) Every finite dimensional rational representation is semisimple.

ii) Every rational representation is semisimple.

iii) The coordinate ring $F[G]$ is semisimple under the right (or the left) action.

iv) If G is a closed subgroup of $GL(V)$ then all the tensor powers V^n are semisimple.

A group satisfying the previous properties is called a **Linearly reductive group***.*

Proof. i) implies ii) by the abstract representation theory. Clearly ii) implies iii) and iv), iii) implies i) by the previous lemma and the fact that direct sums and submodules of semisimple modules are semisimple.

Finally assume iv) we want to deduce iii).

Let A, B be the coordinate rings of the space $End(V)$ of all matrices and of the group $GL(V)$, the coordinate ring $F[G]$ is a quotient of B as ring and as representation thus it suffices to show that B is semisimple.

We consider the action of G by right multiplication on these spaces.

The algebra $B = \cup_{i=0}^{\infty} d^{-i}A$ where d is the determinant function. Clearly d generates a 1-dimensional representation of $GL(V)$ and of G under right action, let us indicate it by L, and $d^{-i}A = L^{-\otimes i}A$.

Thus if A is a semisimple representation under G so is B. It remains to show that A is semisimple under the assumptions of iv).

The space of endomorphisms $End(V)$, as a $G \times G$ module is isomorphic to $V \otimes V^*$ so the ring A is isomorphic to $S[V^* \otimes V] = S[V^{\oplus m}]$ as right G module, if $m = dim\, V$.

As representation $S[V^{\oplus m}]$ is a quotient of the tensor algebra $\oplus_n (V^{\oplus m})^n$ which in turn is isomorphic to a direct sum of tensor powers of V, the claim follows.

Let then G be a linearly reductive group. For every irreducible representation U we have that $U^* \otimes U$ appears in $F[G]$ as $G \times G$ submodule and also as the isotypic component of type U. It follows that

Theorem. *If G is a linearly reductive group we have only countably many non isomorphic irreducible representations and*

(15.2.1) $$F[G] = \oplus_i U_i^* \otimes U_i \quad (as \quad G \times G \quad modules)$$

where U_i runs over the set of all non isomorphic irreducible representations of G.

Remark. Observe that this formula is the exact analogue of the decomposition formula for the group algebra of a finite group.

Given a linearly reductive group, as in the case of finite groups an explicit description of the decomposition 15.2.1 implies a knowledge of its representation theory.

We need some condition to recognize that an algebraic group is linearly reductive. There is a very simple sufficient condition which is easy to apply. This has been proven in 14.1 we recall the statement.

Theorem. *Given a subgroup $G \subset GL(V) = GL(n,\mathbb{C})$, let $G^* := \{g^* = \overline{g}^t | g \in G\}$. If $G = G^*$ then all tensor powers $V^{\otimes m}$ are completely reducible under G. In particular if G is an algebraic subgroup then it is linearly reductive.*

As a consequence one easily verifies:

Corollary. *The groups $GL(n,\mathbb{C})$, $SL(n,\mathbb{C})$, $O(n,\mathbb{C})$, $SO(n,\mathbb{C})$, $Sp(n,\mathbb{C})$, D are linearly reductive (D denotes the group of invertible diagonal matrices).*

We should finally remark that, from the Theory developed

Proposition. *if $G \subset GL(V)$ is linearly reductive all its irreducible representations, up to tensor product with powers of the deeterminat can be found as subrepresentations of $V^{\otimes n}$ for some n.*

We will apply this idea to classify irreducible representations of classical groups.

15.3 The simplest example of linearly reductive group is the torus T isomorphic to the product of n copies of the multiplicative group which can be viewed as the group D of invertible diagonal matrices. Its coordinate ring is the ring of Laurent polynomials $F[T] = F[x_i, x_i^{-1}]$ in n variables. A basis of $F[T]$ is given by the monomials:

$$x^{\underline{m}} = x_1^{m_1} x_2^{m_2} \dots x_n^{m_n} \tag{15.3.1}$$

as $\underline{m} = (m_1, m_2, \dots, m_n)$ varies in the free abelian group $\mathbb{Z}^n$.

The 1-dimensional subspace $x^{\underline{m}}$ is a subrepresentation, and under right action if $t = (t_1, t_2, \dots, t_n)$ we have

$$(x^{\underline{m}})^t = (x_1 t_1)^{m_1} (x_2 t_2)^{m_2} \dots (x_n t_n)^{m_n} = t^{\underline{m}} x^{\underline{m}} \tag{15.3.2}$$

Thus the irreducible characters of T are the elements of a free abelian group, called the **character group**.

Proposition. *Every representation of T has a basis in which the action is diagonal.*

A vector generating a T stable subspace is called a **weight vector** *and the corresponding character or eigenvalue the* **weight***.*

Proof. This is the consequence of the fact that every rational representation is semisimple and that the irreducible representations are 1 dimensinal. □

The weights of the representation can of course appear with any multiplicity, the corresponding character can be identified to a Laurent polynomial $\sum_{\underline{m}} c_{\underline{m}} x^{\underline{m}}$ with the $c_{\underline{m}}$ positive integers.

One should remark that weihts are a generalization of degrees of homogeneity, let us illustrate it in the simple case of a vector space $V = U_1 \oplus U_2$.

To such a decomposition of a space corresponds a (2-dimensional) torus T formed by the linear transformations $(u_1, u_2) \to (xu_1, yu_2)$. The decompositions of the various spaces one constructs from V associated to the givendirect sum decomposition are just weigth space decompositions. For instance

$$S^n(V) = \oplus_{i=0}^n S^i(U_1) \otimes S^{n-i}(U_2), \ \wedge^n(V) = \oplus_{i=0}^n \wedge^i (U_1) \otimes \wedge^{n-i}(U_2)$$

both $S^i(U_1) \otimes S^{n-i}(U_2), \ \wedge^i(U_1) \otimes \wedge^{n-i}(U_2)$ are weigth spaces of weigth $x^i y^{n-i}$.

15.4 The importance of tori in representation theory comes from a sequence of structure theorems which we want to mention, referring to texts in algebraic groups for their proofs.

Definition. *A linear group is called* **unipotent** *if and only if its elements are all unipotent (i.e. all the eigenvalues are 1).*

The main structure theorem is:

Theorem. *In characteristic 0 a linear algebraic group is linearly reductive if and only if it does not contain any closed unipotent normal subgroup.*

A special role in the theory of linear algebraic groups play the notions of *maximal torus* and *Borel subgroup.*

Definition. *A subgroup of an algebraic group is called a maximal torus if it is a closed subgroup, a torus as algebraic group and maximal with this property.*

A subgroup of an algebraic group is called a Borel subgroup if it is closed connected and solvable and maximal with this property.

The main structure theorem is:

Theorem. *All maximal tori are conjugate. All Borel subgroups are conjugate.*

We illustrate this theorem for classical groups giving an elementary proof of the first part.

Example 1 $GL(V)$ In the general linear group of a vector space V a maximal torus is given by the subgroup of diagonal matrices a Borel subgroup by the subgroup of upper triangular matrices.

The set of maximal tori is in 1-1 correspondence with the set of all decompositions of V as direct sum of 1-dimensional subspaces while the set of Borel subgroups is in one to one correspondence with the set of *maximal flags* i.e. the set of sequences $V_1 \subset V_2 \subset \dots V_{n-1} \subset V_n = V$ of subspaces of V with $dimV_i = i$ (assuming $n = dimV$).

Example 2 $SO(V)$ In the special orthogonal group of a vector space V, equipped with a non degenerate symmetric form a maximal torus is given as follows.

If $dimV = 2n$ is even we say that a basis $e_1, f_1, e_2, f_2, \dots, e_n, f_n$ is hyperbolic if the 2 dimensional subspaces V_i spanned by e_i, f_i are mutually orthogonal and the matrix of the form on the vectors e_i, f_i is:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

For such a basis we get a maximal torus of matrices which stabilyze each V_i and, restricted to V_i in the basis e_i, f_i has matrix:

$$\begin{pmatrix} \alpha_i & 0 \\ 0 & \alpha_i^{-1} \end{pmatrix}.$$

The set of maximal tori is in 1-1 correspondence with the set of all decompositions of V as direct sum of 1-dimensional subspaces spanned by hyperbolic bases. The set of Borel subgroups is in one to one correspondence with the set of *maximal isotropic flags* i.e. the set of sequences $V_1 \subset V_2 \subset \dots V_{n-1} \subset V_n$ of subspaces of V with $dimV_i = i$ and such that the subspace V_n is isotropic for the form. To such a flag one associates the maximal flag $V_1 \subset V_2 \subset \dots V_{n-1} \subset V_n = V_n^{\perp} \subset V_{n-1}^{\perp} \cdots \subset V_2^{\perp} \subset V_1^{\perp} \subset V$ which is clearly stable under the subgroup fixing the given isotropic flag.

If $dimV = 2n+1$ is odd we take bases of the form $e_1, f_1, e_2, f_2, \dots, e_n, f_n, u$ with $e_1, f_1, e_2, f_2, \dots, e_n, f_n$ hyperbolic and u orthogonal to $e_1, f_1, e_2, f_2, \dots, e_n, f_n$. As maximal torus we can take the same type of subgroup which now fixes u. The analogue statement holds for Borel subgroups except that now $V_1 \subset V_2 \subset \dots V_{n-1} \subset V_n \subset V_n^{\perp} \subset V_{n-1}^{\perp} \cdots \subset V_2^{\perp} \subset V_1^{\perp} \subset V$ is a maximal flag.

Example 3 $Sp(V)$ In symplectic group of a vector space V equipped with a non degenerate skew-symmetric form a maximal torus is given as follows. Let $dimV = 2n$ we say that a basis $e_1, f_1, e_2, f_2, \dots, e_n, f_n$ is symplectic if the 2 dimensional subspaces V_i spanned by e_i, f_i are mutually orthogonal and the matrix of the form on the vectors e_i, f_i is:

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

For such a basis we get a maximal torus of matrices which stabilyze each V_i and, restricted to V_i in the basis e_i, f_i has matrix:

$$\begin{pmatrix} \alpha_i & 0 \\ 0 & \alpha_i^{-1} \end{pmatrix}$$

The set of maximal tori is in 1-1 correspondence with the set of all decompositions of V as direct sum of 1-dimensional subspaces spanned by symplectic bases. The set of Borel subgroups is again in one to one correspondence with the set of *maximal isotropic flags* i.e. the set of sequences $V_1 \subset V_2 \subset \dots V_{n-1} \subset V_n$ of subspaces of V with $dimV_i = i$ and such that the subspace V_n is isotropic for the form.

To such a flag one associates the maximal flag $V_1 \subset V_2 \subset \dots V_{n-1} \subset V_n = V_n^{\perp} \subset V_{n-1}^{\perp} \cdots \subset V_2^{\perp} \subset V_1^{\perp} \subset V$ which is clearly stable under the subgroup fixing the given isotropic flag.

proof. We use the fact that a torus action on a vector space decomposes in a direct sum of irreducible representations which implies immediately that, any torus in the general linear group has a basis in which it is diagonal and hence the maximal tori are the ones described.

For the other two cases we exploit the fact that, given two eigenspaces relative to 2 characters χ_1, χ_2 these subspaces are orthogonal under the given invariant form unless $\chi_1\chi_2 = 1$.

For instance assume we are in the symmetric case (the other is identical), given two eigenvectors u_1, u_2 and an element of the maximal torus $(u_1, u_2) = (tu_1, tu_2) = (\chi_1\chi_2)(t)(u_1, u_2)$.

It follows that, if $\chi_1\chi_2 =\neq 1$, the two eigenvectors are orthononal.

By the non degenerate nature of the form when $\chi_1\chi_2 = 1$ the two eigenspaces relative to the two characters must be in perfect duality since they are orthogonal to the remaining weight spaces.

It easily follow the existence of a hyperbolic basis for which the given torus is contained in the torus associated before.

All hyperbolic bases are conjugate under the orthogonal group, under the special orthogonal group two hyperbolic bases are conjugate provided eventually we exchange in a pair e_i, f_i the two vectors (leaving the corresponding torus unchanged), so the statement for maximal tori is complete in all cases.

The discussion of Borel subgroups is a little subtler, here one has to use the basic fact:

Theorem Lie-Kolchin. *A connected solvable group G of matrices is conjugate to a subgroup of upper triangular matrices.*

There are various proofs of this statement which can be found in the literature at various levels of generality. The main step is to prove the existence of a common eigenvector for G from which the statement follows immediately by induction.

A particularly slick proof follows immediately from a stronger theorem.

Theorem Borel. *Given an action of a connected solvable group G on a projective variety there exists a fixed point.*

The projective variety to which this theorem has to be applied to obtain the Lie Kolchin Theorem is the flag variety, which is easily seen to be projective (cf. Chap. 4).

Given the theorem of Lie Kolchin the study of Borel subgroups is immediate. For the linear group it is clearly a restatement of this theorem, for the other groups let G be a connected solvable group of linear transformations fixing the form.

We do the symmetric case since the other is similar but simpler. Let u be an eigenvecor of G and $u^\perp$ its orthogonal subspace which is necessarily G stable.

If u the space $u^\perp/\mathbb{C}u$ is equipped with the induced symmetric form for which G acts again a a group of orthogonal transformations and we can apply induction.

In the case u not isotropic we have a direct sum orthogonal decomposition $V = u^\perp \oplus \mathbb{C}u$ and again can apply induction finding a maximal isotropic flag stabilized by G in $u^\perp$.

In particular, unless $dimu^\perp = 1$ (in which case we see that G is reduced to 1) we can find an isotropic eigenvector, by induction, and we are back to the previous case.

For a connected reductive group we have furthermore the following important facts.

Theorem. *Let G be a connected reductive group then.*

(1) *The set of semisimple elements is dense in G.*
(2) *Each semisimple element is contained in a maximal torus.*

We leave as exercise to verify these statements directly for classical groups.

§16 Characters.

16.1 We want to deduce some of the basic theory of characters of finite groups with some comments to compact groups and reductive groups also.

Definition. *Given a linear representation $\rho : G \to GL(V)$ of a group G, where V is a vector space over a field F we define its* **character** *to be the following function on G.*

$$\chi_\rho(g) := tr(\rho(g)).$$

Here tr is the usual trace.

We say that a character is irreducible if it comes from an irreducible representation.

Some properties are immediate.

Proposition. *1) $\chi_\rho(g) = \chi_\rho(aga^{-1}),\ \forall a, g \in G$.*

This means that the character is constant on conjugacy classes, such a function is called a **class function***.*

2) Given two representations ρ_1, ρ_2 we have:

$$\chi_{\rho_1 \oplus \rho_2} = \chi_{\rho_1} + \chi_{\rho_2},\ \chi_{\rho_1 \otimes \rho_2} = \chi_{\rho_1}\chi_{\rho_2},$$

3) Furthermore if ρ is unitarizable we have that the character of the dual representation ρ^* is the conjugate of χ_ρ:

$$\rho^* = \overline{\rho}.$$

Proof. Let us show 3) since the others are clear. If ρ is unitarizable there is a basis in which the matrices $\rho(g)$ are unitary, hence in the dual representation we obtain the conjugate matrix which is equal to the inverse transposed and $tr(\rho^*(g)) = tr(\overline{\rho(g)}) = \overline{tr(\rho(g))}$.

16.2 To understand the deeper properties of characters we state the following for compact groups.

Proposition. *Let $\rho : G \to GL(V)$ be a complex finite dimensional representation of a compact group G (in particular a finite group), then:*

$$dim_{\mathbb{C}} V^G = \int \chi_{\rho(g)} dg.$$

Proof. Let us consider the operator $\pi := \int \rho(g)dg$, we claim that it is the projection operator on V^G. In fact if $v \in V^G$:

$$\pi(v) = \int \rho(g)(v)dg = \int vdg = v.$$

Otherwise:

$$\rho(h)\pi(v) = \int \rho(h)\rho(g)vdg = \int \rho(hg)vdg = \pi(v)$$

by left invariance of the Haar integral.

Then $dim_{\mathbb{C}} V^G = tr(\pi) = tr(\int \rho(g)dg) = \int tr(\rho(g))dg = \int \chi_{\rho(g)} dg$ by linearity of the Trace and of the integral.

The previous proposition has an important consequence.

Theorem (Orthogonality of characters). *Let χ_1, χ_2 be the characters of two irreducible representations ρ_1, ρ_2 then:*

$$\int \chi_1(g)\overline{\chi}_2(g)dg = \begin{cases} 0 & \text{if } \rho_1 \neq \rho_2 \\ 1 & \text{if } \rho_1 = \rho_2 \end{cases}.$$

Proof. Let V_1, V_2 the the spaces of the two representations consider $hom(V_2, V_1) = V_1 \otimes V_2^*$.

As representation it has character $\chi_1(g)\overline{\chi}_2(g)$ from 16.1.

We have seen that $hom_G(V_2, V_1) = (V_1 \otimes V_2^*)^G$ hence $dim_{\mathbb{C}} hom_G(V_2, V_1) = \int \chi_1(g)\overline{\chi}_2(g)dg$ from the previous proposition.

Finally by Schur's lemma and the fact that V_1, V_2 are irreducible, it follows the theorem.

In fact a more precise theorem holds. Let us consider the Hilbert space of L^2 functions on G, inside we consider the subspace $L^2_c(G)$ of class functions then.

Theorem. *The irreducible characters are an orthonormal basis of $L^2_c(G)$.*

Let us give the proof for finite groups.

For a finite group G decompose the group algebra in matrix blocks according to 14.8 as $\mathbb{C}[G] = \oplus_i^m M_{h_i}(\mathbb{C})$.

The m blocks correspond to the m irreducible representations and the m irreducible characters are the composition of the projection to a factor $M_{h_i}(\mathbb{C})$ followed by the ordinary trace.

A function $f = \sum_{g \in G} f(g)g \in \mathbb{C}[G]$ is a class function if and only if $f(ga) = f(ag)$ for all $a, g \in G$. This means that f lies in the center of the group algebra.

The space of class functions is identified to the center of $\mathbb{C}[G]$.

The center of a matrix algebra $M_h(\mathbb{C})$ is formed by the scalar matrices thus the center of $\oplus_i^m M_{h_i}(\mathbb{C})$ equals $\mathbb{C}^{\oplus m}$.

It follows that the number of irreducible characters equals the dimension of the space of class functions. Since the irreducible characters are orthonormal they are a basis.

As corollary we have:

Corollary. *The number of irreducible representations of a finite group* G *equals the number of conjugacy classes in* G.

§17 The Peter-Weyl Theorem.

17.1 The theory for compact groups requires some basic functional analysis, let us give it assuming basic facts on compact operators.

If G is a locally compact group it has a left invariant Haar measure, this allows to define the *convolution* product which is the generalization of the product of elements of the group algebra.

The convolution product is defined first of all on the space of L^1 functions by the formula

$$(f * g)(x) := \int_G f(y)g(y^{-1}x)dy = \int_G f(xy)g(y^{-1})dy \tag{17.1.1}$$

Also we shall need when G is compact the continuous inclusion maps

$$C^0(G) \subset L^2(G) \subset L^1(G)$$

the 3 spaces have respectively the uniform (L^∞), L^1, L^2 norms the inclusions decrease norms since the L^1 norm of f equals the Hilbert scalar product of $|f|$ with 1 so by Schwarz inequality $|f|_1 \leq |f|_2$ while $|f|_2 \leq |f|_\infty$ by obvious reasons.

Lemma. *If* G *is compact then the space of* L^2 *functions is also an algebra under convolution.*

Let us define by $\|f\|$ the L^2 norm of a function In fact we have a special type of normed algebra, a C^*-algebra), since we have the norm inequality

$$\|f * g\| \leq \|f\|\|g\|$$

inserire

The convolution is related to module theory since, if we consider a linear representation ρ of G as operators on a space V we want to define an action of $L^1(G)$ on V by $fu := \int_G f(g)\rho(g)v$.

We want to see that under fairly general hypotheses this is well defined and gives a module structure under convolution.

Lemma. *Given a unitary representation of a group G any two irreducible finite dimensional non isomorphic submodules are orthogonal.*

Proof. Let V_1, V_2 be the two submodules, the Hermitian form induces a G linear map $j : V_1 \to (\overline{V}_2)^*$ by the formula $j(v)(w) := (v, w)$, since the representation is unitary the conjugate of such a representation is isomorphic to the dual hence $V_2 = (\overline{V}_2)^*$ and by irreducibility j must be 0. $\square$

Lemma-Definition. For a continuous function $f \in C^0(G)$ the following are equivalent:

(1) The space spanned by the left translates $f(gx),\ g \in G$ is finite dimensional.
(2) The space spanned by the rigth translates $f(xg),\ g \in G$ is finite dimensional.
(3) The space spanned by the bitranslates $f(gxh),\ g, h \in G$ is finite dimensional.
(4) There is a finite expansion $f(xy) := \sum_{i=1}^k u_i(x)v_i(y)$.

A function satisfying the previuos conditions is called a representative function.

(5) Moreover in the expansion 4) the functions u_i, v_i can be taken as representative functions.

Proof. Assume 1) and let $u_i(x),\ i = 1, \dots, m$ be a basis of the space spanned by the functions $f(gx),\ g \in G$.

Writing $f(gx) = \sum_i v_i(g)u_i(x)$ one easily sees that the coefficients $v_i(g)$ are continuous functions and thus 4) follows,. 4) is a simmetric property and clearly implies 1),2).

In the expansion $f(xy) := \sum_{i=1}^k u_i(x)v_i(y)$ we can take the functions v_i to be a basis of the space spanned by the left translates of f hence they are representative functions, one easily sees that the u_i are in the span of the rigth translates and so they are also representative functions. Now we can apply to them the previous part and see that we must have $u_i(xg) = \sum_j c_{ji}(g)w_j(x)$ for some functions w_j.

As for 3) we have again

$$f(gxh) = \sum_i v_i(g)u_i(xh) = \sum_i v_i(g) \sum_j c_{ji}(h)w_j(x)$$

$\square$

Proposition. *The set $\mathcal{T}_G$ of representative functions is an algebra spanned by the matrix coefficients of the finite dimensional continuous representations of G.*

Proof. We leave the simple proof to the reader. $\square$

If G, K are 2 topological groups we have that

Proposition. *Under multiplication $f(x)g(y)$ we have an isomorphism*

$$\mathcal{T}_G \otimes \mathcal{T}_K = \mathcal{T}_{G \times K}$$

Proof. The multiplication map of functions on two distinct spaces to the product space is always an isomorphism of the tensor product of the space of functions to the image, so we only have to prove that the space of representative functions of $G \times K$ is spanned by the functions $\psi(x,y) := f(x)g(y).f(x) \in \mathcal{T}_G,\ g(y) \in \mathcal{T}_K$.

Using the property 4) of the definition of representative function we have that if $f(x_1x_2) = \sum_i u_i(x_1)v_i(x_2),\ g(y_1,y_2) = \sum_k w_k(y_1)z_k(y_2)$ then

$$\psi((x_1,y_1)(x_2,y_2)) = \sum_{i,k} u_i(x_1)w_k(y_1)v_i(x_2)z_k(y_2)$$

conversely if $\psi(x,y)$ is representative writing $(x,y) = (x,1)(1,y)$ one immediately sees that ψ is in the span of the product of representative functions. $\square$

Before discussing the basic Theorem of Peter Weyl let us recall that a bounded operator T on a Hilbert space $\mathbb{H}$ is compact if it maps bounded sets ito relatively compact sets. If T is Hermitian it has a special spectral decomposition (Fredholm's Theory) the non zero eigenvalues form a discrete set $\hat{T}$ converging to 0 and each non zero eigenvalue λ has a finite dimensional eigenspace H_λ.

The Hilbert space $\mathbb{H}$ has a Hilbert space decomposition $\mathbb{H} = \mathbb{H}_0 \oplus_{\lambda \in \hat{T}} H_\lambda$ i.e. these subspaces are orthogonal and every vector in $\mathbb{H}$ is a convergent serieas in elements out of the given subspaces.

We have seen that for every finite dimensional irreducible representation V of G the space of matrix coefficients $V^* \otimes V$ appears in the space $C^0(G)$ of continuous functions on G and that, for distinct irreducible representations V_1, V_2 the corresponding spaces of matrix coefficients are orthogonal in the L^2 norm.

Since the Hilbert space $L^2(G)$ is separable it follows again that we can have only countably many spaces $V_i^* \otimes V_i$ with V_i irreducible and we claim

Theorem Peter-Weyl. *The direct sum $\oplus_i V_i^* \otimes V_i$ is dense in $L^2(G)$.*

In other words every L^2 function f on G can be developed uniquely as $f = \sum_i u_i$ with $u_i \in V_i^* \otimes V_i$.

Proof. The first point is to identify the direct sum $\mathcal{T} := \oplus_i V_i^* \otimes V_i$ with the space of *representative functions* and this follows from the Theorem 14.15 and the previous Lemma.

Next we must show that the representative functions are dense.

For this we take a continuous function $\phi(x)$ with $\phi(x) = \phi(x^{-1})$ and claim that the convolution map $R_\phi : f \to f * \phi := \int_G f(y)\phi(y^{-1}x)dy$ is a compact operator and that its image is in the closure of $\mathcal{T}$.

The fact that the operator is compact follows from a standard fact that it maps the set of functions of norm 1 into a bounded and equicontinuous set which is relatively compact in $C^0(G)$ and hence also in $L^2(G)$, the boundedness is clear by the norm inequality.

By construction convolution is G equivariant for the left action hence it follows that the eigenspaces of this operator are G stable.

For a compact operator the eigenspaces relative to non 0 eigenvalues are finite dimensional hence we obtain that the image of R_ϕ (spanned by the eigenvectors of non 0 eigenvalue) is in the closure of $\mathcal{T}$.

The next step is to show that given a continuous function f, as ϕ varies one can approximate f with elements in the image of R_ϕ as close as possible.

Given $\epsilon > 0$ take an open set U neighborhood of 1 such that $|f(x)-f(y)| < \epsilon$ if $xy^{-1} \in U$ and take ϕ with support in U and positive with integral 1 then $|f - f * \phi| < \epsilon$ since

$$|f(x)-(f*\phi)(x)| = |\int_G f(x)\phi(y^{-1}x)dy - \int_G f(y)\phi(y^{-1}x)dy| = |\int_{y^{-1}x\in U}(f(x)-f(y))\phi(y^{-1}x)dy| \leq \int_{y}$$

□

17.2 We draw some consequences of the Peter Weyl Theorem.

Corollary. *For any continuous representation of G in a Hilbert space H.*

A vector v such that the elements gv, $g \in G$ span a finite dimensional vector space is called a finite vector.

The set of finite vectors is dense.

Proof. By module theory, if $u \in H$ the set $\mathcal{T}u$ spanned by applying the representative functions is made of finite vectors, but by continuity $u = 1u$ is a limit of these vectors. □

We can now apply the theory developed to class functions, a class function is a function which is invariant under the action of G embedded diagonally in $G \times G$, i.e. $f(x) = f(g^{-1}xg)$ for all $g \in G$.

Develop $f = \sum_i f_i$ with $f_i \in V_i^* \otimes V_i$. By the invariance property and the uniqueness of the development it follows that each f_i is invariant, i.e. a class function.

We know that in $V_i^* \otimes V_i$ the only invariant functions under the diagonal action are the multiples of the corresponding character hence we see that

Corollary. *The irreducible characters are an orthonormal basis of the Hilbert space of L^2 class functions.*

§18 Linearly reductive groups.

18.1 We discuss now linearly reductive groups.

Recall the theorem on maximal tori for reductive groups. One of the implications of this theorem is the fact that the character of a representation M of G is determined by its restriction to a given maximal torus T. On M the group T acts as a direct sum of irreducible 1-dimensional characters in $\hat{T}$ and thus the character of M can be expressed as a sum of these characters with non negative coefficients, expressing their multiplicities.

Consider next the action by conjugation of G on itself. It is the restriction to G, embedded diagonally in $G \times G$ of the left and right action.

Let $Z[G]$ denote the space of regular functions f which are invariant under conjugation.

From the decomposition 15.2.1, $F[G] = \oplus_i U_i^* \otimes U_i$ it follow that the space $Z[G]$ decomposes as direct sum of the spaces $Z[U_i]$ of conjugation invariant functions in $U_i^* \otimes U_i$. We claim that:

Lemma. *$Z[U_i]$ is 1-dimensional, generated by the character of the representation U_i.*

Proof. Since U_i is irreducible and $U_i^* \otimes U_i = End(U_i)^*$ we have by Schur's lemma that $Z[U_i]$ is 1-dimensional, generated by the element corresponding to the trace on $End(U_i)$.

Now we follow the identifications, an element u of $End(U_i)^*$ gives the matrix coefficient $u(\rho_i(g))$ where $\rho_i : G \to GL(U_i) \subset End(U_i)$ denotes the representation map.

We obtain the function $\chi_i(g) = tr(\rho_i(g)$ as the desired invariant elment.

Corollary. *For a linearly reductive group the irreducible characters are a basis of the conjugation invariant functions.*

We have seen that the characters are determined by their restriction to a maximal torus T, the fact that a character is a class function implies a further symmetry. Let N_T denote the normalizer of T, it acts on T by conjugation and a class function restricted to T is invariant under this action. There are many important theorems about this action, first.

Theorem. *T equals its centralizer and N_T/T is a finite group, called the Weyl group and denoted by W.*

Let us illustrate this theorem for classical groups.

We always exploit the same idea.

Let T be a torus contained in the linear group of a vector space V.

Decompose $V := \oplus_\chi V_\chi$ in weight spaces under T and let $g \in GL(V)$ be a linear transformation normalizing T.

g induces by conjugation an automorphism of T wich we still denote by g which permutes the characters of T by the formula $\chi^g(t) := \chi(g^{-1}tg)$.

We thus have, for $v \in V_\chi,\ t \in T$, $tgv = gg^{-1}tgv = \chi^g(t)gv$.

We deduce that $gV_\chi = V_{\chi^g}$. In particular g permutes the weight spaces.

In all the previous cases we have thus a homomorphism from the normalizer of the maximal torus to the group of permutations of the weight spaces. Let us first analize the kernel of this homomorphism, which we denote by N_T^0, in the 4 cases.

1. Let D be the group of all diagonal matrices. It is exactly the full subgroup of linear transformations fixing the 1-dimensional weight spaces generated by the given basis vectors.

An element in N_D^0 by definition fixes all these subspaces and thus in this case $N_D^0 = D$.

2. Even orthogonal group. Again the space decomposes into 1-dimensional eigenspaces spanned by the vectors e_i, f_i and one immediately verifies that a diagonal matrix g given by $ge_i = \alpha_i e_i,\ gf_i = \beta_i f_i$ is orthogonal if and only if $\alpha_i\beta_i = 1$. Again $N_T^0 = T$.

3. Odd orthogonal group. Similar to the previous case except that now we have an extra non isotropic basis vector u and g is orthogonal if furthermore $gu = \pm u$. It is special orthogonal only if $gu = u$. Again $N_T^0 = T$.

4. Symplectic group. Identical to 2.

Now for the full normalizer. In case 1, in the general linear group there is contained the symmetric group S_n acting as permutations on the given basis.

It follows that $N_D = D \ltimes S_n$. In the special linear group we still have an exact sequence $0 \to D \to N_D \to S_n \to 0$ but this does not split and of course only the even permutations are in the special linear group.

2. In the even orthogonal case $dimV = 2n$ the characters come in opposite pairs and clearly the normalizer permutes this set of n pairs.

In the same way as before we see now that the symmetric group S_n permuting simulataneosly with the same permutation the elements $e_1, e_2, \ldots, e_n;\ f_1, f_2, \ldots, f_n$ is formed of special orthogonal matrices.

The kernel of the map $N_T \to S_n$ is formed by matrices diagonal of 2×2 blocks.

Each two by two block, is in the orthogonal group of 2-dimensional space and it is clearly the semidirect product of the torus part $\begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}$ with the permutation matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

In the special orthogonal group only an even number of permutation matrices $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ can appear. It follows that the Weyl group is the semidirect product of the symmetric group S_n with the subgroup of index 2 of $\mathbb{Z}/(2)^n$ formed by the $n-$tuples $a_1, \ldots, a_n$ with $\sum_{i=1}^n a_i = 0$.

3. The odd special orthogonal group is slightly different, we use the notations of 15.3. Now one has also the possibility to act on the basis $e_1, f_1, e_2, f_2, \ldots, e_n, f_n, u$ by -1 on u and this corrects the fact that the determinant of an element defined on $e_1, f_1, e_2, f_2, \ldots, e_n, f_n$ may be 1.

We deduce then that the Weyl group is the semidirect product of the symmetric group S_n with $\mathbb{Z}/(2)^n$.

4. The symplectic group. The discussion starts as in the even orthogonal group except now that, in order to get a symplectic matrix we have to choose the matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ on a 2-dimensional space.

This matrix has determinant 1 and again we deduce that the Weyl group is the semidirect product of the symmetric group S_n with $\mathbb{Z}/(2)^n$.

We have to discus now the action of the Weyl group on the characters of a maximal torus.

In the case of the General Linear group a diagonal matrix X with entries $x_1, \ldots, x_n$ is conjugated by a permutation matrix σ which maps $\sigma e_i = e_{\sigma(i)}$ by $\sigma X \sigma^{-1} e_i = x_{\sigma(i)} e_i$ thus the action of S_n on the *characters* x_i is the usual permutation of variables.

For the orthogonal groups and the symplectic group one has the torus of diagonal matrices of the form $x_1, x_1^{-1}, x_2, x_2^{-1}, \ldots, x_n, x_n^{-1}$

Besides the permutations of the variables we have now also the inversions $x_i \to x_i^{-1}$, except that, for the even orthogonal group one has to restrict to products of only an even number of inversions.

The analysis we have made suggests to interpret the characters of the classical groups as particular symmetric functions. In the linear group case the character group of the maximal torus can be viewed as the polynomial ring $\mathbb{C}[x_1, \ldots, x_n][d^{-1}]$ with $d := \prod_{i=1}^n x_i$ inverted.

d is the n^{th} elementary symmetric function and thus the invariant elements are the polynomial in the elementary symmetric functions $\sigma_i(x),\ i = 1, \ldots, n-1$ and $\sigma_n(x)^{\pm 1}$.

In the case of the inversions we make a remark. Consider the ring $A[t, t^{-1}]$ of Laurent polynomials over a commutative ring A. An element $\sum_i a_i t^i$ is invariant under $t \to t^{-1}$ if and only if $a_i = a_{-i}$. We claim then that it is a polynomial in $u := t + t^{-1}$.

In fact $t^i + t^{-i} = (t + t^{-1})^i + r(t)$ where $r(t)$ has lower degree and one can work by induction.

We deduce that

Theorem. *The ring of invariants of* $\mathbb{C}[x_1, x_1^{-1}, \ldots, x_n, x_n^{-1}]$ *under* $S_n \ltimes \mathbb{Z}/(2)^n$ *is the polynomial ring in the elementary symmetric functions* $\sigma_i(u)$ *in the variables* $u_i := x_i + x_i^{-1}$.

Proof. We can compute the invariants in two steps. First the invariants under $\mathbb{Z}/(2)^n$ which by the previous argument are the polynomials in the u_i and then under the action of S_n which permutes the u_i. The claim follows. $\square$

For the even orthogonal group we need a different computation since now we only want the invariants under a subgroup. Let $H \subset \mathbb{Z}/(2)^n$ be the subgroup defined by $\sum_i a_i = 0$¿

Start from the monomial $x_1 x_2 \ldots x_n$ and consider all monomials $x_1^{\epsilon_1} x_2^{\epsilon_2} \ldots x_n^{\epsilon_n}$ where the elments $\epsilon_i = \pm 1$. We define next $E := \sum_{\prod_{i=1}^n \epsilon_i = 1} x_1^{\epsilon_1} x_2^{\epsilon_2} \ldots x_n^{\epsilon_n}$, $\overline{E} := \sum_{\prod_{i=1}^n \epsilon_i = -1} x_1^{\epsilon_1} x_2^{\epsilon_2} \ldots x_n^{\epsilon_n}$, E is clearly invariant under H and $E + \overline{E}$, $E\overline{E}$ are invariant under $\mathbb{Z}/(2)^n$.

We claim that any H invariant is generated of the form $a + bE$ where a, b are $\mathbb{Z}/(2)^n$ invariants.

Consider the set of all Laurent monomials which is permuted by $\mathbb{Z}/(2)^n$. A basis of invariants under $\mathbb{Z}/(2)^n$ is clearly given by the sums of the vectors in each orbit, similarly for the H invariants. Now let K be the stabilizer of an element of the orbit, which thus has $\frac{2^n}{|K|}$ elements. The stabilizer in H is $K \cap H$ hence a $\mathbb{Z}/(2)^n$ orbit is either an H orbit or it splits into 2 orbits, according whether $K \not\subset H$ or $K \subset H$.

We get H invariants which are not $\mathbb{Z}/(2)^n$ invariants from the last type of orbits.

A monomial $\prod x_i^{h_i}$ is stabilized by all the inversions in the variables x_i which have exponent 0 thus the only case in which the stabilizer is contained in H is when all the variables x_i appear. Let $S^i_{h_1,\dots,h_n}$, $i=1,2$ be the sum on the two orbits under H, for one of the two indeces i we may assume that $h_k>0$, say for S^1. The multiplication $S^1_{h_1-1,\dots,h_n-1}S^1_{1,1,\dots,1}$ gives rise to $S^1_{h_1,\dots,h_n}$ plus terms which are lower in the lexicographic ordering of the h_i's and $S^1_{1,1,\dots,1}=E$ thus by induction we assume that the lower terms are of the required form.

Also by induction $S^1_{h_1-1,\dots,h_n-1}=a+bE$ and so $S^1_{h_1,\dots,h_n}=(a+bE)E=(a+b(E+\overline{E}))E-b(E\overline{E})$ has the required form.

We can now discuss the invariants under the Weyl group. Again the ring of invariants under H is stabilized by S_n which acts by permuting the elements u_i and fixes the element E. We deduce that the ring of W invariant is formed by elements of the form $a+bE$ where a,b are polynomials in the elementary symmetric functions in the elements u_i.

It remains to understand the quadratic equation satisfied by E over the ring of symmetric functions in the u_i.

E satisfies the relation $E^2-(E+\overline{E})E+E\overline{E}=0$ and so we must compute the symmetric functions $E+\overline{E}$, $E\overline{E}$.

We easily see that $E+\overline{E}=\prod_{i=1}^n(x_i+x_i^{-1})$ which is the n^{th} elementary symmetric function in the u_i's. As for $E\overline{E}$

§19 Induction and restriction.

19.1 We collect now some general facts about representations of groups.

First of all let H be a group, $\phi: H\to H$ an automorphism and $\rho: G\to GL(V)$ a linear representation.

Composing with ϕ we get a new representation V^ϕ given by $G@>\phi>>G\xrightarrow{\rho}GL(V)$, it is immediately verified that, if ϕ is an inner automorphism V^ϕ is equivalent to ϕ.

Let now $H\subset G$ be a normal subgroup, every element $g\in G$ induces by inner conjugation in G an automorphism ϕ_g of H.

Given a representation M of G and an H submodule $N\subset M$ we clearly have that $gN\subset M$ is again an H submodule and canonically isomorphic to N^{ϕ_g}, it depends only on the coset gH of G.

In particular assume that, M is irreducible as G module and N is irreducible as H module, then all the submodules gN are irreducible H modules and $\sum_{g\in G/H}gN$ is a G submodule hence $\sum_{g\in G/H}gN=M$.

We want in particular apply this when H has index 2 in $G=H\cup uH$, we shall then use the canonical sign representation ϵ of $\mathbb{Z}/(2)=G/H$.

Theorem. *1) Given an irreducible representation N of H it extends to a representation of G if and only if N is isomorphic to N^{ϕ_u}, in this case it extends in two ways up to the sign representation.*

2) An irreducible representation M of G restricted to H either remains irreducible or splits into 2 irreducible representations $N \oplus N^{\phi_u}$ according to wether M is not or is isomorphic to $M \otimes \epsilon$.

Proof. Let $h_0 = u^2 \in H$. If N is also a G representation the map $u : N \to N$ is an isomorphism with N^{ϕ_u}, conversely let $t : N \to N = N^{\phi_u}$ be an isomorphism so that $tht^{-1} = \phi_u(h)$, then $t^2ht^{-2} = h_0hh_0^{-1}$.

Since N is irreducible we must have $h_0^{-1}t^2 = \lambda$ is a scalar.

We can substitute t with $t\sqrt{\lambda}^{-1}$ and can thus assume that $t^2 = h_0$ (on N).

It follows that mapping $u \to t$ one has the required extension of the representation it also is clear that the choice $-t$ is the other possible choice changing the sign of the representation.

2) From our previous discussion if $N \subset M$ is an irreducible H submodule then $M = N + N^{\phi_u}$ and we clearly have two cases $M = N$ or $M = N \oplus N^{\phi_u}$.

In the first case tensoring by the sign representation changes the representation while in the second we can represent M as the set $N \oplus N$ of pairs (n_1, n_2) over which H acts diagonally while $u(n_1, n_2) := (h_0n_2, n_1)$.

Similarly $M \otimes \epsilon$ is $N \oplus N$ of pairs (n_1, n_2) over which H acts diagonally while $u(n_1, n_2) := -(h_0n_2, n_1)$.

Then it is immediately seen that the map $(n_1, n_2) \to (n_1, -n_2)$ is an isomorphism of the two structures. □

One should compare this property of the possible splitting of irreducible representations with the similar feature for conjugacy classes.

With the same notations as before.

Exercise. Given a conjugacy class C of G contained in H it is either a unique conjugacy class in H or it splits into 2 conjugacy classes permuted by exterior conjugation by u. The second case occurs if and only if the stabilizer of an element in the conjugacy class is contained in H.

19.2 Let now G be a group H a subgroup and N a representation of H (over some field k).

One considers $k[G]$ as a rigth $k[H]$ module and forms $k[G] \otimes_{k[H]} N$ which is a representation under G by the left action of G on $k[G]$.

Definition. $k[G] \otimes_{k[H]} N$ is called the representation induced from N from H to G, it is also denoted by $Ind_H^G N$.

Exercise. If $G \supset H \supset K$ are groups and N is a K module we have

$$Ind_H^G(Ind_K^H N)) = Ind_K^G N$$

The representation $Ind_H^G N$ is in a natural way described by $\oplus_{g \in G/H} gN$ where by $g \in G/H$ we mean that g runs over a choice of representatives of cosets. The action of G on such a sum is easily described.

There is a similar construction by forming $\hom_{k[H]}(k[G], N)$ where now $k[G]$ is considered as a left $k[H]$ module by the right action and $\hom_{k[H]}(k[G], N)$ is a representation under G by the action of G deduced from the left action on $k[G]$.

$$\hom_{k[H]}(k[G], N) := \{f : G \to N | \, f(gh) = hf(g)\}, \ (gf)(k) := f(g^{-1}k).$$

If G is a finite group one has a $G \times G$ isomorphism between $k[G]$ and its dual (3.1) and we obtain an isomorphism

$$k[G] \otimes_{k[H]} N = k[G]^* \otimes_{k[H]} N = \hom_{k[H]}(k[G], N)$$

for algebraic groups and rational representations it is better to take the point of view of the representation $\hom_{k[H]}(k[G], N)$.

If H is a closed subgroup of G one can define $\hom_{k[H]}(k[G], N)$ as the set of regular maps $G \to N$ which are $H-$ equivariant.

The regular maps from an affine algebraic variety V to a vector space U can be identified to $A(V) \otimes U$ where $A(V)$ is the ring of regular functions on V hence if V has an action under an algebraic group H and U is a rational representation f H the space of H equivariant maps $V \to U$ is identified to the space of invariants $(A(V) \otimes U)^H$.

Assume now that G is linearly reductive and let us invoke the decomposition 15.2.1. $k[G] = \oplus_i U_i^* \otimes U_i$ hence

$$\hom_{k[H]}(k[G], N) = (k[G] \otimes N)^H = \oplus_i U_i^* \otimes (U_i \otimes N)^H$$

finally in order to compute $(U_i \otimes N)^H$ remark that $(U_i \otimes N)^H = \hom_H(U_i^*, N)$.

Assume then that N is irreducible and that H is also linearly reductive, it follows from Schur's Lemma that the dimension of the space $\hom_H(U_i^*, N)$ equals the multiplicity of N in the representation U_i^*. We deduce thus

Theorem Frobenius reciprocity. *The multilicity with which an irreducible representation V of G appears in* $\hom_{k[H]}(k[G], N)$ *equals the multiplicity with which N appears in V as representation of H.*

§20 The unitary trick.

20.1 There are several ways in which linearly reductive groups are connected to compact Lie groups, the use of this (rather strict) connection goes under the name of unitary trick.

This use is in fact made in many different ways and here we want to discuss it with particular reference to the examples which we are studying of classical groups.

The first clue is the Proposition 14.1, $U(n,\mathbb{C})$ is a maximal compact subgroup of $GL(n,\mathbb{C})$ any other maximal compact subgroup of $GL(n,\mathbb{C})$ is conjugate to $U(n,\mathbb{C})$. The way in fact in which $U(n,\mathbb{C})$ sits in $GL(n,\mathbb{C})$ is very special and common to maximal compact subgroups of linearly reductive groups.

The first remark is the polar decomposition theorem

Theorem. *Every invertible matrix X is uniquely expressible in the form*

$$X = Ae^B$$

where A is unitary and B is Hermitian.

Proof. Consider $X^*X := \overline{X}^t X$ which is clearly a positive Hermitian matrix.

If $X = Ae^B$ then $X^*X = e^B A^* A e^B = e^{2B}$. Conversely by decomposing the space in eigenspaces it is clear that a positive Hermitian matrix is uniquely of the form e^{2B} with B Hermitian, hence there is a unique B with $X^*X = e^{2B}$ and, setting $A := Xe^{-B}$ we see that A is unitary. □

The previous Theorem has two corollaries, both of which are sometimes used as unitary trick, the first of algebrogeometric nature the second topological.

Corollary. *$U(n,\mathbb{C})$ is Zariski dense in $GL(n,\mathbb{C})$.*

$GL(n,\mathbb{C})$ is homeomorphic to $U(n,\mathbb{C}) \times \mathbb{R}^{n^2}$ in particular $U(n,\mathbb{C})$ is a deformation retract of $GL(n,\mathbb{C})$.

Proof. The first part follows from the fact that one has the exponential map $X \to e^X$ from complex $n \times n$ matrices to $GL(n,\mathbb{C})$, in this holomorphic map the two subspaces iH, $\mathcal{H}$ of antihermitian and hermitian matrices map to the two factors of the polar decomposition i.e. unitary and positive hermitian matrices.

Since $M_n(\mathbb{C}) = \mathcal{H} + i\mathcal{H}$ any two holomorphic functions on $M_n(\mathbb{C})$ coinciding on $i\mathcal{H}$ necessarily coincide and so by exponential and the connectedness of $GL(n,\mathbb{C})$ the same holds in $GL(n,\mathbb{C})$, two holomorphic functions on $GL(n,\mathbb{C})$ coinciding on $U(n,\mathbb{C})$ necessarily coincide. □

The main general theorem which we shall only illustrate in the case of classical groups is

Theorem. *Given a lineary reductive group G there exists a faithful finite dimensional linear representation of G on a space V and a Hilbert space structure on V such that G is self adjoint.*

The unitary elements of G form a maximal compact subgroup K and we have a canonical polar decomposition $G = Ke^{i\mathfrak{k}}$ where $\mathfrak{k}$ is the Lie algebra of K.

All maximal compact subgroups of G are conjugate in G.

20.2 Let us treat now the other linearly reductive groups that we know, we want to show that in all cases the polar decomposition induces a polar decomposition with similar properties.

First the diagonal group $T = (\mathbb{C}^*)^n$ decomposes as $U(1, \mathbb{C})^n \times (\mathbb{R}^+)^n$ and the multiplicative group $(\mathbb{R}^+)^n$ is isomorphic under logarithm to the additive group of $\mathbb{R}^n$.

It is easily seen that this group does not contain any non trivial compact subgroup hence if $K \subset T$ is compact by projecting to $(\mathbb{R}^+)^n$ we see that $K \subset U(1, \mathbb{C})^n$.

Hence $U(1, \mathbb{C})^n$ is the unique maximal compact subgroup of T.

The orthogonal group $O(n, \mathbb{C})$ contains $O(n, \mathbb{R})$ which is compact since the defining equations $XX^t = 1$ exhibit it as a closed and bounded subset of $\mathbb{R}^{n^2}$.

Let X be orthogonal and decompose it as $X = Ae^B$ the polar decomposition.

Consider

$$X = (X^{-1})^t = (A^{-1})^t e^{-B^t}$$

by the uniqueness of the polar decomposition we have $A^t = (A^{-1}),\ B = -B^t$.

Since A is unitary it follows that A is real hence $A \in O(n, \mathbb{R})$ and also $\overline{B} = -B$ hence $B = iC$ where C is a real and antisymmetric matrix.

Next we want to show that $O(n, \mathbb{R})$ is maximal compact and unique up to conjugation. inserire

The symplectic group and quaternions

We consider now the vector space $\mathbb{H}^n = \oplus_{i=1}^n e_i \mathbb{H}$ over the quaternions (acting on the rigth) with basis e_i.

Write $\mathbb{H}^n := \mathbb{C} + j\mathbb{C}$ with the commutation rules $j^2 = -1,\ j\alpha := \overline{\alpha} j, \forall \alpha \in \mathbb{C}$. As a rigth vector space over $\mathbb{C}$ this has as basis $e_1, e_1 j, e_2, e_2 j, \dots, e_n, e_n j$. For a vector $u := (q_1, q_2, \dots, q_n) \in \mathbb{H}^n$ define $\|u\| := \sum_{i=1}^n q_i \overline{q}_i$ and let $Sp(n, \mathbb{H})$ be the group of quaternionic linear transformations preserving the given norm, it is easily seen that this group can be described as the group of $n \times n$ matrices $X := (q_{ij})$ with $\overline{X}^t = X^{-1}$ where $\overline{X}^t$ is the matrix with $\overline{q}_{ji}$ in the ij entry.

This is again clearly a closed bounded group hence compact.

Return to $\mathbb{H}^n = \mathbb{C}^{2n}$ rigth multiplication by j induces an antilinear transformation with matrix a diagonal matrix J of 2×2 blocks of the form

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

The group $Sp(n, \mathbb{H})$ can then be thougth as the subgroup of the unitary group $U(2n, \mathbb{C})$ commuting with the operator j.

Now if on a complex vector space we have a linear operator X with matrix A and an antilinear operator Y with matrix B it is clear that both XY and YX are antilinear with matrices AB and $B\overline{A}$ respectively, in particular the two operators commute if and only if $AB = B\overline{A}$, we apply this now to $Sp(n, \mathbb{H})$ and see that it is formed by those matrices X in $U(2n, \mathbb{C})$ such that $XJ = J\overline{X} = J(X^{-1})^t$. Its Lie algebra $\mathfrak{k}$ is formed by the antihermitian matrices Y with $YJ = J\overline{Y}$.

Take then $Sp(2n, \mathbb{C})$ to be the symplectic group associated to this matrix J and $X \in Sp(2n, \mathbb{C})$ if and only if $X^tJ = JX^{-1}$, thus we have that

$$Sp(n, \mathbb{H}) = U(2n, \mathbb{C}) \cap Sp(2n, \mathbb{C}).$$

Take the polar decomposition $X = Ae^B$ and compute it for a matrix in $Sp(2n, \mathbb{C})$ remarking that $J^t = -J,\ J^2 = -1$ hence if B is hermitian so is JBJ^{-1}.

We deduce $AJe^{J^{-1}BJ} = (A^{-1})^tJe^{-B^t}$ and, by the uniqueness of the decomposition

$$AJ = (A^{-1})^tJ = \overline{A}J,\ J^{-1}BJ = -B^t = -\overline{B}$$

hence $A \in Sp(n, \mathbb{H})$ while $B \in i\mathfrak{k}$.

§21 Stone Weierstrass approximation and Tannaka Krein duality.

21.1 Given two rational representations M, N of G we consider them as continuous representations of K then.

Lemma. *1)* $\hom_G(M, N) = \hom_K(M, N)$.

2) An irreducible representation V of G remains irreducible under K.

Proof. 1) It is enough to show that $\hom_K(M, N) \subset \hom_G(M, N)$.

If $A \in \hom_K(M, N)$ the set of elements $g \in G$ commting with A is clearly an algebraic subgroup of G containing K, since K is Zariski dense in G the claim follows.

2) is clearly a consequence of 1). □

The next step is to understand that the space of regular functions on a linearly reductive group G restricts to a maximal compact subgroup K isomorphically to the space of representative functions.

In fact it is enough to see two things, first since the compact group K is Zariski dense in G the restriction to K of the algebraic functions is injective and clearly equivariant with respect to the left and rigth action of K.

Next we claim that the restriction to K of the space of algebraic functions is dense in the space of continuous functions, in fact by the Stone Weierstrass approximation Theorem it is enough to show that this space is closed under conjugation and it separates points.

Since the algebraic functions on G separate points in G their restriction to K still separate points.

We have a specific representation of G as a self adjoint group of matrices in which K is the unitary matrices, the entries of the matrices generate together with the inverse of the determinant the algebra of regular functions on G.

Then for the second part remark that we have that since K is made of unitary matrices the conjugates of the entries of the matrices in K are entries of the inverse matrix and hence induced by regular functions on G, similarly the conjugate of the inverse of the determinant is the determinant on K again a regular function.

Corollary. *The category of finite dimensional rational representations of G is equivalent to the category of continuous representations of K.*

Proof. By the previous remarks it is enough to show that every irreducible continuous representation of K is induced by a rational irreducible representation of G.

For this notice that the representative functions of K contain the space of regular functions of G as submodule. If there exists an irreducible representation of K which is not in the list of the rational representations of G its space of matrix coefficients is necessarily ortthogonal in the L^2 norm to the regular functions. Since these are dense in the space of continuous functions this is absurd and the claim follows. □

21.2 We now prove

Lemma. *Given a compact group K if $f_1(x), f_2(x)$ are representative functions of K also $f_1(x)f_2(x)$ is representative.*

If $f(x)$ is representative so is $f(x^{-1})$ and $f(xy)$ is representative as function on $K \times K$.

Proof. f is representative if and only if $f(xy) = \sum_{i=1}^{n} u_i(x)v_i(y)$ for some functions u_i, v_i which are also representative, from which all claims follow easily for instance $f(x_1x_2y_1y_2) = \sum u_i(x_1x_2)v_i(y_1y_2)$ and since u_i, v_i are representative again

$$\sum u_i(x_1x_2)v_i(y_1y_2) = \sum a_{ij}(x_1)b_{ij}(x_2)c_{ih}(y_1)d_{ih}(y_2)$$

from which the claim follows. □

By Proposition 14.15 it follows that the space of representative functions $\mathcal{T}_{K\times K}$ for the group $K \times K$ equals

$$\oplus_{i,j\in\hat{K}}(V_i^* \otimes V_i) \otimes (V_j^* \otimes V_j) = \mathcal{T}_K \otimes \mathcal{T}_K$$

with $\hat{K}$ we have denoted the set of isomorphism classes of irreducible representations of K.

We want then describe the Hopf algebra structure on $\mathcal{T}_K$ which for simplicity we will denote by A. A Hopf algebra structure consists of several operations on A.

(1) A is a commutative and associative algebra under multiplication with 1, we set $m : A \otimes A \to A$ to be the multiplication.

(2) The map $\Delta : f \to f(xy)$ from A to $A \otimes A$ is called a **coalgebra structure**, it is **coassociative** $f((xy)z) = f(x(yz)$ or

$$\begin{array}{ccc} A & \xrightarrow{\Delta} & A \otimes A \\ {\scriptstyle\Delta}\downarrow & & \downarrow{\scriptstyle 1_A \otimes \Delta} \\ A \otimes A & \xrightarrow{\Delta \otimes 1_A} & A \otimes A \otimes A \end{array}$$

is commutative. In general Δ is not cocommutative (i.e. $f(xy) \neq f(yx)$).

(3) $(fg)(xy) = f(xy)g(xy)$ i.e. Δ is amorphism of algebras but also m is a morphism of coalgebras since $m(f(x) \otimes g(y)) = f(x)g(x)$ we see that m is a morphism of coalgebras

$$\begin{array}{ccc} A \otimes A & \xrightarrow{m} & A \\ {\scriptstyle\Delta \otimes \Delta}\downarrow & & \downarrow{\scriptstyle\Delta} \\ (A \otimes A) \otimes (A \otimes A) & \xrightarrow{m \otimes m \circ 1_A \otimes \tau \otimes 1_A} & A \otimes A \end{array}$$

where $\tau(a \otimes b) = b \otimes a$, is commutative.

(4) The map $S : f(x) \to f(x^{-1})$ is called an **antipode**.

Clearly S is a homomorphism of the algebra structure but also $f(x^{-1}y^{-1}) = f((yx)^{-1}$ hence S is an antihomomorphism of the coalgebra structure.

(5) We have the **counit** map $\epsilon : f \to f(1)$ mapping $\epsilon : A \to \mathbb{C}$ which is an algebra homomorphism and with respect to the coalgebra structure we have $f(x) = f(x1) = f(1x)$ or

$$1_A \otimes \epsilon \circ \Delta = \epsilon \otimes 1_A \circ \Delta = 1_A$$

$f(xx^{-1}) = f(x^{-1}x) = f(1)$ or

$$\epsilon = m \circ 1_A \otimes S \circ \Delta = m \circ S \otimes 1_A \circ \Delta$$

(6) It is convenient to think also of the unit element as a map $\eta : \mathbb{C} \to A$ satisfying

$$m \circ (1_A \otimes \eta) = 1_A = m \circ (\eta \otimes 1_A), \ \epsilon\eta = 1_{\mathbb{C}}$$

All the previous properties except the axioms on commutativity or cocommutativity can be taken for the axiomatic definition of a Hopf algebra.

The case of interest to us is when A, as algebra, is the coordinate ring of an affine algebraic variety V i.e. A is finitely generated commutative and without nilpotent elements.

In this case, since to give a morphism between two affine algebraic varieties is equivalent to give a morphism in the opposite direction between their coordinate rings and since $A \otimes A$ is the coordinate ring of $V \times V$ it easily follows that the given axioms, translate, on the coordinate ring, the axioms of an algebraic group structure on V.

Now let K be a linear compact group we claim that its ring of representative functions is finitely generated (obviously without nilpotent elements).

In fact it is generated by the coordinates of the matrix representation and the inverse of the determinant since, by a previous argument they generate a dense subalgebra in the algebra of functions.

Theorem Tannaka-Krein duality. *To any linear compact group K there is canonically associated a linear algebraic group G having as regular functions the representative functions of K.*

G is linearly reductive with the same irreducible reprsentations of K, and K is maximal compact in G.

Proof. We only need to prove the second part of the statement. It is a question of formulating all the usual facts on representations in the dual language of the coordinate ring. inserire $\square$